SISSA

Scuola Internazionale Superiore di Studi Avanzati - Trieste

DOCTORAL THESIS

Axions from Strings

CANDIDATE : Marco Gorghetto
ADVISOR : Prof. Giovanni Villadoro

OPPONENTS : Prof. Asimina Arvanitaki
Prof. Javier Redondo

ACADEMIC YEAR 2018 – 2019

SISSA - Via Bonomea 265 - 34136 TRIESTE - ITALY

Acknowledgments

First of all I would like to thank Giovanni Villadoro. I hope that his insistent scientific honesty and reliability, and his ability to deeply understand a broad variety of subjects will always serve as an example for me in the future. His guidance and support were crucial during my PhD work and essential for the results in this thesis. I would also like to thank Ed Hardy for the fruitful collaboration in the last years, for his dedication and determination in progressing with ongoing projects and in guiding me in the early stages of my research. I am also grateful to Alex Azatov for collaborations at an the early stage my PhD.

The SISSA professors deserve special acknowledgments for their constant effort in offering a solid teaching program, and for their support and close interactions during the first year. I am particularly grateful to Matteo Bertolini, Andrea Romanino and Marco Serone for their help and patience in answering questions outside lectures, and for the quality of their courses. Finally, I am grateful to the ICTP for hospitality during the last three years and to all the ICTP staff and for the wonderful research environment they provide.

Abstract

In this thesis we first calculate the relation between QCD axion mass and decay constant at high precision, employing effective chiral lagrangian techniques. In particular we include the $O(\alpha_{\rm em})$ and NNLO corrections in the chiral expansion, which amount to 0.65(21)% and -0.71(29)% respectively. Both corrections are one order of magnitude smaller than the known NLO ones, confirming the very good convergence of the chiral expansion and its reliability. Using the latest estimates for the light quark masses, the current uncertainty is dominated by the one of the low-energy constant ℓ_7. When combined with possible improvements on the light quark mass ratio and ℓ_7 from lattice QCD, the axion mass could be determined with per-mille accuracy as a function of the its decay constant.

Second, with the goal of predicting the numerical value of the axion decay constant (and therefore the axion mass) based on the axion DM relic abundance, we study the system of axion strings that forms in the early Universe if the Peccei-Quinn symmetry is restored after inflation. Understanding the dynamics of this system is a necessary condition for a reliable prediction of the axion DM abundance in the post-inflationary scenario. Using numerical simulations, we establish the existence of an asymptotic solution to which the system is attracted independently of the initial conditions. We study in detail the properties of this solution, including the average number of strings per Hubble patch, the distribution of loops and long strings, the way that different types of radiation are emitted, and the shape of the spectrum of axions produced. We find clear evidence of logarithmic violations of the scaling properties of the attractor solution. We also find that, while most of the axions are emitted with momenta of order Hubble, most of the axion energy density is contained in axions with energy of order the string core scale, at least in the parameter range available in the simulation. While such a spectrum would lead to a negligible number density of relic axions from strings when extrapolated to the physical parameter region, we show that the presence of small logarithmic corrections to the spectrum shape could completely alter such a conclusion. A detailed understanding of the evolution of the axion spectrum is therefore crucial for a reliable estimate of the relic axion abundance from strings.

List of publications

The content of this thesis is based on the following publications:

[1] M. Gorghetto and G. Villadoro, ‘Topological Susceptibility and QCD Axion Mass: QED and NNLO corrections,” JHEP **1903** (2019) 033 [arXiv:1812.01008 [hep-ph]].

[2] M. Gorghetto, E. Hardy and G. Villadoro, “Axions from Strings: the Attractive Solution,” JHEP **1807** (2018) 151 [arXiv:1806.04677 [hep-ph]].

During his PhD studies the author has also performed research on a follow-up of [2], and on the next-to-leading-order correction to the axion mass at finite temperature in eq. 55, in collaboration with the same authors (still work in progress). Partial results are available but left for future publications.

Contents

1 Introduction

With the discovery of the Higgs boson, the Standard Model (SM) of particle physics is in principle a consistent quantum field theory up to energies much larger than the electroweak (EW) scale. While the Higgs particle (or any other electroweak physics restoring perturbative unitarity) was a necessary requirement for the theory to be self-consistent at energies of the order of the EW scale, there is unfortunately no guarantee that new physics beyond the SM lies near the EW scale, nor at energy scales that will ever be probed by collider experiments.

In the absence of a clear indication that traditional collider searches will able to advance our knowledge towards a more fundamental theory of Nature, all the experimental observations and theoretical puzzles that contradict the SM expectations must be scrutinized seriously. First of all, the SM coupled to gravity cannot explain gravitational interactions at the quantum level, but unfortunately any effect of a theory of quantum gravity is likely to appear at energies close to the Planck mass and therefore not accessible with today's experiments. On the other hand, the astrophysical and cosmological observation of dark matter (DM), the baryon-antibaryon asymmetry of the Universe and the naturalness problem of some of the SM parameters could very well find an explanation by new physics lying at energies small enough to be detectable in the near future. It is therefore motivated to examine in detail the simplest models that can give a convincing answer to some (of all) of the mentioned issues, in particular highlighting the most possible model-independent experimental signatures.

Most of the ideas for new physics try to explain the unexpected size of three SM parameters: the cosmological constant, the Higgs mass and Strong-CP phase. While the first only arises when gravity is included, the last two are related to the SM only and could in principle find a particle physics explanation at low enough energies. The Higgs mass hierarchy is traditionally solved by new physics at the TeV scale (e.g. supersymmetry, composite Higgs, extra-dimensions), whose generic predictions start however to be in tension with the latest Run-2 LHC results. On the other hand, the simplest and most robust solution to the smallness of the Strong-CP phase, the QCD axion [1–7], is still largely unconstrained and is at the same time an excellent DM candidate [8–10]. Consequently, a QCD axion that makes up the entire measured DM relic abundance is one of the best motivated scenarios for physics beyond the SM. In addition, numerous experiments aimed at detecting axions are currently running or in development. These will be sensitive to a substantial proportion of the relevant parameter space, and, if an axion is discovered, they could measure its mass and couplings precisely (see e.g. [11, 12]). Finally, the fact axion-like particles seem to be a ubiquitous feature of String Theory compactifications [13] motivates even more the existence of this particle.

Since most of the axion experiments work on resonance, the possible discovery of the QCD axion is strongly related to the numerical value of its mass m_a. A solid theoretical prediction for m_a is therefore crucial to pin down the axion or to even possibly rule it out as a solution of the Strong CP problem. The main goal of this work is to make a step towards this prediction on the basis of the axion DM abundance produced in the early universe.

In the absence of significant sources of explicit breaking of the Peccei-Quinn (PQ) symmetry that gives rise to the axion, the axion mass can be in principle determined at high precision as a function of only one Lagrangian parameter, the axion decay constant f_a, which dictates the scale of spontaneous breaking of PQ. Since in the axion DM experiments the mass sensitivity is 10^{-6} (bounded by the velocity dispersion of the DM particles), in the case of a discovery m_a will be known with very high precision, which can be eventually translated into the same precision on f_a only if the relation between the two is known very well. When combined with measurements of the axion couplings and possibly the information of the axion relic abundance, such precision could be used to learn about the dynamics of new physics at much higher scales, as well as physics of the early universe including inflation, reheating and pre-BBN evolution. In chapter 2 we will therefore derive the relation between m_a and f_a in the most precise way available to date, making use of effective chiral perturbation theory techniques [14–16]. As we will see, next-to-next-to leading order (NNLO) corrections in the chiral expansion and electromagnetic corrections are crucial to estimate the axion mass with less than percent accuracy. Using the latest estimates for the light quark masses the current uncertainty is dominated by the one of the low-energy constant ℓ_7 of the next-to-leading order (NLO) chiral Lagrangian. When combined with possible improvements on the light quark mass ratio and ℓ_7 from lattice QCD, the computation will allow to determine the QCD axion mass as a function of f_a with per-mille accuracy.

A prediction for the actual numerical value of f_a (and therefore of m_a) could in principle come from the production of axions in the early universe. As we will discuss in chapter 3, the dynamics by which QCD axion DM is produced, and its final relic abundance, depends on the cosmological history of the Universe (see e.g. [17, 18]). In particular, if the PQ symmetry was broken after inflation, the axion field had initially random fluctuations over the present day observable Universe. Instead, if the PQ symmetry was broken during inflation and never subsequently restored, the axion field was initially homogeneous. In this case the axion relic abundance is incalculable because it depends on the value of the axion field after inflation in our particular Hubble patch, which is however not known.[1]

In this work we will be interested in the class of models in which the PQ breaking happens after inflation. This includes theories in which inflation happened at a scale above the axion decay constant f_a, and also those with inflation at a lower scale but which were reheated to a temperature above f_a. In this case, assuming a standard cosmological history, the relic abundance is calculable in terms of the the axion decay constant f_a due to the random initial conditions.[2] As a result, there is in principle a unique calculable prediction for the axion mass if it is to make up the complete DM density in such models, which, as mentioned, would be extremely valuable for experimental axion searches.

However, as described in chapter 4, in this scenario the mechanisms by which DM axions are produced are complex, and calculating the relic abundance is challenging [22–26]. The random initial axion field after PQ symmetry breaking leads to the formation of axion strings [27–29]. These are topologically stable field configurations that wind around the $U(1)$ vacuum manifold of the broken PQ symmetry as a loop in physical space is traveled. As we will discuss futher, interactions between strings are thought to result in a scaling regime where the length of string per Hubble volume, measured in Hubble lengths remains approximately of $\mathcal{O}(1)$ as the Universe expands [30–35]. To maintain such a regime the string network must release energy. This dominantly happens through the production of axions, which form a potentially significant fraction of the total relic abundance [36–43].

The string system persists until the temperature of the Universe drops to around the QCD scale, when the axion mass turns on and becomes cosmologically relevant, leading to the formation of domain walls, each one attached to strings. The subsequent dynamics depends on the anomaly coefficient between QCD and the PQ symmetry, N_W, which is equal to the number of minima that are generated in the axion potential [23, 44]. $N_W = 1$ corresponds to the scenario in which the domain walls are automatically unstable and decay, destroying the string network and releasing further DM axions in the process [23, 44, 45]. If $N_W > 1$ the domain walls are generically stable, and the model is not phenomenologically viable, unless further explicit breaking of the PQ symmetry is introduced [46–49]

---

[1] As we will discuss further in sec. 3, there is only an upper bound on the axion mass for which it can make up the full relic abundance, coming from isocurvature constraints [19–21].

[2] As we will discuss later the relation between the relic abundance and the mass of the axion might be affected for extreme choices of the model parameters.

(which might reintroduce the strong CP problem) or the $\mathbb{Z}_{N_W}$ symmetry is gauged [50].

In this work we only consider the string network before the axion mass turns on. An understanding of this stage of its evolution is crucial, both to calculate the relic abundance of axions produced at such times and to set the appropriate initial conditions when analyzing the system once the axion mass becomes relevant. The important properties of the network, which we aim to determine, include: the average length of string per Hubble volume, and the way that this is distributed in loops of different lengths; the rate of energy release into axions; and the spectrum of axions emitted. Due to the complex evolution and interactions of the strings, and later the domain walls, an accurate analytic calculation of the axion relic abundance appears hopeless in this scenario, and instead some form of numerical simulation is required. The most direct approach, which we will follow, is to simulate a complete UV theory that gives rise to the axion numerically on a discrete lattice [51].

As we will discuss at length in chapter 4, the main drawback of numerical simulations (as opposed to analytic estimates) is the impossibility to investigate the physical parameter region. This is essentially due to the huge hierarchy of scales in the system at the QCD crossover. On the one hand, strings have a core with microscopic size, of the order of $1/m_r$, where m_r is the mass of the heavy modes that UV complete the axion theory (usually of order f_a for $O(1)$ parameters). On the other hand, the Hubble horizon at the QCD crossover, which sets the average separation between strings, is a cosmological distance of order $H \sim \Lambda_{QCD}^2/M_P$. While the ratio between the two is $m_r/H \sim 10^{30}$ in the physically relevant region of parameter space, to avoid numerical artifacts simulations can only scan parameters such that $m_r/H \sim 10^3$.

Making physically relevant predictions about the system at the time of QCD crossover therefore requires that results from simulations are extrapolated over a vast difference in scale separations. What makes such an extrapolation not obviously hopeless is the possible existence of an attractor in the evolution, an understanding of which would allow a controlled extrapolation to be made. A key point of this work is that this is an extremely delicate process. In particular, we stress that a careful analysis of which features of the string network are being assumed to remain constant, or change, between $\log(m_r/H) = \log(10^3) \sim 6$ and $\log(m_r/H) = \log(10^{30}) \sim 70$ is required, and that naive extrapolations can lead to misleading results. More specifically, the number density of axions produced by the string network is very sensitive to the evolution of the axion energy density spectrum: as we will show, axion relic abundance will be determined by the late-time behavior of this observable and a small time-dependence in it could completely alter any result obtained in a small-scale simulation. It is also likely that a time-dependence on the spectrum leads to a nontrivial dependence of the subsequent string-domain wall dynamics on the numerical value of $\log(m_r/H)$ at $H = m_a$, which suggests that small-scale simulations might not be able to capture it properly.

Turning to the structure of this thesis: In chapter 2 we will review the Strong CP problem and the axion solution, and derive some low energy properties of this particle including its mass at high precision. In chapter 3 we describe how DM axions are produced in the early universe, in particular focusing on their non-thermal production through the misalignment mechanism. Chapter 4 is the core of the thesis and will be devoted to the production of axions from strings. In this chapter we will first give an overview of the dynamics of strings and domain walls in the early universe, and then discuss the evolution of the axion string system based on results from numerical simulations, including the possible extrapolation to the physical parameter region and possible predictions for the relic abundance. We will finally conclude in Chapter 5.

In appendix A we report the formulas for the result of the axion mass with the explicit quark mass dependence, suitable to be used in lattice simulation fits. In appendix B we give all the details of the numerical extraction of the values of the low energy constants used in the results for the axion mass. Additional technical details about axion string simulations may be found in Appendix C, and an extensive analysis of the systematic errors is given in Appendix D. In Appendix E we present a detailed analysis of the distribution of energy into different components, in Appendix F we provide further evidence for the existence of an attractor solution, and in Appendix G we give details of how we fit the parameters of the string network. Finally, in Appendix H we analyze whether the properties of the global strings that we simulate are converging to those of local strings.

2 The QCD axion mass at high precision

In this chapter we will provide a basic introduction on the motivation and the physics underlying the QCD axion, and describe its properties in the most possible model-independent way. In particular, using effective chiral perturbation theory techniques, we will derive several low-energy properties of this particle, including its potential and interactions with other light Standard Model fields. We will then turn our focus to the axion mass and determine its relation with the decay constant in the most precise way available to date, which takes into account NNLO corrections in the chiral expansion and the effect of the electromagnetic interactions.

2.1 The Strong CP problem

In a modern perspective the SM is considered an effective field theory valid up to (possibly very) high energies, and therefore at the renormalizable level its Lagrangian is expected to include all the dimension four or less operators compatible with its gauge symmetries. In particular, the gluon topological term

$$\mathcal{L}_{\theta_0} = \frac{\theta_0}{32\pi^2} G^a_{\mu\nu}\tilde{G}^{a\mu\nu}, \tag{1}$$

where θ_0 is a dimensionless parameter, $G_{\mu\nu}$ is the gluon field strength and $\tilde{G}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}G^{\rho\sigma}$ its dual, is an allowed classically marginal operator. Although $\mathcal{L}_{\theta_0}$ is gauge invariant, it breaks P and T (but preserves C) because $\epsilon_{\mu\nu\rho\sigma}$ a pseudotensor, and so a $\theta_0 \neq 0$ in eq. (1) would introduce extra (in principle unsuppressed) CP violation to the SM [5].

While in pure Yang–Mills the coefficient θ_0 parameterizes physically inequivalent theories, in the presence of fermions charged under $SU(3)_c$ (like our quarks) θ_0 is not physical because depends on how the phase of those fermion fields is chosen. Indeed, the quark field redefinition $q_L \to e^{i\alpha}q_L$ (where q_L are all the 4×3 left-handed Weyl spinors Q_L, u^c_R, d^c_R) is abelian and chiral and therefore anomalous under $SU(3)_c$, and modifies the path integral measure proportionally to the gluon topological term as

$$[d\overline{q}][dq] \to [d\overline{q}][dq] \exp\left(i\frac{12\alpha}{32\pi^2}G^a_{\mu\nu}\tilde{G}^{a\mu\nu}\right), \tag{2}$$

effectively shifting θ_0 in eq. (1) as $\theta_0 \to \theta_0 + 12\alpha$. The field redefinition above leaves the rest of the SM Lagrangian invariant, except for introducing a factor of $e^{2i\alpha}$ into the quark Yukawa matrices Y_u, Y_d which are the only non chiral-singlets[3] of the SM. As a result, only the combination $\theta = \theta_0 - \arg(\det(Y_u Y_d))$ is independent of the phase choice and therefore physical. Note that this holds only if all eigenvalues of Y_u and Y_d do not vanish[4], like in the SM. In particular, by properly choosing the phase α one can always assume (as we will do from now on) that all the contribution to θ comes from θ_0. Moreover θ and $\theta + 2\pi$ are physically indistinguishable, because a 2π shift of θ can be always reabsorbed by the redefinition $q \to e^{i\alpha}q$ of any of the quarks with $\alpha = -2\pi$, which does not change the Yukawa matrices.

The topological term in eq. (1) is actually a total derivative [52] because

$$G^a_{\mu\nu}\tilde{G}^{a\mu\nu} = \epsilon^{\mu\nu\rho\sigma}\partial_\mu\left(A^a_\nu G^a_{\rho\sigma} - \frac{1}{3}f_{abc}A^a_\nu A^b_\rho A^c_\sigma\right), \tag{3}$$

---

[3] Note that we cannot define a transformation law for the Higgs h so to leave both Yukawa terms $Y_u^{ij}u^c_{R\,i}Q_{L\,j}h^c + Y_d^{ij}d^c_{R\,i}Q_{L\,j}h$ invariant under the chiral field redefinition, because $h^c = \epsilon h^*$.

[4] If there were at least one vanishing eigenvalue, a redefinition with $\alpha = -\theta_0/2$ for the linear combination of quarks corresponding to the null eigenvalue would cancel θ_0 while not changing the rest of the Lagrangian, making θ_0 unphysical.

where A^a_μ is the gluon field, so it does not affect the classical equations of motion, whose solutions will be θ-independent. Moreover, it does not contribute to any order in perturbation theory, because a total derivative provides a Feynman rule proportional to the sum of the external momenta, whose insertion in any diagram vanishes by momentum conservation. Nevertheless, several SM observables such as the vacuum energy[5], the electric dipole moment of the neutron, the pion mass and the proton-neutron mass splitting have nontrivial θ-dependence. This dependence can be qualitatively understood as the contribution to the SM path integral of gauge configurations with nonvanishing topological charge $Q = \frac{1}{32\pi^2}\int d^4x G^a_{\mu\nu}\tilde{G}^{a\mu\nu}$, even if the expansion series around such solutions of the equations of motion are nonperturbative and badly divergent, and so in general not well defined.

As an example, we discuss the dependence on θ of the vacuum energy of QCD, $E(\theta)$, defined as

$$e^{-V_4 E(\theta)} = \int [d\phi] e^{-S[\phi] - i\frac{\theta}{32\pi^2}\int G^a_{\mu\nu}\tilde{G}^a_{\mu\nu}} \ , \tag{4}$$

where V_4 is the Euclidean 4-voulme and $S[\phi]$ is the Euclidean QCD action at $\theta = 0$ and the functional integration is over all the fields ϕ. First, $E(\theta)$ is periodic of period 2π, $E(\theta) = E(\theta + 2\pi)$, and is symmetric under $\theta \to -\theta$ as can be seen by redefining the fields to their CP conjugates $\phi \to \phi^{CP}$ and using $[d\phi] = [d\phi^{CP}]$ and $S[\phi] = S[\phi^{CP}]$ for QCD[6]. Moreover, $\theta = 0$ is the absolute minimum of $E(\theta)$ because

$$e^{-V_4 E(\theta)} = |e^{-V_4 E(\theta)}| \leq \int [d\phi] \left| e^{-S[\phi] - i\frac{\theta}{32\pi^2}\int G^a_{\mu\nu}\tilde{G}^a_{\mu\nu}} \right| = \int [d\phi] e^{-S[\phi]} = e^{-V_4 E(0)}, \tag{5}$$

where we used the Schwartz inequality. This implies $E(\theta) \geq E(0)$, which also means that the parity symmetry $\theta \to -\theta$ is not spontaneously broken (result known as Vafa-Witten theorem [53]).

To know the exact functional form of $E(\theta)$ and show that it is indeed non-trivally $E(\theta) \neq 0$, one would need to systematically evaluate the integral in eq. (4). In QCD this integral might be estimated through saddle point approximations around solutions of the classical equations of motion with finite action, including those with nonvanishing topological charge Q, called instantons [54, 55]. The finiteness of the action requires these solutions to be pure gauge at infinity, i.e. $A_\mu \to iU^\dagger \partial_\mu U$ for $|x|^2 \to \infty$ for some $U(\hat{x}) \in SU(3)_c$, and therefore they are identified by maps $U : S^3 \to SU(3)$, where the generic direction in Euclidean space is an element $\hat{x} \in S^3$. Maps of this kind that can be continuously deformed into each other (therefore leaving the action finite) fall into equivalence classes labeled by an integer $k \in \mathbb{Z}$ (because the homotopy group $\pi_3(SU(3)) = \mathbb{Z}$), and the topological invariant associated to each equivalence class of $U(\hat{x})$ is the topological charge itself [56], $Q = k$, which is so an integer [7]. The classical Euclidean action can be written as

$$\frac{1}{4g_s^2}\int G^a_{\mu\nu}G^a_{\mu\nu} = \frac{1}{8g_s^2}\int \left(G^a_{\mu\nu} \mp \tilde{G}^a_{\mu\nu}\right)^2 \pm \frac{8\pi^2}{g_s^2} Q \ , \tag{6}$$

and, in the class of functions with fixed positive (negative) Q, the minimum of the action is obtained for (anti)self-dual field configurations $G^a_{\mu\nu} = \pm\tilde{G}^a_{\mu\nu}$ and is $\frac{8\pi^2}{g_s^2}|Q|$. Note that fields satisfying the (anti)self-duality condition are automatically solutions of the classical equations of motion because

---

[5] The vacuum energy is a physical observable only when the theory is coupled to gravity.

[6] Actually the full SM Lagrangian (and possibly new physics beyond the SM) is not invariant under CP due to the CKM phase, so $S_{SM}[\phi] \neq S_{SM}[\phi^{CP}]$, but the breaking of CP is small enough that the result holds to high precision.

[7] Indeed, although Q seems to depend on A_μ in the whole space, since $G_{\mu\nu}\tilde{G}_{\mu\nu}$ is a total derivative Q can be written as a surface integral at spatial infinity and so it depends only on A_μ at $|x|^2 \to \infty$, i.e. on $U(\hat{x})$.

$D_\mu G^a_{\mu\nu} = \pm D_\mu \tilde{G}^a_{\mu\nu} = 0$ thanks to the Bianchi identity, and are called k-instantons $A^{(k)}_\mu$ if $Q = k$. In particular, single instanton solutions $A^{(1)}_\mu$ can be found by integrating the equation $G^a_{\mu\nu} = \tilde{G}^a_{\mu\nu}$ with the condition $Q = 1$ and there is an infinite number of them depending on 5 parameters, the position x^μ_0 and the size ρ, because of the translational and scale invariance of the classical action.

The path integral in eq. (4) is then written as the sum of saddle points around $A^{(k)}_\mu$. From eq. (6) the leading order contribution of $A^{(\pm k)}_\mu$ is given by $\exp(-\frac{8\pi^2}{g_s^2}|k| - i|k|\theta) + c.c$, and for g_s small enough the contributions from higher and higher instantons are more and more suppressed and the leading dependence on θ is provided by $k = \pm 1$, on which we will focus. One can show that calculating the Jacobian of the fluctuations around the saddles $A^{(\pm 1)}_\mu$ leads to [54]

$$e^{-V_4 E(\theta)} \supset \int \frac{d\rho}{\rho^5} A(g_s) e^{-\frac{8\pi^2}{g_s^2(1/\rho)} - i\theta} + c.c. \ , \tag{7}$$

where $A(g_s) \propto g_s^{-8}$, and the integration in $d\rho$ as well as the factor ρ^{-5} come from the integration over the zero modes of the kinetic operator and can be intuitively understood by the fact that one has to integrate over all gauge field configurations (so over all ρ's) and by dimensional analysis. The function $g_s(\mu)$ is the one-loop running coupling and in eq. (7) is evaluated at $\mu = 1/\rho$, which is intuitive because $1/\rho$ is the characteristic energy scale of the instanton. As mentioned, the introduction of fermion fields with nonvanishing Yukawa (or nonvanishing mass) makes the coefficient of the topological term depend on their phase, and the phase-independent combination that must enter in eq. (7) is $\exp(-i\theta_0)\det(\rho M_f)$, where $M_f = vY_{u,d}$ is the (in general complex) fermion mass matrix and ρ appears for dimensional reasons. With a field redefinition we can make $\det(\rho M_f)$ real and substitute θ_0 with θ, so that the only phase in eq. (7) comes from $e^{-i\theta}$. In the end, the dependence on ρ and θ in the integrand on eq. (7) is

$$\frac{1}{\rho^5} e^{-\frac{8\pi^2}{g_s^2(1/\rho)}} \det(\rho M_f) \cos\theta = \det M_f \, (\Lambda_{QCD})^{\frac{11}{3}N - \frac{2}{3}N_f} \, \rho^{\frac{11}{3}N + \frac{1}{3}N_f - 5} \cos\theta \ , \tag{8}$$

where we used the one-loop running $g_s(\mu) = \frac{8\pi^2}{\beta_0 \log(\mu/\Lambda_{QCD})}$ with $\beta_0 = 11N/3 - 2N_f/3$, being $N = 3$ for QCD and N_f the number of light quarks. The integral $\int d\rho \, \rho^{6+N_f/3}$ is infrared divergent and dominated by larger and larger instanton sizes, in particular by sizes $\rho \gtrsim \Lambda_{QCD}^{-1}$ for which the coupling $\exp(-8\pi^2/g_s^2(1/\rho)) = (\rho\Lambda_{QCD})^{\beta_0}$ is in the nonperturbative regime[8]. The result is that the saddle point approximation about instanton solutions is not under control and cannot be systematically evaluated, and in the following we will indeed show that in the special case of QCD with light quarks the actual form of $E(\theta)$ is different from the simple $\cos\theta$, as instead eq. (8) would suggest. In any case, this discussion shows that instanton configurations, though incalculable, are likely to contribute to the path integral and give rise to a non-trivial dependence on θ to the vacuum energy.

Since $E(\theta) = E(-\theta)$, any effect of θ appears in the vacuum energy at leading order at least through θ^2 and rapidly disappears for $\theta \ll 1$. The effect of a non-zero θ-parameter is most visible in parity-odd observables, like the electric dipole moment (EDM) of the neutron d_n, which provides the strongest bound on θ to date. The neutron EDM represents the effective coupling between the neutron n and the photon $F_{\mu\nu}$ due to the different distribution of positive and negative electric charges inside the neutron and is formally defined as the effective operator $d_n \, \overline{n}\gamma^{\mu\nu}\gamma_5 n \, F_{\mu\nu}$. Since it is odd under parity,

---

[8] An infrared cut-off to the ρ integral would make the integral converge. As we will discuss in sec. 2.5, this cut-off could be provided by the inverse temperature if we consider the finite-temperature version of the theory, or by the vacuum expectation value of an Higgs field if the gauge group is Higgsed [54].

its coefficient d_n will get quantum-mechanically a non-trivial dependence on θ. This can be very roughly estimated as [57, 58]

$$d_n \sim e|\theta|\frac{m_\pi^2}{m_n^3} \sim 10^{-2}|\theta|\ e \cdot \mathrm{GeV}^{-1}\ , \tag{9}$$

where m_n and m_π are the neutron and pion masses. Eq. (9) holds because the dependence on θ should disappear if any of the quark Yukawa (or masses) vanishes, and this requires $d_n \propto M_q \propto m_\pi^2$, or if the electric change e vanishes. The additional powers of m_n can be understood by dimensional analysis and considering that $m_n \sim$ GeV is the only other scale in the problem. The most recent experimental bound [59] on $|d_n| < 3.6 \cdot 10^{-26}\ e \cdot \mathrm{cm} = 1.8 \cdot 10^{-12}\ \mathrm{e} \cdot \mathrm{GeV}^{-1}$ constrains θ to be very tiny:

$$\theta \lesssim 10^{-10}\ . \tag{10}$$

According to the 't Hooft naturalness criterion, a dimensionless parameter like θ can be very small if a symmetry is restored when the parameter is set to zero (because in this case the quantum corrections are proportional to the parameter itself). This would be true in QCD alone, since in that case CP (and P) would be recovered. However, CP is already broken in the full SM by the CKM phase and is not restored in the limit $\theta \to 0$, which does not guarantee the naturalness of this parameter[9].

Moreover, θ is the sum of two contributions, one of which comes from Yukwava matrices that are known to have an $O(1)$ CP violating phase, and so the natural expectation is $\theta \sim O(1)$, or just a bit smaller not to change significantly nuclear physics[10]. The question why θ is instead so small is the 'Strong CP problem'.

2.2 The Peccei-Quinn solution

Among the various solutions of the Strong CP problem proposed so far, the so-called Peccei-Quinn (PQ) solution [1] is arguably the most robust and model independent, because it only relies on the requirement of a new $U(1)$ global symmetry with a chiral anomaly under $SU(3)_c$ and on the low energy properties of QCD - in particular on its vacuum structure.

The basic idea is that θ-angle is unphysical if there exists a global and continuous transformation, parameterized by a parameter α, that leaves the UV theory is invariant up to a shift of the gluon topological term $\alpha G^a_{\mu\nu}\tilde{G}^a_{\mu\nu}$. The easiest way to realize this setup is to introduce a $U(1)_{PQ}$ global symmetry acting on the fields as $\phi \to \phi_\alpha$ under which the UV action is invariant, $S_{UV}[\phi_\alpha] = S_{UV}[\phi]$, but anomalous under $SU(3)_c$, so that the path integral measure changes as $[d\phi_\alpha] = [d\phi]\exp(i\alpha G^a_{\mu\nu}\tilde{G}^{a\mu\nu})$.

Unfortunately, such a global symmetry is not present in the SM (not even at the renormalizable level). Indeed the $U(3)^5$ global flavor symmetry acting on the SM Weyl spinors $Q_L, u^c_R, d^c_R, l_L, e_R$ which the SM enjoys in the limit of vanishing Yukawa $Y_u = Y_d = Y_e = 0$ is explicitly broken by Y_u and Y_d (and Y_e) into the baryon number and the individual lepton numbers, $U(1)_B \times U(1)_{L_i}$, all of which are not anomalous under $SU(3)_c$. If one of the quark Yukawa's vanished (for instance the up-quark Yukawa), however, the residual symmetry group would be enhanced to $U(1)_B \times U(1)_{L_i} \times U'(1)$ where

---

[9] This implies that, even if θ is set to zero, it will be generated in the effective action proportionally to the CKM phase (which means θ runs with the energy scale). However, it has been shown [60, 61] that the radiative corrections to θ are extremely suppressed, at least 6 orders of magnitude below the bound in eq. (10). As a result, θ might be considered technically natural for all practical purposes in the SM. Note however that new CP violating physics could in general produce radiatively a θ angle bigger than 10^{-10}, and a tuning will be necessary.

