
How the Axion Paves the Way Beyond the Standard Model

Emmanouil Koutsangelas



München 2023

How the Axion Paves the Way Beyond the Standard Model

Emmanouil Koutsangelas

Dissertation
an der Fakultät für Physik
der Ludwig-Maximilians-Universität
München

vorgelegt von
Emmanouil Koutsangelas
aus Stuttgart

München, den 05. Mai 2023

Erstgutachter: Prof. Dr. Gia Dvali

Zweitgutachter: Prof. Dr. Andreas Weiler

Tag der mündlichen Prüfung: 03. Juli 2023

ZUSAMMENFASSUNG

Das Axion stellt einen vielversprechenden Forschungsbereich dar, da es eine Verbindung zwischen verschiedenen Fragen der Hochenergiephysik aufzeigt. Darüber hinaus floriert das experimentelle Axion-Programm derzeit. Das Verständnis der Axion-Landschaft ist daher entscheidend für die Physik jenseits des Standardmodells. Wir zeigen, dass die kosmologische Restriktion der Axion-Skala durch eine frühe Phase stark gekoppelter QCD vermieden werden kann. Die meisten Ansätze auf Theorien beruhen, bei denen die starke Kopplungskonstante durch den Erwartungswert eines skalaren Feldes bestimmt wird. Alternativ zeigen wir, dass eine frühe Phase stark gekoppelter QCD in den KSVZ- und DFSZ-Modellen durch die Modifikation der Kopplungskonstante während der kosmologischen Inflation entstehen kann. Für beide Modelle verändert die Physik, die für die stark gekoppelter QCD verantwortlich ist, das Minimum des Axions kaum. Die Effizienz der Relaxation wird somit durch Parameter der Theorie und der Dauer der Inflation kontrolliert. Als Nächstes berücksichtigen wir allgemeine Konsistenzanforderungen der Quantengravitation. Dadurch wird zum einen das CP-Problem der QCD zu einem Konsistenzproblem erhoben und zum anderen ergibt sich ein CP-Problem in jeder Yang-Mills Gruppe. Wir betrachten Theorien mit dunklen Yang-Mills Sektoren und untersuchen die phänomenologischen Auswirkungen der Einführung eines Axions pro dunkler Yang-Mills Gruppe, wie es die quantengravitative Konsistenz erfordert. Wir führen unsere Berechnungen für einen reinen Yang-Mills Sektor und N exakte Standardmodellkopien durch. Dabei identifizieren wir eine Beschränkung der dunklen Confinement-Skala, eine Beschränkung auf N , sowie eine Beschränkung der inflationären Hubble-Skala durch Isocurvature-Fluktuationen. Ebenfalls diskutieren wir kompakte Objekte, die kollektiv aus Teilchen verschiedener dunkler Sektoren bestehen, und zeigen wie Kinetic-Mixing zur Existenz von zwei verschiedenen Axion-Zuständen führt. Zuletzt berechnen wir systematisch die Axion-Photon-Kopplung für nicht-minimale DFSZ-Modelle. Dadurch können wir jedes berechnete Modell klassifizieren und die für Axion-Experimente relevanten Verteilungen untersuchen. Wir finden Kopplungen, die fast drei Größenordnungen größer sind als die der minimalen Modelle. Die meisten möglichen Axion-Photon-Kopplungen liegen jedoch in der Nähe der Werte, die von den minimalen Modellen vorgegeben werden. Wir quantifizieren dies, indem wir eine A-priori-Wahrscheinlichkeit für nicht-minimale DFSZ-Modelle einführen und 68% und 95% Grenzen sowie zweiseitige Bänder angeben. Wir vergleichen unsere Ergebnisse mit dem KSVZ-Fall und finden ähnliche Werte, sowie ein sehr spezifisches Muster bei den Verteilungen. Um bevorzugte Modelle zu identifizieren, diskutieren wir Flavor Changing Neutral Currents und das Domain-Wall-Problem als mögliche Auswahlkriterien. Es ist möglich, eine große Anzahl von nicht-minimalen DFSZ-Modellen mit einer Domain-Wall-Nummer von Eins zu konstruieren und damit das Domain-Wall-Problem zu vermeiden. Diese Untergruppe hat auch eine signifikant erhöhte Axion-Photon-Kopplung im Vergleich zu den minimalen DFSZ-Modellen.

ABSTRACT

The axion represents a promising avenue of investigation by providing a common theme among various questions of high-energy physics. Furthermore, its experimental program is currently thriving, with until recently unreachable regions of the axion parameter space being probed. Understanding the axion landscape is thus crucial for beyond the Standard Model physics. We demonstrate that the cosmological bound on the invisible axion scale can be avoided by an early phase of strong QCD. While most approaches rely on theories where the strong coupling constant is determined through the expectation value of some scalar field, we show that an early phase of strong QCD emerges in the benchmark KSVZ and DFSZ models by the modification of the running coupling during inflation. For both models the physics that is responsible for making QCD strong does not displace the axions minimum by too much, so that the efficiency of the relaxation is controlled by parameters of the theory and the number of inflationary e-folds. Next, taking into account general consistency requirements of Quantum Gravity not only elevates the strong CP problem in QCD from a small-value puzzle to a consistency problem, but also induces a “strong CP problem” in every Yang-Mills group. We consider theories with dark Yang-Mills sectors and investigate general phenomenological implications of including one axion per dark Yang-Mills group, as demanded by quantum gravitational consistency. In particular, we carry out computations for a pure YM sector and N exact Standard Model copies, and identify phenomenological consequences such as: a cosmological constraint on the dark confinement scale, a bound on N , as well as no tightening of the bound on the inflationary Hubble scale arising from isocurvature perturbations. We also discuss the phenomenon of compact objects collectively made from particles of different dark sectors, and demonstrate how intersector communication through axion kinetic mixing leads to the existence of two distinct axion states. In the last project, we systematically calculate the axion-photon coupling for non-minimal DFSZ models. Thereby we can classify every calculated model and study the resulting distributions, relevant for various axion experiments. We find couplings almost three orders of magnitude larger than the ones of the minimal models. Most of the possible axion-photon couplings, however, lie in the vicinity of the values dictated by the minimal models. We quantify this by introducing a theoretical prior probability distribution for DFSZ-type axions and giving 68% and 95% lower bounds as well as two-sided bands. We compare our results with the KSVZ case, for which a similar analysis has been conducted. Both display similar values as well as a very specific pattern. In order to identify preferred models, we discuss the role of flavor changing neutral currents and the domain wall problem as selection criteria. It is possible to construct a large number of non-minimal DFSZ models with a domain wall number of unity and thereby avoid the domain wall problem. This subset also has a significantly enhanced axion-photon coupling compared to the minimal DFSZ models.

PROJECTS AND PUBLICATIONS

This thesis is based on completed and ongoing projects to which I have contributed during my research conducted at the Ludwig–Maximilians–Universität München and the Max Planck Institute for Physics from June 2019 to April 2023. The projects resulted in a series of paper [1–4] that have already been published or are in preparation. The authors are listed alphabetically by convention in particle physics and share the principal authorship. Although some new aspects were added and discussions were rephrased for a better understanding, a substantial part of this thesis is an *ad verbatim* reproduction with respect to equations, figures and tables of these papers:

1. G. Dvali, E. Koutsangelas, and F. Kuhnel, “Compact Dark Matter Objects via N Dark Sectors,” *Phys. Rev. D* **101** (2020) 083533, [arXiv:1911.13281](#) [[astro-ph.CO](#)]
2. E. Koutsangelas, “Removing the cosmological bound on the axion scale in the Kim-Shifman-Vainshtein-Zakharov and Dine-Fischler-Srednicki-Zhitnitsky models,” *Phys. Rev. D* **107** no. 9, (2023) 095009, [arXiv:2212.07822](#) [[hep-ph](#)]
3. J. Diehl and E. Koutsangelas, “Dine-Fischler-Srednicki-Zhitnitsky-type axions and where to find them,” *Phys. Rev. D* **107** no. 9, (2023) 095020, [arXiv:2302.04667](#) [[hep-ph](#)]
4. M. Ettengruber and E. Koutsangelas, in preparation

The following paper in preparation is not included in this thesis:

5. A. Alexandre, G. Dvali, and E. Koutsangelas, in preparation

ACKNOWLEDGEMENTS

During my PhD research in the group of the renowned scientist Prof. Dr. Gia Dvali, I had the chance and the pleasure to interact with a large number of people who have contributed in no small way to the successful completion of this thesis. This journey has been particularly challenging due to the unprecedented circumstances of the Covid-19 pandemic. Therefore, I would like to take this opportunity to express my sincere gratitude to all those who played a part in this chapter of my life, whether through engaging discussions, valuable feedback, or even simple words of encouragement.

First and foremost, I offer my deepest appreciation to my supervisor Prof. Dr. Gia Dvali, whose passion and scientific work sparked my interest in particle physics and cosmology. I want to thank him for all the time he invested in me, for his priceless advice, and his consistent support, which was crucial in navigating the ups and downs of my PhD. His lectures, seminars, and our discussions have been an invaluable contribution to my understanding of physics and the time as his PhD student taught me qualities that will benefit me beyond graduation. Additionally, I want to acknowledge his understanding during the challenging times of the Covid-19 pandemic.

It is my pleasure to thank my collaborators Florian Kühnel, Johannes Diehl, Manuel Ettengruber and Ana Alexandre for the various topics that we have explored during the last years. It was a real pleasure to work and learn by your side. Moreover, I am most grateful to Georg Raffelt for his personal and professional mentoring, which has granted me numerous insights into axion phenomenology and the craft of academic writing.

I would like to express my appreciation to Frank Steffen and the International Max Planck Research School on Elementary Particle Physics for the interesting seminars, young scientist workshops at castle Ringberg, block courses, and the financial support. In particular, I want to thank our secretaries Herta Wiesbeck-Yonis and Vera Kudrin, and the rest of the MPP administrative staff who on a daily basis tackle questions for us PhD students in a field vastly more complex than theoretical physics: German bureaucracy.

Many thanks go out to my colleagues from LMU and the MPP. You provided an amazing work atmosphere and I will surely miss the great discussions while having lunch, coffee or ice cream: Marc Syväri, Thomas Steingasser, Max Bramberger, Felix Palm, Anamaria Hell, Christoph Müller-Salditt, Max Brinkmann, Bence Temesi, Anja Stuhlfauth, Giacomo Contri, Giordano Cintia, Maximilian Bachmaier, Lasha Berezhiani, Silvia Zanolì, Thomas Raml, Andrei Kovtun, Otari Sakhelashvili, Juan Sebastian Valbuena Bermudes, Alessandro Ratti, Luc Schnell, Christian Kneißl, Antonela Matijašić, Dominik Fuchs and many others. I would like to extend a very special thank you to Annika Stenzl. Thank you for sharing so much with me, for always having my back, and for being relentlessly optimistic, especially when I was not. The time spent with you in the Mathebau will always remain one of my favorite memories about my PhD journey.

Moreover, I am grateful to Gia Dvali, Andreas Weiler, Anja Stuhlfauth, Lea Wassermann, Annika Stenzl and Maximilian Bachmaier for proof-reading and improving parts of this thesis. I also wish to express my gratitude to Gia Dvali, Andreas Weiler, Otmar Biebel and Jan von Delft for being part of my PhD committee.

The last lines are dedicated to my family. I am immensely grateful to Dennis Schäfer, Nahom Solomon, Max Böhler, Ario Sharghi, Christoph Simon, Jakob Johnson and Marvin Müller. Through almost all my life you were the people I knew I could always count on, no matter how difficult times were. You guys are like family to me and I consider myself unbelievably lucky to have experienced such friendship in this life. Lea Wassermann, thank you for your selfless support and unconditional love given to me during the past years. My deepest gratitude goes to my mum, my stepfather, and my grandparents. You raised me with so much dedication and love. There are no words to describe how grateful I am for everything you have done for me. I love you.

CONTENTS

Zusammenfassung	i
Abstract	iii
Projects and Publications	v
Acknowledgements	vii
List of Abbreviations	xiii
1 Introduction	1
1.1 High Energy Physics and the Axion	2
1.2 Outline	5
2 The Strong CP Problem	7
2.1 Low Energy QCD	8
2.1.1 Chiral Symmetry Breaking	8
2.1.2 The Chiral Lagrangian	9
2.1.3 The $U(1)_A$ Problem	11
2.1.4 The Role of θ	12
2.2 Instantons	15
2.2.1 Semi-Classical Theory of Tunneling	15
2.2.2 Tunneling in YM Theory and the Role of Topology	16
2.2.3 The YM Vacuum	20
2.2.4 The YM Vacuum Energy	21
2.2.5 Tunneling in the Presence of Fermions	24
2.3 Topological Three-Form Formalism	26
2.3.1 Lessons from the Schwinger Model	26
2.3.2 Topological Susceptibility in QCD	31
2.3.3 Three-Form Higgs Effect	32
2.3.4 Higgsing of Multiple Three-Forms	34
2.4 $\bar{\theta}$ and the Strong CP Problem	36
2.4.1 Observing $\bar{\theta}$ via the nEDM	36
2.4.2 Constraining $\bar{\theta}$ via Quantum Gravity	38

3	The QCD Axion	41
3.1	The PQ Mechanism	42
3.2	The Low-Energy EFT of the Axion	42
3.2.1	Effective Lagrangian	43
3.2.2	Axion Potential and Mass	44
3.2.3	Axion-Photon Coupling	45
3.3	Axion Models	46
3.3.1	PQWW-Model	46
3.3.2	DFSZ Model	48
3.3.3	KSVZ Model	50
3.3.4	Two-Form Axion	51
4	Cosmological Bounds on the Axion	53
4.1	Axion Production	54
4.1.1	Misalignment Mechanism	55
4.1.2	Decay of Topological Defects	58
4.2	The Axion Window	63
4.2.1	Isocurvature Fluctuations	64
4.2.2	Primordial Gravitational Waves	64
4.2.3	Thermal Fluctuations	65
4.2.4	Astrophysics	65
4.2.5	Black Hole Super-radiance	66
4.3	Early Relaxation using Inflation	67
4.3.1	General Idea	67
4.3.2	The QCD Scale during Inflation	68
4.3.3	Location of Minima during Inflation	71
4.3.4	The Modified Axion Window	73
4.4	Summary and Discussion	75
5	Axions and Hidden YM Groups	79
5.1	Cosmological Implications	80
5.1.1	Cosmological Framework	81
5.1.2	Misalignment: Exact SM Copies	82
5.1.3	Misalignment: SM-like Dark Sectors	87
5.1.4	Misalignment: Pure YM Dark Sector	88
5.1.5	Isocurvature Perturbations	89
5.2	Compact Dark Matter	91
5.2.1	N -MACHOs	91
5.2.2	N -Boson Stars	94
5.3	Kinetic Mixing between Axions	96
5.3.1	Kinetic Mixing as Intersector Interaction	96
5.3.2	Modification of Axion Physics	98
5.4	Summary and Discussion	101

6	Non-Minimal DFSZ Models	103
6.1	DFSZ-Type Models	105
6.1.1	Identifying the Axion	105
6.1.2	The PQ Charges	107
6.1.3	Multiplicity	108
6.1.4	Selection Criteria	109
6.2	Anomaly Ratio Distributions	111
6.2.1	Example: $n_D = 3$	111
6.2.2	Choices for a Statistical Interpretation	114
6.2.3	Results for $n_D = 4 - 7$	114
6.2.4	Extrapolation to $n_D > 7$	118
6.2.5	Constructing extreme $ g_{a\gamma} $	120
6.2.6	Comparison with KSVZ-Type Models	122
6.3	Implications for Axion Searches	124
6.3.1	$C_{a\gamma}$ Cumulative Distribution Function	124
6.3.2	Experimental Constraints	125
6.4	Summary and Discussion	130
7	Conclusion and Outlook	133
A	Conventions and Fundamentals	137
A.1	Units	137
A.2	Indices and Metric Signature	137
A.3	2π and Fourier Transformation	138
A.4	Gauge Theories	139
A.5	The Standard Model	140
A.6	Higher-Forms	141
A.7	Euclidean Formulation of YM Theories	142
	List of Figures	148
	List of Tables	149
	Bibliography	151

LIST OF ABBREVIATIONS

2HDM	Two Higgs doublet model
BBN	Big Bang nucleosynthesis
BPST	Belavin-Polyakov-Schwarz-Tyupkin
CDF	cumulative distribution function
CGS	centimeter-gram-second
CL	Confidence Level
CMB	Cosmic Microwave Background
CP	Charge conjugation parity
DFSZ	Dine-Fischler-Srednicki-Zhitnitsky
EFT	Effective field theory
FCNC	Flavor changing neutral current
FLRW	Friedmann-Lemaître-Robertson-Walker
GR	General Relativity
GUT	Grand unified theory
KSVZ	Kim-Shifman-Vainshtein-Zakharov
KZM	Kibble-Zurek Mechanism
LHC	Large Hadron Collider
LIGO	Laser Interferometer Gravitational-Wave Observatory
LSE	Linear system of equations
MACHO	Massive Compact Halo Object
nEDM	Neutron electric dipole moment
PDF	Probability Density Function
PQ	Peccei-Quinn
PQWW	Peccei-Quinn-Weinberg-Wilczek
QCD	Quantum Chromodynamics
QED	Quantum Electrodynamics
QFT	Quantum field theory
SM	Standard Model
SU(n)	Special unitary group of degree n
SUSY	Supersymmetry
TOV	Tolman-Oppenheimer-Volkoff
U(n)	Unitary group of degree n
UV	Ultraviolet ($\hat{=}$ high-energy)
VEV	Vacuum expectation value
YM	Yang-Mills

CHAPTER ONE

INTRODUCTION

During the 20th century, modern physics witnessed two of its greatest accomplishments: Einstein's general theory of relativity (GR) and the Standard Model of particle physics (SM). Together, by classifying all known particles and their interactions, these two theories provide an exceptionally successful description of the basic building blocks and fundamental forces of nature [6–8]. This allows to explain physical phenomena from tiny distances of 10^{-16} cm all the way up to the size of the observable universe at 10^{28} cm in agreement with all experimental tests [9]. The SM, in particular, has demonstrated experimental accuracy to an impressive degree of 11 digits, establishing it as the most precise scientific theory ever tested [10].

Despite of their tremendous success, the SM and GR are not the final step in our understanding of nature. There are still several unanswered questions of high physical importance that either stem from observed phenomena or from problems within these two theories. Examples for the former include: What is the mechanism behind the neutrino masses [11, 12]? What is dark matter made of [13–15]? Why is the cosmological constant so unnaturally small [16–18]? Examples for the latter are: What is the correct theory of high-energy quantum gravity [19]? What protects the Higgs mass against high-energy contributions [20]? Why does the strong interaction not violate the charge conjugation parity (CP) symmetry [21–23]? All these questions are difficult or even impossible to answer by means of the SM and GR.

Although this may seem like an unprecedented situation, it is not. Our theories of nature are provisional, meaning that they have been tested and found to be valid over a limited range of energies and distances. It is unclear whether they hold true in more extreme circumstances. Many theories have been surpassed by new ones at higher energies, and we anticipate that this pattern will persist. Seen in this light, the unanswered questions of the previous paragraph indicate the existence of theories that go beyond. But the question is how to go beyond.

1.1 High Energy Physics and the Axion

The research presented in this thesis focuses on the last of the mentioned questions, which goes under the name of the strong CP problem. Astonishingly, this problem is closely intertwined with several of the other mentioned problems. Let us briefly elaborate on this connection and why we consider it so important for the way to new physics.

The Strong CP Problem. The non-Abelian nature of quantum chromodynamics (QCD) leads to many new features compared to an Abelian theory as quantum electrodynamics (QED). In particular, QCD has a plethora of vacua that are labeled by a continuous angular parameter called the θ -angle [24]. This non-trivial vacuum structure manifests itself in the form of CP violating processes that for instance result in an electric dipole moment for the neutron (nEDM) proportional to θ [21]. The nEDM has not been observed experimentally, thus putting a bound on the θ -angle, i.e. $\theta \ll 10^{-10}$ [22]. The strong CP problem is the question of why θ is so small. This does not present a naturalness problem because, in contrast to say the Higgs mass, the θ -angle is not sensitive to “ultraviolet” (UV) physics. In fact, its quantum corrections from the SM are many orders below the bound [25]. Thus, in this common view the strong CP problem is not really a problem but rather a small-value puzzle.

High-Energy Quantum Gravity. GR is a non-renormalizable theory and, as such, it was thought for a long time that it cannot be quantized. From a modern point of view, however, GR is an effective field theory (EFT). Within this framework, GR provides a well-behaved theory of low-energy quantum gravity with a (maximal) cut-off given by the Planck scale M_P [26, 27]. Past this cut-off, the true nature of quantum gravity remains a mystery. From the point of view of EFT, this lack of knowledge poses no problem for low-energy physics. Any corrections to GR will be suppressed by higher powers of M_P , so that high energy quantum gravity completely decouples on the electroweak scale. Therefore, it may seem reasonable to ignore quantum gravity in fundamental particle physics or even remain agnostic about the nature of high-energy quantum gravity. But what if high-energy quantum gravity provides self-consistency conditions? In this case ignoring it would result in the consideration of theories that are inconsistent in a real world where gravity is present. The impact could be dramatic.¹

A compelling argument for such a self-consistency condition was presented by Dvali et al. [30–33]. They suggested that the concept of a cosmological constant is incompatible with the quantum resolution of the de Sitter metric as a coherent state of gravitons.² The reason for this is that the gravitons are subject to quantum-scattering, which inevitably will result in the self-destruction of the coherent state. This is in contradiction with a de

¹In the context of string theory this question started the so called swampland program, namely the systematic search of low-energy theories that are incompatible with string theory [28, 29].

²Representing classical gravitational backgrounds, such as black holes and cosmological space-times, in terms of graviton coherent states may sound exotic. In fact, it is not. Any classical field configuration is a limit of quantum states with large occupation numbers of particles. For example, a classical electromagnetic wave is fundamentally a coherent state of photons. It is thus anticipated that this phenomenon will be present in any microscopic theory that accounts for the quantum corpuscular structure of the cosmological background.[30]

Sitter “vacuum” and thus a cosmological constant [31]. Fundamentally, this outcome is rooted in the S-matrix formulation of quantum gravity [33]. The unique limit, in which quantum back-reactions of scattered gravitons on the hypothetical de Sitter vacuum with curvature radius R_{dS} and cosmological constant Λ vanish, is given by $\Lambda \rightarrow \infty$, $M_{\text{P}}^{-1} \rightarrow 0$ such that $\Lambda M_{\text{P}}^{-1} = R_{\text{dS}}^{-2} = \text{finite}$. In this limit, however, the coupling of gravitons with wavelength $\lambda \lesssim R_{\text{dS}}$ goes as $\alpha_{\text{gr}} = 1/M_{\text{P}}^2 \lambda^2 \rightarrow 0$. This indicates that a consistent theory of gravity based on the S-matrix formulation cannot accommodate a de Sitter vacuum without decoupling gravity. Consequently, de Sitter cannot be regarded as a vacuum but only as an excited (coherent) state constructed on top of a true S-matrix vacuum of Minkowski with $\Lambda = 0$.

The connection between these two problems is established by the fact that every QCD vacuum angle with $\theta \neq 0$ behaves as a cosmological constant [34, 35]. Taking the quantum gravity imposed S-Matrix exclusion of de Sitter serious, the number of viable vacua reduces to exactly one: the CP conserving vacuum at $\theta = 0$. This not only elevates the strong CP problem in QCD from a small-value puzzle to a real problem, but makes a vacuum-selection mechanism a necessity. Such a mechanism is given by the Peccei-Quinn (PQ) solution, which essentially introduces a non-linearly realised $U(1)_{\text{PQ}}$ that is anomalous with respect to QCD [36, 37]. The PQ mechanism solves the strong CP problem as the vacuum angle θ gets relaxed to zero by the pseudo-Goldstone boson of the PQ symmetry, namely the axion [38, 39].

As a pseudo-Goldstone boson, the axion has properties that allow the connection to the question about the *true nature of dark matter*. This question arises from plenty of evidence that the matter visible to us only represents a subdominant fraction of the total amount present in the universe [40]. Instead, most of the matter appears to be a cold, collisionless fluid that only interacts gravitationally [13–15]. The invisibility of this component has led to the name dark matter. While a remarkably successful effective description of the dark matter is provided by the cosmological Lambda-Cold-Dark-Matter model [41, 42], a fundamental description in the form of a particle is still missing. As such any particle candidate must fulfill three crucial properties in addition to behaving as a cold, collisionless fluid [43]. First of all, it needs to be stable, otherwise it would decay and not be present in the universe. Secondly, it must be very weakly interacting via the forces described by the SM. Lastly, there must be a production mechanism that results in the observed abundance. The axion fulfills all of these properties, making it of high interest not only to particle physics but also to cosmology and astrophysics. Its production via the so called misalignment mechanism, in particular, puts a severe bound on the parameter space of the axion. The first presented project in this thesis provides evidence on the non-robustness of this cosmological bound by extending on Dvali’s scenario [44].

Going back to the strong CP problem viewed in the light of consistency, it is not exclusive to QCD but extends to any non-Abelian Yang-Mills (YM) group. Every YM group includes vacuum angles that lead to unacceptable de Sitter-type vacua and must therefore be removed [45]. Using the PQ mechanism to achieve this, requires one axion per

YM group because a single axion cannot set multiple θ -parameters to zero, even when all θ parameters are set equal [46, 47]. This is due to the emergence of new effective vacuum angles. Allowing these vacuum angles to persist would conflict with the requirement of a vanishing cosmological constant, hence requiring the use of one axion per YM group. Even without the requirement of a vanishing cosmological constant, the necessity of one axion per YM group still arises. With the new effective vacuum angles the question of the smallness of θ is shifted to the smallness of these new angles. Thus, one can say that the need for one axion per YM group is in fact a condition required for solving the strong CP problem. The second project covered in this thesis investigates the phenomenological consequences of the necessary multitude of axions in theories with hidden non-Abelian YM group.

Theories with a large number of non-Abelian YM groups are of high interest for the mentioned question about the Higgs mass stability. This question goes under the name of the *hierarchy problem* and is basically about one of the most important lessons that GR and the SM teach us, i.e. that physics is scale-dependent. While on our everyday scales Newton's law of gravity provides an excellent description, it must be replaced by Einstein's GR in the vicinity of very heavy objects such as stars or black holes. Similarly, while at energies way below the W -boson mass Fermi's theory provides an excellent description of the weak interaction, for energies comparable to the W -boson mass the electroweak part of the SM must be used. As such the question arises why the Planck mass M_P , as the fundamental scale of gravity, is so much higher than the electroweak scale $v_{EW} \sim 10^2$ GeV. Equivalently, by looking at the ratio between Newton's and Fermi's constants, i.e. $G_N/G_F \sim (v_{EW}/M_P)^2$, this can be rephrased as the question why gravity is so much weaker than the weak force. This is the Hierarchy problem [20]. Formulated in this way, the Hierarchy problem sounds more like a puzzle than a problem but the actual problem can be identified by considering the Higgs mass. Being the mass of a scalar field, it receives quantum corrections proportional to the cutoff of the theory. If these corrections were smaller than the observed value, there would be no problem. However, taking M_P as the cutoff (or any fundamental scale above TeV) requires fine-tuning and thus renders the Higgs mass unnatural. In the presence of a large number of particle species, such as the gluons from additional non-Abelian YM groups or the various fields from copies of the SM, the true fundamental scale of gravity can be pushed all the way down to the electroweak scale. This would nullify the Hierarchy problem by having the Planck scale coincide with the electroweak scale or, in other words, having strong gravity effects in the vicinity of the energy scales probed by the LHC [48, 49].

The requirement to solve the strong CP problem does not specify all of the axion's properties. As a pseudo-Goldstone, all axion low-energy couplings depend on the UV realization of the PQ mechanism, thus requiring model building to make concrete predictions [50]. This is usually achieved by the two large classes of invisible axion models, the Dine-Fischler-Srednicki-Zhitnitsky-type (DFSZ-type) [51, 52] and Kim-Shifman-Vainshtein-Zakharov-type (KSVZ-type) models [53, 54]. The former adds Higgs singlets and doublets to the SM, while the latter adds Higgs singlets and heavy quarks

(for a review see [55]). Even though the minimal models of each type, adding only one Higgs singlet and one Higgs doublet for DFSZ and one Higgs singlet and one uncharged quark for KSVZ, define benchmark models, in principle there is a plethora of non-minimal models. An identification of all these models and systematic approach that allows to extract a prediction from all these models at the same time would be desirable. The goal of the third project is to do exactly this for the DFSZ-type axions.

We see that by predicting the existence of a new light field, the PQ solution opens up a vast landscape of new research directions and challenges. Hence, this solution does not present the end of the story around the strong CP problem, but in a sense a new beginning. Our personal answer to the question on how to go beyond the SM is that the axion represents a promising avenue of investigation by, remarkably, providing a common theme among various questions of high-energy physics. Therefore, understanding this landscape in great detail is essential to pave the way beyond the SM. Of course this answer is highly subjective and its correctness can only be determined retrospectively. However, scientific progress relies on attempting various approaches based on personal interpretation and understanding. So we take our answer as the starting point for this thesis and shall try to delve into the questions surrounding the axion and its potential role in going beyond the SM.

1.2 Outline

While Chapter 2 and Chapter 3 provide a didactic introduction of the theoretical background, Chapter 4, Chapter 5, and Chapter 6 are devoted to individual projects. The latter three chapters feature an introduction and conclusion that provide readers with an overview of the material covered. In addition, these chapters are self-contained and can be read on their own, allowing readers to select the chapters that are of relevance to them. This comes with the price that if one chooses to read the thesis from start to finish, there may be some overlap in the material presented in each chapter. Finally, it is worth noting that standard quantum field theory (QFT) notation has been summarized in Appendix A to avoid interrupting the reading flow. With these things in mind, let us briefly outline the topics covered in each individual chapter.

Chapter 2 provides a detailed introduction to the strong CP problem. Apart from establishing the standard language that is used in axion physics, including chiral perturbation theory and instantons, we make the connection to not so well-known aspects of the axion, namely as the Schwinger model and the higher-form formalism. We try to be very didactic so that this and the next chapter can be used as a suitable starting point for students who are interested in this topic and want to acquire an understanding of the concepts involved.

Chapter 3 introduces the PQ mechanism and discusses how the axion solves the strong CP problem in the various formalisms. Moreover, we write down the low-energy EFT of the axion and introduce the four models that are used throughout the rest of this thesis:

The Peccei-Quinn-Weinberg-Wilczek (PQWW) model, the Kim-Shifman-Vainshtein-Zakharov (KSVZ) model, the Dine-Fischler-Srednicki-Zhitnitsky (DFSZ) model, and the two-form implementation of the axion.

Chapter 4 places the axion within a cosmological framework and discusses the various implications, such as cosmological bounds on the axion scale arising from overproduction and isocurvature perturbations. We discuss a known mechanism to avoid these bounds based on an early phase of strong QCD and show that this mechanism can be implemented into the KSVZ and DFSZ models by using the modification of the running coupling during inflation. In particular, we illustrate how in both models the physics that makes QCD strong during inflation does not displace the axions minimum by too much, so that the efficiency of the relaxation is controlled by parameters of the theory and the number of inflationary e-folds. Moreover, we identify the parameter space compatible with all cosmological constraints when the axion abundance is dominated by inflationary quantum and post-inflationary thermal fluctuations.

Chapter 5 considers theories with hidden YM (sub-)groups and identifies general phenomenological consequences of including one axion per YM group, as demanded from the strong CP problem understood as a consistency problem of every YM group. In particular, we carry out computations for two simple models, namely a pure YM sector and N exact Standard Model copies. We illustrate phenomenological consequences such as: a cosmological constraint on the dark confinement scale, a bound on N , as well as a potential tightening of the bound on the inflationary Hubble scale arising from isocurvature perturbations. Furthermore, we discuss the phenomenon of axion stars collectively made from axions of different dark sectors, as well as the possibility of axion kinetic mixing and its implications for axion states and stability.

Chapter 6 presents a systematic way to calculate the axion-photon coupling for a large number of DFSZ-type models, thereby allowing the classification of these models and the study of the resulting distributions. We show that most of the possible axion-photon couplings lie in the vicinity of the values dictated by the minimal models by introducing a theoretical prior probability distribution for DFSZ-type axions and giving 68% and 95% lower bounds as well as two-sided bands. Moreover, we compare the DFSZ-type distributions with the ones from the KSVZ case and discuss the role of flavor changing neutral currents and the domain wall problem as possible selection criteria. We find a large number of non-minimal DFSZ models with a domain wall number of unity. Intriguingly, these models have an enhanced axion-photon coupling compared to the minimal DFSZ models.

Chapter 7 gives general conclusions that complement the individual summaries provided in Chapters 4 – 6. Additionally, we provide a glimpse into potential theoretical studies in the future that could expand upon the ideas presented in this thesis.

CHAPTER TWO

THE STRONG CP PROBLEM

The PQ mechanism was proposed to solve the strong CP problem almost 50 years ago [36, 37]. For this reason it may seem that comprehending how this mechanism works is only relevant for historical purpose. However, gaining an understanding of how the PQ mechanism resolves the strong CP problem offers valuable insights into the axion. First of all, knowing which axion properties are universal and which are model-dependent is crucial in identifying the axion parameter space. Second, in complex models it is easy to undermine the axion's ability to solve the strong CP problem. For these reasons we take the perspective that it is necessary, rather than purely historical, to understand the details behind the PQ mechanism. This requires understanding the physics underlying the strong CP problem first, to which we dedicate this whole chapter.

The standard starting point for introducing the strong CP problem is Weinberg's $U(1)_A$ problem [56]. Then 't Hooft's instanton solution is usually discussed. While solving the $U(1)_A$ problem, instantons result in the strong CP problem in the sense of a small-value puzzle [57–59]. The first two sections of this chapter stick to this standard way to establish the language commonly used in axion physics. From there, further not so commonly known aspects of the strong CP problem are discussed that are important for some projects on which this thesis is based. These aspects are the $1 + 1$ dimensional Schwinger model [60, 61] as a toy model of the strong CP problem [62], the topological formulation of the strong CP problem using higher forms [63], and the role of gravity [31, 64].

2.1 Low Energy QCD

In QCD, the fundamental degrees of freedom are gluons and quarks. However, the situation changes when the theory is considered at low energies. Due to quantum corrections, the strength of the coupling constant becomes stronger with lower energies until the strong coupling regime is entered and confinement takes place [65, 66]. The scale this happens is called the confinement scale Λ_{Conf} . Additionally, the quark sector of QCD exhibits an approximate chiral symmetry that is spontaneously broken by the condensation of the quarks. The scale at which this happens is called the chiral symmetry breaking scale Λ_χ . It is still not clear as to whether or not chiral symmetry breaking and confinement are separate dynamics or inherently linked in QCD [67], but for the sake of this thesis we will set $\Lambda_{\text{Conf}} = \Lambda_\chi \equiv \Lambda_{\text{QCD}}$. Although it is fascinating by itself that the conformal symmetry of the classical theory is broken upon quantization, the crucial point is that below Λ_{QCD} the perturbative degrees of freedom are no longer gluons and quarks but mesons and baryons. The theory of those degrees of freedom and how it arises from QCD can be explained by the framework of chiral perturbation theory, to which we turn as a tool to understand the $U(1)_A$ problem. This section is mainly based on [68], further sources and insights on this topic can be found in the references cited therein.

2.1.1 Chiral Symmetry Breaking

To simplify the discussion, we will disregard the strange quark for the time being, so that the number of active quarks is $n_F = 2$. In addition, we assume that the remaining light quarks have zero mass, which is a reasonable approximation since $m_u, m_d \ll \Lambda_{\text{QCD}}$. With these assumptions, the QCD Lagrangian simplifies to

$$\mathcal{L}_{\text{QCD}} = i\bar{\Psi}^i \gamma^\mu D_\mu \Psi_i - \frac{1}{4} G_{\mu\nu}^a G^{a,\mu\nu} + \theta \frac{g_s^2}{32\pi^2} G_{\mu\nu}^a \tilde{G}^{a,\mu\nu}, \quad (2.1)$$

where θ is a parameter called the vacuum angle and $\tilde{G}^{a,\mu\nu}$ is the dual field strength defined as

$$\tilde{G}^{a,\mu\nu} = \frac{1}{2} \epsilon^{\mu\nu\alpha\beta} G_{\alpha\beta}^a. \quad (2.2)$$

The last term has been added since it fulfills all the symmetry requirements of QCD. Later in this chapter, we will demonstrate that it can be expressed as a total derivative, implying that it is a boundary term. For the sake of this section, we will thus make the naive conclusion that this term is non-physical and can be neglected.

In addition to the local $SU(3)_c$ color symmetry, this Lagrangian has a global $U(2)_L \times U(2)_R$ flavor symmetry under which

$$\Psi_{L,i} \longrightarrow L_i^j \Psi_{L,j}, \quad L \in U(2)_L, \quad (2.3)$$

$$\Psi_{R,\bar{i}} \longrightarrow R_{\bar{i}}^{\bar{j}} \Psi_{R,\bar{j}}, \quad R \in U(2)_R. \quad (2.4)$$

This symmetry is chiral since the left and right handed fermions transform differently. In

the following it will be advantageous to decompose this symmetry as

$$U(2)_L \times U(2)_R \simeq SU(2)_L \times SU(2)_R \times U(1)_V \times U(1)_A .$$

Since this is per se a symmetry of QCD, the following question arises: What do the charges of these subgroups correspond to?

- The axial $U(1)_A$ symmetry is anomalous, thus it is not expected that the low-energy particles will carry a corresponding charge.
- The charges of the vector $U(1)_V$ are nothing else than the quark numbers or equivalently the baryon numbers.
- The vector subgroup obtained by setting $R = L$ is the isospin symmetry under which hadrons and mesons have non-trivial representations. For example the proton and the neutron form a doublet, while the three pions π^0, π^+, π^- form a triplet.
- In contrast, there is no sign of classification in terms of the axial subgroup $SU(2)_A$, which is obtained by setting $R = L^\dagger$. It must be spontaneously broken.

The vector subgroups seem clear since they are still good symmetries in the infrared. On the other hand, the breaking of the axial subgroups gives rise to many questions. In the following, we discuss two of them in detail.

2.1.2 The Chiral Lagrangian

The first question to be discussed is how the subgroup $SU(2)_A$ is broken. This question can be elegantly answered by means of effective field theory, meaning that knowing the relevant degrees of freedom and symmetries, we can construct an effective Lagrangian that captures the relevant physics.

First, the order parameter has to be identified. Obviously, it must be a Lorentz scalar and an $SU(3)_c$ singlet to not disturb these symmetries. Furthermore, since there are no fundamental scalars, composite fields must be used of which the simplest possibility is $\bar{\Psi}_{L,i} \Psi_R^j$. Finally, the vacuum expectation value (VEV) has to be invariant under $SU(2)_V$ but not under $SU(2)_A$, so that the simplest possibility is $\langle \bar{\Psi}_{L,i} \Psi_R^j \rangle = v^3 \delta_i^j$, where $v \equiv \Lambda_{\text{QCD}}$. A small note is appropriate here: This VEV vanishes at tree level and to all orders in perturbation theory due to the flavor symmetry. In other words, there is no local operator in QCD satisfying the required symmetries that allow the transition of a left-handed fermion into a right-handed fermion. Therefore, the origin of chiral symmetry breaking has to be of non-perturbative nature.

Next, the order parameter needs to be expanded around the vacuum. Since the vacuum manifold for the considered breaking is a three-sphere S^3 , Goldstone's theorem implies that the Goldstone bosons must take values on it. As a result, there is a freedom in choosing the coordinates to parametrize the sphere. In the following the typical choices are summarized:

- Linear parametrization: This yields the linear σ -model of Levy and Gell-Mann [69]. Even though omitting the non-linear terms gives a renormalizable theory, this parametrization fails due to phenomenology – the massive radial component σ was never found.
- Stereographic projections: This leads to the non-linear σ -model of Levy and Gell-Mann [69]. From the point of view of the linear sigma-model, this parametrization can be found by taking a double scaling limit such that σ becomes infinitely heavy and is thus pushed out of the physical spectrum. In this limit the potential becomes a constraint that, when solved explicitly, gives the Lagrangian of the non-renormalizable non-linear σ -model.
- Exponential map: In this parametrization, the non-linearity is encoded in the exponential map, which expanded gives rise to infinite interactions between the Goldstones. Hence, this model is not renormalizable. It is most suitable for our purpose and thus will be used in the following.

The expansion around the VEV using the exponential map yields

$$\bar{\Psi}_L \Psi_R = v^3 \exp\left(\frac{i2\pi^a(x)T^a}{f_\pi}\right) \equiv v^3 U(x), \quad (2.5)$$

where $T^a = \sigma^a/2$ and the factor of 2 is merely for convenience. The most general Lagrangian for the low-energy effective field $U(x)$ that is consistent with $SU(2)_L \times SU(2)_R$ is then

$$\mathcal{L} = \frac{1}{4} f_\pi^2 \text{tr}(\partial_\mu U^\dagger \partial^\mu U) = \frac{1}{2} \partial_\mu \pi^a \partial^\mu \pi^a + \dots, \quad (2.6)$$

where the dots denote higher derivative terms that are suppressed with powers of ∂_μ/f_π . Furthermore, there are no terms without derivatives since $U(x)$ is unitary.

Before going to the second big question, let us add the masses of the quarks, i.e.

$$\mathcal{L}_m = -\bar{\Psi}_L M \Psi_R + \text{h.c.} \quad (2.7)$$

This term is minimized by $U = \text{Id}_2$ so that the expansion is left unaltered. After chiral symmetry breaking, the effective mass term is

$$\mathcal{L}_m = v^3 \text{tr}(MU + M^\dagger U^\dagger) = -\frac{v^3}{f_\pi^2} \text{tr}(M) \pi^a \pi^a + \dots, \quad (2.8)$$

from which the mass of the pions, $m_\pi^2 = 2v^3 \text{tr}(M)/f_\pi^2$, also known as the Gellmann-Oakes-Renner relation, can be read off. According to this relation, all pions have the same mass but the mass degeneracy is removed when electromagnetism is incorporated. To see this, the derivative is promoted into a covariant derivative in order to incorporate the minimal coupling to the photon. The resulting $AA\pi\pi$ and $A\pi\pi$ interactions induce an additional contribution to the mass of the charged pion of the form $\delta m_{\text{EM}}^2 \sim e^2 \text{tr}(QUQU)$ at one-loop order, while the neutral pion is unaffected.

This is also apparent from symmetry considerations. At the classical level, the

minimal coupling preserves a $U(1)_L \times U(1)_R$ subgroup of the $SU(2)_L \times SU(2)_R$ chiral symmetry (these $U(1)$ groups have nothing to do with $U(1)_V$ or $U(1)_A$). Thus, after chiral symmetry breaking, the mass of two mesons will be protected from further contributions from electromagnetism, namely the neutral pion and the η' meson, which will be discussed in a moment.

This whole song and dance can be easily generalized to the case with $n_F = 3$, i.e. including the strange-quark. The flavor symmetry is then given by $U(3)_L \times U(3)_R$ and the additional Goldstone bosons are the four kaons and the η -meson. Note that $SU(3)$ has two Casimirs in contrast to $SU(2)$, which has only one (in general, $SU(N)$ has $N - 1$ Casimirs since the amount of Casimirs is equal to the dimension of the center for simple and semi-simple Lie groups). One of these Casimirs is the isospin I arising from the $SU(2)$ subgroup, while the other one is the strangeness S . Adding the quark masses results in heavier masses for the kaons due to the heavier strange quark. The electromagnetic coupling again produces an additional contribution to the masses of the charged particles via one-loop effects. In the symmetry language, the only difference is that the preserved subgroup of the chiral symmetry $SU(3)_L \times SU(3)_R$ is now $(U(1) \times SU(2))_L \times (U(1) \times SU(2))_R$, leading to five meson masses without electromagnetic contribution: $\pi^0, K^0, \bar{K}^0, \eta, \eta'$, i.e. all the neutral particles. In what follows, we work with an arbitrary amount of flavors n_F .

2.1.3 The $U(1)_A$ Problem

The second big question regards the $U(1)_A$ symmetry [56]. It was not known at that time that $U(1)_A$ was anomalous, so a ninth Goldstone had to be included in the expansion around the VEV, i.e.

$$U(x) = \exp\left(\frac{i2\pi^a(x)T^a}{f_\pi} + \frac{i\pi^9}{f_9}\right), \quad (2.9)$$

where $f_\pi \neq f_9$ since there is no symmetry forcing them to be equal. Due to $U(x)$ now being an element of $U(2)_V$ instead of $SU(2)_V$, another invariant exists that must be included in the effective Lagrangian, leading to

$$\begin{aligned} \mathcal{L} &\supset \frac{1}{4}f_\pi^2 \text{tr}(\partial_\mu U^\dagger \partial^\mu U) + \frac{1}{4}F^2 \partial_\mu (\det U^\dagger) \partial^\mu \det U \\ &= \frac{1}{2} \partial_\mu \pi^a \partial^\mu \pi^a + \frac{1}{2} \partial_\mu \pi^9 \partial^\mu \pi^9 + \dots, \end{aligned} \quad (2.10)$$

where $F^2 = (2f_9^2 - n_F f_\pi^2)/n_F^2$ is fixed by canonical normalization. Again activating the quark masses leads to a mass of all the (pseudo-)Goldstone bosons. Diagonalizing the mass matrix to identify the mass eigenstates, it can be shown that the mass of the η' as the ninth Goldstone cannot exceed $m_{\eta'} \lesssim \sqrt{3}m_\pi$. However, experimentally it is known that $m_{\eta'} \gg m_\pi$. In addition, lattice calculations reveal that $m_{\eta'} \rightarrow 0$ as $m_{u/d} \rightarrow 0$, which should be the case as a consequence of Goldstone's theorem. The absence of the ninth Goldstone boson or, in other words, the η' being too heavy, was called the $U(1)_A$ problem by Weinberg [56].

It did not take long until 't Hooft [57–59] realized that $U(1)_A$ was not a symmetry as a consequence of the Adler–Bell–Jackiw anomaly [70–72], i.e.

$$\partial_\mu J_5^\mu = \frac{n_F g_s^2}{16\pi^2} G_{\mu\nu}^a \tilde{G}^{a\mu\nu}, \quad (2.11)$$

where J_5^μ is the current of $U(1)_A$ and n_F is the number of active quarks. The associated shift in the Lagrangian was, like the θ -term in (2.1), a boundary term and it was initially treated like a mathematical artifact. He showed, however, that non-perturbative configurations in the spectrum of QCD, named instantons [73], make the anomaly physical by inducing an explicit breaking term in the effective Lagrangian. This term was coined the 't Hooft vertex or 't Hooft determinant. Consequently, η' receives an additional contribution to its mass that does not vanish in the chiral limit,

$$\mathcal{L}'_{\text{tHooft}} = -\kappa \det(\bar{\Psi}_R \Psi_L) + \text{h.c.} \longrightarrow -\kappa e^{-i\frac{\eta'}{f_{\eta'}}} \langle \det(\bar{\Psi}_R \Psi_L) \rangle + \text{h.c.} . \quad (2.12)$$

Here, κ is a constant that was calculated by 't Hooft [58] and will be discussed later.

In summary, 't Hooft's instanton mechanism solved the $U(1)_A$ problem by realizing that the Adler–Bell–Jackiw anomaly has physical effects due to non-perturbative configurations called instantons. Thus, it can be said that instantons generate a correlation between the boundary and the bulk of our spacetime. This statement should make us question our naive conclusion about the θ -term in (2.1). Since instantons are treated explicitly in the next section, the remainder of this section will assume their existence and look at the θ -term using chiral perturbation theory.

2.1.4 The Role of θ

Let us now perform chiral symmetry breaking but without neglecting the θ -term to see its consequences. For that consider the full QCD Lagrangian, i.e.

$$\mathcal{L}_{\text{QCD}} = i\bar{\Psi}\gamma^\mu D_\mu\Psi - \bar{\Psi}M\Psi - \frac{1}{4}G_{\mu\nu}^a G^{a,\mu\nu} + \theta\frac{g_s^2}{32\pi^2}G_{\mu\nu}^a \tilde{G}^{a\mu\nu}, \quad (2.13)$$

where M is without loss of generality chosen to be diagonal and the quark masses are denoted by m_f . It is important to note that massive fermions give rise to another contribution to the vacuum angle. The mass appearing in the mass term of the Lagrangian can in general be complex but usually one can get rid of the phase by a chiral redefinition,

$$\Psi_f \longrightarrow e^{i\alpha_f\gamma_5}\Psi_f \quad \text{or} \quad \begin{aligned} \Psi_{L f} &\longrightarrow e^{-i\alpha_f}\Psi_{L f} \\ \Psi_{R f} &\longrightarrow e^{+i\alpha_f}\Psi_{R f} \end{aligned} .$$

This redefinition, however, is anomalous in the presence of gauge fields and hence affects the θ -term via the chiral anomaly [74, 75],

$$\begin{aligned}\theta &\longrightarrow \theta + 2 \sum_f \alpha_f, \\ \arg m_f &\longrightarrow \arg m_f + 2\alpha_f.\end{aligned}\tag{2.14}$$

Exploiting this through a suitable redefinition of the quarks, the physics encoded in the QCD vacuum angle can be transferred to the fermion sector. Vice versa the physics encoded in the phase of the quark masses can be moved to the gluon sector. In either case, since a redefinition of integration variables in the path integral, such as the fermionic fields, cannot change physics, physical quantities cannot depend on θ or M separately but only on the combination

$$\exp(-i\theta) \prod_f m_f \quad \text{or} \quad \begin{aligned}\bar{\theta} &= \theta - \sum_f \arg m_f \\ &= \theta - \arg \det M\end{aligned}\tag{2.15}$$

In the following, let us choose a redefinition that moves the vacuum angle to the quark masses and also ensures that both quark masses have an equal phase. This is achieved by $\alpha_f = -\arg m_f/2 - \bar{\theta}/2n_f$, so that

$$M \longrightarrow e^{-i\bar{\theta}/n_F} M\tag{2.16}$$

Regarding chiral symmetry breaking, the crucial point of having $\bar{\theta}$ in the quark sector is that the minimum is no longer given by the identity. Taking $n_F = 2$ for simplicity and minimizing the potential, instead yields that the minimum is located at

$$V = \begin{pmatrix} e^{i\phi} & \\ & e^{-i\phi} \end{pmatrix}, \quad \text{with} \quad \phi \equiv \arctan \left(\frac{m_u - m_d}{m_u + m_d} \tan \frac{\bar{\theta}}{2} \right).\tag{2.17}$$

The most convenient way to expand around this minimum is via

$$U = \begin{pmatrix} e^{i\frac{\phi}{2}} & \\ & e^{-i\frac{\phi}{2}} \end{pmatrix} \exp \left(\frac{2i\pi^a(x)T^a}{f_\pi} + \frac{i\pi^9}{f_9} \right) \begin{pmatrix} e^{i\frac{\phi}{2}} & \\ & e^{-i\frac{\phi}{2}} \end{pmatrix} \equiv \sqrt{V} U_0 \sqrt{V},\tag{2.18}$$

where U_0 is the solution when $\bar{\theta} = 0$. The final chiral Lagrangian including the 't Hooft

determinant is then

$$\begin{aligned}
\mathcal{L} &= \frac{1}{4} f_\pi^2 \text{tr}(\partial_\mu U_0^\dagger \partial^\mu U_0) + \frac{1}{4} F^2 \partial_\mu (\det U_0^\dagger) \partial^\mu \det U_0 + \kappa (\det U_0 + \text{h.c.}) \\
&\quad + v^3 \text{tr}(e^{-i\theta/2} M V U_0 + \text{h.c.}) \\
&= \frac{1}{2} \partial_\mu \pi^3 \partial^\mu \pi^3 + \frac{1}{2} \partial_\mu \pi^9 \partial^\mu \pi^9 + \kappa \cos\left(\frac{\pi^9}{f_9}\right) \\
&\quad + 2m_u v^3 \cos\left(\phi - \frac{\bar{\theta}}{2} + \frac{\pi^3}{f_\pi} + \frac{\pi^9}{f_9}\right) + 2m_d v^3 \cos\left(\phi + \frac{\bar{\theta}}{2} + \frac{\pi^3}{f_\pi} - \frac{\pi^9}{f_9}\right), \quad (2.19)
\end{aligned}$$

where in the second step any constants in the 't Hooft determinant have been absorbed into κ and only the neutral pion π^3 has been taken into account. The mass eigenstates η' and π^0 can be deduced by diagonalizing the mass term that emerges from expanding the cosines. We parameterize these eigenstates by $\eta'/f_{\eta'} \pm \pi^0/f_{\pi^0} \equiv \pi^3/f_\pi \pm \pi^9/f_9$ with suitable definitions of $f_{\eta'}$ and f_{π^0} and continue to work with the non-expanded cosines. The mass term stemming from the quark masses can then be decomposed using trigonometric identities, yielding the potential

$$\begin{aligned}
V(a, \pi^0) &= -2v^3 \left[m_u \cos\left(\frac{\eta'}{f_{\eta'}} + \phi - \frac{\bar{\theta}}{2}\right) + m_d \cos\left(\frac{\eta'}{f_{\eta'}} + \phi + \frac{\bar{\theta}}{2}\right) \right] \cos\left(\frac{\pi^0}{f_{\pi^0}}\right) \\
&\quad - 2v^3 \left[m_u \sin\left(\frac{\eta'}{f_{\eta'}} + \phi - \frac{\bar{\theta}}{2}\right) + m_d \sin\left(\frac{\eta'}{f_{\eta'}} + \phi + \frac{\bar{\theta}}{2}\right) \right] \sin\left(\frac{\pi^0}{f_{\pi^0}}\right) \\
&\quad - \kappa \cos\left(\frac{\eta'}{f_{\eta'}}\right). \quad (2.20)
\end{aligned}$$

This is the key relation of this section from which a couple of interesting observations can be made about the vacuum of the theory. Defining $z \equiv m_u/m_d$ and taking $\bar{\theta} \ll 1$ for a simpler relation between ϕ and $\bar{\theta}$ (experimentally, this turns out to be a very good approximation as will be explained in Sec. 2.4), the vacuum energy density can be obtained by setting the two fields on their ground states $\eta' = 0$ and $\pi^0 = 0$, which yields

$$\epsilon_0(\bar{\theta}) - \epsilon_0(0) = -2v^3 \left[m_u \cos\left(\frac{m_d \bar{\theta}}{m_u + m_d}\right) + m_d \cos\left(\frac{m_u \bar{\theta}}{m_u + m_d}\right) \right]. \quad (2.21)$$

From this expression, it is apparent that the presence of $\bar{\theta}$ leads to a reduction in the vacuum energy. This is a sign that $\bar{\theta}$ is related to tunneling since these transitions are known to reduce the vacuum energy. The relation to tunneling will become clearer in Sec. 2.2 where the role of $\bar{\theta}$ is analyzed using instantons. Furthermore, the lowest value is located at $\bar{\theta} = 0$, showing that this value for the vacuum angle is very special. This statement will appear many times in this chapter.

For the second interesting observation, we set one of the quark masses to zero, say m_u . This causes $\bar{\theta}$ to vanish from the Lagrangian and, consequently, from the vacuum energy density. What happened? The key point can be extracted by defining the (unphysical) field $\tilde{\eta}' \equiv \phi + \eta'$. In (2.20), the term proportional to m_u vanishes and in the term proportional to m_d we have expanded around the minimum ϕ . So effectively

what happened is that $\tilde{\eta}'$ acquired the VEV ϕ . According to (2.17) with $m_u = 0$, the VEV is now located at $\phi = -\bar{\theta}/2$, thus being located exactly where θ vanishes from the Lagrangian. This is also not changed when nucleons are also taken into account. In that case the $\tilde{\eta}'$ boson will again appear at the same term as $\bar{\theta}$, thus acquiring a VEV that after expansion around it removes $\bar{\theta}$.

According to (2.15), it is necessary that $\bar{\theta}$ vanishes in the presence of a massless quark. However, chiral perturbation theory can be used to reveal the special role η' plays in this, namely the role of what is called the axion. It renders $\bar{\theta}$ dynamical so that it relaxes to the true minimum at $\bar{\theta} = 0$. We will discuss the axion in detail in Chapter 3.

2.2 Instantons

The $U(1)_A$ problem was resolved by 't Hooft's instanton mechanism [57–59], which revealed that instantons [73] make the Adler–Bell–Jackiw anomaly [70–72] physical. In Sec. 2.1 we included the emerging 't Hooft determinant without any derivation to analyze some features of θ and reveal the special role η' plays. The goal of this section is not only to show how the 't Hooft determinant emerges but also to discuss the physics behind the θ -term from a more conceptual point of view. This section is mainly based on [76], further sources and a far more details on this topic can be found there and in the references cited therein.

2.2.1 Semi-Classical Theory of Tunneling

The instantonic field configuration is deeply connected with the notion of tunneling. Before treating instantons in YM theories, it is helpful to start with a simpler example to get the general idea. Consider the simple example of a real scalar field Φ in $1 + 1$ dimensions with a double-well potential. Furthermore, we change to the Euclidean formulation by Wick rotating to imaginary time, $x_0 = -ix_2$, such that the path integral becomes

$$Z = \int [D\Phi] \exp(-S_E[\Phi]) , \quad (2.22)$$

with

$$S_E[\Phi] = \int d^2x_E ((\partial_i\Phi)^2 + V(\Phi)) . \quad (2.23)$$

The reason we go to Euclidean spacetime is because tunneling processes are inherently quantum – we cannot rely on our minimal action principle since the solutions of the EOMs represent classical trajectories in our field theoretic framework. However, in the Euclidean formulation the potential is flipped, $V(\Phi) \rightarrow -V(\Phi)$, meaning that the minima become maxima and vice versa. Thus, the solution of the EOM that corresponds to a tunneling process between the two minima in Minkowski spacetime becomes a motion through a classical region in Euclidean spacetime – we can rely on our minimal action principle again. This non-perturbative solution is called an instanton. It is important to

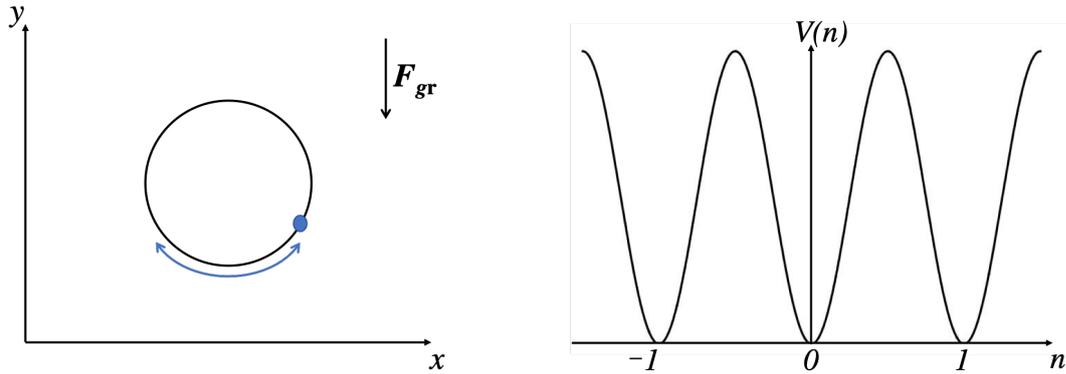


Figure 1: A Simple system to understand the role of topology in YM theories. Left: A particle moving along an upright circle and subject to a constant gravitational force. Right: Potential of this system with n denoting the minima. Classically, the particle oscillates around one of the vacua. Quantum mechanically, the vacuum is a superposition of all the minima, inducing non-trivial effects.

note that, as solutions of the Euclidean EOMs, instantons merely represent the leading order contribution to tunneling processes. The higher contributions corresponding to quantum corrections, however, can become very important in systems that are inherently quantum. This makes instantons merely a tool to understand tunneling processes in the semi-classical limit $\hbar \rightarrow 0$.