[10] Anthropic arguments might be invoked to explain why θ has to be be slightly smaller than $O(1)$, $\theta \sim 10^{-2}$, but they are not enough to explain the huge suppression in eq. (10).

$U'(1)$ is the (anomalous) chiral phase redefinition of the quark with vanishing Yukawa, and it would be possible to identify $U'(1)$ with $U(1)_{PQ}$.

As a consequence, the PQ solution requires new physics beyond the SM. Whatever this new physics is, the 't Hooft anomaly matching condition [62] gives us an handle on how the low energy theory should look like. Indeed, for consistency of the theory, if there is an anomalous global symmetry in the UV, the low energy theory - even when the gauge group confines - should include massless fermions with suitable quantum numbers to reproduce the same anomaly coefficient (in the case of $U(1)_{PQ}$ they should be colored). Moreover, if the symmetry with the global anomaly is spontaneously broken, the Nambu-Golstone boson (NGB) associated to the breaking can play the role of the massless fermions which are therefore not necessary[11]. Since massless colored fermions are actually experimentally excluded, the only option is that $U(1)_{PQ}$ is spontaneously broken and in the low energy theory its NGB, called (QCD) *axion* $a(x)$, reproduces the anomaly by transforming nonlinearly under PQ as $a \to a + \alpha f_a$ and coupling to the gluon topological term as

$$\frac{\alpha_s}{8\pi}\frac{a}{f_a}G^a_{\mu\nu}\tilde{G}^{a\mu\nu} \ , \tag{11}$$

where $\alpha_s = g_s^2/4\pi$ and we used canonically normalized kinetic terms for the gauge field. The axion decay constant f_a is defined by eq. (11) and is in general related to the scale of spontaneous breaking of $U(1)_{PQ}$. The rest of the low energy Lagrangian is invariant under the shift $a \to a + \alpha f_a$ in order to reproduce the anomaly in the UV. Note that the axion could be part of a fundamental scalar or, in theories where the PQ symmetry only acts on fundamental fermions, a composite state made of those fermions[12] [63, 64]. Moreover, since the coupling in eq. (11) breaks the shift symmetry, a mass for the axion will be generated proportionally to f_a^{-1} by quantum corrections, and the axion will be a pseudo NGB (pNGB).

At energies below the scale of PQ breaking, the SM+axion effective Lagrangian will be [65]

$$\mathcal{L}_{IR} = \mathcal{L}_{SM}^{\theta=0} + \theta\frac{\alpha_s}{8\pi}G^a_{\mu\nu}\tilde{G}^{a\mu\nu} + \frac{1}{2}(\partial_\mu a)^2 + \frac{\alpha_s}{8\pi}\frac{a}{f_a}G^a_{\mu\nu}\tilde{G}^{a\mu\nu} + \cdots \ , \tag{12}$$

where the dots represent additional terms that do not break the shift symmetry (and proportional to inverse powers of f_a). While the coupling to gluon eq. (12) is model independent and can be argued on the basis of anomaly matching only, the rest of the low energy Lagrangian depends on the particular model under consideration. Couplings that can appear at order f_a^{-1} in the left hand side (l.h.s.) of (12) include

$$\frac{g^0_{a\gamma\gamma}}{4}aF_{\mu\nu}\tilde{F}^{\mu\nu} + \frac{\partial_\mu a}{2f_a}\sum_i c_i^0\overline{f}_i\gamma^\mu\gamma_5 f_i, \tag{13}$$

where $g^0_{a\gamma\gamma} \propto f_a^{-1}$ represents the coupling to photons[13] and the second term is the coupling to the fermion axial current, where f_i could be either a quark or a lepton. Note that in general there could be

---

[11] This is what actually happens in QCD for the axial symmetry spontaneously broken by the quark condensate and anomalous under $U(1)_{\rm em}$, whose anomaly is reproduced at energies lower than Λ_{QCD} by the coupling of the π_0 through the Wess-Zumino-Witten term.

[12] This last case is indeed what happens for the π_0 in QCD.

[13] More precisely, at energies just below the PQ breaking scale the relevant coupling is with the $SU(2)_L \times U(1)_Y$ gauge bosons. All these couplings do not break the shift symmetry because the terms $\theta_W W_{\mu\nu}\tilde{W}^{\mu\nu}$ and $\theta_B B_{\mu\nu}\tilde{B}^{\mu\nu}$ can be always redefined away in the SM by doing a baryon number transformation, which leaves the SM lagrangian invariant and is anomalous under $SU(2)_L \times U(1)_Y$ since left and right-handed quarks have different quantum numbers under this group.

an explicit non-derivative coupling of the axion to quarks in the quark mass matrix $\overline{q}e^{ic_q\gamma_5 a/f_a}M_q q$, but it can be rotated away through the field redefinition $q \to e^{-(c_q/2)i\gamma_5 a/f_a}q$ that will change the axion-gluon coupling (due to the anomaly) and the axion-quark couplings (due to the quark kinetic term). In particular, f_a will be defined by eq. (11) in the frame where $c_q = 0$. Although the coefficients $g^0_{a\gamma\gamma}$ and c^0_i in eq. (13) are model dependent, their value is important experimentally because they allow to probe and possibly detect the QCD axion. Note moreover that all axion couplings are proportional to inverse powers of f_a, so the axion rapidly becomes invisible if $f_a \gg v_{EW}$.

Astrophysical observations based the neutrino emission from the supernova 1987A give a lower bound on $f_a \gtrsim 10^8$ GeV [66]. On the other hand, an upper bound to $f_a \lesssim 2 \times 10^{17}$ GeV comes from black hole superradiance [13, 67]. Current direct experimental bounds on f_a from the ADMX experiment (which relies on the axion being DM) exclude a region of f_a around 10^{12} GeV for large enough couplings to photons [68]. Within the quoted bounds f_a can be thought as a free parameter, but, as we will see chapters 3 and 4, f_a will be bounded from above by DM overproduction depending on the cosmological scenario.

As an example, the KSVZ [4, 5] and DFSZ [6, 7] models are the two simplest UV-complete models implementing the PQ solution.

- In the KSVZ model the anomalous $U(1)_{PQ}$ is obtained by introducing N new quarks charged under $SU(3)_c$ and an additional complex scalar field Φ, acquiring an expectation value v and giving a mass of the same order to the new quarks. At low energies, the axion is the angular part of Φ and its decay constant is $f_a = v/N$, while the couplings in eq. (13) are $g^0_{a\gamma\gamma} = \frac{E}{N}\frac{\alpha_{\rm em}}{2\pi f_a}$ where $E = 2\,{\rm Tr}[Q_{PQ}Q^2_{\rm em}]$ (often called anomaly coefficient), being Q_{PQ} and $Q_{\rm em}$ the charge matrices of the new fermions under $U(1)_{PQ}$ and $U(1)_{\rm em}$, and $c^0_i = 0$ because the SM quarks are not charged under PQ.

- In the DFSZ model the scalar sector of the SM is enlarged with a 2HDM and an additional complex scalar field, which allows the anomalous $U(1)_{PQ}$ to be generated by the SM fermions themselves and to be spontaneously broken at $f_a \gg v_{EW}$ to escape observational bounds[14]. The axion will be a linear combination of the phase of the scalar fields and $g^0_{a\gamma\gamma}$ has the same form as in the previous model but with $E/N = 8/3$. The coupling to SM fermions is non-vanishing and given by $c^0_u = \frac{1}{3}\cos^2\beta$ and $c^0_d = c^0_e = \frac{1}{3}\sin^2\beta$ for up, down and electron, where $\tan\beta$ is the ratio between the two Higgs vevs.

In the following we will take a model-independent approach in the study of the axion properties at low energies, and parameterize our ignorance about the UV physics via the coefficients in eq. (13).

Note that the axion effective potential $V(a)$ at the lowest order in f_a^{-1}, given by $e^{iV_4 V(a)} = \int [d\phi \neq a] e^{i\int d^4x \mathcal{L}_{IR}(\phi)}$ with $a = $ const, is instead model independent, since $a = $ const in $\mathcal{L}_{IR}$ of eqs. (12) and (13) leaves only the term that does not break the shift symmetry, i.e. axion-gluon coupling. Moreover, given that θ can be removed via the field redefinition $a \to a - \theta f_a$, from the definition of vacuum energy $E(\theta)$ in eq. (4) written in Minkowski coordinates it is easy to see that the axion potential coincides with the vacuum energy, $V(a) = E(a/f_a)$. This in turn shows that $V(a)$ is an even function with period $2\pi f_a$ and minimized at $a = 0$, solving the Strong CP problem (this is of course a consequence of the initial requirement of PQ invariance up to the gluon topological term).

---

[14] Adding only the 2HDM would still be enough to obtain an anomalous $U(1)_{PQ}$, but f_a would be of the order of v_{EW}, as originally observed in [2, 3, 69].

2.3 The QCD axion at low energies

As reviewed in the previous section, in the absence of massless colored fermions all PQ models should have an axion in their low energy spectrum, and at energies below the PQ breaking scale the effective Lagrangian will be given by eqs. (12) and (13). Studying the properties of this particle at energies below the QCD scale, in particular its mass and couplings with light SM fields, might seem challenging since QCD runs into strong coupling, and non-perturbative methods might be required to understand the effect of the axion-gluon interaction in eq. (12). As we will see, fortunately, the possibility of phase redefining the axion-gluon coupling into the quark mass matrix will allow us to embed the axion in the low-energy theory of QCD - the chiral Lagrangian. The axion properties can be then inferred at high precision in terms of the low energy parameters of the chiral Lagrangian itself, which are actually known from experiments involving QCD states only or from lattice QCD. In this section we will derive the axion-pion effective action, and from it the axion low-energy properties.

Without loss of generality, the axion+SM Lagrangian at energies just above the QCD confinement scale can be written as [65, 97]

$$\mathcal{L}_{QCD} = \frac{1}{2}(\partial_\mu a)^2 + \overline{q} i\gamma^\mu D_\mu q - (\overline{q}_L M_a q_R + \text{h.c.}) + \frac{\partial_\mu a}{2f_a}\sum_i c_i \overline{q}_i \gamma^\mu \gamma_5 q_i + \frac{1}{4} g_{a\gamma\gamma}\, a F_{\mu\nu}\tilde{F}^{\mu\nu} + \dots\,, \tag{14}$$

where we defined $M_a = e^{iQ_a \frac{a}{2f_a}} M_q e^{iQ_a \frac{a}{2f_a}}$ with $M_q = \text{diag}(m_u, m_d, \dots)$ the quark mass matrix for $N_f = 2, 3$ flavors and Q_a an arbitrary $N_f \times N_f$ matrix with unit trace $\text{Tr}[Q_a] \equiv \langle Q_a \rangle = 1$, taken to be diagonal for simplicity. Moreover, $c_i = c_i^0 - (Q_a)_{ii}$ or in matrix form $c = c^0 - Q_a$ and $g_{a\gamma\gamma} = g^0_{a\gamma\gamma} - \frac{\alpha_{em}}{2\pi f_a} 6\langle Q_a Q_{\text{em}}^2 \rangle$, where $Q_{\text{em}} = \text{diag}\,(2/3, -1/3, \dots)$ is the matrix of quark charges.

The Lagrangian in eq. (14) is straightforwardly obtained from eqs. (12) and (13) by performing the field redefinitions $a \to a - \theta f_a$ and $q \to e^{i\gamma_5 Q_a \frac{a}{2f_a}} q$. The former, being a symmetry, eliminates the θ-term from the Lagrangian, while the latter, as mentioned below eq. (13), eliminates the axion-gluon coupling while introducing a non-derivative coupling of the axion to quarks (as well as modifying $g^0_{a\gamma\gamma}$ because the field redefinition is in general anomalous under $U(1)_{\text{em}}$). The dots in eq. (14) are terms not dependent on axion and quark fields and include higher dimensional operators obtained integrating out the Higgs sector of the SM and heavy quarks, while $D_\mu = D^g_\mu - ieA_\mu Q_{\text{em}}$ is the $SU(3)_c \times U(1)_{\text{em}}$ gauge covariant derivative.

Below the QCD scale the relevant degrees of freedom are the pions (and kaons), which are the pNGBs of the global $SU(N_f)_L \times SU(N_f)_R$ flavor symmetry $q_L \to L q_L$ and $q_R \to R q_R$ that the Lagrangian in eq. (14) enjoys in the limit of vanishing quark masses and electric change (and $c_i = 0$). This symmetry is spontaneously broken (by the quark condensate) to its diagonal subgroup, and the CCWZ coset construction [70, 71] can be employed to build the effective Lagrangian for the Goldstone modes. These are embedded into the matrix $U = e^{i\Pi/F}$, where F is a constant (conventionally called F_0 for $N_f = 3$) and Π parameterizes the coset as

$$\Pi_{N_f=2} = \begin{pmatrix} \pi^0 & \sqrt{2}\pi^+ \\ \sqrt{2}\pi^- & -\pi^0 \end{pmatrix}\,, \qquad \Pi_{N_f=3} = \begin{pmatrix} \pi^0 + \eta/\sqrt{3} & \sqrt{2}\pi^+ & \sqrt{2}K^+ \\ \sqrt{2}\pi^- & -\pi^0 + \eta/\sqrt{3} & \sqrt{2}K^0 \\ \sqrt{2}K^- & \sqrt{2}\overline{K}^0 & -2\eta/\sqrt{3} \end{pmatrix}\,, \tag{15}$$

and transform under the full group as $U \to LUR^\dagger$. In the presence of nonvanishing M_a and Q_{em}, the global symmetry is formally recovered if M_a and Q_{em} are treated as external sources that transform as spurions under the full symmetry group. In particular, writing the term $eA_\mu \overline{q}\gamma^\mu Q_{\text{em}} q$ in the covariant derivative of eq. (14) as $A_\mu \bar{q}_L \gamma^\mu Q_L q_L + A_\mu \bar{q}_R \gamma^\mu Q_R q_R$ (we will set $Q_L = Q_R = eQ_{\rm em}$ eventually), the symmetry is restored for $c_i = 0$ if

$$M_a \to L M_a R^\dagger \,, \quad Q_L \to L Q_L L^\dagger \,, \quad Q_R \to R Q_R R^\dagger \,. \tag{16}$$

Moreover, in the presence of nonzero c_i, the axial current can be embedded into the quark covariant derivative as

$$\hat{D}_\mu q = (D_\mu + i\gamma_5 \alpha_\mu) q \,, \qquad \text{with} \qquad \alpha_\mu = \frac{\partial_\mu a}{2f_a} \operatorname{diag}(c_u, c_d, \dots) = \frac{\partial_\mu a}{2f_a} c \,, \tag{17}$$

and the symmetry is restored if the global $SU(N_f)_L \times SU(N_f)_R$ transformation is promoted to local and the axial current α_μ transforms as a spurion as

$$\begin{aligned} \alpha_\mu &\to L \alpha_\mu L^\dagger + i\partial_\mu L L^\dagger \\ \alpha_\mu &\to R \alpha_\mu R^\dagger - i\partial_\mu R R^\dagger \,. \end{aligned} \tag{18}$$

As a result, the generating functional

$$\mathcal{Z}[M_a, Q_L, Q_R, \alpha_\mu] = \int [d\bar{q}][dq][dG] e^{i\int d^4x \, \mathcal{L}_{QCD}} \tag{19}$$

is invariant under the local transformations in eqs. (16) and (18). Crucially, the same functional can be rewritten as a path integral over the pions Π of the axion-pion effective Lagrangian $\mathcal{L}_{\rm EFT}$, i.e. $\mathcal{Z}[M_a, Q_L, Q_R, \alpha_\mu] = \int [d\Pi] \exp\left(i \int \mathcal{L}_{\rm EFT}\right)$, and the symmetries in eqs. (16) and (18) will therefore strongly constrain the form of $\mathcal{L}_{\rm EFT}$. In writing the functional $\mathcal{Z}$ we are regarding the axion as a non-dynamical external field, and $\mathcal{L}_{\rm EFT}$ will therefore not take into account the effect of axion loops. However, since every axion interaction is proportional to f_a^{-1}, diagrams containing these loops are suppressed by two additional powers of f_a with respect to the the same diagrams without axion loops and they are not relevant when working at the lowest order in f_a^{-1} for the observable under consideration. In any case, next to leading order corrections in f_a^{-1} would be affected by the higher dimensional operators in the dots of eq. (14).

$\mathcal{L}_{\rm EFT}$ will be the most general Lorentz-invariant Lagrangian constructed out of U and the external sources, and invariant under the local symmetry group. If we are interested to low enough energies and the external sources are small, it makes sense to expand $\mathcal{L}_{\rm EFT}$ in powers of derivatives of U and of external sources. This expansion defines the chiaral Lagrangian, and its coefficients are called low-energy constants (LECs) and can be fitted from experimental data (or lattice QCD). The order of the expansion is conventionally classified by powers of the momentum p: in particular ∂_μ carries one power of momentum, and it has been shown (see e.g. [15, 112, 113]) that a consistent expansion requires M_a to be considered of $O(p^2)$, and $Q_{L,R}$ and α_μ of $O(p)$. When $Q_{L,R} = eQ_{\rm em}$ appear, it is sometimes conventional to quote $O(e)$ instead of $O(p)$.

Due to Lorentz invariance the first non-trivial term in $\mathcal{L}_{\rm EFT}$ is of $O(p^2)$ and reads [14–16, 112, 113]

$$\mathcal{L}_{p^2} = \frac{F^2}{4} \left[\left\langle D_\mu U^\dagger D^\mu U \right\rangle + 2B \left\langle U^\dagger M_a + M_a^\dagger U \right\rangle + ZF^2 \left\langle Q_R U Q_L U^\dagger \right\rangle \right] \,, \tag{20}$$

where F is the same constant that appears in $U = e^{i\Pi/F}$ to have canonically normalized kinetic terms, and B and Z are two additional LECs (conventionally called B_0 and Z_0 for $N_f = 3$). The derivative in eq. (20) is defined by $D_\mu U = \partial_\mu U - i\{\alpha_\mu, U\}$ and is covariant thanks to eq. (18). Note that $\mathcal{L}_{\rm EFT}$ contains also the axion kinetic term and the coupling to photons of eq. (14) that just drop out of $\mathcal{Z}$.

The first term in $\mathcal{L}_{p^2}$ includes the pion derivative self-interactions and the axion-pion derivative interactions. On the other hand, the axion and axion-pion nonderivative interactions are all encoded in the second term that depends only on the quark masses, thanks to the fact that we could fully reabsorb the nonderivative axion coupling into the quark mass term only. In particular, the axion-pion potential, including axion and pion masses, is contained in this second term. Finally, the last term when evaluated for $Q_L = Q_R = eQ_{\rm em}$ provides an electromagnetic splitting between the masses of $\pi^\pm$ and π^0 by contributing to the $\pi^\pm$ potential, but does not affect the axion and π^0 potentials because they are neutral under $U(1)_{\rm em}$. However, as we will see, higher orders in e^2 will affect also these neutral particles.

2.3.1 The axion potential

By expanding the Lagrangian in eq. (20) it is straightforward to calculate the pion and axion potentials at the lowest order in the chiral expansion. As noted in [65], there is in general a mass mixing between axion and π^0 because

$$\mathcal{L}_{p^2} \supset -\frac{a}{2f_a} BF \, {\rm Re} \langle \Pi \{Q_a, M_q\} \rangle \ , \tag{21}$$

which can be avoided at tree level by choosing $Q_a = M_q^{-1}/\langle M_q^{-1}\rangle = (m_u + m_d)^{-1}\,{\rm diag}(m_d, m_u)$, as we will do from now on. We will express the results in terms of the physical pion masses and decay constant f_π, this last one defined by $\langle 0|a_\mu(0)|\pi(p)\rangle = i\sqrt{2}p_\mu f_\pi$ in pure QCD, i.e. at $\alpha_{\rm em} = e^2/(4\pi) = 0$, being $a_\mu(x)$ an axial external current and $\pi = \pi^\pm, \pi^0$. They are given in terms of the LECs by

$$\begin{aligned} m_{\pi^0}^2 &= B(m_u + m_d) \ , \\ m_{\pi^\pm}^2 &= B(m_u + m_d) + 2Ze^2F^2 \ , \\ f_{\pi^\pm} &= f_{\pi^0} = F \ , \end{aligned} \tag{22}$$

and at this order the decay constants of all pions coincide, $f_{\pi^\pm} = f_{\pi^0} = f_\pi$. The axion-neutral pion potential around the vev $\langle \pi^\pm \rangle = 0$ (this must vanish because $U(1)_{\rm em}$ is not spontaneously broken) is

$$V(a, \pi^0) = -m_{\pi^0}^2 f_\pi^2 \sqrt{1 - \frac{4m_u m_d}{(m_d + m_u)^2} \sin^2\left(\frac{a}{2f_a}\right)} \cos\left(\frac{\pi^0}{f_\pi} - \phi_a\right) \tag{23}$$

with

$$\tan\phi_a = \frac{m_u - m_d}{m_u + m_d} \tan\left(\frac{a}{2f_a}\right) \ . \tag{24}$$

This shows that π^0 gets the nontrivial vev $\langle \pi^0 \rangle = f_\pi \phi_a$, and it can be integrated out leading to the axion potential [72]

$$V(a) = -m_{\pi^0}^2 f_\pi^2 \sqrt{1 - \frac{4m_u m_d}{(m_d + m_u)^2} \sin^2\left(\frac{a}{2f_a}\right)} \ . \tag{25}$$

As expected, $V(a) = E(a/f_a)$ is an even and periodic function with minimum at zero, but the dependence on a and on the quark masses is different from the simple $\det M_q$ and $\cos(a/f_a)$ estimated in the instanton calculation in eq. (8). Indeed, the dependence is nonanalytic in the quark masses (though in both cases disappears when at least one mass vanishes) and the potential contains higher harmonics. Contrary to the instanton calculation, the chiral expansion can be systematically evaluated and is under control: two flavor chiral perturbation theory converges very fast since the expansion parameter $m_u/m_s \sim 5\%$ is small, and indeed it was shown in [97] that the NLO corrections (which are of $O(p^4)$) to eq. (25) are only few percent.

2.3.2 Axion-pion interactions

From the first and second term of eq. (20) one can also extract the axion-pion derivative and non-derivative interactions, which determine for example the rate of the process $\pi\pi \to \pi a$, in principle contributing to production of thermal axions[15] in the early universe. These interactions have been determined for the first time in ref. [73] for hadronic axions, for which the model dependent coefficients c_i^0 vanish. The leading interactions are proportional to f_a^{-1} and do not get any contribution from the nonderivative term of eq. (20), because

$$\left\langle e^{i\frac{a}{2f_a}Q_a} M_q e^{i\frac{a}{2f_a}Q_a} e^{-i\frac{\Pi}{F}} + \text{h.c.} \right\rangle \supset -2\frac{a}{f_a}\left\langle M_q Q_a \Pi^3 \right\rangle \tag{26}$$

vanishes since $M_q Q_a \propto \mathbb{1}$ and $\langle \Pi^{2n+1} \rangle = 0$. The derivative interactions are

$$C_{a\pi}\frac{\partial_\mu a}{f_a f_\pi}\left(\partial^\mu \pi_+ \pi_- \pi_0 + \partial^\mu \pi_- \pi_+ \pi_0 - 2\pi_+ \pi_- \partial^\mu \pi_0\right) , \tag{27}$$

where

$$C_{a\pi} = \frac{c_d - c_u}{3} = \frac{c_d^0 - c_u^0}{3} + \frac{1-z}{3(1+z)} . \tag{28}$$

Being derivative, these interactions are model dependent, but are proportional only to the traceless part of $c = \text{diag}(c_u, c_d, \dots)$ because

$$\left\langle D_\mu U^\dagger D^\mu U \right\rangle \supset 2\text{Re}\left\langle -i\partial^\mu U^\dagger \{c, U\} \right\rangle \frac{\partial_\mu a}{2f_a} \tag{29}$$

vanishes if $c \propto \mathbb{1}$, since $U^\dagger \partial_\mu U$ can be written in terms of commutators of Π and so is traceless. As shown in [74, 75] and reviewed in [97], only the singlet part of the quark axial current acquires an anomalous dimension. This means that only the singlet combination of coefficients $c_\Sigma = \frac{1}{N_f}\sum_{i=1}^{N_f} c_i$ runs with the energy scale, while the $N_f - 1$ independent combinations $c_i - c_\Sigma$, i.e. the traceless part of c, do not run. In particular, the coefficients c_u^0 and c_d^0 in $C_{a\pi}$ in eq. (28) can be calculated at an arbitrary energy scale (e.g. at f_a where they are generated).

2.4 Topological Susceptibility of QCD and Axion Mass

As mentioned at the end of sec. 2.2, at leading order in f_a^{-1} the axion potential $V(a)$ coincides with the vacuum energy $E(a/f_a)$. Therefore at this order the axion mass can be related to the topological susceptibility of QCD χ_{top} by $m_a^2 = \chi_{\text{top}}/f_a^2$, where χ_{top} is defined as $\chi_{\text{top}} = d^2\mathcal{Z}(\theta)/d\theta^2|_{\theta=0}$ being

---

[15] As we will see in sec. 3.1, however, the thermal abundance of QCD axions through this process is important only for relatively small values of $f_a \lesssim 10^6$ GeV, ruled out by astrophysical observations.

$\mathcal{Z}(\theta)$ the vacuum generating functional in eq. (4). Corrections to the relation $m_a^2 = \chi_{\rm top}/f_a^2$ are of order m_π^2/f_a^2 and negligible for $f_a \gtrsim 10^8$ GeV, so from now on we will often trade m_a^2 for $\chi_{\rm top}$. In this section we will evaluate this observable to percent accuracy, in particular taking into account $O(p^6)$ corrections in the chiral expansion and the effect of nonvanishing electric charge at the first nontrivial order, $O(e^2p^2)$. We will often omit technical details in the derivations, which are similar to those in the previous sections but applied to $O(p^4)$ and $O(p^6)$ Lagrangians.

From the potential in eq. (25) we can get the leading order axion mass, derived for the first time in ref. [3],

$$(m_a^2)^{\rm LO} = \frac{\chi_{\rm top}^{\rm LO}}{f_a^2} = \frac{z}{(1+z)^2}\frac{m_{\pi^0}^2 f_\pi^2}{f_a^2} \approx \left[5.7\,\mu{\rm eV}\,\frac{10^{12}\,{\rm GeV}}{f_a}\right]^2 \tag{30}$$

where $z = m_u/m_d \sim 0.47$, and we used the values of m_{π^0} and f_π in tab. 6 in app. B. Eq. (30) shows that for the values of $f_a \gtrsim 10^8$ GeV allowed by astrophysical constraints the axion is a very light particle (even lighter that neutrinos), and therefore including it in the effective theory of QCD together with the pions is self-consistent. The NLO correction to eq. (30) comes from the $O(p^4)$ Lagrangian of [15] and one-loop diagrams constructed from $\mathcal{L}_2$. The calculation is simple and has been performed in [97], and leads to

$$\begin{aligned} \chi_{\rm top}^{\rm NLO} &= \chi_{\rm top}^{\rm LO}\left[1+\delta_1\right] , \\ \delta_1 &= 2\frac{m_{\pi^0}^2}{f_\pi^2}\left[h_1^r - h_3^r - \ell_4^r - (1-2\Delta^2)\ell_7\right] , \qquad \Delta \equiv \frac{1-z}{1+z} \ . \end{aligned} \tag{31}$$

The parameter Δ vanishes for equal quark masses and thus measures the amount of strong isospin breaking. The constants ℓ_i^r and h_i^r are the renormalized $O(p^4)$ LECs of [15] and depend on the renormalization group (RG) scale $\overline{\mu}$ of the chiral Lagrangian and m_{π^0} and $f_{\pi^0} = f_{\pi^\pm} = f_\pi$ are the physical pion mass and decay constant (which include NLO corrections). In particular, the combinations ℓ_7^r and $h_1^r - h_3^r - \ell_4^r$ are separately scale invariant. Moreover, at this order there is no contribution from loop diagrams after expressing the overall factor $m_{\pi^0}^2 f_\pi^2$ in eq. (30) in terms of NLO quantities.

A numerical estimate of m_a^2 requires a precise determination of the light quark mass ratio z and the $O(p^4)$ LECs. For z the latest lattice estimates are $z = 0.485(20)$ [99] with three dynamical quarks and $z = 0.513(31)$ [100] and $z = 0.453(16)$ [101], with four dynamical quarks. By combining them with the older $z = 0.470(56)$ [102] (also with four dynamical quarks), we get the following improved estimate

$$z = \left(\frac{m_u}{m_d}\right)^{\overline{\rm MS}}(2~{\rm GeV}) = 0.472(11)\,. \tag{32}$$

We should warn the reader that here the error has been computed by simply propagating the uncertainties quoted by each collaboration, since a proper combination is not yet available. In the remainder we will use the value in eq. (32) as reference, however we will always report separately the uncertainty originating from z and the total one so that it can easily be rescaled if needed.

For the LECs appearing in eq. (31) we use the values

$$h_1^r - h_3^r - \ell_4^r = -0.0049(12)\,, \qquad \ell_7 = 0.0065(38)\,. \tag{33}$$

The first combination is computed using the matching of the 2- and 3- flavor LECs (in particular to L_8^r) as described in ref. [97] and using the latest FLAG estimate $L_8^r = 0.00055(15)$ [103], while the second value is taken from the direct lattice simulation of ref. [104]. These values give

$$\delta_1 = -0.042(13)\,, \tag{34}$$

where the error is dominated by the one from ℓ_7. Combining everything together we get the values for the topological susceptibility and the axion mass at NLO

$$\chi_{\rm top}^{1/4} = 75.46(29)\ {\rm MeV}\,, \qquad m_a = 5.69(2)_z(4)_{\ell_i^r}\ \mu{\rm eV}\ \frac{10^{12}\ {\rm GeV}}{f_a}\,. \tag{35}$$

The dominant error comes from the one of the NLO LECs, in particular ℓ_7, which also controls the strong isospin breaking effect in the pion mass splitting, indeed poorly known. An improvement on this quantity would directly translate into an equivalent improvement in our knowledge of $\chi_{\rm top}$ and thus m_a. Conversely, improvements in the direct computation of $\chi_{\rm top}$ on the lattice could be used to better determine both z and ℓ_7.

A natural question to ask is how much an advance in our knowledge of the light quark masses and the NLO LECs can increase the precision of $\chi_{\rm top}$, before other unknown corrections need to be considered. Among the latter, the most relevant are the NNLO corrections of the chiral expansion and $\mathcal{O}(\alpha_{\rm em})$ electromagnetic (EM) corrections. The firsts do not only determine the ultimate precision reachable with eq. (31) but also measure the convergence and reliability of the chiral expansion. Of course the size of the NNLO corrections is only relevant in the chiral expansion approach and does not represent a source of uncertainty for lattice simulations[16], which contain the full non-perturbative result. The EM corrections, on the other hand, are common to both approaches and so far have never been considered. As we will show in the next section, their size is smaller with the choice made in eq. (30) of using the value of the neutral pion mass in the LO formula. Even with this choice, however, the value of the EM corrections is just below the size of the present uncertainties for $\chi_{\rm top}$, which means that further improvements cannot ignore them.

2.4.1 QED corrections

While the QCD axion has a vanishing electric charge, its mass can receive $\mathcal{O}(\alpha_{\rm em})$ corrections from several sources. Indeed, the leading order formula (30) involves a number of quantities that can introduce potentially large EM corrections depending on the way they are defined and extracted by experiments.

- The pion masses for the neutral and the charged states are degenerate at leading order, but differ at higher orders due to isospin and EM effects. The latter largely dominate this difference, which amounts to $m_{\pi^+} - m_{\pi^0} = 4.5936(5)$ MeV, i.e. around 4% of the total mass. The main effect comes from the charged pion mass (see eq. (22)), whose corrections are $\mathcal{O}(e^2)$, while those in the neutral pion mass start at $\mathcal{O}(e^2p^2)$. Therefore, depending on which pion mass is used in eq. (30), the axion mass can vary by 4%, which is more than the quoted uncertainties of the previous section. The expectation that the neutral pion mass should be used to minimize EM effects is the correct one. Indeed the pion mass entering in the leading order formula can be understood as arising from the mixing between the axion and the neutral pion state.
- In QCD the pion decay constant f_π is not unambiguously defined when EM interactions are turned on. In chiral perturbation theory, on the other hand, $\alpha_{\rm em}$ can be controlled analytically and it is possible to define f_π unambiguously. The best determination of f_π at the moment comes from (radiative) leptonic pion decays $\pi^+ \to \mu\nu_\mu(\gamma)$ where both experimental and theoretical

---

[16] On the other hand, lattice simulations have to face a number of systematic uncertainties which are not present in the chiral expansion such as finite volume, finite lattice spacing effects, explicit chiral symmetry breaking, etc., some of which require delicate and careful analyses.

uncertainties are small [96]. As we will discuss in more detail below, the EM corrections to $\Gamma_{\pi^+\to\mu\nu(\gamma)}$ are dominated by a calculable short distance contribution. The long distance hadronic contribution (which is of the same order of the EM corrections we want to compute for the axion mass) is subleading but dominates the current error of f_π. Given the importance of such corrections for our computation, we revisit their estimate and analyze their interplay with the genuine corrections to the axion mass. An alternative determination of f_π could be obtained from the neutral pion decay $\pi^0 \to \gamma\gamma$, however both the theoretical and experimental uncertainties are not competitive with the charged pion channel [96].

- While the light quark mass ratio $z = m_u/m_d$ at leading order is renormalization group (RG) invariant with respect to QCD corrections, it is not with respect to the QED ones [105]. This introduces an $\mathcal{O}(\alpha_{\rm em})$ ambiguity in the tree-level formula of the axion mass that should be removed by the sub-leading EM corrections:

$$\frac{\partial \log z}{\partial \log \mu} = \frac{6\alpha_{\rm em}}{4\pi}\left[\left(\frac{2}{3}\right)^2 - \left(-\frac{1}{3}\right)^2\right] = \frac{\alpha_{\rm em}}{2\pi} \tag{36}$$

A change of $\mathcal{O}(1)$ in the renormalization scale introduces a shift of $\mathcal{O}(10^{-3})$ in z that can be taken as a lower bound to the order of magnitude of the expected EM corrections to the axion mass.

We start by reporting the result for the computation of the leading EM corrections to the topological susceptibility, which begin at $\mathcal{O}(e^2p^2)$ in the chiral expansion once the leading order term is written in terms of the physical[17] neutral pion mass m_{π^0} (including EM corrections) and the physical charged pion decay constant f_{π^+} (defined in pure QCD, i.e. at $\alpha_{\rm em} = 0$):

$$\chi_{\rm top} = \frac{z}{(1+z)^2}\, m_{\pi^0}^2 f_{\pi^+}^2 \left[1 + \delta_e + \dots\right], \tag{37}$$

$$\delta_e = e^2\left[\frac{20}{9}(k_1^r + k_2^r) - 4k_3^r + 2k_4^r + \frac{8}{3}\Delta\, k_7^r - \frac{Z}{4\pi^2}\left(1 + \log\left(\frac{m_\pi^2}{\overline{\mu}^2}\right)\right)\right], \tag{38}$$

where dots in eq. (37) represent the non-EM corrections discussed previously and in the next section, the coefficients k_i^r are the $N_f = 2$ EM low-energy constants from [106], and the dependence on the RG scale $\overline{\mu}$ cancels against that from the k_i^r coefficients. As anticipated before, once the LO formula is written in terms of $m_{\pi^0}^2$, the EM corrections start at $\mathcal{O}(e^2p^2)$ (the δ_e term). In particular the EM pion mass splitting effects parametrized by (see eq. (22))

$$Z = \frac{m_{\pi^+}^2 - m_{\pi^0}^2}{2e^2 f_{\pi^+}^2} + \dots \simeq 0.81\,. \tag{39}$$

are loop suppressed. Although the value for the couplings k_i^r is not known directly, it can been inferred, as in [107], using their relation to the $N_f = 3$ constants K_i^r, which have been estimated in refs. [108, 109] using various techniques including sum rules and vector meson dominance. The values for the k_i^r we use are taken from [107] (with K_9^r from [109]) and reported in tab. 1. Because of the model dependence of such estimates we decided to assign a conservative 100% uncertainty to each

---

[17] Whenever m_π appears in the following formulas, it can be equivalently understood as m_{π^0} or m_{π^+} because the difference will be accounted by higher orders in either e^2 or p^2. For the numerical estimates we used $m_\pi = m_{\pi^0}$.

k_1^r	k_2^r	k_3^r	k_4^r	k_7^r	
8.4	3.4	2.7	1.4	2.2	$\times 10^{-3}$

Table 1: *Numerical values of the $N_f = 2$ EM LECs k_i^r at the scale $\overline{\mu} = 770$ MeV extracted using their relation to K_i^r. To the k_i^r it is assigned a conservative* 100% *uncertainty.*

LEC, i.e. we use the mentioned values as an order of magnitude estimate of their size. Substituting the numerical values we find

$$\delta_e = 0.0065(21)\,. \tag{40}$$

While we have assigned 100% uncertainties to the LECs k_i^r, the uncertainty on δ_e only amounts to 30% because the dominant contribution comes from the last term in eq. (38).

As discussed before, the QED RG scale dependence from the quark mass ratio z in the leading order formula (30) must be reabsorbed by $\mathcal{O}(\alpha_{\rm em})$ corrections. Indeed the EM LECs $k_{5,7}^r$ have the non-trivial UV-scale μ dependence[18]:

$$\mu\frac{\partial}{\partial\mu}k_5^r = -\frac{3}{5}\frac{1}{(4\pi)^2}\,, \qquad \mu\frac{\partial}{\partial\mu}k_7^r = -\frac{3}{4}\frac{1}{(4\pi)^2}\,. \tag{41}$$

It is easy to check that the variation of k_7^r reabsorbs the dependence induced by the variation of z in the leading order formula (in the $N_f = 3$ case the RG scale dependence is reabsorbed by K_9^r). In fact, the light quark mass ratio z and the constants $k_{5,7}^r$ cannot be determined independently and only the RG invariant combination enters physical quantities. The numerical value of k_7^r in tab. 1 is of the same order of the scale dependence in eq. (41), which therefore dominates its determination. In any case, the current uncertainties on the quark mass ratio z are still bigger than the effects from the scale dependence in z, and therefore bigger than the effects from k_7^r.

To complete the computation of $\chi_{\rm top}$ we need the value of the pion decay constant f_{π^+} at $\alpha_{\rm em} = 0$. Currently the best determination comes from the charged pion leptonic decay, which according to the PDG [96] provides $f_{\pi^+} = 92.28(9)$. This estimate however involves EM corrections of the same order of δ_e, so that a consistent calculation of $\chi_{\rm top}$ within the chiral expansion should consider the two sources of EM corrections together. In more details f_{π^+} is related to the EM inclusive pion decay rate via

$$\Gamma_{\pi^+\to\mu\nu(\gamma)} = \frac{G_F^2|V_{ud}|^2 m_{\pi^+} m_\mu^2 f_{\pi^+}^2}{4\pi}\left(1-\frac{m_\mu^2}{m_{\pi^+}^2}\right)^2\left[1+\delta_\Gamma^{\rm loc}+\delta_\Gamma^{\rm had}\right] \tag{42}$$

where the δ_Γ terms computed in [110] are the $\mathcal{O}(\alpha_{\rm em})$ corrections, which we split into two terms: the local contribution $\delta_\Gamma^{\rm loc}$ and the IR one $\delta_\Gamma^{\rm had}$, which parametrizes the hadronic form factors and depends on the chiral LECs. Explicitly they read:

$$\begin{aligned}
\delta_\Gamma^{\rm loc} &= \frac{\alpha_{\rm em}}{\pi}\left[\log\left(\frac{m_Z^2}{m_\rho^2}\right)+F\left(\frac{m_\mu^2}{m_{\pi^+}^2}\right)-\frac{m_\mu^2}{m_\rho^2}\left(\bar c_2\log\left(\frac{m_\rho^2}{m_\mu^2}\right)+\bar c_3+\bar c_4\right)+\frac{m_{\pi^+}^2}{m_\rho^2}\bar c_{2t}\log\left(\frac{m_\rho^2}{m_\mu^2}\right)\right],\\
\delta_\Gamma^{\rm had} &= e^2\left[\frac{8}{3}(K_1^r+K_2^r)+\frac{20}{9}(K_5^r+K_6^r)-\frac{4}{3}X_1^r-4(X_2^r-X_3^r)-\tilde X_6^{r,\rm eff}\right.\\
&\qquad\left.+\frac{1}{(4\pi)^2}\left(2-3Z-Z\log\left(\frac{m_K^2}{\overline{\mu}^2}\right)+(3-2Z)\log\left(\frac{m_\pi^2}{\overline{\mu}^2}\right)\right)\right],
\end{aligned}$$

---

[18] This can be derived by computing the operators generated in the chiral Lagrangian by an RG transformation of the quark mass matrix in terms of the EM charge spurions $Q_{L,R}$.

$\bar{c}_2$	$\bar{c}_3$	$\bar{c}_4$	$\bar{c}_{2t}$
5.2	-10.5	1.69	0

K_1^r	K_2^r	K_3^r	K_4^r	K_5^r	K_6^r	K_9^r	K_{10}^r	X_1^r	X_2^r	X_3^r	$\tilde{X}_6^{r,\text{eff}}$	
-2.7	0.7	2.7	1.4	12	2.8	-1.3	4	-3.7	3.6	5	13	$\times 10^{-3}$

Table 2: *Top: Numerical values of the $\bar{c}_i$ constants appearing in $\delta_\Gamma^{\text{loc}}$ from [110]. Bottom: Numerical values of the $N_f = 3$ radiative and leptonic LECs from [108, 109, 111] at the scale $\bar{\mu} = 770$ MeV. All constants are assigned a conservative* 100% *uncertainty.*

$$F(x) \equiv \frac{3}{2}\log x + \frac{13-19x}{8(1-x)} - \frac{8-5x}{4(1-x)^2}x\log x - \left(2 + \frac{1+x}{1-x}\log x\right)\log(1-x) - 2\frac{1+x}{1-x}\mathrm{Li}_2(1-x)\,. \tag{43}$$

The constants $\bar{c}_i$ have been estimated in a model dependent way in ref. [110], and for this reason we assign a conservative 100% uncertainty to them. The corresponding numerical values from [110] are reported in tab. 2. The constants K_i^r and X_i^r are the $N_f = 3$ radiative and leptonic LECs respectively, defined in refs. [112,113] (except for $\tilde{X}_6^{r,\text{eff}}$ defined in ref. [111]). We use the values estimated in ref. [108] for $K_{1,\ldots,6}^r$, in ref. [109] for $K_{9,10}^r$, and in ref. [111] for X_i^r, which we report in tab. 2 and to which we associate conservatively a 100% uncertainty.[19]

Numerically the size of the EM corrections to $\Gamma_{\pi^+\to\mu\nu(\gamma)}$ amounts to

$$\delta_\Gamma^{\text{loc}} + \delta_\Gamma^{\text{had}} = 0.0177(38)\,, \tag{44}$$

very close to the PDG estimate (0.0176(21)) but with larger error (here we have been more conservative). Note that, although the uncertainties of all LECs have been taken $\mathcal{O}(1)$, the result has a 20% accuracy, since the first term in $\delta_\Gamma^{\text{loc}}$ largely dominates over all the others. Combining eqs. (42) and (44) we get $f_{\pi^+} = 92.26(18)$.

Since some of the LECs appearing in $\delta_\Gamma^{\text{had}}$ are common with some of those appearing in δ_e, the topological susceptibility χ_{top} should be written directly in terms of the $\Gamma_{\pi^+\to\mu\nu(\gamma)}$ rather than f_{π^+}, i.e.

$$\chi_{\text{top}} = \frac{z}{(1+z)^2}m_{\pi^0}^2 \frac{\Gamma_{\pi^+\to\mu\nu(\gamma)}}{\frac{G_F^2|V_{ud}|^2 m_{\pi^+} m_\mu^2}{4\pi}\left(1 - \frac{m_\mu^2}{m_{\pi^+}^2}\right)^2}\left[1 + \delta_e - \delta_\Gamma^{\text{loc}} - \delta_\Gamma^{\text{had}} + \ldots\right]\,. \tag{45}$$

The combination $\delta_e - \delta_\Gamma^{\text{had}}$ can be written either all entirely in terms of $N_f = 3$ LECs, or in a hybrid way in terms of $N_f = 2$ k_i^r and $N_f = 3$ X_i^r (because the $N_f = 2$ leptonic LECs are not available):

$$\begin{aligned}\delta_e - \delta_\Gamma^{\text{had}} &= e^2\left[2k_4^r - 4k_3^r + \frac{8}{3}\Delta\, k_7^r - 4k_9^r + \frac{4}{3}X_1^r + 4(X_2^r - X_3^r) + \tilde{X}_6^{r,\text{eff}} - \frac{2(1+Z) + (3+2Z)\log\frac{m_\pi^2}{\bar{\mu}^2}}{(4\pi)^2}\right] \\ &= e^2\left[2K_4^r - 4K_3^r + \frac{8}{3}\Delta(K_9^r + K_{10}^r) + \frac{4}{3}X_1^r + 4(X_2^r - X_3^r) + \tilde{X}_6^{r,\text{eff}}\right. \\ &\quad \left. - \frac{1}{(4\pi)^2}\left(2(1+Z+\Delta Z) + (3+2Z)\log\frac{m_\pi^2}{\bar{\mu}^2} + 2\Delta Z\log\frac{m_K^2}{\bar{\mu}^2}\right)\right]\,.\end{aligned} \tag{46}$$

---

[19] The various estimates of the K_i^r in refs. [108,109] and references therein are not always compatible with each other, hence our conservative choice for the error, which is supposed to take those model-dependent deviations into account.

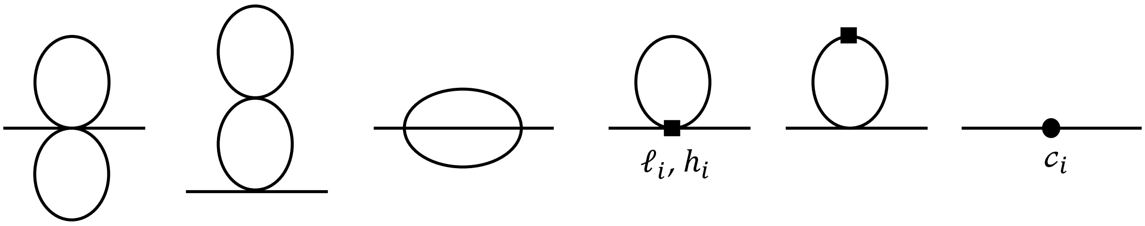

Figure 1: *One-particle-irreducible diagrams for the axion and pion 2-point functions at NNLO.*

In this way we get

$$\delta_e - \delta_\Gamma^{\rm had} - \delta_\Gamma^{\rm loc} = 0.024(6)\,, \tag{47}$$

which in combination with eq. (45) can be used to evaluate the EM contribution to $\chi_{\rm top}$.

The direct extraction of f_{π^+} from lattice QCD simulations is not competitive with the estimate above. However, recently the EM corrections to $\Gamma_{\pi^+\to\mu\nu(\gamma)}$ have been computed in a preliminary study on the lattice [114], giving

$$\delta_\Gamma^{\rm loc} + \delta_\Gamma^{\rm had} = 0.0169(15)\,, \tag{48}$$

which is in very good agreement with eq. (44). Accidentally this value is very close to the PDG one, both in size and uncertainty. Eq. (48) implies $f_{\pi^+} = 92.30(7)$. Given the compatibility of the chiral and the lattice results, and the fact that the latter has better precision and less model dependence, we will use eq. (40) and this lattice estimation for f_{π^+}, bearing in mind that numerically this choice is also equivalent to using the PDG determination.

2.4.2 NNLO corrections

Given the smallness of the expansion parameter, the $N_f = 2$ chiral expansion is expected to converge very fast and the NNLO corrections to be only a few percent with respect to the NLO ones. Nevertheless, depending on the magnitude of the low-energy coefficients, they might be competitive with the EM corrections discussed in the previous section. Their estimation is therefore essential for a precise calculation of $\chi_{\rm top}$.

The NNLO corrections to the topological susceptibility receive contributions from the diagrams in fig. 1, which correspond to: 1) the two-loop diagrams constructed from $\mathcal{O}(p^2)$ vertices, 2) the one-loop diagrams generated by $\mathcal{O}(p^4)$ vertices and 3) the tree-level graphs from the $\mathcal{O}(p^6)$ Lagrangian.[20] At $\alpha_{\rm em} = 0$, the full NNLO result is

$$\begin{aligned}
\chi_{\rm top}^{\rm NNLO} &= \frac{z}{(1+z)^2}\, m_{\pi^0}^2 f_{\pi^+}^2 \left[1+\delta_1+\delta_2\right]\,, \qquad (49)\\
\delta_2 &= \frac{m_\pi^4}{f_\pi^4}\left[32(c_6^r+2c_{19}^r) + \frac{2\bar\ell_3+4\bar\ell_4+3\log\frac{m_\pi^2}{\overline{\mu}^2}-2}{4(4\pi)^4}\log\frac{m_\pi^2}{\overline{\mu}^2} + \frac{3\ell_7}{8\pi^2}\log\frac{m_\pi^2}{\overline{\mu}^2}\right.\\
&\qquad +\frac{h_1^r-h_3^r-\ell_4^r-\ell_7}{8\pi^2}\bar\ell_3 - \frac{1}{(4\pi)^4}\left(\bar\ell_4^2-\bar\ell_3\bar\ell_4+\frac{25}{16}\right)\\
&\qquad +\left(32\left(c_9^r-6c_{10}^r-4\left(c_{11}^r+c_{17}^r+c_{18}^r\right)-6c_{19}^r\right)-\frac{7\ell_7}{8\pi^2}\log\frac{m_\pi^2}{\overline{\mu}^2}\right.
\end{aligned}$$

---

[20] Note that even with a quark field redefinition that avoids the tree-level mixing between the axion and the neutral pion, i.e. $Q_a = M_q^{-1}/\langle M_q^{-1}\rangle$ this mixing arises at one loop, producing effects at NNLO in $\chi_{\rm top}$.

$10^6 c_6^r$	$10^6 c_7^r$	$10^6 c_8^r$	$10^6 c_9^r$	$10^6 c_{10}^r$	$10^6 c_{11}^r$	$10^6 c_{17}^r$	$10^6 c_{18}^r$	$10^6 c_{19}^r$
1.0(3.8)	−2.2(2.0)	2.1(1.1)	−0.3(1.2)	−1.0(1.1)	−0.1(0.7)	6.4(3.8)	−2.2(5.9)	−0.1(9.5)

Table 3: *Numerical values of the $N_f = 2$ LECs in δ_2 at the scale $\overline{\mu} = 770$ MeV extracted by combining the lattice results of [104] and the matching with the $N_f = 3$ LECs of [121] (see app. B for more details).*

$$+8(h_1^r - h_3^r - \ell_4^r)\ell_7 - 12\ell_7^2 + \frac{2\overline{\ell}_3 + 2\overline{\ell}_4 - 1}{8\pi^2}\ell_7\Bigg)\Delta^2 + 20\ell_7^2\,\Delta^4\Bigg] , \tag{50}$$

with δ_1 given in eq. (31), $\overline{\ell}_i$ defined in [15] and the c_i^r being the $\mathcal{O}(p^6)$ $N_f = 2$ LECs introduced in [116, 117]. Note that the charged and neutral pion decay constants (defined, as mentioned, at $\alpha_{\text{em}} = 0$) differ only at two loops due to isospin breaking effects, so f_π in δ_1 and δ_2 can be understood either as f_{π^+} or f_{π^0}, being the difference accounted by $\mathcal{O}(p^8)$ terms. The scale dependence of the combinations of c_i^r in eq. (50) is fully reabsorbed by the $\log \frac{m_\pi^2}{\overline{\mu}^2}$ and $\log^2 \frac{m_\pi^2}{\overline{\mu}^2}$ terms, and since $h_1^r - h_3^r - \ell_4^r$, ℓ_7 and $\overline{\ell}_{3,4}$ are scale invariant, the scale dependence of δ_2 cancels separately in each line of eq. (50).

While the numerical value of most of the $\mathcal{O}(p^4)$ LECs is reasonably well known, the determination of the c_i^r is in much worse shape. In fact, only few combinations of c_i^r can be extracted directly because there are not enough experimental observables to fit all the $N_f = 2$ Lagrangian parameters.[21] Recent partially-quenched lattice QCD simulations [104] provided results for some of the combinations of c_i^r appearing in eq. (50). For the remaining ones we matched the relevant combinations to the $N_f = 3$ LECs, for which some estimates exist [121] (taken with a conservative 100% error). In this way we have been able to extract an order of magnitude estimate for all the c_i^r appearing in eq. (50), which we report in tab. 3 (see app. B for more details).

The LECs in tab. 3 and eq. (33), combined with the values $\overline{\ell}_3 = 2.81(49)$ and $\overline{\ell}_4 = 4.02(25)$ from [104], lead to the following numerical result for the NNLO corrections:

$$\delta_2 = -0.0071(01)_z(23)_{\ell_i^r}(19)_{c_i^r} = -0.0071(29)\,. \tag{51}$$

While the uncertainty from z is very small, those from ℓ_i^r (of which ℓ_7 provides the largest contribution) and c_i^r have similar size. Notice that although the relative uncertainties of the c_i^r are large, they only have a milder impact on the final uncertainty of δ_2, because numerically δ_2 receives bigger contributions from the $\mathcal{O}(p^4)$ LECs and the non-local contributions. Moreover, the isospin-breaking terms in δ_2 (the last two lines in eq. (50)), which are suppressed by powers of $\Delta^2 \approx 0.1$, contribute less than 20% to the final result and are within the uncertainty of δ_2. As a consequence, the precision on the LECs is still not enough for the result to be sensitive to isospin breaking corrections. Finally notice that δ_2 is numerically of the same order of the EM corrections in eq. (40), but with opposite sign. Therefore, both have to be considered for a sub-percent estimate of χ_{top}.

---

[21] In fact, of the c_i^r that appear in eq. (50), only c_6^r has been extracted semi-directly from experiments, in particular from the pion scalar form factor [118], with some phenomenological modeling. As explained in app. B, since on its numerical value there is still disagreement [118–120], we will not use it in our analysis.

2.4.3 Final Results and Axion Mass

We can now combine the analysis of the previous sections and estimate of the topological susceptibility to $\mathcal{O}(p^6, e^2p^2)$. The final result reads

$$\chi_{\text{top}} = \frac{z}{(1+z)^2}\, m_{\pi^0}^2 f_{\pi^+}^2 \left[1 + \delta_1 + \delta_2 + \delta_e\right] , \tag{52}$$

where the $\mathcal{O}(p^4)$ contribution δ_1 is given in eq. (31), the $\mathcal{O}(p^6)$ contribution δ_2 in eq. (50) and the $\mathcal{O}(e^2p^2)$ contribution δ_e in eq. (38). For completeness, in app. A we also report χ_{top} expressed in terms of quark masses and bare chiral Lagrangian parameters.

Substituting our numerical estimates, the final results for the topological susceptibility and the axion mass read

$$\chi_{\text{top}}^{1/4} = 75.44(34)\ \text{MeV}\,, \qquad m_a = 5.691(51)\,\mu\text{eV}\,\frac{10^{12}\ \text{GeV}}{f_a}\,. \tag{53}$$

Notice how these values almost coincide with the NLO ones in eq. (35), since both NNLO and EM corrections are comparable but smaller than the present uncertainties of the NLO estimate, and, having opposite sign, they tend to cancel each other. This result confirms the reliability of the NLO estimate in [97]. It is instructive to deconstruct the contributions at each order with the various uncertainties: for the axion mass case they read

$$m_a = \Big[\underbrace{5.815(22)_z(04)_{f_\pi}}_{\text{LO}}\,\underbrace{-0.121(38)_{\ell_i^r}}_{\text{NLO}}\,\underbrace{-0.022(07)_{\ell_i^r}(05)_{c_i^r}}_{\text{NNLO}}\,\underbrace{+0.019(06)_{k_i^r}}_{\text{EM}}\Big]\,\mu\text{eV}\,\frac{10^{12}\ \text{GeV}}{f_a}\,, \tag{54}$$

where the reported uncertainties on each contributions come from those of z, from the EM corrections in the extraction of f_{π^+}, and from those of the various LECs in the NLO (ℓ_i^r and h_i^r), in the NNLO (c_i^r) and in the EM (k_i^r) chiral Lagrangians.

Several comments are in order. First notice how, while NLO corrections are almost two orders of magnitude smaller than the LO result, NNLO are barely one order of magnitude below the NLO ones. On one side this means that the chiral expansion is nicely converging and, given the current uncertainties on z and the LECs, the NLO result is enough. On the other side, the size of the NNLO corrections is such that they cannot be ignored in future improvements of m_a.

EM corrections are of similar size, slightly less than 0.5% and within present uncertainties. The numerical estimate of the EM corrections has been carried out using the lattice QCD results for f_{π^+} extracted from eqs. (42) and (48) with δ_e in eq. (40), since these values are more model-independent.[22] However, one could have also used eqs. (45) and (47) obtaining essentially the same central value although with an error twice as large. As for NNLO corrections, they must be considered should the uncertainties coming from z and the NLO LECs decrease. As commented before, the size of these corrections also represents the ultimate precision that can be reached in lattice estimates which do not include EM corrections.

We conclude by noticing that, if the uncertainties in m_u/m_d and the NLO LECs (in particular ℓ_7) are reduced by a factor of few (which is not unreasonable) these results could be used to determine the axion mass (and $\chi_{\text{top}}^{1/4}$) with per-mille accuracy.

---

[22] A very similar result would follow using the PDG value for f_{π^+}, which, as mentioned, is very close to the lattice estimate both in size and error.

2.5 Topological Susceptibility and Axion Mass at finite temperature

As we will see in sec. 3.2, the cosmological evolution of the axion field depends on how the axion potential (in particular its mass) behaves at finite temperature, which is determined by finite-temperature QCD properties. We can separate the study of the temperature dependence in three regimes, and only on two of them we will have analytical control.

For $T \ll \Lambda_{QCD}$ the finite temperature effects are expected to be exponentially suppressed by Boltzmann factors $e^{-m_\pi^2/T^2}$ due to the mass gap of QCD. Even if QCD is nonperturbative in this regime, finite-temperature chiral perturbation theory may be used to study effects at small enough T. Indeed it has been shown in ref. [97] that if $T < m_\pi$ the whole potential is extremely close to its zero temperature form in eq. (25) and for $T \sim 150$ MeV the potential is $O(1)$ different (but in this regime the chiral expansion breaks down and might be not reliable).

For $T \sim \Lambda_{QCD}$, more precisely at $T = T_c \sim 160$ MeV, QCD undergoes a phase transition (actually thought to be a crossover) where the density of quark condensates $\langle qq \rangle$ could be considered the order parameter. The density of condensates diminishes fast as the temperature increases and the only way to understand the behavior of physical observables like χ_{top} is to use non-perturbative methods, e.g. lattice QCD [76, 77].

For $T \gg \Lambda_{QCD}$, QCD becomes perturbative and the saddle point approximation of the path integral in eq. (4) (which defines the axion potential and χ_{top}) about instanton configurations should be in principle under control and can be systematically evaluated. Indeed, one can show that at finite temperature the one-loop approximation around $k = \pm 1$ in eq. (7) gets modified by the introduction of the term $-\frac{2\pi^2}{g_s^2} m_d^2(T)\rho^2$ (among others) in the exponential [78], where $m_d(T) = \frac{1}{3} g_s^2 T^2 (3 + N_f/2)$ is the Debye mass of the gluon. This term can be interpreted as an effective mass for the gluon arising from its interactions with the charged thermal plasma of unconfined quarks, and acts as an exponential cut-off for the integration over ρ in eq. (7) at values $\rho \sim (\pi T)^{-1}$, making the integral converge. Note that large size instantons are screened and only those with ρ smaller than $(\pi T)^{-1}$ (for which the coupling $g_s(1/\rho \sim \pi T)$ is perturbative if T is large enough) contribute. Moreover, at large enough temperatures it is indeed meaningful to consider the contribution to the path integral only from $k = \pm 1$ instantons as in eq. (7), since the saddles around instantons with higher and higher k are more and more suppressed by the exponential $\exp(-8\pi^2 |k|^2 / g_s^2)$, where g_s will be the coupling $g_s(\pi T)$ because of similar temperature cutoffs. Higher loop corrections around $k = \pm 1$ instantons will be instead only suppressed by powers of $g_s^2(\pi T)$.

Restricting to the semiclassical approximation around $k = \pm 1$ and performing the integral in eqs. (7) and (8), we get

$$E(\theta, T) \propto -\det M_f \, (\Lambda_{QCD})^{4-N_f} \left(\frac{\Lambda_{QCD}}{T} \right)^{7+\frac{1}{3}N_f} \cos\theta \, . \tag{55}$$

The result is that at large enough temperatures (e.g. at 1 GeV when $N_f = 3$) the topological susceptibility (and so m_a^2) decreases as $\sim T^{-8}$, while the only harmonic present in the potential is $\cos\theta$. As the temperature gets closer to Λ_{QCD}, higher order corrections in the perturbative expansion around $k = \pm 1$ and – later because of the exponential suppression – saddles around higher order instantons will become relevant, changing in principle the dependence on T in eq. (55) and switching on higher harmonics $\cos(k\theta)$, until when the saddle point approximation breaks down completely[23].

---

[23] Note that also the dependence on the quark masses will have to change from $\det M_q$ to that of eq. (23), which is non-analytic.