2.2.2 Tunneling in YM Theory and the Role of Topology

With this simple example in 1+1 dimensions, let us move on to YM theories. For simplicity, we consider the pure YM Lagrangian with gauge group $SU(2)$ and without the θ -term, i.e.

$$\mathcal{L}(A) = -\frac{1}{4}G_{\mu\nu}^a G^{a\mu\nu}. \quad (2.24)$$

It is not obvious that this theory has non-trivial tunneling processes because the direction in which the tunneling is taking place is hidden in the infinite dimensional space of fields and the potential along this direction is a result of non-trivial topology [76].

An excellent analogy to understand the role of topology in YM theories is given by a particle moving along an upright circle and subject to a constant gravitational force (Fig. 1, left panel). Classically the system is not very interesting since the particle will oscillate around the bottom of the circle. Quantum mechanically, however, the story changes due to the periodicity of the potential (Fig. 1, right panel). The true vacuum of the system is not one of the single, so-called perturbative vacua but a superposition of all of them. This is independent of the fact that all the perturbative vacua lie at the same position in the physical position space, i.e. the bottom of the circle. Therefore, the naive perturbative treatment is not complete. We need to take into account the influence of tunneling through the barrier induced by the non-trivial topology to capture the whole picture.

To show the existence of such tunneling processes in YM theories, we change into the Hamiltonian formulation. While the classical Hamiltonian is easy to find via Legendre transformation of the action, the vanishing momentum conjugate of the scalar potential leads to a vanishing commutator and thus makes canonical quantization problematic, just like in QED. This problem happens as we have included non-physical polarizations at the beginning. An easy fix to this, is adding a term to the action that fixes the gauge.³ But then the next problem appears. Picking the Lorenz gauge due to being Lorentz invariant, does not fix the gauge freedom completely and results in negative norm states coming from the time-like polarization. Thus, we are either left with the Gupta-Bleuler formalism of enforcing the Lorenz gauge merely on the physical states $\langle \Psi | \partial_\mu A^\mu | \Psi \rangle$, or using the non-Lorentz covariant temporal gauge $A_0 = 0$ and imposing Gauss's law $\text{div} E^a = \rho^a$ by hand since it no longer appears as an equation of motion. Both of these solutions are rather impractical to perform calculations in the case of YM theory (in contrast to e.g. QED where it is to some degree possible) but for the purpose of this section it is sufficient. In the following, we pick the latter but do not impose Gauss's law so that we work with an undergauged theory. The corresponding Hamiltonian in the temporal gauge reads

$$\mathcal{H} = \frac{1}{2} \int d^3x (E_i^a E^{ai} + B_i^a B^{ai}) . \quad (2.25)$$

First we need to identify the zero-energy states of this theory. Naively, one would expect that these are given by $A_i(\vec{x}) = 0$ but due to gauge freedom the zero-energy states are pure gauge configurations

$$A_i(\vec{x})|_{\text{vac}} = \frac{i}{g} U(\vec{x}) \partial_i U^\dagger(\vec{x}) , \quad (2.26)$$

where $U(\vec{x})$ belongs to $\text{SU}(2)$. For the corresponding action to be finite, we must have

$$\lim_{|\vec{x}| \rightarrow \infty} U(\vec{x}) = \text{const} . \quad (2.27)$$

This boundary conditions compactifies our three-dimensional space, making it topologically equivalent to S^3 . Since the group manifold of $\text{SU}(2)$ is also S^3 , the functions $U(\vec{x})$ realize a mapping $S^3 \rightarrow S^3$. These mappings can intuitively be classified according to the number of times one sphere is wound around the other sphere or, in other words, the functions $U(\vec{x})$ are forming distinct classes $U_n(\vec{x})$ that are labeled by the winding number $n \in \mathbb{Z}$.

The crucial point here is that matrices belonging to a class with given n can be related to each other via continuous gauge transformations. At the same time, no continuous gauge transformation can change n . In order to pass from one value of n to another, one must pass through configurations that are not pure gauge. In other words, there is an energy barrier between the field configurations with different values on n . Thus, n is a

³Such terms look like they explicitly break the gauge symmetry, but gauge symmetries can never be broken due to not being a real symmetry but a redundancy in the description. This becomes manifest in the Stückelberg formulation.

topological quantum number that labels the distinct vacua of the theory, similar to the right panel of Fig. 1.

Mathematically, the winding number is called the Pontryagin-index and it being non-trivial follows from the third homotopy group of the three-sphere being non-trivial,

$$\pi_3(S^3) = \mathbb{Z} . \quad (2.28)$$

Furthermore, there exists an expression for the winding number in terms of the gauge group elements $U(\mathbf{x})$, i.e.

$$n = -\frac{1}{24\pi^2} \int d^3x \epsilon^{ijk} \text{tr} \left[(U\partial_i U^\dagger)(U\partial_j U^\dagger)(U\partial_k U^\dagger) \right] , \quad (2.29)$$

where the integral goes over any coordinates parametrizing the three-sphere. It can explicitly be checked that this expression is locally invisible and additive, thus satisfying all the properties of a topological quantum number.

The existence of many vacua as a consequence of non-trivial mappings between S^3 and the gauge group $SU(2) \simeq S^3$ can make the impression that it is only present in theories with an $SU(2)$ gauge symmetry, but this is not true. There is a theorem by Raoul Bott that states that any continuous mapping of S^3 into a simple Lie group G can be continuously deformed to a mapping into an $SU(2)$ subgroup of G . This means that every conclusion made so far and every mathematical equality including numerical factors equally apply for any simple group G , in particular $SU(N)$.

With this expression for the winding number, the existence of instantons in YM theories can be shown by finding solutions of the Euclidean EOMs with the boundary condition

$$\lim_{x_4 \rightarrow \pm\infty} \lim_{|\mathbf{x}| \rightarrow \infty} A_i(\vec{x}, x_4) = \frac{i}{g} U_\pm(\vec{x}, x_4) \partial_i U_\pm^\dagger(\vec{x}, x_4) , \quad (2.30)$$

where $U_\pm(\vec{x}, x_4)$ have winding number n_\pm and all quantities are in Euclidean formulation (see Appendix A.7). Using (2.29), the total change in the winding number can be expressed as [68]

$$\begin{aligned} \nu \equiv n_+ - n_- &= \frac{ig^3}{24\pi^2} \int dS_\mu \epsilon^{\mu\nu\rho\sigma} \text{tr} [A_\nu A_\rho A_\sigma] = \frac{g^2}{32\pi^2} \int dS_\mu \tilde{C}^\mu \\ &= \frac{g^2}{32\pi^2} \int d^4x_E \partial_\mu \tilde{C}^\mu = \frac{g^2}{32\pi^2} \int d^4x_E E , \end{aligned} \quad (2.31)$$

where the following topological quantities have been used,

$$\begin{aligned} C_{\mu\nu\rho} &\equiv 2\text{tr} \left[A_\mu G_{\nu\rho} - \frac{2}{3} ig A_\mu A_\nu A_\rho \right] , \\ \tilde{C}^\mu &\equiv \epsilon^{\mu\nu\rho\sigma} C_{\nu\rho\sigma} , \\ E &\equiv \partial_\mu \tilde{C}^\mu = G_{\mu\nu}^a \tilde{G}^{a\mu\nu} . \end{aligned} \quad (2.32)$$

The first object is the Chern-Simons three-form. The second object is the dual of the

Chern-Simons three-form, called the Chern-Simons current. From the last expression of the first row in (2.31) it follows that this is the direction in field space where the tunneling is taking place (again see right panel of Fig. 1 for the analogy). The last object is the Pontryagin density, which by its definition expresses the (non-)conservation of the Chern-Simons current.

From the last equality in (2.31) it follows that for a nontrivial value of ν , it is sufficient to find a configuration for which the integral over the Pontryagin density does not vanish. For this purpose, the Bogomol'nyi trick is applied, i.e.,

$$\begin{aligned} S_E[A] &= -\frac{1}{4} \int d^4x_E G_{ij}^a G^{a ij} = -\frac{1}{8} \int d^4x_E (G_{ij}^a \pm \tilde{G}_{ij}^a)^2 \mp \frac{1}{4} \int d^4x_E E \\ &\geq \mp \frac{1}{4} \int d^4x_E E = \frac{8\pi^2}{g^2} \nu . \end{aligned} \quad (2.33)$$

The equality holds iff the gauge field is dual or self-dual, i.e.

$$G_{ij}^a = \pm \tilde{G}_{ij}^a . \quad (2.34)$$

Any solution of this first-order equation is a minimum of the action $S[A]$ and thus a solution of the Yang-Mills EOM. This also immediately follows from the Bianchi identity. The corresponding action given by (2.33) is the instanton action.

For $\nu = 1$ an explicit solution can be found. This solution is the so called BPST instanton, named after Belavin, Polyakov, Schwarz, and Tyupkin and it takes the form

$$A_\mu^a(x) = \frac{2}{g} \eta_{a\mu\nu} \frac{(x - x_0)_\nu}{(x - x_0)^2 + \rho^2} , \quad (2.35)$$

where x_0 denotes the location of the instanton, ρ denotes its size, and $\eta_{a\mu\nu}$ is called the 't Hooft symbol,

$$\eta_{a\mu\nu} = \begin{cases} \epsilon_{a\mu\nu}, & \mu, \nu = 1, 2, 3 , \\ -\delta_{a\nu}, & \mu = 4 , \\ \delta_{a\mu}, & \nu = 4 , \\ 0, & \mu, \nu = 4 . \end{cases} . \quad (2.36)$$

It is apparent that A_μ^a falls off as slowly as $1/x$ but what is most intriguing about this solution is that the color and Lorentz indices are entangled. This is a consequence of the instanton solution spontaneously breaking both global color and Lorentz symmetry, but preserving a diagonal combination.

The BPST instanton is the most important solution for two reasons. First, it is the leading order configuration since higher ν are suppressed as follows from the instanton action (2.33). Secondly, multi-instantons with $\nu \neq 1$ can be expressed in terms of many BPST instantons. However, it is not the most general, since all instantons have the same orientation in color space. The most general solution – the so-called ADHM construction named after Atiyah, Drinfel'd, Hitchin and Manin – assigns 8 Moduli per instanton for

SU(2): 4 for its position, 1 for its size, and 3 orientations in color-Lorentz-space. In the general case of SU(N) there are $4N$ moduli per instanton and so $4N|\nu|$ moduli altogether in the multi-instanton solution.

2.2.3 The YM Vacuum

The existence of instantons has dramatic consequences. According to quantum mechanics, when a theory has degenerate vacua separated by a finite energy barrier, the actual vacuum state is described by a superposition of the degenerate “pre-vacua”. The necessity for this can be seen by looking at the tunneling amplitude between two vacua $|n\rangle$ and $|n'\rangle$, which is generically given by

$$\langle n|H|n'\rangle = e^{-S_E} . \quad (2.37)$$

Here H denotes the Hamiltonian and S_E the Euclidean action of the configuration that mediates between the vacua. Since the vacuum of a consistent theory is per definition (quasi-)stable, the Hamiltonian must be diagonalized to give the true vacuum.

In QFT this is slightly different. Since every point in space must perform this tunneling procedure, the tunneling amplitude must be considered in the infinite volume limit. Thus, the true vacuum is only different from a pre-vacuum if the action of the configuration, which mediates between the different pre-vacua, does not vanish in the infinite volume limit. In YM theory the vacua $|n\rangle$ correspond to different winding numbers and the tunneling amplitude (2.37) is given by

$$\langle n|H|n'\rangle = e^{-|\nu|\frac{g^2}{8\pi^2}} . \quad (2.38)$$

Since the action is independent of x , it does not vanish in the infinite volume limit and the true vacuum of the theory is a superposition of $|n\rangle$, i.e.

$$|\theta\rangle = \sum_{n=-\infty}^{\infty} e^{in\theta} |n\rangle . \quad (2.39)$$

This is called a θ -vacuum. The phase including the parameter θ results from imposing the periodic Bloch boundary condition on the wave function, i.e. $|n+1\rangle = \exp(i\theta)|n\rangle$, which is the most natural boundary condition to impose on a periodic system.

As a result of the periodic boundary condition, the vacuum is labeled by a parameter θ . This has the interesting property that for any operator \mathcal{O} acting on the physical Hilbert space,

$$\langle \theta|\mathcal{O}|\theta'\rangle = 0 . \quad (2.40)$$

This means that the Hilbert spaces are orthogonal or, in other words, different values of θ correspond to different superselection sectors. This makes the value of θ a global constant that once fixed cannot be changed.

The need for θ -vacua can also be seen in a path integral formulation [77]. Per definition, any mediation between the different winding numbers must be included in the path integral. Denoting the corresponding weight factor with $f(\nu)$, the path integral becomes

$$Z = \sum_{\nu} f(\nu) \int [DA] e^{iS[A]}, \quad (2.41)$$

Next consider the expectation value of a local observable \mathcal{O} placed within a large Euclidean volume Ω ,

$$\langle \mathcal{O} \rangle = \frac{\sum_{\nu} f(\nu) \int_{\nu} D\Phi e^{iS_{\Omega}} \mathcal{O}}{\sum_{\nu} f(\nu) \int_{\nu} D\Phi e^{iS_{\Omega}}}. \quad (2.42)$$

Suppose that Ω is split into two (very large) volumes Ω_1 and Ω_2 , with \mathcal{O} in the volume Ω_1 . In this case the expectation value in (2.42) becomes

$$\langle \mathcal{O} \rangle = \frac{\sum_{\nu_1, \nu_2} f(\nu_1 + \nu_2) \int_{\nu_1} D\Phi e^{iS_{\Omega_1}} \mathcal{O} \int_{\nu_2} D\Phi e^{iS_{\Omega_2}}}{\sum_{\nu_1, \nu_2} f(\nu_1 + \nu_2) \int_{\nu_1} D\Phi e^{iS_{\Omega_1}} \int_{\nu_2} D\Phi e^{iS_{\Omega_2}}}. \quad (2.43)$$

For the cluster decomposition principle to hold, which essentially is the requirement of S -matrix locality (or, in other words, that distant experiments are uncorrelated), the factors including Ω_2 must cancel. This is only possible if

$$f(\nu_1 + \nu_2) = f(\nu_1)f(\nu_2), \quad (2.44)$$

which is uniquely solved by

$$f(\nu) = e^{i\theta\nu}. \quad (2.45)$$

Plugging in the definition of ν from (2.31) results in the appearance of the θ -term in the Lagrangian,

$$\mathcal{L}_{\theta} = \theta \frac{g^2}{32\pi^2} G_{\mu\nu}^a \tilde{G}^{a\mu\nu}. \quad (2.46)$$

Even though the θ -term in the Lagrangian has so far been neglected in this whole section, once instantons were properly taken into account, the term appeared anyway. This shows that, while the θ -term is irrelevant in the classical YM Lagrangian due to being a boundary term, the quantum theory makes its inclusion a necessity. However, the θ -term can even classically have implication in some theories. For instance, this is the case for QED in the presence of a magnetic monopole. On this background, the θ -term cannot be neglected since there is a classical configuration creating a “new boundary” that is not located at infinity. Thus, the correct conclusion would be that θ is never zero. It is at most irrelevant on a chosen background.

2.2.4 The YM Vacuum Energy

After establishing that the true YM vacuum is parametrized by the global parameter θ , the question is how the vacuum energy density depends on θ . Using the path integral, a few general properties can be derived easily [55, 78]:

- $E(\theta) = E(\theta + 2\pi)$
- $E(0) \leq E(\theta)$
- $E(\theta) = E(-\theta)$

These properties also hold when fermions are included. While in the last section the vacuum energy was calculated in chiral perturbation theory, in this section the vacuum energy will be calculated using instanton calculus. For this purpose a weight factor encoding the influence of the instantons in the functional integral will be derived – this is called the instanton measure. The calculation was first performed by 't Hooft [58, 76] and is straight forward but cumbersome. Here we will outline the procedure and estimate the final result. All quantities are still in the Euclidean formulation.

To include quantum fluctuations of the gauge field near the instanton, the following expansion is performed,

$$A_\mu^a = (A_{\text{inst}})_\mu^a + a_\mu^a . \quad (2.47)$$

Up to quadratic order in this expansion, the action is then given by

$$S_E[A] = \frac{8\pi^2}{g^2} - \frac{1}{2} \int d^4x_E a_\mu^a L^{ab\mu\nu} [(A_{\text{inst}})_\mu^a] a_\nu^b , \quad (2.48)$$

where the zeroth order, which we will denote as S_0 in the following, is the instanton action (2.33) and

$$L^{ab\mu\nu} [(A_{\text{inst}})_\mu^a] = (D^2 \delta^{\mu\nu} - D^\mu D^\nu) \delta^{ab} - g \epsilon^{abc} G_{\text{inst}}^{c\mu\nu} . \quad (2.49)$$

Integrating out the small fluctuations of the gauge field around the instanton, naively results in $(\det L)^{-1/2}$, but this is where the technicalities need to be taken into account. These are the following:

Gauge-Fixing Fixing the gauge in order to quantize introduces ghosts and therefore an additional ghost determinant. We will absorb this contribution into an overall constant.

Non-zero modes The product of the non-zero eigenmodes diverges and thus needs regularization, where usually the Pauli-Villars regulator M_{UV} is used. This effectively means

$$\det L \rightarrow \frac{\det L}{\det(L + M_{\text{UV}}^2)} \quad (2.50)$$

Zero-modes The zero-modes must be excluded from $\det L^{-1/2}$ to avoid infinities. Instead, each zero-mode gives rise to an integral over the corresponding modulus times a Jacobian, which produces $\sqrt{S_0}$ per collective coordinate. Additionally, each Jacobian will be accompanied by M_{UV} , due to the exclusion of the zero-modes after regularization.

As already mentioned, for $\text{SU}(2)$ the general instanton solution has 8 collective coordinates: x_0 (the position of its center), ρ (the instanton size), and three Euler angles

θ, ϕ, ψ , which specify the orientation of the instanton in either the color $SU(2)$ or the dotted $SU(2)$ of the Lorentz group $SU(2) \times SU(2)$. Performing the above steps yields [58, 76]

$$\begin{aligned} d\mu_{\text{inst}} &= \text{const} \times \left(M_{\text{UV}} S_0^{1/2}\right)^8 \int d^4 x_0 d\rho \rho^3 \exp(-S_0 + \Phi_1) \int d\phi d\theta d\psi \sin\theta \\ &= \text{const} \times \left(-\frac{8\pi^2}{g^2}\right)^4 \int \frac{d^4 x_0 d\rho}{\rho^5} \exp\left(\frac{8\pi^2}{g^2} + 8 \ln(M_{\text{UV}} \rho) + \Phi_1\right), \end{aligned} \quad (2.51)$$

where in the second step the integration over the Euler angles has been absorbed into the overall constant and the Pauli-Villars regulator M_{UV} has been moved to the exponent. The function Φ_1 parametrizes the non-zero mode contribution and is given by

$$\Phi_1 = -\frac{2}{3} \ln(M_{\text{UV}} \rho) + \text{const}. \quad (2.52)$$

Plugin this into the expression for the instanton measure (2.51), the prefactor of the logarithm becomes the renormalization coefficient $b_0 = 8 - 2/3 = 22/3$. Thus, the whole term can be absorbed by the coupling to give the renormalized coupling $g(\rho)$. The final expression for $SU(2)$ can then be written as

$$d\mu_{\text{inst}} \equiv \int \frac{dx_0^4 d\rho}{\rho^5} d(\rho), \quad (2.53)$$

$$d(\rho) = \text{const} \times \left(\frac{8\pi^2}{g^2}\right)^4 \exp\left(-\frac{8\pi^2}{g^2(\rho)}\right), \quad (2.54)$$

where $d(\rho)$ is called the instanton density. The reason why the renormalized coupling appears only in the exponent is because only the first order quantum corrections have been included. A careful calculation to second order would result in the appearance of the renormalized coupling also in the non-exponential factor, as expected from the renormalizability of Yang Mills theory.

It should be noted that the ρ dependence of the instanton density at leading order is given by $d(\rho) \sim \rho^{b_0}$, which is a step function for large values of ρ . Since at large ρ the gauge coupling constant becomes strong, the control over the theory is lost and semi-classical methods are no longer reliable. This is the reason why instantons are of no help in solving the confinement problem in QCD.

The generalization to $SU(N)$ is straight forward. The relation between $d\mu_{\text{inst}}$ remains the same, while the power of the prefactor becomes $2N$ due to the presence of $4N$ zero-modes. Furthermore, the ρ dependence of the instanton density becomes $d(\rho) \sim \rho^{b_0} = \rho^{11N/3}$, which is an even steeper function for large values of ρ .

While one instanton has been considered so far, it is necessary to analyze an ensemble of n_+ instantons and n_- anti-instantons in order to describe the YM vacuum. In general, there are highly non-trivial interactions between the instantons and the other degrees of freedom that need to be taken into account. Instead, the dilute instanton gas approximation [79] will be used in the following, which assumes the (anti-)instantons to

be well separated and unaffected by each other. Even though the instanton gas is not a good approximation in real QCD due to the presence of large instantons, it provides a basic intuition.

In the large 4-volume limit, the relation between the energy density of the QCD vacuum and the Euclidean path integral $Z_E(\theta)$ is given by [80]

$$Z_E(\theta) = \lim_{V_4 \rightarrow \infty} \exp(-V_4 \rho(\theta)) . \quad (2.55)$$

At the same time the generating functional can be brought to the following form by taking the (anti-)instanton measure to be dominated by the instantons of size $\rho \sim 1/\Lambda_{\text{QCD}}$,

$$\begin{aligned} Z_E(\theta) &= \sum_{n_+, n_-} \prod \int d\mu_{\text{inst}} \int d\mu_{\text{anti-inst}} \\ &= \sum_{n_+} \frac{1}{n_+!} (K \Lambda_{\text{QCD}}^4 V_4)^{n_+} e^{-\frac{8\pi^2 n_+}{g^2(\Lambda_{\text{QCD}})}} e^{i\theta n_+} \times (\text{anti-instanton}) \end{aligned} \quad (2.56)$$

$$= \exp \left(2K \Lambda_{\text{QCD}}^4 V_4 e^{-\frac{8\pi^2}{g^2(\Lambda_{\text{QCD}})}} \cos \theta \right) , \quad (2.57)$$

where K encodes all the constants appearing in the evaluation of the instanton measure and the θ -term was included. Hence, the energy density of the QCD vacuum depends on θ via

$$\rho(\theta) = -2K \Lambda_{\text{QCD}}^4 e^{-\frac{8\pi^2}{g^2(\Lambda_{\text{QCD}})}} \cos \theta , \quad (2.58)$$

which is negative, as expected from the general property that tunneling lowers the ground-state energy.

2.2.5 Tunneling in the Presence of Fermions

So far, only the pure YM Lagrangian has been considered but in real QCD fermions are present. This gives rise to the question what influence fermions have on instantons and if the statements made in this section still apply. In order to answer this question the path integral over the fermions will be evaluated in the presence of a fixed-size instanton.

Since QCD is vector-like, the analysis is limited to Dirac spinors. The corresponding action in Euclidean space is given by

$$S_E^{\text{ferm}} = \int d^4 x_E \bar{\psi} (-i\gamma^\mu D_\mu - im) \psi . \quad (2.59)$$

Using Grassmann calculus the corresponding integral yields

$$\det(i\gamma^\mu D_\mu - im) = \prod_n (\lambda_n + im) , \quad (2.60)$$

where the real numbers λ_n denote the eigenvalues of the Hermitian operator $i\gamma^\mu D_\mu$.

Solving the eigenvalue equation of $i\gamma^\mu D_\mu$ explicitly, it is found that among the solutions there are two zero-modes per Dirac fermion, one in ψ and one in $\bar{\psi}$. Consequently, in the presence of massless fermions the determinant in (2.60) vanishes and the tunneling probability becomes zero.

This is peculiar because the presence of fermions does not change the non-trivial topology in the space of gauge fields. Therefore, it should not change the picture of tunneling along this direction. The solution of this puzzle lies again in the chiral anomaly,

$$\partial_\mu J_5^\mu = \frac{n_F g_s^2}{16\pi^2} G_{\mu\nu}^a \tilde{G}^{a\mu\nu} . \quad (2.61)$$

Integrating both sides over the Euclidean four-volume and evaluating the integrals in the presence of the instanton, results in

$$Q_5(t = \infty) - Q_5(t = -\infty) = 2\nu . \quad (2.62)$$

The instanton changes the chiral charge of each massless fermion by two units, meaning that a left-handed fermion has to become right-handed or vice versa. In fact, since the number of zero-modes is equal to the change in the chiral charge, this is nothing else than a special case of the famous Atiya-Singer index theorem,

$$\nu_+ - \nu_- = \nu , \quad (2.63)$$

where ν_+ (ν_-) is the number of the normalizable zero-modes of positive (negative) chirality. To sum up, there is nothing wrong with the tunneling picture along the direction of non-trivial topology. The tunneling still takes place in the presence of massless fermions, but in such a way that the index theorem holds.

For the index theorem to be incorporated into the path integral, an operator with exactly this feature must be induced after integrating out the instantons. In the presence of an instanton located at x_1 , this is achieved by [57]

$$\int D\psi D\bar{\psi} e^{S[A,\psi]+J\bar{\psi}\psi} \iff \kappa \int D\psi D\bar{\psi} e^{S[0,\psi]+J\bar{\psi}\psi} \det(\bar{\psi}_R(x_1)\psi_L(x_1)) . \quad (2.64)$$

Generalizing this to n_+ instantons and n_- anti-instantons at the locations x_i , and integrating over their locations x_i yields an additional term in the action,

$$\begin{aligned} \sum_{n_\pm=0} \frac{\kappa^{n_++n_-}}{n_+!n_-!} e^{i\bar{\theta}(n_+-n_-)} \left(\int d^4x \det(\bar{\psi}_R(x)\psi(x)_L) \right)^{n_+} \left(\int d^4x \det(\bar{\psi}_R(x)\psi(x)_L) \right)^{n_-} \\ = \exp \left(\int d^4x \left(\kappa e^{i\bar{\theta}} \det(\bar{\psi}_R\psi_L) + \text{h.c.} \right) \right) . \end{aligned} \quad (2.65)$$

This effective vertex between (anti-)instantons and $2n_F$ fermion lines is the already mentioned 't Hooft vertex (see Fig. 2). In contrast to (2.12) in the previous section, where θ was moved to the fermion sector via a suitable redefinition, here we moved the phase

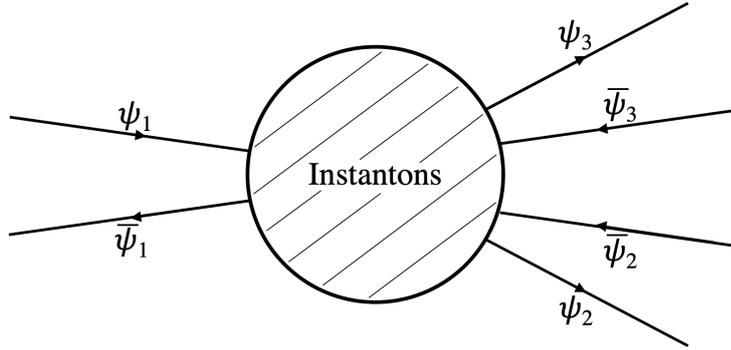


Figure 2: 't Hooft vertex for $n_F = 3$. In its essence, this non-perturbative multi-fermion vertex is a manifestation of the Atiyah-Singer index theorem, arising from zero-modes of the Dirac equation in the instanton background.

of the quark masses to the gluon sector. Thus, $\bar{\theta}$ appears in front of the determinant.

2.3 Topological Three-Form Formalism

Instantons play a crucial role in addressing not only the $U(1)_A$ problem via the 't Hooft vertex, but also in revealing the non-trivial vacuum structure of YM theories. This structure emerges from a compact direction in the infinite-dimensional field space, where an energy barrier exists. The compact nature of this direction leads to an infinite number of degenerate vacua labeled by the winding number n . Instantons correspond to finite-action tunneling processes through the periodic potential so that the physical vacua $|\theta\rangle$ are superpositions of n . As such, the angular parameter θ serves as a label for the non-trivial vacuum structure.

Our previous discussion on instantons and chiral perturbation theory provided the standard framework for understanding the nature of θ and the η' meson. However, there exists an additional language based on the topological quantities introduced in the previous section. This language is known as the topological three-form formalism [45, 46, 81] and will be the primary focus of this section. To provide a more accessible introduction to the topological three-form formalism, we will first explore its application in the context of the Schwinger model [60, 61]. This simplified model allows us to gain a deeper understanding of the topological features of the theory, which we can then apply to QCD [62].

2.3.1 Lessons from the Schwinger Model

The Schwinger model is basically massless QED in 1+1 dimensional space-time [60, 61], i.e.

$$S = \int d^2x \left(-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \bar{\psi} i \gamma^\mu (\partial_\mu + ieA_\mu) \psi + \frac{1}{2} \theta \epsilon^{\mu\nu} F_{\mu\nu} \right), \quad (2.66)$$

where the last term is the corresponding θ -term. In 1+1 dimensional space-time, the dual of the field strength tensor is a (pseudo) scalar given by $F^{\mu\nu} = \epsilon^{\mu\nu} E$. This scalar is essentially the electric field present on the off-diagonal elements of the tensor when represented as a matrix.

Let us ignore the electron for a moment and focus on the photon kinetic term and the θ -term. The field equations written in terms of E are

$$\partial_\mu F^{\mu\nu} = \epsilon^{\mu\nu} \partial_\mu E = 0. \quad (2.67)$$

Therefore, the solutions are given by $E = \kappa$, where κ represents an integration constant, indicating that the solution is a uniform electric field extending throughout the entire space.⁴ Since the vacuum energy in this model is simply $\rho = E^2/2$, a non-zero electric field indicates a non-zero vacuum energy.

To see the role of θ in this simple model, we write the action in terms of E , i.e.

$$S = \int d^2x \left(\frac{1}{2} E^2 - \theta E \right). \quad (2.68)$$

We see that by redefining the electric field as $E = \theta + \bar{E}$, we can remove the θ -term. The new solution to the field equations is $\bar{E} = \kappa - \theta$. This is a crucial observation. The vacuum angle θ is essentially some contribution into a background effective field, so setting θ to zero does not remove the background electric field. Rather it is the integration constant that encodes the physics of a vacuum angle. In the following, we set θ to zero and keep in mind that the physics of the vacuum angle is encoded in the integration constant κ .

Let us now add the massless electrons. From Sec. 2.1 we know that adding a massless fermion in the presence of a gauge symmetry should make θ unphysical. To see this, we formally integrate out the fermions via functional integration, yielding the effective action

$$S_{\text{eff}} = -\frac{1}{4} \left(\int d^2x F_{\mu\nu} F^{\mu\nu} - i \ln \det[\gamma^\mu (\partial_\mu + ieA_\mu)] \right). \quad (2.69)$$

The evaluation of the determinant can either be performed by solving the eigenvalue problem of the operator inside the determinant or, since the Schwinger model is quite simple, diagrammatically. The only non-vanishing Feynman diagram is the vacuum polarization graph (see Fig. 3). Due to Lorentz covariance and Ward-identities, the resulting polarization operator must be of the form $\Pi^{\mu\nu} \propto (g^{\mu\nu} - k^\mu k^\nu / k^2)$. The coefficient can be calculated after regularizing the diagram and is given by $e^2/2\pi$. Thus, the effective Lagrangian becomes

$$\mathcal{L}_{\text{eff}} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{e^2}{2\pi} A_\mu \left(g^{\mu\nu} - \frac{\partial^\mu \partial^\nu}{\square} \right) A_\nu, \quad (2.70)$$

⁴The integration constant κ has nothing to do with the constant in front of the 't Hooft determinant of the previous section.

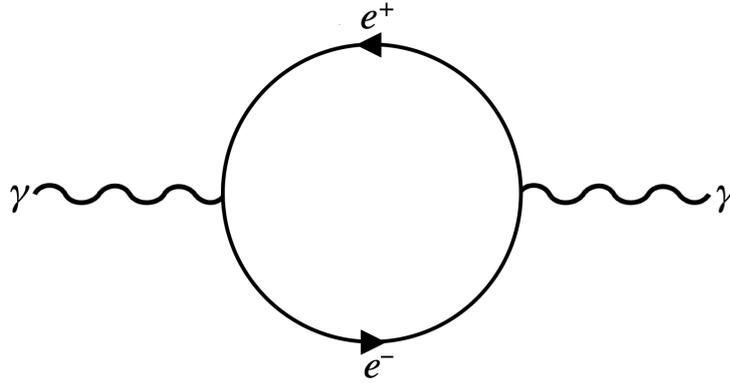


Figure 3: The vacuum polarization diagram that generates the polarization operator $\Pi^{\mu\nu} \propto (g^{\mu\nu} - k^\mu k^\nu / k^2)$.

where $1/\square$ is a symbolic notation for the Green's function of \square and the photon receives a mass $m^2 = e^2/\pi$. In 1+1 dimensions the photon field can be decomposed as $A_\mu = (\partial_\mu \phi + \epsilon_{\mu\nu} \partial^\nu \eta')/m$ as follows from Helmholtz's theorem. By plugging the Helmholtz decomposition into the field equations of the photon, we find that the first part of the decomposition decouples so that we are left with

$$(\square + m^2)\eta' = 0. \quad (2.71)$$

By also plugging in the Helmholtz decomposition into the field strength tensor, we find the relation between E and η' to be $\square\eta' = -mE$. Taking the d'Alembertian from the left in (2.71) and using the relation between E and η' , we find

$$(\square + m^2)E = 0. \quad (2.72)$$

The dual photon is now massive so that a non-zero constant electric field is no longer a solution. The vacuum angle indeed becomes unphysical by adding a massless fermion. This is intuitive since in the presence of a charged massless fermion the vacuum configuration will always be such that the electric field is completely screened by the fermion.

To make the connection with Sec. 2.1 where in the massless quark limit the anomaly was crucial to make θ unphysical, we use the 1+1 dimensional duality relation between axial and vector currents,

$$J_\mu^5 = \epsilon_{\mu\nu} J^\nu. \quad (2.73)$$

This relation is a consequence of geometry, which follows from the 1+1 dimensional gamma matrix relation,

$$\gamma_\mu \gamma_5 = \epsilon_{\mu\nu} \gamma^\nu. \quad (2.74)$$

In 3+1 dimensions the calculation of a triangle diagram between two vector- and one axial current reveals that the axial and vector current cannot be conserved simultaneously. Since the conservation of the vector current is crucial for gauge symmetry, the axial symmetry must be anomalous in the quantum theory. In 1+1 dimensions the same

holds true but the responsible diagram is the vacuum polarization (see Figure 3) with a vector and an axial current on each end. From the gamma matrix relation follows that the correlator between the axial current and the vector current is given by ${}^5\Pi_\mu^\nu = \epsilon_{\mu\alpha}\Pi^{\alpha\nu}$. Demanding the conservation of the vector symmetry by gauge invariance results in the divergence of the axial current becoming anomalous [82],

$$\partial^\mu J_\mu^5 = \frac{e}{\pi} F . \quad (2.75)$$

From there the generation of the mass-gap of E follows from the field equations,

$$\partial_\mu F^{\mu\nu} = \epsilon^{\mu\nu}\partial_\mu E = eJ^\nu , \quad (2.76)$$

where J^ν is the fermionic vector current. Bringing the epsilon-tensor to the other side gives

$$\partial_\alpha E = -eJ_\alpha^5 . \quad (2.77)$$

Taking another divergence results in (2.72), meaning that the anomaly is also in the Schwinger model responsible for the mass generation.

In order to understand intuitively why the constant electric field configuration disappears, we express the theory in the Stückelberg formulation, i.e.

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{m^2}{2}\left(A_\mu - \frac{1}{m}\partial_\mu B\right)^2 . \quad (2.78)$$

Here, B denotes the Stückelberg field, which represents the longitudinal mode of the photon. Let us dualize the Stückelberg field by first imposing the Bianchi identity in the Lagrangian via the Lagrange-multiplier η' ,

$$\mathcal{L}_{\text{dualize}} = \eta'\epsilon^{\mu\nu}\partial_\mu\partial_\nu B . \quad (2.79)$$

Integrating B out via the equations of motion gives

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}\partial_\mu\eta'\partial^\mu\eta' + \frac{m}{2}\eta'\epsilon_{\mu\nu}F^{\mu\nu} , \quad (2.80)$$

or equivalently

$$\mathcal{L} = \frac{1}{2}E^2 + \frac{1}{2}\partial_\mu\eta'\partial^\mu\eta' + m\eta'E . \quad (2.81)$$

The equation of motion for the photon field can then be written as

$$\partial_\mu E = m\partial_\mu\eta' , \quad (2.82)$$

so that $E = m(\eta' - \kappa)$ with κ again an integration constant. Plugging this back into the Lagrangian results in the expression we are looking for,

$$\mathcal{L} = \frac{m^2}{2}(\eta' - \kappa)^2 + \frac{1}{2}\partial_\mu\eta'\partial^\mu\eta' . \quad (2.83)$$

We observe that η' has its VEV exactly where the background electric field, expressed via the integration constant κ , vanishes. Incorporating θ just shifts κ , so that the VEV will be located where the combination of θ and κ vanishes. This picture of course is equivalent to the screening by the charged massless fermion but here the screening is performed by a boson. The equivalence between the fermion ψ and the η' boson is a special feature of 1+1 dimensional theories, namely bosonization.

The pseudo-scalar mode η' of the photon essentially plays the same role as the η' meson in QCD. Its mass is induced by the anomaly and it makes the vacuum angle unphysical by acquiring a VEV such that θ vanishes. Hence, every aspect of η' in the massless quark limit that we have discussed in the previous sections is included in the Schwinger model. However, in the Schwinger model we have the intuitive picture that the vacuum angle is encoded in the background electric field, which gets screened by massless fermions or, equivalently, by the scalar η' . Before showing that a similar picture exists in QCD, let us introduce the notion of a topological susceptibility.⁵

The susceptibility is a quantity that describes how a system is deformed by a source, which is assumed to be a small perturbation of the original system. Given a source term in the Lagrangian of the form $\mathcal{L}_{\text{source}} \sim \phi \mathcal{O}$, where ϕ denotes a source and \mathcal{O} some operator, the susceptibility is defined as

$$\langle \mathcal{O}, \mathcal{O} \rangle_{q \rightarrow 0} \equiv \lim_{q \rightarrow 0} \int d^d x e^{iqx} \langle \mathcal{T}[\mathcal{O}(x)\mathcal{O}(0)] \rangle. \quad (2.84)$$

If this value is non-zero, the system responds on the source and thus depends on it in some way.

In the Schwinger model we are interested in the topological susceptibility regarding the θ -term,

$$\langle \epsilon^{\mu\nu} F_{\mu\nu}, \epsilon^{\mu\nu} F_{\mu\nu} \rangle_{q \rightarrow 0} = \lim_{q \rightarrow 0} \int d^2 x e^{iqx} \langle \mathcal{T}[\epsilon^{\mu\nu} F_{\mu\nu}(x) \epsilon^{\rho\sigma} F_{\rho\sigma}(0)] \rangle, \quad (2.85)$$

which is just the two-point function of E . Here, the operator $\mathcal{O} = \epsilon^{\mu\nu} F_{\mu\nu}$ is a topological quantity, hence the two-point function is referred to as the topological susceptibility. As the topology of a system affects the vacuum energy (as discussed in Sec. 2.2), it is also referred to as the topological vacuum susceptibility.

Without a massless fermion, the field equations are solved by a constant electric field, so that the topological susceptibility is non-zero and the vacuum energy depends on θ . Since we can express the field strength in terms of A_μ , this statement is equivalent to

$$\langle AA \rangle_{q \rightarrow 0} \sim \frac{1}{q^2}, \quad (2.86)$$

which is the propagator of a massless field. From this we can make the crucial observation

⁵In the Schwinger model, the same dynamics and the same degrees of freedom that generate the mass are also responsible for the anomaly. In the 3+1 dimensional theory we must posit the anomaly separately from the mass generating dynamics [62].

that any physics making θ unphysical must remove the pole at $q = 0$, which in the presence of gauge-invariance is uniquely achieved by making A massive. This is another way to understand why θ becomes unphysical, when the massless fermion or the Stückelberg field is added.

2.3.2 Topological Susceptibility in QCD

Let us apply the argument based on the topological susceptibility to QCD. There, the topological susceptibility is given by

$$\langle G\tilde{G}, G\tilde{G} \rangle_{q \rightarrow 0} = \lim_{q \rightarrow 0} \int d^4x e^{iqx} \langle T[G\tilde{G}(x)G\tilde{G}(0)] \rangle.$$

The problem in QCD is that this quantity is very hard to calculate. As we mentioned in Sec. 2.2, 't Hooft's mechanism is semi-classical in the sense that instantons represent the extremal trajectory in the semi-classical treatment of tunneling transitions. Below the confinement scale, instantons are no longer suitable degrees of freedom so that the contribution from quantum correction cannot be neglected. However, Edward Witten and Veneziano demonstrated that in the large N limit of QCD, i.e. in the limit with a large number of colors, the topological susceptibility is not zero [83–85]. This claim is supported by other formalisms, such as lattice computations [86, 87](see also [88]), so that there is strong evidence for the topological susceptibility of QCD being non-zero. Using the topological quantities defined in (2.32), a constant topological susceptibility in QCD is equivalent to

$$\langle C, C \rangle_{q \rightarrow 0} \sim \frac{1}{q^2}, \quad (2.87)$$

where C is the Chern-Simons three-form. From this correlator we observe that the physicality of θ automatically leads to a pole at $q = 0$. Consequently, any physics making θ unphysical must remove this pole. Requiring gauge-invariance, this is uniquely achieved by making C massive [63].

In more physical terms, the pole at $q = 0$ corresponds to a non-zero propagator of Coulomb-type, meaning that C propagates a long-range interaction [89]. Since a three-form in four dimensions has no propagating degrees of freedom, the long range interaction takes the form of a constant four-form “electric field” $F_{\mu\nu\rho\sigma}$. Making C massive then means that the theory is moved from the Coulomb phase to the Higgs phase and, consequently, the four-form electric field gets screened.

We have exactly the same situation as in the Schwinger model but instead of the Photon field A in the correlator, we have the Chern-Simons three-form C . This is of not unexpected because A_μ in 1+1 dimensions plays the role of the Chern-Simons one-form and $\epsilon^{\mu\nu} F_{\mu\nu}$ plays the role of the Pontryagin density. In the following, we will set the constant value of the topological susceptibility to 1 for simplicity and demonstrate the three-form Higgs effect using EFT as first pointed out [63].

2.3.3 Three-Form Higgs Effect

The difference in QCD is that C and E are composite operators of the gluon field matrix A_μ^a . As such, the behavior of the gluons under the QCD gauge transformation translates into

$$C \rightarrow C + d\Omega, \quad E \rightarrow E, \quad (2.88)$$

where Ω is a two-form. While above the confinement scale C and E are not elementary and thus can be treated as a formal notation, this is no longer true below the confinement scale. In that regime C obtains the role of an elementary degree of freedom and E can be thought of the dual of its field strength [89],

$$F_{\mu\nu\rho\sigma} = \partial_{[\mu} C_{\nu\rho\sigma]} = \epsilon_{\mu\nu\rho\sigma} E. \quad (2.89)$$

Let us use EFT to make the analogy to the Schwinger model clearer. The effective Lagrangian of $C_{\nu\rho\sigma}$ is

$$\begin{aligned} \mathcal{L}_{\text{eff}} &= -\frac{1}{2 \cdot 4! \Lambda_{\text{QCD}}^4} (F_{\mu\nu\rho\sigma})^2 + \frac{1}{4!} \theta \epsilon^{\mu\nu\rho\sigma} F_{\mu\nu\rho\sigma} + \dots \\ &= \frac{1}{2 \Lambda_{\text{QCD}}^4} E^2 - \theta E + \dots, \end{aligned} \quad (2.90)$$

where the dots correspond to higher dimensional operators and in the second equality we used (2.89). The higher dimensional operators either carry higher powers of derivatives, which vanish in the $q \rightarrow 0$ limit, or have higher powers in E , which have no influence on the existence of the pole at $q = 0$ (see [90] for a discussion of sub-leading terms). Thus, we ignore these higher order terms in the following.

Like in the Schwinger model, the equations of motion are solved by a constant four-form electric field,

$$E = \Lambda_{\text{QCD}}^4 \kappa, \quad (2.91)$$

where κ is an integration constant that induces an additional CP violation beyond θ . The θ -term can be removed by redefining E , so that the solution of the equations of motion becomes $E = \Lambda_{\text{QCD}}^4 (\kappa - \theta)$. The vacuum angle is thus again a contribution into the background electric field. Therefore, it is not sufficient to get rid of θ alone but the constant solution for E must be prevented to make the vacuum angle unphysical. For simplicity, we set $\theta = 0$ in the following.

Let us now see how the anomaly is incorporated into the effective Lagrangian in the massless quark limit. We set the prefactors to unity and rewrite the anomalous current in terms of the Pontryagin density, i.e.

$$\partial_\mu J_5^\mu = G_{\mu\nu}^a \tilde{G}^{a\mu\nu} = E. \quad (2.92)$$

The anomaly in 3+1 dimensions generates the following interaction in the effective

Lagrangian [91]

$$\Delta\mathcal{L} = \frac{1}{\Lambda_{\text{QCD}}^2} E \frac{\partial_\mu}{\square} J_5^\mu . \quad (2.93)$$

With this interaction included, the equation of motion for $C_{\mu\nu\rho}$ becomes

$$\partial_\nu \left(E + \Lambda_{\text{QCD}}^2 \frac{\partial_\mu}{\square} J_5^\mu \right) = 0 . \quad (2.94)$$

Inserting the anomaly from (2.92),

$$(\square + \Lambda_{\text{QCD}}^2)E = 0 , \quad (2.95)$$

we see that the equation of motion for E is that of a massive field, which has no pole at $q = 0$, again like in the Schwinger model.

The relation between the massive three-form, which propagates one degree of freedom, and the massless three-form is evident in the Stückelberg formulation. The Higgsing comes down to the appearance of a two-form Stückelberg field, in contrast to a pseudo-scalar in the Schwinger model. Since in 3+1 dimensions a two-form is dual to a pseudo-scalar, this pseudo-scalar in the massless quark limit is the η' -meson. To explicitly see this, let us insert the current for the massless quark,

$$J_5^\mu = \bar{\Psi} \gamma^\mu \gamma_5 \Psi , \quad (2.96)$$

which after chiral symmetry breaking effectively becomes $J_5^\mu = f_{\eta'} \partial^\mu \eta'$. Introducing $m^2 = \Lambda_{\text{QCD}}^4 / f_{\eta'}^2$ in the following, the effective Lagrangian becomes

$$\mathcal{L}_{\text{eff}} = -\frac{1}{2 \cdot 4! \Lambda_{\text{QCD}}^4} (F_{\mu\nu\rho\sigma})^2 + \frac{f_{\eta'}^2}{2m^2} (\partial_\mu \eta')^2 - \frac{f_{\eta'}}{4! \Lambda_{\text{QCD}}^2} \eta' \epsilon^{\mu\nu\rho\sigma} F_{\mu\nu\rho\sigma} ,$$

where we took into account the η' kinetic term, which arises from $J_\mu^5 J_5^\mu / 2m^2$. From this Lagrangian there are two ways to explicitly see how a mass is generated.

First, by dualizing the η' and canonically normalizing, we arrive at the following three-form Stückelberg Lagrangian [46],

$$\mathcal{L}_{\text{eff}} = -\frac{1}{2 \cdot 4!} (F_{\mu\nu\rho\sigma})^2 + \frac{m^2}{2 \cdot 3!} (C_{\mu\nu\rho} - \partial_{[\mu} B_{\nu\rho]})^2 . \quad (2.97)$$

We see that C becomes massive by eating the two-form $B_{\mu\nu}$, which is the η' dual and plays the role of the Stückelberg-field.

Alternatively, by dualizing the three-form instead and integrating out E via its equations of motion,

$$\partial_\nu E = \Lambda_{\text{QCD}}^2 f_{\eta'} \partial_\nu \eta' , \quad (2.98)$$

one arrives at the following (canonically normalized) η' Stückelberg Lagrangian [46],

$$\mathcal{L}_{\text{eff}} = \frac{1}{2}(\partial_\mu \eta')^2 - \frac{m^2}{2}(\eta' - \kappa)^2, \quad (2.99)$$

where κ is the same integration constant as in (2.91). Reintroducing θ would effectively change $\kappa \rightarrow \kappa - \theta$. We see that the η' VEV is exactly at $\kappa - \theta$. This is also apparent in (2.98), where the η' perfectly screens E , not just θ .

So far we considered the massless quark limit. Let us now include the quark mass and see how the situation changes. The mass induces a further term in the divergence of the axial current,

$$\partial_\mu J_5^\mu = G_{\mu\nu}^a \tilde{G}^{a\mu\nu} + 2i \sum_f m_f \bar{\psi}_f \gamma_5 \psi_f = E + 2(m_d - m_u) \Lambda_{\text{QCD}}^3 \frac{\eta'}{f_{\eta'}}. \quad (2.100)$$

Going through the same calculation again but including the additional term arising from the quark masses, results in a second mass term in (2.99) where η'^2 and not $(\eta' - \kappa)^2$ appears. Hence, the VEV is no longer located where the integration constant vanishes. To put it differently, the η' can only partially screen the four-form electric field, leaving a residual contribution.

2.3.4 Higgsing of Multiple Three-Forms

The above analysis can be generalized to N non-Abelian YM groups, where for concreteness we choose N SU(3) groups. Without a massless quark the situation is exactly the same as in the single three-form case. Each SU(3) brings its own Chern-Simons three-form C_i , $i = 1, 2$, and a pole at $q = 0$ in the correlators $\langle C_i, C_i \rangle$. This again means that the equations of motion of each C_i will be solved by a constant four-form electric field,

$$E_i = \Lambda_i^4 (\kappa_i - \theta_i), \quad (2.101)$$

so that we have N different θ -vacuum structures. This is not changed if kinetic mixing terms of the form $E_i E_j$ are introduced in the effective Lagrangian, since we can transform into a basis with canonical kinetic terms.

Let us now consider two YM sectors and add one massless quark or, equivalently, one η' with the naive intention to make both vacuum angles unphysical. This should be done by making the axial symmetry anomalous with respect to both gauge groups, i.e.

$$\partial_\mu J_5^\mu = \alpha_1 G_{1\mu\nu}^a \tilde{G}_1^{a\mu\nu} + \alpha_2 G_{2\mu\nu}^a \tilde{G}_2^{a\mu\nu} = \alpha_1 E_1 + \alpha_2 E_2, \quad (2.102)$$

where the α_i are dimensionless parameters determining the coupling strength. Inserting this into the effective interaction,

$$\Delta\mathcal{L} = \frac{\alpha_1}{\Lambda_1^2} E_1 \frac{\partial_\mu}{\square} J^\mu + \frac{\alpha_2}{\Lambda_2^2} E_2 \frac{\partial_\mu}{\square} J^\mu, \quad (2.103)$$

the equations of motion for the different C_i become

$$\partial_\nu E_1 + \Lambda_1^2 \partial_\nu \frac{\partial_\mu}{\square} J^\mu = 0 , \quad (2.104)$$

$$\partial_\nu E_2 + \Lambda_2^2 \partial_\nu \frac{\partial_\mu}{\square} J^\mu = 0 . \quad (2.105)$$

Plugging in (2.102), this can be rewritten as

$$(\square + m^2) \left(\alpha_1 E_1 + \alpha_2 E_2 \right) = 0 , \quad (2.106)$$

$$\square \left(\frac{\alpha_2}{\Lambda_1^4} E_1 - \frac{\alpha_1}{\Lambda_2^4} E_2 \right) = 0 , \quad (2.107)$$

where $m^2 = (\alpha_1^2 \Lambda_1^4 + \alpha_2^2 \Lambda_2^4) / f_{\eta'}^2$. Thus, one superposition of four-form electric fields becomes massive, while the other one remains massless. Including the θ_i , the latter's equation of motion is solved by

$$\frac{\alpha_2}{\Lambda_1^4} E_1 - \frac{\alpha_1}{\Lambda_2^4} E_2 = \alpha_2 \kappa_1 - \alpha_1 \kappa_2 , \quad (2.108)$$

where κ_i are again integration constants inducing additional CP violation beyond the existed θ_i .

Let us again insert the single axial current (2.96) and perform the same steps as in the single three-form case. The effective Lagrangian takes the form

$$\begin{aligned} \mathcal{L}_{\text{eff}} = & - \frac{1}{2 \cdot 4! \Lambda_1^4} (F_1^{\mu\nu\rho\sigma})^2 - \frac{1}{2 \cdot 4! \Lambda_2^4} (F_2^{\mu\nu\rho\sigma})^2 + \frac{f_{\eta'}^2}{2m^2} (\partial_\mu \eta')^2 \\ & - \frac{f_{\eta'}}{4!} \eta' \left(\frac{\alpha_1}{\Lambda_1^2} \epsilon_{\mu\nu\rho\sigma} F_1^{\mu\nu\rho\sigma} + \frac{\alpha_2}{\Lambda_2^2} \epsilon_{\mu\nu\rho\sigma} F_2^{\mu\nu\rho\sigma} \right) . \end{aligned} \quad (2.109)$$

By dualizing the η' and canonically normalizing, we arrive at the following three-form Stückelberg Lagrangian,

$$\begin{aligned} \mathcal{L}_{\text{eff}} = & - \frac{1}{2 \cdot 4!} (F_1^{\mu\nu\rho\sigma})^2 - \frac{1}{2 \cdot 4!} (F_2^{\mu\nu\rho\sigma})^2 \\ & + \frac{m^2}{2 \cdot 3!} \left(\alpha_1 C_1^{\mu\nu\rho} + \alpha_2 C_2^{\mu\nu\rho} - \partial^{[\mu} B^{\nu\rho]} \right)^2 . \end{aligned} \quad (2.110)$$

Again, we see that only one superposition of three-forms can eat the two-form $B_{\mu\nu}$ and become massive, leaving the theory with one pole at $q = 0$.

Alternatively, by dualizing the three-form and integrating out both E_i via their equations of motion,

$$\partial_\nu E_i = \Lambda_i^2 f_{\eta'} \partial_\nu \eta' , \quad (2.111)$$

we arrive at the following (canonically normalized) axion Stückelberg Lagrangian,

$$\mathcal{L}_{\text{eff}} = \frac{1}{2}(\partial_\mu \eta')^2 - \frac{m^2}{2}(\eta' - \kappa_1)^2 + \frac{m^2}{2}(\eta' - \kappa_2)^2, \quad (2.112)$$

where the κ_i are the same integration constants as previously. We observe that the η' VEV cannot be chosen such that both κ_i disappear. It is important to realize that the situation is not altered when all the α_i and θ_i in the Lagrangian are set equal by some symmetry. In that case it may seem that a single η' with the above anomalous current can remove both θ_i but this is not true. The solution of the massless superposition's equation of motion still gives rise to the integration constant $\kappa = \alpha_2 \kappa_1 - \alpha_1 \kappa_2$, spontaneously breaking the assigned symmetry and making the vacuum angle physical again. To conclude this section:

- One η' cannot remove more than one vacuum angle. In order to make the vacuum angles of N non-Abelian gauge groups unphysical, N η' are required.
- Any attempt to change the η' mass by adding non-Abelian gauge groups, results in the η' no longer removing all vacuum angles.

2.4 $\bar{\theta}$ and the Strong CP Problem

So far we focused on understanding the θ -parameter by using different languages but we did not mention any observable consequences. The observable consequences, however, are the origin of the strong CP problem as formulated almost 50 years ago. In this section we discuss the observable consequences and formulate the strong CP problem in the traditional sense as well as in a modern formulation using consistence arguments of quantum gravity.

2.4.1 Observing $\bar{\theta}$ via the nEDM

The crucial property of the $\bar{\theta}$ -term for that purpose is that its CP violating as can be seen from the epsilon tensor. Alternatively, this can be seen by writing $G\tilde{G} \propto \vec{E}^a \cdot \vec{B}^a$: while both fields are unchanged under charge conjugation, the B -field is parity odd in contrast to the parity even E -field. But how can this contribute to an observable?

The key to this question lies in the Wigner-Eckert theorem [92, 93]. Without delving into technical details, the theorem essentially asserts that for a quantum system possessing a well-defined directional property (e.g. angular momentum) and an observable associated with a direction (e.g. a vector-valued operator), there exist strict limitations on how the orientation of the observable and the orientation of the state can interact with each other. For instance, in the case of a spin 1/2 particle all observables that are described by a vector must be collinear to the spin angular momentum vector \vec{S} . For a charged, spin 1/2 particle in an electric field \vec{E} , the simplest non-relativistic interaction including a vector-valued observable is

$$\mathcal{H} = -\vec{E} \cdot \vec{d} = -d \vec{E} \cdot \vec{S}, \quad (2.113)$$

where \vec{d} is the intrinsic electric dipole moment and we used the Wigner-Eckert theorem in the second equality. This interaction is CP violating and thus, due to the universality of the Wigner-Eckert theorem, the QCD vacuum angle and all other sources of CP violation contribute into the electric dipole moment.