Since the first nontrivial correction to eq. (55) has not been yet computed, it is not known at which temperature the validity of eq. (55) stops. However, lattice QCD simulations seem to suggest that the dependence on T^{-8} (and the single harmonic $\cos\theta$) holds up to very small temperatures, just above the phase transition, although the overall size of χ_{top} is not yet well reproduced by the simulations (see [90, 91] for pure Yang Mills and [92–95] for QCD).

For the estimates in the following chapter we will assume the simplified functional form

$$\chi_{\text{top}}(T) = m_a^2(T) f_a^2 = \begin{cases} \chi_{\text{top}}(0) & \text{if } T < \Lambda_{QCD} \\ \chi_{\text{top}}(0)\left(\frac{\Lambda_{QCD}}{T}\right)^{\alpha} & \text{if } T > \Lambda_{QCD} \end{cases} \tag{56}$$

where $\chi_{\text{top}}(0)$ is the zero temperature value discussed in the previous section and $\alpha \sim 8$, which will capture the main effects, but we will still leave the power α to be a free parameter.

3 QCD axion dark matter

Having introduced a new light particle such as the QCD axion in the theory, we need to study its cosmological implications and in particular how much energy density in the form of this field is left today in the Universe. We will consider the minimal scenario where the QCD axion (and the inflation) are the only relevant particles beyond the SM, and possible new physics that UV completes the axion theory does not change significantly the cosmology. In this way we can extrapolate the SM cosmology up to sufficiently high temperatures; in particular the Universe is radiation dominated up to temperatures larger than Λ_{QCD}. In this context the only free Lagrangian parameter will be the axion decay constant f_a.

We will show in this chapter that, depending on the hierarchy between f_a and the scale of inflation, there is a large window of values of f_a such that the QCD axion relic density is in the form of pressureless matter. Moreover, a contribution to the relic energy density in this field will be unavoidable for all values of f_a. While one might naively guess that a population of relic axions is still relativistic today (because they are extremely light) and so behaves as radiation, we will show that the mechanism by which they are produced automatically guarantees that they are nonrelativistic.

Note that the axion is very weakly coupled to all SM particles and very hard to detect because its couplings are suppressed by powers of f_a. In particular, its main decay channel is to photons, and its decay rate $\Gamma_{a\gamma\gamma} = \frac{m_a^3}{4\pi}(\frac{g_{a\gamma\gamma}}{4})^2$ is dictated by the coupling (see eq. (14) and below)

$$g_{a\gamma\gamma} = \frac{\alpha_{\rm em}}{8\pi}\left[\frac{E}{N} - \frac{2}{3}\frac{m_u + 4m_d}{m_u + m_d}\right] \tag{57}$$

where the second term in the bracket[24] is about -2. Taking E/N of order 1, the lifetime is $\tau = \Gamma_{a\gamma\gamma}^{-1} \approx 10^{18}\tau_u(f_a/10^9\,\mathrm{GeV})^5$ where τ_u is the age of the Universe, and therefore the axion is absolutely stable on cosmological scales. Given all these properties, axions are naturally excellent candidates to make up a part or all of the cold dark matter (CDM) of our Universe.

Axions are produced in the early universe in at least two ways. As we will see, there is a contribution of (warm) thermal axions which turns out to be suppressed in the region $f_a \gtrsim 10^8$ GeV that is not ruled out by astrophysical constraints. On the other hand, nonrelativistic axions are produced non-thermally (and automatically) through the so-called misalignment mechanism, which will be our focus.

3.1 Thermal QCD axions

If the interaction rate $\Gamma = \langle n\sigma v\rangle$ of the QCD axion with the SM thermal bath has ever been bigger than the Hubble parameter H, axions would have been in thermal equilibrium with the SM and possibly produced by thermal freeze-out (see [79–81]). In the following we will estimate the relic abundance of axions from this process and understand whether they are possibly still relativistic today.

As explained in sec. 2.5, at high enough temperatures axions are massless (so relativistic) and their number density is $n = \zeta(3)T^3/\pi^2$ and $v \sim 1$. The cross section σ of axions with the SM particles depends on the temperature, and in general also on the model. We consider only the model-independent coupling to gluons, which at temperatures larger than Λ_{QCD} mediates at tree level the process $gg \to ag$ whose cross section is $\sigma \sim (\frac{\alpha_s}{8\pi}\frac{1}{f_a} \times g_s)^2 \sim 10^{-5} f_a^{-2}$ (the first factor comes from the axion-gluon trilinear vertex, the second from an additional trilinear QCD vertex). At temperatures of order Λ_{QCD}, the relevant coupling is to pions (see eq. (27)) and the relevant process is $\pi\pi \to \pi a$,

---

[24] A NLO computation [97] gives $-1.92(4)$ for the second term in the bracket.

whose cross section is $\sigma \sim \frac{1}{4\pi}(\frac{E}{f_\pi f_a})^2 \sim \frac{1}{4\pi}\frac{1}{f_a^2}$ and we neglected the model dependent coefficients c_i^0. In total, the interaction rate is $\Gamma \sim 10^{-5}T^3/f_a^2$ for $T \gg \Lambda_{QCD}$ and $\Gamma \sim 10^{-2}T^3/f_a^2$ for $T \sim \Lambda_{QCD}$.

In radiation domination the Hubble parameter is $H^2 = (\pi^2/90)g(T)T^4/M_P^2$ where $g(T)$ is the effective number of relativistic degrees of freedom at the temperature T.

- In the regime $T \gg \Lambda_{QCD}$ the condition $\Gamma > H$ translates into $T \gtrsim f_a^2/(10^{13}$ GeV) where we used $g(T) \sim 100$ since we are assuming only SM degrees of freedom. At temperatures below the bound above, axions decouple from the thermal bath and evolve as free particles. Depending on f_a and on whether the Universe has even been hot enough axions were therefore in equilibrium with the plasma.
- For $T \sim \Lambda_{QCD}$ the equilibrium condition is $T \gtrsim f_a^2/(10^{16}$ GeV), which is satisfied only for $f_a \lesssim 10^7$ GeV, i.e. in the parameter region excluded by astrophysical bounds. We are therefore interested in the previous case only.

In the first case axions will decouple at a temperature T_d bigger than Λ_{QCD} at which $g(T_d) \sim 100$. The temperature of these axions today is $T_a = T_d(s_0/s_d)^{1/3}$ where s is the entropy density and we used $s \propto T^{-3}$ for decoupled species (like axions). Therefore $T_a = T_0(g(T_0)/g(T_d))^{1/3}$ since $s = (2\pi^2/45)g(T)T^3$ where $g(T_0) = 43/11$ and $T_0 = 2\cdot 10^{-4}$ eV is the temperature of the CMB photons today, and so $m_a < T_a$ if $f_a \gtrsim 10^{11}\,\text{GeV}(g(T_d)/100)^{1/3}$. The result is that for values of f_a lower than 10^{11} GeV these axions are nonrelativistic today.

The relic abundance $\Omega_a^{\text{th}} = \rho_a/\rho_c$ of thermal axions is

$$\Omega_a^{\text{th}} = \frac{m_a n_a^0}{\rho_\gamma/\Omega_\gamma} = \Omega_\gamma \frac{m_a \frac{\zeta(3)}{\pi^2}T_a^3}{2\times\frac{\pi^2}{30}T_0^4} = \Omega_\gamma \frac{m_a}{T_0}\frac{15\zeta(3)}{\pi^4}\frac{g(T_0)}{g(T_d)} = 10^{-8}\left(\frac{100}{g(T_d)}\right)\left(\frac{10^{12}\,\text{GeV}}{f_a}\right) \tag{58}$$

where we assumed that the comoving number density of axions is conserved from the freeze-out time (also when they turn nonrelativistic). From the previous equation we see that for $f_a \gtrsim 10^8$ GeV the thermal relic density is negligible compared to the energy density in dark matter $\Omega_{dm} \sim 0.24$ measured today. However this thermal population of axions might still leave an imprint as a component of relativistic energy density at the time of the CMB decoupling, usually measured as the effective number of neutrinos ΔN_{eff}. While the bounds on ΔN_{eff} are still large, these might tighten in the near future and further constrain this scenario (see e.g. [82]).

3.2 Misalignment QCD axions

The relic energy density of any scalar field with a nontrivial potential (such as the QCD axion) gets a contribution via the misalignment mechanism, which is based on the fact that the initial value of the field in the early universe is not in general aligned with the minimum of the potential [8–10] (see e.g. [17, 18] for reviews). Depending on the initial field value and the Universe evolution, the field will start oscillating around its minimum and the energy stored in these oscillations will in principle contribute to the energy budget of our Universe today (and possibly behave as pressureless matter).

In this section we will assume that the axion field $a(t, \vec{x})$ is spatially homogeneous at the beginning of its evolution. This is always the case unless inflation happens at a scale H_I bigger than f_a or the reheating temperature T_R if bigger than f_a, so that quantum fluctuations during inflation or finite temperature effects do not have the chance of restoring the PQ symmetric phase in any region of the universe. Indeed, in this case due to the exponential expansion of the Universe during inflation, the axion field will reach at the end of inflation a constant value over distances of order $e^N H_I^{-1}$, where $N \gtrsim 60$ is the number of e-foldings, which is much larger than the Hubble distance. We will discuss in the next chapter the more complicated case where either H_I or T_R are bigger than f_a, in which the PQ is restored after inflation (and then broken again).

The equations of motion for the axion will then be

$$\ddot{a} + 3H\dot{a} + \frac{\partial V(a,t)}{\partial a} = 0 \ , \tag{59}$$

where $H = \dot{R}/R = 1/2t$ is the Hubble parameter in radiation domination[25], $R(t) \propto \sqrt{t}$ is the scale factor and $H^2 = (\pi^2/90)g(T)T^4/M_P^2$. The dependence of the potential on time comes from its dependence on temperature in eqs. (55) and (56). Moreover, if the initial condition $a(t_0)$ at the time $t = t_0$ is sufficiently near the minimum of the potential, we can approximate $\partial_a V(a,t)$ with $m_a^2(t)a$.

Eq. (59) corresponds to an harmonic oscillator with a friction term decreasing in time. At early times the Hubble friction dominates, $H \gg m_a(t)$, the last term in eq. (59) is negligible and the solution is $a(t) = a(t_0)$ where t_0 is the initial time (for all $\dot{a}(t_0)$), i.e. the axion stays frozen at its initial value. At times where $H \sim m_a(t)$ the axion starts feeling the potential and moves away from $a(t_0)$. Finally, at late times $H \ll m_a(t)$ and eq. (59) is an underdamped harmonic oscillator with frequency $m_a(t)$, and $a(t)$ will oscillate around the minimum $a = 0$ with an amplitude decreasing less and less since Hubble decreases.

In the limit $H \ll m_a(t)$ and $\dot{m}_a(t) \ll m_a^2(t)$ which is verified at sufficiently late times (when $m_a(t)$ is practically constant), the energy density of the field averaged over one oscillation period behaves like pressureless matter and the comoving number density of nonrelativistic quanta is conserved in time. Indeed, in the limit above the solution of eq. (59) is approximated by $a(t) = A(t)\cos(\int^t m_a(t')dt')$ where $A(t) \propto R(t)^{-3/2}$ varies in time much slower than than m_a. It follows that $\langle \dot{a}^2(t)\rangle = \langle m_a^2(t)a^2(t)\rangle \left[1 + O(H^2/m_a^2)\right]$ where $\langle\cdot\rangle$ is the average over one oscillation. The energy density $\rho_a = T_{00}$ and the pressure $P_a = -T_0{}^i$ of the field are

$$\rho_a = \frac{1}{2}\dot{a}^2 + \frac{1}{2}m_a^2 a^2 \ , \qquad P_a = \frac{1}{2}\dot{a}^2 - \frac{1}{2}m_a^2 a^2 \tag{60}$$

and so $\langle\rho_a\rangle = \langle m_a^2 a^2\rangle \left[1 + O(H^2/m_a^2)\right]$ and $\langle P_a\rangle/\langle\rho_a\rangle = O(H^2/m_a^2)$, which in the small H/m_a limit is the equation of state of a pressure-less fluid. Moreover, using the fact that $\langle H\rangle = H$ and $\langle \dot{m}_a/m_a\rangle = \dot{m}_a/m_a$ up to $O(H^2/m_a^2)$ corrections we get $\langle\dot{\rho}\rangle = \langle\rho\rangle(\dot{m}_a/m_a - 3\dot{R}/R)$. Integrating this equation in time, it is easy to see that the average number density $n_a = \langle\rho\rangle/E$ ($E = m_a$ because only the zero mode with momentum $k = 0$ is excited) satisfies

$$n_a = \frac{\langle\rho\rangle}{m_a} \propto R^{-3} \ . \tag{61}$$

The result is that the axion energy density redshifts as matter (as soon as m_a becomes constant) and the comoving number density $n_a R^3$ is conserved. An homogeneous axion field oscillating in time is interpreted as a collection of axions at rest, whose energy density indeed behaves as in eq. (61). The energy stored in these oscillations will last until today (though redshifted) and will make up at least a part the CDM we observe.

To compute the number density of axions leftover today we would need to solve numerically eq. (59) with the proper time-dependence of the potential. However, we can get a rough analytic estimate by assuming that the late time solution (therefore the conservation of the number density) holds from

---

[25] We are assuming radiation domination up to the time of reheating, but what matters for the final abundance is up to temperatures just above T_* defined in eq. (62), when the axion starts oscillating.

the time t_* at which the axion starts oscillating[26], roughly set by $m_a(t_*) = H(t_*)$. Given the mass dependence on temperature in eq. (56) with $\chi_{\rm top}(0) \approx \Lambda_{QCD}^2$, the corresponding temperature is

$$\frac{T_*}{\Lambda_{QCD}} = \left[\frac{M_P}{f_a}\frac{3\sqrt{10}}{\pi\sqrt{g(T_*)}}\right]^{\frac{2}{4+\alpha}} , \tag{62}$$

and $g(T) \sim 60$ at $T = 1$ GeV (but it depends on the exact value of T especially near the QCD phase transition). From eq. (62), T_* depends on f_a and is generally bigger than Λ_{QCD} unless f_a is very close to M_P, and the smaller is α the bigger is the separation between T_* and Λ_{QCD}. Note moreover that eq. (62) is not a good approximation if $T_* \sim \Lambda_{QCD}$ because the single power-law in the temperature-dependence of m_a in principle breaks down; however, the fact that T_* is parametrically above Λ_{QCD} makes eq. (62) self-consistent. In addition, the poor knowledge of the temperature dependence of m_a around the QCD crossover does not affect the conservation of the axion number density, that holds as soon as $H \ll m_a(t)$ and $\dot{m}_a(t) \ll m_a^2(t)$, and so it is somewhat unimportant if f_a is small enough.

Similarly to eq. (58), using the conservation of the comoving entropy and number densities from t_* to today, the contribution to the energy density from these axions is

$$\Omega_a^{\rm mis} = \frac{m_a n_a^0}{\rho_\gamma/\Omega_\gamma} = \Omega_\gamma \frac{m_a}{T_0}\frac{86}{33}\frac{n_a^*}{s_*} \tag{63}$$

where T_0 is the temperature of the CMB photons. We have $n_a^* = \frac{1}{m_a^*}\langle m_a^{*2} a^2(t_*)\rangle = \frac{1}{2}H(t_*)\theta_0^2 f_a^2$, where we used $a(t_0) = a(t_*)$ and we parameterized $a(t_0) = \theta_0 f_a$ with $|\theta_0| \leq \pi$ for periodicity. Expressing $H(t_*)$ and s_* in terms of T_* and using the explicit form of T_* we get finally

$$\Omega_a^{\rm mis} = 0.1 k_\alpha \left(\frac{\theta_0}{2.15}\right)^2 \left(\frac{f_a}{10^{12}\,{\rm GeV}}\right)^{1+\frac{1}{2+\alpha/2}} \left(\frac{60}{g(T_*)}\right)^{\frac{1}{2}-\frac{1}{4+\alpha}} , \tag{64}$$

where we omitted $O(1)$ factors, that include the uncertainty from the transient period when $m_a \sim H$, and k_α ranges from 3 to 5 for α from 3 to 8.

Therefore, misalignment axions could explain the whole CDM abundance if θ_0 is of $O(1)$ and $f_a \sim 10^{12}$ GeV. However, the value of f_a such that the QCD axion makes up the whole relic density is not uniquely determined, since is a function on the (unknown) initial misalignment angle θ_0, which in principle depends on the specific Hubble patch in the Universe. In particular, the correct relic abundance can be obtained also for larger values of f_a as long as[27] θ_0 is smaller than 1. As a consequence, this scenario does not provide the axion relic density only in terms of Lagrangian parameters (i.e. of f_a), and so it is not predictive.

In this discussion we neglected the inharmonicities of the axion potential. These are however important if $|\theta_0| \sim \pi$, because they delay the oscillations and therefore produce a bigger DM abundance, and a self-consistent analysis should take them into account.

Note that in writing eq. (59) we omitted the interactions with all SM particles. This is justified because they are suppressed by powers of f_a. In this regard, as discussed in [83], we are allowed to treat misalignment axions separately from the thermal bath because their thermalization rate is very

---

[26] This is a good approximation because both m_a and H change power-like with time, so the transient period where m_a and H are comparable is very short, which suggests that it does not have a big impact in the estimate.

[27] The determination of the relic abundance becomes a cosmographical question and anthropic arguments could be invoked to explain why $\theta_0 \ll 1$ is acceptable in the Hubble patch where we live. Note that even if θ_0 was known, there would still be a milder uncertainty coming from α and $g(T_*)$, which is however small is $T_* \gg \Lambda_{qCD}$.

small. We are also neglecting the interactions with possible new particles that UV complete the axion theory at the scale f_a because they are heavy enough (with respect to Hubble) not to be produced, at least at the time when the axion starts oscillating.

Moreover, θ_0 cannot be chosen arbitrarily close to zero or to π in this scenario because of quantum fluctuations during inflation. Indeed, any massless scalar field fluctuates during inflation with a variance of $H_I/(2\pi)$, which means that the minimum displacement of $a(t_0)$ from any point (in particular from $a(t_0) = 0$ and $a(t_0) = \pi$) is bounded by $H_I/(2\pi)$. As a consequence, it must be $|\theta_0| \geq H_I/(2\pi f_a)$ and similarly $|\pi - \theta_0| \geq H_I/(2\pi f_a)$, which excludes high scale inflation models if θ_0 is tuned near 0 or π.

More importantly, the quantum fluctuations of the axion during inflation provide isocurvature perturbation in the CMB spectrum [85], which actually give the strongest upper bound on H_I for a fixed value of f_a. The quantum fluctuations imprinted to the axion are isocurvature because they do not change the total energy density of the Universe during inflation, being the axion negligible in the energy budget during inflation. On the other hand, these perturbations produce spatial anisotropies of the axion field at distances $\gtrsim H_I^{-1}$, and therefore different patches of the Universe will have different of the axion number densities when the axion starts oscillating. In particular, the modes of order Hubble at recombination will provide anisotropies in the axion dark matter densities at the time of the CMB decoupling, which are in addition to those coming from the initial quantum fluctuations of the inflaton (or equivalently of the metric). These additional anisotropies in turn leave a measurable imprint in the temperature fluctuation of the CMB spectrum usually denoted by $(\delta T/T)_{iso}$, on which there is a limit $\alpha_{iso} = (\delta T/T)_{iso}/(\delta T/T)_{tot} < 0.02$ [84]. One can compute $(\delta T/T)_{iso}$ as a function of H_I and f_a using $\delta a \sim H_I/(2\pi)$ and $(\delta T/T)^2_{tot} \sim 10^{-9}$, and find that the Hubble parameter for which the amount of isocurvature fluctuations matches α_{iso} is [85, 86]

$$H_I^{iso} = 10^8 \text{ GeV} \left(\frac{\alpha_{iso}}{0.02} \frac{0.24}{\Omega_a^{\text{mis}}} \frac{f_a}{10^{12}\,\text{GeV}} \right)^{\frac{1}{2}}, \tag{65}$$

where dropped the dependence of the axion DM relic density on α (the dependence on α is anyway mild). This means that $H_I < H_I^{iso}$ if the value of α_{iso} is an upper bound. If axions make up the whole relic density, high scale inflationary models are therefore exclude in this scenario. Note that one can require Ω_a^{mis} to be smaller than the observed DM abundance to relax this bound. However, there is always a lower bound on the abundance produced via the misalignment mechanism because, as discussed, $\theta_0 \geq H_I/(2\pi f_a)$.

4 Axions from Strings

In the previous chapter we studied the production of axion dark matter in the cosmological scenario where both H_I and T_R are smaller than f_a, in which case the axion field $a(t, \vec{x})$ could be considered constant in every Hubble patch (up to small anisotropies due to the quantum fluctuation of the axion and the inflaton during inflation). If instead $T_R > f_a$, the PQ symmetric phase is generically restored in the Universe after inflation by finite temperature effects (and then broken again when T drops below f_a). As we will discuss further below, this means that the axion field had random fluctuations over the observable horizon. Similarly, if $H_I > f_a$, quantum fluctuations during inflation are big enough not to be neglected anymore. In this scenario one needs to analyze the evolution of the spatially inhomogeneous axion field, in particular of all Fourier modes. While this might be doable at least in the regime where the axion potential is negligible (or linear) because in this case all the modes are decoupled, being the axion a NGB the field $a(t, \vec{x})$ takes values on a circle and topologically nontrivial configurations, called axion strings [22–29], are allowed and generically present. As we will see, a self-consistent treatment of these configurations in a field theory framework requires to study at the same time the axion field and the UV degrees of freedom that resolve the singularities associated to the center of the strings. The result is that the equations of motion one needs to solve are highly nonlinear, and analytical methods are unfortunately not applicable anymore.

In this chapter we will first give an overview on the dynamics of axions strings and domain walls in the early universe. While some of the features we will describe are only qualitative (mainly due to disagreements in the literature), the overall picture is accepted. We will then analyze the dynamics of axion strings numerically, and discuss the implications for the axion DM abundance.

4.1 Formation and dynamics of strings and domain walls

Except in particular cases, see e.g. symmetry non-restoration at finite temperature [87, 88], at high enough temperature a spontaneously broken $U(1)$ symmetry such as the PQ symmetry is restored [89]. The prototypical example of this phenomenon is the theory of a complex scalar field ϕ with a mexican-hat-like potential $V(\phi) = \frac{\lambda}{4}(|\phi|^2 - v^2)^2$. At finite temperature the potential gets (among others) the contribution $m_T^2|\phi|^2 = \frac{\lambda}{6}T^2|\phi|^2$, interpreted as a thermal mass for ϕ. Therefore, at finite temperature the vev $\langle\phi\rangle_T$ depends on T and, if $T > \sqrt{3}v$, $\langle\phi\rangle_T = 0$, which means that the $U(1)$ symmetry of the complex scalar field is restored. At $T \leq \sqrt{3}v$ the vev $\langle\phi\rangle_T$ becomes non-zero and a phase transition (with order parameter $\langle\phi\rangle_T$ itself) happens. While at temperatures bigger than the critical temperature the relevant degrees of freedom are the real and imaginary parts of ϕ, after the phase transition the angular and radial parts of ϕ become the mass eigenstates. In particular, since different Hubble patches of the Universe are causally disconnected, at the time of the phase transition the angular mode will be in principle generated with different values in different patches. These patches subsequently reenter the horizon and so, after a few Hubble times from the phase transition, the angular mode (which has no potential) will have random fluctuations within every Hubble patch [27, 28]. This prototypical example applies to the case of axion theories if the angular mode is identified with the axion and the radial part with some of the UV degrees of freedom of the fundamental theory. Note that neglecting the axion potential is a good approximation at high temperatures, in particular at $T \sim f_a \gg \Lambda_{QCD}$.

Since the axion is a NGB, it appears in the fundamental theory as a periodic variable with period v, which is a model-dependent parameter, in the previous case being for instance the period of the angular variable of the PQ field ϕ. In general $v = N_W f_a$ differs from f_a by an integer factor N_W (for

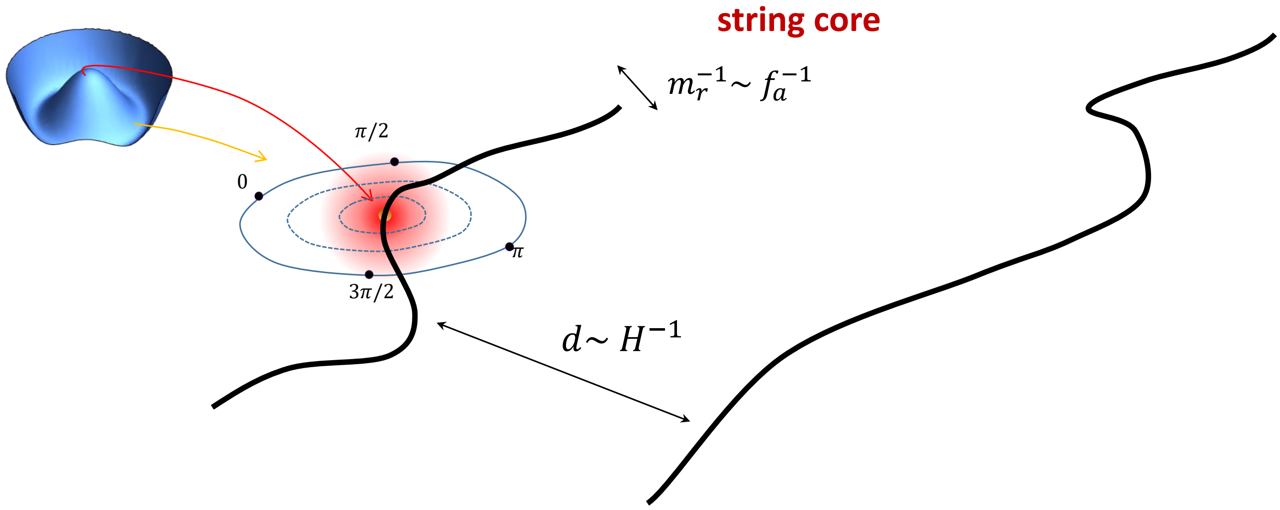

Figure 2: *A generic axion string configuration in the early universe. Around the string the axion (normalized to v) wraps the fundamental domain* $[0, 2\pi]$ *once.*

KSVZ-like models [4,5], $N_W = N$ is the number of fermions charged under PQ). Given its periodicity and the random fluctuations described above, there will be loops in coordinate space around which the axion wraps nontrivially the domain $[0, 2\pi v]$. By continuity in at least one point somewhere inside the loop the axion field assumes all the values between 0 and $2\pi v$, i.e. it is singular (see fig. 2). Applying the same argument to the whole space, it is easy to see that there must be a line of points, called axion string, where the axion is singular and around which it shifts of $2\pi v$. While in the NGB theory alone the energy density of this configuration is UV divergent due to the singularity, when considering the heavy degrees of freedom that complete the axion theory this UV divergence is cut-off by their mass m_r, of the order of f_a for $O(1)$ parameters. In the case of a a single complex scalar field with a mexican-hat potential, as detailed in the next section, typical solutions of the equations of motion have the radial mode (whose mass is m_r) at the maximum the potential for a distance m_r^{-1}, called string core, and at the minimum of the potential at spatial infinity.

The properties of the axion strings are critically affected by the fact that they come from a global (as opposed to local) $U(1)$ symmetry. As we will see in the next section, this means that the energy per unit length, called string tension μ, of an isolated string gets a contribution from the core where the radial is on the top of the potential, but also a logarithmic infrared divergence[28]. In the early Universe the divergence is cut-off by the distance to a string of opposite orientation, or for a small loop its diameter, both of which are typically of order of the Hubble length H^{-1}, which means that at a first approximation $\mu \sim \pi f_a^2 \log \frac{m_r}{H}$, see e.g. [22]. Although single string configurations are stable, when taken together strings can interact between each other, for e.g. forming loops that subsequently shrink. The result of this process is the reduction of the total string length and the conversion of the energy stored in the strings into primarily axion radiation (the only light degree of freedom) with momentum of order the Hubble horizon.

As we will discuss in much more detail, after being formed at the PQ phase transition (or due to

---

[28] The reason of this divergence is the same as for a charged wire in electrodynamics and is related to the nonvanishing gradient of the axion field at arbitrary distances from the center of the string.

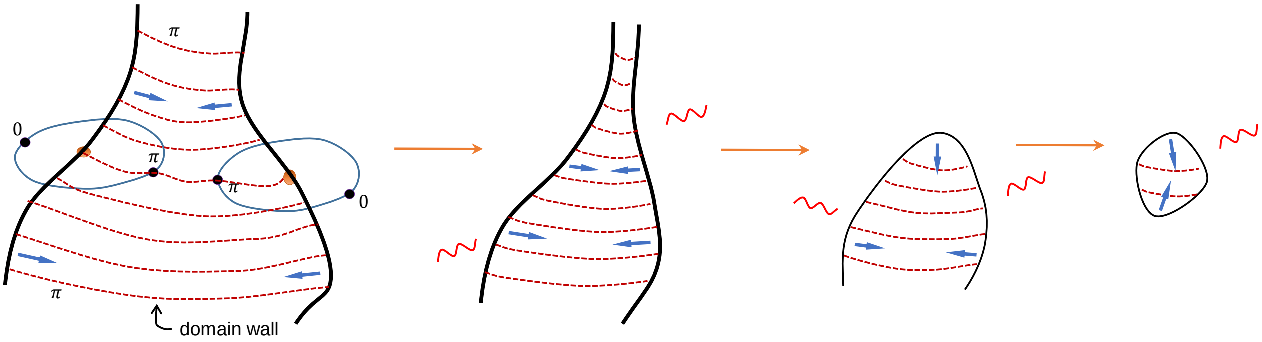

Figure 3: *Domain wall configurations with* $N_W = 1$*. Domain walls (in red) are attached to strings and in* $O(1)$ *Hubble times the full system shrinks into axion radiation.*

inflationary quantum fluctuations if $H_I > f_a$) a string network is thought to enter a scaling regime [30–35] where there is approximately a constant number of strings per Hubble volume. To maintain such a regime the string network must constantly release energy in radiation. After the axion gets a mass at the QCD crossover, this radiation will mostly become nonrelativistic, and is interpreted as axion particles that make up a part of the CDM relic abundance we observe today.

The scaling regime lasts until Hubble and the axion mass become comparable, at which time the axion mass becomes cosmologically relevant and cannot be neglected anymore in the evolution of the system, see eq. (59). As discussed in the previous chapter, at this time the temperature of the Universe T_* in eq. (62) is about Λ_{QCD}, and $\log(m_r/H) \gtrsim 70$ if $m_r \sim f_a$ and given $H \sim \Lambda^2_{QCD}/M_P$ and the lower bound on f_a from astrophysics. At this time the axion potential term leads to the formation of a system of domain walls attached to strings [23, 44]. Indeed, the spatial points where the axion field assumes the value $(2n+1)\pi v/N_W$, with $n = 0, \dots, N_W - 1$ integer, form 2-dimensional surfaces where the axion on the top of its QCD potential (see fig. (3)), called domain walls.[29] Given that m_a is the characteristic scale of the potential, the regions with distance $\lesssim m_a^{-1}$ from the points $(2n+1)\pi v/N_W$ have more potential energy density w.r.t. the rest of the space, which means that domain walls have a thickness of order m_a^{-1}. While at $T = T_*$ the thickness is comparable to Hubble by definition of T_*, it rapidly decreases with time (but still remaining much bigger than the string cores). The energy stored in the domain walls is usually measures in terms of their tension (energy per unit area), which for the case of the QCD axion (with cosine potential) is $\sigma = 8 m_a f_a^2$, while $\sigma \sim 9 m_a f_a^2$ for the zero-temperature potential [97].

If $N_W = 1$ each string is attached to just one domain wall centered in $a = \pi f_a$, and the configuration of strings and domain walls is automatically unstable. Indeed, since the system evolves to minimize the domain wall (and string) tension, the domain wall pulls the strings together as in fig. 3, which therefore annihilate in $O(1)$ Hubble times. The energy stored in strings and domain walls is emitted in additional radiation. On the other hand, if $N_W > 1$ the system is topologically stable because on each

---

[29] Formally speaking, a domain wall is a solution $\varphi(x)$ of the field equations that interpolates between two vacua φ_0 and φ_1, i.e. $\varphi(x_0) = \varphi_0$ and $\varphi(x_1) = \varphi_1$ for some x_0 and x_1. Since φ_0 and φ_1 are minima, between x_0 and x_1 there will be a region (a 'wall') where the field is at a maximum and therefore where some potential energy is stored. In the case under consideration, for $N_W = 1$ the vacua are $a = 0$ and $a = 2\pi f_a$, which correspond to the same point in field space, so x_0 and x_1 can be the same spatial point, as indeed happens by definition for axion strings, see fig. (3). For $N_W > 1$ instead domain walls interpolate between different vacua.

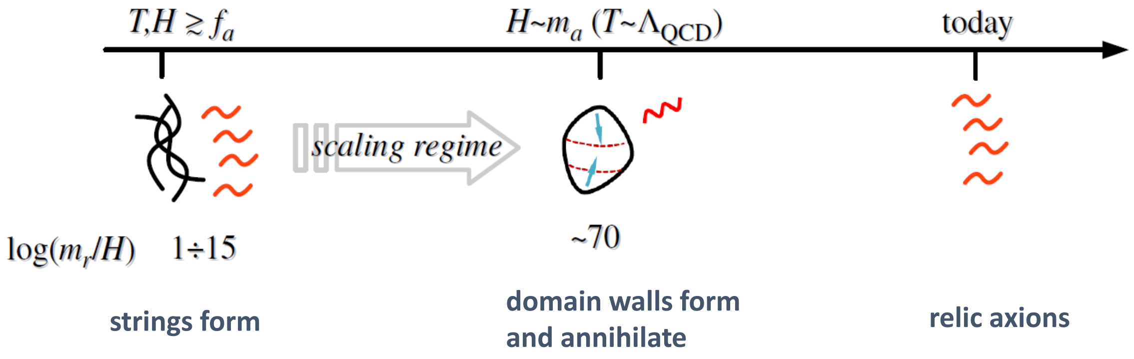

Figure 4: *Evolution of string and domain walls in the scenario* $N_W = 1$*.*

string end $N_W > 1$ domain walls, each of which tends to pull the string in a different direction resulting in an equilibrium. Therefore strings and domain walls do not annihilate and, as the Universe expands, more and more domain walls will reenter the horizon, eventually dominating the energy density of the Universe. Thus, this scenario is ruled out unless one introduces additional breaking of the PQ symmetry (that might reintroduce the Strong CP problem). This would remove the degeneracy of the N_W vacua that the axion scans in a loop around the string, and make the pull from the tension of one domain wall win over the others. In the following we will consider $N_W = 1$ so that the PQ breaking scale $v = f_a$, however, since we will focus on the early evolution of the string network, the general case can be recovered by simply replacing f_a with $v = N_W f_a$.

Summarizing, a network of axion strings is formed when $T \sim f_a$ as a result of the thermal PQ phase transition or if $H \sim f_a$ from inflationary quantum fluctuations (see fig. (4)). At this time $\log(m_r/H)$ ranges from ~ 15 in the first case to ~ 1 in the second. Soon after its formation, the network enters a scaling regime until the axion mass is comparable to Hubble (at about the QCD crossover), at which time $\log(m_r/H) \sim 70$ and domain walls form and annihilate the system of strings and domain walls. During the whole process, in particular during the scaling regime and at the time of the domain wall annihilation, energy is converted from strings and domain walls to axion radiation. After the QCD crossover the axion gets a mass and most of the radiation becomes nonrelativistic. The comoving number density of axions is conserved, and the coherent axion field behaves as cold DM.

Soon after the domain walls form, their thickness becomes small compared to Hubble and therefore in most of the volume of the Universe there will be no domain walls (or strings). Historically, this has motivated to treat separately the axion field away from string and domain walls, which is expected to relax to the minimum through the misalignment mechanism. In particular, due to the presence of strings the axion field will scan all the values in the fundamental domain $[-\pi f_a, \pi f_a]$, and contrary to the discussion of sec. 3.2 there is no free parameter in the initial conditions to adjust. A rough estimate of the contribution to the relic abundance might be calculated considering the field to be homogeneous (even if it is not) and using eq. (64), with θ_0^2 given by the average of θ^2 in the interval $[-\pi, \pi]$, i.e. $\frac{1}{2\pi}\int_{-\pi}^{\pi}\theta^2 = \pi^2/3$, which is $\overline{\theta} \approx 2.15$ when considering also the inharmonicities of the potential. As as result, all the values of f_a bigger than 10^{12} GeV are excluded in this scenario otherwise $\Omega_a^{\rm mis}|_{\theta_0=\overline{\theta}}$ would be bigger than the measured dark matter energy density.

In this estimate we neglected strings and domain walls, which however store a relevant part of the energy density of the field (e.g. the energy density in the gradient). In particular, the evolution of the string network during the scaling regime and the subsequent decay of the string-domain wall network is expected to inject additional energy in the form of radiation. Moreover, these waves are expected to have momentum of order the Hubble horizon at T_* and therefore will become non-relativistic soon after the QCD crossover, because of the increase of the axion mass. While the radiation emitted by strings during the scaling regime might be treated separately, as soon as the domain walls form there is no clear distinction between axions produced from the misalignment and from the string-domain wall network. The naive expectation is that the abundance is at least $\Omega_a^{\rm mis}|_{\theta_0=\bar{\theta}}$, but it is most likely much more because of the contribution from topological defects.

While this scenario has no free parameters in the initial conditions and therefore Ω_a should be only a function of f_a (up to the milder dependence on α and $g(T_*)$), computing the abundance is challenging. In principle one has to solve (e.g. numerically) the full equations of motion for the inhomogeneous axion field, including a possible UV completion that resolves the core. In practice, however, the huge scale separation $\log(m_r/H) \sim 70$ in the physically relevant regime presents an immediate problem in attempting to study the system using numerical simulations. To resolve the dynamics of the strings, a 2-dimensional slice of the lattice perpendicular to a string must contain at least a few grid points inside the string core, and to capture the interactions and dynamics of strings a few Hubble volumes must be simulated (we show this by analyzing the systematic errors in simulations in Appendix D). Given the computational power available, the largest grids that can be simulated have $N^3 \sim 1000^3$ lattice points. Consequently, the maximum scale separations that can be directly studied correspond to $\log(m_r/H) \sim 6$. In this system, the tension of strings, and their couplings[30] to axion radiation and heavy degrees of freedom, are far from the physically relevant values. Indeed, even if these could be adjusted to the physical values by modifying the UV theory this would not be sufficient. For example, the properties of the string network depend on whether collapsing string loops rebound and oscillate many times before disappearing, and if strings that approach each other recombine. In order to accurately capture such dynamics, processes on all scales between Hubble and the string core size must be resolved.

In the next sections we consider the simplest UV completion of the QCD axion theory and study the string network before the axion mass turns on. An understanding of this stage of its evolution is crucial, both to calculate the relic abundance of axions produced at such times and to set the appropriate initial conditions when analysing the system once the axion mass becomes relevant. In particular, we will calculate the number density n_a of axion waves produced by strings during the scaling regime, which will be relativistic because at those times the axion is massless. As a reference, we will directly compare n_a with the misalignment number density defined by convention as $n_a^{\rm mis} = 2n_*|_{\theta_0=\bar{\theta}} = H(t_*)\bar{\theta}^2 f_a^2$, where n_* is the number density in eq. (63) at the time instant $t = t_*$ when the axion mass becomes cosmologically relevant. In the following analysis, a possible prediction for the DM relic density from strings will then implicitly assume that all the axions produced during the scaling regime (as we will see, only the last instants will matter) will become nonrelativistic at t_* with the same number density[31] n_a, which is therefore conserved.

---

[30] The large scale separation between m_r and H also suppresses the coupling between strings and axions by the same logarithmic factor [38], and is expected to render emission of heavy modes associated to the theory's UV completion irrelevant.

[31] While this is true (up to $O(1)$ transient factors) if axions are free and have momentum of order Hubble, there might be nontrivial effects coming from the fact that the axion potential is bounded from above and from the possible interaction of these waves with domain walls.

4.2 Axion Strings and Simulations

We consider a complex scalar field ϕ taken to have the $U(1)_{\rm PQ}$ invariant Lagrangian

$$\mathcal{L} = |\partial_\mu \phi|^2 - V(\phi)\,, \qquad \text{with} \quad V(\phi) = \frac{m_r^2}{2f_a^2}\left(|\phi|^2 - \frac{f_a^2}{2}\right)^2 , \tag{66}$$

in a spatially flat Friedmann-Robertson-Walker background. The metric is $ds^2 = dt^2 - R^2(t)dx^2$, and the Universe is assumed to expand as in radiation domination, so the scale factor $R(t) \propto t^{1/2}$, and the Hubble parameter $H \equiv \dot{R}/R = 1/(2t)$.

The potential $V(\phi)$ leads to ϕ getting a $U(1)_{\rm PQ}$ breaking vacuum expectation value (VEV) $|\langle\phi\rangle|^2 = f_a^2/2$. We decompose

$$\phi(x) = \frac{r(x) + f_a}{\sqrt{2}}\, e^{i\frac{a(x)}{f_a}} , \tag{67}$$

into the radial field $r(x)$, which has a mass m_r, and the axion field $a(x)$, which has a period $2\pi f_a$. Since we focus on the properties of the system at temperatures above the QCD crossover, the PQ breaking axion potential generated by QCD can be neglected and the axion is massless (at lower temperatures this must be added to eq. (66)).

The average Hamiltonian density $\rho_{\rm tot} = \langle T_{00}\rangle$ of the complex field ϕ is

$$\rho_{\rm tot} = \left\langle |\dot{\phi}|^2 + |\nabla\phi|^2 + V(\phi) \right\rangle , \tag{68}$$

where $\dot{\phi} = d\phi/dt$, ∇ is the gradient with respect to the physical spatial coordinates $R(t)x$, and $\langle A\rangle \equiv \lim_{V\to\infty} \frac{1}{V}\int_V d^3x\, A$ is the spatial average of A. After decomposing ϕ as in eq. (67),

$$\begin{aligned}\rho_{\rm tot} = &\left\langle \frac{1}{2}\dot{a}^2 + \frac{1}{2}|\nabla a|^2 \right\rangle + \left\langle \frac{1}{2}\dot{r}^2 + \frac{1}{2}|\nabla r|^2 + V(r) \right\rangle \\ &+ \left\langle \left(\frac{r^2}{2f_a^2} + \frac{r}{f_a}\right)\left(\dot{a}^2 + |\nabla a|^2\right) \right\rangle ,\end{aligned} \tag{69}$$

where $V(r) = \frac{m_r^2}{8f_a^2} r^2 \left(r + 2f_a\right)^2$. The terms on the first line of eq. (69) correspond to the kinetic and potential parts of the axion and radial modes' energies, and the term on the last line is the interaction energy between the two. In the small field limit $|r|/f_a \to 0$, the Hamiltonian can be approximated as the sum of that from decoupled axions and radial modes (i.e. by the first line of eq. (69)). However, away from this limit interaction terms between the two fields are not negligible and make the axion-radial system strongly coupled.

Note that the field's equation of motion

$$\ddot{\phi} + 3H\dot{\phi} - \nabla^2\phi + \phi\frac{m_r^2}{f_a^2}\left(|\phi|^2 - \frac{f_a^2}{2}\right) = 0\,, \tag{70}$$

does not depend on the ratio m_r/f_a directly. Indeed the dependence on the two scales f_a and m_r can be reabsorbed by rescaling respectively the field $\phi \to \phi f_a$ and the space-time coordinates $t \to t/m_r$ and $x \to x/m_r$. Therefore, up to a trivial field rescaling, the physics is only sensitive to the ratio $m_r/H = 2m_r t$.

The equations of motion in eq. (70) admit solitonic string-like solutions. As mentioned in the previous section, these are topologically non-trivial configurations that contain loops in space where the axion field a wraps the fundamental domain $[0, 2\pi f_a]$ non-trivially. The prototype of such solutions is a static, infinite, string lying along the z-axis. In cylindrical coordinates (ρ, θ, z) this is given by

$$\phi(x) = \frac{f_a}{\sqrt{2}} g(m_r \rho) e^{i\theta} \ , \tag{71}$$

where g is a profile function that satisfies $g(\rho) = C\rho + \mathcal{O}(\rho^3)$ for $\rho \to 0$ and $g(\rho) = 1 - \rho^{-2} + \mathcal{O}(\rho^{-4})$ for $\rho \to \infty$. The string core is defined as the region in which ϕ is close to the maximum of its potential, i.e. when $r/f_a \sim -1$, which corresponds to points at a distance less than m_r^{-1} away from the centre of the string $\rho = 0$. In this part of space the axion-radial mode system is strongly coupled, and all of the terms in eq. (69) contribute to the string energy density. However, for a single string configuration the axion energy density diverges logarithmically for $\rho \to \infty$ due to the angular gradient $\frac{1}{2}\langle |\nabla a|^2 \rangle$. Consequently, the total string energy is dominantly in this component, and is mostly stored away from the string core. In the early Universe this leads to the string tension

$$\mu \sim \pi f_a^2 \log\left(\eta \frac{m_r}{H}\right) \ , \tag{72}$$

where η is a $O(1)$ numerical factor that takes into account the average curvature of the string system.

To analyze the dynamics of the string system in the early Universe, we numerically integrate the equations of motion, eq. (70), in 3+1 dimensions. Starting from suitable initial conditions, for example ϕ random with sufficiently large fluctuations, axion strings automatically form and evolve. In doing so, we are assuming that solutions of *classical* equations of motion capture the physics of strings and axion radiation. This is justified because strings are themselves intrinsically classical and the relevant part of the axion radiation has large occupation number.

The complex scalar ϕ is discretized on a lattice with approximately $N^3 = 1250^3$ grid points, and evolved in fixed steps of conformal time $\tau \sim \sqrt{t}$. The simulation is carried out with periodic boundary conditions, and in comoving coordinates x, so that the comoving distance between grid points remains constant and the physical distance between grid points grows $\sim \sqrt{t}$. Further details of the algorithm used are given in Appendix C. As the system is evolved forward, the number of Hubble lengths contained in the box side decreases $\sim 1/\sqrt{t}$ and the number of lattice points inside a string core also decreases $\sim 1/\sqrt{t}$, as shown in Figure 5. The maximum accessible scale separation corresponds to an upper bound on the final time that can be simulated. Other possible sources of numerical uncertainty include the time step used in the simulation and the way that the contribution of the string energy is excluded from the calculation of the energy in free axions, and a full analysis is given in Appendix D.

The simulations that we carry out with $m_r/H \sim 1$ can be interpreted in terms of a theory with $m_r \approx f_a$ at a time when $H \approx f_a$, however this is not a physically relevant part of parameter space. First, for such models the scale separation simulated is far from that at the time of the QCD crossover, when the majority of axions produced by the string network are emitted. Additionally, this regime does not correspond to a system that is realizable even at early times: when $H \sim f_a$, the temperature of the Universe is $T \sim \sqrt{f_a M_{\rm Pl}}$, where $M_{\rm Pl}$ is the Planck mass, and a physical theory will be in the unbroken, PQ symmetric, phase. However, by studying the potential eq. (66) without including finite temperature effects, we can use the results obtained to extrapolate the properties of the string system to low temperatures, when such corrections are actually negligible. In particular, at the time of the QCD crossover $T \sim$ GeV and finite temperature corrections to the potential of the radial scalar are irrelevant.

Alternatively, simulations at a scale separation $m_r/H \sim 1$ can be interpreted as studying a model with $m_r \ll f_a$ at a time when the Hubble parameter is $H \ll f_a$. Axion theories with a light radial

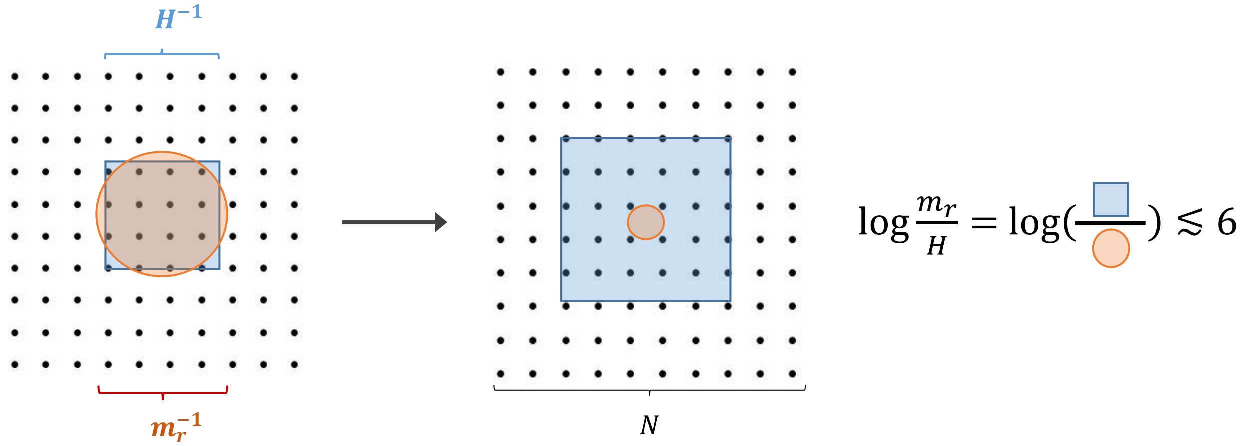

Figure 5: An illustration of how the size of a string core, shaded red, and a Hubble volume, shaded blue, evolve relative to the lattice points in our simulations, where N is the number of lattice points in a spatial dimension. Requiring that the simulation contains at least a few Hubble volumes and that a string core contains at least ~ 1 lattice point constrains the maximum scale separation that can be studied.

mode are less commonly considered, but in this case the simulations are directly analyzing a physically realizable point in parameter space. In particular, taking $m_r \sim 10^{-18}\,\text{GeV}$, the Hubble parameter is that at the time of the QCD crossover (although such a light scalar is excluded by fifth-force experiments).[32] Meanwhile, theories in an intermediate regime, with a radial degree of freedom that has a mass $10\ \text{keV} \lesssim m_r \lesssim f_a$, are not experimentally excluded by evolution of stars and fifth-force experiments. These correspond to scale separations at the time of the QCD crossover that are larger than can be reached with simulations, so that extrapolation is still required, but which are smaller than if $m_r \approx f_a$.

The Lagrangian in eq. (66) includes only one heavy scalar degree of freedom, and it is clearly not the most generic that can arise in UV-complete axion theories. However, when $H \ll m_r$ the dynamics of axion strings and radiation are expected to be largely independent of which massive degrees of freedom are included, since these only get excited when strings interact over distances of order m_r^{-1}, e.g. when loops shrink or long strings intersect. The energy of a global string configuration remains logarithmically divergent in more complicated theories, since this comes from the axion angular gradient, which is always present. In principle one could also include interactions with other fields that ϕ is coupled to — e.g. SM fields — in eq. (66), but away from the string cores the couplings of axions to these are suppressed by powers of f_a, and are negligible. Meanwhile the radial mode can have order 1 couplings to SM fields, however it is expected to decouple from the dynamics of the strings and axions at large scale separations. At early times, when the temperature is high, interactions of strings with the visible sector thermal plasma could modify their dynamics [134], however these effects will also be negligible at temperatures around the QCD crossover.

As well as the physical system, the literature has often used a deformed theory in which the mass

---

[32] In such models the PQ phase transition still happens at $T \sim f_a$, and finite temperature corrections to the potential of the radial mode are not important, provided that it is weakly coupled to states in the thermal bath.

of the radial mode in eq. (66) is replaced with a time dependent one

$$m_r\left(t\right) = m_i \frac{R(t_i)}{R(t)} = m_i \sqrt{\frac{t_i}{t}} \ , \tag{73}$$

where m_i is the mass of the radial mode at the initial time t_i. This is equivalent to a theory with a quartic coupling $\lambda \sim m_r^2(t)/f_a^2$ that decreases with time, and is known as the *fat string* scenario. The size of the string cores increases as $m_r^{-1}(t) \sim t^{1/2}$, and the number of lattice points inside a string core remains constant throughout a simulation. The maximum scale separation that can be simulated is unchanged compared to the physical potential, however the time taken to reach a particular scale separation (starting from $H \sim m_r$) is increased, and simulations can be run for longer before arriving at the upper bound. As a result, energy left over from the initial conditions is redshifted more, and there is more time available for properties of the string system to reach their asymptotic behavior (there are additional benefits that will be seen in Section 4.4). Despite these advantages, we stress that by making the potential time dependent the equations of motion of the system are changed by an order 1 amount. Consequently, although the dynamics of axion strings might remain qualitatively similar to those of the physical Lagrangian this is not guaranteed, and the numerical values of the parameters of the scaling solution are not expected to be the same in the two cases. We perform simulations using both techniques and discuss their advantages and disadvantages.

In both the fat string and the physical scenarios, the axions produced in the simulations are massless, and their energy densities redshift as $\sim 1/R\left(t\right)^4$. Meanwhile, the radial modes produced are highly non-relativistic. In simulations with the physical Lagrangian their energy density redshifts as $\sim 1/R\left(t\right)^3$. In simulations of the fat string Lagrangian the scalar mass decreases with time, and the energy density of these states redshifts as $\sim 1/R\left(t\right)^4$, the same as axions.

4.3 The Scaling Solution

It has long been claimed that a system of axion strings is driven towards a particular solution, which is independent of its initial conditions [27–29]. Indeed, this feature is crucial for making predictions about the properties of the string system at late times, and in particular of the axion relic abundance from strings, that do not depend on the dynamics of the system at early times, which are model dependent.

The existence of such an attractor solution is simple to motivate qualitatively. Strings can lose the energy stored in their length by radiating axions and radial modes. Therefore bends in strings with curvature larger than the Hubble scale tend to straighten, and closed string loops smaller than the horizon are expected to disappear, emitting radiation. Additionally, long strings (or equivalently string loops larger than the horizon) can interact when they enter each other's horizon through a process called recombination: when strings cross they can recombine into a new configuration with a lower tension, and similarly a region of high curvature in a long string can split off forming an isolated loop. The net effect is a reduction of the total string length and the production of smaller loops, string segments with larger curvatures, and radiation. The rate at which such processes occur depends on the density of strings within the horizon. Below some critical density recombination is inefficient. In this case, the number of strings in each Hubble patch increases as the Universe expands and new strings enter the horizon. On the other hand, above a critical density recombination becomes efficient, reducing the number of strings within the horizon. As a result, the density of strings is pushed towards a particular (not necessarily time independent) value. Other statistical properties of the network, such as the distribution of the string density in loops of different length, are expected to converge similarly to a common behavior.

We define the average number of strings per Hubble patch $\xi(t)$ as

$$\xi(t) \equiv \lim_{L\to\infty} \frac{\ell_{\rm tot}(L)\, t^2}{L^3} \, , \tag{74}$$

where $\ell_{\rm tot}(L)$ is the total length stored in strings in a volume L^3. Hence the energy density of strings is

$$\rho_s(t) = \xi(t) \frac{\mu_{\rm eff}(t)}{t^2} \, , \tag{75}$$

where, given eq. (72), the effective string tension $\mu_{\rm eff}(t)$ is expected to be

$$\mu_{\rm eff}(t) = \gamma(t)\, \mu_0 \log\left(\frac{m_r\, \eta(t)}{H\sqrt{\xi(t)}} \right) , \tag{76}$$

with $\mu_0 = \pi f_a^2$, and $\gamma(t)$ the effective boost factor associated to the kinetic energy of the string configuration, which, as we will see, is $\mathcal{O}(1)$. The factor $m_r/(H\sqrt{\xi})$ is anticipated to capture the main time dependence of $\mu_{\rm eff}(t)$ since the logarithm is cut-off by the average distance between strings ($\propto t/\sqrt{\xi}$). The remaining time dependence is encoded in the factor $\eta(t)$, which takes into account the non-trivial shape of the strings and is expected either to be constant or to have at most a very mild time dependence. Indeed, it will be a non-trivial check of the string network's properties that the energy density in strings extracted from the simulations is well reproduced by eqs. (75) and (76).

The existence of the scaling law (75) with constant $\xi(t) = \xi_0$ can easily be understood for local strings. For these the string tension is localized on the core and is constant $\mu_{\rm eff}(t) = \mu$. Neglecting the core size the problem only has one scale, $H = 1/2t$, which completely fixes the scaling law for the energy density $\rho_s(t) = \xi_0 \mu/t^2$. The presence of a single scale suggests that during the scaling regime all the properties of the string configuration should be scale invariant.

On the other hand, for the global case several properties of the strings, including their tension and their coupling to axions [38], depend logarithmically on the core size m_r, therefore logarithmic corrections to the scaling law can be expected. To account for these effects we leave an explicit time dependence both in $\xi(t)$ and in $\eta(t)$, besides the one contained in m_r/H inside $\mu_{\rm eff}(t)$.

In the rest of the section we will establish the existence of the (approximate) scaling solution for axion string networks, and study its properties in detail. In particular we will present results from numerical simulations demonstrating the presence of the attractor, its independence from the initial conditions, the behavior of the parameter $\xi(t)$, and the distribution of loops and long strings during the scaling regime.

4.3.1 The Attractive Solution

The existence of the attractor can be tested by studying if different statistical properties of the string network converge to the same values independently of the initial conditions of the field. As a representative example, here we focus on the evolution of the average number of strings per Hubble patch $\xi(t)$. In Appendix F we present additional results showing that the number density, the total and the instantaneous spectrum of axions emitted from the string network also clearly converge, regardless of the initial conditions. The convergence is particularly evident in the instantaneous emission spectrum, which depends only on the string configuration at a particular moment and has no memory of earlier times.

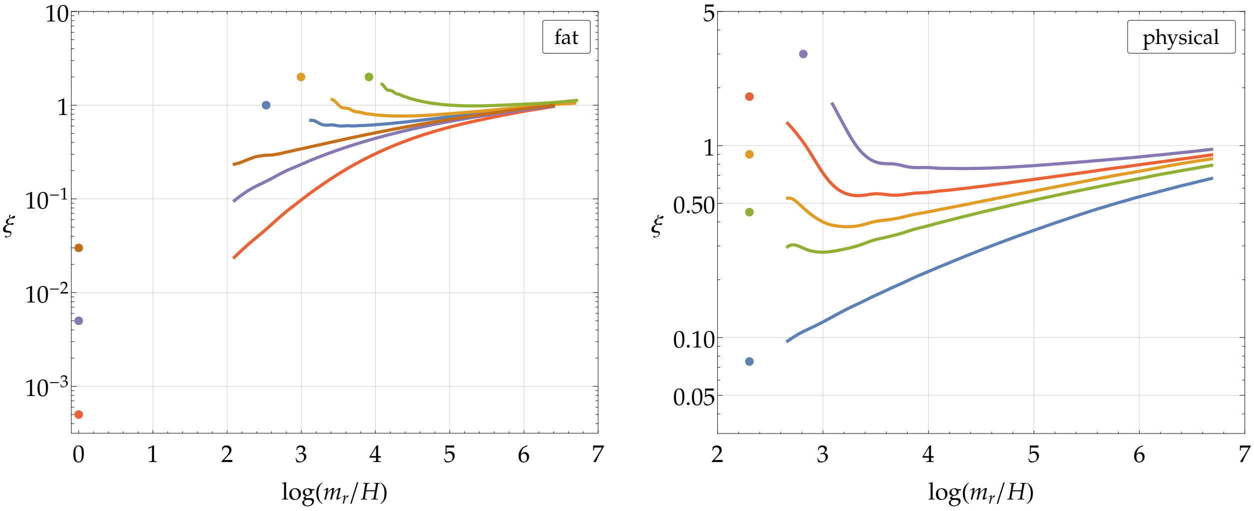

Figure 6: The evolution of the average number of strings per Hubble patch $\xi(t)$ as a function of time (here represented by Hubble) using the fat string trick (for which $m_r \propto t^{-1/2}$, left) and for the physical case ($m_r = \text{const}$, right) for different initial conditions. Each curve corresponds to the average of many simulations with the same initial value of ξ. The position of the coloured dots indicates the initial time and the value of ξ for each simulation set. The number of simulations has been taken large enough that the statistical uncertainties are smaller than the thickness of the curves.

We set the initial conditions in two different ways. In the first, we just generate sets of random fields. In the second, we construct initial conditions with a fixed number of strings by evolving random configurations until the total string length in the box reaches a required value, and then we reset the clock rescaling Hubble (more details about the procedure can be found in Appendix C). Besides allowing us to start simulations with a predetermined density of strings, the second method produces initial field configurations that have less primordial background radiation (although this would redshift away anyway) and more suitable for a cleaner study of the instantaneous spectrum. In Appendix F we show that the properties and evolution of the string network are independent of the way in which the initial conditions are set.

In Figure 6 we show the evolution of ξ with time in simulations, starting from initial conditions with different numbers of strings using the second method described above. ξ has been computed at different time shots from its definition eq. (74), using the algorithm described in Appendix C. As discussed in Section 4.2, it takes a longer time to get to the same value of $\log(m_r/H)$ in the fat string scenario than for the physical theory. As a result the attractor regime is reached at smaller values of the log in the fat string case (here and in the rest of the paper we sometimes use the short-hand notation "log" to refer to $\log(m_r/H)$).