Since the measurement for protons is quite hard due to charge contamination and the theoretical value for leptons is very small, the nEDM d_n is the most suitable to measure the QCD vacuum angle. The calculation of the nEDM can be performed using various frameworks, e.g. chiral perturbation theory [21], QCD sum-rules [94], holography [95], and lattice QCD [96]. In the following, we will simply estimate the nEDM by dimensional reasoning and EFT.

The above Hamiltonian interaction can be written in terms of a Lorentz invariant Lagrangian operator as follows,

$$\mathcal{L} = -d_n \frac{i}{2} \bar{n} \sigma_{\mu\nu} \gamma_5 n F^{\mu\nu}, \quad (2.114)$$

where n denotes the neutron and $\sigma_{\mu\nu}$ are the Lorentz generators for the anti-symmetric representation. This is a dimension 5 operator and hence the dipole moment as the Wilson coefficient must naively be suppressed by $d_n \sim 1/m_n$. Next, it needs to pick-up an imaginary part in order to contribute to the real nEDM, which can only originate from the phase of the quark mass. This motivates $d_n \sim \bar{\theta}$. In addition, from (2.15) we see that $d_n \rightarrow 0$ when $\bar{\theta} \rightarrow 0$ or $m_f \rightarrow 0$, so that a reasonable suppression factor is $d_n \sim m_q/m_n$. Lastly, being a dipole, the operator must be generated via an electromagnetic loop, i.e. $d_n \sim e/(8\pi^2)$. The final effective contribution to d_n can thus be estimated as

$$|d_n| \sim \frac{e}{8\pi^2} \frac{\bar{\theta}}{m_n} \frac{m_q}{m_n} \sim 10^{-18} \bar{\theta} \text{ e cm}. \quad (2.115)$$

The most precise calculation is based on QCD sum-rules and yields

$$d_n = 10^{-16} \bar{\theta} \text{ e cm}. \quad (2.116)$$

At this time, the best measurement of the nEDM is $d_n \lesssim 10^{-26} \text{ e cm}$, constraining the value of $\bar{\theta}$ to be smaller than 10^{-10} .

This unnatural smallness of $\bar{\theta}$ is known as the Strong CP problem. While it indeed looks like a naturalness problem, this is actually misleading. The quantum corrections to $\bar{\theta}$ are much smaller than the measured bound [25]. Furthermore, in the SM only $\bar{\theta}$ is an observable, not its two components θ and $\arg \det M_f$, so there is no reason to expect $\bar{\theta}$ to be large due to the CP violation in the Yukawa sector. Thus, from the point of view of naturalness it is rather a small value puzzle than a problem. In our experience, however, there is usually a deeper reason behind the smallness of parameters.

2.4.2 Constraining $\bar{\theta}$ via Quantum Gravity

The situation dramatically changes when gravity is taken into account, in particular quantum gravity (QG). This might seem hopeless in an EFT sense since QG corrections to any of our low-energy observables should be extremely small due to the huge suppression by the Planck scale. However, QG can become important if it forces consistency criteria upon low-energy theories.

A compelling argument for such a self-consistency requirement was proposed by Dvali et al [30–33], which states that the notion of a cosmological constant is not compatible with the quantum interpretation of the de Sitter metric as a coherent state of gravitons. The framework where this argument arises is called the corpuscular theory of gravity. In this picture gravitational systems are treated as condensates of gravitons. Describing classical gravitational backgrounds, such as black holes and cosmological space-times, using graviton coherent states may seem unconventional, but it is not. In fact, any classical field configuration can be viewed as a limit of quantum states with a large occupation number of particles. As an example, a classical electromagnetic wave is essentially a coherent state of photons. Therefore, it is expected that this is also the case for any microscopic theory that considers the quantum corpuscular structure of the cosmological background [30]. Consequently, gravitational systems possess the same phenomena as other many particle systems.

The phenomenon in this framework we are particularly interested in is the so called quantum breaking. In a many body system the constituents are subject to classical scattering but also to quantum scattering. The classical break time t_{cl} is the time scale after which the system changes its properties due to classical scattering. The quantum break time t_Q , in contrast, is the time after which the system changes due to quantum scattering and the relation between the two is

$$t_Q = \frac{t_{cl}}{\alpha}, \quad (2.117)$$

where α is the dimensionless (quantum) coupling strength.

In the case of a black hole, which is a localized object, this effect leads to the departure from the classical behavior after t_Q , meaning that the black hole can no longer be described by the solution of Einstein's field equations. In other words, the mean field approximation breaks down. In the case of de Sitter, due to the space-time being non-local, quantum breaking is non-physical and is thus not allowed to take place. Hence, the so called quantum breaking bound results in the inconsistency of de Sitter space vacuum. To put it differently, the quantum nature of gravitons leads to their scattering and ultimately causes the collapse of the coherent state, rendering it incompatible with a de Sitter “vacuum” and the existence of a cosmological constant [31].

This result is fundamentally based on the S-matrix formulation of quantum gravity and is thus far more general [33]. In such a formulation, the only limit in which the quantum back-reaction of scattered gravitons on the hypothetical de-Sitter vacuum,

characterized by a curvature radius R_{dS} and cosmological constant Λ , vanishes (or in other words the quantum breaking time goes to infinity) is given by:

$$\Lambda \rightarrow \infty, \quad \frac{1}{M_{\text{P}}} \rightarrow 0, \quad \text{such that} \quad \frac{\Lambda}{M_{\text{P}}} = \frac{1}{R_{\text{dS}}^2} = \text{finite}. \quad (2.118)$$

In this limit, however, the coupling of gravitons with wavelength $\lambda \lesssim R_{\text{dS}}$ goes as

$$\alpha_{\text{gr}} = \frac{1}{M_{\text{P}}^2 \lambda^2} \rightarrow 0. \quad (2.119)$$

This implies that a self-consistent theory of gravity based on the S-matrix formulation cannot include a de-Sitter vacuum without gravity decoupling. Therefore, de Sitter cannot be considered as a vacuum state but only an excited (coherent) state constructed on top of a true S-matrix vacuum of Minkowski. Notably, this is a unique feature of gravity, as the stability of the de Sitter “vacuum” and the quantum coupling of gravitons are both determined by the same parameter, M_{P} .

What does the inconsistency of a cosmological constant have to do with the strong CP problem? As we have seen in Sec. 2.2, the different vacua of QCD belong to different superselection sectors that are labeled by $\bar{\theta}$. The energy of these vacua depends on $\bar{\theta}$ with the global minimum being located at $\bar{\theta} = 0$ as a consequence of the Vafa-Witten theorem [34, 35]. Thus, all vacua with $\theta \neq 0$ are of de Sitter type and are thus forbidden by the quantum breaking bound or, more generally, by the S-Matrix exclusion of de Sitter. This not only elevates the strong CP problem in QCD from a small-value puzzle to a real problem but makes a vacuum-selection mechanism, which chooses the CP conserving vacuum at $\theta = 0$, a necessity. Such a mechanism is given by the PQ solution, which is the subject of the next chapter.

CHAPTER THREE

THE QCD AXION

We have seen that in the massless quark limit the η' meson removes the vacuum angle from the theory. However, in QCD the fit to light quark masses on the lattice has ruled out the possibility of massless quarks by 20σ [97]. Using the information from the previous chapter, we can understand how massive quarks spoil the solution. In chiral perturbation theory, the VEV of the η' meson is no longer located at the point where the vacuum angle is removed. In the three-form formalism, the η' cannot fully align with the background electric field to screen it. The issue in both languages stems from the anomaly, where the inclusion of the quark mass term results in an additional contribution to the anomalous current. From this statement, a solution is almost obvious: We need to introduce a new global axial symmetry $U(1)_{\text{PQ}}$, which is non-linearly realized and anomalous only with respect to QCD. This is the PQ solution and its pseudo-Goldstone is the axion. In this light, the massless quark solution is nothing else than the PQ mechanism with the η' playing the role of the axion.

In this section, we formally explain the PQ solution from the perspectives of the formalisms used in Chapter 2. Then we write down the low-energy EFT of the axion and derive the axion mass, the axion-fermion couplings, and the axion-photon coupling. Lastly, we discuss explicit models, i.e. models which provide a UV completion for the axion EFT, that will be used in latter chapters of this thesis. These are the PQWW model [36–39], the DFSZ-type models [51, 52], the KSVZ-type models [53, 54], and the two-form implementation [45, 63].

3.1 The PQ Mechanism

Essentially, the PQ mechanism is the introduction of a global axial symmetry $U(1)_{\text{PQ}}$, which is non-linearly realized and anomalous only with respect to a certain YM group. In the low-energy EFT, all these properties result in the appearance of the axion a , a pseudo-scalar with the anomalous coupling

$$\mathcal{L}_a \supset \left(\frac{a}{f_a} + \bar{\theta} \right) \frac{g^2}{32\pi^2} G\tilde{G}. \quad (3.1)$$

Here, f_a is the axion decay constant, G the field strength of the associated YM group, \tilde{G} its dual, and g the YM coupling.

In Chapter 2, we observed that the vacuum angle enters into the vacuum energy. Including the axion, which transforms as $a \rightarrow a + \kappa$ with a constant κ (the non-linearly realized $U(1)_{\text{PQ}}$), the vacuum angle becomes dynamical. Consequently, the vacuum energy acts as a potential for a , allowing it to relax to the unique minimum at $a(x) = 0$ as dictated by the Vafa-Witten theorem. The particular potential can be calculated either in chiral perturbation theory or using instanton techniques (see Sec. 3.2).

We want to stress that, for the PQ mechanism to work, the only contribution to the anomaly must be the one proportional to the Pontryagin density of the associated YM group. Any additional explicit breaking results in a residual vacuum angle, thus shifting the question of the smallness of $\bar{\theta}$ to the smallness of this new angle or, equivalently, to the smallness of the breaking parameter. This is not a fundamental issue per se but it renders the PQ solution no longer a real solution. However, it becomes a fundamental issue when quantum gravity is incorporated. With additional explicit breaking terms, the PQ solution cannot remove all de Sitter-type vacua from the theory, so that the theory is equally incompatible with the S-matrix exclusion of de Sitter as before. In this light, any additional contribution into the anomaly is strictly forbidden.

An exception to this is provided by Abelian anomalies. There, the vacuum angle is not physical due to the lack of field configurations, such as instantons, that result in a non-trivial topological susceptibility.

3.2 The Low-Energy EFT of the Axion

The PQ solution is intriguing in the sense that it predicts the existence of a new light pseudoscalar particle. This section focuses on formulating the low-energy axion EFT, which provides a good starting point for obtaining an overview of the axion properties and understanding how the UV model dependency enters [50]. We have based our notation and presentation on [55], but made some minor modifications and clarifications to better suit the specific focus of this thesis.

3.2.1 Effective Lagrangian

The low-energy axion EFT follows from the required properties of the PQ mechanism. As a consequence of the $U(1)_{\text{PQ}}$ being non-linearly realized, the axion couplings have the form

$$\mathcal{L}_a \supset \frac{\partial_\mu a}{v_a} J_{\text{PQ}}^\mu + \frac{a}{v_a} \partial_\mu J_{\text{PQ}}^\mu, \quad (3.2)$$

where J_μ^{PQ} denotes the PQ current and v_a the PQ scale. The PQ current depends on the global charges χ_ϕ of the fields transforming under $U(1)_{\text{PQ}}$, so a chiral fermion ψ_L results in the contribution $J_{\text{PQ}}^\mu|_{\psi_L} = -\chi_{\psi_L} \bar{\psi}_L \gamma^\mu \psi_L$. The anomaly of the PQ current, including the electromagnetic anomaly and considering arbitrary quark representations, is given by

$$\partial_\mu J_{\text{PQ}}^\mu = \frac{g^2 \mathcal{N}}{16\pi^2} G\tilde{G} + \frac{e^2 \mathcal{E}}{16\pi^2} F\tilde{F}, \quad (3.3)$$

where \mathcal{N} and \mathcal{E} denote the color and the electromagnetic anomaly coefficients, respectively. We constrain ourselves to operators up to dimension 5 that play a role in this thesis, thus omitting numerous couplings and mixings like axion-nucleon couplings and axion- η' mixing (for these, see [55, 98]). Correspondingly, the effective Lagrangian is

$$\begin{aligned} \mathcal{L}_a = & \frac{1}{2} \partial_\mu a \partial^\mu a + \frac{a}{v_a} \frac{g^2 \mathcal{N}}{16\pi^2} G\tilde{G} + \frac{a}{v_a} \frac{g^2 \mathcal{E}}{16\pi^2} F\tilde{F} - \bar{q}_L M_q q_R + \text{h.c.} \\ & - \frac{\partial_\mu a}{v_a} [\chi_{q_L} \bar{q}_L \gamma^\mu q_L + \chi_{q_R} \bar{q}_R \gamma^\mu q_R + \chi_{\ell_L} \bar{\ell}_L \gamma^\mu \ell_L + \chi_{\ell_R} \bar{\ell}_R \gamma^\mu \ell_R], \end{aligned} \quad (3.4)$$

where q denotes the quarks, ℓ the leptons, M_q is the diagonal quark mass matrix, and we used the axion shift-symmetry to remove the θ -parameter. The $G\tilde{G}$ -term is canonically normalized by defining

$$f_a \equiv \frac{v_a}{2\mathcal{N}}. \quad (3.5)$$

For convenience, we perform the field-dependent axial transformation of the quark fields,

$$q \rightarrow e^{i\gamma_5 \frac{a}{2f_a} Q_a} q, \quad (3.6)$$

where Q_a is a generic matrix with $\text{tr}(Q_a) = 1$ acting on the quark fields. This redefinition removes the $G\tilde{G}$ -term but induces a shift in the $F\tilde{F}$ -term through the electromagnetic anomaly, a shift in the axion-quark interaction through the quark kinetic terms, and a change in the quark mass term. Taking all these changes into account, the effective Lagrangian becomes

$$\begin{aligned} \mathcal{L}_a = & \frac{1}{2} \partial_\mu a \partial^\mu a + \frac{1}{4} g_{a\gamma} a F\tilde{F} - \bar{q}_L M_a q_R + \text{h.c.} \\ & + c_q \frac{\partial_\mu a}{2f_a} \bar{q} \gamma^\mu \gamma_5 q + c_l \frac{\partial_\mu a}{2f_a} \bar{l} \gamma^\mu \gamma_5 l, \end{aligned} \quad (3.7)$$

where we introduced

$$g_{a\gamma} = \frac{\alpha}{2\pi f_a} \left(\frac{\mathcal{E}}{\mathcal{N}} - 6 \operatorname{tr}(Q_a Q^2) \right) \quad \text{with} \quad Q = \operatorname{diag}(2/3, -1/3, -1/3), \quad (3.8)$$

$$c_q = \frac{\chi_{qL} - \chi_{qR}}{2\mathcal{N}} - Q_a, \quad (3.9)$$

$$c_\ell = \frac{\chi_{\ell L} - \chi_{\ell R}}{2\mathcal{N}}, \quad (3.10)$$

$$M_a = e^{i\frac{a}{2f_a}Q_a} M_q e^{i\frac{a}{2f_a}Q_a}. \quad (3.11)$$

By utilizing this effective Lagrangian, the consequences of the axion can be investigated, making it a crucial tool for any phenomenological analysis. In the following, we discuss the various terms.

3.2.2 Axion Potential and Mass

Let us start by calculating the axion potential emerging from (3.7). In principal, the axion potential can either be obtained through instanton techniques or chiral perturbation theory. However, since we have already computed the vacuum energy in these formalisms and the axion is essentially a dynamical θ , we can use that the vacuum energy acts as potential for a . For simplification, we again only consider the two light quarks in the following.

The vacuum energy emerging from instantons in the dilute instanton gas approximation is given by (2.58), giving rise to the axion potential

$$V(a) = \Lambda_{\text{QCD}}^4 \left(1 - \cos\left(\frac{a}{f_a}\right) \right), \quad (3.12)$$

where we approximated $Ke^{-\frac{8\pi^2}{g^2(\Lambda_{\text{QCD}})}} \sim 1$ [58]. Expanding the cosine, yields the axion mass

$$m_a^2 = \frac{\Lambda_{\text{QCD}}^4}{f_a^2} \implies m_a = 4 \times 10 \left(\frac{10^{12} \text{ GeV}}{f_a} \right) \mu\text{eV}. \quad (3.13)$$

Although instantons may not provide reliable degrees of freedom at energies below the QCD scale, the potential captures the basic qualitative properties at leading order in a/f_a . For this reason, it is a suitable potential when axion self-interactions are negligible.

A more reliable potential can be obtained from chiral perturbation theory. There, the vacuum energy (2.21) gives rise to the axion potential

$$V(a) = -2v^3 \left[m_u \cos\left(\frac{m_d}{m_u + m_d} \frac{a}{f_a}\right) + m_d \cos\left(\frac{m_u}{m_u + m_d} \frac{a}{f_a}\right) \right], \quad (3.14)$$

which corresponds to the choice $Q_a = M_q^{-1}/\operatorname{tr}M_q^{-1} = \operatorname{diag}(m_d, m_u)/\operatorname{tr}M_q$. It should be noted that according to (2.20), this choice does not induce any mixing with the pions.

Expanding then results in the mass,

$$m_a^2 = \frac{2\Lambda_{\text{QCD}}^3}{f_a^2} \frac{m_u m_d}{(m_u + m_d)^2} \sim \frac{\Lambda_{\text{QCD}}^3 m_u}{f_a^2} \implies m_a = 5.7 \left(\frac{10^{12} \text{ GeV}}{f_a} \right) \mu\text{eV}. \quad (3.15)$$

An alternative expression follows from the choice $Q_a = \text{diag}(1, 1)/2$, giving rise to the potential

$$V(a) = -\frac{m_\pi^2 f_\pi^2}{(1+z)} \sqrt{1+z^2+2z \cos\left(\frac{a}{f_a}\right)}, \quad (3.16)$$

where $z \equiv m_u/m_d$. Expanding in a/f_a , results in the same expression for the mass as the first choice of Q_a . However, higher order terms differ because $Q_a = \text{diag}(1, 1)/2$ induces a mixing with the pions, which we have neglected here.

The crucial point regarding the axion mass is that it does not directly depend on a particular UV model. The form $m_a \propto f_a^{-1}$ follows from the anomaly requirement of the PQ mechanism, hence one can say that this form is dictated by the necessity to solve the strong CP problem. Nevertheless, the UV dependency can enter indirectly by UV models preferring different values of f_a .

3.2.3 Axion-Photon Coupling

To obtain the axion-photon coupling, let us use $Q_a = M_q^{-1}/\text{tr}M_q^{-1}$, so no axion-pion mixing has to be taken into account. Inserting this choice for Q_a in (3.8), yields

$$g_{a\gamma} = \frac{\alpha}{2\pi f_a} \left(\frac{\mathcal{E}}{\mathcal{N}} - \frac{2}{3} \frac{4m_d + m_u}{m_d + m_u} \right) = \frac{\alpha}{2\pi f_a} \left(\frac{\mathcal{E}}{\mathcal{N}} - 1.92(4) \right), \quad (3.17)$$

where we included next-to-leading-order chiral corrections in the numerical value [99].⁶

Let us now define the anomaly coefficients \mathcal{E} and \mathcal{N} . These can be decomposed into the contributions from each irreducible representation ϕ ,

$$\mathcal{N} = \sum_{\phi} \mathcal{N}_{\phi}, \quad \mathcal{E} = \sum_{\phi} \mathcal{E}_{\phi}, \quad (3.18)$$

where each contribution is given by (for a comprehensive review, see [75]),

$$\mathcal{N}_{\phi} = \chi_{\phi} \dim(I_{\phi}) \text{T}(C_{\phi}), \quad (3.19)$$

$$\begin{aligned} \mathcal{E}_{\phi} &= \chi_{\phi} \dim(C_{\phi}) \text{tr}(Q_{\phi}^2) \\ &= \chi_{\phi} \dim(C_{\phi}) \dim(I_{\phi}) \left(\frac{1}{12} (\dim(I_{\phi})^2 - 1) + Y_{\phi}^2 \right). \end{aligned} \quad (3.20)$$

In these expressions, the χ_{ϕ} denote the PQ charges, C_{ϕ} and I_{ϕ} the color and isospin representations, $\text{T}(C_{\phi})$ the color Dynkin index with normalization $\text{T}(3) = 1/2$, Q_{ϕ} the electromagnetic charge generator, Y_{ϕ} the hypercharge, and in the second equality for

⁶The number in parenthesis denotes the error in the last digit.

\mathcal{E} we simply used the Gellmann-Nishijima relation $Q_\phi = I_\phi^{(3)} + Y_\phi$. It should be noted that the contributions from right handed fermions come with an overall minus due to the different sign in the projector P_R in the evaluation of the triangle diagram.

3.3 Axion Models

We have observed that the axion-photon coupling has two contributions. The second term comes from QCD mixing effects. The first term stems from UV physics and is thus model dependent. As evident from (3.9) and (3.10), such a model dependency appears in any axion coupling, revealing that it is a universal characteristic of axion-couplings. In fact, this characteristic is fundamentally rooted in the axion's Goldstone nature [50]. The important consequence is that UV models are needed to make concrete predictions about the couplings of the axion. In this section, we present the benchmark models relevant for this thesis and discuss general aspects of axion model building. These benchmark models are:

- PQWW (Peccei-Quinn-Weinberg-Wilczek) [36–39]
Induce anomaly with respect to QCD by requiring SM fermions to carry PQ charges. As a consequence, the SM scalar sector must be expanded by a second Higgs doublet. This model is phenomenologically excluded since the PQ scale is connected to the electroweak scale.
- DFSZ (Dine-Fischler-Srednicki-Zhitnitsky) [51, 52]
Same basic idea as PQWW model but the SM scalar sector is expanded by a second Higgs doublet and a Higgs singlet. In this way the PQ scale is decoupled from the electroweak scale, resulting in an invisible axion model.
- KSVZ (Kim-Shifman-Vainshtein-Zakharov) [53, 54]
Induce anomaly with respect to QCD by a new heavy quark and a new Higgs singlet that guarantees the PQ invariance. In this way, only these two new fields carry PQ charges instead of any SM fermions.
- Two-Form Axion [45, 46]
Introduce the axion as a two-form that plays the role of the Stückelberg field for the massive Chern-Simons three-form. Below the PQ scale, the two-form is dual to the ordinary axion, but UV completion and implications are dramatically different.

3.3.1 PQWW-Model

The crucial ingredient in developing an axion model is to make the $U(1)_{PQ}$ anomalous with respect to QCD or, in other words, to have $\mathcal{N} \neq 0$. This is not possible with a single Higgs doublet when the SM fermions are required to carry PQ charges. To see this, let us assume that the up- and down-type quarks carry PQ charges. Their contribution to

the color anomaly results in $\mathcal{N} = (\chi_u + \chi_d)/2$, where

$$\chi_u = \chi_{u_L} - \chi_{u_R} , \quad (3.21)$$

$$\chi_d = \chi_{d_L} - \chi_{d_R} . \quad (3.22)$$

In the SM, the Higgs doublet couples to both quark types but with a relative complex conjugation. Therefore, we have $\chi_d = -\chi_u$, resulting in a vanishing color anomaly. Another way to see this, is by realizing that there is no Goldstone left in the SM to play the role of the axion. When the electroweak symmetry is broken, three Higgs doublet degrees of freedom render the W and Z bosons massive by becoming their longitudinal modes, while the fourth degree of freedom becomes the Higgs field. For these reasons, a minimal implementation with charged SM fermions requires a second Higgs doublet.

A general prediction of a two Higgs doublet model (2HDM) are flavor changing neutral currents (FCNCs) [100]. In the SM, these are only present at higher orders as a consequence of the Glashow-Iliopoulos-Maiani mechanism [101] and are thus suppressed, in agreement with experiments [102, 103]. Therefore, new physics should not induce FCNCs at tree level or provide mechanisms to suppress FCNC reactions. The simplest way to achieve this is by imposing the Glashow-Weinberg condition [104, 105], meaning that every fermion type must couple to a single Higgs. We denote the doublet coupling to the up-type quarks as H_u and the one coupling to the down-type quarks as H_d . The leptons then acquire their mass via couplings to either of the Higgs doublets. With these couplings the hypercharge of each doublet is fixed to be $Y_{H_d} = Y_{H_u} = -1/2$.

The kinetic and the Yukawa sector have a $U(1)_u \times U(1)_d$ symmetry, which must be identified with $U(1)_Y \times U(1)_{PQ}$ in the end. For this reason, none of the initial $U(1)$ symmetries must be explicitly broken in the scalar potential. This fixes the potential to have the form

$$V(H_u, H_d) = \sum_{i,j \in \{u,d\}} -\mu_i^2 H_i^\dagger H_i + a_{ij} (H_i^\dagger H_i) (H_j^\dagger H_j) + b_{ij} (H_i^\dagger H_j) (H_j^\dagger H_i) , \quad (3.23)$$

where the matrices a_{ij} and b_{ij} are real and symmetric. The coefficients can be restricted to give rise to a proper scalar potential⁷, so that for the electroweak symmetry breaking both doublets acquire a VEV in the standard basis,

$$\langle H_d \rangle = v_d \exp \left(i \frac{a_d}{v_d} \right) \begin{pmatrix} 0 \\ 1 \end{pmatrix} , \quad \langle H_u \rangle = v_u \exp \left(i \frac{a_u}{v_u} \right) \begin{pmatrix} 1 \\ 0 \end{pmatrix} . \quad (3.24)$$

The superposition that becomes the longitudinal mode of the Z boson is

$$b = \frac{1}{v_b} \sum_{i \in \{u,d\}} Y_i v_i a_i = \frac{1}{v} (v_d a_d - v_u a_u) , \quad (3.25)$$

⁷A scalar potential that is bounded from below and renormalizable is called a proper scalar potential. For these there exists a region in parameter space where all doublets get a VEV in the standard basis.

where v is the electroweak scale and $v_b^2 = \sum Y_i^2 v_i^2 = v^2/4$. The axion is uniquely defined by being the orthogonal superposition, i.e.

$$a = \frac{1}{v_a} \sum_{i \in \{u,d\}} \chi_i v_i a_i = \frac{1}{v} (v_u a_d + v_d a_u), \quad (3.26)$$

where $\chi_d = v_u/v_d$ and $\chi_u = v_d/v_u$ and, consequently, $v_a^2 = \sum \chi_i^2 v_i^2 = v^2$.

The problem with the PQWW axion is that the PQ scale coincides with the electroweak scale. Such a relatively strongly coupled Goldstone is not only contradicting to astrophysics (see Chapter 4) but also to many experiments conducted around the time of the axions proposal.

3.3.2 DFSZ Model

One way to address the issues of the PQWW model is to decouple the PQ scale from the electroweak scale. The PQ scale can then be pushed to higher scales, rendering the axion “invisible”. This can most easily be achieved by introducing a Higgs singlet whose angular mode will contribute into the axion. The resulting model is the DFSZ-model and it provides the first example of an invisible axion model [51, 52].

The DFSZ model is similar to the original PQWW model, as the fermionic fields in the SM are charged under the PQ symmetry. To make the PQ symmetry anomalous with respect to QCD, it is necessary for the scalar sector to feature two doublets denoted as H^d and H^u . Moreover, a singlet is introduced to make the axion undetectable by decoupling the PQ scale from the electroweak scale [51, 52]. We denote the singlet as S in the following. We again impose the Glashow-Weinberg condition, so that H^u couples to the up-type quarks and H^d to the down-type quarks. This fixes the hypercharges to be $Y_{H^d} = Y_{H^u} = -1/2$. With two doublets there is a freedom in choosing which doublet couples to the leptons but we will leave this unspecified for now.

The kinetic and the Yukawa sector have a global $U(1)_u \times U(1)_d \times U(1)_S$ symmetry. To ensure a well-defined PQ current and to prevent the existence of Goldstone bosons with decay constants of the electroweak scale order, it is necessary to explicitly break down this symmetry to $U(1)_{\text{PQ}} \times U(1)_Y$. With two doublets this is uniquely achieved by the term $H_u H_d S^\dagger S^\dagger$ (up to hermitian conjugation of both singlets), which is equivalent to the condition $\chi_u + \chi_d = 2\chi_S$. By means of a proper scalar potential, each scalar field develops a VEV v_f , where we introduced the index $f = u, d, S$. Expanding around these VEVs yields,

$$H_d \supset \frac{v_d}{\sqrt{2}} e^{i \frac{a_d}{v_d}} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad H_u \supset \frac{v_u}{\sqrt{2}} e^{i \frac{a_u}{v_u}} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad S \supset \frac{v_S}{\sqrt{2}} e^{i \frac{a_S}{v_S}}.$$

Here, we neglected any degrees of freedom not containing the axion. Each angular mode a_f transforms under a PQ transformation as $a_f \rightarrow a_f + \kappa_f \chi_f v_f$ with the χ_f denoting

the PQ charges and the κ_f being constants. After spontaneous symmetry breaking, the corresponding PQ current is then

$$\begin{aligned} J_\mu^{\text{PQ}} \Big|_a &\supset -\chi_S S^\dagger i \partial^\mu S - \chi_u H_u^\dagger i \partial^\mu H_u - \chi_d H_d^\dagger i \partial^\mu H_d + \text{h.c.} \\ &= \chi_u v_u \partial_\mu a_u + \chi_d v_d \partial_\mu a_d + \chi_S v_S \partial_\mu a_S . \end{aligned} \quad (3.27)$$

By requiring $J_\mu^{\text{PQ}}|_a = v_a \partial_\mu a$ and $a \rightarrow a + \kappa v_a$ under the PQ transformation, the axion field is defined as a linear combination of all scalar angular modes,

$$a = \frac{1}{v_a} (\chi_u v_u a_u + \chi_d v_d a_d + \chi_S v_S a_S) , \quad (3.28)$$

$$v_a^2 = \chi_u^2 v_u^2 + \chi_d^2 v_d^2 + \chi_S^2 v_S^2 , \quad (3.29)$$

similar to the PQWW model. We observe that the VEV of the singlet decouples the PQ scale from the electroweak scale.

Again, the axion must be orthogonal to the superposition that becomes the longitudinal mode of the Z boson. This is equivalent to the orthogonality requirement between the PQ current defined in (3.27) and the weak hypercharge current $J_\mu^Y|_a = \sum_f Y_f v_f \partial_\mu a_f$, implying

$$\sum \chi_f Y_f v_f^2 = 0 . \quad (3.30)$$

Solving this condition together with the other condition resulting from the explicit breaking term, allows to express χ_u and χ_d in terms of χ_S , which is otherwise unconstrained in the original DFSZ model and can be conveniently chosen to be $\chi_S = 1$. With this choice, we find

$$\chi_u = 2 \frac{v_d^2}{v^2} , \quad \chi_d = 2 \frac{v_u^2}{v^2} . \quad (3.31)$$

With the axion identified, the low energy theory can be constructed using standard methods. Inverting (3.29) allows to express the scalar angular modes in terms of the axion. Since our focus is on terms that include the axion, this essentially amounts to the substitution

$$\frac{a_f}{v_f} \rightarrow \chi_f \frac{a}{v_a} . \quad (3.32)$$

The Lagrangian can then be brought to the form of the axion EFT in (3.4) by performing the field-dependent chiral redefinition of the fermion fields,

$$\psi \rightarrow \exp \left(-i \gamma_5 \chi_f \frac{a}{2v_a} \right) \psi . \quad (3.33)$$

In the model under consideration, the ratio between the electromagnetic and color anomaly coefficients can be expressed as [106]

$$\frac{E}{N} = \frac{2}{3} + 2 \frac{\chi_u + \chi_e}{\chi_u + \chi_d} . \quad (3.34)$$

As we already mentioned, with two doublets there is a freedom in choosing which doublet couples to the leptons. If H^d couples to the leptons, which is referred to as the DFSZ-I model, we have $\chi_e = \chi_d$. Alternatively, if \tilde{H}^u couples to the leptons, this is referred to as the DFSZ-II model and we have $\chi_e = -\chi_u$. For the anomaly ratio this means

$$\text{DFSZ - I : } \frac{E}{N} = \frac{8}{3}, \quad (3.35)$$

$$\text{DFSZ - II : } \frac{E}{N} = \frac{2}{3}. \quad (3.36)$$

Finally, note that using (3.31) the color anomaly is $\mathcal{N} = 3$ and thus $f_a = v_a/6$.

3.3.3 KSVZ Model

The phenomenological issues of the PQWW model can also be addressed by building a model where the QCD anomaly is not generated by the SM fermions. Instead, it is generated by a new fermion Q , which must transform non-trivially under QCD. The simplest possibility is given by a fermion that is a QCD fundamental, a $SU(2)_L$ singlet, and has zero hypercharge, i.e. $Q \sim (3, 1, 0)$. Adding a Higgs singlet $S \sim (1, 1, 0)$ to guarantee a PQ invariant Yukawa coupling, results in the KSVZ model, the second benchmark invisible axion model [53, 54].

The Lagrangian of this model is

$$\mathcal{L} = |\partial_\mu S|^2 + \bar{Q}i\gamma^\mu D_\mu Q - y_Q S \bar{Q}_L Q_R + \text{h.c.} - V(S). \quad (3.37)$$

It features a $U(1)_{\text{PQ}}$ symmetry under which Q_L carries charge $1/2$, Q_R carries charge $-1/2$, and S consequently carries charge $+1$. Moreover, $V(S)$ is a Mexican hat potential resulting in the spontaneous breaking of the $U(1)_{\text{PQ}}$,

$$S \supset \frac{v_s}{\sqrt{2}} e^{i\frac{a}{v_a}}. \quad (3.38)$$

Performing a field-dependent chiral redefinition of Q to remove the axion from the quark sector, again results in the axion EFT (3.4). The SM fermions carry no PQ charges in the KSVZ model, so that the axion-photon coupling is the only coupling of the axion with the SM fields. Regarding this coupling, since in the KSVZ model Q is electrically neutral, the anomaly ratio vanishes,

$$\text{KSVZ : } \frac{E}{N} = 0, \quad (3.39)$$

and the axion-photon coupling is fully determined by the QCD mixing. In this model, the color anomaly is $\mathcal{N} = 1/2$ and thus $f_a = v_a$.

3.3.4 Two-Form Axion

The models discussed so far realize the PQ mechanism by introducing the global $U(1)_{\text{PQ}}$ symmetry by hand with no additional explicit breaking. However, there exists a folk-theorem stating that no global continuous symmetries are allowed in quantum gravity (see for instance [107]). The general consensus arising from this folk-theorem is that global continuous symmetries only appear as approximate symmetries after the spontaneous breaking of gauge symmetries. Consequently, it is expected that also the PQ symmetry emerges as a global approximate symmetry with a sufficiently small explicit breaking to guarantee $\bar{\theta} \lesssim 10^{-10}$. With this believe the above models merely shift the question of a small $\bar{\theta}$ to the question of a small breaking parameter, which is called the “quality problem” of the PQ symmetry. The situation gets even worse when the S-matrix exclusion of de Sitter is taken into account, which requires the exact cancellation of any vacuum angle.

An approach avoiding these issues is provided by implementing the PQ mechanism at an intrinsic field theory level.⁸ By introducing the axion as a two-form that plays the role of the Stückelberg field for the massive Chern-Simons three-form, the axion shares the gauge redundancy of gluons and is thus protected against arbitrary UV physics. While the two-form axion provides a different approach in the implementation of the PQ mechanism, it is not a UV model. Nevertheless, a fair amount of consequences arise just from the assumption of such an intrinsic implementation.

To begin with, there is no renormalizable UV-completion in terms of a complex scalar field for a two-form. Consequently, it is likely that a UV completion has its origin in a theory of quantum gravity. This implies that in the two-form implementation f_a is related to the UV cutoff of quantum gravity, which in the simplest case would mean $f_a \sim M_{\text{P}}$. As such, the two-form implementation provides an example for an indirect influence on the axion-mass by favoring certain values of f_a .

Next, in such an implementation the PQ mechanism is only spoiled by deformations of the theory that put it out of the Higgs phase. Such a deformation is given by the introduction of additional Chern-Simons three-forms, as demonstrated in Sec. 2.3. This gives a clear picture on which UV contributions spoil the PQ mechanism and how to avoid these in model building.

Lastly, with an exact cancellation of $\bar{\theta}$, a measurement of the nEDM would indicate the source of CP violation to be new physics. In the ordinary implementation of the PQ mechanism, the source could not be distinguished from a partially screened $\bar{\theta}$ arising from some additional explicit breaking of the PQ symmetry. This gives a clear interpretation on potential measurements of the nEDM. Alternatively, this allows to exclude the two-form axion by experimentally identifying any type of PQ violation.

We want to stress that the two-form axion results in the same low-energy EFT, due to the duality between two-forms and pseudo-scalars in 3+1 dimensions. This is

⁸This is usually the case for axion-like particles arising in string theory. There, string compactifications generate $U(1)$ symmetries that are often spontaneously broken at the string scale [108].

in agreement with Sec. 2.3, where we introduced the topological three-form formalism in the context of the η' mesons as a language that allows to understand the strong CP problem from a different perspective. The duality on which this is based, however, only holds true for energies below f_a .

CHAPTER FOUR

COSMOLOGICAL BOUNDS ON THE AXION

So far we have seen that the axion is an excellently motivated particle candidate beyond the Standard Model. It emerges as the leading solution to the strong CP problem, which due to quantum gravity is promoted from a small-value puzzle to a consistency problem, necessitating a mechanism to remove the QCD vacuum angle. However, due to its Goldstone nature, the axion also provides an excellent dark matter candidate for the following reasons [43]: (1) it is stable, (2) it is very weakly interacting via the forces described by the SM, (3) it is subject to misalignment, a production mechanism that is able to produce the required amount of axions to account for the observed dark matter abundance [109–111]. Because of that the axion’s appeal extends beyond particle physics and encompasses fields such as astrophysics and cosmology. The axion experimental program is currently thriving, with previously unreachable regions of the axion parameter space being probed. Therefore, understanding the axion’s parameter space is crucial, and the project presented in this chapter strives to contribute to this aim.

Let us direct the attention to the region in the parameter space where the axion decay constant fulfills $f_a \gtrsim 10^{12}$ GeV. This region is usually excluded from the axion window because of overproduction from the misalignment production mechanism. If this bound was taken as cast in stone, a non-measurement in the classical window 10^9 GeV $\lesssim f_a \lesssim 10^{12}$ GeV would result in the axion’s exclusion. Alternatively, a measurement at higher values of f_a would be interpreted as an axion-like particle instead of the QCD axion. However, if it were possible to circumvent this cosmological bound, the interpretation of the measurements would differ considerably. This emphasizes the significance of comprehending the imposed bounds, particularly the possibility to evade them.

The objective of our work is to do exactly that, i.e. to provide evidence for the non-robustness of the cosmological bound in question by studying how Dvali’s mechanism

[44] can be realized in the benchmark axion models presented in Sec. 3.3. The cosmological bound is primarily based on the critical assumption that the initial misalignment angle θ_1 is of order $\mathcal{O}(1)$ and not fine-tuned. While fine-tuning itself is not problematic, it is typically considered undesirable in the absence of an explanation for its origin. Consequently, if a mechanism existed that yielded very small values of θ_1 , the parameter space with $f_a \gtrsim 10^{12}$ GeV would become natural. This would be of particular interest for string axion-like particles [108] or even the pure gauge two-form implementation of the axion [46, 112] since both favor large values of f_a . Our work is complementary to other studies of Dvali's early relaxation mechanism that have been conducted in the literature. Examples include [113] and [114], which explore its implementation in the context of supersymmetry (SUSY). In [115], the implications of Higgs inflation on the mechanism were studied. Additionally, for the related concept of low-scale inflation, which focuses on constructing a concrete QCD-scale inflation model with successful reheating, we refer to [116]. A list of alternative ideas to expand the axion window can be found in [55]. We also note that a complementary study of the same mechanism was presented in [117].

This chapter is structured as follows. To begin with, in Sec. 4.1 we introduce the two main production mechanisms for the axion, i.e. the misalignment mechanism and the decay of topological defects. Next, in Sec. 4.2 we briefly review the different phenomena that constrain the parameter space of the axion, resulting in the axion window. We specifically focus on the initial misalignment angle θ_1 and the impact of thermal and inflationary fluctuations. In Sec. 4.3, we examine the possibility of early relaxation. We begin by demonstrating how inflation can cause a shift in the QCD scale and lead to an early phase of strong QCD. Using this information, we determine the initial misalignment angle in both the KSVZ and DFSZ models by considering the inflation-induced shift in the axion's minimum. Finally, we calculate the energy density and identify the feasible region of parameter space within our model. We conclude by summarizing our findings in Sec. 4.4.

4.1 Axion Production

The axion is mainly produced by the misalignment mechanism or by the decay of topological defects. The former is basically the release of energy that is stored in the axion field since its emergence. Intriguingly, the produced axions are non-thermal as required for a dark matter candidate. The latter is based on cosmic strings and domain walls that emerge from the spontaneous breaking of the PQ symmetry. These can decay into axions, thus contributing to its cosmological abundance. Both of this production mechanism are introduced in this section.

4.1.1 Misalignment Mechanism

In the standard FLRW cosmological framework, the axion emerges at the temperature $T_{\text{PQ}} \sim v_a \gg \Lambda_{\text{QCD}}$ as the $U(1)_{\text{PQ}}$ symmetry is spontaneously broken [36, 38, 39]. For convenience, let us in the following work with the dimensionless axion field $\theta(x) = a(x)/f_a$. For large temperatures, instantons provide good degrees of freedom so that we can describe the axion potential by instantons coupled to the thermal bath. In the dilute instanton gas approximation [79], the temperature-dependent potential takes the form [118]

$$V(\theta) = m_a^2(T) f_a^2 (1 - \cos(\theta)) , \quad (4.1)$$

where for the axion mass we use

$$m_a(T) \equiv \frac{(\Lambda_{\text{QCD}}^3 m_u)^{\frac{1}{2}}}{f_a} \begin{cases} \gamma \left(\frac{\Lambda_{\text{QCD}}}{T} \right)^4 & : T > \Lambda_{\text{QCD}} , \\ 1 & : T \lesssim \Lambda_{\text{QCD}} . \end{cases} \quad (4.2)$$

Here, γ encodes QCD and active quark physics, which for the Standard Model roughly are of order 10^{-2} [79]. We observe that the axion potential is essentially flat in the beginning due to its Goldstone nature. With decreasing temperature due to the expansion of the universe, the potential becomes more significant until the temperature becomes of order $\Lambda_{\text{QCD}} \sim 0.2 \text{ GeV}$. At this point, the potential becomes independent of the surrounding thermal environment and remains unchanged.

The equation of motion for the axion in the Friedmann–Lemaître–Robertson–Walker (FLRW) background are given by

$$\ddot{\theta} + 3H(t)\dot{\theta} - \frac{1}{R^2(t)}\Delta\theta + \frac{V'(\theta)}{f_a^2} = 0 , \quad (4.3)$$

where $R(t)$ denotes the scale factor and $H(t)$ the Hubble parameter. We can get an intuitive picture by performing the following two simplification. First, as we are only concerned with small values of θ , we neglect higher-order terms in the potential expansion. Secondly, we will only focus on the zero-mode of θ . Incorporating these simplifications, the equation of motion reduces to that of a damped harmonic oscillator,

$$\ddot{\theta} + 3H(t)\dot{\theta} + m_a^2(T(t))\theta = 0 . \quad (4.4)$$

For $T \gg \Lambda_{\text{QCD}}$, the Hubble friction dominates and the axion is essentially frozen out. As long as this situation persists, the axion behaves as dark energy. Since dark energy does not redshift, the energy remains stored within the axion field. Eventually, as

$$m_a(T(t)) = 3H(t) , \quad (4.5)$$

the axion can release the built-up energy and begin performing coherent oscillations around the vacuum. After this point, the axion's equation of state changes from that of dark energy to that of non-relativistic matter, resulting in the axion contributing to

dark matter. For convenience, let us use the universe's temperature instead of cosmic time in the following, and denote the temperature when the oscillations begin as T_1 . The oscillations commence during radiation domination, so that the Hubble parameter can be expressed as

$$H(T) = \sqrt{\frac{\rho}{3M_{\text{P}}^2}} \sim \frac{T^2}{M_{\text{P}}} . \quad (4.6)$$

Together with the axion mass defined by the first row of (4.2), this results in the oscillations commencing at the temperature

$$T_1 \sim \left(\frac{10^{12} \text{ GeV}}{f_a} \right)^{\frac{1}{6}} \text{ GeV} . \quad (4.7)$$

It is worth noting that for f_a values greater than approximately $2 \times 10^{17} \text{ GeV}$, the oscillations commence after the axion has attained its low-temperature mass, resulting in a distinct expression for T_1 . Moreover, in the range of $2 \times 10^{15} \text{ GeV} \lesssim f_a \lesssim 2 \times 10^{17} \text{ GeV}$, there is a slight caveat due to the dilute instanton gas approximation's breakdown. Nevertheless, we will employ (4.7) for all $f_a < M_{\text{P}}$ for the purpose of illustration.

We can estimate the contribution to today's dark matter fraction as follows. The initial energy density of these oscillations at T_1 is

$$\rho_1 = \frac{1}{2} f_a^2 m_a^2(T_1) \theta_1^2 , \quad (4.8)$$

where θ_1 is the initial misalignment angle. This is nothing else than the amplitude of the axion oscillation at the moment when they commence.

Assuming the change in the mass to be in the adiabatic regime, the number of zero-modes per co-moving volume is conserved as the axions are essentially decoupled. Today's energy density is thus given by

$$\rho_0 = \rho_1 \frac{m_a(T_0)}{m_a(T_1)} \left(\frac{R(T_1)}{R(T_0)} \right)^3 , \quad (4.9)$$

and the associated cold dark matter fraction is

$$\frac{\Omega_a}{\Omega_{\text{CDM}}} \sim \left(\frac{f_a}{10^{12} \text{ GeV}} \right)^{7/6} \left(\frac{\theta_1}{1} \right)^2 . \quad (4.10)$$

We can observe from this basic estimate that the maximal value of f_a depends on θ_1 . For “non-special” initial conditions, i.e. $\theta_1 \approx \mathcal{O}(1)$, the upper limit for f_a is approximately 10^{12} GeV . For rather special initial conditions on the other hand, i.e. $\theta_1 \ll 1$, the upper limit for f_a exceeds 10^{12} GeV .

It is important to distinguish between two scenarios that are determined by the time when the PQ symmetry is broken. We can quantify these scenarios by introducing the following scales:

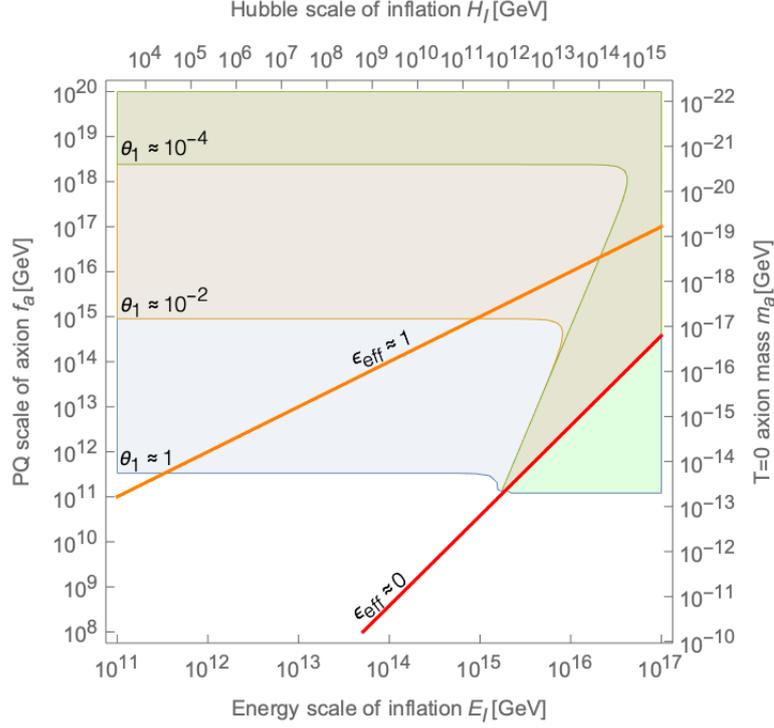


Figure 4: Cosmological bound on the axion scale f_a depending on the inflationary Hubble scale H_I . The thick red line represents $f_a = \max\{T_{\text{GH}}, T_{\text{RH}}\}$ in the case of inefficient reheating, separating the pre-inflationary and the post-inflationary scenarios. The case of very efficient reheating with $\epsilon_{\text{eff}} \approx 0$ is depicted in orange to show the influence of the thermalization efficiency on the bound. We neglect the contribution from the decay of topological defects in the post-inflationary scenario due to large uncertainties in their determination [119–121].

- The Gibbons-Hawking temperature during inflation, $T_{\text{GH}} = H_I/2\pi$, where H_I is the inflationary Hubble parameter.
- The maximum thermalization temperature [122]

$$T_{\text{Max}} = \epsilon_{\text{eff}} E_I \sim \epsilon_{\text{eff}} (M_{\text{P}} H_I)^{\frac{1}{2}}, \quad (4.11)$$

where ϵ_{eff} is used as a dimensionless efficiency parameter with $0 \lesssim \epsilon_{\text{eff}} \lesssim 1$. This should not be confused with the reheating temperature which can be somewhat lower [123].

The pre-inflationary scenario occurs when PQ symmetry is spontaneously broken during or before inflation, and is never restored afterward, i.e. $f_a > \max\{T_{\text{GH}}, T_{\text{RH}}\}$. Since the axion potential is essentially flat during its emergence, in each Hubble patch one initial value is chosen randomly from a uniform distribution on the range $[-\pi, \pi]$. Inflation then enlarges each patch so that after inflation has ended the universe starts with one homogeneous value θ_1 . This render θ_1 a free parameter in this scenario.

In contrast, the post-inflationary scenario occurs when PQ symmetry is either broken or restored after inflation, i.e. when $f_a < \max\{T_{\text{GH}}, T_{\text{RH}}\}$. In this scenario, patches with various initial values come into causal contact, resulting in all possible values appearing

in the current universe. Consequently, the initial misalignment angle in the present-day universe can be defined as the average over the uniform distribution, $\theta_1 \equiv \sqrt{\langle \theta_1^2 \rangle} = \pi/\sqrt{3}$.

We show the cosmological bound for both scenarios in Fig. 4. The red line represents $f_a = \max\{T_{\text{GH}}, T_{\text{Max}}\}$, thus separating the two scenarios in the case of very inefficient reheating ($\epsilon_{\text{eff}} \approx 0$). The post-inflationary scenario occurs below the red line, while the pre-inflationary scenario takes place above it. We also present the line that separates the two scenarios in the case of very efficient reheating ($\epsilon_{\text{eff}} \approx 1$) to demonstrate the influence of the ϵ_{eff} parameter. In the pre-inflationary scenario, we include the impact of inflationary fluctuations (discussed in Sec. 4.2.1), which cause the deviation from the horizontal lines.

4.1.2 Decay of Topological Defects

The pre-inflationary and post-inflationary scenarios are different in regard to topological defects, in particular domain walls and cosmic strings. In the pre-inflationary model, these defects are inflated away and are thus irrelevant to the subsequent evolution. In the post-inflationary scenario, however, these defects are present and add to the energy density of the axion through their decay. Therefore, we briefly explain the formation of these topological defects as a result of the PQ symmetry undergoing spontaneous breaking. This phenomenon can be understood by the Kibble-Zurek Mechanism (KZM), which describes the non-equilibrium dynamics and the creation of topological defects in a system undergoing a continuous phase transition caused by a finite rate change of a control parameter λ [124–126].

Kibble-Zurek Mechanism

Let us focus on continuous second-order phase transitions, where the KZM was developed. These phase transitions are characterized by the divergence of the equilibrium correlation length ξ and the equilibrium relaxation time τ , which usually take the form of a power-law,

$$\xi(\epsilon) = \frac{\xi_0}{|\epsilon|^\nu}, \quad \tau(\epsilon) = \frac{\tau_0}{|\epsilon|^{z\nu}}. \quad (4.12)$$

Here, ν is the correlation length critical exponent, z the dynamical critical exponent, and ϵ is the reduced distance parameter, i.e. a conveniently normalized distance to the critical point λ_c ,

$$\epsilon \equiv \frac{\lambda_c - \lambda}{\lambda_c}. \quad (4.13)$$

In terms of ϵ , the system is initially prepared in the high-symmetry phase ($\epsilon < 0$) and is forced to face a spontaneous symmetry breaking scenario as the critical point is crossed towards the degenerate vacuum manifold ($\epsilon > 0$).

In order to describe the dynamics during the phase transition, λ is assumed to be time-dependent. In the proximity of the critical point, this time-dependence can usually

be linearized,

$$\lambda(t) = \lambda_c(1 - \epsilon(t)) . \quad (4.14)$$

Taking a linear quench symmetric around the critical point,

$$\epsilon(t) = \frac{t}{\tau_Q} , \quad (4.15)$$

the distance parameter varies linearly in the interval $[-\tau_Q, \tau_Q]$ and is characterized by the quench time τ_Q . This whole approximation captures the essence of the non-equilibrium dynamics involved in the phase transition at a finite rate and is often referred to as the adiabatic-impulse approximation. This leads to the dynamics being split into three stages (see Fig. 5)

- Initial stage (dynamics adiabatic and high symmetry):
 $\lambda < 0$ and $|\lambda| \gg \lambda_c$: $\epsilon \gg 0 \implies \tau(\epsilon) \longrightarrow 0$
- Intermediate stage (dynamics frozen):
 $\lambda \sim \lambda_c$: $\epsilon \sim 0 \implies \tau(\epsilon) \longrightarrow \infty$
- Final stage (dynamics adiabatic and low symmetry):
 $\lambda > 0$ and $\lambda \gg \lambda_c$: $\epsilon \gg 0 \implies \tau(\epsilon) \longrightarrow 0$

The crucial point in the KZM is that the evolution of the system or the order parameter does not stop. The microstate still evolves as the Hamiltonian dictates and local thermodynamic equilibrium of the microscopic degrees of freedom may be maintained. However, the order parameter does not follow its equilibrium value, it can merely catch up locally. This is the reason for the formation of topological defects.

The freeze-out time \hat{t} , i.e. the boundary between stages, can be estimated by comparing the relaxation-time $\tau(t)$ with the inverse Hubble rate during radiation domination,

$$\tau(t) \stackrel{!}{\sim} t \implies \hat{t} \sim (\tau_0 \tau_Q^{z\nu})^{\frac{1}{1+z\nu}} . \quad (4.16)$$

The equilibrium value of the order parameter corresponds to the instantaneous value of ϵ in the interval $[-\hat{\epsilon}, \hat{\epsilon}]$,

$$\hat{\epsilon} = |\epsilon(\hat{t})| = \left(\frac{\tau_0}{\tau_Q} \right)^{\frac{1}{1+z\nu}} , \quad (4.17)$$

but it should be kept in mind that the true value lags behind this value.

The average length of domains set by the value of the equilibrium correlation length at $\hat{\epsilon}$ is given by

$$\hat{\xi} \equiv \xi(\hat{\epsilon}) = \xi_0 \left(\frac{\tau_Q}{\tau_0} \right)^{\frac{\nu}{1+z\nu}} . \quad (4.18)$$

This can be recast as an estimate for the density of topological defects n ,

$$n \sim \frac{\hat{\xi}^d}{\hat{\xi}^D} = \frac{1}{\xi_0^{D-d}} \left(\frac{\tau_Q}{\tau_0} \right)^{(D-d)\frac{\nu}{1+z\nu}} , \quad (4.19)$$

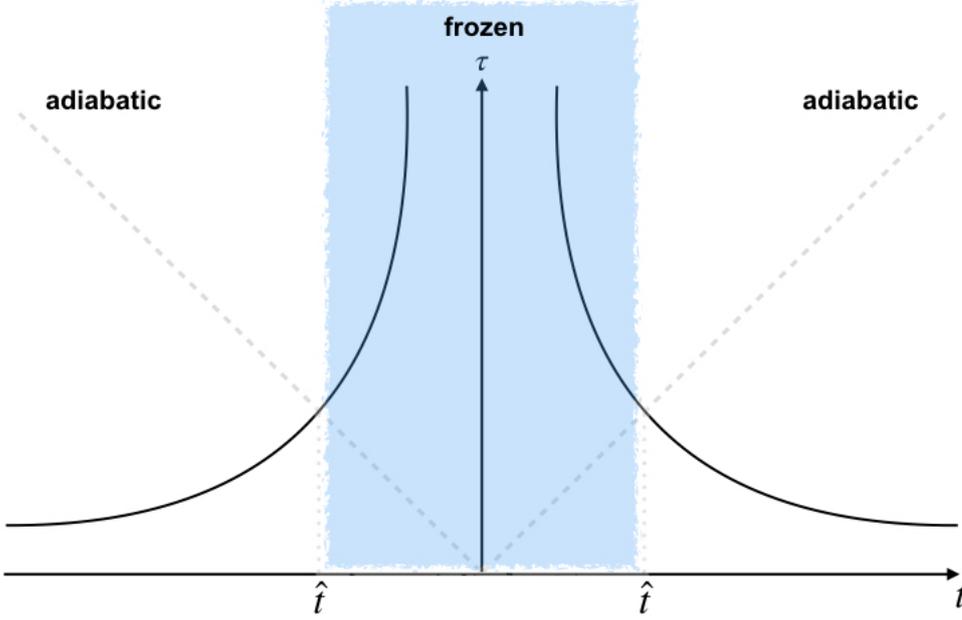


Figure 5: Schematic representation of the freeze-out captured by the adiabatic-impulse approximation.

where D denotes the dimension of space and d the dimension of the defect (for example in $D = 3$ a cosmic string has $d = 1$ and a domain wall has $d = 2$).

Cosmic Axion Strings

The vacuum manifold of the spontaneously broken $U(1)_{PQ}$ is isomorphic to S^1 . Since the first homotopy group of this vacuum manifold is $\pi_1(S^1) = \mathbb{Z}$, the low energy theory includes cosmic strings. For instance in either the KSVZ or the DFSZ model, the spontaneous breaking occurs via a standard Higgs potential of the singlet S with a coefficient λ in front of the quartic term. The energy per length L of the emerging string is given by

$$\mu_{\text{String}} \approx \pi f_a^2 \ln \left(\frac{L}{\delta} \right). \quad (4.20)$$

Here, δ denotes the thickness of the string and is given by $\delta \sim m^{-1} \sim (\sqrt{\lambda} v_a)^{-1}$, where m is the mass of the radial degree of freedom of the Higgs singlet.