In both the fat string and the physical models, the convergence towards a common value of ξ is manifest. In the fat string case, the initial values of ξ span more than three orders of magnitude and, by the end of the simulations, they lead to the same value of ξ to within 10%. For the physical case the convergence is a little slower but it is still clear. In simulations starting at $H = m_r$ it is not possible to initially have more than one string per Hubble patch. As a result, to achieve initial conditions with a clear overdensity of strings we started such simulations later, when $H < m_r$. The corresponding

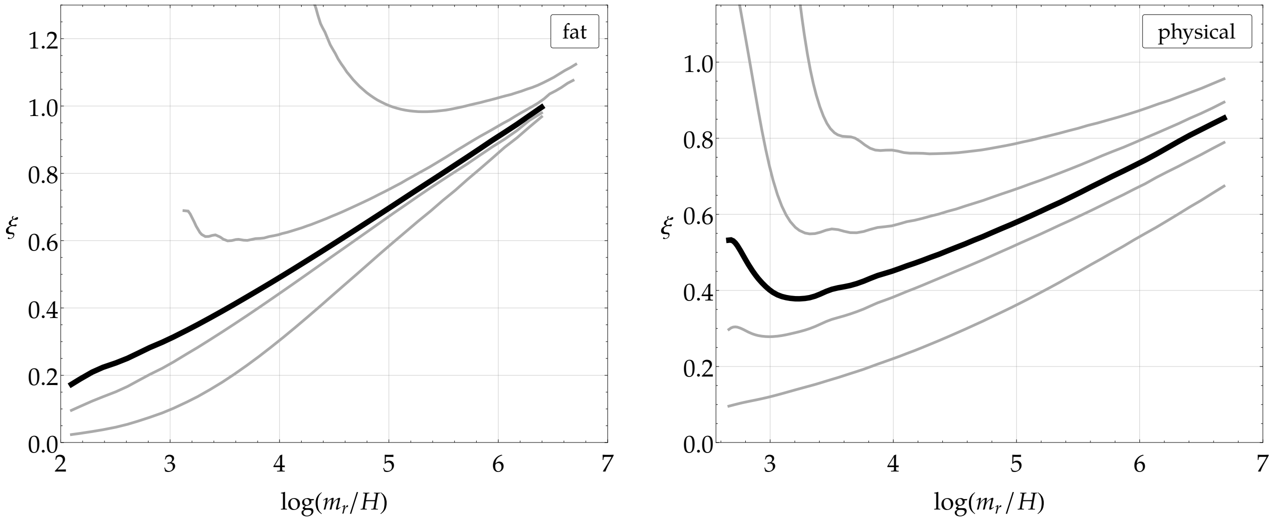

Figure 7: The growth of $\xi(t)$ with time ($\log(m_r/H)$) for different initial conditions using the fat string trick (left) and in the physical case (right). The black curves correspond to the initial conditions that are the closest to the attractor solution.

data in the Figure 6 have the initial points (the coloured dots) at larger values of $\log(m_r/H)$.

The network of global strings was first studied using field theoretic computer simulations in [41,42], and more recently over a longer time range in [135]. Ref. [136,138] showed evidence for the existence an attractor for the fat string system, respectively in two and three dimensions. Our simulations have a similar time range and constitute an independent check of the convergence of ξ in the fat string system starting from a wide range of initial conditions, and a demonstration of the attractor's existence for the physical case. Further, in Appendix F, we show that other properties of the network, including the spectrum of axions emitted, also converge.

4.3.2 Scaling Violation

Having shown that the attractor solution exists, we now turn to study its properties. One prominent feature is that, although different boundary conditions converge to a common value of ξ, this value does not seem to be constant in time. To see the change more clearly, in Figure 7 we show the plots of Figure 6 on a linear scale. A growth of ξ linear in $\log(m_r/H)$ is evident both for the fat string and for the physical cases. The fact that simulations with an overdensity of strings first rapidly evolve to smaller values of ξ, converging to the attractor, and then start increasing again is particularly convincing. This strongly indicates that the growth is an intrinsic property of the scaling solution, rather than the sign that the attractor has not yet been reached.

The behavior shown in Figure 7 is compatible with the asymptotic form

$$\xi(t) = \alpha \log\left(\frac{m_r}{H}\right) + \beta \,. \tag{77}$$

In particular, at late times β is subleading, and the value of $\xi(t)$ is dominated by the logarithmic term. From Figure 7 it can be seen that the coefficient α is universal, independent of the initial conditions. Indeed, the derivatives $t\,\xi'(t) = \partial\xi/\partial[\log(m_r/H)]$ of the curves tend towards a common value α more rapidly than ξ itself.

We can use the convergence of the slopes to a constant value to select the optimal initial conditions, i.e. those for which the scaling regime is reached the earliest. The corresponding lines are those plotted in solid black in Figure 7, and curves starting from different boundary conditions reach the same constant slope at later times. Considering only simulations that reach the scaling regime (i.e. those that show a sufficiently large region of constant slope) we extract estimates for α

$$\alpha_{\text{fat}} = 0.22(2)\,, \qquad \alpha_{\text{phys}} = 0.15(5)\,. \tag{78}$$

Here the errors clearly have no statistical significance, but rather they represent our educated, conservative, guesses about the uncertainty. Plots showing the behavior of the slope for different boundary conditions, and more details about the fit, can be found in the Appendices F and G. We do not report the values of β, since the uncertainty on these from different initial conditions is larger, and they are irrelevant for the physically interesting values of the log.

As mentioned, a logarithmic violation of the scaling behavior is not completely unexpected, since several properties of the string network, including the string tension and coupling to axions, have a similar dependence. If such behavior is maintained at later times, as seems plausible, the average number of strings per Hubble patch will grow substantially for values of the log relevant to the QCD axion. For example, if $m_r \sim f_a$ eq. (78) would imply $\xi = 10.5 \pm 3.5$ at the physically relevant separation, $\log(m_r/H) \approx 70$. This value is an order of magnitude larger than that found in refs. [41, 42, 135] from numerical simulations on smaller grids, and the value that is usually assumed in rough estimates of the axion abundance produced by strings. Meanwhile, the extrapolated value of ξ that we obtain is only a factor of 2 larger than that recently obtained in ref. [136], which used the fat string trick and a different UV completion of the core. We stress that although the fat string system shows the same qualitative linear growth with the log as physical strings, the numerical parameters are somewhat different in the two cases. This is not surprising, and in extrapolating to the physical scale separation it is important to study the physical system, not just the fat string case.

The logarithmic enhancement in ξ was first observed and studied in the 2+1 dimensional simulations of ref. [137], but it was missed in the 3+1 dimensional ones of ref. [41, 42, 135], perhaps due to the use of smaller grids (which limited the time range of their simulations) and the choice of overdense initial conditions. Indeed, the combination of these two factors can produce a fake plateau at the intermediate values of the logs that were analysed (see e.g. the top curve of the right hand plot in Figure 7).

Conversely, the later simulations of the fat string case in ref. [138], made on larger grids, also observe a logarithmic increase, in agreement with our results. Finally, in the recent analysis of ref. [136], which use fat strings and a different UV completion of the core to partly include the effects of a large scale separation, the value of ξ increases with the log. The growth rate is not clear and the authors suggest that part of it may be due to spurious H/m_r effects. However, the results of this reference for purely global strings are in agreement with our numerical fit, and correspond to initially overdense networks.

We will resist the temptation to interpret the log-increase of $\xi(t)$ in terms of the reduction of the string-axion coupling or the increase of the string tension. In fact a similar growth also seems to be present for local strings, and can be seen in Figure 3 of ref. [136] and is hidden in Figure 7 of ref. [139].[33] Neglecting gravity, the only way that local strings can maintain a scaling regime is through the production of heavy modes associated to the core scale. The latter then cannot be neglected and its presence allows for a violation of the scaling law $\rho_s = \xi_0 \mu / t^2$ argued before. As we will discuss

---

[33] We have carried out preliminary simulations of local strings to study this effect, and these appear to confirm such a growth.

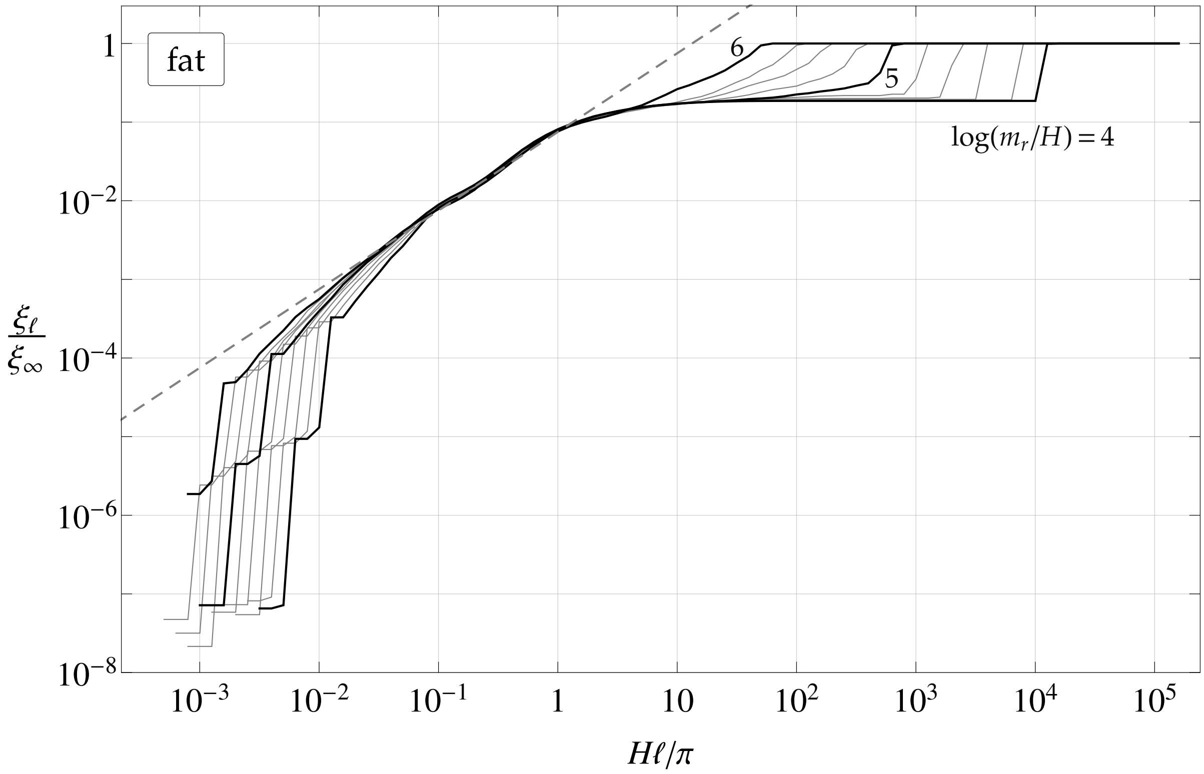

Figure 8: The fraction of the total string length ξ_ℓ/ξ_∞ that is contained in loops smaller than ℓ for different time shots.

below, the way that global strings lose energy is not so different from the one above, which might explain the similarity in the way that the scaling is violated.

4.3.3 Long *vs* Short: the scale-invariant distribution of loops

In order to further characterize the attractor solution, and to better understand its properties, we also study how the total string length per Hubble patch ξ is distributed over different loop sizes. If we call $dn_\ell/d\ell$ the loop number density, i.e. the number of loops per unit volume and per unit of loop length, then the quantity

$$\xi_\ell \equiv t^2 \int_0^\ell d\ell' \, \ell' \, \frac{dn_{\ell'}}{d\ell'} \,, \tag{79}$$

represents the contribution to ξ from loops of size smaller than ℓ and, in particular, $\xi_\infty = \xi$.

We have performed a large number of simulations using the fat string trick to acquire enough statistics to study ξ_ℓ as a function of time and loop size ℓ. The initial conditions were fixed as for the thick line in Figure 7 left, so that the system started close to the scaling solution.

The results for the ratio ξ_l/ξ_∞, which gives the percentage of string length contained in loops of size smaller than l, are plotted in Figure 8, and reveal several features of the string network. Most of the string length, more than 80%, is contained in loops much larger than Hubble, of order of the full box size (in fact it seems that most of the string length is contained in a single loop wrapped around the simulation box multiple times). This leads to the abrupt increase in the right part of the plot, which saturates ξ_ℓ to its asymptotic value ξ. Less than 20% of the total contribution to ξ is contained in loops of size of order H^{-1} and smaller, which results in the appearance of the plateau with $\xi_\ell/\xi_\infty \approx 0.2$ at $H\ell \gg 1$.[34] On the left of the plot the UV cut-off corresponding to the smallest possible loops, of order the core size, is also visible. As the Universe expands the physical size of the simulation box in units of Hubble shrinks, and as a result ξ_ℓ saturates its asymptotic value ξ at smaller and smaller values of $H\ell$. At the same time, the value of m_r/H grows so the UV cut-off moves to the left.

The lines in Figure 8 corresponding to different times approximately overlap for values of ℓ sufficiently far from the UV and IR cutoffs. Therefore, since ξ_∞ grows logarithmically with time, the corresponding growth in ξ_ℓ is homogeneous in ℓ. This signals that the logarithmic increase of ξ is equally distributed over all scales, and that the ratio between long and short strings stays constant in time (see also Figure 21 in Appendix D). The fact that ξ_ℓ/ξ_∞ remains constant in time for $\ell \lesssim 10\pi/H$ also shows that the number of loops of a particular length per Hubble patch does not change, apart from this logarithmic increase. As loops shrink and disappear (or recombine with other strings) they are replaced at the same rate by larger loops themselves shrinking, or by new loops being produced from interactions of long strings, which is an indication that the attractor solution has been reached. When, at the final times, the Hubble scale becomes of order of the box size there is no longer a sharp distinction between long and short strings.

Another feature of the loop distribution is evident from the plot: for loop lengths smaller than Hubble $\xi_\ell \propto \ell$ (the dashed line in the plot), so $d\xi_\ell/d\ell \propto \ell dn_\ell/d\ell = \text{const}$. This means that the number of loops in each logarithmic interval of length is constant, over almost two orders of magnitude. Equivalently, the 10% of the full string length contained in sub-Hubble loops is equally distributed over all loop sizes on a linear scale. This approximate power law seems to become a better fit as time progresses, suggesting that it is an intrinsic property of the attractor solution, and further confirming that the attractor regime has been reached within the time range of the simulation.[35]

Our analysis suggests that, in the infinite volume limit, the distribution of string length in the attractor solution is of the form depicted in Figure 9. Roughly 80% of the string length per Hubble patch is contained in long strings (infinite string loops), while the remaining 20% is distributed in loops ranging from the core to the Hubble size, with equal numbers of loops in each decade of length. The total string length grows logarithmically according to eq. (77) with the relative 4:1 ratio fixed.

In the literature the parameter ξ is sometimes defined restricting to long strings only. However, the loop distribution that we observe implies that the two definitions only differ by 20%, and more importantly the factor of proportionality is constant in time.

Since we do not directly use the quantitative behavior of ξ_ℓ in our subsequent estimates of the final axion abundance, we have only performed this analysis in the fat string case. The similarity in the behavior of all of the other properties studied suggests that the picture for the physical case should be qualitatively similar.

4.4 The Spectrum

We now discuss how the string network radiates energy, and study in detail how the energy in the system is split among the different components (strings, axions, radial modes), and the way that the

---

[34] The fact that only approximately 10% of the string length is contained in sub-Hubble loops was mentioned in [136], which matches our more detailed analysis.

[35] At late times a similar power law behavior is also present in field theoretic simulations of local strings, see ref. [140].

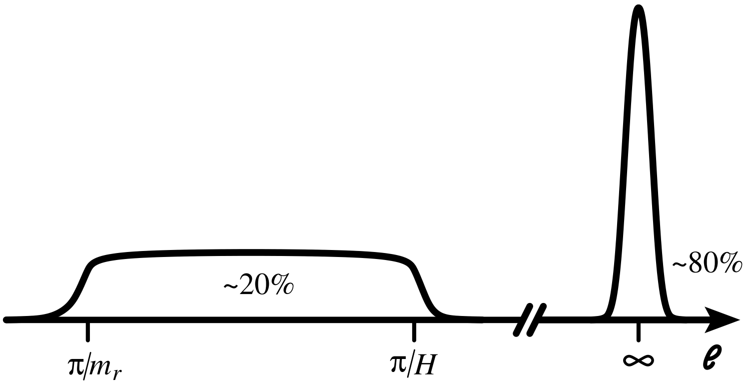

Figure 9: A cartoon of the distribution of string loops, $d\xi_\ell/d\ell$, in the scaling regime. A constant fraction of the total length is in sub-Horizon loops, with equal numbers of loops in each logarithmic interval down to the scale of the string cores.

energy in axions is distributed in modes of different frequencies. We then analyze the evolution of the axion number density, and the extrapolations of this to the physically relevant point. However, before turning to the results of our simulations, we first establish the physically relevant quantities and describe how these affect the axion relic abundance.

The conservation of energy implies that, in order to maintain the scaling regime, the string network must constantly lose energy into radiation. This is because, in the absence of interactions, the number of strings per Hubble patch would increase fast as more strings re-enter the horizon. To keep ξ approximately constant, the excess string length must be destroyed, emitting energy.

The rate at which the system of interacting strings releases energy can be calculated by comparing the energy density in the scaling regime (parameterized by eq. (75)), to that of a "free" network of strings. By free we mean that long strings remain essentially at fixed comoving coordinates, so that $\xi(t) \propto R^2(t) \propto t$. The energy density of such a system is

$$\rho_s^{\rm free}(t) \propto \frac{\mu(t)}{R^2(t)} \propto \gamma_{_f}(t)\frac{\log\left(m_r\, d(t)\right)}{t}\,, \tag{80}$$

where $\mu(t) = \gamma_{_f}(t)\pi f_a^2 \log(m_r\, d(t))$ is the tension of free strings with the corresponding γ-factor, and $d(t) \propto 1/R(t) \propto 1/\sqrt{t}$ parametrizes the average distance between strings.

We consider a network of free strings that has the same string configuration as the interacting system with energy $\rho_s(t)$ given by eq. (75) at a time t_0. The energy of such a system is

$$\rho_s^{\rm free}(t) = \frac{\xi(t_0)\mu(t)}{t_0 t}\,, \tag{81}$$

with $\mu(t_0) = \mu_{\rm eff}(t_0)$, so that this matches that of the interacting network at $t = t_0$. The string energy density of the interacting system in the scaling regime decreases faster than $\rho_s^{\rm free}$, and the difference corresponds to energy that is released. The rate Γ at which the interacting network emits energy into radiation is therefore given by

$$\Gamma = \left[\dot\rho_s^{\rm free}(t_0) - \dot\rho_s(t_0)\right]_{t_0=t} = \rho_s\left[2H - \frac{\dot\xi}{\xi} - \frac{\mu_0}{\mu_{\rm eff}}\left(H + \frac{\dot\eta}{\eta} - \frac{1}{2}\frac{\dot\xi}{\xi}\right) + \frac{\Delta\dot\gamma}{\gamma}\right]$$

$$\Gamma \xrightarrow{\log\gg 1} 2H\rho_s = \frac{\xi(t)\mu_{\rm eff}(t)}{t^3} \,. \tag{82}$$

where $\Delta\gamma = \gamma_f - \gamma$ is expected to have only a mild dependence on time, as we will see below. The last limit holds at late times when the log is sufficiently large, as is the case in the physically relevant regime.[36]

The rate of energy loss from strings can be split into two contributions $\Gamma = \Gamma_a + \Gamma_r$, corresponding to the rate of energy transfer to axions and radial modes respectively. The corresponding continuity equations for the axion and radial energy densities are then

$$\begin{aligned} \dot\rho_a + 4H\rho_a &= \Gamma_a + \dots \,, \\ \dot\rho_r + zH\rho_r &= \Gamma_r + \dots \,, \end{aligned} \tag{83}$$

where z is a factor ranging from 3 (for non-relativistic radial modes) to 4 (for relativistic radial modes, and in the fat string case), and the dots represent subleading contributions from energy transfer via axion-radial mode interactions.

At sufficiently late times in the scaling system's evolution, most of the energy released by the string network is expected to go into axions, since radial modes are heavy relative to Hubble and harder to excite. The energy density in axions is therefore mostly fixed by conservation of energy. In contrast, the number of axions produced depends on the energy spectrum with which they are emitted by the string network.[37] Indeed it is the nature of the axion spectrum that is the source of the largest uncertainty in the relic abundance of axions produced from strings, and this has been the subject of disagreement for many years [25,39,43,134]. Before reporting the arguments underlying the different possibilities, we first review how the axion number density depends on the properties of the spectrum.

Since strings typically have curvature of order Hubble, the natural expectation — always assumed in the literature, but never confirmed in simulations — is that the spectrum of axions emitted at each instant is peaked at momenta of order the Hubble parameter at that time. Meanwhile production of modes with momentum below Hubble or above the string core scale is expected to be strongly suppressed. Between these scales an approximate power law is usually assumed, which determines the hardness of the spectrum. If the spectrum is soft, meaning that it is sharply peaked in the IR scale (around Hubble in this case), a relatively large number of axions will be released to account for the total energy lost by strings. If the spectrum is harder, with a larger UV tail, fewer axions will be produced, although each will be more energetic. The expectation that the attractor solution is approximately scale invariant corresponds to a prediction that the location of the spectrum's peak relative to Hubble, and the power law fall off, are constant up to possible logarithmic corrections.

From eq. (83), if we neglect the energy emitted in radial modes (and axion-radial interactions), which as we will see is a small fraction, we have that

$$\frac{1}{R^4(t)}\frac{\partial}{\partial t}\Big(R^4(t)\rho_a(t)\Big) = \Gamma_a(t) \approx \Gamma(t) \,, \tag{84}$$

---

[36] To derive Γ we used eq. (80) although this only applies for long strings. Free sub-Hubble loops are expected to redshift as non-relativistic matter, however since as we saw in the previous section, they represent only a small fraction of the total energy density of strings the corresponding correction to Γ is small. Similarly the kinetic energy represented by the γ-coefficients would also redshift differently. We have absorbed this effect into the time dependence of the boost factors, which is expected to disappear at late time.

[37] Regardless of their initial energy, cosmic expansion will redshift the momenta of the emitted axions so that most of them are non-relativistic soon after the QCD crossover. Consequently, the spectrum affects the subsequent phenomenology and the final relic abundance only through its impact on the axion number density at that time.

and therefore the energy density in axions at a time t, when they are still massless is

$$\rho_a(t) = \int^t dt' \left(\frac{R(t')}{R(t)}\right)^4 \Gamma(t') \quad \xrightarrow{\log\gg 1} \quad \frac{\xi(t)\mu_{\rm eff}(t)}{3t^2} \log\left(\frac{m_r}{H}\right) , \tag{85}$$

that is, integral of the energy emitted at each previous instant, appropriately redshifted (and we omitted the initial time in the integral since it is dominated by late times).

We also introduce the differential energy transfer rate

$$\frac{\partial\Gamma}{\partial k}\left[k,t\right] : \quad \Gamma(t) = \int dk \, \frac{\partial\Gamma}{\partial k}\left[k,t\right] \ , \tag{86}$$

which depends only on the axion momenta k, the time, and the core size m_r. It is convenient to further split this up as

$$\frac{\partial\Gamma}{\partial k}\left[k,t\right] = \frac{\Gamma(t)}{H(t)} F\left[\frac{k}{H},\frac{m_r}{H}\right] , \qquad \int dx \, F\left[x,y\right] = 1 \, , \tag{87}$$

where the function $F(x,y)$ fully characterises the shape of the spectrum (through the variable x), and its time dependence (through the variable y). Combining eqs. (85), (86) and (87) we get a formula for the axion spectral energy density

$$\frac{\partial\rho_a}{\partial k}\left[k,t\right] = \int^t dt' \frac{\Gamma'}{H'} \left(\frac{R'}{R}\right)^3 F\left[\frac{k'}{H'},\frac{m_r'}{H'}\right] , \tag{88}$$

where primed quantities are computed at the time t', the redshifted momentum is defined as $k' = kR/R'$, and we have left a possible time dependence in the core mass scale m_r to include the fat string case. Eq. (88) is just the time integral of the instantaneous spectra appropriately redshifted, and the change in power of the redshift factors compared to eq. (85) is due to the extra power of k in the differential spectrum.

The total number density of axions is therefore given by

$$n_a(t) = \int \frac{dk}{k}\frac{\partial\rho_a}{\partial k} = \int^t dt' \frac{\Gamma'}{H'} \left(\frac{R'}{R}\right)^3 \int \frac{dx}{x} F\left[x,\frac{m_r'}{H'}\right] . \tag{89}$$

To see how the number density depends on the shape of the spectrum we consider an analytic form that reproduces the theoretical expectation:

$$F\left[x,y\right] = \begin{cases} \frac{1}{x_0}\left(\frac{x_0}{x}\right)^q \frac{q-1}{1-\left(\frac{x_0}{y}\right)^{q-1}} & x_0 < x < y \\ 0 & x < x_0 \vee x > y \ , \end{cases} \tag{90}$$

i.e. a single power law $1/k^q$ with an IR cutoff at $k = x_0 H$ and a UV one at $k = yH \approx m_r$ (the extra factors are required to have the right normalisation). Substituting this into eq. (89) and taking the large time limit, which corresponds to keeping the leading log contributions, we get

$$n_a(t) \approx \frac{8H\xi(t)\mu_{\rm eff}(t)}{x_0} \frac{1-q^{-1}}{1-(2q-1)e^{(1-q)\log(m_r/Hx_0)}} \, . \tag{91}$$

Given the large size of the log, the last factor strongly depends on whether the power q is larger, equal, or less than 1 (see Figure 10). In fact we can rewrite the expression above as

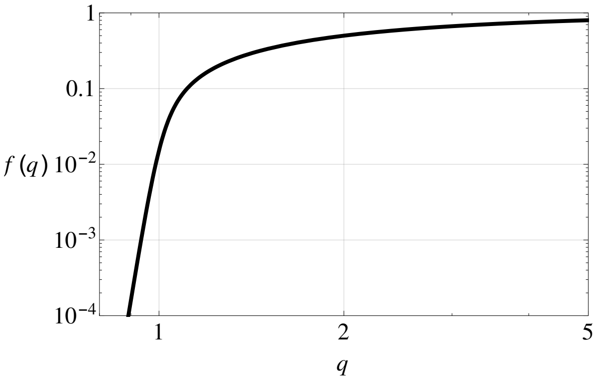

Figure 10: The dependence of the axion number density, relative to its value in the limit $q \to \infty$, $f(q) \equiv n_a/(8H\xi\mu_{\rm eff}/x_0)$ on the power q of the spectrum of axions emitted by the string network, at the physically relevant scale separation $\log(m_r/H) \approx 70$.

$$n_a(t) \approx \frac{8H\xi(t)\mu_{\rm eff}(t)}{x_0} \times \begin{cases} 1 - 1/q & q > 1 \\ \frac{1}{\log\left(\frac{m_r}{Hx_0}\right)} & q = 1 \\ \frac{1-q}{q(2q-1)}\left[\frac{Hx_0}{m_r}\right]^{1-q} & \frac{1}{2} < q < 1\,, \end{cases} \tag{92}$$

in which the last factor varies from $\mathcal{O}(1)$ for $q > 1$ to $\mathcal{O}(10^{-2})$ for $q = 1$, and is exponentially small (in terms of the log) for $q < 1$.

The dependence of the axion number density on q can be easily understood in terms of our previous qualitative discussion. For $q > 1$ the spectrum is soft, most of the energy is emitted with momenta of order Hubble, and the final number density is of order the total energy density contained in strings, $H^2\xi\mu_{\rm eff}$, divided by the average axion momentum $\mathcal{O}(H)$. For $q = 1$ energy is equally distributed in logarithmic intervals of momentum. Therefore, although most of the axions are still emitted with momenta of order H, the total number of axions emitted is smaller by a factor of the log. For $q < 1$ the spectrum is UV dominated, and the majority of the energy is distributed to axions with large momentum so that the axion number density is power suppressed by the UV scale.

The different behavior of the number density for different choices of q can be linked to the change in the average momentum of the axions in the spectrum. If we define the inverse average momentum as

$$\langle k^{-1}\rangle = \frac{1}{\rho_a}\int \frac{d\rho_a}{dk}\frac{dk}{k}\,, \tag{93}$$

the number and energy densities are related via $n_a = \langle k^{-1}\rangle\rho_a$. Depending on whether q is larger or smaller than unity the average momentum is parametrically of order H or m_r respectively.

The huge ratio of scales m_r/H in the physically relevant part of parameter space results in an enormous range of possible values of n_a. It is therefore clear that understanding the spectrum is of paramount importance if results obtained at the values of the log accessible in the simulations are to be extrapolated to the physical values. In particular, even a small change in the behavior of the spectrum could change the extrapolated value of the relic abundance by many orders of magnitude.

We can now identify the main source of disagreement in the literature. Refs. [25, 26, 39, 40] claim that at late times, when the scale separation is large, the coupling of strings to axions is small and the rate that axions are emitted is suppressed, and as a result the dynamics of axion strings are close to those of local strings. If this is the case, the expectation based on the Nambu–Goto effective theory is that loops will oscillate many time before emitting their energy, producing a spectrum that is sharply peaked at small frequencies, of order their initial size, i.e. Hubble. Consequently, they predict that $q > 1$, and that the number density of axions produced by strings in the scaling regime will dominate over the contribution from misalignment, here taken as a reference value $n_a^{\rm mis} = \theta_0^2 H f_a^2$, with $\theta_0^2 \approx 5$.[38] Setting $\mu_{\rm eff} = \pi f_a^2 \log$ following the theoretical expectation in eq. (76), $\log \approx 70$ and $x_0 \approx 10$ we get

$$\frac{n_a^{q>1}}{n_a^{\rm mis}} \approx \frac{8\pi}{\theta_0^2 x_0} \xi \log \sim 30\xi \,, \tag{94}$$

which ranges from 30 to 300 depending if ξ is taken to be close to 1 or 10.

Conversely, refs. [134, 141] claim that string loops do not oscillate, but instead efficiently shrink emitting all of their energy at once and producing a spectrum with $q = 1$. In this case the number density from strings is suppressed

$$\frac{n_a^{q=1}}{n_a^{\rm mis}} \approx \frac{8\pi}{\theta_0^2 x_0} \xi \sim 0.5\,\xi \,, \tag{95}$$

and can even be subleading with respect to that from misalignment if ξ is taken to be 1.

In the rest of this section we present a detailed analysis of the spectrum obtained from simulations with the aim of understanding which of these possibilities is more likely. We therefore postpone further discussion to the end of the section, when we compare our findings with the existing literature.

To analyze the spectrum emitted by the scaling solution we fixed the initial conditions in simulations to be as close to the attractor as possible, corresponding to the black curves in Figure 7. This isolates the radiation emitted in the scaling regime as much as possible, and reduces contamination from pre-existing radiation. In Appendix F we show that, starting with different initial conditions, the spectrum and number density converge to those of the scaling solution, so that the results we obtain do not depend on this convenient choice.

4.4.1 Energy Budget

In analysing the distribution of energy in the scaling solution, and the rate at which axions are produced by strings, it is useful to organise the total energy density stored in the complex scalar field into three components, namely

$$\rho_{\rm tot} = \rho_s + \rho_a + \rho_r \,. \tag{96}$$

Here $\rho_{\rm tot} = \langle T_{00} \rangle$ is the total energy density of the scalar field as given by the average Hamiltonian density eq. (68); ρ_s is the contribution contained in strings; ρ_a is the energy density in axion particles; and ρ_r is that in radial modes. At early times axions, strings, and radial modes are strongly coupled to each other so this separation is ill-defined, but at later times the individual components decouple and the separation becomes meaningful. In the scaling regime, the theoretical expectation is that when the energy density in strings is parameterized as in eq. (76), $\eta(t)$ will be of order 1 and vary only slowly with time.