The evolution of string networks is an extremely complex problem involving many different length scales, but all analyses hint that the network reaches a self-similar attractor solution on large scales in which all the properties and length scales describing the network scale with time [119, 120]. In this regime, the strings energy density can be estimated by

$$\rho_{\text{string}} \sim \frac{\mu_{\text{String}} L}{L^3} = \frac{\mu_{\text{String}}}{\gamma^2 t^2}, \quad (4.21)$$

where γ is the factor of proportionality between the characteristic length and time.

In an FLRW background, this “scaling solution” represents a dissipative solution to the equations of motion, meaning that the string decays into various forms of radiation in order to maintain its self-similarity. Possible forms of radiation include the scalar of which the string is formed, gauge fields to which the scalar is coupled to, and gravitational waves. Before the QCD phase transition, axions are practically massless and, neglecting gravitational radiation, represent the only form of radiation in the present case. Thus, the time-evolution of a string is given by

$$\frac{1}{R(t)^2} \frac{d}{dt} (\rho_{\text{string}} R(t)^2) = \dot{\rho}_{\text{string}} + 2H\rho_{\text{string}} = -P_{\text{string}} \quad , \quad (4.22)$$

$$\frac{1}{R(t)^4} \frac{d}{dt} (\rho_a R(t)^4) = \dot{\rho}_a + \underbrace{4H\rho_a}_{\text{Cosmo. redshift}} = \underbrace{P_{\text{string}}}_{\text{axion emission}} \quad . \quad (4.23)$$

Inserting (4.21) into (4.22) and assuming the universe to be radiation dominated, gives

$$P_{\text{string}} = \frac{\pi f_a^2}{\gamma^2 t^3} \left[\ln \left(\frac{\gamma t}{\delta} \right) - 1 \right] \quad . \quad (4.24)$$

Since $E_a(t) = \rho_a(t)R(t)^4$ is the comoving energy of radiated axions, (4.23) can be rewritten as

$$\frac{d}{dt} E_a = R(t)^4 P_{\text{string}} \quad . \quad (4.25)$$

Expressing the average energy of emitted axions at t' as $\langle E_a(t') \rangle = R(t')\bar{\omega}_a(t')$, the comoving number of radiated axions can be estimated as

$$\begin{aligned} N_a(t > t_{\text{QCD}}) &= \int_{t_{\text{PQ}}}^{t_{\text{QCD}}} dt' \frac{1}{\langle E(t') \rangle} \frac{d}{dt} E_a(t') \\ &= \int_{t_{\text{PQ}}}^{t_{\text{QCD}}} dt' \frac{R(t')^3}{\bar{\omega}_a(t')} \frac{\pi f_a^2}{\gamma^2 t'^3} \left[\ln \left(\frac{\gamma t'}{\delta} \right) - 1 \right] \quad . \end{aligned} \quad (4.26)$$

In the literature there is a controversy on the determination of $\bar{\omega}_a$. Naively, $\bar{\omega}_a$ should be comparable to the horizon scale at time t , i.e. $\bar{\omega}_a \sim 2\pi/t$. However, turbulent decay processes can make it large, so that it requires a cut-off at the horizon and string scale, i.e. $\bar{\omega}_a \sim 2\pi \ln(\gamma t/\delta)/t$. Modern simulations yield intermediate results [121]. Depending on the choice of $\bar{\omega}_a$, the final result is often stated as

$$\Omega_a h^2 = \Omega_{a,\text{mis}} h^2 (1 + \alpha_{\text{dec}}) \quad , \quad (4.27)$$

with $0.16 \leq \alpha_{\text{dec}} \leq 186$ [55].

String-Wall System

The situation changes when at $T \sim T_{\text{QCD}}$ the axion potential is “switched on”. Let us use the potential in (4.1) but with the canonically normalized axion and with the axion

decay constant expressed as $f_a = v_a/2\mathcal{N}$,

$$V(a) = \frac{m_a^2 v_a^2}{2\mathcal{N}} \left(1 - \cos \left(2\mathcal{N} \frac{a}{v_a} \right) \right). \quad (4.28)$$

The periodicity of the potential results in discrete vacua. The number of these vacua in a single 2π -loop is the domain wall number N_{DW} , which can be read off to be $N_{\text{DW}} = 2\mathcal{N}$. In a symmetry language, this potential explicitly breaks the $U(1)_{\text{PQ}}$ symmetry into a $\mathbb{Z}_{N_{\text{DW}}}$ subgroup under which the axion transforms as $a \rightarrow a + 2\pi n f_a$ with $n \in \mathbb{Z}$. The DW number can also be determined from this transformation as the value of n that corresponds to a single loop with a circumference of $2\pi v_a$, which again results in $N_{\text{DW}} = 2\mathcal{N}$. The $\mathbb{Z}_{N_{\text{DW}}}$ is spontaneously broken by one of the N_{DW} vacua. Due to this spontaneous breaking of a discrete symmetry [127], domain walls emerge and attach themselves to the already present strings in such a way that the intersection point of domain walls is given by the cosmic strings.

After the QCD phase transition, the tension of the strings dominates over the tension of the domain walls σ_{DW} . However, after the time t_{DW} , defined by $\sigma_{\text{DW}} \sim \mu_{\text{string}}/t_{\text{DW}}$, the tension becomes comparable. the fate of the string-wall system then depends on the value of N_{DW} . In particular, there are two cases:

- $N_{\text{DW}} = 1$, short lived domain walls

One domain wall is attached to each string. These configurations are unstable and decay soon after due to the dominating domain wall tension. The energy density at t_{QCD} is $\rho_{\text{DW}} \sim \sigma_{\text{DW}} L^2/L^3 \sim \sigma_{\text{DW}}/t_{\text{QCD}}$. Assuming that their energy density is well approximated by extrapolating the previous energy densities to t_{DW} , the total energy density can be estimated to be

$$\rho_{\text{string-wall}} \sim \frac{\sigma_{\text{DW}}(t_{\text{DW}})}{t_{\text{DW}}} + \frac{\mu_{\text{string}}(t_{\text{DW}})}{t_{\text{DW}}^2}. \quad (4.29)$$

The total number density of axions produced by the decay of string-wall systems is thus

$$n_{a,\text{dec}}(t) = \frac{\rho_{\text{string-wall}}}{\bar{\omega}_a} \left(\frac{R(t_{\text{DW}})}{R(t)} \right)^3, \quad (4.30)$$

which can be used to calculate $\Omega_{a,\text{dec}} h^2$.

- $N_{\text{DW}} > 1$, long lived domain walls

These configurations are stable because the strings are sustained by the tension of walls from multiple directions. For $t > t_{\text{DW}}$ the strings energy density can be ignored and the total energy density is given by $\rho_{\text{DW}}(t) \sim \sigma_{\text{DW}}/t \propto R(t)^{-2}$. Therefore, domain walls redshift slower than dust ($\propto R(t)^{-3}$) and radiation ($\propto R(t)^{-4}$) during radiation domination, leading to an overclosure of the universe. This is the so called domain wall problem [128] (see [129] for a review).

Before moving to the next section, let us discuss a caveat in theories where the axion is a linear combination of angular modes a_f from three or more scalars [130]. For

concreteness, we use the DFSZ model as an example. There, the explicit breaking of each angular mode results in a residual cyclic symmetry, $a_f \rightarrow a_f + 2\pi n_f v_f$, where n_f is an integer. To account for these residual symmetries, we apply the cyclic symmetry transformations to both sides of the first equation in (3.29) and define N_{DW} as previously mentioned,

$$N_{\text{DW}} = 2\mathcal{N} \frac{\sum n_f \chi_f v_f^2}{\sum \chi_f^2 v_f^2}. \quad (4.31)$$

To ensure that N_{DW} is an integer, we need to demand that the fraction in the expression for N_{DW} is also an integer, which can be achieved by setting it equal to one without loss of generality. One possible way to do this is to require $n_f = \chi_f$, which imposes the compactness of each $U(1)$ and is thus highly restrictive. However, there is a less restrictive alternative. We can use the orthogonality condition to eliminate one of the v_f terms in the numerator and denominator of the fraction, and then compare the resulting terms with the same v_f^2 . Let us consider DFSZ₂ for simplicity, where $f = u, d, S$, and demand that the fraction is equal to one. We find

$$n_S = \chi_S, \quad (4.32)$$

$$n_u + n_d = \chi_u + \chi_d = 2\chi_S, \quad (4.33)$$

where in the second equality we used the PQ invariance from the unique V_{eb} term $H_u H_d S S$ (up to complex conjugation of both singlets). We observe that for N_{DW} to be an integer in the presence of residual cyclic symmetries, χ_S and certain combinations of PQ charges must also be integers. This is not very restrictive for the DFSZ model, however, it can become more restrictive for more complicated models as the ones we will consider in Chapter 6. There, we will encounter models where χ_S has a minimum integer value larger than one.

In summary, when the axion is a linear combination of fields, the DW number can still be expressed as $N_{\text{DW}} = 2\mathcal{N}$, but with the condition that the fraction in (4.31) is equal to one. This requirement implies that χ_S must be an integer, but not necessarily equal to one. In regard of the domain wall problem, it is convenient to choose the minimal possible value of χ_S . Therefore, we define the domain wall number as

$$N_{\text{DW}} = \min \text{positive integer } \{2\mathcal{N}\} \quad (4.34)$$

and use this definition for the remainder of this thesis.

4.2 The Axion Window

The axion production mechanism results in an upper bound on the scale f_a or, equivalently, a lower bound on the axion mass m_a . However, this is not the only bound on the parameters of the axion. In this section, we give a brief overview over the various bounds and discuss how model-dependent these are. All bounds are shown in Fig. 6, where the

non-excluded region is the axion window (more details on the various bounds can for instance be found in [55]).

4.2.1 Isocurvature Fluctuations

In the pre-inflationary scenario, the axion exists during inflation as an essentially massless field. As such, it is subject to quantum fluctuations. Let us write the axion field as $\theta = \langle \theta \rangle + \delta\theta$, where $\langle \theta \rangle = \theta_1$ and $\delta\theta$ denote the quantum fluctuations. These fulfill $\langle \delta\theta \rangle = 0$ and have a standard deviation of

$$\sigma_\theta \sim \sqrt{\langle \delta\theta^2 \rangle} \sim \frac{H_1}{2\pi f_a}. \quad (4.35)$$

These fluctuations displace the axion value during inflation, resulting in the modification of (4.10) by $\theta_1^2 \rightarrow \theta_1^2 + \sigma_\theta^2$. For large values of H_1 the quantum fluctuations can dominate over the initial misalignment angle, resulting in the deviation from the horizontal lines in Fig. 4.

In addition, if the axion acts as a spectator field during inflation, its fluctuations will not be adiabatic but of isocurvature-type. Isocurvature fluctuations have a distinct signature in the temperature and polarization fluctuations of the CMB, which places constraints on the parameter space of axions. Assuming the fluctuations to be normal distributed in the regime of small θ_1 , where anharmonic corrections of the axion potential can be ignored, the amplitude of the axions isocurvature fluctuations is given by [131]

$$\Delta_a(k_0) = \frac{\delta\Omega_{\text{DM}}}{\Omega_{\text{DM}}} = \frac{\Omega_a}{\Omega_{\text{DM}}} \frac{\delta \ln \Omega_a}{\delta\theta_1} \sigma_\theta = \frac{\Omega_a}{\Omega_{\text{DM}}} \frac{H_1}{\pi\theta_1 f_a},$$

The latest experimental bound on uncorrelated isocurvature perturbations by Planck is [40]

$$\frac{\Delta_a^2(k_0)}{\Delta_{\mathcal{R}}^2(k_0) + \Delta_a^2(k_0)} < 0.038 \quad \text{at 95\% CL}, \quad (4.36)$$

where $k_0 = 0.050 \text{ Mpc}^{-1}$ and $\Delta_{\mathcal{R}}(k_0) = (2.196 \pm 0.060) \times 10^{-9}$ denotes the measured amplitude of scalar perturbations. This implies

$$H_1 \lesssim 10^7 \left(\frac{\Omega_a}{\Omega_{\text{CDM}}} \right)^{-1} \left(\frac{f_a}{10^{12} \text{ GeV}} \frac{\theta_1}{1} \right) \text{ GeV}. \quad (4.37)$$

We want to stress that this bound only applies in the pre-inflationary scenario when the axion is a massless spectator field during inflation. Any violation of these conditions reduces or even nullifies this bound.

4.2.2 Primordial Gravitational Waves

The pre-inflationary scenario depends on the scale of inflation and, as such, is subject to the constraints on these. The most robust and model-independent constraint comes

from tensor fluctuations, which develop into primordial gravitational waves. Their amplitude $\Delta_t = 2H_1^2/(\pi^2 M_{\text{P}}^2)$ only depends on the expansion rate during inflation and thus represents a bound of the parameter space to the right. It is convenient to express the bound in terms of the tensor-to-scalar ratio, $r \equiv \Delta_t/\Delta_{\mathcal{R}}$, which is just the tensor amplitude normalized with respect to the scalar amplitude. The bound given by the Planck-BICEP2 joint analysis is $r < 0.063$ [40], implying

$$H_1 < 6.1 \times 10^{13} \text{ GeV} . \quad (4.38)$$

4.2.3 Thermal Fluctuations

After inflation has ended and the universe is reheated, the axion is essentially frozen out until the QCD phase transition. Throughout this period, the axion is exposed to thermal fluctuations that cause it to deviate from its value at the end of inflation [44]. Since the potential is essentially flat during that time, its movement can be considered a random walk with a step size of $\Delta\theta \sim T/f_a$ per Hubble time. In the pre-inflationary setting, the axion is present right from the onset of the thermal universe at the reheating temperature T_{RH} . Consequently, the maximum deviation until the QCD phase transition is

$$\sigma_{\theta}^{\text{thermal}} \sim \frac{T_{\text{RH}}}{f_a} . \quad (4.39)$$

The influence of the thermal fluctuations on the misalignment can be captured by the same replacement as the quantum fluctuations, i.e. $\theta_1^2 \rightarrow \theta_1^2 + (\sigma_{\theta}^{\text{thermal}})^2$. In the pre-inflationary scenario, $f_a \gg T_{\text{RH}}$, so that thermal fluctuations are negligible compared to $\theta_1 \sim \mathcal{O}(1)$. However, as the quantum fluctuations, they can become important for $\theta_1 \ll 1$ and can even dominate for certain values of the parameters.

4.2.4 Astrophysics

A lower bound on f_a is obtained by studying astrophysical consequences of the axion [132]. In particular, the axions-photon or any axion- matter couplings induce a new cooling channel that significantly alters the stellar evolution and cooling efficiency of various stars. The reason for such efficient cooling is the weakness of the couplings due to the suppression by f_a . Consequently, a produced axion within the star is very likely to leave it. This is in contrast to photons, which get scattered a huge amount of time until they leave the star. While this is a strong simplification for the complicated physics that take place in stars, it provides an intuitive picture that captures the essence. The most stringent bounds come from the cooling of red giants and the neutrino flux of SN 1987A, excluding for the benchmark models the region

$$f_a \lesssim 10^9 \text{ GeV} . \quad (4.40)$$

It is important to note, that this bound only holds for the benchmark models or models with similar couplings. As presented in Sec. 3.2.1, all axion couplings get contributions

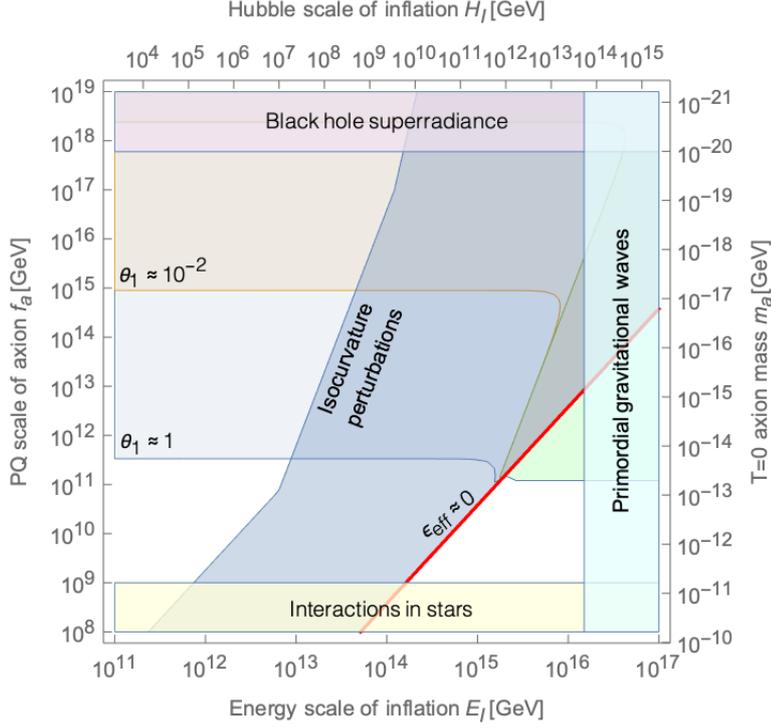


Figure 6: Axion window with the bounds coming from isocurvature perturbations (darker blue), black hole super-radiance (purple), interactions in stars (yellow), primordial gravitational waves (cyan) and overproduction (light blue or brown). See [55] for further details on each constraint.

from several terms, allowing for cancellations or at least strong suppression in more complicated models. These are called astrophobic axion models and they allow for lower values of f_a .

4.2.5 Black Hole Super-radiance

The last bound we want to mention is black hole super-radiance [133, 134]. This phenomenon describes the amplification of a wave’s amplitude by scattering from a black hole, extracting angular momentum and energy from the black hole in the process. For massive bosons like the axion, the boson mass acts as a mirror that traps the incoming wave to the vicinity of the black hole. The super-radiance phenomenon can then occur repeatedly, leading to an instability that drastically reduces the angular momentum of the black hole and leads to potential gravitational wave and γ -ray signatures. In particle terms, the phenomenon of super-radiance leads to the axions binding to the black hole, forming a “gravitational atom”. Angular momentum and energy conservation lead to an exponential growth of the occupation number of the atomic levels, essentially forming an axion Bose-Einstein condensate cloud around the black hole. When the attractive axion self-interactions dominate the gravitational binding energy, the cloud collapses in a phenomenon called “bosenova”. By analyzing the masses and spins of observed black holes, a limit can be set on the mass of the boson. In the case of the axion, this excludes

the range where

$$6 \times 10^{17} \text{ GeV} \lesssim f_a \lesssim 10^{19} \text{ GeV} . \quad (4.41)$$

This bound can be lowered by future experiments such as Advanced Laser Interferometer Gravitational-Wave Observatory (LIGO), which can detect gravitational wave signals resulting from a bosonova of a QCD axion cloud with f_a as low as the grand unified theory (GUT) scale.

4.3 Early Relaxation using Inflation

The upper boundary of the axion window at $f_a \sim 10^{12}$ GeV is crucially based on the initial misalignment angle θ_1 being $\mathcal{O}(1)$. However, smaller values of θ_1 emerge naturally when the axion is able to relax to its minimum in the early universe. Such an early relaxation mechanism as proposed by Dvali [44], allows to shift the bound on the axion scale to $f_a \gg 10^{12}$ GeV.

4.3.1 General Idea

As proposed in [44], we use inflation to generate a temporal phase of strong QCD. Such a phase occurs during inflation when the QCD scale is larger than the inflationary Gibbons-Hawking temperature,

$$\Lambda_{\text{QCD}} \gtrsim H_{\text{I}} . \quad (4.42)$$

As in the late universe, the QCD phase transition results in the axion developing its potential and relaxing to its minimum.

The efficiency of this relaxation is controlled by the axion mass m_a , the inflationary Hubble parameter H_{I} and the number of e-folds \mathcal{N}_e . If for instance $m_a \gtrsim H_{\text{I}}$, the Hubble friction is overcome and the relaxation to the minimum is very efficient. As an additional feature of this scenario, the axion does not generate any isocurvature perturbations, as these perturbations only arise for massless particles with respect to the inflationary Hubble parameter. If $m_a \lesssim H$, the relaxation occurs when the phase of strong QCD lasts long enough. To see this, we solve the axion equation of motion in this case, yielding

$$\theta(t_f) \sim \theta(t_i) \exp\left(-\mathcal{N}_e \frac{m_a^2}{3H_{\text{I}}^2}\right) . \quad (4.43)$$

where $\theta(t_i) \sim \mathcal{O}(1)$ denotes the initial misalignment angle during inflation. For the relaxation to be efficient, the initial phase of strong QCD must last for

$$\mathcal{N}_e \gg \frac{3H_{\text{I}}^2}{m_a^2} . \quad (4.44)$$

In both situations the potential vanishes at the end of this early phase of strong QCD and the axion remains frozen out in the vicinity of its potential minimum. When the

“ordinary” QCD phase transition activates the axion potential again in the late universe, the resulting amplitude of the oscillations is naturally small, i.e. $\theta_1 \ll 1$.

4.3.2 The QCD Scale during Inflation

The condition (4.42) makes the impression that early relaxation can only take place for small values of H_I . This impression, however, is based on the assumption that parameters such as m_a or Λ_{QCD} are unaltered during inflation and the history of the universe. While this assumption seems natural, in fact, it is not.

In generic supergravity and superstring theories, the strong coupling constant is usually fixed by the VEV(s) of some scalar field(s). The minima of those scalars during inflation, in general, do not coincide with their late universe minima due to thermal and quantum fluctuations but also due to a displacement of the minima themselves. This could result in $\alpha_s \gg 1$ during inflation, thus making QCD strong with a Λ_{QCD} different from its late universe value [44].

The change of the QCD scale can also occur indirectly from the effect of displaced scalar VEVs. In the SM, for instance, the VEV of the Higgs H can be displaced as well. For example, consider the case of massive chaotic inflation, i.e. $V(\Phi) = m^2\Phi^2/2$, together with the Higgs-portal coupling

$$\Delta V = -\frac{1}{2}\kappa\Phi^2 H^\dagger H, \quad (4.45)$$

where κ is some positive constant that fulfills $\kappa \lesssim 10^{-6}$ for quantum corrections to be negligible [135]. From the Friedmann-equations in the slow-roll approximation we obtain $3H_I^2 M_{\text{P}}^2 = m^2\Phi^2/2$, so that the potential becomes

$$\Delta V = -\frac{1}{2}\kappa \left(6\frac{H_I^2}{m^2}\right) H^\dagger H. \quad (4.46)$$

We observe that the Higgs receives a curvature of order H_{Inf} , displacing its VEV. Alternatively, a displacement occurs when higher dimensional operators of the inflaton and the Higgs are included, e.g.

$$\Delta V = \frac{\lambda}{\Lambda^2}\Phi^4 H^\dagger H, \quad (4.47)$$

where $\lambda < 1$ is a coupling constant and Λ is some UV cutoff, say M_{P} . In large field inflation, where the Higgs takes value of order of M_{P} , the Higgs VEV becomes of order of $\lambda^{1/2}M_{\text{P}}$. Motivated by these examples, the Higgs VEV can be significantly larger during inflation. This would have a considerable impact on the masses of the quarks, since these are directly proportional to the Higgs VEV. This change in turn affects the running of α_s and, consequently, Λ_{QCD} . According to (4.2), a changed value of Λ_{QCD} automatically results in a modified value of m_a .⁹

⁹Even though we will focus on the implications from the changed value Λ_{QCD} in the early universe on

Let us now turn to the influence of inflation on Λ_{QCD} in the KSVZ and the DFSZ models (see Sec. 3.3). Since in these two models the strong coupling constant is not controlled by the VEV of some moduli field, the shift comes purely from the displaced Higgs VEV. The first order β -function of QCD in dimensional regularization is given by

$$\beta \equiv \mu \frac{d\alpha_s}{d\mu} = -\frac{\alpha_s^2}{2\pi} b, \quad (4.48)$$

where

$$b = 11 - \frac{2}{3}n_f. \quad (4.49)$$

Here, n_f denotes the number of active quarks in the energy range under consideration. In the late universe, given the measured value $\alpha_s(M_Z) = 0.1179$ [137], the differential equation can be solved in the energy range $m_b \leq E \leq m_t$ where $n_f = 6$. Taking the value of the solution at m_b as new initial condition, the differential equation can be solved in the energy range $m_c \leq E \leq m_b$ where $n_f = 5$, and so on. The QCD scale is identified with the scale where $\alpha_s = 1$ and is $\Lambda_{\text{QCD}} \sim 0.28 \text{ GeV}$ at first loop. Note that a heavy quark needs to be included in the running of the KSVZ model. Since this heavy quark is obtaining its mass from the singlet S , its mass is assumed to be roughly of order of f_a . This influences the running only above f_a , so that it plays no role in the determination of the QCD scale in the late universe.

During inflation the quark masses are much larger due to the shifted Higgs VEV. Consequently, the quark are integrated out at much higher energies. Since the running of α_s becomes steeper with less active quarks, the emerging Landau pole is located at a much higher energy than with the smaller quark masses in the late universe. Let us indicate parameters during inflation with the label ‘‘Inf’’ in the following.

In order to define the quark masses during inflation, we take into account the running of the Yukawa couplings in addition to the changed value of the Higgs VEV. Usually, the mass of the (heavy) quarks is defined at the corresponding mass scale in the $\overline{\text{MS}}$ scheme, i.e. $m_q \sim v_H y_q(m_q)$, where v_H is the Higgs VEV and y_q is the corresponding Yukawa coupling. We take this as the defining relation and solve it numerically by making the approximation $y_q^{\text{Inf}} \sim y_q^{\text{today}}$. Incorporating the Yukawa coupling in this way, gives a slightly more precise definition of the quark masses than only using the shifted Higgs VEV, but it turns out that the final influence on the QCD scale is negligible.

Moreover, in the ordinary low energy situation the measured value of $\alpha_s(M_Z)$ served as the initial condition for the running. This condition is no longer applicable during inflation due to the changed quark masses. Hence, we need to use as initial condition the value of α_s at a higher scale that is not affected by the larger quark masses. This, for instance, could be the GUT scale $\Lambda_{\text{GUT}} \sim 10^{16} \text{ GeV}$ or the Planck scale M_{P} . Since the value of α_s at such a large scale crucially depends on the field content, the particular value is very model dependent. In the DFSZ model without additional fields, the

the axion, it should be noted that this could have interesting consequences on other theories. For example, this would allow for a production of primordial black holes from the confinement of our QCD as discussed in [136].

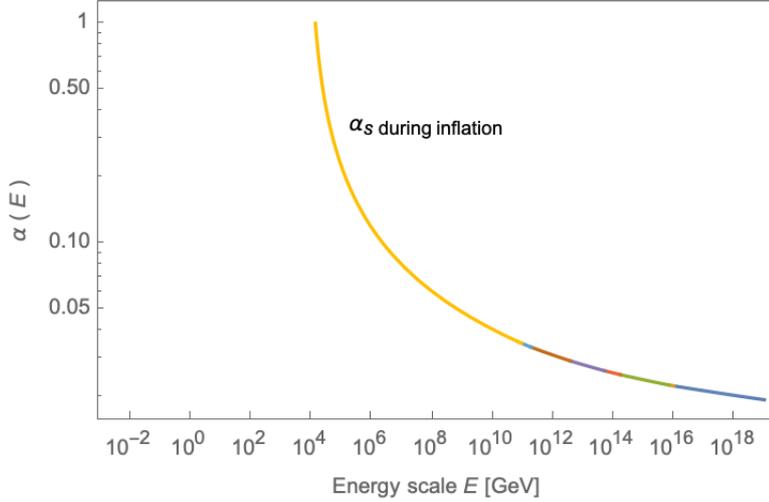


Figure 7: Running of α_s during inflation in the KSVZ model with the inflationary Higgs VEV chosen to be $v_H^{\text{Inf}} \sim M_{\text{P}}$ and the initial condition $\alpha_s^{\text{Inf}}(M_{\text{P}}) = \alpha_s(M_{\text{P}})$. The different colors indicate the parts with a different number n_f of active quarks. We see that the QCD scale is much larger compared to its value in the late universe.

running is purely controlled by the known quarks and we have $\alpha_s(\Lambda_{\text{GUT}}) \sim 0.0224$ or $\alpha_s(M_{\text{P}}) \sim 0.0200$. For the KSVZ model, the heavy quark can have an influence if its mass is smaller than the energy at which the new initial condition is defined. Taking its mass to be not too far off from the GUT scale, i.e. $m_Q \sim \Lambda_{\text{GUT}}$, its influence is rather small and the initial conditions of the DFSZ and the KSVZ models coincide. It should also be noted that the KSVZ singlet could be coupled to the inflaton itself, leading to a shift of the mass m_Q during inflation.

In order to determine the QCD scale, we solve (4.48) in the same manner as in the late universe case but incorporate the changes we just discussed. For illustration, let us use the following set of parameter in the KSVZ model. First, the mass of the heavy KSVZ quark is fixed at $m_Q \sim \Lambda_{\text{GUT}}$. Next, the late time value of α_s at M_{P} is used as initial condition. Lastly, the value of the inflationary Higgs VEV is chosen to be $v_H^{\text{Inf}} \sim M_{\text{P}}$. We depict the resulting running of α_s in Fig. 7. As can be read of, with these parameters the QCD scale is located at $\Lambda_{\text{QCD}}^{\text{Inf}} \sim 10^5$ GeV. We repeat this procedure for various values of the inflationary Higgs VEV. The resulting values of the QCD scale are plotted in blue in Fig. 8 and can be fitted by

$$\Lambda_{\text{QCD}}^{\text{Inf}}(v_H) \sim 10^5 \left(\frac{v_H}{M_{\text{P}}} \right)^{5/14} \text{ GeV} . \quad (4.50)$$

Our analysis shows that for the given choice of parameters and initial conditions the QCD scale during inflation is significantly changed. This effect is known in the literature [115]. The different exponent in our calculation comes from including the running of Yukawa couplings.

The analysis can be repeated for the DFSZ model. By comparing the running of

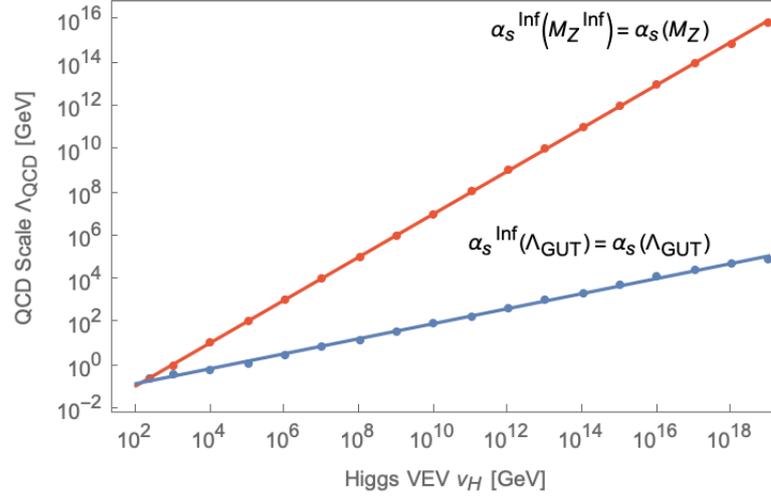


Figure 8: Values of the QCD scale for various values of the inflationary Higgs VEV v_H^{Inf} and the corresponding linear fit. The red points have $\alpha_s^{\text{Inf}}(m_Z^{\text{Inf}}) = \alpha_s(m_Z)$ as initial condition for the running while the blue points have $\alpha_s^{\text{Inf}}(\Lambda_{\text{GUT}}) = \alpha_s(\Lambda_{\text{GUT}})$. In both the KSVZ quark mass was fixed to be of order of $f_a \sim \Lambda_{\text{GUT}}$.

both models (with the fixed mass of the KSVZ quark $m_Q \sim f_a$), we find that the KSVZ quark has negligible influence on the running of α_s for the parameter space region where $f_a > 10^{12}$. This conclusion is not changed when a coupling between the inflaton and the KSVZ singlet is included so that f_a gets shifted during inflation. To be more precise, the influence on the running gets even more negligible.

In order to demonstrate the impact of the initial condition on the value of the QCD scale, we solve the ODE with an alternative initial condition, say $\alpha_s^{\text{Inf}}(m_Z^{\text{Inf}}) = \alpha_s(m_Z)$. Of course additional fields are required for this but let us calculate the quark controlled running to illustrate the influence of the initial condition. The resulting values are plotted in red in Fig. 8 and this time can be fitted by

$$\Lambda_{\text{QCD}}^{\text{Inf}}(v_H) \sim 10^{15} \left(\frac{v_H}{M_{\text{P}}} \right) \text{ GeV} . \quad (4.51)$$

4.3.3 Location of Minima during Inflation

While the concept of early relaxation seems intuitive, its implementation must take into account that the minima of the axion during both strong QCD phases must coincide. If this was not the case, the axion would relax to its early time minimum but when the ordinary QCD phase transition takes place, the axion is not necessarily located close to its true minimum. Fine-tuning would then be necessary, making the proposed mechanism no more advantageous than simply fine-tuning $\theta_1 \ll 1$. For instance, this is the case when the axion mass is dominated by a temporal additional contribution in the early universe. The crucial word here is additional. The axion then relaxes to its minimum at that time but the minimum is in general very different from the true minimum in the

late universe.

In order to achieve the coincidence of the minima, it is necessary that the axions potential in the early universe arises in the same way as in the late universe, namely by non-perturbative QCD effects. This, however, is not enough since physics that give rise to this early phase of strong QCD can also change the location of the minimum. Thus, the necessary condition is that the early phase of QCD emerges in such a way that the axions minimum during that phase sufficiently coincides with the minimum in the late universe. In the following, we calculate the displacement of the minimum for the KSVZ and the DFSZ axions during inflation. We perform the calculation with all parameters taken to be different during inflation and, for a first approximation, assume a step-function like transition from the inflationary values to the later ones. To provide a clearer representation, we will use the label “0” to indicate a value in the late universe, in addition to the label “Inf” that we have used so far to represent a value during inflation.

The location of the minima is seen most conveniently by considering the axion Lagrangian in the presence of the θ -term,

$$\mathcal{L} \supset \left(\bar{\theta} + \frac{\tilde{a}}{f_a} \right) \frac{\alpha_s^2}{8\pi} G_{\mu\nu}^a \tilde{G}^{a\mu\nu}, \quad (4.52)$$

where $\bar{\theta}$ is the CP violating parameter with the contribution from QCD and the quark masses,

$$\bar{\theta} = \theta_{\text{QCD}} + \arg \det y_u y_d. \quad (4.53)$$

After QCD gets strong, this gives rise to a potential that has its minimum at $\langle \tilde{a} \rangle = -\bar{\theta} f_a$ [78] as dictated by the Vafa-Witten theorem. The physical axion field is then identified as

$$a = \tilde{a} - \langle \tilde{a} \rangle. \quad (4.54)$$

Due to the values of the parameters evolving throughout the cosmological evolution, the minima are different and thus the physical axion field itself. Regarding the original field \tilde{a} , however, it depends on the particular axion model.

In the case of the KSVZ axion, the definition of \tilde{a} is the same at any time since it does not depend on any additional parameter. Hence, we can write down the axion VEV in the late universe and during inflation as

$$\langle \tilde{a} \rangle_0 = -\bar{\theta}_0 f_0^a, \quad (4.55)$$

$$\langle \tilde{a} \rangle_{\text{Inf}} = -\bar{\theta}_{\text{Inf}} f_{\text{Inf}}^a. \quad (4.56)$$

Using these, the resulting displacement of the vacua $\Delta\theta$ is given by

$$\Delta\theta_{\text{KSVZ}} \equiv \left| \frac{\langle \tilde{a} \rangle_{\text{Inf}} - \langle \tilde{a} \rangle_0}{f_0^a} \right| = \left| \bar{\theta}_0 - \bar{\theta}_{\text{Inf}} \frac{f_{\text{Inf}}^a}{f_0^a} \right|. \quad (4.57)$$

Let us now turn to the DFSZ axion. Here the situation is more involved due to the original axion field \tilde{a} depending on the VEVs of the two Higgs doublets H_u and H_d and

the singlet S . For concreteness, let us take the DFSZ-I model. Even though Equations (4.55) and (4.56) are also valid for the DFSZ model, calculating the displacement by just subtracting the VEVs from each other, as in the KSVZ case, does not give the correct initial amplitude for today's axion \tilde{a}_0 . The reason is that the axion potential in the late universe emerges along \tilde{a}_0 . Hence, it is the projection on \tilde{a}_0 that results in the correct amplitude for the late time oscillations,

$$\Delta\theta_{\text{DFSZ}} \equiv \left| \frac{\langle \tilde{a} \rangle_{\text{Inf}} |_{\tilde{a}_0} - \langle \tilde{a} \rangle_0}{f_a^a} \right| = \left| \bar{\theta}^0 - \bar{\theta}^{\text{Inf}} \frac{f_a^{\text{Inf}}}{f_a^0} P \right|. \quad (4.58)$$

Here, we introduced the projection factor P , which is given by

$$\begin{aligned} P &= \left(\frac{\chi_u^{\text{Inf}} v_u^{\text{Inf}}}{v_a^{\text{Inf}}} \right) \left(\frac{\chi_u^0 v_u^0}{v_a^0} \right) + \left(\frac{\chi_d^{\text{Inf}} v_d^{\text{Inf}}}{v_a^{\text{Inf}}} \right) \left(\frac{\chi_d^0 v_d^0}{v_a^0} \right) + \left(\frac{v_S^{\text{Inf}}}{v_a^{\text{Inf}}} \right) \left(\frac{v_S^0}{v_a^0} \right) \\ &\sim \left(\frac{v^{\text{Inf}} v^0}{v_a^{\text{Inf}} v_a^0} \right) + \left(\frac{v_S^{\text{Inf}}}{v_a^{\text{Inf}}} \right). \end{aligned} \quad (4.59)$$

In the second row we used $v_S^0 \sim v_a^0$ for today's axion to be invisible and, in addition, we assumed both doublets to get similar shifting from the inflaton so that $v_u^{\text{Inf}} \sim v_d^{\text{Inf}} \sim v^{\text{Inf}}$. The first term is strongly suppressed compared to the second one due to the factor $v^0/v_a^0 \ll 1$. Hence, the difference between the KSVZ and DFSZ models comes down to the factor $P \sim v_s^{\text{Inf}}/v_a^{\text{Inf}} \lesssim 1$.

Within the KSVZ and the DFSZ models $\bar{\theta}$ is per se a free parameter but let us assume it takes its natural value following from its renormalization [25],

$$\bar{\theta} \sim \left(\frac{\alpha}{\pi} \right)^2 s_1^2 s_2 s_3 \sin \delta \frac{m_s^2 m_c^2}{m_W^2} \sim 10^{-16}. \quad (4.60)$$

Here m_s is the mass of the strange quark, m_c is the mass of the charm quark, and m_W is the mass of the W-boson. Furthermore, the $s_k \equiv \sin \phi_k$ are the mixing angles and δ is the CP odd phase of the CKM matrix in the original parameterization of Kobayashi and Maskawa. Evaluated during inflation, the running is not that much altered since the factor including the masses is independent of the Higgs VEV, the running of the Yukawa couplings is very slow, and the parameters of the CKM matrix are unchanged. Hence, it is reasonable to take $\bar{\theta}_{\text{Inf}} \sim \bar{\theta}_0 \sim 10^{-16}$. To conclude, the displacement is extremely small for both benchmark models.

4.3.4 The Modified Axion Window

Since the axion's inflationary vacuum essentially coincides with its late time vacuum for the KSVZ and the DFSZ models, the value of θ_1 is determined by the efficiency of the relaxation. Thus, for sufficiently long inflation, the region with $H_I \lesssim \Lambda_{\text{QCD}}^{\text{Inf}}$ becomes viable. Without further field content than the one necessary for the KSVZ and the DFSZ models, the maximum inflationary QCD scale is $\Lambda_{\text{QCD}}^{\text{Inf}} \sim 10^5 \text{ GeV}$.

Including all fluctuations, the final axion abundance is given by

$$\frac{\Omega_a}{\Omega_{\text{CDM}}} \sim \left(\frac{f_a}{10^{12} \text{ GeV}} \right)^{7/6} \left[\theta(t_f)^2 + \sigma_\theta^2 + \left(\sigma_\theta^{\text{thermal}} \right)^2 \right]. \quad (4.61)$$

When the relaxation is extremely efficient, it creates an intriguing situation where the dark matter density is dominated by quantum or thermal fluctuations, which in turn depend on the reheating temperature T_{RH} and on the inflationary Hubble scale H_{I} . In the case when either of the types of fluctuations dominates, we can use (4.61) to find the parameter space where the axion makes up all the dark matter, i.e.

$$H_{\text{I}} \sim 10^{13} \left(\frac{f_a}{10^{12} \text{ GeV}} \right)^{5/6} \text{ GeV} : \quad H_{\text{I}} \gg T_{\text{RH}} \quad (4.62)$$

$$T_{\text{RH}} \sim 10^{12} \left(\frac{f_a}{10^{12} \text{ GeV}} \right)^{5/6} \text{ GeV} : \quad H_{\text{I}} \ll T_{\text{RH}}. \quad (4.63)$$

When quantum fluctuations dominate over the post-inflationary thermal fluctuations, it is not possible to reach values for f_a that are much larger than 10^{12} GeV. This is due to the upper bound on H_{I} from inflationary gravitational waves as dictated by (4.38). Hence, for such an efficient relaxation the cosmological bound can only be avoided by large reheating temperatures. In particular, using the maximum thermalization temperature defined in (4.11) as an upper bound for the reheating temperature, allows to give a constraint on f_a in terms of H_{I} ,

$$f_a \lesssim 10^{15} \left(\frac{\epsilon_{\text{eff}}}{1} \right)^{6/5} \left(\frac{H_{\text{I}}}{10^{11} \text{ GeV}} \right)^{3/5} \text{ GeV}. \quad (4.64)$$

We show the maximum of this bound by a blue line in Fig. 9 for $\epsilon_{\text{eff}} \sim 1$ (top) and for $\epsilon_{\text{eff}} \sim 10^{-1}$ (bottom). The viable parameter space is defined by the region enclosed by the blue and red lines. For $\epsilon_{\text{eff}} \sim 10^{-2}$ there is already no more parameter space left. Overall we can say, that efficient relaxation of θ_1 results in a very narrow allowed region.

This region lies in the isocurvature constrained region of the parameter space. Hence, it is only allowed if the axion avoids the isocurvature constrain for which there are several known possibilities (see [55] for an overview). One of those possibilities is given by the fact that if the axion acquires a sufficiently high mass the constrain is relaxed [138]. In this case, the QCD scale during inflation has to fulfill not only $H_{\text{I}} \lesssim \Lambda_{\text{QCD}}^{\text{Inf}}$ in order to make QCD strong but the more restrictive condition

$$H_{\text{I}} \lesssim m_a^{\text{Inf}} \sim \begin{cases} \frac{\sqrt{(\Lambda_{\text{QCD}}^{\text{Inf}})^3 m_u^{\text{Inf}}}}{f_a^{\text{Inf}}} & : m_u^{\text{Inf}} \lesssim \Lambda_{\text{QCD}}^{\text{Inf}}, \\ \frac{(\Lambda_{\text{QCD}}^{\text{Inf}})^2}{f_a^{\text{Inf}}} & : m_u^{\text{Inf}} > \Lambda_{\text{QCD}}^{\text{Inf}}. \end{cases} \quad (4.65)$$

Here we use different expressions for the axion mass depending on whether there are light quarks or not. In the case when there are no light quarks, as described by the lower

equation, we used the axion mass derived from the dilute instanton gas approximation. Alternatively, by plugging (4.65) into (4.64), we can derive a condition on $\Lambda_{\text{QCD}}^{\text{Inf}}$,

$$\Lambda_{\text{QCD}}^{\text{Inf}} \gtrsim \begin{cases} 10^{15} \left(\frac{f_a}{10^{16} \text{ GeV}} \right)^{16/9} \left(\frac{M_{\text{P}}}{v_H^{\text{Inf}}} \right)^{1/3} \left(\frac{1}{\epsilon_{\text{eff}}} \right)^{4/3} \text{ GeV} & , \\ 10^{14} \left(\frac{f_a}{10^{16} \text{ GeV}} \right)^{4/3} \left(\frac{1}{\epsilon_{\text{eff}}} \right) \text{ GeV} & . \end{cases} \quad (4.66)$$

If a model with some $\Lambda_{\text{QCD}}^{\text{Inf}}$ satisfies this condition, then the initial misalignment angle will be dominated by thermal fluctuations, giving rise to the viable region depicted in Fig. 9.

In the KSVZ and DFSZ models, where $\Lambda_{\text{QCD}}^{\text{Inf}}(v_H^{\text{Inf}})$ is given by (4.50), it is not possible to satisfy condition (4.66). Without additional field content that modifies the running of the strong coupling, $\Lambda_{\text{QCD}}^{\text{Inf}}$ is simply not high enough. Thus, in these two models such an efficient relaxation must be avoided. For instance, this can be achieved by requiring $\theta_1 \gtrsim \sigma_\theta$. Using (4.43), this translates into

$$\mathcal{N}_e \lesssim \frac{3H_{\text{I}}^2}{(m_a^{\text{Inf}})^2} \ln \left(\frac{2\pi f_a}{H_{\text{I}}} \right) . \quad (4.67)$$

For $H_{\text{I}} \sim \Lambda_{\text{QCD}}^{\text{Inf}} \sim 10^5 \text{ GeV}$ and $f_a \sim 10^{16} \text{ GeV}$, this bound gives $\mathcal{N}_e \lesssim 10^{22}$.

4.4 Summary and Discussion

In this chapter, we showed that there is a viable part of the parameter space where the early relaxation mechanism can successfully be incorporated into the KSVZ and DFSZ axion models. For clarity, we first recapitulated the early relaxation mechanism proposed in [44] which can be summarized as follows. When QCD becomes strong during inflation, i.e. when $H_{\text{I}} \lesssim \Lambda_{\text{QCD}}^{\text{Inf}}$, the axion potential is activated prior to the time when the ordinary misalignment mechanism occurs. In this way it can relax to the minimum so that, when the ordinary QCD phase transition takes place, the axion is located close to its minimum. Hence, the initial misalignment angle naturally takes small values.

As we pointed out, a Higgs portal between the inflaton and our Higgs doublet or, alternatively, higher order operators of those fields, can result in much larger quark masses during inflation than in today's universe. This changes the running of the strong coupling constant and thus can result in a larger value of the QCD scale. By using $\alpha_s^{\text{Inf}}(M_{\text{GUT}}) = \alpha_s(M_{\text{GUT}})$ as initial condition for the running as dictated by the minimal field content of the KSVZ and DFSZ models, we found that $\Lambda_{\text{QCD}}^{\text{Inf}}$ can be as large as 10^5 GeV . But using additional field content such that $\alpha_s^{\text{Inf}}(M_Z^{\text{Inf}}) = \alpha_s(M_Z)$ would be the initial condition, we found that $\Lambda_{\text{QCD}}^{\text{Inf}}$ can be as large as 10^{15} GeV .

Furthermore, we showed that the displacement of the axion minimum during inflation is completely negligible for the KSVZ and DFSZ models. Since we used the same non-perturbative QCD effects as in the late universe to make QCD strong and to generate the

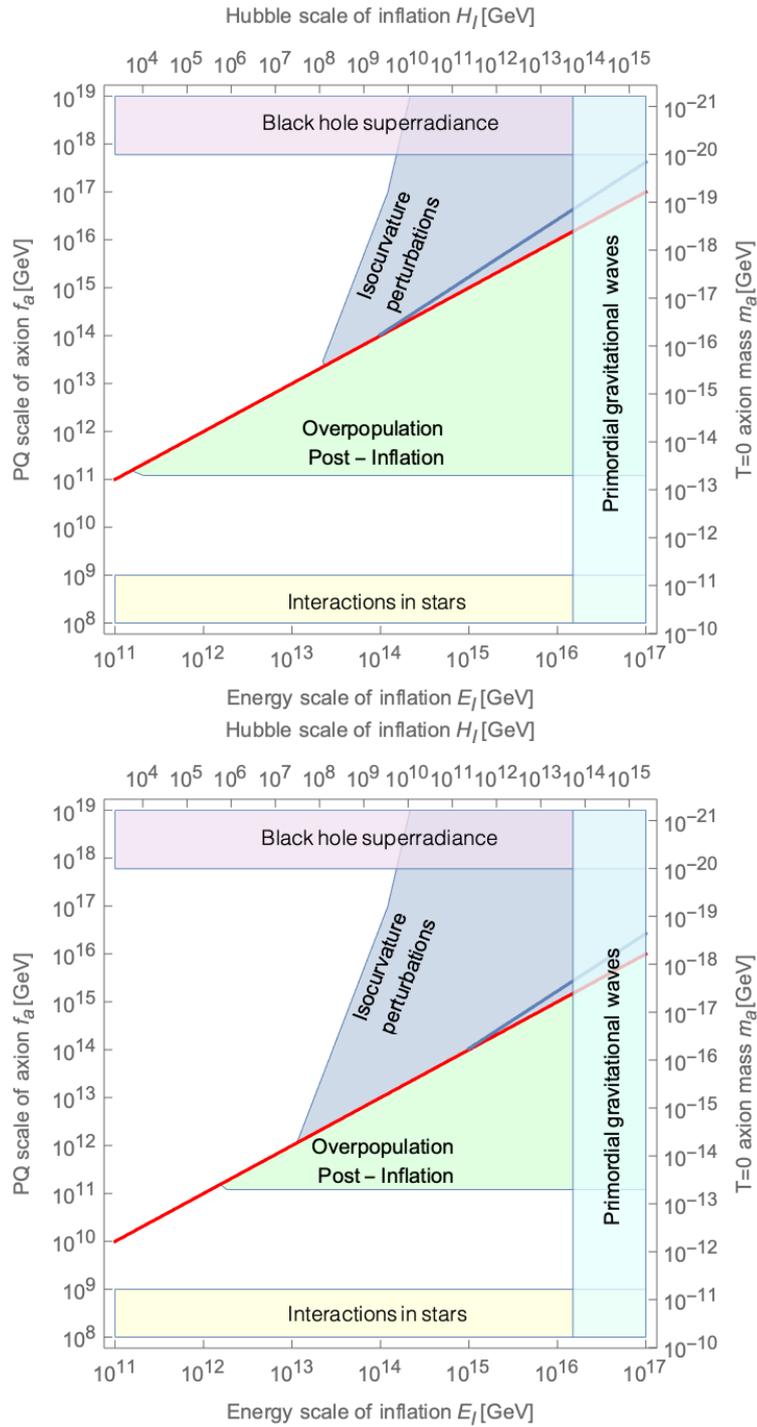


Figure 9: Top: Constraints on the axion for very efficient reheating, $\epsilon_{\text{eff}} \sim 1$. The blue line depicts the upper bound for our minimal scenarios. Hence, the region enclosed by the red and the blue line defines the parameter space where the axion dynamically relaxes to the minimum during inflation and later makes up all the dark matter. Bottom: Same as top but with $\epsilon_{\text{eff}} \sim 10^{-1}$.

axion potential, the relaxation is guaranteed to suppress the initial misalignment angle. The resulting suppression depends on the duration of the early phase of strong QCD.

In particular, we pointed out that when $m_a^{\text{Inf}} \gtrsim H_I$ or inflation lasts long enough, the axion is diluted to an extent where the initial misalignment angle is dominated by inflationary quantum fluctuations and post-inflationary thermal fluctuations. The former cannot result in the axion making up all the dark matter due to the bound on H_I induced by inflationary tensor perturbations. Hence, in this interesting case only strong thermal fluctuations can result in the axion making up all the dark matter for values of $f_a \gtrsim 10^{12}$ GeV. The required amount of thermal fluctuations is controlled by the reheating temperature, which can be bounded by above by the maximum thermalization temperature defined in (4.11). For very efficient reheating, i.e. $10^{-1} \lesssim \epsilon_{\text{eff}} \lesssim 1$, we found that this results in a relatively narrow region of applicability, while for less efficient reheating there is no viable parameter space. In particular we discovered that the point with $f_a \sim 10^{16}$ GeV and $H_I \sim 10^{13}$ GeV lies in that allowed region, which is very interesting due to the PQ scale coinciding with the GUT scale and the inflationary Hubble parameter being very close to the experimental constraint.

Since this narrow region of applicability is furthermore constrained by isocurvature perturbations, these must be avoided in addition. For this purpose, we required the axion mass m_a^{Inf} to be larger than the inflationary Hubble parameter H_I , resulting in a constraint on $\Lambda_{\text{QCD}}^{\text{Inf}}$. Unfortunately, this constraint is not satisfied for the KSVZ and DFSZ models where the modification of the QCD scale comes only from the displacement of the Higgs VEV. Without a modification of the running coupling from additional fields or a direct enhancement via a moduli field, only the region with $H_I \lesssim 10^5$ GeV becomes viable in the minimal models.

CHAPTER FIVE

AXIONS AND HIDDEN YM GROUPS

In Sec. 2.4, we explained how the strong CP problem of QCD is promoted to a consistency problem due to the S-matrix exclusion of de Sitter. From this perspective of quantum gravitational consistency, the strong CP problem is not limited to QCD but extends to all non-Abelian YM groups. Each YM group contains a vacuum angle that leads to unacceptable de Sitter-type vacua, necessitating its removal [45]. To achieve this using the PQ mechanism, one axion is required per YM group because a single axion cannot set multiple θ -parameters to zero, as pointed out in Sec. 2.3.4. In the projects presented in this chapter, we consider theories in which dark matter includes one or more YM sectors, add the necessary axions, and study phenomenological consequences. To our knowledge, axions have not been taken into account in these theories because there was no necessity for them from the understanding of the strong CP problem as a small value puzzle. However, the situation is essentially similar to the string theory inspired axiverse [108], even though the axions in our framework arise from a motivation that is completely different. The phenomenology of several axions in the context of the axiverse was for instance studied in [139].

To be precise, we consider theories containing N sectors labeled by $i = 1, \dots, N$, where each sector contains at least one Yang-Mills subgroup. For convenience, we choose our sector to be labeled by $i = 1$. All the other sectors collectively form the dark matter and we will refer to them as hidden or dark sectors. The sectors interact via gravity and additional non-gravitational interactions, which we encode in \mathcal{L}_{mix} . Thus, the total Lagrangian can be written as

$$\mathcal{L} = \sum_{i=1}^N \mathcal{L}_i + \mathcal{L}_{\text{mix}} . \quad (5.1)$$

The range of potential models is remarkably wide, including mirror dark matter [140, 141], the many-species solution to the hierarchy problem [48, 49], and string theory

inspired models [142–145]. All these models predict the existence of various particles and fundamental forces in the hidden sectors, which interact very weakly with our sector’s particles. This makes the particles from the dark sectors effectively invisible to most of our current detection methods. Nonetheless, they could still account for a significant portion of the dark matter in the universe.

To provide concrete examples, we will examine two simple models: a pure YM sector and N exact SM copies. The former has been utilized to model the dark matter due to the interesting dynamics of YM theories. For instance, these dynamics can result in a potential phase transition that generates stochastic gravitational waves during the early universe [146, 147]. The latter is a realization of the many-species solution to the hierarchy problem, which yields many fascinating phenomena when applied to model the dark matter [148], such as neutron and neutrino oscillations [149, 150], modified black hole physics [151], and compact dark matter objects [152]. Although we focus on these two models for our calculations, our discussion aims to be as general as possible and highlight the model-independent predictions.

By adding to the dark YM sectors the necessary axions, we find several phenomenological consequences such as constraints on the parameter space of the dark sectors, potential new small structures, and new experimental signatures that we briefly summarize here. For a pure YM sector the misalignment mechanism and the stability requirement of the axions result in a constraint on the dark confinement scale. For N exact SM copies the misalignment mechanism in combination with the species bound places an upper bound on N . In the pre-inflationary scenario for multiple axions, the additive influence of isocurvature fluctuations does not necessarily tighten the single axion bound on the inflationary Hubble scale and we identify regions in the parameter space where the isocurvature bound is avoided. The collective phenomenon leading to N -MACHOs does not take place for axions due to their attractive self-interaction. Thus, there are no collective axion stars with a $1/\sqrt{N}$ suppressed mass spectrum. Lastly, intersector communication through axion kinetic mixing leads to the existence of two distinct axion states and an additional constraint from the stability requirement of the axions.

This chapter is organized as follows. To begin with, in Sec. 5.1 we define the cosmological framework of the N dark sectors. We then discuss how to implement the PQ mechanism in each sector and show the bounds arising from the misalignment production and isocurvature perturbations of N axions. In Sec. 5.2 we discuss collective phenomena that result in compact dark matter objects in theories with N dark sectors and study if this can lead to interesting collective bound states of multiple axions. In Sec. 5.3 we study the consequences of kinetic mixing between N axions and point out the differences for different types of experiments. Finally, in Sec. 5.4 we summarize our results.

5.1 Cosmological Implications

The axion provides an excellent dark matter candidate due to its Goldstone nature and is subject to several cosmological constraints, as discussed in Sec. 4.2. In this section, we

ask the question how these constraints change in the presence of several axions.

5.1.1 Cosmological Framework

Let us start by clarifying the cosmological behavior of N sectors. We neglect intersector interactions except gravity for now by setting $\mathcal{L}_{\text{mix}} = 0$ and discuss them separately in Sec. 5.3.

In order for these models to be viable in a cosmological context, the dark sectors must exhibit characteristics of a cold and pressureless fluid, forming stable dark matter halos. Additionally, if the dark sectors involve any massless fields, they must avoid the bound on the number of massless particles during nucleosynthesis. These requirements can appear challenging to achieve, particularly when non-trivial interactions are present within the dark sectors. We will use the approach proposed in [148, 153] to address these requirements.

In that approach, inflation is used to populate and reheat the sectors in different ways. The key aspect is that the reheating process in the dark sectors is significantly less efficient compared to the visible sector. This leads to a low reheating temperature in the dark sectors, which is then used to avoid the nucleosynthesis bound and prevent cooling mechanisms from collapsing dark matter halos. The behavior throughout cosmological history is in general altered, despite potential similarities to the visible sector. For large N , each sector is so sparsely populated that the pressureless fluid-like behavior arises from kinematics, meaning that particles from the same sector rarely interact with each other. The whole situation is unchanged even when starting with equal sectors. The cosmological evolution breaks this symmetry in a sense by resulting in different energy densities and temperatures. The field theoretic parameters of each sector, however, remain the same.

Let us now add an axion to each YM sector via the PQ mechanism as described in Chapter 3. The implementation of the $U(1)_{\text{PQ}}$ in the dark sectors depends on the model, but there is more freedom compared to our SM sector due to a lack of bounds. For exact SM copies, the KSVZ or DFSZ models or the two-form implementation can be used. For SM-like sectors with the same field content and group structure but different parameters, even the original PQWW model can be used. In pure YM sectors, the KSVZ scenario, 2-form implementation, or massless quark solution can be implemented. The crucial point here is that the cosmological evolution is essentially identical for all options, only the values of f_i^a differ. Hence, we will remain independent of any concrete realization of the PQ mechanism and treat f_i^a as free parameters. However, we will discuss consequences specific to certain realizations if necessary.

Before going to the misalignment mechanism, let us discuss the range of N . The lowest value of interest is $N = 2$, which in the case of a SM copies corresponds to the ordinary mirror dark matter [140, 141]. This value is not viable if the dark SM copy is supposed to make up all the dark matter since in this case the dark matter halos would have collapsed. By increasing the number of SM copies, the particle density in

each sector is reduced. This leads to a suppressed dissipation efficiency and the lowest phenomenologically viable value of N being of $\mathcal{O}(10)$ [149, 154]. An upper bound on N arises from the fact that in the presence of N_s particle species, the gravitational cutoff is not given by the Planck scale M_{P} but by

$$M_* \sim \frac{M_{\text{P}}}{\sqrt{N_s}} . \quad (5.2)$$

In the case of SM copies, since no strong gravity effects have been measured up to the energy scale of roughly 1 TeV, this results in the upper bound $N \lesssim 10^{32}$. We will keep N as a free parameter throughout this work and focus on $N = 2$ and large values of N .

5.1.2 Misalignment: Exact SM Copies

The misalignment mechanism, which is required not to produce more dark matter than observed, imposes a very strict constraint on the axion scale in the case of a single axion [109–111]. In this section, we extend this constraint to multiple axions and derive limits on the parameter space of the dark sectors. For a clearer presentation, let us begin with the case of N exact SM copies, where the equality between the sectors can be ensured by a discrete symmetry [149, 155]. For other compositions of the dark sectors, we then highlight the differences in the calculation and provide the final outcome.

In the standard axion scenario, the $U(1)_{\text{PQ}}$ symmetry is spontaneously broken at $T^{\text{PQ}} \sim f_a$, whereas in our framework, this occurs in each sector at temperatures $T_i^{\text{PQ}} \sim f_a$. The axions a_i emerge as the corresponding Goldstone bosons. Let us again work with the dimensionless field $\theta_i(x) \equiv a_i(x)/f_a$ in the following. Each axion receives a potential from its corresponding QCD, which we again model by the dilute instanton gas coupled to the cosmological thermal bath [118],

$$V_i(\theta_i) = m_a^2(T_i) f_a^2 (1 - \cos(\theta_i)) , \quad (5.3)$$

where the Temperature-dependent masses are given by

$$m_a(T_i) \equiv \frac{(\Lambda_{\text{QCD}}^3 m_u)^{\frac{1}{2}}}{f_a} \begin{cases} \beta \left(\frac{\Lambda_{\text{QCD}}}{T_i} \right)^4 & : T_i > \Lambda_{\text{QCD}} , \\ 1 & : T_i \lesssim \Lambda_{\text{QCD}} . \end{cases} \quad (5.4)$$

For the SM, the parameter β encodes QCD and active quark physics at the temperature T and is roughly 10^{-2} [79].

In an FLRW background with $R(t)$ denoting the scale factor and $H(t)$ the Hubble parameter, each axion satisfies the equation of motion

$$\ddot{\theta}_i + 3H(t)\dot{\theta}_i - \frac{1}{R^2(t)}\Delta\theta_i + \frac{V'_i(\theta_i)}{f_a^2} = 0 . \quad (5.5)$$

Let us as in the single axion case make the following two simplifications. First, only the leading order in θ_i of the potential is taken into account. Secondly, only the zero mode

of θ_i is considered. With these simplifications, each equation of motion reduces to that of a damped harmonic oscillator,

$$\ddot{\theta}_i + 3H(t)\dot{\theta}_i + m_a^2(T_i(t))\theta_i = 0 . \quad (5.6)$$

At $T_i \gg \Lambda_{\text{QCD}}$, the axions are basically frozen out due to the flat potential. At $T_i \sim \Lambda_{\text{QCD}}$ the corresponding axion potential gets significant but it is not until the Hubble friction is overcome before the axion starts performing coherent oscillations around the vacuum. This moment is defined by

$$m_a(T_i(t_{\text{osc}})) = 3H(t_{\text{osc}}) , \quad (5.7)$$

for each axion. From that moment on, its equation of state does no longer correspond to that of dark energy but to that of non-relativistic matter, making the corresponding axion contribute into the dark matter.

For the model under consideration, this implies that the axions in the dark sectors are essentially created with their zero-temperature potential due to the very low temperature in the dark sectors. In contrast, our axion is created with an essentially flat potential due to the temperature suppression in (5.4). Since the zero-temperature mass is larger than the high-temperature mass, the axion from our sector will hence start oscillating later than the other axions. In order to quantify this delay, let us use our sectors temperature $T_1 \equiv T$ as a clock instead of the cosmic time t . In terms of T , the Hubble parameter during radiation domination is given by

$$H(T) = \sqrt{\frac{\rho_{\text{tot}}}{3M_{\text{P}}^2}} \sim \frac{T^2}{M_{\text{P}}} , \quad (5.8)$$

where ρ_{tot} denotes the total energy density, which is dominated by our sector. Using (5.8) and the axion masses defined by (5.4), the condition (5.7) results in the oscillations commencing when

$$T_{\text{osc},1} \sim \left[\frac{\beta M_{\text{P}} \Lambda_{\text{QCD}}^{\frac{11}{2}} m_u^{\frac{1}{2}}}{3f_a} \right]^{\frac{1}{6}} \sim 4 \times 10^{-1} \left(\frac{10^{12} \text{ GeV}}{f_a} \right)^{\frac{1}{6}} \text{ GeV} , \quad (5.9)$$

in our sector and

$$T_{\text{osc},i} \sim \left[\frac{M_{\text{P}} \Lambda_{\text{QCD}}^{\frac{3}{2}} m_u^{\frac{1}{2}}}{3f_a} \right]^{\frac{1}{2}} \sim 2 \times 10^1 \left(\frac{10^{12} \text{ GeV}}{f_a} \right)^{\frac{1}{2}} \text{ GeV} , \quad (5.10)$$

in the dark sectors with $i \neq 1$. It should be kept in mind that both moments in time are expressed in terms of our sectors temperature. While we left f_a as a free parameter there is a small caveat for $f_a \gtrsim 6 \times 10^{17} \text{ GeV}$. Above of this value, our axion has also reached its zero-temperature mass before overcoming the Hubble friction [156]. Consequently, all

axions start oscillating at the temperature dictated by (5.10).

The initial energy density of each oscillation is

$$\rho_{a_i}(T_{\text{osc},i}) = \frac{1}{2} f_a^2 m_{a_i}^2(T_{\text{osc},i}) \theta_i^2(T_{\text{osc},i}), \quad (5.11)$$

where $\theta_i(T_{\text{osc},i}) \equiv \theta_i^{\text{ini}}$ is each sectors initial misalignment angle.

The value of θ_i^{ini} is determined by whether the pre- or post-inflationary scenario takes places. To remind the reader, in the post-inflationary all possible values of θ_i^{ini} appear in the Hubble patch of the universe, and averaging over a uniform distribution leads to $\theta_i^{\text{ini}} = \pi/\sqrt{3}$ for all i . In the pre-inflationary scenario, the initial misalignment angle can be an arbitrary value, although a reasonable choice in the absence of any explanation for special initial conditions is $\theta_i^{\text{ini}} \sim \mathcal{O}(1)$. Moreover, for a sufficiently large number of dark sectors, assuming a uniform distribution among them leads to an average value of $\theta_i^{\text{ini}} \sim \mathcal{O}(1)$. Let us adopt $\theta_i^{\text{ini}} \sim \mathcal{O}(1)$ for the time being and revisit this issue later, taking into account the discussion of Chapter 4.

During the period of radiation domination when Big Bang nucleosynthesis occurred, it is crucial that the energy density of the dark sector axions does not dominate the energy density of the universe. The energy density contributed by the relativistic degrees of freedom at $T_{\text{osc},i}$ is approximated by $\rho_{\text{rad}}(T_{\text{osc},i}) \sim g_*(T_{\text{osc},i})(T_{\text{osc},i})^4$, where the effective number of relativistic species is $g(T_{\text{osc},i}) \sim 10^2$. To satisfy the condition $\sum_{i=2}^N \rho_{a_i}(T_{\text{osc},i}) \ll \rho_{\text{rad}}(T_{\text{osc},i})$, we find the constraint

$$N \lesssim \frac{g_*(T_{\text{osc},i})(T_{\text{osc},i})^4}{\Lambda_{\text{QCD}}^3 m_u(\theta_i^{\text{ini}})^2} \sim 10^{12} \left(\frac{10^{12} \text{ GeV}}{f_a} \right)^2 \left(\frac{1}{\theta_i^{\text{ini}}} \right)^2. \quad (5.12)$$

As it turns out, a more stringent constraint on the value of N can be derived at present temperature. Assuming the changes in the mass to be adiabatic, the number of axion zero modes per co-moving volume is conserved. Hence, the current energy density for each sector is

$$\rho_{a_i}(T_{\text{today}}) = \rho_{a_i}(T_{\text{osc},i}) \frac{m_{a_i}(T_{\text{today}})}{m_{a_i}(T_{\text{osc},i})} \left(\frac{T_{\text{today}}}{T_{\text{osc},i}} \right)^3 \quad (5.13)$$

We normalize with respect to the critical energy density $\rho_{\text{cr}} \sim M_{\text{P}}^2 H_{\text{today}}^2$ to receive the corresponding axion fraction Ω_{a_i} in the present universe. By comparing Ω_{a_i} with the observed dark matter fraction Ω_{DM} from the latest Planck mission [40], we find for our sector

$$\frac{\Omega_{a_1}}{\Omega_{\text{DM}}} \sim 0.54 \left(\frac{\beta^{-\frac{1}{6}} m_u^{\frac{5}{12}} \Lambda_{\text{QCD}}^{\frac{7}{12}}}{10^{-2} \text{ GeV}} \right) \left(\frac{f_a}{10^{12} \text{ GeV}} \right)^{\frac{7}{6}} \left(\frac{\theta_1^{\text{ini}}}{1} \right)^2, \quad (5.14)$$

while for each dark sector we get

$$\frac{\Omega_{a_i}}{\Omega_{\text{DM}}} \sim 0.01 \left(\frac{m_{u_i}^{\frac{1}{4}} \Lambda_{\text{QCD}}^{\frac{3}{4}}}{10^{-2} \text{ GeV}} \right) \left(\frac{f_a}{10^{12} \text{ GeV}} \right)^{\frac{3}{2}} \left(\frac{\theta_i^{\text{ini}}}{1} \right)^2 \quad : i \neq 1. \quad (5.15)$$

To summarize, differences in reheating temperatures cause axion oscillations to commence at different times, leading to distinct axion densities even with identical field theoretic parameters. Although a single mirror sector has a minimal cosmological impact owing to its lower axion density relative to our axion, the presence of numerous mirror sectors leads to the accumulation of densities, and their combined impact cannot be ignored. In the following discussion, we will examine the parameter space and investigate the consequences of this collective effect.

In the single axion case, the scale f_a is restricted to lie in the classic axion window, $10^9 \text{ GeV} \lesssim f_a \lesssim 10^{12} \text{ GeV}$ (see Sec. 4.2). The lower bound on f_a is determined by astrophysical considerations, where a non-trivial interaction with matter contributes to the cooling of stars. Meanwhile, the upper bound on f_a is determined by (5.14) and the requirement that our axion does not exceed the observed density of dark matter. These limitations remain unchanged even in the presence of axions from dark sectors. However, the total number of axions must not surpass the observed dark matter density, i.e.

$$\sum_{i=1}^N \frac{\Omega_{a_i}}{\Omega_{\text{DM}}} \lesssim 1. \quad (5.16)$$

The result in the following inequality for the viable parameter space,

$$N \lesssim 10^2 \left(\frac{10^{12} \text{ GeV}}{f_a} \right)^{3/2} \left(\frac{1}{\theta_i^{\text{ini}}} \right)^2 \left[1 - 0.54 \left(\frac{f_a}{10^{12} \text{ GeV}} \right)^{\frac{7}{6}} \left(\frac{\theta_1^{\text{ini}}}{1} \right)^2 \right]. \quad (5.17)$$

Fig. 10 depicts the viable regions in the N - f_a plane for various initial values of θ_i^{ini} . The bold lines correspond to the equality in (5.16) and indicate the parameter combinations where the axions constitute all of the dark matter. For instance, following the blue line on the graph, our axion comprises the entirety of the dark matter when $f_a \sim 10^{12} \text{ GeV}$, whereas for $f_a \sim 10^9 \text{ GeV}$, approximately $N \sim 10^6$ axions from the dark sectors make up the dark matter. A similar behavior is observed for the other depicted values of θ_i^{ini} . We can quantify this behavior by calculating the fraction of our axion along these lines. The result is shown in Fig. 11. There, we also observe that for a fixed θ_i^{ini} , lower f_a values result in a dominant contribution from dark sector axions while larger f_a values lead to a dominant contribution from our axion.

Let us now discuss the range of the initial misalignment angles, taking into account our results from Chapter 4. While $\theta_i^{\text{ini}} \sim \mathcal{O}(1)$ most certainly applies when the PQ symmetry is broken after inflation, it does not necessarily when the PQ symmetry is broken during inflation. In this scenario, the initial misalignment angle is an initial condition, and thus can take any value. While $\theta_i^{\text{ini}} \sim \mathcal{O}(1)$ appears to be the most

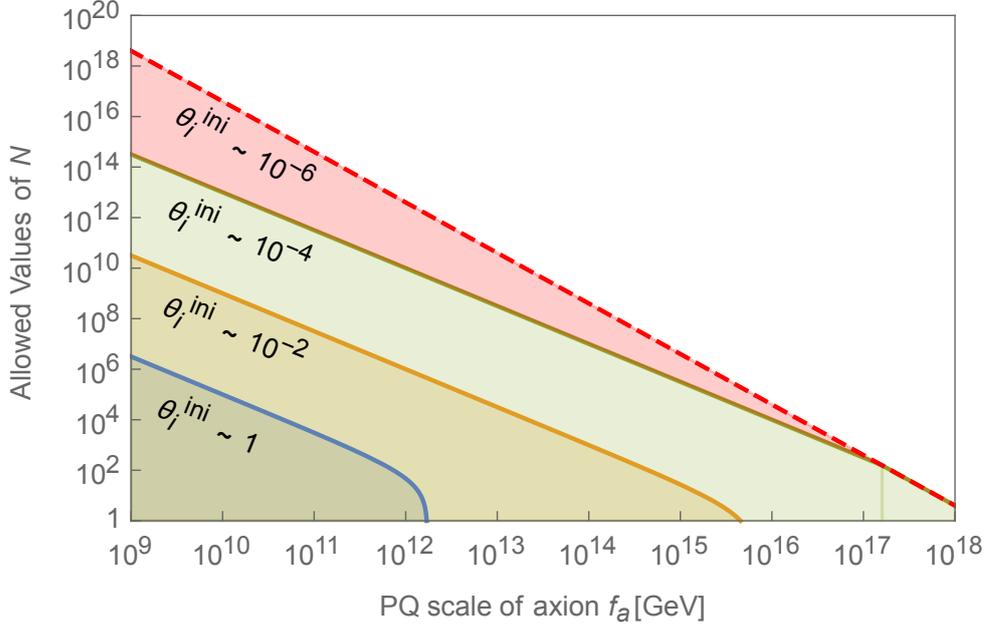


Figure 10: The allowed parameter space with different initial misalignment angles θ_i^{ini} according to (5.17). The values on the thick lines correspond to the case where the dark matter is entirely composed of axions. The dashed line presents the species bound, meaning that along this line the gravitational cutoff M_* and f_a coincide.

plausible, a strong QCD phase during inflation can dynamically result in $\theta_i^{\text{ini}} \ll 1$ [44]. Although such a phase may appear exotic, it arises naturally for small inflationary Hubble scales (refer to [2, 157] for specific implementations). If such a phase or a similar mechanism is present in each sector, the parameter space can extend to $f_a \gg 10^{12}$ GeV and $N \gg 10^6$. However, note that for a given N , the PQ scale is limited from above by (5.2), which we display as a dashed line in Fig. 10. Therefore, for sufficiently small values of θ_i^{ini} within a given f_a -window, axions cannot compose all of the dark matter.

If such a phase existed in each sector, it would allow for $N > 10^6$, but this raises the question of whether this is sufficient to address the Hierarchy problem. As discussed in the previous section, $N \sim 10^{32}$ copies of the SM would provide a solution to the Hierarchy problem. However, this requires $f_a \sim \text{TeV}$ and an early phase of strong QCD, which is in conflict with the lower bound on the PQ scale from astrophysical considerations and direct detection in our sector ($f_a \gtrsim 10^9$ GeV). Thus, a large number of exact SM copies as a solution to the Hierarchy problem seems to be excluded. However, proposals such as the clockwork mechanism [158] allow for the separation of the PQ scale from the suppression factor of the couplings, which may render the part of the parameter space that solves the Hierarchy problem viable.

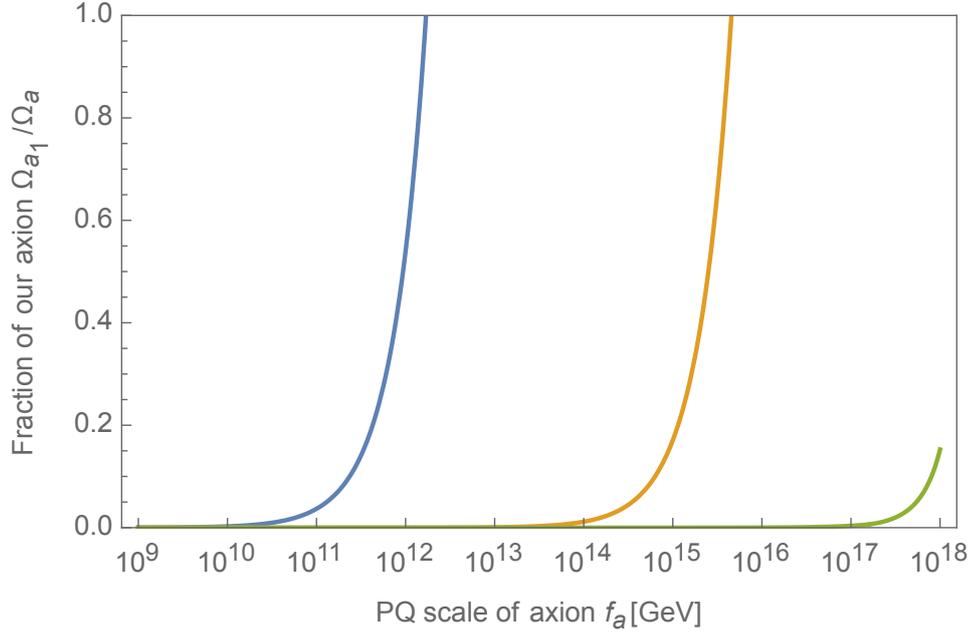


Figure 11: The fraction of our axion energy density with respect to the total axion energy density in the case when the axions make up all the dark matter. For larger f_a our axion is dominating, while for lower f_a the axions from dark sectors dominate.

5.1.3 Misalignment: SM-like Dark Sectors

Another way that does not rely on small values of $\theta_i(T_{\text{osc}})$ is given by relaxing the assumption of N exact copies of the SM. Although specific predictions require a defined profile in the discrete space of sectors, there will be some general consequences that are worth mentioning.

To begin with, the PQ symmetry could be broken at different temperatures $T_i^{\text{PQ}} \sim v_i$. Therefore, some axions could appear during inflation while others would appear after inflation. This would result in the constraints from each scenario to only apply to the corresponding axions, e.g. only the axions that appeared after inflation will be subject to the domain wall problem while those appearing during inflation would lead to isocurvature perturbations. Moreover, different values of v_i change the axions mass, which is also changed by different values of Λ_i^{QCD} and m_{u_i} . This would have the crucial consequence that some axions never start to oscillate, resulting in the bound from misalignment only holding for the sectors with oscillating axions. In other words, if there were N_{mis} sectors with parameters roughly of the order of our SM parameters and $N - N_{\text{mis}}$ sectors with any parameters such that oscillation does not take place until today, the bounds in (5.17) would only hold for N_{mis} instead of N . Hence, N could be large enough to solve the Hierarchy problem.

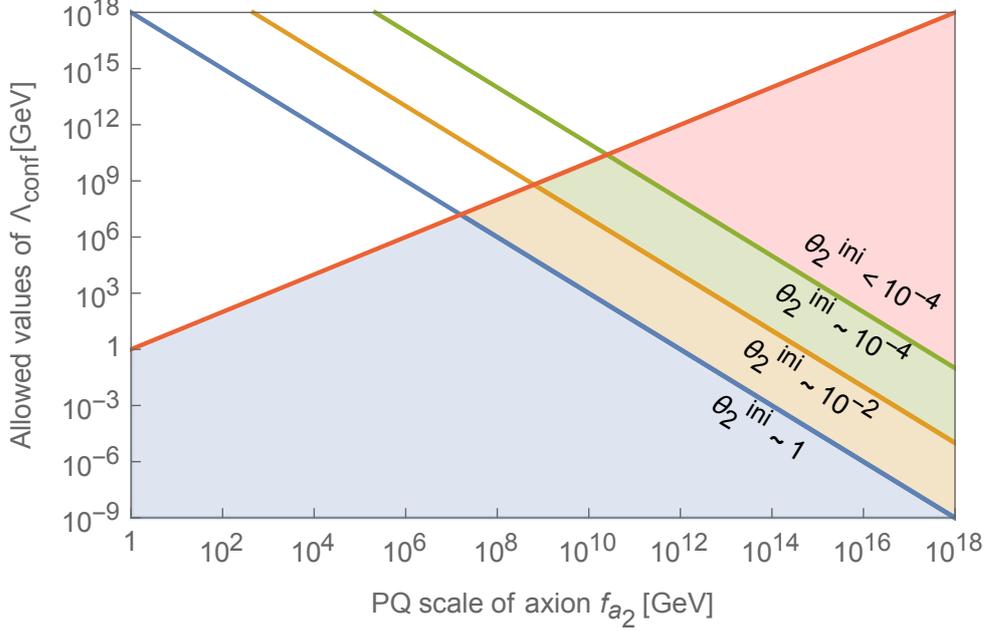


Figure 12: The allowed region for a single dark, pure YM sector with confinement scale Λ_{conf} , axion scale f_{a_2} , and no intersector interactions. The red line represents the perturbative unitarity bound $\Lambda_{\text{conf}} \lesssim f_{a_2}$. The region with $\theta_2^{\text{ini}} \sim 1$ is favored by minimality, since in this scenario there is no need for additional physics that results in a small misalignment angle.

5.1.4 Misalignment: Pure YM Dark Sector

Finally, let us turn to the case where each dark sectors is based on a pure YM group $SU_i(N_C)$. The difference compared to the case with exact SM copies would be the absence of light quarks. Without light quarks, chiral perturbation techniques can no longer be used to calculate the zero-temperature mass of the axion. Instead, the mass would have to be calculated by using alternatives such as large N_C methods. As before, we use as an approximation the extrapolated result of the dilute instanton gas at finite temperatures [118],

$$m_{a_i}(T^i) \equiv \frac{(\Lambda_{\text{conf}}^i)^2}{f_{a_i}} \begin{cases} \left(\frac{\Lambda_{\text{conf}}^i}{T^i}\right)^4 & : T^i > \Lambda_{\text{conf}}^i, \\ 1 & : T^i \lesssim \Lambda_{\text{conf}}^i. \end{cases} \quad (5.18)$$

Here, the confinement scales of the dark sectors are denoted by $\Lambda_{\text{conf},i}$ and the factor of β is absent because of the absence of light quarks. In contrast to the case of exact SM copies, the $\Lambda_{\text{conf},i}$ can be smaller than the dark sector temperatures. Hence, the dark sector axion can in principle be created with an essentially flat potential, although this requires quite low confining temperatures. We will only consider the scenario where the dark axions are produced with their zero-temperature mass.

From there, the calculation is vastly the same as in the exact SM case. The oscillations

in the dark sectors commence at

$$T_{\text{osc},i} \sim \Lambda_{\text{conf}}^i \left(\frac{M_{\text{P}}}{3f_{a_i}} \right)^{\frac{1}{2}},$$

which results in the dark sector density of

$$\frac{\Omega_{a_i}}{\Omega_{\text{DM}}} \sim 10^{-18} \left(\frac{\Lambda_{\text{conf}}^i}{\text{GeV}} \right) \left(\frac{\theta_i^{\text{ini}}}{1} \right)^2 \left(\frac{f_{a_i}}{\text{GeV}} \right)^{\frac{3}{2}}. \quad (5.19)$$

Let us discuss the parameter space for the particularly interesting case of a single, pure YM sector. For the dark sector axion to not result in an abundance of dark matter, we find

$$\Lambda_{\text{conf}}^i \lesssim 10^{18} \text{ GeV} \left(\frac{1}{\theta_i^{\text{ini}}} \right)^2 \left(\frac{\text{GeV}}{f_{a_i}} \right)^{\frac{3}{2}}, \quad (5.20)$$

where the equality is valid when the dark axion makes up all the dark matter. The viable region is shown in Fig. 12. An interesting difference arises between the invisible axion and the two-form realizations of the PQ mechanism. In the case of the KSVZ axion, the decay constant f_{a_i} is essentially unconstrained. Since there is no requirement for additional physics to yield a small misalignment angle in this dark sector, we can assume $\theta_i^{\text{ini}} \sim \mathcal{O}(1)$. Therefore, in order for the dark confinement scale to be below the Planck scale, the axion scale must satisfy $f_{a_i} \gtrsim 1 \text{ GeV}$. For the two-form realization, the value of $f_{a_i} \sim M_{\text{P}}$ is strongly favored, which leads to $\Lambda_{\text{conf}}^i \sim 1 \text{ eV}$.

5.1.5 Isocurvature Perturbations

When one of the PQ symmetries is broken during inflation and never restored afterwards, the corresponding axion field is subject to quantum fluctuations. Let us write the axion fields as $\theta_i = \langle \theta_i \rangle + \delta\theta_i$, where $\langle \theta_i \rangle = \theta_i^{\text{ini}}$ and $\delta\theta_i$ denotes the quantum fluctuations. These fulfill $\langle \delta\theta_i \rangle = 0$ and have a standard deviation of

$$\sigma_{\theta_i} \sim \sqrt{\langle \delta\theta_i^2 \rangle} \sim \frac{H_{\text{I}}}{2\pi f_{a_i}}. \quad (5.21)$$

Moreover, if none of the axions drives inflation, their fluctuations will not be of adiabatic but of isocurvature-type. Since these lead to a unique imprint in the temperature and polarisation fluctuations of the CMB, they give rise to a constraint on the axion's parameter space.

One could worry that in the presence of several axions, the amount of isocurvature perturbations is significantly enhanced due to the contributions being additive. This would tighten the existing constraint, making it much more severe. Let us show that this is not necessarily the case.

Assuming the fluctuations to be normal distributed in the regime of small θ_i , where anharmonic corrections of the potentials are negligible, the collective amplitude of the

axions isocurvature fluctuations is given by [131]

$$\begin{aligned}\Delta_a(k_0) &= \frac{\delta\Omega_{\text{DM}}}{\Omega_{\text{DM}}} = \frac{\sum_i \Omega_{a_i}}{\Omega_{\text{DM}}} \frac{\delta \ln \Omega_{a_i}}{\delta \theta_i^{\text{ini}}} \sigma_{\theta_i} \\ &= \sum_i \frac{\Omega_{a_i}}{\Omega_{\text{DM}}} \frac{H_{\text{I}}}{\pi \theta_i^{\text{ini}} f_{a_i}},\end{aligned}\quad (5.22)$$

where $\theta_i^{\text{ini}} \ll \sigma_{\theta_i}$ was used. The latest experimental bound on uncorrelated isocurvature perturbations by Planck is [40]

$$\beta(k_0) \equiv \frac{\Delta_a^2(k_0)}{\Delta_{\mathcal{R}}^2(k_0) + \Delta_a^2(k_0)} < 0.038 \quad \text{at 95\% CL}, \quad (5.23)$$

where $k_0 = 0.050 \text{Mpc}^{-1}$. This translates to a constrain on H_{I} ,

$$H_{\text{I}} \lesssim 10^7 \text{ GeV} \sum_i \frac{\Omega_{\text{DM}}}{\Omega_{a_i}} \left(\frac{f_{a_i}}{10^{12} \text{ GeV}} \right) \left(\frac{\theta_i^{\text{ini}}}{1} \right). \quad (5.24)$$

For a single dark YM sector and N exact SM copies, the generated isocurvature perturbations are either dominated by our sector or the dark sector(s).

For the former, if the dark matter is dominated by the axion from the dark sector, we can eliminate θ_2^{ini} by using the equality of (5.20). The bound on H_{I} is then expressed in terms of the dark confinement scale,

$$H_{\text{I}} \lesssim 10^7 \text{ GeV} \left(\frac{f_{a_i}}{10^{12} \text{ GeV}} \right)^{\frac{1}{4}} \left(\frac{1 \text{ GeV}}{\Lambda_{\text{conf}}} \right)^{\frac{1}{2}}. \quad (5.25)$$

In particular, for a two-form axion from such a dark sector, i.e. with $f_{a_2} \sim M_{\text{P}}$ and the minimal possible dark confinement scale $\Lambda_{\text{conf}} \sim \text{eV}$, the bound reduces to $H_{\text{I}} \lesssim 10^{13} \text{ GeV}$. This shows that such an axion is effectively not plagued by an isocurvature constraint.

For N exact SM copies, when the dark matter is collectively composed by the axions from the copies, using (5.17) to eliminate the initial misalignment angles in (5.24) yields

$$H_{\text{I}} \lesssim 10^7 \text{ GeV} \left(\frac{f_{a_i}}{10^{12} \text{ GeV}} \right)^{\frac{1}{4}} \left(\frac{10}{\sqrt{N}} \right). \quad (5.26)$$

We observe that for $N \sim 10^2$ the bound remains relatively unaffected. On the other hand, for values required to solve the Hierarchy problem, i.e. $N \sim 10^{32}$ and $f_a \sim \text{TeV}$, the bound naively tightens to $H_{\text{I}} \lesssim 10^{-10} \text{ GeV}$. However, in this case, the dark confinement scale and the axion mass exceed the inflationary Gibbons-Hawking temperature $T_{\text{I}} \sim H_{\text{I}}$. Consequently, the dark YM sector becomes strongly coupled during inflation and the axion develops a substantial mass, meaning that the requirements for the development of isocurvature perturbations are no longer given. Of course, this requires avoiding the astrophysical bounds on f_a by a mechanism such as clockworking (see discussion at the end of Section 5.1.2).

5.2 Compact Dark Matter

Bosonic dark matter is known to be able to form dense clumps, which in the case of the axion are called axion stars. The expression “star” in this context is used to denote an object sustained by hydrostatic equilibrium, regardless of its emission of light. Such objects as well as other variations of compact dark matter structures have been long studied and provides an interesting portal to the dark matter arising for instance from gravitational waves, gravitational lensing, and large scale structure formation [159–162]. In particular, models with N dark sectors provide interesting compact structures that are made collectively from the N dark sectors and have a mass spectrum suppressed by $1/\sqrt{N}$. In this section, we outline the collective effect leading to these exotic structures that we put forward in [1] and apply it to scalars such as axions.

5.2.1 N -MACHOs

In the case of N identical SM sectors with gravity as the only intersector interaction, these objects can be intuitively understood as follows. When the particles from the N sectors are put within a certain volume, gravitational forces will attempt to collapse the volume. However, the pressure exerted by each sector will oppose this collapse, eventually reaching an equilibrium where pressure and gravity balance each other out. Since each particle can only interact with a small fraction of the present particles, the pressure is immensely reduced compared to the single sector case. At the same time, gravitational forces remain unaffected. Consequently, the equilibrium configuration will have much smaller masses and radii.

To quantify this, we consider the condition for hydrostatic equilibrium in general relativity in the presence of N sectors (see [163] for a pedagogical treatment of the single sector case). The Einstein-Hilbert action without a cosmological constant reads

$$\mathcal{S} = \int \sqrt{-g} d^4x \left(\frac{M_{\text{P}}^2}{2} \mathcal{R} + \mathcal{L}_{\text{tot}} \right), \quad (5.27)$$

where $g = \det g_{\mu\nu}$, \mathcal{R} is the Ricci scalar and $\mathcal{L}_{\text{tot}} \equiv \sum_{j=1}^N \mathcal{L}_j$ denotes the total matter Lagrangian. Varying this action with respect to the metric, yields Einstein’s field equations,

$$G_{\mu\nu} \equiv \mathcal{R}_{\mu\nu} - \frac{1}{2} \mathcal{R} g_{\mu\nu} = \frac{1}{M_{\text{P}}^2} (T_{\text{tot}})_{\mu\nu}, \quad (5.28)$$

where $G_{\mu\nu}$ is the Einstein Tensor and $(T_{\text{tot}})_{\mu\nu}$ the total energy-momentum tensor. For a general Lagrangian \mathcal{L}_j , the energy-momentum tensor is given by

$$(T_j)_{\mu\nu} = -\frac{2}{\sqrt{-g}} \frac{\partial(\sqrt{-g} \mathcal{L}_j)}{\partial g^{\mu\nu}} = -2 \frac{\partial \mathcal{L}_j}{\partial g^{\mu\nu}} + \mathcal{L}_j g_{\mu\nu}. \quad (5.29)$$

Summing over these, give the total energy-momentum tensor

$$(T_{\text{tot}})_{\mu\nu} = \sum_{j=1}^N T_{k\mu\nu} , \quad (5.30)$$

which is the source in the field equations. Considering each sector to behave as a perfect fluid, the energy-momentum tensor in the rest-frame of the fluid can be written as

$$\begin{aligned} (T_{\text{tot}})_{\mu\nu} &= \text{diag}(-\rho_{\text{tot}}, P_{\text{tot}}, P_{\text{tot}}, P_{\text{tot}}) \\ &= \text{diag}\left(-\sum_{j=1}^N \rho_j, \sum_{j=1}^N P_j, \sum_{j=1}^N P_j, \sum_{j=1}^N P_j\right) , \end{aligned} \quad (5.31)$$

where ρ_{tot} and P_{tot} are the energy density and the pressure measured by an observer in that frame. Neglecting rotation, it is reasonable to make a spherical ansatz for compact objects, i.e.

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu = -B(r)dt^2 + A(r)dr^2 + r^2 d\theta^2 + r^2 \sin^2(\theta) d\phi^2 , \quad (5.32)$$

where r, θ, ϕ denote the standard spherical coordinates and $A(r), B(r)$ are the standard coefficients of the spherical ansatz. Plugging this ansatz into the field equations yields the Tolman–Oppenheimer–Volkoff (TOV) equation [164]:

$$\frac{dP_{\text{tot}}(r)}{dr} = -\frac{M(r)\rho_{\text{tot}}(r)}{8\pi M_{\text{P}}^2 r^2} \left(1 + \frac{P_{\text{tot}}(r)}{\rho_{\text{tot}}(r)}\right) \left(1 + \frac{4\pi r^3 P_{\text{tot}}(r)}{M(r)}\right) \left(1 - \frac{M(r)}{4\pi M_{\text{P}}^2 r}\right)^{-1} , \quad (5.33)$$

where

$$M(r) = \int_0^r 4\pi r'^2 \rho_{\text{tot}}(r') dr' . \quad (5.34)$$

The three brackets encode general relativistic corrections, which we neglect in the following. In other words, we consider the Newtonian limit, in which the TOV equation reduces to

$$r^2 \frac{dP_{\text{tot}}(r)}{dr} = -\frac{M(r)\rho_{\text{tot}}(r)}{8\pi M_{\text{P}}^2} , \quad M(r) = \int_0^r 4\pi r'^2 \rho_{\text{tot}}(r') dr' . \quad (5.35)$$

Dividing the left equation by $\rho_{\text{tot}}(r)$ and then differentiating, allows to combine both equations into a single second-order equation,

$$\frac{d}{dr} \left(\frac{r^2}{\rho_{\text{tot}}(r)} \frac{dP_{\text{tot}}(r)}{dr} \right) = -\frac{r^2 \rho_{\text{tot}}(r)}{2M_{\text{P}}^2} . \quad (5.36)$$

To proceed, we need to specify the type of pressure by choosing an equation of state. A simple model for the behavior of any known stellar structure is given by a polytropic equation of state,

$$P_{\text{tot}} = K_{\text{tot}} \rho_{\text{tot}}^\gamma , \quad (5.37)$$

where γ and K_{tot} are constants. A polytropic EOS describes an object with uniform

entropy, which for instance is the case if it has effectively zero temperature or is in convective equilibrium. While these properties are idealistic, they result in the right order of magnitude. With a polytropic equation of state, the TOV equation becomes the Lane-Emden-equation, which is known to have a stable finite radius solution for $\gamma > 6/5$ [163]. This radius is given by

$$R \equiv \left(\frac{2M_{\text{P}}^2 K_{\text{tot}} \gamma}{\gamma - 1} \right)^{1/2} \rho_{\text{tot}}(0)^{(\gamma-2)/2} \xi_0, \quad (5.38)$$

resulting in the mass

$$M = 4\pi \rho_{\text{tot}}(0)^{(3\gamma-4)/2} \left(\frac{2M_{\text{P}}^2 K_{\text{tot}} \gamma}{\gamma - 1} \right)^{3/2} \xi_0^2 |\theta'(\xi_0)|. \quad (5.39)$$

Here, ξ_0 and $-\xi_0^2 |\theta'(\xi_0)|$ are numerical coefficients which for instance can be found in [163].

Let us now turn to the N dependence of the above solution. Consider for each sector a generic polytropic equation of state with polytropic index γ and polytropic constant K . By summing over the contribution from each sector, we find the total pressure to be

$$P_{\text{tot}} = \sum_{j=1}^N P_j = K \sum_{j=1}^N \rho_j^\gamma = \frac{\sum_{j=1}^N \rho_j^\gamma}{\rho_{\text{tot}}^\gamma} K \rho_{\text{tot}}^\gamma \equiv K_{\text{tot}} \rho_{\text{tot}}^\gamma,$$

which again describes a polytropic equation of state. From this we observe that the influence of the N sectors is encoded in the quantity

$$\frac{K_{\text{tot}}}{K} = \frac{\sum_{j=1}^N \rho_j^\gamma}{\rho_{\text{tot}}^\gamma}. \quad (5.40)$$

For sufficiently large N and identical copies, we can assume the densities to be roughly equal, i.e.

$$\rho_{\text{tot}} = \sum_{j=1}^N \rho_j \equiv N \rho, \quad (5.41)$$

where we denoted a single sector without a label. With this, (5.40) simplifies to

$$\frac{K_{\text{tot}}}{K} = \frac{1}{N^{\gamma-1}}. \quad (5.42)$$

Inserting the last expression for K_{tot} into (5.38) and (5.39), we find

$$M \sim \frac{M_0(\rho(0))}{\sqrt{N}}, \quad (5.43a)$$

$$R \sim \frac{R_0(\rho(0))}{\sqrt{N}}, \quad (5.43b)$$

where M_0 and R_0 depend on the particular pressure. Thus, if DM consists of N identical sectors, there exist stellar objects in the spectrum that are lighter by a factor of $1/\sqrt{N}$ compared to the single sector case. For a single dark sector, the difference is almost negligible but for large values as required by the N species solution to the hierarchy problem the impact is dramatic. In [1] we used the Fermi pressure to counteract gravity and called the resulting objects N -MACHOs.

5.2.2 N -Boson Stars

For an axion arising from a dark YM sector, the physics of axion stars are essentially equal to those of our sectors, merely the values of the axion mass and decay constant change. However, the existence of multiple dark sectors may lead to structures that are collectively formed from the axions of the different sectors. The analysis outlined in Sec. 5.2.1 was based on the TOV equation, which characterizes the hydrostatic equilibrium of a self-gravitating, spherically symmetric object made of a perfect fluid. Unlike perfect fluids, the pressure of a scalar field of axion stars cannot be described by a simple equation of state due to a non-trivial self-interaction. Consequently, the equilibrium structure of axion stars cannot be determined solely by the TOV equation. Instead, it requires simultaneously solving the Klein-Gordon equation and Einstein's field equations.

Let us limit ourselves to so called dilute axion stars. To identify such configurations, we follow the standard approach (see for instance [162]) and impose a set of approximations to simplify the relevant equations while still maintaining a high level of accuracy. We aim to maintain a maximal level of generality in our calculations to allow for the application to scalars with various potentials. First, we can utilize non-relativistic effective field theory since the axions produced through misalignment are non-relativistic. In this framework, it is convenient to introduce the complex field $\psi_i(x)$ for each axion through

$$a_i(t, \vec{x}) = \frac{1}{\sqrt{2m_{a_i}}} [\psi_i(t, \vec{x})e^{im_{a_i}t} + \psi_i^*(t, \vec{x})e^{-im_{a_i}t}] . \quad (5.44)$$

By taking the non-relativistic limit of each Klein-Gordon equation and expressing the axion field in terms of ψ_i , we obtain the time-dependent Gross-Pitaevskii equation. To furthermore separate the time dependence, we make the standard separation ansatz $\psi_i(\vec{x}, t) = A_i(\vec{x})e^{-i\omega_i t}$, resulting in the time-independent Gross-Pitaevskii equation,

$$m_{a_i} (\Phi + V_i'(|\psi_i|^2)) - \frac{1}{2m_{a_i}} \frac{\Delta\psi_i}{\psi_i} = E_i . \quad (5.45)$$

Next, the cosmic axion condensate has a relatively low mean mass density, so that we can use Newtonian gravity described by the Poisson equation,

$$\Delta\Phi = \frac{\sum_i m_{a_i} |\psi_i|^2}{M_{\text{P}}^2} . \quad (5.46)$$

Within these approximations, the Gross-Pitaevskii-Poisson equations provide a simplified

set of equations that can accurately describe dilute axion stars. To rewrite this system as a single hydrodynamic equation, we first take the Laplacian of the Gross-Pitaevskii equation and plug in the Poisson equation. Secondly, we perform the Madelung transformation to express everything in terms of the (pseudo) density $\rho_i = m_{a_i} |\psi_i|^2$ and the pressure, which is related to the potential via $\nabla V_i' = \nabla p_i / \rho_i$. This results in the fundamental equation of hydrostatic equilibrium (5.36) with quantum effects taken into account,

$$-\nabla \cdot \left(\frac{\nabla p_i}{\rho_i} \right) + \frac{1}{2m_{a_i}^2} \Delta \left(\frac{\Delta \sqrt{\rho_i}}{\sqrt{\rho_i}} \right) = \frac{\rho_{\text{tot}}}{M_{\text{P}}^2}. \quad (5.47)$$

The first term describes the hydrodynamic pressure, which can either be attractive or repulsive, while the second term characterizes the repulsive quantum pressure due to the Heisenberg uncertainty principle. The total density $\rho_{\text{tot}} = \sum_i \rho_i$ appears on the right hand side since gravity couples to all axions.

The last approximation regards the (pseudo) density ρ , which we take to be small compared to the cosmic condensate density $m_{a_i}^2 f_{a_i}^2$. The effective potential $V_i(\rho_i)$ in this regime is dominated by the leading term in its power series. For the instantonic potential from (5.3), the leading order interaction is the quartic term of a_i . Therefore, in terms of ρ_i the potential can be approximated as

$$V_i(\rho_i) \sim -\frac{1}{m_{a_i}^2 f_{a_i}^2} \rho_i^2. \quad (5.48)$$

This is equivalent to the negative polytropic pressure

$$p_i \sim -\frac{1}{m_{a_i}^2 f_{a_i}^2} \rho_i^2 \equiv K_i \rho_i^{\gamma_i}, \quad (5.49)$$

with polytropic constant $K_i \sim -1/(m_{a_i}^2 f_{a_i}^2)$ and polytropic index $\gamma_i = 2$. With this equation of state, (5.47) becomes

$$-\nabla \cdot \left(\gamma_i K_i \rho_i^{\gamma_i-2} \nabla \rho_i \right) + \frac{1}{2m_{a_i}^2} \Delta \left(\frac{\Delta \sqrt{\rho_i}}{\sqrt{\rho_i}} \right) = \frac{\rho_{\text{tot}}}{M_{\text{P}}^2}. \quad (5.50)$$

For the sake of illustration, let us assume that all densities are equal, i.e. $\rho_i = \rho_{\text{tot}}/N$, as collective effects are most pronounced when all sectors have equal densities [1]. We can then express (5.50) in terms of the total density,

$$-\nabla \cdot \left(\frac{\gamma_i K_i \rho_{\text{tot}}^{\gamma_i-2}}{N^{\gamma_i-1}} \nabla \rho_{\text{tot}} \right) + \frac{1}{2m_{a_i}^2} \Delta \left(\frac{\Delta \sqrt{\rho_{\text{tot}}}}{\sqrt{\rho_{\text{tot}}}} \right) = \frac{\rho_{\text{tot}}}{M_{\text{P}}^2}. \quad (5.51)$$

We observe that in the presence of N axions the pressure from short-range interactions is suppressed by N^{γ_i-1} while the quantum pressure is unaffected.

The influence of N on the solution can be understood by considering the following limiting cases:

- Neglecting the hydrostatic pressure, the equilibrium is between the gravitational

attraction and the repulsion by the quantum pressure. Since both are independent of N , the final configuration is unchanged from the single axion case.

- Neglecting the quantum pressure, the equilibrium is between the gravitational attraction and the hydrostatic pressure (which must be repulsive). From here on the analysis is exactly the same as in Sec. 5.2.1. Since the hydrostatic pressure is strongly suppressed by N^{γ_i-1} , the radius and mass of the final configuration is suppressed by $1/\sqrt{N}$.
- Neglecting the gravitational attraction, the equilibrium is between the hydrostatic pressure (which must be attractive) and the repulsion by quantum pressure. This would again result in an altered mass spectrum, however, such equilibria are always unstable [165].

We can see that stable collective configurations with a suppressed mass spectrum can only exist for repulsive interactions. This is not the case for the axion. Therefore, dilute axion stars would consist of several axions but would not differ in mass or radius compared to the single sector case. The situation is different for scalars with a repulsive interaction (see for instance [166, 167]), which we will not consider further in this thesis.

5.3 Kinetic Mixing between Axions

In this section we discuss non-gravitational intersector interactions by restoring \mathcal{L}_{mix} in (5.1). For concreteness, we start with N exact SM copies and discuss the pure YM sector afterwards.

5.3.1 Kinetic Mixing as Intersector Interaction

Without the additional axions, the possible renormalizable interactions of N exact SM copies that are compatible with gauge-, Lorentz-, and the underlying discrete symmetry are photon kinetic mixing, a Higgs portal coupling, and neutrino mass mixing. At the non-renormalizable level additional oscillations such as neutron oscillations become possible as well. The effects of these interactions have been vastly discussed in the literature (see [168] for a review). Regarding cosmology, the primary impact of say photon kinetic mixing and the Higgs portal coupling is that they can result in thermal equilibrium between our sector and the dark sectors prior to BBN. This would lead to inconsistencies with nucleosynthesis or result in an excessive amount of dark matter particles.

As our focus is not on those topics, we will not elaborate further on their influence and focus on axion kinetic mixing, which in the model under consideration is described by the terms

$$\mathcal{L}_{\text{mix}} = \epsilon \sum_{i \neq j} \partial_\mu a^i \partial^\mu a^j . \quad (5.52)$$

The axions that are produced by misalignment are non-thermal, thus they cannot transfer heat between sectors and thermalize them. However, depending on the UV theory that gives rise to the kinetic mixing term, there can appear bounds from BBN.

In the standard PQ implementation via the KSVZ model, a singlet scalar is added to make the axion invisible. Denoting this singlet in each sector as Φ_i allows for the following dimension six scalar portal couplings between the sectors

$$\mathcal{L}_{\text{mix}}^{\text{UV}} = \frac{1}{M^2} \sum_{i \neq j} (\Phi_i^\dagger \partial_\mu \Phi_i) (\Phi_j \partial^\mu \Phi_j^\dagger) + \text{h.c.} , \quad (5.53)$$

where M is some cut-off scale. The singlets acquire the VEVs f_{a_i} via a proper scalar potential that spontaneously break the $U(1)_{\text{PQ}}$ symmetries. This results in the kinetic mixing described in (5.52) with

$$\epsilon \equiv \frac{f_a^2}{2M^2} . \quad (5.54)$$

Similar operators that result in axion kinetic mixing are also possible in DFSZ-type models, where extra Higgs doublets are required in addition to the singlet.

In addition, (5.53) results in kinetic mixing between the radial modes of the singlets with the same kinetic mixing parameter ϵ . Depending on the particular values of the f_{a_i} and the reheating temperature, the radial modes can be relativistic for some time in the early universe. Then, a bound on ϵ would appear from the requirement to avoid thermalization of the dark sectors before BBN. Since we are mostly interested in large values of N , let us focus on a model independent bound arising from unitarity. The operators in (5.53) lead to loop diagrams, such as the one depicted in Fig. 13, with a large number of scalars involved in the loop. For m such loops, the amplitude of this process scales as $\mathcal{M} \sim N^m (\Delta p/M)^{2m+2} \lesssim \epsilon (N\epsilon)^m$, where Δp is the momentum transfer of the process. For the process to obey unitarity, the amplitude should not exceed unity, thereby resulting in the constraint

$$\epsilon \lesssim \frac{1}{N} . \quad (5.55)$$

In the alternative two-form implementation of the PQ mechanism, axion kinetic mixing appears via non-diagonal mass terms of the three-form action,

$$\mathcal{L}_{\text{mix}} \sim \frac{1}{f_a^2} \sum_{i \neq j} \left(C_{\mu\nu\rho}^i - \partial_{[\mu} B_{\nu\rho]}^i \right) \left(C_{\mu\nu\rho}^j - \partial_{[\mu} B_{\nu\rho]}^j \right) . \quad (5.56)$$

Mixing of this kind is not forbidden by gauge symmetries and can possibly originate from virtual black hole exchange, giving rise to suppression by powers of M_{P} . The reason for this suppression is that micro black holes cannot be universally coupled, which requires gravitational suppression of inter-sector transitions at the fundamental level [47]. Since the UV origin of this PQ implementation is unknown, we do not know if a unitarity bound arises in the same way as in the standard implementation. In the remainder of this section we will study the consequences of the axion kinetic mixings in the pseudo-scalar formulation. We will stay agnostic of their origin, but use the unitarity bound described

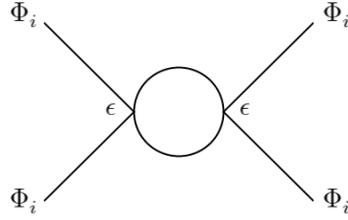


Figure 13: Loop induced by the operators in (5.53). Each vertex contributes the effective coupling $\Delta p^2/M^2$, but for maximum momentum transfers of order f_a the effective coupling becomes ϵ .

by (5.55).

5.3.2 Modification of Axion Physics

The presence of the extra term in the Lagrangian requires a change of basis to express it in a canonical form. As a result, there is a mismatch between the “sector basis”, which corresponds to the labels of the species, and the “canonical kinetic basis”, in which the propagator is in canonical form. So the task is to find the relation between these two bases.

In our case the mismatch appears in the axion sector. The first step is to express the kinetic part in (5.52) as

$$\mathcal{L} \supset \begin{pmatrix} \partial_\mu a_1 \\ \vdots \\ \partial_\mu a_N \end{pmatrix}^T \begin{pmatrix} 1 & \epsilon & \dots & \epsilon \\ \epsilon & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \epsilon \\ \epsilon & \dots & \epsilon & 1 \end{pmatrix} \begin{pmatrix} \partial_\mu a_1 \\ \vdots \\ \partial_\mu a_N \end{pmatrix}. \quad (5.57)$$

We can rewrite the matrix, which we shall call K from now on, in the following way,

$$K = \begin{pmatrix} 1 - \epsilon & 0 & \dots & 0 \\ 0 & 1 - \epsilon & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 - \epsilon \end{pmatrix} + \epsilon \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix}. \quad (5.58)$$

In this way, the problem reduces to the diagonalization of a matrix of just ones. The transformation matrix S that diagonalizes K is

$$S = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & -1 & 0 & \dots & 0 \\ 1 & 0 & -1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & 0 & \dots & 0 & -1 \end{pmatrix}. \quad (5.59)$$

The first row give rise to the eigenstate of the form,

$$\tilde{a}_L = \frac{1}{\sqrt{N}}a_1 + \sqrt{\frac{N-1}{N}}a_h, \quad (5.60)$$

where we have introduced the notation

$$a_h = \frac{1}{\sqrt{N-1}} \sum_{i=2} a_i \quad (5.61)$$

due to later convenience. The eigenstate \tilde{a}_H corresponds to the eigenvalue of $1 + (N-1)\epsilon$. Because the matrix S is not a unitary matrix we still have to find a convenient basis.

From all other rows, we see that we have $N-1$ degenerate eigenstates v_i of the eigenvalue $1 - \epsilon$. These v_i are the columns of S . Due to this degeneracy, we can reduce the $N \times N$ problem to a 2×2 problem by defining a superposition of these degenerate eigenstates,

$$\tilde{a}_H = \sqrt{\frac{N-1}{N}}a_1 - \frac{1}{\sqrt{N}}a_h. \quad (5.62)$$

The eigenvalue of this superposition is again $1 - \epsilon$. We see that \tilde{a}_H is a collective expression made out of all former a_i . Inverting the eigenstates, we find a_1 expressed in the canonical kinetic basis as

$$a_1 = \sqrt{\frac{N-1}{N}}\tilde{a}_H + \frac{1}{\sqrt{N}}\tilde{a}_L. \quad (5.63)$$

Expressing (5.57) in terms of \tilde{a}_L and \tilde{a}_H , there is no more mixing but each term is multiplied by the corresponding eigenvalue. In order to have canonical kinetic terms, the states \tilde{a}_L and \tilde{a}_H need to be redefined by

$$\tilde{a}_H \rightarrow \frac{1}{\sqrt{1-\epsilon}}\tilde{a}_H, \quad \tilde{a}_L \rightarrow \frac{1}{\sqrt{1+(N-1)\epsilon}}\tilde{a}_L. \quad (5.64)$$

One effect of this redefinition is that the mass terms in the Lagrangian become

$$\mathcal{L}_{\text{mass}} \sim \begin{pmatrix} \tilde{a}_H \\ \tilde{a}_L \end{pmatrix}^T \begin{pmatrix} \frac{m_a^2}{1-\epsilon} & 0 \\ 0 & \frac{m_a^2}{1+(N-1)\epsilon} \end{pmatrix} \begin{pmatrix} \tilde{a}_H \\ \tilde{a}_L \end{pmatrix}. \quad (5.65)$$

with m_a being the mass induced by the PQ mechanism. We observe that kinetic mixings leads to a splitting of the masses of the axions. $N-1$ axions are degenerated and one light axion whose mass is suppressed by the number of copies. In other words, kinetic mixing of many copies of the axion leads to two different detectable axion states with different masses. The relation of these masses is

$$\frac{m_L}{m_H} = \left(\frac{1-\epsilon}{1+(N-1)\epsilon} \right)^{\frac{1}{2}} \sim \frac{1}{\sqrt{2}}, \quad (5.66)$$

where we assumed natural values of ϵ with respect to the unitarity bound, i.e. $\epsilon \sim N^{-1}$, and $N \gg 1$ in the similarity.

With the relation between the species basis and the canonical kinetic basis, we can turn to the phenomenological implications of this model. Due to the Goldstone nature of the axion, its lowest order couplings to fermions and gauge bosons have the generic form

$$g_{ao} a_1 \mathcal{O}, \quad (5.67)$$

where the whole UV dependency is encoded in g_{ao} . After the diagonalization and field redefinition of (5.64) the expression becomes

$$g_{ao} \left(\sqrt{\frac{N-1}{N}} \frac{1}{\sqrt{1-\epsilon}} \tilde{a}_H + \frac{1}{\sqrt{N}} \frac{1}{\sqrt{1+(N-1)\epsilon}} \tilde{a}_L \right) \mathcal{O}. \quad (5.68)$$

We see that by having N kinetically mixed axions, our sector couples to N axions instead of one. These N axions come in two categories: $N-1$ axions encoded in \tilde{a}_H that behave exactly the same, and one special axion \tilde{a}_L . The single axion coupling g_{ao} gets modified by one of the following factors,

$$f_H(N, \epsilon) = \sqrt{\frac{N-1}{N}} \frac{1}{\sqrt{1-\epsilon}}, \quad (5.69)$$

$$f_L(N, \epsilon) = \frac{1}{\sqrt{N}} \frac{1}{\sqrt{1+(N-1)\epsilon}}. \quad (5.70)$$

These demonstrate a different behavior regarding the two parameters N and ϵ . Imposing the unitarity constraint (5.55) on ϵ , the coupling of \tilde{a}_H is essentially indistinguishable from a single ordinary axion from our sector, while \tilde{a}_L shows a suppression by $1/\sqrt{N}$.

Let us focus on the axion-photon coupling, which becomes

$$g_{a\gamma}^{H/L} = \frac{\alpha}{2\pi f_a} \left(\frac{\mathcal{E}}{\mathcal{N}} - \frac{2}{3} \frac{4m_d + m_u}{m_u + m_d} \right) f^{H/L}(N, \epsilon). \quad (5.71)$$

In Fig. 14, we show the band in which the light axion state would reside relative to the original KSVZ prediction for $N = [1, 10^6]$ (as motivated by Sec. 5.1). In the framework of many axions, the next step after discovery of the ordinary axion would be to search for a second light weakly coupled state. After a potential discovery of the axion, m_H and $g_{a\gamma}^H$ would be known. The ratio between $g_{a\gamma}^L$ and $g_{a\gamma}^H$ is given by

$$\frac{g_{a\gamma}^L}{g_{a\gamma}^H} \sim \frac{1}{\sqrt{N}}. \quad (5.72)$$

Since the masses of both states almost coincide, this means that after measuring the coupling and the mass of the first axion, the properties of the second are uniquely determined by N . For example, if $N = 10^6$ would be realized the couplings are related by $g_{a\gamma}^L \sim 10^{-3} g_{a\gamma}^H$. We depict this in Fig. 14. The included projected sensitivities of

the strong CP problem to a consistency problem. As such, an axion becomes a necessary component not only in QCD but in every YM group. We focused on two models, namely N exact copies of the SM and a single pure YM sector.