The way that we actually compute the various components in eq. (96) is as follows: The axion energy density is calculated from the spatial average of $\dot{a}^2$ away from the core of the strings (close

---

[38] In this expression for $n_a^{\rm mis}$, H is the Hubble parameter when the axion mass becomes cosmologically relevant, which is around the time of the QCD crossover although the exact value depends on f_a.

to the string cores, the motion of strings gives a significant contribution to $\dot{a}^2$). By using $2\langle\frac{1}{2}\dot{a}^2\rangle$ rather than $\langle\frac{1}{2}\dot{a}^2+\frac{1}{2}\left|\nabla a\right|^2\rangle$ we avoid the part of the energy that corresponds to the strings' tension, which is mostly contained in $\left|\nabla a\right|^2$. We have checked that our results for ρ_a redshift as expected (i.e. as relativistic matter) and that they are robust against different types of string-core masking (more details about our procedure for screening strings and the consistency tests can be found in Appendix D). The radial energy density is computed by averaging the part of the Hamiltonian density that involves only the radial mode, i.e. $\langle\frac{1}{2}\dot{r}^2+\frac{1}{2}\left|\nabla r\right|^2+V(r)\rangle$, again away from the string cores. Finally ρ_s is simply extracted from the difference $\rho_s=\rho_{\rm tot}-\rho_a-\rho_r$, which avoids double-counting energy contributions in the other components. The string energy density defined in this way includes the energy density stored in the axion-radial interactions, corresponding to the terms $(r/f_a+r^2/2f_a^2)(\partial a)^2$ in the Hamiltonian (second line of eq. (69)), only part of which (that in regions close to string cores) genuinely contributes to the string energy. The remainder corresponds to interaction energy between axion and radial modes. Such interactions could in principle trigger parametric resonance and have a substantial effect, however we have checked that they only give negligible oscillating corrections to the energy densities in our simulations.[39]

We can compare ρ_s extracted in this way to the prediction obtained using the theoretical expectation for the string tension, based on the typical separation between strings, and the measured values of $\xi(t)$. In particular, we compute the effective tension $\mu_{\rm eff}=\rho_s(t)t^2/\xi(t)$ from the definition eq. (75), using $\xi(t)$ and $\rho_s(t)$ from the simulation. This can be compared to the theoretically expected form

$$\mu_{\rm th}=\mu_0\log\left(\frac{m_r\,\eta_c}{H\sqrt{\xi}}\right)\,, \tag{97}$$

which is obtained by replacing, in eq. (76), $\gamma(t)$ with 1 and $\eta(t)$ with $\eta_c=1/\sqrt{4\pi}$ as a reference (we choose $1/\sqrt{4\pi}$ somewhat arbitrarily based on the average distance between strings if they were all parallel, but any roughly similar value would also be theoretically reasonable).

In Figure 11, we plot the ratio $\mu_{\rm eff}/\mu_{\rm th}$ as a function of time, for the fat string and the physical cases. The closeness of $\mu_{\rm th}$ and $\mu_{\rm eff}$ over the entire time range that the system is in the scaling regime is highly non-trivial. These quantities could have differed by orders of magnitude, or had different time dependences, but instead the naive theoretical prediction reproduces the results from the simulation to within 20% throughout. As well as showing that our method of computing ρ_s is meaningful, it is a strong sign that eq. (75) with $\mu_{\rm eff}$ replaced by the theoretical expectation eq. (97) correctly captures the dynamics of the string system. This includes the logarithmic growth of both $\xi\,(t)$ and the string tension due to the increasing scale separation, as well as the variation of ρ_s compared to ρ_a and ρ_r with time. The closeness of $\mu_{\rm th}$ and $\mu_{\rm eff}$ also shows that relativistic effects are not large. This is expected on the basis that the energy density is dominated by long strings, whose motion is damped by the Hubble expansion; indeed $\gamma_c>1$ would increase the ratio $\mu_{\rm eff}/\mu_{\rm th}$ which is instead smaller than (and close to) unity.[40]

The small difference between $\mu_{\rm eff}$ and $\mu_{\rm th}$ is not worrisome. In fact we do not have a reliable way to compute η analytically since its value is determined by the loop distribution and the shape of the strings, which can also depend on time. Further, the parametrization $\mu=\mu_0\log\left(m_r\eta/H\sqrt{\xi}\right)$ with an IR cutoff $\sim m_r/H\sqrt{\xi}$ and constant η applies only to long strings (for loops with radius smaller than Hubble, which make up less than 10% of ξ, the tension is expected to be cutoff at smaller values). A

---

[39] The axion-radial interaction terms are small because $|r|/f_a\ll 1$ away from string cores, and the amplitude of radial waves is rapidly decreased by redshifting.

[40] The analysis in the fat strings case of refs. [136, 138] indeed finds values for the γ factors close to 1.

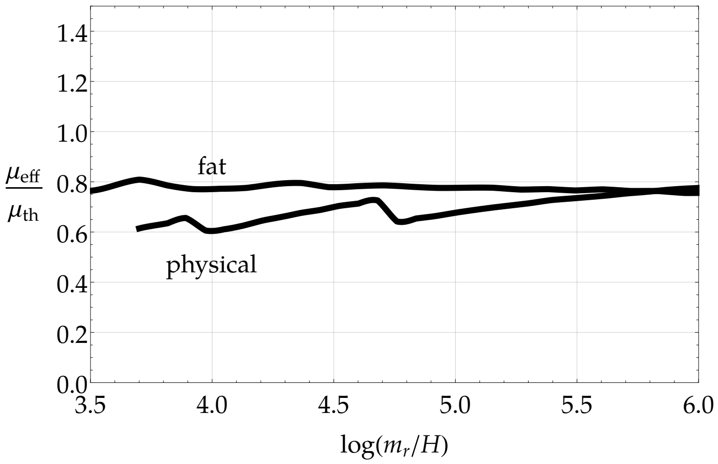

Figure 11: The closeness of $\mu_{\rm eff} = t^2\rho_s(t)/\left(\xi(t)\right)$ to the theoretical prediction $\mu_{\rm th}$, defined in eq. (97), (plotted in terms of the ratio of these two quantities) is a non-trivial check that our procedure to extract the string energy density ρ_s is reliable. More importantly, it shows that the relation eq. (75) can be used to predict the energy in the string network for a given string density and time, by replacing $\mu_{\rm eff}$ by $\mu_{\rm th}$, and that the theoretically predicted logarithmic growth in the string tension is seen in simulations.

simple modification of η can fix the ratio $\mu_{\rm eff}/\mu_{\rm th} = 1$ at all times, however for the moment we are not interested in such a detailed understanding of $\mu_{\rm eff}$, and we content ourselves with the degree of agreement obtained in Figure 11.

Turning to consider the energy in axions and radial modes, in Figure 12 we show the proportion of the total energy that is in the three components of eq. (96) as a function of time, in the fat string and the physical cases. We plot this only for times corresponding to the range $\log(m_r/H) = 3.5 \div 6$ for the fat string case, and $\log(m_r/H) = 4 \div 6$ for the physical one. Data at later time may not be safe from finite volume effects (see Appendix D) while data at earlier times is not representative of the scaling regime since there is not yet a sufficient hierarchy between Hubble and the core scale to consider these decoupled (as will be clear when we analyze the axion spectrum, shortly).

First, we note that in the fat string case all three components redshift as $1/t^2$, up to logarithmic corrections. Therefore the time dependence in Figure 12 is only a result of the increase in energy in strings (due to the factor $\xi(t)\mu_{\rm eff}(t)$ growing) and energy transfer from strings to axions and radial modes. Indeed, using eqs. (82) and (85) these two effects combined predict that $\rho_a/\rho_s \propto \log(m_r/H)$, which explains the relative change in ρ_a and ρ_s. The proportion of energy in radial modes appears to be approximately constant in time, accounting for around 13% of the energy budget.

The situation is less clear for the physical case. The radial mode now redshifts as non-relativistic matter $1/t^{3/2}$, slower than the other components. If we use again eqs. (83) and (82), and assume constant energy transfer rates Γ_i, then $\rho_r/\rho_a \propto t^{1/2}/(\xi(t)\mu_{\rm eff}(t)\log(m_r/H))$. Therefore, if the rates remain constant until late times, energy in radial modes will eventually dominate that in axions. However, over the time range plotted in Figure 12 this expression predicts that ρ_r/ρ_a remains approximately constant, which matches our results from simulations. The fact that the string energy density is not decreasing with respect to that in axions, as happens in the fat string case, may be an indication that larger scale separations are required for the asymptotic behavior $\Gamma \propto \xi(t)\mu_{\rm eff}(t)/t^3$ in

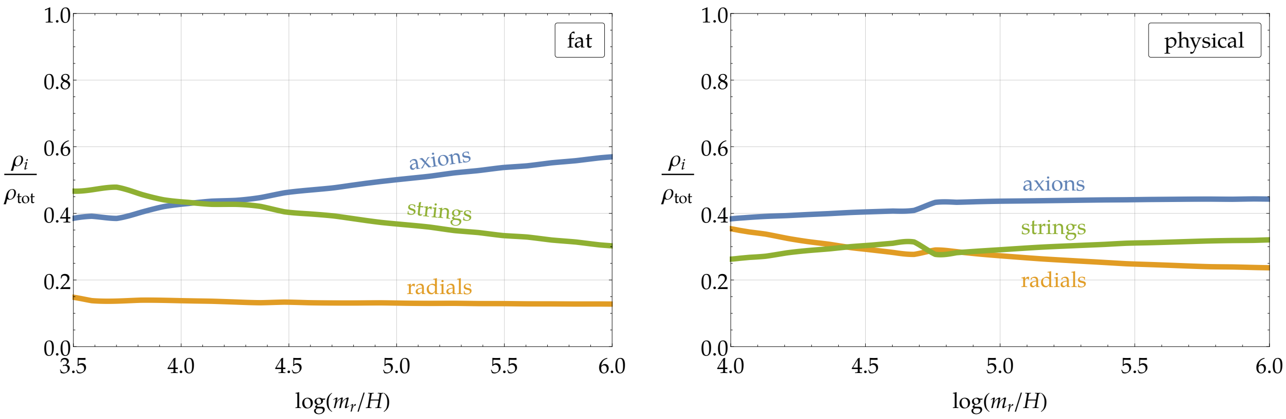

Figure 12: The fraction of the total energy density in free axions, radial modes and strings, as a function of time using the fat string trick (left) and in the physical case (right). The string contribution $\rho_s(t)$ is extracted from the difference $\rho_s = \rho_{\rm tot} - \rho_a - \rho_r$ as explained in the text.

eq. (82) to be reached in the physical case, or due to the shorter time that simulations can be run compared to the fat string scenario.

Perhaps more revealing are the rates at which energy is transferred into axions and radial modes. In Figure 13 we show the time dependence of the ratio $r_a = \Gamma_a/\Gamma$ in the simulations, which we compute from eq. (83) by taking derivatives of the energy densities calculated as above (setting $z = 3$ in the physical case, $z = 4$ in the fat string one, and neglecting the dots, i.e. including possible energy transfer between axions and radial modes in Γ_a). It seems that in both cases the fraction of energy that is transferred to axions is roughly constant, and the value in the fat string scenario is compatible with the approximately constant proportion of the total energy that is in radial modes, plotted in Figure 12. As discussed, in the physical case a constant energy transfer will eventually lead to radial modes dominating the total energy density, but is compatible with the results in Figure 12 for the time range that we can simulate.

Similarly to the fraction of the total energy in strings $\rho_s/\rho_{\rm tot}$, the constant value of r_a in the physical case might only be a transient effect. In fact we expect that the rate of energy transfer from radial modes to axions will increase if the abundance of radial modes grows, so that at sufficiently late times the dominant net effect of energy loss from the string network should be the production of axions. Because of this, any interpretation of the fractional rate r_a computed from simulations in the physical case should be taken with a grain of salt.

To summarize, in both the fat string and the physical cases we find that most of the energy radiated by strings in the scaling regime goes into axions, and these give the largest contribution to the energy density at late times. Meanwhile, the string energy density is well reproduced by the ansatz from the scaling solution, with the theoretically expected string tension and the values of $\xi(t)$ measured in simulations. There is a non-negligible component of the energy density in radial modes, and an approximately constant proportion of the energy emitted by strings goes into such states. This surprising result clashes with the expectation that heavy modes decouple from the evolution of macroscopic soft objects, and that their production is suppressed. We will see that this phenomenon has a close analogue in the rate at which high momentum axions are emitted.

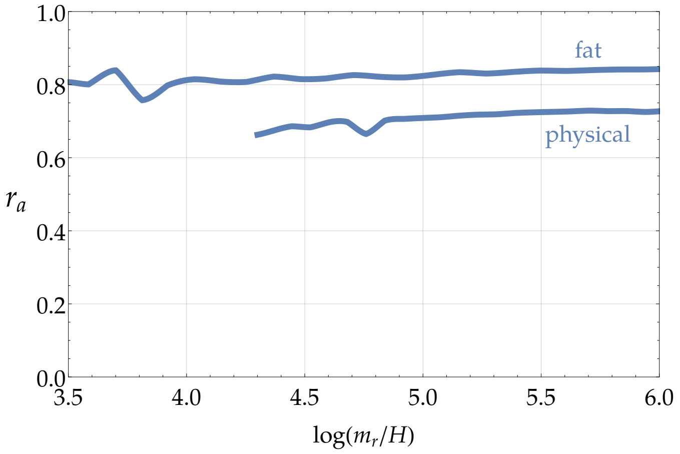

Figure 13: The fraction of the instantaneous energy density emitted by strings that is converted into axions (as opposed to radial modes) $r_a = \Gamma_a/\Gamma$ as a function of time, for the fat string and the physical simulations. At late times more than 70% (80%) of the energy released from the string network in the scaling regime goes into axions in the physical (fat string) case. The rest of the energy goes into the production of massive radial mode states.

4.4.2 Axion Spectrum

We finally arrive at our analysis of the spectrum of axions emitted, the shape of which has a dramatic effect on the axion relic density, as previously discussed.

In Figure 14 we show the differential energy density $\partial\rho_a/\partial k$ at different time shots, for both the fat string and the physical cases. The spectrum is such that $\int dk\ \partial\rho_a/\partial k = \rho_a$, and it has been computed from the Fourier transform of $\dot{a}$, with the strings appropriately masked out, as explained in Appendix E. We plot the spectrum as a function of k/H to highlight that it remains peaked around momenta of order Hubble as this decreases, and we divide it by the factor Hf_a^2 to remove the main time dependence (see eq. (85)).

The spectra have a number of relevant features. We start with the fat string case, for which the core scale m_r evolves in exactly the same way as physical momenta redshift. As a result, radiation produced by the cores with $k \sim m_r$ remains at $k \sim m_r$ at all later times, and contributes to the same part of the differential spectrum regardless of when it was originally produced. Similarly, since the simulations begin with $H \sim m_r$, the typical momentum of pre-existing radiation from the initial conditions is $k \sim m_r$ and states from these early times will subsequently remain at $\sim m_r$. Therefore the part of the spectrum sufficiently below m_r is entirely from genuine radiation produced by strings, and is independent of the physics of their cores.

Figure 14 clearly shows that the spectrum is peaked in the IR at around the Hubble scale, in particular at $k/H \sim 5 \div 10$ for all values of the log larger than 3. There is another small peak in the UV, which is exactly at $k = m_r/2$, the typical frequency of parametric resonance. These modes could be produced by the string core itself, or by the conversion of non-relativistic radial modes into axions. The position of this peak at a particular time identifies the part of the spectrum that is sensitive to core scale (i.e. UV) physics. For times corresponding to $\log(m_r/H) \lesssim 3$ the peaks in the UV and the IR cannot be distinguished, and it is only for $\log(m_r/H) \gtrsim 3.5$ that there starts to be enough of a

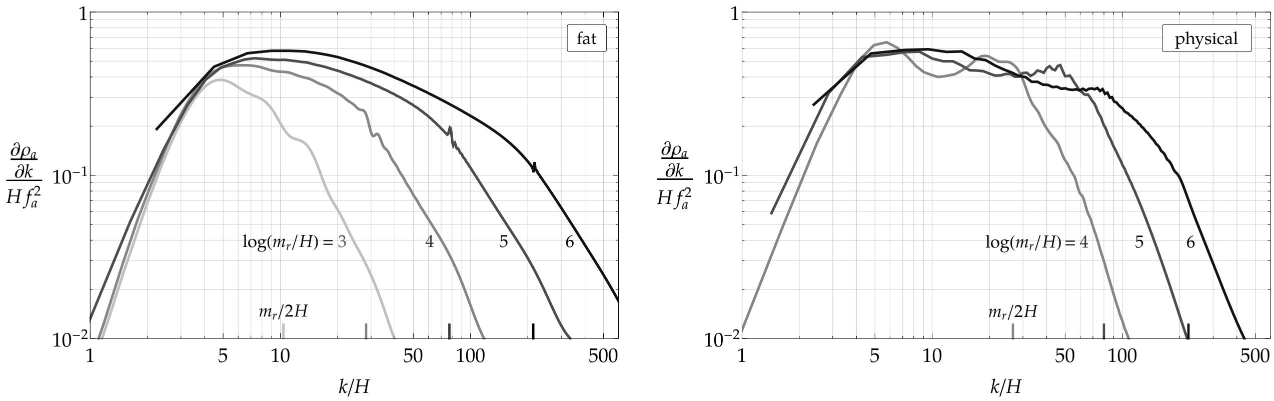

Figure 14: The axion energy density spectrum $\partial\rho_a/\partial k$ as a function of the ratio of the physical momentum k to the Hubble parameter. The results are shown for different times (i.e. different values of $\log(m_r/H)$) for a string system in the scaling regime, in the fat string (left) and the physical (right) scenarios. In both cases the spectrum is dominated by a broad peak at around $k/H = 10$, and emission at lower momenta is suppressed. As a reference, for each time shot we also show the value of k corresponding to the core ($k = m_r/2$), above which there is very little emission, as expected.

hierarchy to justify considering the IR dynamics as decoupled from the core scale physics. For this reason we do not consider smaller values of the log in our analyses involving the spectrum.

At late times the size of the simulated box, $L(t)$, starts cutting off the low momentum part of the spectrum, as can be seen from the interruption of the spectrum at the minimum non-trivial $k = 2\pi/L(t)$ in the final shots. Subsequently this would start altering the IR peak, so we only analyse the spectrum until $\log(m_r/H) = 6$ when this effect is still harmless (a more detailed study of the various systematics can be found in Appendix D).

As expected, the spectrum is power suppressed for momenta smaller than the IR peak or larger than the UV peak — long wavelength modes are inhibited by the horizon, while the high energy ones are suppressed by decoupling. In particular, the spectrum seems to fall as $\sim k^3$ in the far IR, and as $\sim 1/k^2$ in the far UV. The region of physical interest is in between the two peaks. The spectrum reaches a stable form towards the end of the simulation, and in the last Hubble e-folding (i.e. for logs in the interval 5 to 6) its shape remains very similar, modulo the shift of the UV peak. However, even though $\partial\rho_a/\partial k$ is largest in the IR, its area is dominantly in the UV, since the slope between the two peaks is less steep than $1/k$. This can be seen more clearly by plotting $\partial\rho_a/\partial \log k$ on a log scale (shown in Figure 31 in Appendix G). We conclude that although most of the axions produced by the evolution of the string network at late times are soft, the majority of the energy density is contained in UV axions with energy of order the inverse core size.

The situation for the physical case seems similar, although the uncertainties are larger. In fact it is in studying the spectrum that the advantages of the fat string trick are most pronounced. In the physical case, axions produced at early times with momentum of order the core-scale redshift with respect to m_r, which is constant. This means that, although the distance between the IR peak and $m_r/2$ at late times is the same as in the fat string case (for the same value of $\log(m_r/H)$), the spectrum is contaminated by core-scale radiation produced earlier and redshifted down to $k \sim \sqrt{H m_r/2}$. This effectively halves the region of the spectrum that is free from UV dependent contributions. Indeed the spectrum shows that the IR and the UV dynamics are not decoupled before logs of order 4.5 or 5, and for this reason we do not consider quantities that rely on the spectrum at times corresponding to $\log < 4.5$. Similarly, the UV peak is now replaced by a broader feature, caused by the convolution of core-scale radiation emitted at different times and redshifted by different amounts.

Because of these disadvantages, extracting the shape of the spectrum between the peaks is more challenging in the physical case. The rest of the spectrum shows similar features to the fat string case, with a strong suppression of modes below $k \sim (5 \div 10)\, H$ and above $k = m_r/2$ (for the UV modes, this time with a stronger suppression, $\sim 1/k^3$ instead of $\sim 1/k^2$). As for the fat string case, the IR peak seems to dominate the spectrum, especially at late times, while the area is still dominated by the UV region. Therefore, most of the axions produced by the evolution of the string network at late times are again soft, but most of the energy density is in UV axions. However, we stress that the results in the physical case are less clear and should be interpreted with caution.

To the best of our knowledge, the only other serious attempt to extract the axion spectrum in the scaling regime was in ref. [135], based on results of simulations carried out on a somewhat smaller grid than ours. In that paper the authors observe an exponential suppression of the spectrum at large momenta, which they interpreted as indicating a strongly IR peaked distribution. However, the range of momenta that shows such behavior seems to lie above the scale of the core, $m_r/2$. Unfortunately, the region at smaller momenta has been very poorly binned, so that little information on the actual behavior of the spectrum in this region is available.

4.4.3 Instantaneous Emission

In addition to the overall spectrum, the shape of the instantaneous axion spectrum, i.e. the function $F(x, y)$ in eq. (87), is crucial for understanding the properties of the emitted axions, and in particular for inferring the evolution of the axion number density at later times.

We compute F from the spectrum by inverting eq. (88), namely

$$F\left[\frac{k}{H}, \frac{m_r}{H}\right] = \frac{A}{R^3}\frac{\partial}{\partial t}\left(R^3 \frac{\partial \rho_a}{\partial k}\right) \ , \tag{98}$$

where the factor $A = H/\Gamma$ is fixed by requiring that F is normalised to 1. To evaluate the derivative numerically we took the difference of $R^3 \partial\rho_a/\partial k$ between two subsequent time shots (separated by $\Delta\log = 1/4$). The results are shown in Figure 15. Since interactions with radial modes induce small oscillations in time with frequency $\sim m_r$ of the axion energy density (see Appendix E.1) the procedure to extract F is subject to fluctuations at frequencies near the core, as evident in the plots (as explained in Appendix E.1 this effect is more pronounced for physical strings than for fat ones).

For the fat string case we plot F at three time shots, i.e. at three values of the log. At all of these times F has an IR peak at momentum around the Hubble scale (and suppressed emission at lower momenta), and a UV peak at the scale of the string cores. In particular, the IR peak is located around $k/H \approx 5$ and extends to $k/H \approx 10$, while the UV peak is at $k \sim m_r/2$. As a result, the ratio between the maximum and the minimum momenta between which the instantaneous spectrum can show a power law behavior is bounded by $k_{\max}/k_{\min} \lesssim m_r/40H \lesssim N/120$, where N is the number of points in the spatial grid of the simulations (assuming a lattice spacing $a = m_r^{-1}$ and that the simulations are stopped when the box size $L = 3/H$, and also requiring $k/H > 20$ to be safely away from the IR peak). For our simulations this means that $k_{\max}/k_{\min} \sim 10$.

Although the separation of the UV and IR peaks is not yet large for $\log = 4$, at values above 5 an intermediate momentum range can be clearly identified in which the differential spectrum shows a

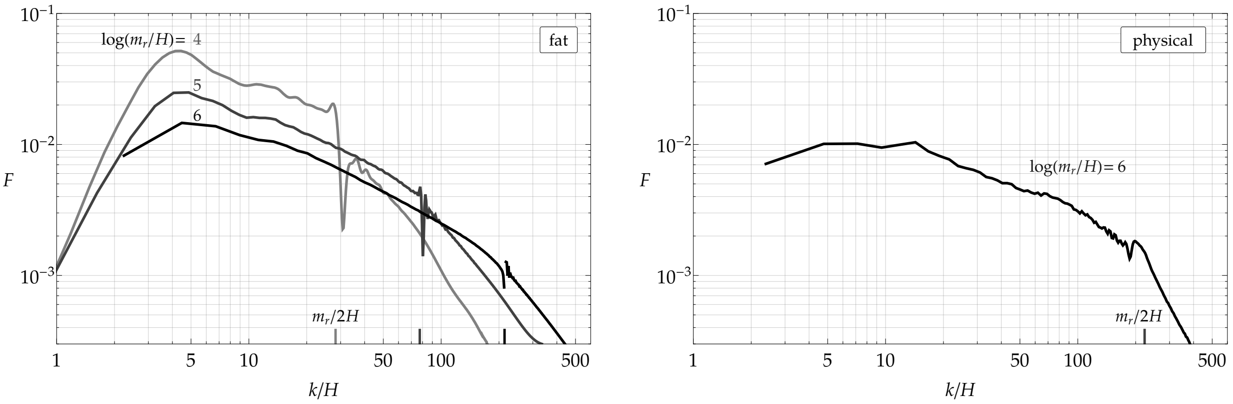

Figure 15: The shape of the spectrum of axions emitted instantaneously by the string network $F(k/H, y)$ for the fat string (left) and the physical (right) cases, for different scale separations, i.e. different values of $y = m_r/H$.

definite power law behavior. In this region the instantaneous spectrum of axion emitted is compatible with a behavior $1/k^q$ with $q \approx 0.7 \div 0.8$. More details about how we extract q, and further plots, can be found in Appendix G. Consistent with our analysis of the convoluted spectrum, this value means that most of the energy released by the string network goes into high energy axions, although most of the axions are soft.

We also note the constant form of F at different times (over a range in which H changes by 2 orders of magnitude) is another demonstration of the scaling behavior of the system. In particular, the IR and UV peaks remain their expected positions, and the intermediate power law with $q \approx 0.75(5)$ is constant with time to within the uncertainty. Although there is no noticeable change in q in the interval of logs between 5 and 6, the uncertainty is relatively large. Consequently, we cannot exclude the possibility that, as for ξ, a logarithmic scaling violation occurs. This could lead to, for example, a behavior $q \sim 0.7 + \epsilon \log(m_r/H)$ for some small constant ϵ. In Appendix F we also show that F is independent of the string systems initial conditions, confirming that we are indeed analyzing the properties of the attractor solution.

As expected, the analysis of the physical case is more difficult. UV modes pollute the convoluted spectrum well below the scale m_r, so useful results only start appearing at late times. Because of this, in Figure 15 we only show F for the last time shot at log $= 6$. Although there are still large fluctuations due to higher frequency modes, an approximate power law can be recognized at this time, with a value of q that is again smaller than 1 and which appears to be similar to in the fat string scenario.

The form of the instantaneous emission found above clashes with either of the theoretical expectations discussed in eqs. (94) and (95), which predicted $q \geq 1$. Similarly to the radial modes, axions with momentum of order the string core scale have not decoupled from the dynamics of strings at the final times $\log(m_r/H) \lesssim 6$, despite the relatively large hierarchy $m_r/H \sim 500$. To investigate this surprising result further, in Appendix H we study the dynamics of the collapse of a single axion string loop. In particular we analyze whether this is converging to the prediction for a Nambu–Goto string in the limit that its initial radius is large compared to its core size. We find that the radius of the axion string loop as a function of time during its initial collapse does indeed get closer to the

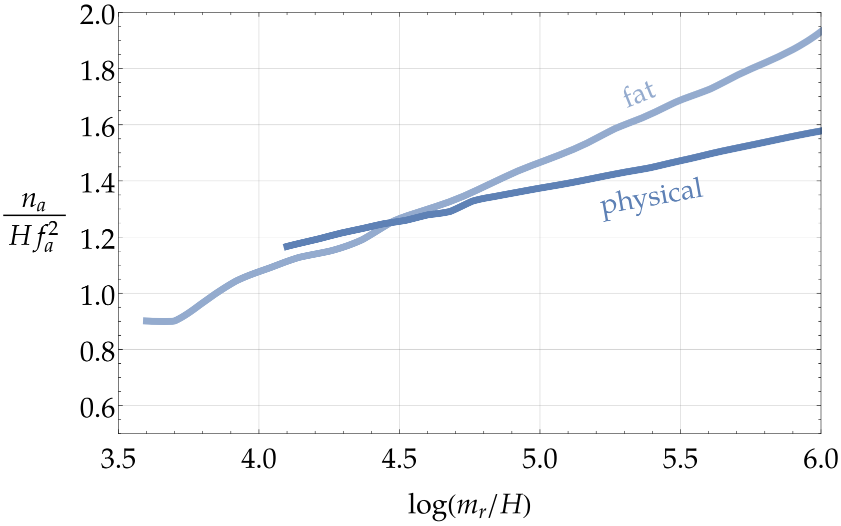

Figure 16: The axion number density, normalised by Hf_a^2, as a function of time for the fat string and the physical cases.

cosine prediction (and is extremely close to the prediction based on the effective theory of global strings coupled to radiation in the limit of large scale separation [38]). At the scale separations that can be simulated ($\log(m_r/R_0) \lesssim 5$ where R_0 is the initial loop radius) the loop does not rebound significantly, as opposed to the case where $\log(m_r R_0) \gg 1$ in which the loop is expected to rebound many times according to arguments based on the use of the Nambu–Goto string action [35]. On the other hand, whether the loop bounces also depends on its dynamics at small radius, a regime in which the Nambu–Goto approximation is not valid. Since it is at such times that energetic modes are more efficiently produced, this result is consistent with our observation that the spectrum also deviates from the expectation based on a loop oscillating many times.

4.4.4 Number density

We are now ready to consider the axion number density. We evaluate this in simulations using the differential spectrum $\partial\rho_a/\partial k$ at different time shots, and eq. (89). The results for n_a, normalised with the factor Hf_a^2, as a function of time are shown in Figure 16 for the fat string and the physical scenarios. In both cases n_a/H increases logarithmically over the time range plotted, although the values and the slopes are different. Therefore, similarly to the fit of $\xi(t)$, our first conclusion is that at the quantitative level the fat string system differs from the physical theory, although they seem to have the same qualitative features.

As we have stressed, the values of the log that can be analyzed with numerical simulations are nowhere near large enough to reach the physically relevant region of parameter space. Consequently, n_a must be extrapolated to extract predictions, and the way that this is carried out has a dramatic impact on the results obtained. Such an extrapolation is a viable possibility because we have shown the existence of an attractor solution, and we have found that the energy of the string network is accurately reproduced by eq. (75) with the theoretical prediction for $\mu_{\rm eff}$, eq. (97). Additionally, given our results for $\xi(t)$ it is plausible that the fit in eqs. (77) and (78) can be extended all the way to the physical scale separation. The remaining component needed in order to predict the axion relic abundance is an understanding of the form of the instantaneous emission spectrum at late times.

As usual we will discuss the fat string case before moving to the physical one. From Figure 16 it is tempting to extrapolate $n_a/\left(Hf_a^2\right)$ linearly in the log, however this procedure is too naive. From eq. (92), a linear behavior is expected only for very large logs and even then only if the power law of the instantaneous emission spectrum is $q=1$ (and $\xi(t)$ also increases linearly with the log). This is not the case in simulations since, as we saw in the previous section, $q\sim 0.75$ and the logs are not large enough to neglect $1/\log$ corrections in eqs. (82) and (89). We conclude that the linear behavior observed is most likely a transient effect. Indeed, should the power q remain constant below 1 the number density would start decreasing exponentially with the log at late times (from eq. (92)). Meanwhile, if q grows to be significantly larger than 1 the number density will increase as the log squared.