We first pointed out that the total axion density from all sectors should not be higher than the observed dark matter density. Using the misalignment mechanism on all axions, for N exact SM copies this results in an upper bound on N . For a single pure YM sector, we find a relation between the dark confinement scale and the PQ scale. Notably, the misalignment mechanism is independent of the thermalization of the dark sectors and thus our findings apply even for large values of N when the sectors are very dilute.

Furthermore, we showed that the contribution of additional axions to the isocurvature fluctuations does not necessarily tighten the bound on H_1 . In fact, both models exhibit parameter space regions where the bound is essentially not present. For N exact SM copies, the viable parameter space requires either a moderate number of sectors, i.e. $N \lesssim 10^2$, or a mechanism that allows for smaller PQ scales.

In models with several dark sectors it is possible that particles from different sectors collectively form structures with a mass spectrum suppressed by $1/\sqrt{N}$. While this phenomenon was originally studied with fermions, we showed that it also applies to bosons with a repulsive self-interaction. Since axions have an attractive self-interactions, this phenomenon cannot take place. Therefore, the presence of multiple axions does not modify the mass spectrum of axion stars.

Lastly, we studied non-gravitational communication between the dark sectors from axion kinetic mixing. We showed that a mass splitting into $N - 1$ degenerate states and one light state emerges. Interestingly, both states have almost similar masses and the lighter has couplings weaker by a factor of $1/N^{1/2}$ with respect to the heavier axion states. If two axions were discovered, this would allow to determine N . For a single YM sector the axion masses are different, so that the axions from the dark sector could decay into photons. Furthermore, the kinetic mixing parameter ϵ is constrained by the requirement to not reheat the dark sector prior to BBN. We are currently investigating these two phenomena in order to improve the phenomenological bounds that we have found so far.

CHAPTER SIX

NON-MINIMAL DFSZ MODELS

In Sec. 3.3, we pointed out that the PQ solution does not specify the axion low-energy couplings as a result of the axions Goldstone nature [50]. UV models are thus needed to make concrete predictions about the couplings of the axion. Two commonly utilized models for this purpose are the DFSZ [51, 52] and KSVZ [53, 54] invisible axion models. So far, we have only addressed the minimal versions of these models. However, there exists a plethora of non-minimal generalizations that cannot be overlooked, even though minimal models are generally preferred due to reasons such as Occam’s razor and predictiveness. It is possible that a physical principle renders the minimal model invalid, so knowledge of non-minimal models is crucial. In addition, comprehending the landscape of non-minimal models enables us to answer questions such as the extent to which the parameter space is excluded in the absence of experimental signatures. For these reasons, it would be beneficial to identify all non-minimal models and develop a systematic approach that enables us to extract predictions from all of them concurrently. The objective of the project discussed in this chapter is to achieve exactly this for DFSZ-type axions.

We achieve this by exploiting the unique property of the axion-photon coupling (3.17), that its UV physics are fully encoded in the ratio between the electromagnetic and the QCD anomaly coefficients [50]. For convenience, let us introduce the following definition,

$$g_{a\gamma} = \frac{\alpha}{2\pi f_a} \left[\frac{\mathcal{E}}{\mathcal{N}} - 1.92(4) \right] \equiv \frac{\alpha}{2\pi f_a} \mathcal{C}_{a\gamma}, \quad (6.1)$$

where by $\mathcal{C}_{a\gamma}$ we denote the dimensionless part. In contrast to the other couplings, this ratio is not influenced by unknown VEVs or mixing angles due to the nature of anomalies. Instead, it is determined solely by the representations of the fields. In the case of DFSZ-type models, this entails fixing the PQ charges of the SM fermions, which are not free but determined by linear consistency and phenomenology conditions [55]. By systematically solving the associated linear system of equations (LSE), we are able to

compute the anomaly ratio and consequently the axion-photon coupling for a vast range of DFSZ-type models.

In addition to computing the anomaly ratios for numerous DFSZ-type models, we are also able to determine how many distinct models yield the same anomaly ratio. We employ this concept of multiplicity to assign a probability to each anomaly ratio. By analyzing the resulting distributions, we can extract various intriguing insights for the axion experimental program. One of our key observations is that the values advocated by the minimal DFSZ models, i.e. $\mathcal{E}/\mathcal{N} = 2/3$ and $\mathcal{E}/\mathcal{N} = 8/3$, are statistically favored, despite the existence of numerous other possible anomaly ratios for DFSZ-type models. While this confirms the potential experimental significance of these values, we argue that a non-observation at these values still leaves a considerable amount of the axion parameter space viable. We quantify this statement by defining an axion band and lower bounds for $g_{a\gamma}$.

A similar investigation has previously been conducted for KSVZ-type axions: the identification and categorization of these models are outlined in [106], while the statistical analysis is presented in [170]. Additionally, in [106], a section is dedicated to DFSZ-type axions. There, by approximating the maximum possible anomaly ratio, the authors contend that the majority of realistic DFSZ-type models lie within the same range as the favored KSVZ-type models. With our study, we not only provide a more accurate estimate of the maximum feasible anomaly ratio, which turns out to be larger than the prior estimate, but we also conduct a detailed comparison between the two classes, which enables us to gain a better understanding of their relation.

This chapter is organized as follows. First, in Sec. 6.1, we define the non-minimal DFSZ models under consideration, which we refer to as DFSZ-type models. We place a particular emphasis on the determination of the PQ charges and potential phenomenological selection criteria. Furthermore, we provide a general procedure for determining all possible anomaly ratios and their multiplicities for a given number of Higgs doublets. In Sec. 6.2, we apply this approach to models with three to nine Higgs doublets, discussing the problems that arise for a large number of doublets and comparing our results with those of KSVZ-type models. Next, Sec. 6.3 explores the experimental implications of our findings by estimating the necessary sensitivities for axion searches. Finally, in Sec. 6.4, we summarize our results.

6.1 DFSZ-Type Models

We consider DFSZ-type models that extend the original DFSZ model by allowing for more Higgs doublets. Hence, the scalar sector consists of at least one extra Higgs doublet to make the PQ symmetry anomalous with respect to QCD, and one singlet to make the axion undetectable by decoupling the PQ scale from the electroweak scale. The inclusion of more Higgs doublets results in more freedom in assigning the PQ charges to the fermions, leading to a larger number of possible anomaly ratios.

In the DFSZ-type models, the anomaly of the PQ current is solely determined by the difference between the PQ charges of left- and right-handed fermions. To simplify the analysis, we assume that the PQ charges of the left-handed fermions are zero. This leaves us with the PQ charges of the right-handed fermions, denoted as $\chi_{u_i}, \chi_{d_i}, \chi_{e_i}$ with i being a generation index. The situation is different for neutrinos, as their left-handed component does not directly contribute to the anomaly ratio \mathcal{E}/\mathcal{N} . However, if there exists a right-handed neutrino in the theory, the left-handed neutrino could indirectly contribute. Since it is uncertain whether the neutrino masses are realized via the type-I seesaw mechanism, which necessitates the inclusion of right-handed neutrinos, we exclude the neutrinos in our analysis by setting their PQ charge to zero in accordance with the other left handed fermions.

Throughout this chapter, we will make several comparisons with KSVZ-type models. We define those as in [106], i.e. as those generalizations of the original KSVZ model that introduce additional heavy fermions in per se arbitrary representations and a singlet scalar to the SM in order to address the strong CP problem.

6.1.1 Identifying the Axion

Let us use the notation DFSZ_{n_D} to refer to DFSZ-type models, where n_D represents the total number of doublets. For concreteness, we consider a DFSZ_{n_D} model with n_D being less than or equal to nine and start by defining the Yukawa sector. To fully exhaust the freedom of PQ charges, we consider a Yukawa sector where each right-handed fermion couples to only one doublet. This approach allows to refer to the doublets as H^{u_i}, H^{d_i} , and H^{e_i} , while the singlet can be denoted as S . The Yukawa sector then takes the form

$$\mathcal{L} \supset -y_{ij}^u H^{u_i} \bar{Q}_L^i u_R^j - y_{ij}^d H^{d_i} \bar{Q}_L^i d_R^j - y_{ij}^e H^{e_i} \bar{E}_L^i e_R^j + \text{h.c.} . \quad (6.2)$$

For $n_D = 9$ each right-handed fermion couples to a different doublet, while for $n_D < 9$ some fermions have to couple to the same doublet. This form of the Yukawa sector fixes the weak hypercharge of the doublets to be

$$-Y_{H^{u_i}} = Y_{H^{d_i}} = Y_{H^{e_i}} = \frac{1}{2} . \quad (6.3)$$

In principle several doublets can couple to the same right-handed fermion. For the time being, we disregard this matter and revisit it in Sec. 6.1.3.

The standard kinetic term for each scalar is invariant under a $U(1)^{n_D+1}$ symmetry. To ensure the PQ current is well-defined and to prevent the existence of Goldstone bosons with decay constants of the electroweak scale order, it is necessary to explicitly break down this symmetry to $U(1)_{\text{PQ}} \times U(1)_Y$. With this requirement in mind, we decompose the potential into two parts,

$$V = V_{\text{moduli}} + V_{\text{eb}} . \quad (6.4)$$

The first term only contains the modulus of each scalar or the modulus of two doublets and hence does not break any of the global $U(1)$ groups explicitly. In contrast, V_{eb} consists of terms that all break the $U(1)^{n_D+1}$ symmetry explicitly. Since the broken symmetry must be identified with $U(1)_{\text{PQ}} \times U(1)_Y$, the number of terms in V_{eb} required is $n_D - 1$.

By choosing the parameters such that we have a proper scalar potential, each scalar field develops a VEV v_f , where we introduced the index $f = u_i, d_i, e_i, S$. Expanding around these VEVs yields,

$$\begin{aligned} H_{d_i} &\supset \frac{v_{d_i}}{\sqrt{2}} e^{i \frac{a_{d_i}}{v_{d_i}}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} , & H_{u_i} &\supset \frac{v_{u_i}}{\sqrt{2}} e^{i \frac{a_{u_i}}{v_{u_i}}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} , \\ H_{e_i} &\supset \frac{v_{e_i}}{\sqrt{2}} e^{i \frac{a_{e_i}}{v_{e_i}}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} , & S &\supset \frac{v_S}{\sqrt{2}} e^{i \frac{a_S}{v_S}} . \end{aligned} \quad (6.5)$$

Here, we only consider angular degrees of freedom that contain the axion. Each angular mode a_f transforms under a PQ transformation as $a_f \rightarrow a_f + \kappa_f \chi_f v_f$ with the κ_f being constants. The PQ current after spontaneous symmetry breaking takes the form

$$\begin{aligned} J_\mu^{\text{PQ}} \Big|_a &\supset -\chi_S S^\dagger i \partial^\mu S - \sum_{f \setminus S} \chi_f H_f^\dagger i \partial^\mu H_f + \text{h.c.} \\ &= \sum_f \chi_f v_f \partial_\mu a_f . \end{aligned} \quad (6.6)$$

By requiring $J_\mu^{\text{PQ}} \Big|_a = v_a \partial_\mu a$ and $a \rightarrow a + \kappa v_a$ under the PQ transformation, the axion field is defined as a linear combination of all scalar angular modes,

$$a = \frac{1}{v_a} \sum_f \chi_f v_f a_f , \quad v_a^2 = \sum_f \chi_f^2 v_f^2 . \quad (6.7)$$

With the axion identified, the low energy theory can be constructed like in the original DFSZ model. Inverting (6.7) allows to express the scalar angular modes in terms of the axion. Since our focus is on terms that include the axion, this essentially amounts to the substitution

$$\frac{a_f}{v_f} \rightarrow \chi_f \frac{a}{v_a} . \quad (6.8)$$

The Lagrangian can then be brought to the form of the axion EFT in (3.4) by performing

the field-dependent chiral redefinition of the fermion fields,

$$\psi \rightarrow \exp\left(-i\gamma_5\chi_f\frac{a}{2v_a}\right)\psi. \quad (6.9)$$

In the considered models, all representations except for the PQ charges are known. Therefore, the ratio between the electromagnetic and color anomaly coefficients can be expressed in as [106]

$$\frac{\mathcal{E}}{\mathcal{N}} = \frac{\sum_i \frac{4}{3}\chi_{u_i} + \frac{1}{3}\chi_{d_i} + \chi_{e_j}}{\frac{1}{2}\sum_i \chi_{u_i} + \chi_{d_i}} = \frac{2}{3} + 2\frac{\sum_i \chi_{u_i} + \chi_{e_i}}{\sum_i \chi_{u_i} + \chi_{d_i}}. \quad (6.10)$$

6.1.2 The PQ Charges

When a new U(1) is introduced, the charges are a priori free. However, there can be phenomenological and theoretical constrains on this choice. This is the case for DFSZ-type models. In particular, the $U(1)^{n_D+1}$ symmetry must be explicitly broken into $U(1)_{\text{PQ}} \times U(1)_Y$ to avoid Goldstones with decay constants of electroweak scale couplings. This explicit breaking must respect the following conditions [55]:

1. Orthogonality between J_μ^{PQ} and the weak hypercharge current J_μ^Y .
2. Invariance under PQ symmetry.
3. Well-definiteness of domain wall number N_{DW} .

Consequently, the PQ charges are determined by the $n_D + 1$ relations following from these requirements. Solving the resulting linear system yields all PQ charges.

Let us discuss these conditions in detail. To begin with, the orthogonality requirement between the PQ current defined in (6.6) and the weak hypercharge current $J_\mu^Y|_a = \sum_f Y_f v_f \partial_\mu a_f$ implies

$$\sum \chi_f Y_f v_f^2 = 0. \quad (6.11)$$

From this relation one can immediately see that in general the PQ charges are not integer numbers, which also follows from the fact that $U(1)_{\text{PQ}}$ is not compact.

Regarding the PQ invariance, the $n_D - 1$ terms in V_{eb} are divided into two categories: those composed of two doublets and two singlets ($HHSS$), and those consisting of four doublets ($HHHH$). To ensure the invisibility of the axion, at least one term of the type $HHSS$ must be present, while the form of the remaining $n_D - 2$ terms is unconstrained in principle. We only consider renormalizable terms, thereby excluding higher-order scalar terms. When choosing the terms in V_{eb} , it is necessary to ensure that they give rise to linearly independent conditions. This means that V_{eb} should have enough terms to render the system exactly solvable and neither underdetermined nor overdetermined. If the system is underdetermined, it will not explicitly break enough of the $U(1)^{n_D+1}$ symmetry, resulting in undesired massless states. Conversely, if the system is overdetermined, it will be inconsistent.

After imposing the conditions of orthogonality and PQ invariance, it becomes possible to solve for all PQ charges in terms of χ_S , which is a singlet and otherwise unconstrained. It is important to note that the value of χ_S is not relevant for the anomaly ratio since it cancels out in the ratio. Therefore, for a given set of terms in V_{eb} , all PQ charges can be expressed in terms of χ_S , allowing for the calculation of the anomaly ratio. This is the main message regarding the PQ charges.

However, there are quantities in which χ_S does not cancel, and one of them is the domain wall number. Since domain wall numbers larger than unity can lead to the domain wall problem, it is helpful to also fix χ_S . Following our discussion in Sec. 4.1.2 regarding the caveats with residual cyclic symmetries in theories where the axion is a superposition of angular modes, we fix χ_S by the definition of the domain wall number in (4.34). In the models under consideration, the domain wall number takes the form $N_{\text{DW}} = q \chi_S$, where q is a rational number. Then, the definition (4.34) uniquely fixes χ_S to be the denominator of q .

6.1.3 Multiplicity

The just described method of fixing the PQ charges allows for a straightforward calculation of all possible anomaly ratios, at least in principle. To define a precise concept of *multiplicity*, however, more specification is required to avoid the possibility of overcounting. The standard consensus in constructing models is to include all terms allowed by given symmetries. Regarding V_{eb} , this implies that potentials that yield the *same* PQ charges should not be considered different, as they can simply be added together. This can be understood in the language of conditions and LSEs. The construction described in Sec. 6.1.2 required $n_D - 1$ terms in the explicit breaking potential. Including fewer terms would result in undesired Goldstone bosons, while incorporating too many independent terms would lead to overdetermined systems with $\chi_f = 0$ for all f , thus failing to solve the strong CP problem. However, one can add additional terms to the potential that give rise to redundant conditions. These potentials correspond to the same solution of the underlying LSE, or in other words, they have the same PQ charges.

This reasoning also affects the construction of the Yukawa sector. While the initial construction couples a single Higgs doublet to each right-handed fermion, in principle, multiple doublets can couple to the same right-handed fermion. Therefore, the Yukawa sector is completed a posteriori for each set of possible PQ charges. For example, if a solution for some LSE is $\chi_{d1} = \chi_{e1}$, then the Yukawa sector for that solution would be

$$\begin{aligned} y_{1j}^d H^{d1} \bar{Q}_L^1 d_R^j &\longrightarrow (y_{1j}^d H^{d1} + \tilde{y}_{1j}^d H^{e1}) \bar{Q}_L^1 d_R^j, \\ y_{1j}^e H^{e1} \bar{E}_L^1 e_R^j &\longrightarrow (\tilde{y}_{1j}^e H^{d1} + y_{1j}^e H^{e1}) \bar{E}_L^1 e_R^j. \end{aligned} \quad (6.12)$$

By adding the potentials and completing the Yukawa sectors for a specific set of PQ charges, we ensure that all possible Yukawa terms compatible with that set are included (such as cross-couplings where up-type doublets couple to down-type fermions). Moreover,

since each set of PQ charges is unique after adding the potentials, the Yukawa sector with all compatible couplings is uniquely determined, and no additional multiplicities need to be taken into account. To conclude, adding the potentials and completing the Yukawa sectors for a specific set of PQ charges specifies one model for the counting of the multiplicity. The final step is to calculate the anomaly ratio for each model and count its multiplicity, which completes the construction procedure.

6.1.4 Selection Criteria

Now that we have established the definition of a model in regard of the multiplicity, we can explore the possibility of applying (phenomenological) selection criteria to identify favored axion models.

In the KSVZ-type models, all selection criteria stem from the existence of new fermions [106]. For instance, if the new fermions are excessively massive and long-lived, they are subject to stringent constraints from BBN and Cosmic Microwave Background (CMB) observations. Furthermore, since their mass is related to the axion decay constant f_a , the precise value of f_a is of crucial importance. Lastly, the presence of additional quarks may significantly alter the running of the QCD coupling constant, potentially violating asymptotic freedom or leading to Landau poles below the Planck scale. Since none of these constraints apply in the DFSZ case, we will not discuss them further (for a detailed discussion, see [106] or [170]).

Next, let us briefly discuss the aspects that are present in both types of invisible axion models, starting with the domain wall problem. As explained in Sec. 4.1.2, models with $N_{\text{DW}} > 1$ result in the domain wall problem. Hence, one can consider imposing $N_{\text{DW}} = 1$ as a selection criterion for axion models. However, there are several ways to avoid the domain wall problem. First of all, it is not present when the PQ symmetry is broken during or before inflation since no domain walls form within the Hubble horizon. In the scenario where the PQ symmetry is broken after inflation, it is possible that the symmetry is not restored at high temperatures, thus avoiding the production of strings and walls [171]. Alternatively, by embedding the discrete subgroup into a continuous group, the different vacua become related via symmetry transformations, resulting in an effective domain wall number of unity [172]. Given these viable solutions, we do not consider $N_{\text{DW}} = 1$ to have a sufficient level of generality to represent a necessary selection criterion for our main analysis. Nonetheless, we calculate the domain wall number for DFSZ₃ to DFSZ₇, demonstrate the influence of this selection criterion, and compare it with the KSVZ case in Sec. 6.3.2.

Another aspect that is present in both types of invisible axion models is the impact of additional Higgs doublets on the running of the electroweak gauge coupling. While the maximal case of $n_D = 9$ appears to enhance unification compared to the SM, the resulting unification scale of $\Lambda_{\text{GUT}} \sim 10^{13}$ GeV leads to a rapid proton decay, rendering the improvement of unification not suitable as a selection criterion. Moreover, for $n_D \sim 50$ asymptotic freedom is spoiled and a Landau Pole emerges below the Planck scale, setting

an upper limit on the number of doublets [106]. Because of these arguments and for the sake of better comparability with the KSVZ case, we do not consider improvement of unification as a selection criterion.

Let us now consider an aspect that is specific to the DFSZ-type models, namely the general feature of multi-Higgs doublet models to include FCNCs. Since FCNC are subject to strong experimental constraints [173], they could potentially limit the number of viable DFSZ-type models. However, as with the domain wall problem, there exist viable solutions to avoid the presence of FCNC (for a review, see [174]):

Natural Flavor Conservation: The easiest way to avoid FCNCs is to impose the Weinberg-Glashow-Paschos condition [105, 175], which requires all right-handed fermions of a given electric charge to couple to only one of the doublets, as in the original models. Imposing this condition effectively sets several Yukawa couplings to zero, which for $n_D > 3$ results in $n_D - 3$ decoupled Higgs doublets. Hence, for DFSZ-type models as we have defined them in the beginning of this section, natural flavor conservation is only possible for $n_D \leq 3$.

Flavor Alignment: A less restrictive possibility is to impose an alignment condition, i.e. requiring the Yukawa matrices of each right-handed fermion to be proportional to one Yukawa matrix. All Yukawa matrices are then simultaneously diagonalised in the fermion mass-eigenbasis, yielding no FCNC at tree level [176–178].

Mass Matrix Ansätze: Another possibility is to take the Yukawa matrices to have a specific texture in flavour space. This allows viable SM mass and mixing phenomenology and sufficient suppression of the tree-level FCNCs [179].

Implementing Natural Flavor Conservation and Mass Matrix Ansätze typically requires imposing (discrete) symmetries that protect the flavor structure from quantum corrections. However, imposing additional symmetries on the scalar potential can spoil the decoupling property of general multi-Higgs doublet models [180]. This implies that the new scalar cannot have arbitrarily large masses, which can result in significant deviations from the measured SM couplings. Thus, in order to avoid FCNC using these solutions, it would be necessary to systematically determine which of our models have discrete symmetries that avoid FCNC while allowing for a decoupling limit. Such an analysis is beyond the scope of this work due to the large number of models and the lack of a catalog of possible symmetries for $n_D > 3$ [174]. Flavor Alignment, on the other hand, preserves the decoupling limit but leaves the flavor structure vulnerable to quantum corrections. However, residual flavor symmetries can sufficiently mitigate this vulnerability [181]. In summary, for the DFSZ-type models, we have identified desirable features for specific models but no selection criteria that have a sufficient level of generality.

Finally, we note that it is possible to enlarge the definition of DFSZ-type axions by including more singlets or doublets beyond $n_D = 9$, which do not couple to SM fermions. Although this does not alter the axion-photon coupling given by (6.10), it can yield very large PQ charges in an indirect way [106, 182]. We limit our analysis to the narrower definition of DFSZ models introduced earlier in this section, which can be regarded as a form of selection criterion.

6.2 Anomaly Ratio Distributions

We demonstrated how for the DFSZ-type axions the calculation of the anomaly ratio reduces to fixing a V_{eb} and solving the resulting LSE. Therefore, to compute all possible anomaly ratios, it is necessary to perform this calculation for every feasible V_{eb} . Moreover, we defined a model to consist of the sum of all V_{eb} that give rise to the same set of PQ charges. By counting how many different models lead to the same anomaly ratio enables us to assign a probability to each anomaly ratio within the considered set of models. With this probability distribution, we can define lower bounds for $|\mathcal{C}_{a\gamma}|$ above which the majority of DFSZ-type axion models can be found.

Our procedure can be summarized as a cooking recipe, which consists of the following steps:

1. Specify the Yukawa sector for a fixed n_D by coupling one doublet to each right-handed fermion. This exhausts the maximal freedom regarding the anomaly ratio.
2. Write down all possible V_{eb} with $n_D - 1$ terms.
3. Solve all associated LSEs to find all possible sets of PQ charges. Underdetermined systems are discarded.
4. Add the potentials of all equal PQ charges to get the most general potential associated with a particular solution. This defines one model for the sake of counting the multiplicity.
5. For each model complete the Yukawa sector by adding all Yukawa terms compatible with the PQ- and hypercharges.
6. For each model calculate the anomaly ratio and count its multiplicity.

We calculate the PQ charges and anomaly ratios numerically using the programming language “Julia” [183]. The “StaticArrays” package [184] allows us to compute the extremely large number of LSEs very fast without heap memory allocation. Since it is not relevant to the acquired solutions, we skip step 5 in practice.

6.2.1 Example: $n_D = 3$

Let us illustrate our approach using DFSZ₃ as an example, where we impose the Weinberg-Glashow-Paschos condition. In this case, there are three possible bilinears, namely $(H_u H_d)$, $(H_u H_e)$, and $(H_d H_e^\dagger)$, along with their complex conjugates. Each bilinear can either couple to the singlet, resulting in 6 different terms of the form $HHSS$, or to another bilinear, resulting in 36 different quadrilinears of the form $HHHH$. For the latter case, after removing terms related by Hermitian conjugation and those resulting in no condition, the number is reduced to 9 (see Tab. 1). For $n_D = 3$, the breaking potential consists of either one $HHSS$ and one $HHHH$ term, or two $HHSS$ terms. The former has 54 possible combinations and the latter has 15 possible combinations, totaling to 69 possible combinations for V_{eb} (see Tab. 2).

The resulting 3×3 LSEs consist of the orthogonality relation, $\chi_u v_u^2 - \chi_d v_d^2 - \chi_e v_e^2 = 0$, and the two conditions coming from the potential. Solving the LSEs yields the PQ charges in terms of χ_S , which can then be fixed by our definition of the domain wall number. We can perform the following two simplifications for the purpose of calculating the anomaly ratio. First, we can set all VEVs equal to one because E and N are independent of them, and secondly, we can leave χ_S unfixed because it cancels in the anomaly ratio after expressing all PQ charges in terms of χ_S .

Of the 69 minimal potentials found, many have no or degenerate solutions. For example, potentials including a bilinear and its Hermitian conjugate at the same time gives rise to an underdetermined system, and the nine quadrilinears only give six unique conditions for PQ charges. A summary of all solutions can be found in Tab. 2 (top). In total this leaves us with only 16 different solutions for the PQ charges, for each of which we have to add all the terms to the potential that give rise to the same set of PQ charges.

The Yukawa sector in this example does not need any completion since it is already fixed by the Weinberg-Glashow-Paschos condition. Hence, it merely remains to plug the different sets of PQ charges into (6.10), yielding the following possible anomaly ratios (see Tab. 2, bottom),

$$\text{DFSZ}_3 : \frac{\mathcal{E}}{\mathcal{N}} = -\frac{4}{3}, \frac{2}{3}, \frac{5}{3}, \frac{8}{3}, \frac{14}{3}. \quad (6.13)$$

Counting the multiplicity, we find that $2/3$ and $8/3$ each appear $2 \times$ with four terms in the potential each, and $-4/3$, $5/3$, as well as $14/3$ each appear $4 \times$ with three or two terms in the potential each. A visualization of this result together with all other n_D values can be found in Fig. 17. For a summary of important statistics in this model see Tab. 3.

It turns out useful in the following to introduce a compact notation that encodes which doublet couples to which of the nine fermions. For this, we assign to the nine fermions a position in a nine dimensional row vector with square brackets and write the subscript of the doublets that couple to a certain fermion to the corresponding position. If one doublet couples to multiple fermions, we use the first subscript in the order presented above. For a more comprehensive notation, we use the fermion type (up-, down- or lepton-type, short u , d or e) and the generation (1 to 3), so that

$$\begin{array}{cccccccccc} u & c & t & d & s & b & e & \mu & \tau & \\ u_1 & u_2 & u_3 & d_1 & d_2 & d_3 & e_1 & e_2 & e_3 & \\ [\cdot, \cdot, \cdot, \cdot, \cdot, \cdot, \cdot, \cdot, \cdot] , & & & & & & & & & \end{array} \quad (6.14)$$

For DFSZ₉, this row vector would be $[u_1, u_2, u_3, d_1, d_2, d_3, e_1, e_2, e_3]$ while for the original DFSZ₂-I model it would be $[u_1, u_1, u_1, d_1, d_1, d_1, d_1, d_1, d_1]$.

6.2.2 Choices for a Statistical Interpretation

We are considering many different solutions for the PQ charges. In Sec. 6.2.1 we just counted the number of models leading to specific anomaly ratios, but in the end we want to translate a catalog of models with specific \mathcal{E}/\mathcal{N} values to a probability distribution of anomaly ratios. To achieve this, we require relative probabilities of the solutions, which are subject to some sort of theoretical prior belief. Multiple decisions about this belief are required, including:

- **The concept of multiplicity** as outlined in Sec. 6.1.3.
- **The relative probability of different Yukawa sectors given a specific n_D .**

A reasonable choice is to demand all solutions with a given n_D to be equally probable. The same applies to different Yukawa sectors. Unfortunately both cannot be true at the same time, because different Yukawa sectors can lead to different amounts of possible solutions. We take the approach of requiring solutions to be equally likely (given equal multiplicity and same n_D). This also implies not applying any “beauty” arguments for Yukawa sectors, e.g. in favor of coupling patterns that are equal for different fermion types.

- **The relative probabilities of different n_D .**

For our total anomaly ratio distribution, we treat the probability of all n_D values $2 \leq n_D \leq 9$ as equal. This implies at the same time that we consider any single solution for e.g. DFSZ₃ (of which there are 16) much more probable than any single solution for e.g. DFSZ₅ (of which there are 9.7×10^4). It would be equally reasonable to additionally penalize models with higher n_D , enhance the probability of models satisfying symmetry arguments, or consider all solutions equally probable. In the latter case, the final histogram would be completely dominated by DFSZ₉ due to the much larger amount of unique solutions.

The arguments above all imply a probabilistic approach to model selection, i.e. nature “selects” one of the possible realizations at random. This notion itself may be subject to critique, but in absence of any decisive underlying physical argument singling out any specific model we deem it to be satisfactory. In Sec. 6.1.4, we outline theoretical arguments that might challenge this view.

Even under the assumption of probabilistic model selection, we acknowledge that any of these choices is to some extent a matter of taste. For this reason, we aim to ensure transparency by providing the raw catalog and generating code as supplementary material, allowing the reader to make their own choices without being constrained by our own preferences.

6.2.3 Results for $n_D = 4 - 7$

Now that we have outlined our assumptions leading to a statistical treatment, we can move on to analyzing the higher numbers of Higgs doublets for which we investigate

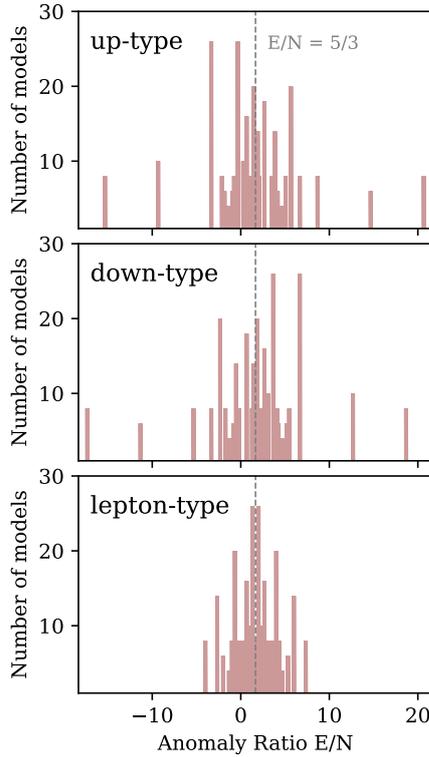


Figure 15: Anomaly ratio distributions for DFSZ-type models with 4 Higgs doublets. Two Higgs couple to the fermions specified in the panels with the other two Higgs covering the remaining two fermion types invariant with respect to fermion generation. For example, the Yukawa sectors $[u1, u1, u3, d1, d1, d1, e1, e1, e1]$, $[u1, u2, u1, d1, d1, d1, e1, e1, e1]$ and $[u1, u2, u2, d1, d1, d1, e1, e1, e1]$ are all equivalent and have anomaly ratio distributions as shown in the top panel. Note that the up-type and down-type cases are mirrored around $5/3$.

multiple Yukawa sectors. However, we will start by focusing on the DFSZ₄ to DFSZ₇ models as we can explicitly calculate all possible solutions for these models.

Fig. 15 provides an overview of the anomaly ratio distributions for DFSZ4 models, grouped by the various Yukawa sectors. Each histogram represents all models of the corresponding coupling type, with the explicit symmetry breaking potential V_{eb} composed of k or more $HHSS$ -terms and $3 - k$ or fewer $HHHH$ -terms. The results are independent of fermion generation, as the construction of Higgs charges as well as (6.10) treat all generations equally.

Histograms for Yukawa sectors with a special coupling to a lepton are symmetric around $5/3$, while the histograms for up- and down-type special couplings are mirrored around $5/3$. The reason for this symmetry is that for every n_D , we consider all possible Yukawa sectors. As can be seen from (6.10), every solution has a corresponding one obtained by applying the transformation

$$\chi_{u_i} \rightarrow -\chi_{d_i}, \quad \chi_{d_i} \rightarrow -\chi_{u_i}. \quad (6.15)$$

This is because up-type and down-type quarks are treated equally in the construction,

except for the sign of their hypercharges. Therefore, for example, all solutions for the Yukawa sector $[u1, u1, u3, d1, d1, d1, e1, e1, e1]$ have a corresponding solution in the Yukawa sector $[u1, u1, u1, d1, d1, d3, e1, e1, e1]$ under the transformation (6.15). Solutions related through this transformation have anomaly ratios that are related by

$$\frac{\mathcal{E}}{\mathcal{N}} \rightarrow \frac{10}{3} - \frac{\mathcal{E}}{\mathcal{N}}, \quad (6.16)$$

which corresponds to a mirror symmetry around $5/3$.

By summing up all nine histograms from Fig. 15 without preferring any Yukawa sector, we obtain the distribution shown in Fig. 17 (second row, left). Due to the symmetries of the nine contributing Yukawa sectors, the distribution is symmetric around $5/3$ as well. The largest number of models coincides with the two possible values for the DFSZ₂ model: $2/3$ and $8/3$. Both of these statements are true for $n_D \in [4, 7]$, as Fig. 17 shows (second row, third row left).

With increasing n_D , we find an increasing number of unique anomaly ratios and more extreme \mathcal{E}/\mathcal{N} values. Anomaly ratios $\mathcal{E}/\mathcal{N} = 5/3 + k$ with $k \in \mathbb{Z}$ are highly favored for $n_D \geq 5$, especially for odd k . We see this very characteristic, peaked spectrum evolving: \mathcal{E}/\mathcal{N} values with high probability tend to have their probabilities shrink with increasing n_D , whereas low probability \mathcal{E}/\mathcal{N} values behave in the opposite manner. In Fig. 17 one can most easily see this evolution at big anomaly ratios $\mathcal{E}/\mathcal{N} \gtrsim 10$.

This trend can be understood from a purely mathematical perspective. The anomaly ratio in (6.10) is a function with nine variables, where each variable's value can be thought of as being drawn from a specific distribution. In Fig. 16 we show the effect of using different distributions for the variables on the outcome of the function. A continuous, flat charge distribution of arbitrary width produces a smooth, fat-tailed \mathcal{E}/\mathcal{N} distribution. If the median of the charges is 0, the median of the distribution is at $5/3$ (Fig. 16 top three rows). Allowing only positive values for the charges shifts the distribution to higher values, with a median of $8/3$ and makes $\mathcal{E}/\mathcal{N} < 0$ impossible (Fig. 16 bottom row). The fewer distinct input values for the charges are used, the more peaked the anomaly ratio structure becomes, i.e. anomaly ratios with high relative probability see their likelihood increased and vice versa. This also leads to fewer possible unique \mathcal{E}/\mathcal{N} values. The continuum limit with its vanishing skewness and positive kurtosis can be approximated in analytic form via a Pearson type VII distribution [185],

$$p\left(\frac{\mathcal{E}}{\mathcal{N}}\right) = \frac{1}{\alpha \text{B}(m - \frac{1}{2}, \frac{1}{2})} \left[1 + \left(\frac{\frac{\mathcal{E}}{\mathcal{N}} - \lambda}{\alpha} \right)^2 \right]^{-m}, \quad (6.17)$$

with reasonable fit parameters $\lambda = 5/3$, $\alpha = 7/4$ and $m = 1$, and Beta function B with $\text{B}(1/2, 1/2) = \pi$.

Following these insights from a mathematical perspective, it can be understood that the histograms for larger n_D should be smoother, considering that there are more unique solutions (Tab. 3). Note, however, that this effect neglects the influence of

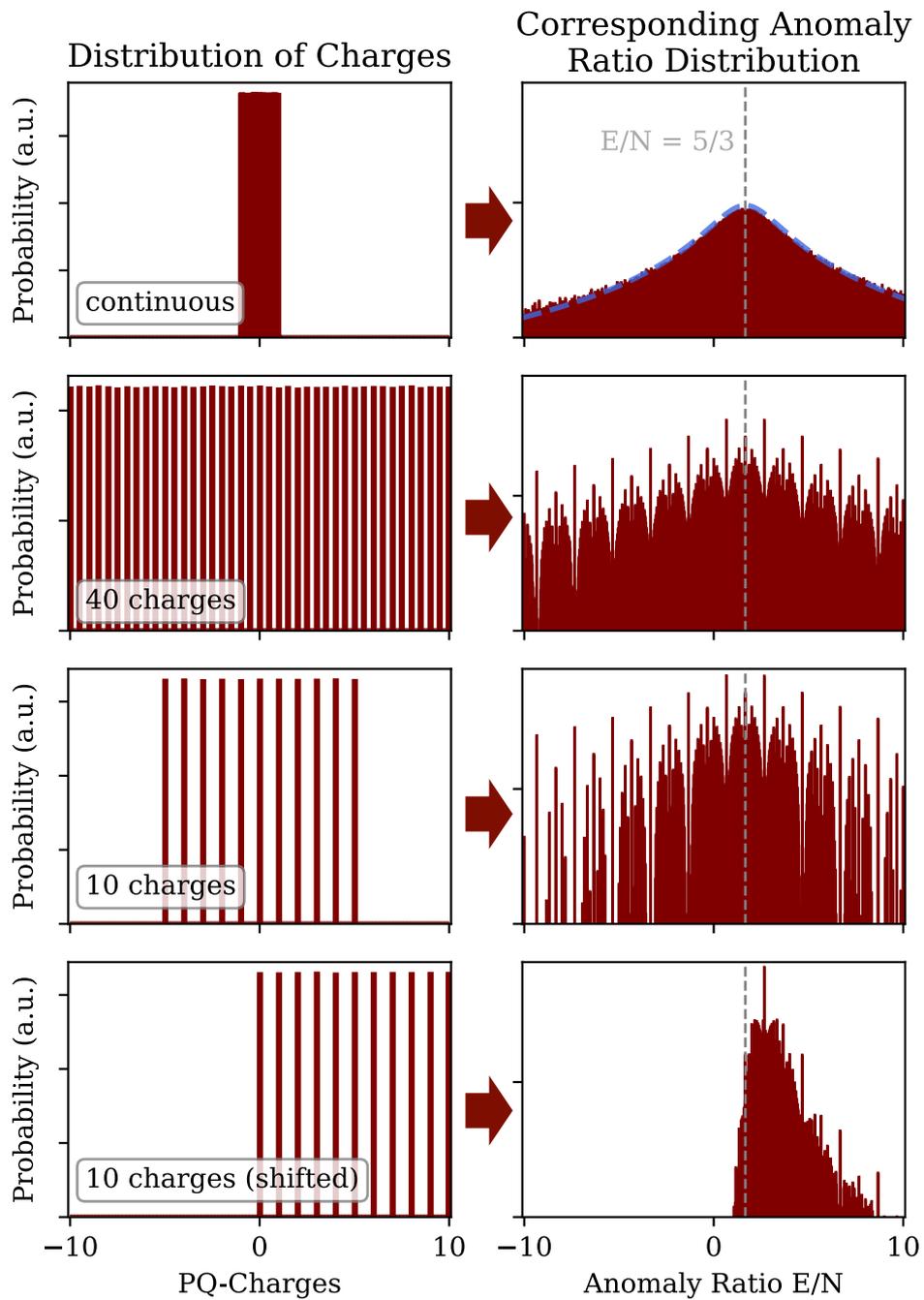


Figure 16: Influence of drawing charges from different distributions on the resulting anomaly ratio distribution, using (6.10). More unique charges lead to a smoother anomaly ratio distribution, irrespective of their distribution. Charge distributions centered around 0 produce anomaly ratio distributions centred around $5/3$. The dashed blue line in the top right panel denotes the fit presented in (6.17).

choosing different probabilities for different solutions. Non-uniform probabilities reduce the effective number of different solutions.¹⁰ Using our approach of adding all possible potential terms for one solution of charges, leads to more comparable probabilities for the charges than if we had separately considered all potentials with the minimal amount of terms to fix the PQ charges (we call these minimal potentials). Therefore the effect of non-uniform charge probabilities is clearly subdominant for DFSZ₅ to DFSZ₇. We expect this to still be the case even for DFSZ₈ and DFSZ₉.

6.2.4 Extrapolation to $n_D > 7$

Our procedure can in principle be applied to any number of Higgs doublets. However, for larger values of n_D , it becomes computationally intensive as it requires solving a vast number of linear systems of LSEs. Let us get an idea of the magnitude by estimating the number of all possible terms for step 2 with an arbitrary n_D . Since the number of possible bilinears is $n_B = \binom{n_D}{2}$ plus their Hermitian conjugate, there are $2n_B$ terms of the form $HHSS$. This results in $(2n_B)^2$ possible quadrilinears, which can be written as a matrix

$$\begin{array}{c} HH \quad (HH)^\dagger \\ HH \quad \left(\begin{array}{cc} A & B \\ C & D \end{array} \right), \\ (HH)^\dagger \end{array} \quad (6.18)$$

where A denotes the submatrix formed by all terms of the form $HHHH$, B the submatrix formed by $HH(HH)^\dagger$, and so on. However, as in the DFSZ₃ example there are several equal terms in this matrix that should not be counted. First of all, the whole matrix is symmetric. Secondly, since Hermitian conjugated terms are equal, D is completely redundant with respect to A . Lastly, B is anti-symmetric, so that the number reduces to n_B^2 quadrilinears.

From the set of all terms, we need to pick $n_D - 1$ terms with at least one being of the form $HHSS$. Therefore, we can pick between 1 and $n_D - 1$ terms of the form $HHSS$, then fill up with $HHHH$ terms, and repeat this for all possible amounts of $HHSS$ terms. The total number of possible V_{eb} can then be estimated by

$$N_{\text{tot}}(n_D) \sim \sum_{j=1}^{n_D-1} \binom{2n_B}{j} \binom{n_B^2}{n_D-1-j}, \quad (6.19)$$

which at the same time is the number of LSEs that needs to be solved. In principle, we can again perform the simplifications used for the DFSZ₃ example, i.e. setting all VEVs to one and not fixing χ_S , but regardless of these simplifications the computation time rises exponentially with n_D . While $N_{\text{tot}}(n_D = 3) = 69$ is easily manageable, for e.g. $n_D = 8$ the number of possibilities becomes $N_{\text{tot}}(n_D = 8) \approx 2 \cdot 10^{16}$. Thus, it is not

¹⁰For instance, this can be understood by considering a charge distribution with 100 unique solutions, in which 10 solutions are 1000× more probable than the other 90. The resulting \mathcal{E}/\mathcal{N} distribution will behave more as if it came only from 10 unique charges than as if it had 100.

possible for us to solve all LSEs for n_D beyond DFSZ₇ due to the high computational requirements.

A potential solution to the computationally prohibitive number of LSEs would be to sample the (minimal) potentials. However, due to step 4 in our approach this is not possible without introducing a bias: Multiple minimal potentials can lead to the same solution, and all of them belong to the same model in our approach, resulting in a lengthy potential that is likely to be found by any sampling algorithm. On the opposite side, there are also models that can just be found with one or two minimal potentials. Therefore, sampling in the space of minimal potentials results in biased sampling in the space of models.

Another way to estimate the DFSZ₈ and DFSZ₉ distributions is given by the following observations. When a sufficiently large number of theories is considered, the plots in Fig. 16 (top, right) and Fig. 16 (third row, right) can be regarded as extreme cases for the anomaly ratio distribution. By “extreme,” we do not mean the upper or lower limits of individual \mathcal{E}/\mathcal{N} bins, as we are dealing with normalized probability measures. Rather, Fig. 16 (top row, right) appears very smooth, while Fig. 16 (third row, right) is highly peaked. To quantify this criterion, we look at the cumulative sum of anomaly ratios below a certain value. We define the smoothness of an anomaly ratio distribution $f(\mathcal{E}/\mathcal{N})$ as

$$\max_x \left| \sum_{\mathcal{E}/\mathcal{N} < x} f(\mathcal{E}/\mathcal{N}) - \sum_{\mathcal{E}/\mathcal{N} < x} c(\mathcal{E}/\mathcal{N}) \right|, \quad (6.20)$$

where $c(\mathcal{E}/\mathcal{N})$ represents the continuous distribution as shown in Fig. 16 (top, right). The smoothness as defined by (6.20) measures the maximum difference in the cumulative sum of the distribution compared to the case of continuous charges and serves as a possible metric for the required task. In Sec. 6.3.1, we will see that this metric is closely related to a relevant observable. The metric ranges from one to zero (by construction for the continuous distribution), with DFSZ₃ having a value of 17%, DFSZ₄ having 5.7%, and DFSZ₇ having only 1.4%.

We aim to place a rough constraint on the smoothness of the anomaly ratio distributions for DFSZ₈ and DFSZ₉. Based on our findings for DFSZ₃ to DFSZ₇, we observe that the anomaly ratio distribution becomes smoother as the number of doublets increases. Additionally, from our analysis of biased sampling for $n_D = 6$ and $n_D = 7$, where the actual distributions can be computed, we find that sampling leads to less smooth distributions. Therefore, we expect the distribution for $n_D = 8$ or $n_D = 9$ to be smoother than their respective sampled distribution and the $n_D = 7$ distribution. In Fig. 17, we utilize the $n_D = 7$ distribution as one estimate, referred to as “limit 1”.

Another way to estimate the smoothness of the DFSZ₈ and DFSZ₉ anomaly ratio distributions is to consider that the difference in smoothness between DFSZ₆ and DFSZ₇ is smaller than that between DFSZ₅ and DFSZ₆. By extrapolating the distributions beyond $n_D = 7$ using the difference between the DFSZ₆ and DFSZ₇ distributions, we can obtain estimates of the anomaly ratio distributions that are smoother than our actual

expectation. These estimates are denoted as “limit 2” in Fig. 17, with one subtraction for $n_D = 8$ and two subtractions for $n_D = 9$. Using the metric described by (6.20), we find that the smoothness values for DFSZ₈ and DFSZ₉ are 0.73% and 0.71%, respectively.

It is important to note that these estimates are not definitive and should be viewed as rough approximations. Although the difference in probability in Fig. 17 appears significant due to the logarithmic axis, both estimates are much closer to the continuous case of Fig. 16 (top, right) than to the peaked case of Fig. 16 (third row, right), indicating that the probability mass is spread across a continuum rather than concentrated at unique \mathcal{E}/\mathcal{N} values.

6.2.5 Constructing extreme $|\mathbf{g}_{a\gamma}|$

Another problem arising from sampling potentials is that it is very unlikely to find the anomaly ratio corresponding to the maximum axion-photon coupling, which we denote as

$$\widehat{\mathcal{E}/\mathcal{N}} = \operatorname{argmax}_{\mathcal{E}/\mathcal{N}} (|\mathcal{E}/\mathcal{N} - 1.92|) . \quad (6.21)$$

This anomaly ratio, however, provides a useful bound for DFSZ-type models. For this reason, we present a procedure on how to estimate it. Before turning to this procedure though, let us note that due to the symmetry around $\mathcal{E}/\mathcal{N} = 5/3$, in absence of selection criteria, $\widehat{\mathcal{E}/\mathcal{N}}$ is not given by the largest possible anomaly ratio but the smallest.

The procedure is based on observations of the LSEs that led to $\widehat{\mathcal{E}/\mathcal{N}}$ for the smaller numbers of doublets. There, we found that any of the LSEs leading to $\widehat{\mathcal{E}/\mathcal{N}}$ of DFSZ₄ can be extended to an LSE leading to $\widehat{\mathcal{E}/\mathcal{N}}$ for DFSZ₅. The same behavior can be seen from DFSZ₅ to DFSZ₆ and in a slightly different form from DFSZ₃ to DFSZ₄. We do not have a rigorous mathematical explanation for this behavior, so applying it to larger n_D is more of an educated guess. However, it turns out to give extreme anomaly ratios, so we use it to systematically provide an estimate for $\widehat{\mathcal{E}/\mathcal{N}}$. The procedure works as follows:

1. Take all LSEs that lead to $\widehat{\mathcal{E}/\mathcal{N}}$ for a number of doublets where all solutions are known, say $n_D = 6$.
2. Add one additional Higgs doublet by specifying its Yukawa sector.
3. Adjust the orthogonality relation appearing in all LSEs.
4. Add one additional relation to the LSEs, solve them and calculate the anomaly ratio.
5. Repeat step 2 – step 4 for every possible relation and every possible Higgs doublet, specified by its Yukawa sector.
6. Extract the LSEs with the smallest anomaly ratio.

This procedure results in extreme negative anomaly ratios. However, we found for DFSZ₉ that taking the resulting LSEs and systematically exchanging one (or more if the computational time is acceptable) of the relations, new LSEs are found that give even

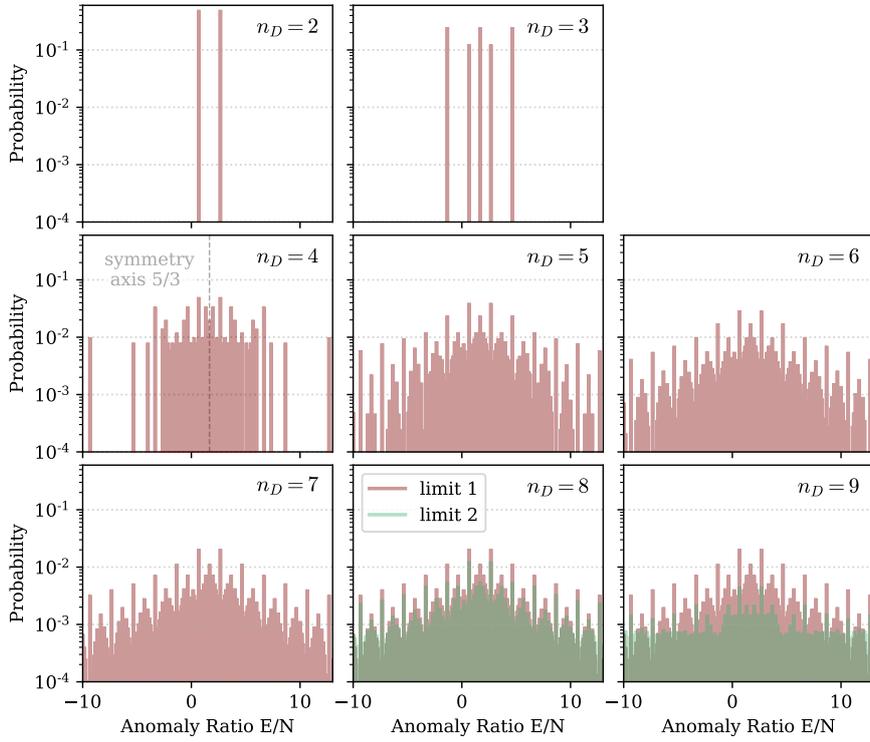


Figure 17: Anomaly ratio distributions for different numbers of Higgs doublets n_D . All histograms are symmetric around $5/3$. $n_D \geq 5$ display a characteristic peaked structure which becomes smoother with increasing n_D . DFSZ₈ and DFSZ₉ could not be fully calculated, the two semi-transparent colors denote the two estimates as discussed in the text. Note that limit 2 only slightly exceeds limit 1 at big absolute anomaly ratios for $n_D = 8$ as well as $n_D = 9$.

Table 3: Important statistics of DFSZ-type models broken down by number of Higgs doublets n_D . We include information on the model with maximal photon coupling $\widehat{\mathcal{E}/\mathcal{N}}$ from (6.21) and the percentage of models that have minimal photon coupling (photophobic, $|\mathcal{E}/\mathcal{N} - 1.92| < 0.04$). ‘x’ denotes values that could not be estimated.

n_D	$\#V_{\text{eb}}$	unique solutions	unique E/\mathcal{N} s	$\widehat{\mathcal{E}/\mathcal{N}}$	% photophobic	% $N_{\text{DW}} = 1$
2	2	2	2	$2/3$	0	0
3	54	16	5	$-4/3$	0	0
4	52614	996	83	$-52/3$	1.4	6.00
5	6.65×10^7	9.7×10^4	432	$-112/3$	1.52	6.64
6 ¹	$\lesssim 4 \times 10^9$	$> 2.19 \times 10^6$	1680	$-238/3$	1.37	5.83
7 ¹	$\lesssim 7 \times 10^{12}$	x	6256	$-466/3$	1.39	5.19
8 ²	$\lesssim 2 \times 10^{16}$	x	> 11617	$< -628/3$	x	x
9 ²	$\lesssim 1 \times 10^{20}$	x	$\gg 14122$	$< -1216/3$	x	x

¹For $n_D \geq 6$, $\#V_{\text{eb}}$ and “unique solutions” are estimates. Number of minimal potentials calculated via (6.19), many of which will be unphysical and not produce valid solutions for PQ charges. “unique solutions” gives the number of solution found in sample, for which data exists.

²For $n_D \geq 8$ we did not calculate all possible models, therefore we have no exact value neither for the number of unique \mathcal{E}/\mathcal{N} , nor for the percentage of photophobic models or models with $N_{\text{DW}} = 1$. $\widehat{\mathcal{E}/\mathcal{N}}$ was estimated as shown in Sec. 6.2.5.

smaller anomaly ratios. In DFSZ₉ for instance, the smallest anomaly ratio we construct in this way is $\mathcal{E}/\mathcal{N} = -1216/3$ and it is generated by the terms

$$\begin{aligned} & (H_{d_2}^\dagger H_{e_1})(H_{d_2}^\dagger H_{d_1}), (H_{u_1} H_{d_1})(H_{u_1} H_{d_2}), (H_{u_3}^\dagger H_{u_1})(H_{u_3}^\dagger H_{u_2}), (H_{e_1}^\dagger H_{d_1})(H_{e_1}^\dagger H_{e_2}), \\ & (H_{e_2}^\dagger H_{d_1})(H_{e_2}^\dagger H_{e_3}), (H_{u_2} H_{d_3})(H_{u_1}^\dagger H_{u_2}), (H_{d_3} H_{u_1})(H_{d_1}^\dagger H_{d_3}), (H_{d_1} H_{u_1})S^\dagger S^\dagger. \end{aligned} \quad (6.22)$$

6.2.6 Comparison with KSVZ-Type Models

In [170], the authors add all anomaly ratios of phenomenologically allowed KSVZ-type models, regardless of the number of quarks and including additive quark representations. Consequently, a model with $N_Q = 9$ quarks, of which there are more than 10^5 , is considered equally probable as a model with only $N_Q = 1$, of which there are only 15. The final distribution is dominated by $7 \lesssim N_Q \lesssim 21$. If we were to apply a similar weighting to our DFSZ-type models, the resulting distribution would be indistinguishable from the DFSZ₉ case. In Sec. 6.2.2, we argued for an alternative approach in which all separate values of n_D are equally probable. Since the raw data from [170] was provided, we can weight their KSVZ data to give equal probability to all values of N_Q , so the 15 models with $N_Q = 1$ combined are equally likely as all $\sim 10^5$ models with $N_Q = 9$ combined. Using this weighting, their data can be compared with our DFSZ results on a fair basis and we show the result in Fig. 18.

Nonetheless, we want to stress that there remain differences: The authors of [170] were able to apply selection criteria, which significantly reduced the number of viable models. As we did not find stringent selection criteria, our catalog contains the complete range of models rather than a preferred subset. Therefore, any comparison between the two model types or a combined axion band should not be considered conclusive, but rather as including all current selection criteria. Also note that in our case, a model with a higher value of n_D is always less probable than a model with a lower value, which can be seen as an appropriate penalty for introducing more degrees of freedom to the model. However, this is not the case for the KSVZ data from [170] when our weighting scheme is applied. As an example, for $N_Q = 28$ they discovered only 510 preferred models, which is considerably lower than for $N_Q = 9$. Consequently, a single model with $N_Q = 28$ is more likely than a single model with $N_Q = 9$ in our weighting scheme.

The impact of giving equal weights to all n_D values can be observed in Fig.18, particularly in the vicinity of $\mathcal{E}/\mathcal{N} = 5/3$. The DFSZ₂ and DFSZ₃ models, which have high relative probabilities, are responsible for the significantly elevated probabilities of the five \mathcal{E}/\mathcal{N} values (compare Fig.17). The effect of the two DFSZ₈ and DFSZ₉ estimates only becomes substantial at low absolute probabilities and above $|\mathcal{E}/\mathcal{N}| \gtrsim 20$.

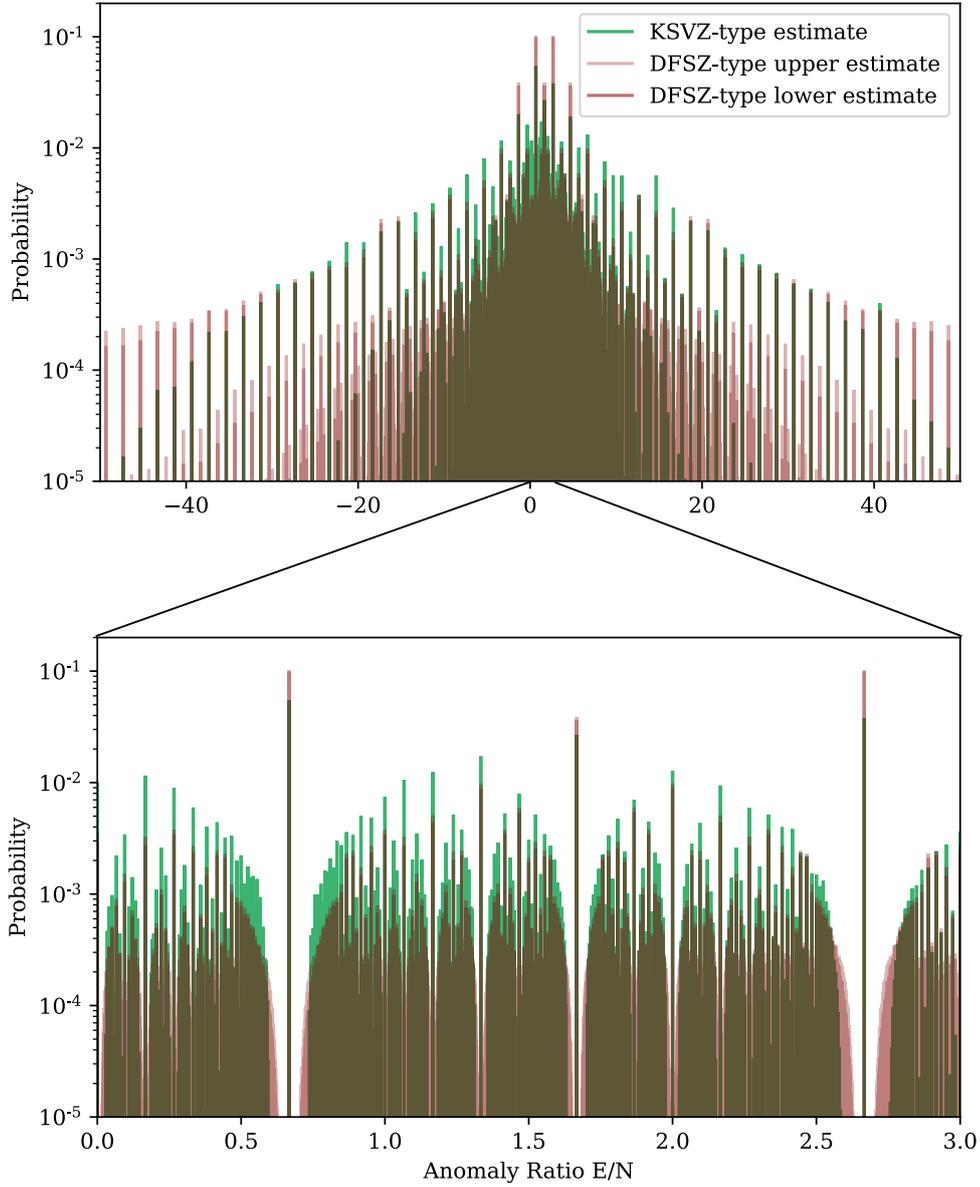


Figure 18: Comparison between anomaly ratio distributions for KSVZ-type and DFSZ-type axion models. The KSVZ-type estimate of [170] includes all phenomenologically allowed models, adding and subtracting quark representations, and assumes every model to be equally likely. Our DFSZ-type results include calculations for DFSZ₂ to DFSZ₇ and estimates for DFSZ₈ and DFSZ₉, giving equal probability to each n_D . For DFSZ-type, the different shades denote maximum and minimum for each bin under the two limits for DFSZ₈ and DFSZ₉ described above.

We find the KSVZ results display a similar peaked structure to the DFSZ case, which is only noticeable in a finely binned histogram. In fact, for \mathcal{E}/\mathcal{N} values not including DFSZ₂ and DFSZ₃, the DFSZ-type histograms are less peaked than the KSVZ-type histograms, with decreased probability at moderately large $|\mathcal{E}/\mathcal{N}|$ and significantly increased probability for $|\mathcal{E}/\mathcal{N}| \gtrsim 40$. However, this pattern does not extend to the maximum axion-photon coupling. We find a maximal $|g_{a\gamma}|$ at $\widehat{\mathcal{E}/\mathcal{N}} > -1216/3$, which is comparable to the KSVZ case for $N_Q \leq 9$ before any phenomenological constraints are imposed ($\widehat{\mathcal{E}/\mathcal{N}} = -1312/3$).

Regarding models with the smallest axion-photon couplings, we define photophobic models in the same way as [170], i.e. by selecting those with an anomaly ratio \mathcal{E}/\mathcal{N} that is compatible with vanishing $g_{a\gamma}$ within 1 sigma theoretical uncertainty. Tab. 3 shows that there is no clear trend towards a higher or lower percentage of photophobic models with increasing n_D . As discussed in Sec. 6.2.3, the distribution of anomaly ratios becomes smoother with increasing n_D , with peaks becoming less pronounced and ratios with low probability becoming more likely. The absence of a clear trend suggests that the photophobic region lies in the middle between these two extremes. In summary, the percentage of photophobic models we find for DFSZ-type models with $n_D \leq 7$ is similar to the KSVZ case.

The probability near the highest peaks is notably reduced in both the KSVZ- and DFSZ-type model probability distributions (Fig. 18, bottom). However, this effect is less pronounced in DFSZ-type models compared to KSVZ-type models because, as mentioned earlier, DFSZ-type models are less peaked when DFSZ₂ and DFSZ₃ are excluded.

Closer inspection reveals that the distribution of KSVZ-type models is not symmetric around $5/3$, unlike the DFSZ-type models which have both the mean and median anomaly ratios at exactly $5/3$. In particular, the mean and median anomaly ratios for KSVZ-type models are $\mathcal{E}/\mathcal{N}|_{\text{mean}} = 1.43$ and $\mathcal{E}/\mathcal{N}|_{\text{median}} = 1.30$, respectively. These values do not change even when considering only the subset of $N_{\text{DW}} = 1$ models. The deviation from $5/3$ in the KSVZ-type models of [170] could potentially arise from the phenomenological selection criteria imposed in their study.

6.3 Implications for Axion Searches

We have so far derived probability mass functions for the anomaly ratio from theoretical assumptions for different DFSZ-type models. Let us now discuss what implications these results have for axion experiments.

6.3.1 $C_{a\gamma}$ Cumulative Distribution Function

In order to comprehend the consequences for axion searches, it is necessary to map the \mathcal{E}/\mathcal{N} distributions into $g_{a\gamma}$ space using (6.1). To eliminate the dependence on the axion mass, we present our findings in terms of the dimensionless quantity $|C_{a\gamma}|$, as defined in (6.1).

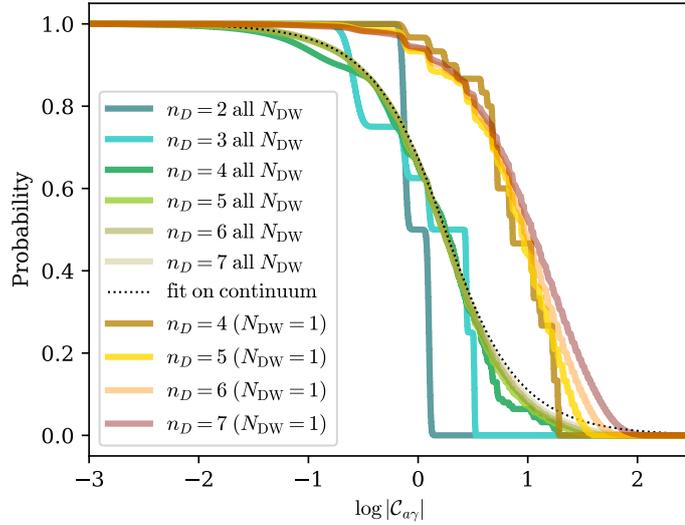


Figure 19: Cumulative probability of models with $|\mathcal{C}_{a\gamma}|$ higher than the indicated values. The plot includes DFSZ-type models of arbitrary domain wall number N_{DW} with DFSZ₃ to DFSZ₇ as well as $N_{\text{DW}} = 1$ models for DFSZ₄ to DFSZ₇ (for smaller n_D no $N_{\text{DW}} = 1$ models exist). The CDFs become smoother with increasing n_D , with DFSZ₆ and DFSZ₇ already being almost indistinguishable. $N_{\text{DW}} = 1$ models have systematically larger $|\mathcal{C}_{a\gamma}|$, shifted by almost one order of magnitude. The dashed line indicates the analytic fit on the continuum limit from (6.23).

Typically, two-sided axion bands centered around the region of maximal probability are provided in this situation [106, 170, 186–188]. However, experiments are usually sensitive to all axion-photon couplings above a certain threshold $|\mathcal{C}_{a\gamma}|^{\text{min}}$, and therefore it is relevant to also report a one-sided limit that must be achieved in order to detect, e.g. 68% of all DFSZ-type models for a given axion mass. For this purpose, we utilize a cumulative distribution function (CDF) plotted against $|\mathcal{C}_{a\gamma}|$, which represents the combined theoretical prior probability of models with $|\mathcal{C}_{a\gamma}|(\text{model}) > |\mathcal{C}_{a\gamma}|^{\text{min}}$.

Since we are treating the anomaly ratio as a random variable derived from a distribution that we aim to determine, we must handle the next-to-leading order QCD corrections $\mathcal{C}_{a\gamma\gamma}^{(0)}$ in the same manner. We model its uncertainty as a normal distribution $\mathcal{N}(1.92, 0.04)$ with a mean of 1.92 and standard deviation of 0.04. This approach smooths out the steps in the CDF that result from high probability \mathcal{E}/\mathcal{N} values, particularly for anomaly ratios near the mean value of $\mathcal{C}_{a\gamma\gamma}^{(0)}$.

6.3.2 Experimental Constraints

The anomaly ratios of the DFSZ₂ and DFSZ₃ models are still the most prominent feature in the probability distribution for all possible DFSZ models. However, reaching sensitivity to these models may be either not necessary or not sufficient because only one value of the anomaly ratio is realized in nature.

Fig. 19 displays the resulting theoretical prior probability of DFSZ-type axion models with $|\mathcal{C}_{a\gamma}|$ above a certain threshold. We group results by possible values of n_D . Let us

first discuss the “all N_{DW} ”-case, in which the domain wall number is not imposed as selection criterion. DFSZ₃ models have zero probability above $\log |\mathcal{C}_{a\gamma}| \gtrsim 0.5$. Should an axion be found above this value that can be determined to be of DFSZ-type, this would imply the existence of $n_D > 3$ Higgs doublets. The impact of the prominent peaks of maximal probability between $\mathcal{E}/\mathcal{N} = -4/3$ or $\mathcal{E}/\mathcal{N} = 14/3$ on the cumulative probability is only minor for theories with $n_D \geq 5$. As the CDFs for DFSZ₆ and DFSZ₇ are already almost indistinguishable, we refrain from additionally plotting our estimates for higher n_D . In fact, the relative difference on $|\mathcal{C}_{a\gamma}|$ exclusion limits between our two ways of estimating the smoothness of the DFSZ₈ and DFSZ₉ distributions is below the percent level. For the purpose of $|\mathcal{C}_{a\gamma}|$ exclusion limits, the two estimates are therefore virtually equivalent. In the following, we use limit 2. An analytical estimate for the cumulative probability distribution can be obtained by utilizing the analytic anomaly ratio fit given in (6.17). This fit can be translated to $|\mathcal{C}_{a\gamma}|$ as,

$$p(|\mathcal{C}_{a\gamma}|) = 1 - \frac{\tan^{-1} \left[\frac{4}{7} \left(|\mathcal{C}_{a\gamma}| - \frac{19}{75} \right) \right] + \tan^{-1} \left[\frac{4}{7} \left(|\mathcal{C}_{a\gamma}| + \frac{19}{75} \right) \right]}{\pi}, \quad (6.23)$$

which is plotted as a dotted line in Fig. 19.

We can compare the complete set of DFSZ₄ to DFSZ₇ models with their subsets that have $N_{\text{DW}} = 1$. The latter models are preferred in the post-inflationary scenario due to cosmological energy density arguments (see Sec. 6.1.4). The $N_{\text{DW}} = 1$ models have $|\mathcal{C}_{a\gamma}|$ values that are almost one order of magnitude higher on average than the complete set. Therefore, they are more easily detectable. Like in the previous case, higher values of n_D tend to have smoother distributions. Hence, we can reasonably introduce our two estimates. Here, the difference between the two estimates with respect to the $|\mathcal{C}_{a\gamma}|$ limits is also below the percent level. Therefore, we use limit 2, the extrapolation estimate, in the following analysis.

Now let us compare the complete set of DFSZ₄ to DFSZ₇ models with their subsets that have $N_{\text{DW}} = 1$. The latter models could be considered preferred in the post-inflationary scenario since they avoid the domain wall problem. The $N_{\text{DW}} = 1$ models have $|\mathcal{C}_{a\gamma}|$ values almost an order of magnitude higher on average than the full set and are therefore much easier to detect. Similar to the “all N_{DW} ”-case, higher n_D values tend to have smoother distributions. It therefore seems reasonable to analogously introduce our two estimates where the difference with respect to the $|\mathcal{C}_{a\gamma}|$ limits between the two is again below the percent level. We again use limit 2 in the following.

Fig. 20 compares the CDFs for DFSZ- and KSVZ-type models. In general, both types are very similar for all values of $|\mathcal{C}_{a\gamma}|$, only for DFSZ-type models with $N_{\text{DW}} = 1$ a significant fraction of models is above $\log |\mathcal{C}_{a\gamma}| \gtrsim 1.5$. The lines of $\mathcal{E}/\mathcal{N} = 2/3$ and $\mathcal{E}/\mathcal{N} = 8/3$ are clearly visible for both, DFSZ-type and KSVZ-type models. The relative difference between the 68% limits of KSVZ- and DFSZ-type axions is only $\sim 3\%$ and $\sim 19\%$ for the 95% limits with the DFSZ limit being higher in the latter case. Taking into account possible effects from diverging theory assumptions, this relative difference can be seen as negligible.

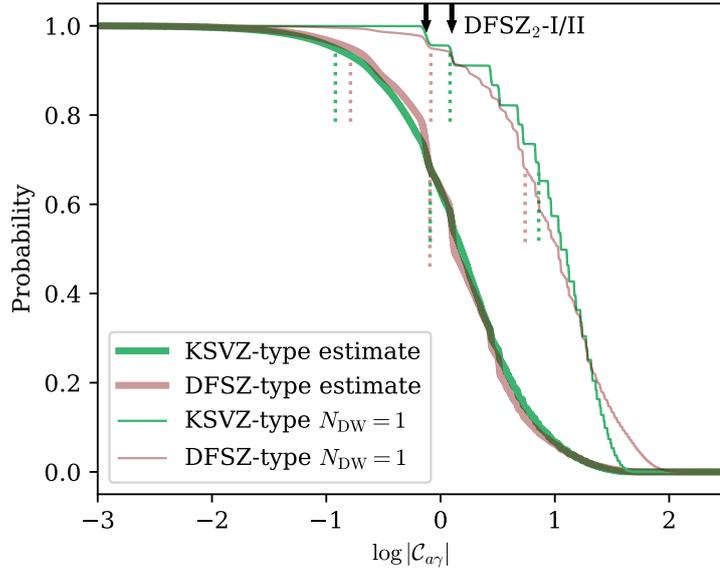


Figure 20: Cumulative probability of models with $|C_{a\gamma}|$ higher than the indicated values for the complete set of DFSZ-type and KSVZ-type models as well as for models with $N_{\text{DW}} = 1$ specifically (thin lines). One sided 95% and 68% limits for both cases are given with coloured vertical dotted lines. The arrows at the top indicate the location of DFSZ₂-I and DFSZ₂-II.

Although exploring various theoretical assumptions is not the focus of this study, it is worth noting that different assumptions can only alter the relative importance of the prominent peaks of DFSZ₂ and DFSZ₃ in the entire set of models. For instance, if we adopt a different definition of multiplicity based on minimal potentials, it leads to a substantial increase in their probability mass, but it does not shift the overall cumulative probability towards higher or lower $|C_{a\gamma}|$ values. Therefore, any variation in theoretical assumptions, except for the model selection criteria, would lie somewhere between the cumulative probabilities of the $n_D = 2$ and the continuum cases.

Translating these limits to $g_{a\gamma}$ over a range of axion masses, we obtain Fig. 22 (top). If an experimental exclusion limit touches the 68% line, it excludes 68% of the probability mass in the model space for a given mass range, assuming the aforementioned assumptions. An experiment reaching down to the 95% line would be sensitive to 95% of the probability for all models within the targeted mass range. We provide these and the 99% limit for DFSZ-type as well as KSVZ-type models and the combined case of $N_{\text{DW}} = 1$. The maximum $\widehat{\mathcal{E}/\mathcal{N}}$ value found for DFSZ₉ is also included in black. Additionally to being excluded by experiments for a large fraction of the m_a range, this model may likely also be subject to phenomenological constraints (see Sec. 6.1.4).

Our study has made it possible to determine one-sided limits or axion bands for the combined KSVZ and DFSZ scenario, assuming that both types of axions are equally probable. The associated probability density function (PDF) for the combined “all N_{DW} ”-case is shown in Fig. 21. In logarithmic space with the chosen binning, the distribution appears to be approximately Gaussian, with several notable peaks at $\mathcal{E}/\mathcal{N} =$

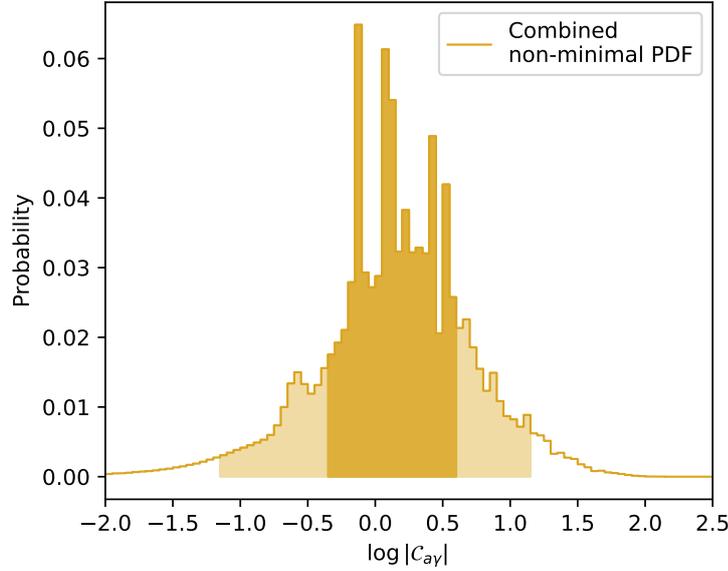


Figure 21: Probability density in $\log |C_{a\gamma}|$ -space of the combined DFSZ-type and KSVZ-type “all N_{DW} ”-case. Central 68% and 95% regions used for Fig. 22 are indicated in different shades of yellow. Note that the underlying distribution is discrete and any illustration will in part depend on the binning chosen.

$5/3, 8/3, 2/3, 14/3$ and $-4/3$ (from left to right). It should be kept in mind, however, that the true underlying distribution is comprised out of a multitude of delta peaks and thus fundamentally discrete. Central 68% and 95% bands from this distribution are used in Fig. 22 (bottom) together with a previous estimate for the same band from [106]. Previous work was either limited to very few extensions of DFSZ-type [106] or the KSVZ case [170]. Despite several caveats, such as the imprecise prediction for DFSZ₈ and DFSZ₉ models or the lack of selection criteria in the DFSZ case, it is still valuable to provide usable data of typical limits and bands for a variety of scenarios. An overview can be found in Tab. 4 and more detailed information is available on the website “zenodo” together with the model catalog (see end of Sec. 6.4 for links).

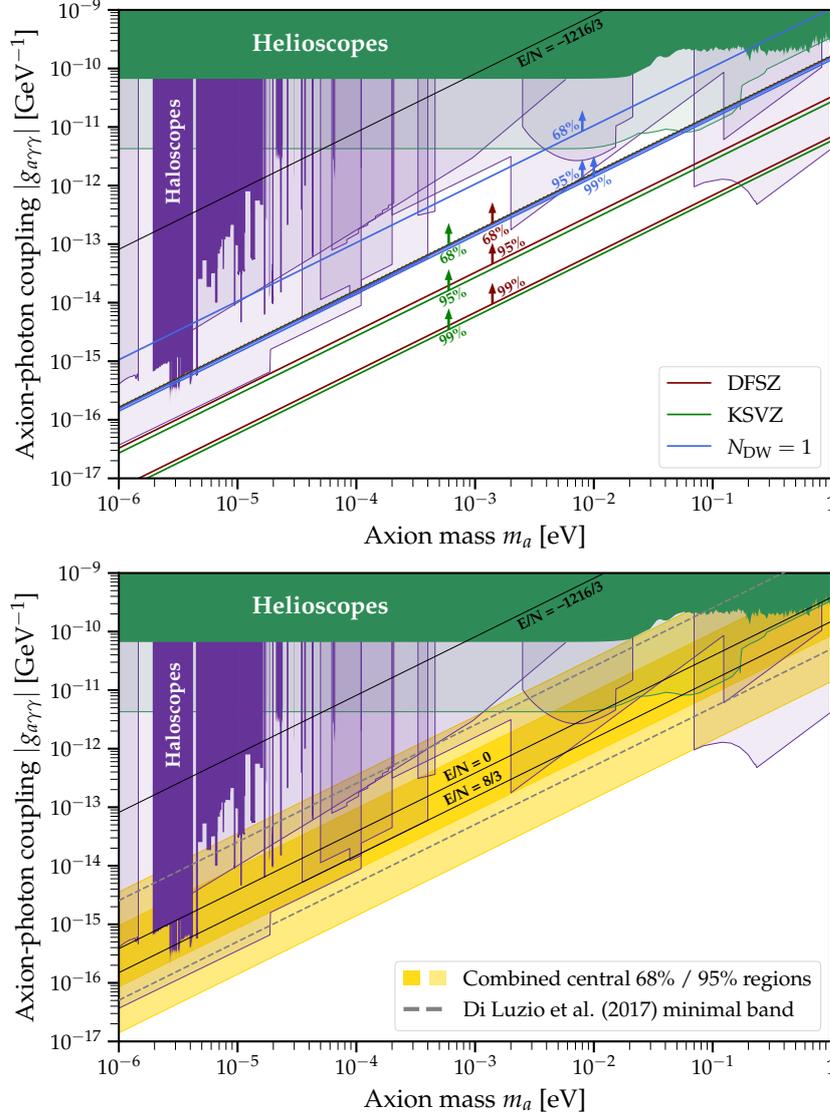


Figure 22: Top: 68%, 95% and 99% limits for the complete preferred KSVZ case [170], our complete DFSZ case (using extrapolation for DFSZ_8 and DFSZ_9) as well as the combined $N_{\text{DW}} = 1$ case. The highest DFSZ-type coupling found is shown in black ($\mathcal{E}/\mathcal{N} = -1216/3$). $\text{DFSZ}_{2\text{-I}}$ and $\text{DFSZ}_{2\text{-II}}$ roughly coincide with the 68% limit of the complete DFSZ case and the 95% limit of the $N_{\text{DW}} = 1$ case, respectively. **Bottom:** Central 68% and 95% regions for the case combining all preferred KSVZ and all DFSZ models together with a previous band from Di Luzio et al. [106] for comparison. We show helioscope limits and forecasts [189–191] in green as well as limits and forecasts from various haloscope experiments [192–235] in purple. For reference we also show the $\mathcal{E}/\mathcal{N} = 0$ and $\mathcal{E}/\mathcal{N} = 8/3$ lines in black. All experimental limits shown here are Frequentist in nature and should therefore only be seen as a rough comparison with respect to our Bayesian prior results. For the full cumulative probability from which the three limits shown in the top panel are taken, see Fig. 20, and for the combined probability density from which the band in the bottom panel is derived, see Fig. 21. (Plotted with tools by O’Hare [169].)

Table 4: $|\mathcal{C}_{a\gamma}|$ lower prior limits for selected combinations of models. All limits shown are one sided, so a central 68% band can be constructed with values given for 16% and 84% and similar for 95%. The label “KSVZ” denotes re-weighted results from [170], while “DFSZ” denotes results from this paper. Both are combined with equal probability for the case “Combined”. The combination only considering models with domain wall number of unity is shown as “ $N_{\text{DW}} = 1$ ”.

$ \mathcal{C}_{a\gamma} $	68% band		95% band				
	one sided limit	68%	95%	16%	84%	2.5%	97.5%
KSVZ		0.833	0.135	4.684	0.427	15.274	0.068
DFSZ		0.809	0.164	4.529	0.482	19.272	0.08
Combined		0.819	0.148	4.593	0.451	17.285	0.074
$N_{\text{DW}} = 1$		5.294	0.769	22.773	2.733	36.729	0.731

6.4 Summary and Discussion

This study aimed to determine the axion-photon coupling for a large number of DFSZ-type models and provide limits required to be reached in order to be sensitive to a certain fraction of the probability mass of these models. We began by clearly defining the DFSZ-type models that under consideration and discussed various selection criteria, including the absence of FCNCs and the domain wall problem, to identify the preferred models. In contrast to the KSVZ-type models, where all selection criteria follow from cosmological bounds on the new fermions, for DFSZ-type axions we did not find criteria with a sufficient level of generality, only desirable features.

Next, we developed a recipe to calculate the axion-photon coupling for DFSZ-type models. This recipe relies on the fact that the PQ charges in DFSZ-type models are not arbitrary but are constrained by linear consistency and phenomenology relations. Consequently, the process of DFSZ-type model building reduces to solving LSEs. By systematically going through each Yukawa sector and solving all possible LSEs, we were able to derive all possible anomaly ratios for up to seven Higgs doublets, and hence all possible axion-photon couplings.

Furthermore, by essentially counting how many models lead to a certain anomaly ratio and establishing relative probabilities of these models, we were able to assign probabilities to each anomaly ratio. For this counting of models, we considered as a model the Lagrangian that arises by combining different potentials resulting in the same set of PQ charges and by adding the Yukawa couplings compatible with the resulting set of PQ charges. In this way, we took into account the general consensus that all terms allowed by symmetry must be included and avoided overcounting. The obtained anomaly ratio distributions have their median at $\mathcal{E}/\mathcal{N} = 5/3$, their maximum values at $\mathcal{E}/\mathcal{N} = 2/3$ and $\mathcal{E}/\mathcal{N} = 8/3$, and a shape that is reminiscent of the distributions found in KSVZ-type models. These findings can be attributed to the discrete nature of the underlying sets of PQ charges, which can be thought of as charge distributions with uniform probability and symmetry around zero.

Although our recipe can be applied to DFSZ-type models with any number of Higgs

doublets, the computational time required becomes impractical for eight or more doublets. Attempting to estimate these cases through simple sampling of potential terms results in a significant bias. Therefore, we developed estimations for $n_D > 7$ based on the expected smoothness of the distributions. We have also used an incremental construction procedure to identify a maximal anomaly ratio that exceeds previous estimates in [106] by more than a factor of two.

Regarding the axion experimental program, the anomaly ratio distributions confirm the experimental importance of the values dictated by the minimal DFSZ models, namely $\mathcal{E}/\mathcal{N} = 2/3$ and $\mathcal{E}/\mathcal{N} = 8/3$, since they are also favored for every number of Higgs doublets (with the exception of $n_D = 3$, where we imposed the Weinberg-Glashow-Paschos condition). However, the distributions also indicate that there is a considerable amount of viable parameter space above and below these values. Therefore, a non-observation at these favored values is insufficient to declare the axion excluded, while an observation above these values would suggest the presence of at least two additional Higgs doublets from the DFSZ-type perspective. The statistical analysis also reveals that KSVZ and DFSZ models set comparable requirements for experiment sensitivity.

Intriguingly, for $n_D \geq 4$ we found a subset of models with domain wall number $N_{\text{DW}} = 1$. These avoid the domain wall problem and thus make DFSZ-type models theoretically more viable in post-inflationary scenarios. This subset even displays a significantly enhanced axion-photon coupling compared to the minimal scenarios for both invisible axion classes, hence rendering these models easier to probe.

In conclusion, we want to emphasize that our analysis is valuable for axion searches regardless of the statistical interpretation. By providing all possible \mathcal{E}/\mathcal{N} values for up to seven doublets and a full catalog for up to five doublets, one can perform hypothesis testing with the compatible models in the case of a detection. Since all \mathcal{E}/\mathcal{N} values for preferred KSVZ models are also known, this could be used for model comparison between these two classes. Thus, whether or not one adopts a statistical viewpoint, our analysis presents another step forward in the understanding and mapping of the landscape of axion models.¹¹

¹¹Our generating code can be found at <https://github.com/jhbdiehl/DFSZforest>, the model catalogs and axion limits/ bands together with usable Bayesian theory priors at <https://doi.org/10.5281/zenodo.7656939>.

CHAPTER SEVEN

CONCLUSION AND OUTLOOK

In this thesis, we investigated the axions potential role in going beyond the SM by studying various aspects of its landscape. Taking this avenue of investigation is motivated by the axion providing a common theme among various questions of high-energy physics. Thus, the axion holds significant potential for playing a key role in beyond the SM physics.

We began by providing a detailed introduction into the strong CP problem and the axion solution. The aim of this introduction is twofold: First, to establish the standard languages employed in axion physics, namely chiral perturbation theory and instantons. Secondly, to set up a connection between these languages and less well-known aspects of the axion through the exploration of the Schwinger model and the higher-form formalism. In this way, the most suitable formalism can be chosen to solve a particular problem. As part of this introduction, we outlined the low-energy EFT of the axion and the various benchmark UV models that we used throughout the rest of the thesis.

We then placed the axion within an FLRW framework and discussed the model dependency of the arising cosmological bounds. In particular, we showed that there is a viable part of the parameter space where Dvali's early relaxation mechanism can successfully be incorporated into the KSVZ and DFSZ axion models by only using the modification of the running coupling during inflation. We achieved this via much heavier quarks resulting from a Higgs portal between the inflaton and our Higgs doublet or, alternatively, higher order operators of those fields. The heavier quarks change the running of the strong coupling constant and thus result in a larger value of the QCD scale. Thus, QCD can become strong in both models for $H_I \lesssim 10^5$ GeV, vastly expanding the viable parameter space. We also illustrated how in both models the physics that makes QCD strong during inflation does not displace the axions minimum, so that the efficiency of the relaxation is controlled by parameters of the theory and the number of

inflationary e-folds. Intriguingly, we identified a parameter space compatible with all cosmological constraints when the axion abundance is dominated by inflationary quantum and post-inflationary thermal fluctuations. Unfortunately, this region is not realized in the minimal models.

Next, we pointed out that the strong CP problem, as a quantum gravitational consistency problem, is not special to QCD but arises in every YM group. We then considered hidden sectors with YM subgroups and identified general phenomenological consequences of including one axion per YM group as required to solve “each strong CP problem”. In particular, we carried out computations for two simple models, i.e. a pure YM sector and N exact Standard Model copies, and identified phenomenological consequences such as: a cosmological constraint on the dark confinement scale, an upper bound on N , as well as no necessary tightening of the bound on the inflationary Hubble scale arising from isocurvature perturbations. Furthermore, we discuss the phenomenon of axion stars collectively made from axions of different dark sectors, as well as the possibility of axion kinetic mixing and its implications for axion states and stability. Axions cannot collectively form structures with a $1/\sqrt{N}$ suppressed mass spectrum due to their repulsive self-interaction. Kinetic mixing results in $N - 1$ degenerate states that behave as the ordinary axion, and one special state that is slightly lighter by a factor of $1/\sqrt{2}$ and couples weaker by $1/N^{1/2}$.

Lastly, we presented a systematic way to calculate the axion-photon coupling for a large number of non-minimal DFSZ models and studied the resulting distributions. We found that the values dictated by the minimal models are favored for every number of Higgs doublets and that most of the possible axion-photon couplings lie in the vicinity of these values. This confirms the experimental importance of these values. The distributions also reveal that there exists a significant amount of viable parameter space above and below the indicated values. Therefore, a non-observation at these preferred values is not enough to exclude the axion. On the other hand, detection of the axion above these values would imply the existence of at least three Higgs doublets from a DFSZ-type perspective. The statistical analysis also indicates similar requirements in terms of experimental sensitivity for the KSVZ and DFSZ models. It is also worth mentioning that we found a large number of non-minimal DFSZ models with a domain wall number of unity, thus avoiding the domain wall problem. In addition, all these models have an enhanced axion-photon coupling compared to the minimal DFSZ models and are thus easier to detect.

Let us now provide an outlook to future theoretical studies, which build upon the projects presented in this thesis. Regarding the cosmological bound on the axion scale, let us discuss which models could potentially offer a modification of the running so that the narrow region of interest becomes viable. By playing around with simple modifications of the minimal models, such as varying the quark mass of the KSVZ model or adding more heavy quarks, we found that $\Lambda_{\text{QCD}}^{\text{Inf}}$ cannot dramatically exceed 10^5 GeV. An interesting possibility, however, is given by embedding of the PQ symmetry into GUTs with an intermediate breaking. For instance, this could be achieved by $\text{SO}(10)$, which breaks into

SU(5) or the Pati-Salam group. Since the running is steeper in these groups, such an implementation would not only modify the initial condition of the running but would also modify the running itself. This could potentially result in a larger $\Lambda_{\text{QCD}}^{\text{Inf}}$, thus allowing for the implementation of a GUT or M_{P} scale axion without the need of moduli fields. Furthermore, let us make the note that our analysis is automatically valid for the two-form implementation of the axion due to duality below the Peccei-Quinn scale and a minor influence of the KSVZ and DFSZ fields on the running.

The two-form implementation of the axion, in general, offers many interesting questions. Since there is no renormalizable UV-completion in terms of a complex scalar field for a two-form, it is likely that a UV completion has its origin in a theory of quantum gravity. This could be studied in string theory where the Kalb-Ramond two-form is a generic prediction. Alternatively, since the two-form is sourced by fundamental strings, it is interesting to study its dynamics in the presence of confining strings in QCD. It would also be interesting to study how the PQ scale can be lowered, since for $f_a \sim M_{\text{P}}$ the axion avoids any current experimental endeavors. One example to achieve this is by introducing a large number of species as these lower the fundamental scale of gravity.

Moving on to the second project, the situation with multiple axions closely resembles the axiverse proposed by string theory. However, in our case, the axions arise from quantum gravitational consistency in theories with numerous hidden YM groups. Given that hidden YM groups and multiple axions are predictions of string theory, exploring this connection in the context of string theory is a valuable pursuit. Furthermore, all of our axions originate from a single underlying physical principle, namely the S-matrix exclusion of de Sitter. Pushing this idea further, if the dark energy is realized by quintessence, this would suggest the existence of a non-trivial vacuum structure that needs to be eliminated by this scalar field. The requirement of the corresponding particle to behave as dark energy instead of dark matter results in a dark confinement scale below $\sim 10^{-3}$ eV. A potential late phase transition in the universe could have intriguing consequences as pointed out in [236].

In regard of the DFSZ project, there are several potential avenues for extending our analysis. One possibility is to investigate models with a right-handed neutrino or hybrid KSVZ/DFSZ models, which featuring additional Higgs singlets, Higgs doublets, and heavy quarks. Although the resulting anomaly ratios would be different, we anticipate that the shape of the distributions and the exclusion lines for axion mass versus axion-photon coupling would be similar. Nevertheless, it would still be important to perform an explicit analysis to confirm this expectation. Another area of interest is the exploration of other axion couplings, such as the axion-electron coupling. Although the higher-dimensional parameter space due to the VEVs of the Higgs doublets poses a challenge, the perturbative range of the top and bottom Yukawa couplings [237] or other phenomenological constraints could be used to derive reasonable limits. Furthermore, finding a better estimate for the anomaly distribution of eight or more doublets, or developing an unbiased method for computing it, would be desirable.

It is also interesting to mention that with the identification and classification of both

large classes of invisible axion models, a comparison with other classes of axion models, like the two-form implementation of the axion, is possible. There, in contrast to the ordinary invisible axion models, no explicit PQ violating processes are possible there. Hence, should the axion be detected through its axion-photon coupling, we can dramatically limit the number of invisible axion models and search for PQ violating processes in these. If a PQ violating process is experimentally found, this would completely eliminate the two-form axion.

The research presented in this thesis provides a comprehensive examination of various aspects of the axion landscape, shedding light on its potential role in answering fundamental questions in high-energy physics. Our findings suggest that the axion represents a promising avenue for further exploration, with multiple branches of physics intersecting at this unique particle. As such, the research summarized here sets the stage for future investigations.

APPENDIX

A

CONVENTIONS AND FUNDAMENTALS

A.1 Units

Throughout this thesis Lorentz-Heaviside natural units are used, i.e. $c = \hbar = k_B = 1$, which gives all quantities the dimension of energy to some power. In these units the electron is chosen to have rationalized charge $-e$, so that the fine-structure constant is $\alpha = e^2/4\pi \simeq 1/137$.

It is advantageous to interchange between natural units and centimeter-gram-second (CGS) units. In the following table we show the values for this conversion,

Natural Unit	CGS Value
eV	10^5 cm^{-1}
eV	10^{15} s^{-1}
eV	10^4 K
eV	10^{-33} g

A.2 Indices and Metric Signature

- Latin indices (i,j,k, ...) generally run over the three spatial coordinate labels taken 1, 2, 3. Greek indices (μ, ν, ρ, \dots) generally run over the four spacetime coordinate labels 0, 1, 2, 3.
- The Levi-Civita tensor $\epsilon^{\mu,\nu,\rho,\sigma}$ is defined as the totally antisymmetric tensor with $\epsilon^{0123} = 1$.

- For the spacetime metric $\eta_{\mu\nu}$ the “particle physics”, “west coast” or “mostly minus” convention is chosen, i.e., $\eta_{\mu\nu} \equiv \text{diag}(+, -, -, -)$.
- The γ -matrices are defined by the anti-commutator relation of the Clifford algebra, $\{\gamma_\mu, \gamma_\nu\} = 2\eta_{\mu\nu} \mathbb{1}_4$.

Many authors use different conventions, so it can be very helpful to know how to translate between conventions to get those minus signs right. The biggest confusion enters when fermions are involved, since the metric appears in the Clifford algebra and thus affects the gamma matrices. Demanding the Clifford algebra to be independent of the metric convention and the spatial component of the momentum to be defined as in non-relativistic quantum mechanics (see Appendix A.3), the two conventions can be related by

$$\begin{aligned}\eta_{\mu\nu}^{\text{wc}} &= -\eta_{\mu\nu}^{\text{ec}} \\ \gamma_\mu^{\text{wc}} &= -i\gamma_\mu^{\text{ec}} .\end{aligned}$$

Physical quantities will never depend on these conventions, but some equations will change, e.g.

$$\begin{array}{lll} \text{West coast metric} & \partial_\mu F^{\mu\nu} = J^\mu , & (\square + m^2)\Phi = 0 , \quad (i\gamma^\mu \partial_\mu - m)\psi = 0 , \\ \text{East coast metric} & \partial_\mu F^{\mu\nu} = -J^\mu , & (\square - m^2)\Phi = 0 , \quad (\gamma^\mu \partial_\mu + m)\psi = 0 . \end{array}$$

A.3 2π and Fourier Transformation

The origin of all 2π is the relation

$$\delta(x) = \int_{-\infty}^{\infty} dp e^{\pm 2\pi i p x} . \quad (\text{A.1})$$

We choose the convention with negative sign and rescale p in order to remove the 2π factor from the exponent. Therefore, the Fourier transformation of an arbitrary function $f(x)$ is given by

$$f(x) = \int_{-\infty}^{\infty} \frac{d^4 p}{(2\pi)^4} \tilde{f}(p) e^{-ipx} \quad \leftrightarrow \quad \tilde{f}(p) = \int_{-\infty}^{\infty} d^4 x f(x) e^{ipx} . \quad (\text{A.2})$$

This Fourier transform convention is consistent with

$$p_\mu = i\partial_\mu , \quad (\text{A.3})$$

which has spatial components $\vec{p} = -i\vec{\nabla}$, as in quantum mechanics (recalling that contravariant four-vectors have only positive components, while for the four gradient the covariant components are all positive due to the definition $\partial_\mu = d/dx^\mu$).

A.4 Gauge Theories

As usual, the kinetic term of scalars or fermions can be made gauge invariant by the introduction of a gauge covariant derivative. For a field ϕ carrying U(1) charge Q and transforming as

$$\phi \longrightarrow e^{iQ\alpha} \phi, \quad (\text{A.4})$$

we define the gauge covariant derivative as

$$D_\mu \phi = (\partial_\mu - igQA_\mu)\phi, \quad (\text{A.5})$$

where the gauge field A_μ transforms as

$$A_\mu \longrightarrow A_\mu + \frac{1}{g}\partial_\mu \alpha. \quad (\text{A.6})$$

In non-Abelian YM Theories the fundamental gauge group G is a non-Abelian Lie group. For the kinetic term to transform covariantly, a covariant derivative is introduced in the standard way, i.e

$$D_\mu = \partial_\mu - igA_\mu. \quad (\text{A.7})$$

The gauge field A_μ is now Lie algebra valued, meaning that $A_\mu = A_\mu^a T^a$, where T^a are the generators of the algebra and $a = 1, \dots, \dim G$. These satisfy

$$[T^a, T^b] = if^{abc}T^c, \quad \text{Tr } T_R^a T_R^b = \frac{T(R)}{2}\delta^{ab}, \quad (\text{A.8})$$

where f^{abc} are the structure constants and the representation dependent number $T(R)$ is called the Dynkin Index. We fix it to be 1 in the fundamental representation and 2 in the adjoint. The corresponding field strength is defined as

$$\begin{aligned} G_{\mu\nu} &= i[D_\mu, D_\nu] = \partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu] \\ &= (\partial_\mu A_\nu^a - \partial_\nu A_\mu^a + gf^{abc}A_\mu^b A_\nu^c)T^a = G_{\mu\nu}^a T^a. \end{aligned} \quad (\text{A.9})$$

Under gauge transformations the gauge field and the field strength transform as

$$\begin{aligned} A_\mu &\longrightarrow UA_\mu U^{-1} + \frac{i}{g}U\partial_\mu U^{-1} = A_\mu + \frac{1}{g}\partial_\mu \alpha - i[A_\mu, \alpha] \\ &= (A_\mu^a + \frac{1}{g}\partial_\mu \alpha^a - f^{abc}\alpha^b A_\mu^c)T^a, \end{aligned} \quad (\text{A.10})$$

$$G_{\mu\nu} \longrightarrow G_{\mu\nu} - i[G_{\mu\nu}, \alpha] = (G_{\mu\nu}^a - f^{abc}\alpha^b G_{\mu\nu}^c)T^a. \quad (\text{A.11})$$

The locally gauge invariant Lagrangian is then

$$\mathcal{L} = -\frac{1}{4}\text{tr}[G_{\mu\nu}G^{\mu\nu}] = -\frac{1}{4}G_{\mu\nu}^a G^{\mu\nu,a}, \quad (\text{A.12})$$

where g is the analog of e in QCD. Note that there are two frequently used conventions

in the literature. In order to change to the other convention, we simply need to make the following two rescalings,

$$A_\mu \longrightarrow \frac{1}{g} A_\mu, \quad G_{\mu\nu} \longrightarrow \frac{1}{g} G_{\mu\nu}. \quad (\text{A.13})$$

A.5 The Standard Model

The Standard Model is the QFT based on the gauge group $G_{\text{SM}} = \text{SU}(3)_c \times \text{SU}(2)_L \times \text{U}(1)_Y$ that contains all known particles (see Fig. 23 for an overview of the particles, the corresponding charges, and the couplings). The Lagrangian can be split into four sectors:

QCD sector

$$\mathcal{L}_{\text{QCD}} = \sum_q \bar{q}_i i \gamma^\mu (D_\mu)^{ij} q_j - \frac{1}{4} G_{\mu\nu}^a G^{\mu\nu,a}, \quad (\text{A.14})$$

where

$$(D_\mu)q = (\partial_\mu - i g_s A_\mu^a T^a)q. \quad (\text{A.15})$$

Electroweak sector:

$$\mathcal{L}_{\text{EW}} = \sum_\Psi \bar{\Psi} i \gamma^\mu (D_\mu) \Psi - \frac{1}{4} W_{\mu\nu}^a W^{\mu\nu,a} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}, \quad (\text{A.16})$$

where

$$D_\mu \Psi = (\partial_\mu \mathbb{1} - i g W_\mu^a \tau^a - i g' Y B_\mu) \Psi. \quad (\text{A.17})$$

Note that for $\text{SU}(2)_L$ singlets the coupling with the corresponding gauge bosons in the covariant derivative vanishes.

Higgs sector:

$$\mathcal{L}_{\text{Higgs}} = (D_\mu H)^\dagger (D^\mu H) - \frac{\lambda}{2} (H^\dagger H - v^2)^2, \quad (\text{A.18})$$

where the covariant derivative is the same as in the EW sector since the Higgs doublet carries no color charge.

Yukawa sector:

$$\mathcal{L}_{\text{Yukawa}} = -y^{(u)} \bar{Q}_L \tilde{H} u_R - y^{(d)} \bar{Q}_L H d_R - y^{(e)} \bar{L}_L H e_R + \text{h.c.} \quad (\text{A.19})$$

where $\tilde{H} = i\sigma_2 H$.

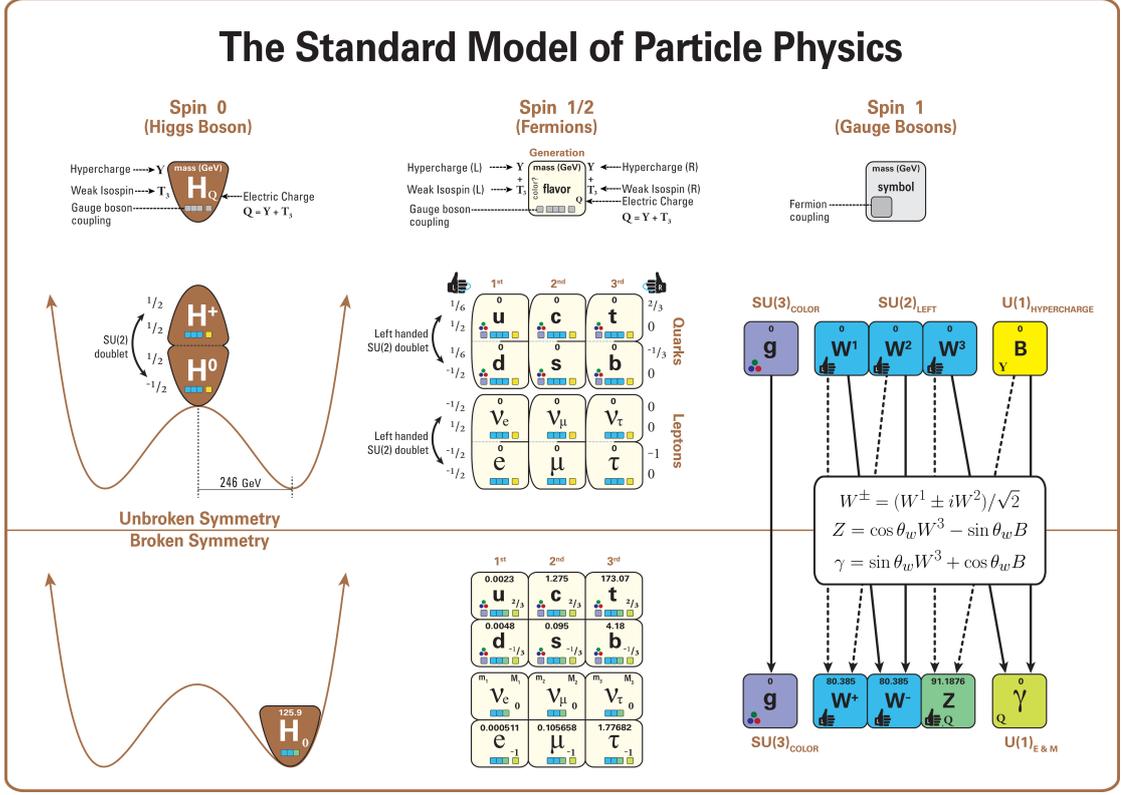


Figure 23: The Standard Model of particle Physics with its particles, the corresponding charges, and the couplings for both, unbroken and spontaneous broken symmetry.

A.6 Higher-Forms

The structure of Maxwell's theory can be generalized to tensor fields of higher rank, called higher form fields. A p -form is a differential form with a corresponding antisymmetric covariant tensor field of rank p . Acting on a p -form with the external derivative d , a $(p+1)$ -form can be constructed. In the context of Maxwell theories, the p -form A plays the role of the higher rank analog of the gauge field, while the corresponding $(p+1)$ -form F can be thought of as its field strength. Choosing a local coordinate system, the components of the field strength can be expressed as

$$F_{\mu_1 \dots \mu_{p+1}} = (dA)_{\mu_1 \dots \mu_{p+1}} \equiv \partial_{[\mu_1} A_{\mu_2 \dots \mu_{p+1}]} . \quad (\text{A.20})$$

The square brackets in this relation indicate antisymmetrization with respect to all indices inside the brackets. There are different conventions for the canonical normalization of p -forms. In this thesis, we normalize the Lagrangian of a p -form field as

$$\mathcal{L} = -\frac{1}{2(p+1)!} (F_{\mu_1 \dots \mu_{p+1}})^2 + \frac{m^2}{2p!} (A_{\mu_1 \dots \mu_p} + \partial_{[\mu_1} a_{\mu_2 \dots \mu_p]})^2 + J^{\mu_1 \dots \mu_p} A_{\mu_1 \dots \mu_p} . \quad (\text{A.21})$$

where $J^{\mu_1 \dots \mu_p}$ represents an antisymmetric tensor current. In order to dualize Lagrangian (A.21) in n dimensions, it is useful to impose the Bianchi-identity via a Lagrange multiplier,

i.e.

$$\mathcal{L}_{\text{dual}} = \frac{1}{(p+1)!} \partial_{\mu_{p+2}} \xi_{\mu_{p+3} \dots \mu_n} \epsilon^{\mu_1 \dots \mu_n} F_{\mu_1 \dots \mu_{p+1}}, \quad (\text{A.22})$$

In this thesis, we have chosen the convention for the contraction of epsilon-tensors such that

$$\epsilon_{i_1 \dots i_k i_{k+1} \dots i_n} \epsilon^{i_1 \dots i_k j_{k+1} \dots j_n} = -k! \delta_{[i_{k+1}}^{j_{k+1}} \dots \delta_{i_n]}^{j_n}, \quad (\text{A.23})$$

so that the Euler-Lagrange equations of a general p -form take a form that is similar to that of Maxwell's theory,

$$\partial_\nu F^{\nu \mu_1 \dots \mu_p} = J^{\mu_1 \dots \mu_p}. \quad (\text{A.24})$$

In principle, one can either work with (coordinate independent) forms or with the corresponding fields. In this thesis we merely use the higher form fields and use the terms “higher forms” and “higher form fields” interchangeably (even though this might be highly confusing in a strict mathematical sense).

A.7 Euclidean Formulation of YM Theories

To go to Euclidean spacetime, we need to go to imaginary time, which induces a similar change on the timelike components of other tensors. For this purpose, we make the following definitions,

$$\begin{aligned} x^i &= \hat{x}_i & A^i &= -\hat{A}_i & D^i &= -\hat{D}_i & G_{ij}^a &= \hat{G}_{ij}^a, \\ x_0 &= -i\hat{x}_4 & A_0 &= i\hat{A}_4 & D_0 &= i\hat{D}_4 & G_{0j}^a &= i\hat{G}_{4j}^a, \end{aligned} \quad (\text{A.25})$$

where

$$\hat{D}_\mu = \frac{\partial}{\partial \hat{x}^\mu} - ig \hat{A}_\mu^a T^a, \quad \hat{G}_{\mu\nu}^a = \frac{\partial}{\partial \hat{x}^\mu} \hat{A}_\nu^a - \frac{\partial}{\partial \hat{x}^\nu} \hat{A}_\mu^a + gf^{abc} \hat{A}_\mu^b \hat{A}_\nu^c. \quad (\text{A.26})$$

The Greek letters now denote indices running from 1 to 4 and we chose to work with merely lower indices, since there is no difference between upper and lower indices in Euclidean spacetime. Furthermore, the Levi-Civita tensor is defined such that $\epsilon_{1234} = 1$. Similarly, we define the fermions as

$$\begin{aligned} \Psi &= \hat{\Psi} & \gamma_0 &= \hat{\gamma}_4, \\ \bar{\Psi} &= -i\hat{\bar{\Psi}} & \gamma^i &= i\hat{\gamma}_i. \end{aligned} \quad (\text{A.27})$$

All put together in the Yang-Mills action yields

$$iS = -\hat{S}, \quad (\text{A.28})$$

with

$$\begin{aligned}
S &= \int d^4x \left\{ -\frac{1}{4} G_{\mu\nu}^a G^{a\mu\nu} + \bar{\Psi}(i\gamma^\mu D_\mu - m)\Psi + \theta \frac{g^2}{32\pi^2} G_{\mu\nu}^a \tilde{G}^{a\mu\nu} \right\}, \\
\hat{S} &= \int d^4\hat{x} \left\{ +\frac{1}{4} \hat{G}_{\mu\nu}^a \hat{G}^{a\mu\nu} + \hat{\Psi}(-i\hat{\gamma}_\mu \hat{D}_\mu - im)\hat{\Psi} + i\theta \frac{g^2}{32\pi^2} \hat{G}_{\mu\nu}^a \hat{\tilde{G}}^{a\mu\nu} \right\}. \quad (\text{A.29})
\end{aligned}$$

When working in the Euclidean formulation in the main part of this thesis, we simply drop the hats on the Euclidean quantities. Instead, we mark the action and the integral measure with an “E” to indicate that all quantities are in the Euclidean formulation as defined in this appendix. We will also write upper indices to mark the contraction, even though there is no difference between covariant and contravariant tensors in Euclidean space-time.

LIST OF FIGURES

1	A Simple system to understand the role of topology in YM theories. Left: A particle moving along an upright circle and subject to a constant gravitational force. Right: Potential of this system with n denoting the minima. Classically, the particle oscillates around one of the vacua. Quantum mechanically, the vacuum is a superposition of all the minima, inducing non-trivial effects.	16
2	't Hooft vertex for $n_F = 3$. In its essence, this non-perturbative multi-fermion vertex is a manifestation of the Atiya-Singer index theorem, arising from zero-modes of the Dirac equation in the instanton background. . . .	26
3	The vacuum polarization diagram that generates the polarization operator $\Pi^{\mu\nu} \propto (g^{\mu\nu} - k^\mu k^\nu / k^2)$	28
4	Cosmological bound on the axion scale f_a depending on the inflationary Hubble scale H_I . The thick red line represents $f_a = \max\{T_{GH}, T_{RH}\}$ in the case of inefficient reheating, separating the pre-inflationary and the post-inflationary scenarios. The case of very efficient reheating with $\epsilon_{\text{eff}} \approx 0$ is depicted in orange to show the influence of the thermalization efficiency on the bound. We neglect the contribution from the decay of topological defects in the post-inflationary scenario due to large uncertainties in their determination [119–121].	57
5	Schematic representation of the freeze-out captured by the adiabatic-impulse approximation.	60
6	Axion window with the bounds coming from isocurvature perturbations (darker blue), black hole super-radiance (purple), interactions in stars (yellow), primordial gravitational waves (cyan) and overproduction (light blue or brown). See [55] for further details on each constraint.	66
7	Running of α_s during inflation in the KSVZ model with the inflationary Higgs VEV chosen to be $v_H^{\text{Inf}} \sim M_P$ and the initial condition $\alpha_s^{\text{Inf}}(M_P) = \alpha_s(M_P)$. The different colors indicate the parts with a different number n_f of active quarks. We see that the QCD scale is much larger compared to its value in the late universe.	70

- 8 Values of the QCD scale for various values of the inflationary Higgs VEV v_H^{Inf} and the corresponding linear fit. The red points have $\alpha_s^{\text{Inf}}(m_Z^{\text{Inf}}) = \alpha_s(m_Z)$ as initial condition for the running while the blue points have $\alpha_s^{\text{Inf}}(\Lambda_{\text{GUT}}) = \alpha_s(\Lambda_{\text{GUT}})$. In both the KSVZ quark mass was fixed to be of order of $f_a \sim \Lambda_{\text{GUT}}$ 71
- 9 Top: Constraints on the axion for very efficient reheating, $\epsilon_{\text{eff}} \sim 1$. The blue line depicts the upper bound for our minimal scenarios. Hence, the region enclosed by the red and the blue line defines the parameter space where the axion dynamically relaxes to the minimum during inflation and later makes up all the dark matter. Bottom: Same as top but with $\epsilon_{\text{eff}} \sim 10^{-1}$ 76
- 10 The allowed parameter space with different initial misalignment angles θ_i^{ini} according to (5.17). The values on the thick lines correspond to the case where the dark matter is entirely composed of axions. The dashed line presents the species bound, meaning that along this line the gravitational cutoff M_* and f_a coincide. 86
- 11 The fraction of our axion energy density with respect to the total axion energy density in the case when the axions make up all the dark matter. For larger f_a our axion is dominating, while for lower f_a the axions from dark sectors dominate. 87
- 12 The allowed region for a single dark, pure YM sector with confinement scale Λ_{conf} , axion scale f_{a_2} , and no intersector interactions. The red line represents the perturbative unitarity bound $\Lambda_{\text{conf}} \lesssim f_{a_2}$. The region with $\theta_2^{\text{ini}} \sim 1$ is favored by minimality, since in this scenario there is no need for additional physics that results in a small misalignment angle. 88
- 13 Loop induced by the operators in (5.53). Each vertex contributes the effective coupling $\Delta p^2/M^2$, but for maximum momentum transfers of order f_a the effective coupling becomes ϵ 98
- 14 Viable axion-photon couplings for the light axion state (yellow band), which arises from equal kinetic mixing of axions from N sectors. Current bounds and predicted sensitivities of future experiments are depicted different colors. The parameter space for $N \lesssim 10$ will be probed by future experiments. The two stars represent a situation when future experiments measure one axion (the upper star) and where we expect the second axion (the lower star). The yellow band is not to be confused with the band from e.g. [3], which includes the set of DFSZ-type or KSVZ-type models. This plot was created with the help of the software [169]. 101

- 15 Anomaly ratio distributions for DFSZ-type models with 4 Higgs doublets. Two Higgs couple to the fermions specified in the panels with the other two Higgs covering the remaining two fermion types invariant with respect to fermion generation. For example, the Yukawa sectors $[u1, u1, u3, d1, d1, d1, e1, e1, e1]$, $[u1, u2, u1, d1, d1, d1, e1, e1, e1]$ and $[u1, u2, u2, d1, d1, d1, e1, e1, e1]$ are all equivalent and have anomaly ratio distributions as shown in the top panel. Note that the up-type and down-type cases are mirrored around $5/3$ 115
- 16 Influence of drawing charges from different distributions on the resulting anomaly ratio distribution, using (6.10). More unique charges lead to a smoother anomaly ratio distribution, irrespective of their distribution. Charge distributions centered around 0 produce anomaly ratio distributions centred around $5/3$. The dashed blue line in the top right panel denotes the fit presented in (6.17). 117
- 17 Anomaly ratio distributions for different numbers of Higgs doublets n_D . All histograms are symmetric around $5/3$. $n_D \geq 5$ display a characteristic peaked structure which becomes smoother with increasing n_D . DFSZ₈ and DFSZ₉ could not be fully calculated, the two semi-transparent colors denote the two estimates as discussed in the text. Note that limit 2 only slightly exceeds limit 1 at big absolute anomaly ratios for $n_D = 8$ as well as $n_D = 9$ 121
- 18 Comparison between anomaly ratio distributions for KSVZ-type and DFSZ-type axion models. The KSVZ-type estimate of [170] includes all phenomenologically allowed models, adding and subtracting quark representations, and assumes every model to be equally likely. Our DFSZ-type results include calculations for DFSZ₂ to DFSZ₇ and estimates for DFSZ₈ and DFSZ₉, giving equal probability to each n_D . For DFSZ-type, the different shades denote maximum and minimum for each bin under the two limits for DFSZ₈ and DFSZ₉ described above. 123
- 19 Cumulative probability of models with $|C_{a\gamma}|$ higher than the indicated values. The plot includes DFSZ-type models of arbitrary domain wall number N_{DW} with DFSZ₃ to DFSZ₇ as well as $N_{\text{DW}} = 1$ models for DFSZ₄ to DFSZ₇ (for smaller n_D no $N_{\text{DW}} = 1$ models exist). The CDFs become smoother with increasing n_D , with DFSZ₆ and DFSZ₇ already being almost indistinguishable. $N_{\text{DW}} = 1$ models have systematically larger $|C_{a\gamma}|$, shifted by almost one order of magnitude. The dashed line indicates the analytic fit on the continuum limit from (6.23). 125
- 20 Cumulative probability of models with $|C_{a\gamma}|$ higher than the indicated values for the complete set of DFSZ-type and KSVZ-type models as well as for models with $N_{\text{DW}} = 1$ specifically (thin lines). One sided 95% and 68% limits for both cases are given with coloured vertical dotted lines. The arrows at the top indicate the location of DFSZ₂-I and DFSZ₂-II. . . 127

-
- 21 Probability density in $\log |\mathcal{C}_{a\gamma}|$ -space of the combined DFSZ-type and KSVZ-type “all N_{DW} ”-case. Central 68% and 95% regions used for Fig. 22 are indicated in different shades of yellow. Note that the underlying distribution is discrete and any illustration will in part depend on the binning chosen. 128
- 22 **Top:** 68%, 95% and 99% limits for the complete preferred KSVZ case [170], our complete DFSZ case (using extrapolation for DFSZ₈ and DFSZ₉) as well as the combined $N_{\text{DW}} = 1$ case. The highest DFSZ-type coupling found is shown in black ($\mathcal{E}/\mathcal{N} = -1216/3$). DFSZ₂-I and DFSZ₂-II roughly coincide with the 68% limit of the complete DFSZ case and the 95% limit of the $N_{\text{DW}} = 1$ case, respectively. **Bottom:** Central 68% and 95% regions for the case combining all preferred KSVZ and all DFSZ models together with a previous band from Di Luzio et al. [106] for comparison. We show helioscope limits and forecasts [189–191] in green as well as limits and forecasts from various haloscope experiments [192–235] in purple. For reference we also show the $\mathcal{E}/\mathcal{N} = 0$ and $\mathcal{E}/\mathcal{N} = 8/3$ lines in black. All experimental limits shown here are Frequentist in nature and should therefore only be seen as a rough comparison with respect to our Bayesian prior results. For the full cumulative probability from which the three limits shown in the top panel are taken, see Fig. 20, and for the combined probability density from which the band in the bottom panel is derived, see Fig. 21. (Plotted with tools by O’Hare [169].) 129
- 23 The Standard Model of particle Physics with its particles, the corresponding charges, and the couplings for both, unbroken and spontaneous broken symmetry. 141

LIST OF TABLES

1	Resulting PQ conditions from quadrilinears, constructed from corresponding bilinears. The lower triangle ('-') are not counted because the order of the bilinears does not matter. The terms 'x' are not to be counted because they are Hermitian to a term that has already been counted and the potential by definition has to include all Hermitian conjugated terms. Terms with 'o' produce only trivial conditions. We are left with 9 distinct quadrilinears, which produce 6 unique conditions.	112
2	All possible solutions for PQ charges of Higgs doublets in terms of χ_S (top) and anomaly ratios (bottom) for the $n_D = 3$ Yukawa sector under consideration. The potential should not produce the same condition twice ('x'), nor does the order of the conditions matter ('-'). 'o' denotes combinations of conditions that do not have a solution. Infinite solutions arise when $\mathcal{N} = 0$	112
3	Important statistics of DFSZ-type models broken down by number of Higgs doublets n_D . We include information on the model with maximal photon coupling $\widehat{\mathcal{E}/\mathcal{N}}$ from (6.21) and the percentage of models that have minimal photon coupling (photophobic, $ \mathcal{E}/\mathcal{N} - 1.92 < 0.04$). 'x' denotes values that could not be estimated.	121
4	$ \mathcal{C}_{a\gamma} $ lower prior limits for selected combinations of models. All limits shown are one sided, so a central 68% band can be constructed with values given for 16% and 84% and similar for 95%. The label "KSVZ" denotes re-weighted results from [170], while "DFSZ" denotes results from this paper. Both are combined with equal probability for the case "Combined". The combination only considering models with domain wall number of unity is shown as " $\mathcal{N}_{\text{DW}} = 1$ ".	130

BIBLIOGRAPHY

- [1] G. Dvali, E. Koutsangelas, and F. Kuhnel, “Compact Dark Matter Objects via N Dark Sectors,” *Phys. Rev. D* **101** (2020) 083533, [arXiv:1911.13281 \[astro-ph.CO\]](#).
- [2] E. Koutsangelas, “Removing the cosmological bound on the axion scale in the Kim-Shifman-Vainshtein-Zakharov and Dine-Fischler-Srednicki-Zhitnitsky models,” *Phys. Rev. D* **107** no. 9, (2023) 095009, [arXiv:2212.07822 \[hep-ph\]](#).
- [3] J. Diehl and E. Koutsangelas, “Dine-Fischler-Srednicki-Zhitnitsky-type axions and where to find them,” *Phys. Rev. D* **107** no. 9, (2023) 095020, [arXiv:2302.04667 \[hep-ph\]](#).
- [4] M. Ettengruber and E. Koutsangelas, in preparation.
- [5] A. Alexandre, G. Dvali, and E. Koutsangelas, in preparation.
- [6] S. Weinberg, *The Quantum Theory of Fields*, vol. 1. Cambridge University Press, 1995.
- [7] *The Rise of the Standard Model: A History of Particle Physics from 1964 to 1979*. Cambridge University Press, 1997.
- [8] S. Weinberg, “The Making of the standard model,” *Eur. Phys. J. C* **34** (2004) 5–13, [arXiv:hep-ph/0401010](#).
- [9] C. M. Will, “The Confrontation between General Relativity and Experiment,” *Living Rev. Rel.* **17** (2014) 4, [arXiv:1403.7377 \[gr-qc\]](#).
- [10] D. Hanneke, S. Fogwell, and G. Gabrielse, “New measurement of the electron magnetic moment and the fine structure constant,” *Phys. Rev. Lett.* **100** (2008) 120801, [arXiv:0801.1134](#).
- [11] **Super-Kamiokande** Collaboration, Y. Fukuda *et al.*, “Evidence for oscillation of atmospheric neutrinos,” *Phys. Rev. Lett.* **81** (1998) 1562–1567, [arXiv:hep-ex/9807003](#).

- [12] A. de Gouvêa, “Neutrino mass models,” *Annual Review of Nuclear and Particle Science* **66** no. 1, (2016) 197–217, <https://doi.org/10.1146/annurev-nucl-102115-044600>.
- [13] F. Zwicky, “Die Rotverschiebung von extragalaktischen Nebeln,” *Helvetica Physica Acta* **6** (Jan., 1933) 110–127.
- [14] V. C. Rubin and J. Ford, W. Kent, “Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions,” **159** (1970) .
- [15] G. Bertone and D. Hooper, “History of dark matter,” *Rev. Mod. Phys.* **90** no. 4, (2018) 045002, [arXiv:1605.04909](https://arxiv.org/abs/1605.04909) [astro-ph.CO].
- [16] **Supernova Search Team** Collaboration, A. G. Riess *et al.*, “Observational evidence from supernovae for an accelerating universe and a cosmological constant,” *Astron. J.* **116** (1998) 1009–1038, [arXiv:astro-ph/9805201](https://arxiv.org/abs/astro-ph/9805201).
- [17] S. Weinberg, “The cosmological constant problem,” *Rev. Mod. Phys.* **61** (Jan, 1989) 1–23. <https://link.aps.org/doi/10.1103/RevModPhys.61.1>.
- [18] J. Martin, “Everything You Always Wanted To Know About The Cosmological Constant Problem (But Were Afraid To Ask),” *Comptes Rendus Physique* **13** (2012) 566–665, [arXiv:1205.3365](https://arxiv.org/abs/1205.3365) [astro-ph.CO].
- [19] S. Carlip, “Quantum gravity: A Progress report,” *Rept. Prog. Phys.* **64** (2001) 885, [arXiv:gr-qc/0108040](https://arxiv.org/abs/gr-qc/0108040).
- [20] C. Csáki and P. Tanedo, “Beyond the Standard Model,” in *2013 European School of High-Energy Physics*, pp. 169–268. 2015. [arXiv:1602.04228](https://arxiv.org/abs/1602.04228) [hep-ph].
- [21] V. Baluni, “CP-nonconserving effects in quantum chromodynamics,” *Phys. Rev. D* **19** (Apr, 1979) 2227–2230. <https://link.aps.org/doi/10.1103/PhysRevD.19.2227>.
- [22] J. M. Pendlebury *et al.*, “Revised experimental upper limit on the electric dipole moment of the neutron,” *Phys. Rev. D* **92** no. 9, (2015) 092003, [arXiv:1509.04411](https://arxiv.org/abs/1509.04411) [hep-ex].
- [23] A. Hook, “TASI Lectures on the Strong CP Problem and Axions,” *PoS TASI2018* (2019) 004, [arXiv:1812.02669](https://arxiv.org/abs/1812.02669) [hep-ph].
- [24] C. Callan, R. Dashen, and D. Gross, “The structure of the gauge theory vacuum,” *Physics Letters B* **63** no. 3, (1976) 334–340. <https://www.sciencedirect.com/science/article/pii/037026937690277X>.
- [25] J. Ellis and M. K. Gaillard, “Strong and weak cp violation,” *Nuclear Physics B* **150** (1979) 141–162. <https://www.sciencedirect.com/science/article/pii/0550321379902979>.

- [26] J. F. Donoghue, “General relativity as an effective field theory: The leading quantum corrections,” *Phys. Rev. D* **50** (1994) 3874–3888, [arXiv:gr-qc/9405057](#).
- [27] C. P. Burgess, “Quantum gravity in everyday life: General relativity as an effective field theory,” *Living Rev. Rel.* **7** (2004) 5–56, [arXiv:gr-qc/0311082](#).
- [28] C. Vafa, “The String landscape and the swampland,” [arXiv:hep-th/0509212](#).
- [29] E. Palti, “The Swampland: Introduction and Review,” *Fortsch. Phys.* **67** no. 6, (2019) 1900037, [arXiv:1903.06239 \[hep-th\]](#).
- [30] G. Dvali and C. Gomez, “Quantum Compositeness of Gravity: Black Holes, AdS and Inflation,” *JCAP* **01** (2014) 023, [arXiv:1312.4795 \[hep-th\]](#).
- [31] G. Dvali and C. Gomez, “Quantum Exclusion of Positive Cosmological Constant?,” *Annalen Phys.* **528** (2016) 68–73, [arXiv:1412.8077 \[hep-th\]](#).
- [32] G. Dvali, C. Gomez, and S. Zell, “Quantum Breaking Bound on de Sitter and Swampland,” *Fortsch. Phys.* **67** no. 1-2, (2019) 1800094, [arXiv:1810.11002 \[hep-th\]](#).
- [33] G. Dvali, “ S -Matrix and Anomaly of de Sitter,” *Symmetry* **13** no. 1, (2020) 3, [arXiv:2012.02133 \[hep-th\]](#).
- [34] E. Witten, “Large n chiral dynamics,” *Annals of Physics* **128** no. 2, (1980) 363–375. <https://www.sciencedirect.com/science/article/pii/0003491680903255>.
- [35] P. Di Vecchia and G. Veneziano, “Chiral dynamics in the large n limit,” *Nuclear Physics B* **171** (1980) 253–272. <https://www.sciencedirect.com/science/article/pii/0550321380903703>.
- [36] R. D. Peccei and H. R. Quinn, “CP conservation in the presence of pseudoparticles,” *Phys. Rev. Lett.* **38** (Jun, 1977) 1440–1443. <https://link.aps.org/doi/10.1103/PhysRevLett.38.1440>.
- [37] R. D. Peccei and H. R. Quinn, “Constraints imposed by CP conservation in the presence of pseudoparticles,” *Phys. Rev. D* **16** (Sep, 1977) 1791–1797. <https://link.aps.org/doi/10.1103/PhysRevD.16.1791>.
- [38] S. Weinberg, “A new light boson?” *Phys. Rev. Lett.* **40** (Jan, 1978) 223–226. <https://link.aps.org/doi/10.1103/PhysRevLett.40.223>.
- [39] F. Wilczek, “Problem of Strong P and T Invariance in the Presence of Instantons,” *Phys. Rev. Lett.* **40** (1978) 279–282.
- [40] **Planck** Collaboration, N. Aghanim *et al.*, “Planck 2018 results. VI. Cosmological parameters,” *Astron. Astrophys.* **641** (2020) A6, [arXiv:1807.06209 \[astro-ph.CO\]](#). [Erratum: *Astron. Astrophys.* 652, C4 (2021)].
- [41] P. J. E. Peebles, *Principles of Physical Cosmology*. 1993.

- [42] S. Weinberg, *Cosmology*. 2008.
- [43] E. A. Baltz, “Dark matter candidates,” *eConf* **C040802** (2004) L002, [arXiv:astro-ph/0412170](https://arxiv.org/abs/astro-ph/0412170).
- [44] G. R. Dvali, “Removing the cosmological bound on the axion scale,” [arXiv:hep-ph/9505253](https://arxiv.org/abs/hep-ph/9505253).
- [45] G. Dvali, “Strong-CP with and without gravity,” [arXiv:2209.14219](https://arxiv.org/abs/2209.14219) [hep-ph].
- [46] G. Dvali, “Three-form gauging of axion symmetries and gravity,” [arXiv:hep-th/0507215](https://arxiv.org/abs/hep-th/0507215).
- [47] G. Dvali and G. R. Farrar, “Strong CP Problem with 10^{32} Standard Model Copies,” *Phys. Rev. Lett.* **101** (2008) 011801, [arXiv:0712.3170](https://arxiv.org/abs/0712.3170) [hep-th].
- [48] G. Dvali, “Black Holes and Large N Species Solution to the Hierarchy Problem,” *Fortsch. Phys.* **58** (2010) 528–536, [arXiv:0706.2050](https://arxiv.org/abs/0706.2050) [hep-th].
- [49] G. Dvali and M. Redi, “Black Hole Bound on the Number of Species and Quantum Gravity at LHC,” *Phys. Rev. D* **77** (2008) 045027, [arXiv:0710.4344](https://arxiv.org/abs/0710.4344) [hep-th].
- [50] M. Srednicki, “Axion couplings to matter: (i). cp-conserving parts,” *Nuclear Physics B* **260** no. 3, (1985) 689–700. <https://www.sciencedirect.com/science/article/pii/0550321385900549>.
- [51] A. R. Zhitnitsky, “On Possible Suppression of the Axion Hadron Interactions. (In Russian),” *Sov. J. Nucl. Phys.* **31** (1980) 260.
- [52] M. Dine, W. Fischler, and M. Srednicki, “A simple solution to the strong cp problem with a harmless axion,” *Physics Letters B* **104** no. 3, (1981) 199–202. <https://www.sciencedirect.com/science/article/pii/0370269381905906>.
- [53] J. E. Kim, “Weak-interaction singlet and strong CP invariance,” *Phys. Rev. Lett.* **43** (Jul, 1979) 103–107. <https://link.aps.org/doi/10.1103/PhysRevLett.43.103>.
- [54] 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.
- [55] 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](https://arxiv.org/abs/2003.01100) [hep-ph].
- [56] S. Weinberg, “The U(1) Problem,” *Phys. Rev.* **D11** (1975) 3583–3593.
- [57] G. ’t Hooft, “Symmetry Breaking Through Bell-Jackiw Anomalies,” *Phys. Rev. Lett.* **37** (1976) 8–11. [[226\(1976\)](https://arxiv.org/abs/226(1976)8)].
- [58] G. ’t Hooft, “Computation of the quantum effects due to a four-dimensional pseudoparticle,” *Phys. Rev. D* **14** (Dec, 1976) 3432–3450. <https://link.aps.org/doi/10.1103/PhysRevD.14.3432>.

- [59] G. 't Hooft, “How instantons solve the $u(1)$ problem,” *Physics Reports* **142** no. 6, (1986) 357–387. <https://www.sciencedirect.com/science/article/pii/0370157386901171>.
- [60] J. Schwinger, “Gauge invariance and mass,” *Phys. Rev.* **125** (Jan, 1962) 397–398. <https://link.aps.org/doi/10.1103/PhysRev.125.397>.
- [61] J. Schwinger, “Gauge invariance and mass. ii,” *Phys. Rev.* **128** (Dec, 1962) 2425–2429. <https://link.aps.org/doi/10.1103/PhysRev.128.2425>.
- [62] G. Dvali, R. Jackiw, and S.-Y. Pi, “Topological mass generation in four dimensions,” *Phys. Rev. Lett.* **96** (2006) 081602, [arXiv:hep-th/0511175](https://arxiv.org/abs/hep-th/0511175).
- [63] G. Dvali, “Three-form gauging of axion symmetries and gravity,” [arXiv:hep-th/0507215](https://arxiv.org/abs/hep-th/0507215).
- [64] G. Dvali, C. Gomez, and S. Zell, “A Proof of the Axion?,” [arXiv:1811.03079](https://arxiv.org/abs/1811.03079) [hep-th].
- [65] D. J. Gross and F. Wilczek, “Ultraviolet behavior of non-abelian gauge theories,” *Phys. Rev. Lett.* **30** (Jun, 1973) 1343–1346. <https://link.aps.org/doi/10.1103/PhysRevLett.30.1343>.
- [66] H. D. Politzer, “Reliable perturbative results for strong interactions?” *Phys. Rev. Lett.* **30** (Jun, 1973) 1346–1349. <https://link.aps.org/doi/10.1103/PhysRevLett.30.1346>.
- [67] N. Evans and K. S. Rigatos, “Chiral symmetry breaking and confinement: separating the scales,” *Phys. Rev. D* **103** (2021) 094022, [arXiv:2012.00032](https://arxiv.org/abs/2012.00032) [hep-ph].
- [68] M. Srednicki, *Quantum field theory*. Cambridge University Press, 1, 2007.
- [69] M. Gell-Mann and M. Levy, “The axial vector current in beta decay,” *Nuovo Cim.* **16** (1960) 705.
- [70] S. L. Adler, “Axial-vector vertex in spinor electrodynamics,” *Phys. Rev.* **177** (Jan, 1969) 2426–2438. <https://link.aps.org/doi/10.1103/PhysRev.177.2426>.
- [71] 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.
- [72] W. A. Bardeen, “Anomalous ward identities in spinor field theories,” *Phys. Rev.* **184** (Aug, 1969) 1848–1859. <https://link.aps.org/doi/10.1103/PhysRev.184.1848>.
- [73] A. A. Belavin, A. M. Polyakov, A. S. Schwartz, and Yu. S. Tyupkin, “Pseudoparticle Solutions of the Yang-Mills Equations,” *Phys. Lett.* **59B** (1975) 85–87. [,350(1975)].

- [74] K. Fujikawa, “Path-integral measure for gauge-invariant fermion theories,” *Phys. Rev. Lett.* **42** (Apr, 1979) 1195–1198. <https://link.aps.org/doi/10.1103/PhysRevLett.42.1195>.
- [75] A. Bilal, “Lectures on Anomalies,” [arXiv:0802.0634](https://arxiv.org/abs/0802.0634) [hep-th].
- [76] M. Shifman, *Advanced Topics in Quantum Field Theory: A Lecture Course*. Cambridge University Press, 2012.
- [77] S. Weinberg, *The quantum theory of fields. Vol. 2: Modern applications*. 1996. <http://www.slac.stanford.edu/spires/find/hep/www?key=3763846>. Cambridge, UK: Univ. Pr. (1996) 489 p.
- [78] C. Vafa and E. Witten, “Restrictions on symmetry breaking in vector-like gauge theories,” *Nuclear Physics B* **234** no. 1, (1984) 173–188. <https://www.sciencedirect.com/science/article/pii/055032138490230X>.
- [79] C. G. Callan, R. Dashen, and D. J. Gross, “Toward a theory of the strong interactions,” *Phys. Rev. D* **17** (May, 1978) 2717–2763. <https://link.aps.org/doi/10.1103/PhysRevD.17.2717>.
- [80] S. Coleman, *Aspects of Symmetry: Selected Erice Lectures*. Cambridge University Press, 1985.
- [81] G. Dvali, S. Folkerts, and A. Franca, “How neutrino protects the axion,” *Phys. Rev. D* **89** no. 10, (2014) 105025, [arXiv:1312.7273](https://arxiv.org/abs/1312.7273) [hep-th].
- [82] E. Farhi and R. Jackiw, “Dynamical gauge symmetry breaking: A collection of reprints,” 1982.
- [83] E. Witten, “Current algebra theorems for the u(1) “goldstone boson”,” *Nuclear Physics B* **156** no. 2, (1979) 269–283. <https://www.sciencedirect.com/science/article/pii/0550321379900312>.
- [84] E. Witten, “Instantons, the Quark Model, and the $1/n$ Expansion,” *Nucl. Phys. B* **149** (1979) 285–320.
- [85] G. Veneziano, “U(1) Without Instantons,” *Nucl. Phys. B* **159** (1979) 213–224.
- [86] E. Meggiolaro, “The Topological susceptibility of QCD: From Minkowskian to Euclidean theory,” *Phys. Rev. D* **58** (1998) 085002, [arXiv:hep-th/9802114](https://arxiv.org/abs/hep-th/9802114).
- [87] **ETM Collaboration** Collaboration, P. Dimopoulos, C. Helmes, C. Jost, *et al.*, “Topological susceptibility and η' meson mass from $N_f = 2$ lattice qcd at the physical point,” *Phys. Rev. D* **99** (2019) 034511, [arXiv:1812.08787](https://arxiv.org/abs/1812.08787).
- [88] J. Kogut and L. Susskind, “How quark confinement solves the $\eta \rightarrow 3\pi$ problem,” *Phys. Rev. D* **11** (Jun, 1975) 3594–3610. <https://link.aps.org/doi/10.1103/PhysRevD.11.3594>.

- [89] M. Lüscher, “The secret long range force in quantum field theories with instantons,” *Physics Letters B* **78** no. 4, (1978) 465–467. <https://www.sciencedirect.com/science/article/pii/0370269378904872>.
- [90] P. Di Vecchia, F. Nicodemi, R. Pettorino, and G. Veneziano, “Large n , chiral approach to pseudoscalar masses, mixings and decays,” *Nuclear Physics B* **181** no. 2, (1981) 318–334. <https://www.sciencedirect.com/science/article/pii/0550321381903564>.
- [91] P. Di Vecchia and G. Veneziano, “Chiral dynamics in the large n limit,” *Nuclear Physics B* **171** (1980) 253–272. <https://www.sciencedirect.com/science/article/pii/0550321380903703>.
- [92] C. Eckart, “The application of group theory to the quantum dynamics of monatomic systems,” *Rev. Mod. Phys.* **2** (Jul, 1930) 305–380. <https://link.aps.org/doi/10.1103/RevModPhys.2.305>.
- [93] J. J. Sakurai and E. D. Commins, “Modern quantum mechanics, revised edition,” *American Journal of Physics* **63** no. 1, (1995) 93–95, <https://doi.org/10.1119/1.17781>. <https://doi.org/10.1119/1.17781>.
- [94] M. Pospelov and A. Ritz, “Theta vacua, QCD sum rules, and the neutron electric dipole moment,” *Nucl. Phys. B* **573** (2000) 177–200, [arXiv:hep-ph/9908508](https://arxiv.org/abs/hep-ph/9908508).
- [95] L. Bartolini, F. Bigazzi, S. Bolognesi, *et al.*, “Neutron electric dipole moment from gauge/string duality,” *Phys. Rev. Lett.* **118** no. 9, (2017) 091601, [arXiv:1609.09513](https://arxiv.org/abs/1609.09513) [hep-ph].
- [96] M. Abramczyk, S. Aoki, T. Blum, *et al.*, “Lattice calculation of electric dipole moments and form factors of the nucleon,” *Phys. Rev. D* **96** no. 1, (2017) 014501, [arXiv:1701.07792](https://arxiv.org/abs/1701.07792) [hep-lat].
- [97] **Particle Data Group** Collaboration, M. Tanabashi, K. Hagiwara, K. Hikasa, *et al.*, “Review of particle physics,” *Phys. Rev. D* **98** (2018) 030001.
- [98] K. Choi, K. Kang, and J. E. Kim, “Effects of η' in Low-energy Axion Physics,” *Phys. Lett. B* **181** (1986) 145–149.
- [99] G. Grilli di Cortona, E. Hardy, J. Pardo Vega, and G. Villadoro, “The QCD axion, precisely,” *JHEP* **01** (2016) 034, [arXiv:1511.02867](https://arxiv.org/abs/1511.02867) [hep-ph].
- [100] G. C. Branco, P. M. Ferreira, L. Lavoura, *et al.*, “Theory and phenomenology of two-Higgs-doublet models,” *Phys. Rept.* **516** (2012) 1–102, [arXiv:1106.0034](https://arxiv.org/abs/1106.0034) [hep-ph].
- [101] S. L. Glashow, J. Iliopoulos, and L. Maiani, “Weak interactions with lepton-hadron symmetry,” *Phys. Rev. D* **2** (Oct, 1970) 1285–1292. <https://link.aps.org/doi/10.1103/PhysRevD.2.1285>.

- [102] **CMS** Collaboration, D. Pedrini, “Search for the flavour-changing neutral current decay $D^0 \rightarrow \mu^+ \mu^-$ in pp collisions at $\sqrt{s} = 7$ TeV with CMS,” in *5th International Workshop on Charm Physics*. 8, 2012. [arXiv:1208.5908 \[hep-ex\]](#).
- [103] **ATLAS, CDF, CMS, D0** Collaboration, E. Yazgan, “Flavor changing neutral currents in top quark production and decay,” in *6th International Workshop on Top Quark Physics*, pp. 285–294. 2014. [arXiv:1312.5435 \[hep-ex\]](#).
- [104] S. L. Glashow and S. Weinberg, “Natural Conservation Laws for Neutral Currents,” *Phys. Rev.* **D15** (1977) 1958.
- [105] E. A. Paschos, “Diagonal neutral currents,” *Phys. Rev. D* **15** (Apr, 1977) 1966–1972. <https://link.aps.org/doi/10.1103/PhysRevD.15.1966>.
- [106] L. Di Luzio, F. Mescia, and E. Nardi, “Window for preferred axion models,” *Phys. Rev. D* **96** no. 7, (2017) 075003, [arXiv:1705.05370 \[hep-ph\]](#).
- [107] R. Kallosh, A. D. Linde, D. A. Linde, and L. Susskind, “Gravity and global symmetries,” *Phys. Rev. D* **52** (1995) 912–935, [arXiv:hep-th/9502069](#).
- [108] A. Arvanitaki, S. Dimopoulos, S. Dubovsky, *et al.*, “String Axiverse,” *Phys. Rev. D* **81** (2010) 123530, [arXiv:0905.4720 \[hep-th\]](#).
- [109] J. Preskill, M. B. Wise, and F. Wilczek, “Cosmology of the invisible axion,” *Physics Letters B* **120** no. 1, (1983) 127–132. <https://www.sciencedirect.com/science/article/pii/0370269383906378>.
- [110] M. Dine and W. Fischler, “The not-so-harmless axion,” *Physics Letters B* **120** no. 1, (1983) 137–141. <https://www.sciencedirect.com/science/article/pii/0370269383906391>.
- [111] L. F. Abbott and P. Sikivie, “A Cosmological Bound on the Invisible Axion,” *Phys. Lett. B* **120** (1983) 133–136.
- [112] O. Sakhelashvili, “Consistency of the dual formulation of axion solutions to the strong cp problem,” *Phys. Rev. D* **105** (Apr, 2022) 085020. <https://link.aps.org/doi/10.1103/PhysRevD.105.085020>.
- [113] K. Choi, H. B. Kim, and J. E. Kim, “Axion cosmology with a stronger QCD in the early universe,” *Nucl. Phys. B* **490** (1997) 349–364, [arXiv:hep-ph/9606372](#).
- [114] R. T. Co, E. Gonzalez, and K. Harigaya, “Axion Misalignment Driven to the Bottom,” *JHEP* **05** (2019) 162, [arXiv:1812.11186 \[hep-ph\]](#).
- [115] H. Matsui, F. Takahashi, and W. Yin, “QCD Axion Window and False Vacuum Higgs Inflation,” *JHEP* **05** (2020) 154, [arXiv:2001.04464 \[hep-ph\]](#).
- [116] F. Takahashi, W. Yin, and A. H. Guth, “QCD axion window and low-scale inflation,” *Phys. Rev. D* **98** no. 1, (2018) 015042, [arXiv:1805.08763 \[hep-ph\]](#).

- [117] G. Dvali and J. B. Gómez, “Axion alignment during inflation,” unpublished Thesis by J. B. Gómez.
- [118] D. J. Gross, R. D. Pisarski, and L. G. Yaffe, “Qcd and instantons at finite temperature,” *Rev. Mod. Phys.* **53** (Jan, 1981) 43–80. <https://link.aps.org/doi/10.1103/RevModPhys.53.43>.
- [119] A. Vilenkin, “Cosmic strings,” *Phys. Rev. D* **24** (Oct, 1981) 2082–2089. <https://link.aps.org/doi/10.1103/PhysRevD.24.2082>.
- [120] R. Davis, “Cosmic axions from cosmic strings,” *Physics Letters B* **180** no. 3, (1986) 225–230. <https://www.sciencedirect.com/science/article/pii/037026938690300X>.
- [121] M. Buschmann, J. W. Foster, and B. R. Safdi, “Early-Universe Simulations of the Cosmological Axion,” *Phys. Rev. Lett.* **124** no. 16, (2020) 161103, [arXiv:1906.00967](https://arxiv.org/abs/1906.00967) [astro-ph.CO].
- [122] M. P. Hertzberg, M. Tegmark, and F. Wilczek, “Axion Cosmology and the Energy Scale of Inflation,” *Phys. Rev. D* **78** (2008) 083507, [arXiv:0807.1726](https://arxiv.org/abs/0807.1726) [astro-ph].
- [123] E. W. Kolb, A. Notari, and A. Riotto, “On the reheating stage after inflation,” *Phys. Rev. D* **68** (2003) 123505, [arXiv:hep-ph/0307241](https://arxiv.org/abs/hep-ph/0307241).
- [124] T. W. B. Kibble, “Topology of cosmic domains and strings,” *Journal of Physics A: Mathematical and General* **9** no. 8, (Aug, 1976) 1387. <https://dx.doi.org/10.1088/0305-4470/9/8/029>.
- [125] T. W. B. Kibble, G. Lazarides, and Q. Shafi, “Walls bounded by strings,” *Phys. Rev. D* **26** (Jul, 1982) 435–439. <https://link.aps.org/doi/10.1103/PhysRevD.26.435>.
- [126] T. Kibble, “Some implications of a cosmological phase transition,” *Physics Reports* **67** no. 1, (1980) 183–199. <https://www.sciencedirect.com/science/article/pii/0370157380900915>.
- [127] A. Vilenkin and A. E. Everett, “Cosmic strings and domain walls in models with goldstone and pseudo-goldstone bosons,” *Phys. Rev. Lett.* **48** (Jun, 1982) 1867–1870. <https://link.aps.org/doi/10.1103/PhysRevLett.48.1867>.
- [128] P. Sikivie, “Axions, domain walls, and the early universe,” *Phys. Rev. Lett.* **48** (Apr, 1982) 1156–1159. <https://link.aps.org/doi/10.1103/PhysRevLett.48.1156>.
- [129] A. Vilenkin and E. P. S. Shellard, *Cosmic Strings and Other Topological Defects*. Cambridge University Press, 7, 2000.
- [130] A. Ernst, A. Ringwald, and C. Tamarit, “Axion Predictions in $SO(10) \times U(1)_{PQ}$ Models,” *JHEP* **02** (2018) 103, [arXiv:1801.04906](https://arxiv.org/abs/1801.04906) [hep-ph].

- [131] T. Kobayashi, R. Kurematsu, and F. Takahashi, “Isocurvature Constraints and Anharmonic Effects on QCD Axion Dark Matter,” *JCAP* **09** (2013) 032, [arXiv:1304.0922 \[hep-ph\]](#).
- [132] G. G. Raffelt, “Astrophysical axion bounds,” *Lect. Notes Phys.* **741** (2008) 51–71, [arXiv:hep-ph/0611350](#).
- [133] A. Arvanitaki and S. Dubovsky, “Exploring the String Axiverse with Precision Black Hole Physics,” *Phys. Rev. D* **83** (2011) 044026, [arXiv:1004.3558 \[hep-th\]](#).
- [134] A. Arvanitaki, M. Baryakhtar, and X. Huang, “Discovering the QCD Axion with Black Holes and Gravitational Waves,” *Phys. Rev. D* **91** no. 8, (2015) 084011, [arXiv:1411.2263 \[hep-ph\]](#).
- [135] K. Kamada, “Inflationary cosmology and the standard model Higgs with a small Hubble induced mass,” *Phys. Lett. B* **742** (2015) 126–135, [arXiv:1409.5078 \[hep-ph\]](#).
- [136] G. Dvali, F. Kühnel, and M. Zantedeschi, “Primordial black holes from confinement,” *Phys. Rev. D* **104** no. 12, (2021) 123507, [arXiv:2108.09471 \[hep-ph\]](#).
- [137] **Particle Data Group** Collaboration, P. Zyla *et al.*, “Review of Particle Physics,” *PTEP* **2020** no. 8, (2020) 083C01.
- [138] K. S. Jeong and F. Takahashi, “Suppressing Isocurvature Perturbations of QCD Axion Dark Matter,” *Phys. Lett. B* **727** (2013) 448–451, [arXiv:1304.8131 \[hep-ph\]](#).
- [139] K. J. Mack and P. J. Steinhardt, “Cosmological Problems with Multiple Axion-like Fields,” *JCAP* **05** (2011) 001, [arXiv:0911.0418 \[astro-ph.CO\]](#).
- [140] S. I. Blinnikov and M. Y. Khlopov, “ON POSSIBLE EFFECTS OF ‘MIRROR’ PARTICLES,” *Sov. J. Nucl. Phys.* **36** (1982) 472.
- [141] E. W. Kolb, D. Seckel, and M. S. Turner, “The Shadow World,” *Nature* **314** (1985) 415–419.
- [142] D. J. Gross, J. A. Harvey, E. Martinec, and R. Rohm, “Heterotic string,” *Phys. Rev. Lett.* **54** (Feb, 1985) 502–505. <https://link.aps.org/doi/10.1103/PhysRevLett.54.502>.
- [143] L. J. Dixon, J. A. Harvey, C. Vafa, and E. Witten, “Strings on Orbifolds,” *Nucl. Phys. B* **261** (1985) 678–686.
- [144] O. Lebedev, H. P. Nilles, S. Raby, *et al.*, “A Mini-landscape of exact MSSM spectra in heterotic orbifolds,” *Phys. Lett. B* **645** (2007) 88–94, [arXiv:hep-th/0611095](#).
- [145] V. Braun, Y.-H. He, B. A. Ovrut, and T. Pantev, “A Heterotic standard model,” *Phys. Lett. B* **618** (2005) 252–258, [arXiv:hep-th/0501070](#).

- [146] J. Halverson, C. Long, A. Maiti, *et al.*, “Gravitational waves from dark Yang-Mills sectors,” *JHEP* **05** (2021) 154, [arXiv:2012.04071 \[hep-ph\]](#).
- [147] W.-C. Huang, M. Reichert, F. Sannino, and Z.-W. Wang, “Testing the dark SU(N) Yang-Mills theory confined landscape: From the lattice to gravitational waves,” *Phys. Rev. D* **104** no. 3, (2021) 035005, [arXiv:2012.11614 \[hep-ph\]](#).
- [148] G. Dvali, I. Sawicki, and A. Vikman, “Dark Matter via Many Copies of the Standard Model,” *JCAP* **0908** (2009) 009, [arXiv:0903.0660 \[hep-th\]](#).
- [149] G. Dvali and M. Redi, “Phenomenology of 10^{32} Dark Sectors,” *Phys. Rev.* **D80** (2009) 055001, [arXiv:0905.1709 \[hep-ph\]](#).
- [150] M. Ettengruber, “Neutrino physics in TeV scale gravity theories,” *Phys. Rev. D* **106** no. 5, (2022) 055028, [arXiv:2206.00034 \[hep-ph\]](#).
- [151] G. Dvali and O. Pujolas, “Micro Black Holes and the Democratic Transition,” *Phys. Rev. D* **79** (2009) 064032, [arXiv:0812.3442 \[hep-th\]](#).
- [152] G. Dvali, E. Koutsangelas, and F. Kuhnel, “Compact Dark Matter Objects via N Dark Sectors,” *Phys. Rev. D* **101** (2020) 083533, [arXiv:1911.13281 \[astro-ph.CO\]](#).
- [153] Z. Berezhiani, D. Comelli, and F. L. Villante, “The Early mirror universe: Inflation, baryogenesis, nucleosynthesis and dark matter,” *Phys. Lett. B* **503** (2001) 362–375, [arXiv:hep-ph/0008105](#).
- [154] N. Arkani-Hamed, S. Dimopoulos, G. R. Dvali, and N. Kaloper, “Many fold universe,” *JHEP* **12** (2000) 010, [arXiv:hep-ph/9911386](#).
- [155] R. Foot, H. Lew, and R. Volkas, “A model with fundamental improper space-time symmetries,” *Physics Letters B* **272** no. 1, (1991) 67–70. <https://www.sciencedirect.com/science/article/pii/037026939191013L>.
- [156] P. Fox, A. Pierce, and S. D. Thomas, “Probing a QCD string axion with precision cosmological measurements,” [arXiv:hep-th/0409059](#).
- [157] F. Takahashi, W. Yin, and A. H. Guth, “QCD axion window and low-scale inflation,” *Phys. Rev. D* **98** no. 1, (2018) 015042, [arXiv:1805.08763 \[hep-ph\]](#).
- [158] D. E. Kaplan and R. Rattazzi, “Large field excursions and approximate discrete symmetries from a clockwork axion,” *Phys. Rev. D* **93** no. 8, (2016) 085007, [arXiv:1511.01827 \[hep-ph\]](#).
- [159] I. Tkachev, “On the possibility of bose-star formation,” *Physics Letters B* **261** no. 3, (1991) 289–293. <https://www.sciencedirect.com/science/article/pii/037026939190330S>.

- [160] J. H. Chang, D. Egana-Ugrinovic, R. Essig, and C. Kouvaris, “Structure Formation and Exotic Compact Objects in a Dissipative Dark Sector,” *JCAP* **03** (2019) 036, [arXiv:1812.07000 \[hep-ph\]](#).
- [161] B. Carr and F. Kuhnel, “Primordial Black Holes as Dark Matter: Recent Developments,” *Ann. Rev. Nucl. Part. Sci.* **70** (2020) 355–394, [arXiv:2006.02838 \[astro-ph.CO\]](#).
- [162] H. Zhang, “Axion Stars,” *Symmetry* **12** no. 1, (2019) 25, [arXiv:1810.11473 \[hep-ph\]](#).
- [163] S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*. New York: Wiley, 1972.
- [164] J. R. Oppenheimer and G. M. Volkoff, “On massive neutron cores,” *Phys. Rev.* **55** (Feb, 1939) 374–381. <https://link.aps.org/doi/10.1103/PhysRev.55.374>.
- [165] P.-H. Chavanis, “Mass-radius relation of Newtonian self-gravitating Bose-Einstein condensates with short-range interactions. I. Analytical results,” **84** no. 4, (2011) 043531, [arXiv:1103.2050 \[astro-ph.CO\]](#).
- [166] L. Berezhiani, G. Cintia, and M. Warkentin, “Core fragmentation in simplest superfluid dark matter scenario,” *Phys. Lett. B* **819** (2021) 136422, [arXiv:2101.08117 \[astro-ph.CO\]](#).
- [167] L. Berezhiani, G. Cintia, and J. Khoury, “Thermalization, Fragmentation and Tidal Disruption: The Complex Galactic Dynamics of Dark Matter Superfluidity,” [arXiv:2212.10577 \[astro-ph.CO\]](#).
- [168] R. Foot, “Mirror dark matter: Cosmology, galaxy structure and direct detection,” *Int. J. Mod. Phys. A* **29** (2014) 1430013, [arXiv:1401.3965 \[astro-ph.CO\]](#).
- [169] C. O’Hare, “cajohare/axionlimits: Axionlimits.” <https://cajohare.github.io/AxionLimits/>, Jan., 2023.
- [170] V. Plakkot and S. Hoof, “Anomaly ratio distributions of hadronic axion models with multiple heavy quarks,” **104** no. 7, (Oct., 2021) 075017, [arXiv:2107.12378 \[hep-ph\]](#).
- [171] G. R. Dvali and G. Senjanovic, “Is there a domain wall problem?,” *Phys. Rev. Lett.* **74** (1995) 5178–5181, [arXiv:hep-ph/9501387](#).
- [172] G. Lazarides and Q. Shafi, “Axion models with no domain wall problem,” *Physics Letters B* **115** no. 1, (1982) 21–25. <https://www.sciencedirect.com/science/article/pii/0370269382905068>.
- [173] A. Pich, “Flavour Physics and CP Violation,” in *6th CERN-Latin-American School of High-Energy Physics*, pp. 119–144. 2013. [arXiv:1112.4094 \[hep-ph\]](#).

- [174] I. P. Ivanov, “Building and testing models with extended Higgs sectors,” *Prog. Part. Nucl. Phys.* **95** (2017) 160–208, arXiv:1702.03776 [hep-ph].
- [175] S. L. Glashow and S. Weinberg, “Natural conservation laws for neutral currents,” *Phys. Rev. D* **15** (Apr, 1977) 1958–1965. <https://link.aps.org/doi/10.1103/PhysRevD.15.1958>.
- [176] M. Y. Gogberashvili and G. R. Dvali, “Hierarchy at Yukawa constants and K_0 anti- K_0 , B_0 anti- B_0 oscillations in the model with two Higgs doublets. (In Russian),” *Sov. J. Nucl. Phys.* **53** (1991) 491–492.
- [177] A. Pich and P. Tuzon, “Yukawa Alignment in the Two-Higgs-Doublet Model,” *Phys. Rev. D* **80** (2009) 091702, arXiv:0908.1554 [hep-ph].
- [178] I. de Medeiros Varzielas and J. Talbert, “FCNC-free multi-Higgs-doublet models from broken family symmetries,” *Phys. Lett. B* **800** (2020) 135091, arXiv:1908.10979 [hep-ph].
- [179] T. P. Cheng and M. Sher, “Mass-matrix ansatz and flavor nonconservation in models with multiple higgs doublets,” *Phys. Rev. D* **35** (Jun, 1987) 3484–3491. <https://link.aps.org/doi/10.1103/PhysRevD.35.3484>.
- [180] S. Carrolo, J. C. Romão, J. a. P. Silva, and F. Vazão, “Symmetry and decoupling in multi-Higgs boson models,” *Phys. Rev. D* **103** no. 7, (2021) 075026, arXiv:2102.11303 [hep-ph].
- [181] A. Peñuelas and A. Pich, “Flavour alignment in multi-Higgs-doublet models,” *JHEP* **12** (2017) 084, arXiv:1710.02040 [hep-ph].
- [182] M. Farina, D. Pappadopulo, F. Rompineve, and A. Tesi, “The photo-philic QCD axion,” *JHEP* **01** (2017) 095, arXiv:1611.09855 [hep-ph].
- [183] J. Bezanson, S. Karpinski, V. B. Shah, and A. Edelman, “Julia: A fast dynamic language for technical computing,” *CoRR* **abs/1209.5145** (2012) , 1209.5145. <http://arxiv.org/abs/1209.5145>.
- [184] “Staticarrays.jl.” <https://juliaarrays.github.io/StaticArrays.jl/>, Aug., 2022.
- [185] K. Pearson, “Mathematical contributions to the theory of evolution. xix. second supplement to a memoir on skew variation,” *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* **216** (1916) 429–457. <http://www.jstor.org/stable/91092>.
- [186] J. Sloan, M. Hotz, C. Boutan, *et al.*, “Limits on axion–photon coupling or on local axion density: Dependence on models of the milky way’s dark halo,” *Physics of the Dark Universe* **14** (2016) 95–102. <https://www.sciencedirect.com/science/article/pii/S2212686416300504>.

- [187] **TASTE** Collaboration, V. Anastassopoulos *et al.*, “Towards a medium-scale axion helioscope and haloscope,” *JINST* **12** no. 11, (2017) P11019, [arXiv:1706.09378](#) [[hep-ph](#)].
- [188] S. Scopel, “Particle dark matter candidates,” *J. Phys. Conf. Ser.* **120** (2008) 042003, [arXiv:0711.2870](#) [[hep-ph](#)].
- [189] 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** no. 3, (2013) 4500604, [arXiv:1212.4633](#) [[physics.ins-det](#)].
- [190] **CAST** Collaboration, V. Anastassopoulos *et al.*, “New CAST Limit on the Axion-Photon Interaction,” *Nature Phys.* **13** (2017) 584–590, [arXiv:1705.02290](#) [[hep-ex](#)].
- [191] E. Armengaud, D. Attié, S. Basso, *et al.*, “Physics potential of the International Axion Observatory (IAXO),” *JCAP* **2019** no. 6, (June, 2019) 047, [arXiv:1904.09155](#) [[hep-ph](#)].
- [192] S. Beurthey *et al.*, “MADMAX Status Report,” [arXiv:2003.10894](#) [[physics.ins-det](#)].
- [193] T. Grenet, R. Ballou, Q. Basto, *et al.*, “The Grenoble Axion Haloscope platform (GrAHal): development plan and first results,” [arXiv:2110.14406](#) [[hep-ex](#)].
- [194] **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](#) [[hep-ex](#)].
- [195] N. Crisosto, P. Sikivie, N. S. Sullivan, *et al.*, “ADMX SLIC: Results from a Superconducting LC Circuit Investigating Cold Axions,” *Phys. Rev. Lett.* **124** no. 24, (2020) 241101, [arXiv:1911.05772](#) [[astro-ph.CO](#)].
- [196] 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** no. 4, (2021) 041301, [arXiv:2101.11290](#) [[astro-ph.CO](#)].
- [197] **QUAX** Collaboration, N. Crescini *et al.*, “Axion search with a quantum-limited ferromagnetic haloscope,” *Phys. Rev. Lett.* **124** no. 17, (2020) 171801, [arXiv:2001.08940](#) [[hep-ex](#)].
- [198] D. Alesini *et al.*, “Galactic axions search with a superconducting resonant cavity,” *Phys. Rev. D* **99** no. 10, (2019) 101101, [arXiv:1903.06547](#) [[physics.ins-det](#)].
- [199] D. Alesini *et al.*, “Search for invisible axion dark matter of mass $m_a = 43 \mu\text{eV}$ with the QUAX- $a\gamma$ experiment,” *Phys. Rev. D* **103** no. 10, (2021) 102004, [arXiv:2012.09498](#) [[hep-ex](#)].

- [200] C. A. Thomson, B. T. McAllister, M. Goryachev, *et al.*, “Upconversion Loop Oscillator Axion Detection Experiment: A Precision Frequency Interferometric Axion Dark Matter Search with a Cylindrical Microwave Cavity,” *Phys. Rev. Lett.* **126** no. 8, (2021) 081803, [arXiv:1912.07751 \[hep-ex\]](#). [Erratum: *Phys.Rev.Lett.* **127**, 019901 (2021)].
- [201] B. T. McAllister, G. Flower, J. Kruger, *et al.*, “The ORGAN Experiment: An axion haloscope above 15 GHz,” *Phys. Dark Univ.* **18** (2017) 67–72, [arXiv:1706.00209 \[physics.ins-det\]](#).
- [202] J. Jeong, S. Youn, S. Bae, *et al.*, “Search for Invisible Axion Dark Matter with a Multiple-Cell Haloscope,” *Phys. Rev. Lett.* **125** no. 22, (2020) 221302, [arXiv:2008.10141 \[hep-ex\]](#).
- [203] A. V. Gramolin, D. Aybas, D. Johnson, *et al.*, “Search for axion-like dark matter with ferromagnets,” *Nature Phys.* **17** no. 1, (2021) 79–84, [arXiv:2003.03348 \[hep-ex\]](#).
- [204] C. P. Salemi *et al.*, “Search for Low-Mass Axion Dark Matter with ABRACADABRA-10 cm,” *Phys. Rev. Lett.* **127** no. 8, (2021) 081801, [arXiv:2102.06722 \[hep-ex\]](#).
- [205] J. L. Ouellet *et al.*, “First Results from ABRACADABRA-10 cm: A Search for Sub- μeV Axion Dark Matter,” *Phys. Rev. Lett.* **122** no. 12, (2019) 121802, [arXiv:1810.12257 \[hep-ex\]](#).
- [206] **HAYSTAC** Collaboration, L. Zhong *et al.*, “Results from phase 1 of the HAYSTAC microwave cavity axion experiment,” *Phys. Rev. D* **97** no. 9, (2018) 092001, [arXiv:1803.03690 \[hep-ex\]](#).
- [207] **HAYSTAC** Collaboration, K. M. Backes *et al.*, “A quantum-enhanced search for dark matter axions,” *Nature* **590** no. 7845, (2021) 238–242, [arXiv:2008.01853 \[quant-ph\]](#).
- [208] C. Hagmann, P. Sikivie, N. Sullivan, and D. Tanner, “Results from a search for cosmic axions,” *Physical Review D* **42** no. 4, (1990) 1297.
- [209] S. DePanfilis, A. Melissinos, B. Moskowitz, *et al.*, “Limits on the abundance and coupling of cosmic axions at $4.5 < m_a < 5.0 \mu\text{eV}$,” *Physical Review Letters* **59** no. 7, (1987) 839.
- [210] **ADMX** Collaboration, N. Du *et al.*, “A Search for Invisible Axion Dark Matter with the Axion Dark Matter Experiment,” *Phys. Rev. Lett.* **120** no. 15, (2018) 151301, [arXiv:1804.05750 \[hep-ex\]](#).
- [211] **ADMX** Collaboration, T. Braine *et al.*, “Extended Search for the Invisible Axion with the Axion Dark Matter Experiment,” *Phys. Rev. Lett.* **124** no. 10, (2020) 101303, [arXiv:1910.08638 \[hep-ex\]](#).

- [212] **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** no. 26, (2021) 261803, [arXiv:2110.06096](#) [[hep-ex](#)].
- [213] M. Baryakhtar, J. Huang, and R. Lasenby, “Axion and hidden photon dark matter detection with multilayer optical haloscopes,” *Phys. Rev. D* **98** no. 3, (2018) 035006, [arXiv:1803.11455](#) [[hep-ph](#)].
- [214] Y. Michimura, Y. Oshima, T. Watanabe, *et al.*, “DANCE: Dark matter Axion search with riNg Cavity Experiment,” *J. Phys. Conf. Ser.* **1468** no. 1, (2020) 012032, [arXiv:1911.05196](#) [[physics.ins-det](#)].
- [215] B. Aja *et al.*, “The Canfranc Axion Detection Experiment (CADEx): search for axions at 90 GHz with Kinetic Inductance Detectors,” *JCAP* **11** (2022) 044, [arXiv:2206.02980](#) [[hep-ex](#)].
- [216] **BREAD** Collaboration, J. Liu *et al.*, “Broadband Solenoidal Haloscope for Terahertz Axion Detection,” *Phys. Rev. Lett.* **128** no. 13, (2022) 131801, [arXiv:2111.12103](#) [[physics.ins-det](#)].
- [217] A. J. Millar *et al.*, “ALPHA: Searching For Dark Matter with Plasma Haloscopes,” [arXiv:2210.00017](#) [[hep-ph](#)].
- [218] M. Lawson, A. J. Millar, M. Pancaldi, *et al.*, “Tunable axion plasma haloscopes,” *Phys. Rev. Lett.* **123** no. 14, (2019) 141802, [arXiv:1904.11872](#) [[hep-ph](#)].
- [219] **TASEH** Collaboration, H. Chang *et al.*, “First Results from the Taiwan Axion Search Experiment with a Haloscope at 19.6 μeV ,” *Phys. Rev. Lett.* **129** no. 11, (2022) 111802, [arXiv:2205.05574](#) [[hep-ex](#)].
- [220] D. Alesini, D. Babusci, C. Braggio, *et al.*, “Search for Galactic axions with a high-Q dielectric cavity,” *Phys. Rev. D* **106** no. 5, (Sept., 2022) 052007, [arXiv:2208.12670](#) [[hep-ex](#)].
- [221] A. P. Quiskamp, B. T. McAllister, P. Altin, *et al.*, “Direct search for dark matter axions excluding ALPogenesis in the 63- to 67- μeV range with the ORGAN experiment,” *Sci. Adv.* **8** no. 27, (2022) abq3765, [arXiv:2203.12152](#) [[hep-ex](#)].
- [222] **HAYSTAC** Collaboration, M. J. Jewell *et al.*, “New results from HAYSTAC’s phase II operation with a squeezed state receiver,” *Phys. Rev. D* **107** no. 7, (2023) 072007, [arXiv:2301.09721](#) [[hep-ex](#)].
- [223] S. Lee, S. Ahn, J. Choi, *et al.*, “Axion Dark Matter Search around 6.7 μeV ,” *Phys. Rev. Lett.* **124** no. 10, (Mar., 2020) 101802, [arXiv:2001.05102](#) [[hep-ex](#)].
- [224] O. Kwon, D. Lee, W. Chung, *et al.*, “First Results from an Axion Haloscope at CAPP around 10.7 μeV ,” *Phys. Rev. Lett.* **126** no. 19, (May, 2021) 191802, [arXiv:2012.10764](#) [[hep-ex](#)].

- [225] Y. Lee, B. Yang, H. Yoon, *et al.*, “Searching for Invisible Axion Dark Matter with an 18 T Magnet Haloscope,” *Phys. Rev. Lett.* **128** no. 24, (June, 2022) 241805, [arXiv:2206.08845 \[hep-ex\]](#).
- [226] J. Kim *et al.*, “Near-Quantum-Noise Axion Dark Matter Search at CAPP around $9.5 \mu\text{eV}$,” *Phys. Rev. Lett.* **130** no. 9, (2023) 091602, [arXiv:2207.13597 \[hep-ex\]](#).
- [227] A. K. Yi *et al.*, “Axion Dark Matter Search around $4.55 \mu\text{eV}$ with Dine-Fischler-Srednicki-Zhitnitskii Sensitivity,” *Phys. Rev. Lett.* **130** no. 7, (2023) 071002, [arXiv:2210.10961 \[hep-ex\]](#).
- [228] A. Berlin, R. T. D’Agnolo, S. A. R. Ellis, and K. Zhou, “Heterodyne broadband detection of axion dark matter,” *Phys. Rev. D* **104** no. 11, (2021) L111701, [arXiv:2007.15656 \[hep-ph\]](#).
- [229] **DMRadio** Collaboration, L. Brouwer *et al.*, “Projected sensitivity of DMRadio-m3: A search for the QCD axion below $1 \mu\text{eV}$,” *Phys. Rev. D* **106** no. 10, (2022) 103008, [arXiv:2204.13781 \[hep-ex\]](#).
- [230] D. Alesini, D. Babusci, D. Di Gioacchino, *et al.*, “The KLASH Proposal,” [arXiv:1707.06010 \[physics.ins-det\]](#).
- [231] Z. Zhang, D. Horns, and O. Ghosh, “Search for dark matter with an LC circuit,” *Phys. Rev. D* **106** no. 2, (2022) 023003, [arXiv:2111.04541 \[hep-ex\]](#).
- [232] K. Nagano, T. Fujita, Y. Michimura, and I. Obata, “Axion Dark Matter Search with Interferometric Gravitational Wave Detectors,” *Phys. Rev. Lett.* **123** no. 11, (2019) 111301, [arXiv:1903.02017 \[hep-ph\]](#).
- [233] H. Liu, B. D. Elwood, M. Evans, and J. Thaler, “Searching for Axion Dark Matter with Birefringent Cavities,” *Phys. Rev. D* **100** no. 2, (2019) 023548, [arXiv:1809.01656 \[hep-ph\]](#).
- [234] J. Schütte-Engel, D. J. E. Marsh, A. J. Millar, *et al.*, “Axion quasiparticles for axion dark matter detection,” *JCAP* **08** (2021) 066, [arXiv:2102.05366 \[hep-ph\]](#).
- [235] D. J. E. Marsh, K.-C. Fong, E. W. Lentz, *et al.*, “Proposal to Detect Dark Matter using Axionic Topological Antiferromagnets,” *Phys. Rev. Lett.* **123** no. 12, (2019) 121601, [arXiv:1807.08810 \[hep-ph\]](#).
- [236] C. S. Lorenz, L. Funcke, E. Calabrese, and S. Hannestad, “Time-varying neutrino mass from a supercooled phase transition: current cosmological constraints and impact on the Ω_m - σ_8 plane,” *Phys. Rev. D* **99** no. 2, (2019) 023501, [arXiv:1811.01991 \[astro-ph.CO\]](#).
- [237] F. Björkeröth, L. Di Luzio, F. Mescia, *et al.*, “Axion-electron decoupling in nucleophobic axion models,” *Phys. Rev. D* **101** no. 3, (2020) 035027, [arXiv:1907.06575 \[hep-ph\]](#).