The different behaviors at large values of the log are shown in Figure 17, assuming that ξ continues to increase logarithmically. It is clear that a naive linear extrapolation leads to significantly different results compared to if q is assumed to remain constant at 0.75. Since we are not able to reliably study the behavior of q with time, we cannot exclude the possibility that a small scaling violation results in q growing to 1 or even larger values, and we also show the axion number densities at large values of the log in these cases. It can be seen that if such a change does occur, the large log behavior of n_a is completely different to both the linear extrapolation and the $q=0.75$ possibilities, leading to much larger number densities. The extrapolations in Figure 17 have been obtained from eq. (89) using a form for F fitted from the results shown in Figure 15 with q modified between the IR and UV peaks, although the results are not very sensitive to the exact shape of F away from the power law region (more details on the form of F used are given in Appendix G).

For the physical case there is even more uncertainty in the value of q, and consequently on the late time number density. In Figure 17 (right) we plot the extrapolated number density for different values of q, together with the naive linear extrapolation and the available values from the simulation. As a comparison the reference number density from misalignment ($n_a^{\rm mis}=\theta_0^2Hf_a^2$ with $\theta_0=2.2$) is also shown. Similarly to the fat string case, if $q=1$ then $n_a/\left(Hf_a^2\right)$ increases proportionally to the log, if $q\gg 1$ it grows with log squared (again assuming that ξ continues to increase logarithmically), while if $q<1$ it is exponentially suppressed in terms of the log.

It follows that if the spectrum remains UV dominated with small q (i.e. with $q\lesssim 0.85$) for large values of the log, the number of axions produced from strings during the scaling regime is negligible. However, if q increases at larger log, the axions produced by strings can easily match or dominate those from misalignment. Indeed, even a very mild log dependence of the power q, for example of the form $q=0.7+0.01\log(m_r/H)$, too small to be excluded by our simulations, would be enough to make the final abundance at $\log=70$ more than 2 orders of magnitude larger than the misalignment contribution.

As discussed in Section 4.2, the results of simulations only depend on the ratio m_r/H, and can be interpreted either in terms of a theory with $m_r\sim f_a$ at an early time when $H\sim f_a$, or a theory with a much smaller $m_r\ll f_a$ at a later time $H\ll f_a$. Fixing the Hubble scale to its value when the axion mass becomes relevant $H\approx m_a\sim\Lambda_{\rm QCD}^2/M_{\rm Pl}$, the values of n_a calculated in our simulations can therefore be used to extract the physically relevant axion number density for theories with extremely small values of m_r (as would occur in a UV completion with a complex scalar field that had a tiny quartic coupling). However, constraints from observations of the evolution of stars and fifth force experiments require that $m_r\gtrsim$ keV in viable theories, corresponding to rather large values of the log (~ 23 for the reference value of $f_a=10^{11}$ GeV), which is still beyond the reach of simulations. Due to this phenomenological requirement, we only begin the extrapolations in Figure 17 at this

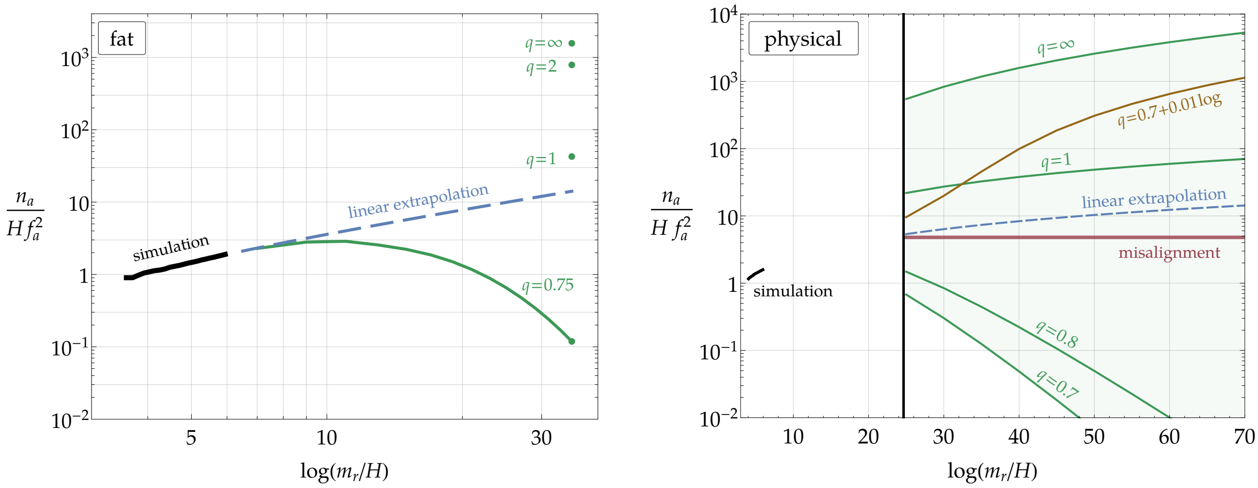

Figure 17: Possible extrapolations of the axion number density to large values of the log with different assumptions about the late time behavior of the power q of the instantaneous emission spectrum. For the fat string case (left) we compare a naive linear extrapolation of n_a to that assuming $q = 0.75$ remains constant. We also show the late time number densities if q changes to 1, 2, or ∞ at an intermediate value of the log. For the physical case (right) we only show extrapolations in the range $\log \gtrsim 23 \div 70$, corresponding roughly to the values of m_r in the range keV$\div 10^{11}$GeV, which are not excluded by fifth force experiments and star cooling bounds. In addition to constant values $q = 0.75$, 1, 2 and ∞, we plot the effect of a small logarithmic scaling violation $q = 0.75 + 0.01 \log{(m_r/H)}$.

scale separation. Meanwhile, the typical axion models with $m_r \sim f_a$ correspond to logs ≈ 70 at the physically relevant time as usual, at the right hand boundary of the plot.

The main message from this section is that an understanding of the late time evolution of the properties of the string network is of paramount importance for a correct extrapolation of the axion abundance to the physically relevant parameter range. If only small-log data are available, very precise computations of the spectrum and its time dependence will be necessary, or if theoretical arguments that $q \to \infty$ in the limit of large scale separation are believed, a prediction for the relic abundance can already be made. In contrast, a naive extrapolation of the axion number density (such as the linear-log one plotted above) could easily give results that are off by several orders of magnitude.

Our final results for the axion number density calculated in simulations are not in disagreement with those obtained by other groups. Despite this, the conclusions that have been reached in the previous literature about the abundance at the physically relevant scale separations are incompatible with each other, primarily due to the different ansatz used in extrapolating. In particular:

- Refs. [25,26,39,40] assume that $q > 1$ based on arguments related to the evolution and emissivity of effective Nambu–Goto strings, which are expected to reproduce the dynamics of global strings at large logs. Assuming $\xi = \mathcal{O}(10)$ they conclude that strings should produce a large number of axions. Our result would only be compatible with this estimate if the power q of the axion spectrum increases at large scale separation, however we are not aware of any reliable evidence that this is indeed the case beyond the Nambu–Goto approximation. In fact, as mentioned, energetic modes are expected to be efficiently produced from collapsing loops when their radius is comparable to the string thickness (or by long strings that are similarly close together), and in this regime the Nambu–Goto effective description is not valid.

- Refs. [134, 141] on the contrary assume that $q = 1$, supported by computations of the axion spectrum produced by collapsing loops at values of the log $\lesssim 5$. Using $\xi = \mathcal{O}(1)$ they conclude that the abundance of axions produced by strings is not larger than that from misalignment. While the value of ξ assumed is compatible with those measured in our simulations, the time evolution of ξ that we observe suggests that larger values need to be used in the physical regime. Additionally, the study of the spectrum from collapsing loops is not sufficient to infer the spectrum emitted by the full string network in the scaling regime. Instead, this is the result of a combination of spectra produced by loops and long strings, and therefore also depends strongly on the distribution and the evolution of the latter. While the value of q observed in our simulation is not far from 1, we have no evidence to support an expectation of an asymptotic value of $q = 1$ at large logs.

- Refs. [135] performed similar simulations of global strings to those that we have carried out. They find a value of $\xi \approx 1$, in the same approximate range as that obtained in our simulations, but they did not observe any linear increase with the log (probably for the reasons discussed in Section 4.3.2). For the reasons discussed in Section 4.4.2, they inferred a very IR dominated spectrum and conclude that strings would efficiently produce a large axion number density, even though their ξ is lower than our extrapolated value. This example shows how a different interpretation of the spectrum can lead to a completely different conclusion, despite the fact that the spectrum measured in the simulation for small logs might be qualitatively in agreement. At small logs, a very detailed study of the spectrum is mandatory to reliably extrapolate the number density to the physical point.

- Finally in refs. [136, 142] a different UV completion of the string cores was introduced in order to simulate fat strings with a large effective tension $\mu_{\text{eff}} \sim 70\mu_0$, despite the actual core size remaining bounded by the lattice spacing size $\log(m_r/H) \sim 6$. While this trick is claimed to remove the need for a large extrapolation connected to the string tension and the effective axion-string coupling, the one associated to the decoupling of the radial modes of the string core remains. The results in ref. [136, 142] confirm the non-trivial dependence of ξ on the logs, although the actual functional dependence is not clear from the analysis. The authors did not present explicit results for the spectrum. The final number density of axions observed, which also includes the contribution from the domain walls and the destruction of the string network owing to the axion mass, turns out to be smaller than that from misalignment. This result seems to be compatible with an underproduction of axions from strings, associated to a spectrum with $q < 1$, in agreement with our results if we assume that q remains below 1 at larger logs.

  However, note that in this case the spectrum (and the way that strings sustain the scaling regime) is dominated by core scale physics. This makes the extrapolation of the string core size from $\log(m_r/H) \sim 6$ to $\log(m_r/H) \sim 70$, as subtle as in the normal case. In fact we cannot exclude the possibility that at larger logs (i.e. relatively thinner strings) the production of energetic modes gets suppressed and q increases. The extrapolation of $\log(m_r/H)$ to the physically relevant values made in ref. [142] is therefore just as sensitive to the uncertainties discussed as it is in conventional simulations. A high precision study of the spectrum is therefore required to identify the correct extrapolation.

5 Conclusions

In this thesis we have studied two important ingredients for a theoretical prediction of the QCD axion mass based on the production of DM in the early universe.

First, in chapter 2 we calculated the relation between axion mass and decay constant at high precision, making use of chiral perturbation theory techniques. The result in eqs. (53) and (54) shows that the chiral expansion is converging very well and that the NNLO corrections (of the same orders of the EM ones) are one order of magnitude below the NLO contributions. While the current uncertainty is of the order of the percent and dominated by the one on the low energy constant ℓ_7, an improvement in ℓ_7 and in the ratio m_u/m_d from lattice QCD studies would improve our estimate up to the per mille level.

After reviewing in chapter 3 the production of axion DM via the misalignment mechanism, in chapter 4 we studied the cosmological scenario where the PQ symmetry is restored after inflation. In this scenario the axion DM abundance can be in principle calculated only as a function of f_a (and so of m_a via eq. (53)), which would give a unique prediction for the axion mass if the QCD axion is to make up the entire DM energy density. Unfortunately, studying the dynamics of the axion field is challenging because of the presence of strings and domain walls.

Using numerical simulations, we have shown that the string network approaches an attractor solution that is independent of its initial conditions, and which is approximately scale invariant. We have also seen that this solution has a number of interesting properties:

- The string length per Hubble patch $\xi(t)$, defined by eq. (74), increases logarithmically with the ratio of the Hubble parameter and the string core size, with best fit parameters given in eqs. (77) and (78).
- At any time more than 80% of the string length is contained in long strings and the rest is in loops of size Hubble and smaller, which follow a scale invariant distribution. While such loops shrink and disappear they are replaced, at the same rate, by loops produced from longer strings.
- The energy density of the string network is determined by eq. (75) with — nontrivially — an effective string tension μ_{eff} that is close to the theoretical expectation eq. (97). This means that to a good approximation, the energy density of the string network in the scaling regime at a particular time can be determined solely from the density of strings $\xi(t)$.
- Over the scale separations that can be simulated, a substantial proportion of the energy emitted by the string network goes into heavy radial modes. Contrary to expectations, the radial modes have therefore not decoupled from the string dynamics by the end of the simulations, despite the relatively large scale separations $m_r/H \sim 500$ reached at these times.
- The instantaneous spectrum of energy emitted into axions, plotted in Figure 15 in terms of F defined in eq. (87), has the theoretically expected shape with a peak at momentum around the Hubble scale. At higher momenta it follows a power law until the string core scale above which emission is suppressed. The during the whole simulation the slope q of the power is such that the majority of energy emitted by strings goes into high momentum axions for the scale separations that we can study (although a greater number of low momentum axions are produced). However, uncertainties are still large enough that a small time dependence of q could change completely the behavior of the spectrum in the physically relevant parameter region.

- In simulations with physical strings, the instantaneous emission spectrum of the scaling solution can only be evaluated close to the end of the simulation because the residual energy from the initial conditions has less time to redshift. The result obtained matches the conclusions drawn from our analysis of the overall spectrum. The fat string model ameliorates these problems, and for this system we see that F remains similar at different times. Although the slope of the power law shows no sign of changing as the scale separation increases, the uncertainty is substantial and scaling violations also here are not excluded.

Since the scale separations that can be simulated are far from the physically relevant values, extrapolation over a vast distance is necessary if predictions relevant to the relic abundance of the QCD axion are to be made. Given our results, it is plausible that ξ continues to grow logarithmically (although we cannot exclude the possibility that it eventually saturates). Such an increase would enhance the number of axions emitted by a factor ~ 10. Having tested that the energy density in strings is well matched by eq. (76) with the theoretical expectation for the string tension, the remaining ingredient necessary for a precise prediction of the relic abundance is the instantaneous axion emission spectrum at large scale separations, in particular the power q in F. While $q < 1$ at the scale separations that we can study, there is space within the current uncertainty for a small logarithmic correction that would result in it increasing to ≥ 1 in the physically relevant part of parameter space. On the other hand, if $q < 1$ persists at large scale separations then the number of axions produced by strings is presumably sensitive to the UV completion of the theory, as a significant proportion of the energy is going into exciting heavy modes, although in this case the number density is suppressed relative to that from misalignment.

To determine whether q has a scale dependence, it is of paramount importance to carry out simulations at larger scale separations. Indeed, even with a relatively modest increase in range, evidence for a change in q might be seen, which would allow an extrapolation.[41] Improvement could come from different directions. One possibility is to simply perform simulations on larger high performance computing clusters, after implementing more efficient parallelization than we have in our present work. A more involved possibility would be to develop an adaptive mesh algorithm, in which the scalar field is simulated on a grid with finer mesh spacing in regions of space where this is required, close to the cores of strings. The mesh must be updated as the strings move, which requires a more sophisticated algorithm, however this approach has proved beneficial in simulations of, for example, astrophysical black hole mergers [144]. It may also be interesting to further study the dynamics of individual strings, or small numbers of strings, which could give an indication of the behavior of the full network. A qualitatively different approach would be to develop a numerical simulation in which the strings themselves were treated as the fundamental degrees of freedom, with parameterized dynamics and interactions. Such an effective theory style simulation could allow much larger scale separations to be studied. In [145] this idea was used to numerically simulate “strings” in 2 dimensions, although the extension to three dimensions is challenging, and a careful analysis would be necessary to ensure that the full dynamics of the underlying theory is capture.

A full calculation of the relic abundance must also include the axions produced when axion mass turns on, at which time domain walls form and the string network is destroyed. As mentioned at the beginning of chapter 4, after the axion mass becomes cosmologically relevant, when $m_a \sim H$, it

---

[41] Encouraging preliminary results on this are already available but will be omitted, and left for a future publication [143]. In particular, logarithmic violations in the spectrum index are actually present and about one order of magnitude bigger than those shown in fig. 17, and point towards an IR dominated spectrum. A log dependence can be also found in other observables that might be related to the spectrum shape, such as the percentage of energy emitted in axions *vs* radial modes, the number of bounces and Lorentz factors of circular and noncircular string loops.

continues to increase fast, and quickly $H \ll m_a \ll f_a$. As a result, this system has three widely separated scales, and is even harder to numerically simulate than the string network alone.

It is however reasonable (but not guaranteed) that the effect of the axions emitted during the scaling regime can be treated separately if they do not interact significantly with the domain walls. If this assumption is verified, evolving a system of axion waves through the QCD crossover with the spectrum extrapolated at $\log = 70$ would be feasible and possibly lead to a solid lower bound for the relic abundance, and so for the axion mass. In any case, even if the complicated nonlinear dynamics does not allow for the separate treatment of the axions produced from strings, the results of this thesis will be still necessary to set the initial conditions of the string system at the domain-wall epoch.

A Results in terms of the quark masses

We provide here the results for pion mass, pion decay constant and topological susceptibility in both $N_f = 2$ and $N_f = 3$ chiral perturbation theory (in this last case in the unbroken isospin limit $m_u = m_d$) in terms of the bare chiral Lagrangian parameters and quark masses. These are intermediate results used to obtain the formulas in the main sections and the matching of app. B.

A.1 Two-flavor results

We start with the neutral pion mass m_{π^0}, calculated at $\mathcal{O}(p^6, p^2e^2)$, and the charged pion decay constant f_{π^+}, defined at $\alpha_{\text{em}} = 0$ and calculated at $\mathcal{O}(p^6)$. These have been calculated in the unbroken isospin limit in [106, 124–127], and read:

$$m_{\pi^0}^2 = M^2\left[1 + \delta_1^M + \delta_2^M + \delta_e^M\right] , \tag{99}$$

$$f_{\pi^+} = F\left[1 + \delta_1^F + \delta_2^F\right] , \tag{100}$$

where $M^2 \equiv B(m_u + m_d)$, B and F are the LECs of the leading order $N_f = 2$ Lagrangian [15], and

$$\delta_1^M = \frac{M^2}{F^2}\left[(2\ell_3^r + (1/2)\kappa\lambda_M) - (2\ell_7)\Delta^2\right] , \tag{101}$$

$$\delta_1^F = \frac{M^2}{F^2}\left[\ell_4^r - \kappa\,\lambda_M\right] , \tag{102}$$

$$\delta_e^M = e^2\left[-\frac{20}{9}\left(k_1^r + k_2^r - \frac{9}{10}(2k_3^r - k_4^r) - k_5^r - k_6^r - \frac{1}{5}(1-3\Delta)k_7^r\right) + 2\kappa Z\,(1+\lambda_M)\right] , \tag{103}$$

where $\kappa \equiv (4\pi)^{-2}$ and $\lambda_M \equiv \log\frac{M^2}{\mu^2}$. The $\mathcal{O}(p^6)$ contributions are

$$\begin{aligned}\delta_2^M = \frac{M^4}{F^4}\Bigg\{ &- 16\,(2c_6^r + c_7^r + 2c_8^r + c_9^r - 3c_{10}^r - 2\,(3c_{11}^r + c_{17}^r + 2c_{18}^r)) \\ &+ (\ell_1^r + 2\ell_2^r + \ell_3^r)\,\kappa + \frac{163}{96}\kappa^2 + \left[-(14\ell_1^r + 8\ell_2^r + 3\ell_3^r)\,\kappa - \frac{49}{12}\kappa^2\right]\lambda_M + \frac{17}{8}\kappa^2\lambda_M^2 \\ &+ \left[16\,(-c_7^r + c_9^r + 3c_{10}^r + 2c_{11}^r) + \kappa(1+5\lambda_M)\ell_7\right]\Delta^2\Bigg\} ,\end{aligned} \tag{104}$$

$$\begin{aligned}\delta_2^F = \frac{M^4}{F^4}\Bigg\{ &8\,(c_7^r + 2c_8^r + c_9^r) + (-(1/2)\,\ell_1^r - \ell_2^r - 2\ell_3^r)\,\kappa - \frac{13}{192}\kappa^2 \\ &+ \left[(7\ell_1^r + 4\ell_2^r - 2\ell_3^r - (1/2)\,\ell_4^r)\,\kappa + \frac{23}{12}\kappa^2\right]\lambda_M - \frac{5}{4}\kappa^2\lambda_M^2 \\ &+ \left[8\,(c_7^r - c_9^r) + \kappa\,(1+\lambda_M)\ell_7\right]\Delta^2\Bigg\} .\end{aligned} \tag{105}$$

At the same order, the topological susceptibility reads

$$\chi_{\text{top}} = \frac{z}{(1+z)^2}M^2F^2\left[1 + \delta_1^\chi + \delta_2^\chi + \delta_e^\chi\right] , \tag{106}$$

where the NLO correction have been first computed in ref. [128]

$$\delta_1^\chi = \frac{M^2}{F^2}\left[2\left(h_1^r - h_3^r - \ell_7 + \ell_3^r\right) - (3/2)\kappa\lambda_M + \left(2\ell_7\right)\Delta^2\right] , \tag{107}$$

while the EM and the NNLO corrections read

$$\delta_e^\chi = e^2\left[\frac{20}{9}\left(k_5^r + k_6^r\right) + \frac{4}{9}(1+3\Delta)\,k_7^r - 2\kappa Z\left(1+\lambda_M\right)\right] , \tag{108}$$

$$\begin{aligned}\delta_2^\chi = \frac{M^4}{F^4}\Big\{&16\left(3c_{10}^r + 6c_{11}^r + 2c_{17}^r + 4\left(c_{18}^r + c_{19}^r\right)\right) - 3\kappa\ell_3^r + 3\kappa\left[-(1/4)\kappa - 3\ell_3^r + 2\ell_7\right]\lambda_M \qquad (109)\\ &- \frac{9}{8}\kappa^2\lambda_M^2 + \left[-48c_{10}^r - 32\left(c_{11}^r + c_{17}^r + 2\left(c_{18}^r + c_{19}^r\right)\right) + \kappa\left(1 - 7\lambda_M\right)\ell_7^r - 4{\ell_7}^2\right]\Delta^2 + 4\Delta^4\ell_7^2\Big\} .\end{aligned}$$

It is a nontrivial consistency check that the dependence on the scale $\overline{\mu}$ cancels separately in any of the previous equations. Moreover, the QED running of the quark masses is compensated by the shift of k_i^r as explained in section 2.4.1, in such a way that both m_{π^0} and $\chi_{\rm top}$ are independent of the QED RG scale μ. Inverting eqs. (99-100) for M and F and plugging the result into eq. (106), we obtain the topological susceptibility $\chi_{\rm top}$ expressed as a function of the physical π^0 mass and f_{π^+} only, as in eq. (52).

A.2 Three-flavor results

In the unbroken isospin limit $m_u = m_d \equiv m$ and at $\alpha_{\rm em} = 0$, the pion mass and decay constant at NNLO in $N_f = 3$ chiral perturbation theory are

$$m_\pi^2 = M_0^2\left[1 + \epsilon_1^M + \epsilon_2^M\right] , \tag{110}$$

$$f_\pi = F_0\left[1 + \epsilon_1^F + \epsilon_2^F\right] , \tag{111}$$

where $M_0^2 \equiv 2B_0 m$, B_0 and F_0 are the LECs of the leading order $N_f = 3$ Lagrangian of [16] and

$$\epsilon_1^M = -\frac{B_0 m_s}{F_0^2}\left\{2\left[\frac{\kappa\lambda_\eta}{9} + 8\left(L_4^r - 2L_6^r\right)\right] + \left[\kappa\left(\frac{\lambda_\eta}{9} - \lambda_0\right) + 16\left(2L_4^r + L_5^r - 4L_6^r - 2L_8^r\right)\right]w\right\} , \tag{112}$$

$$\epsilon_1^F = -\frac{B_0 m_s}{F_0^2}\left\{\left[\frac{\kappa\lambda_K}{2} - 8L_4^r\right] + \left[\kappa\left(2\lambda_0 + \frac{\lambda_K}{2}\right) - 8\left(2L_4^r + L_5^r\right)\right]w\right\} \tag{113}$$

Here m_s is the strange quark mass, $w \equiv m/m_s$, $M_K^2 \equiv B_0(m+m_s)$ and $M_\eta^2 \equiv (4/3)M_K^2 - (1/3)M_0^2$ are the tree-level kaon and eta masses, $\lambda_P \equiv \log\frac{M_P^2}{\overline{\mu}^2}$ for $P = 0, K, \eta$, and L_i^r are the NLO LECs of the $N_f = 3$ chiral Lagrangian of [16]. The terms ϵ_2^M and ϵ_2^F are $\mathcal{O}(p^6)$ and depend on the LECs of the $N_f = 3$ NNLO Lagrangian, C_i^r. Both ϵ_2^M and ϵ_2^F involve the calculation of a two-mass scale sunset integral at non-zero external momentum, and therefore do not admit a closed analytic expression (see [123], where a two-integral representation is provided). However, by employing the analytic form of the two-mass scale sunset integral at vanishing external momentum [123, 129] and the recursion relations for sunset integrals [130–132], ϵ_2^M and ϵ_2^F can be expanded in power series of m. Such an expansion has been calculated for the first time in [130] for ϵ_2^M and [122] for ϵ_2^F up to $\mathcal{O}(m^3)$ and $\mathcal{O}(m^2)$ respectively and is sufficient for the matching of the $N_f = 2$ and $N_f = 3$ LECs discussed in app. B. Since the result of ϵ_2^M and ϵ_2^F turns out to be quite involved, we will avoid reporting here the explicit expressions, for which we refer to [122, 130].

Finally, the topological susceptibility at $\alpha_{\text{em}} = 0$ (first computed to $\mathcal{O}(p^4)$ in [128]) reads

$$\chi_{\text{top}} = \frac{m_u m_d}{m_u + m_d + \frac{m_u m_d}{m_s}} B_0 F_0^2 \left[1 + \epsilon_1^\chi + \epsilon_2^\chi\right] + \mathcal{O}(p^8)\,, \tag{114}$$

where in the unbroken isospin limit $m_u = m_d \equiv m$:

$$\begin{aligned}\epsilon_1^\chi = \frac{B_0 m_s}{F_0^2(1+w/2)} \Bigg\{ &\left[-\kappa\left(\frac{2\lambda_\eta}{9} + \lambda_K\right) + 32L_6^r\right] + \left[-\kappa\left(\frac{5\lambda_\eta}{9} + 3\lambda_0 + 2\lambda_K\right) + 16\left(5L_6^r + 9L_7^r + 3L_8^r\right)\right] w \\ &+ \left[-\kappa\left(\frac{2\lambda_\eta}{9} + \lambda_K\right) + 32L_6^r\right] w^2 \Bigg\}\,,\end{aligned} \tag{115}$$

$$\epsilon_2^\chi = \frac{B_0^2 m_s^2}{F_0^4(1+w/2)} \left[\epsilon_{2,C}^\chi + \epsilon_{2,\log\times\log}^\chi + \epsilon_{2,\log}^\chi + \epsilon_{2,\log\times L}^\chi + \epsilon_{2,L}^\chi + \epsilon_{2,L\times L}^\chi + \epsilon_{2,\text{finite}}^\chi\right]\,. \tag{116}$$

The result of ϵ_2^χ has been conveniently organized into different contributions: $\epsilon_{2,C}^\chi$ are terms containing the LECs C_i^r, $\epsilon_{2,\log}^\chi$ and $\epsilon_{2,\log\times\log}^\chi$ are terms respectively linear and quadratic in the chiral logs $\lambda_{0,K,\eta}$ (but without L_i^r), $\epsilon_{2,L}^\chi$ and $\epsilon_{2,L\times L}^\chi$ are terms linear and quadratic in L_i^r (but independent of chiral logs), and $\epsilon_{2,\log\times L}^\chi$ contains products of chiral logs and LECs. Finally, $\epsilon_{2,\text{finite}}^\chi$ is the remaining constant piece and is automatically scale independent. In particular:

$$\begin{aligned}\epsilon_{2,C}^\chi = 8\{&8\left[C_{20}^r + 3C_{21}^r\right] + 4\left[3C_{19}^r + 7C_{20}^r + 27C_{21}^r + 2C_{31}^r + C_{94}^r + 6\left(C_{32}^r + C_{33}^r\right)\right] w + \\ &4\left[6C_{19}^r + C_{94}^r + 4\left(4C_{20}^r + 9C_{21}^r + C_{31}^r + 3\left(C_{32}^r + C_{33}^r\right)\right)\right] w^2 + \left[8C_{20}^r + 48C_{21}^r + C_{94}^r\right] w^3\}\,,\end{aligned} \tag{117}$$

$$\begin{aligned}\epsilon_{2,\log\times\log}^\chi = \frac{\kappa^2}{2+w} \Bigg\{ &\frac{4}{9}\left[\lambda_\eta(\lambda_\eta - 5\lambda_K)\right] + \frac{1}{3}\left[\frac{46\lambda_\eta^2}{9} + \frac{47\lambda_K^2}{3} - \frac{70\lambda_\eta\lambda_K}{3} - 4\lambda_0^2 + 16\lambda_\eta\lambda_0\right] w \\ &+ \left[10\lambda_0\lambda_\eta + \frac{5\lambda_\eta^2}{3} - \frac{19\lambda_0^2}{3} - 10\lambda_\eta\lambda_K + 18\lambda_0\lambda_K + \frac{14\lambda_K^2}{9}\right] w^2 \\ &+ \left[5\lambda_0\lambda_\eta + \frac{11\lambda_\eta^2}{18} - \frac{71\lambda_0^2}{6} - \frac{46\lambda_\eta\lambda_K}{9} + 12\lambda_0\lambda_K - \frac{7\lambda_K^2}{9}\right] w^3 + \frac{1}{3}\left[\frac{2\lambda_\eta^2}{9} + 2\lambda_0\lambda_\eta - \frac{8\lambda_\eta\lambda_K}{3}\right] w^4 \Bigg\}\,,\end{aligned} \tag{118}$$

$$\begin{aligned}\epsilon_{2,\log}^\chi = &-\frac{\kappa^2}{3}\left\{\frac{2}{3}\left[\frac{10\lambda_\eta}{9} + \lambda_K\right] + \frac{2}{3}\left[\frac{29\lambda_\eta}{3} + 34\lambda_K\right] w + 2\left[\frac{19\lambda_\eta}{9} + 19\lambda_0 + \frac{35\lambda_K}{3}\right] w^2\right. \\ &\left.+2\left[\frac{8\lambda_\eta}{27} - \lambda_0 + \frac{2\lambda_K}{3}\right] w^3\right\}\,,\end{aligned} \tag{119}$$

$$\begin{aligned}\epsilon_{2,\log\times L}^\chi = \frac{8\kappa}{2+w} \Bigg\{ &\left[\frac{16}{9}\left(3L_4^r + L_5^r - 6L_6^r + 3L_7^r - L_8^r\right)\lambda_\eta + 2\left(8L_4^r + 3L_5^r - 16L_6^r - 6L_8^r\right)\lambda_K\right] \\ &+ \Bigg[\frac{8}{9}\left(18L_4^r + 7L_5^r - 36L_6^r - 27L_7^r - 23L_8^r\right)\lambda_\eta + 3\left(20L_4^r + 7L_5^r - 40L_6^r - 24L_7^r - 22L_8^r\right)\lambda_K \\ &+24\left(L_4^r - 2L_6^r\right)\lambda_0\Bigg] w + \Bigg[\frac{20}{3}\left(3L_4^r + L_5^r - 4L_8^r - 6\left(L_6^r + L_7^r\right)\right)\lambda_\eta + \left(82L_4^r + 27L_5^r - 2\left(82L_6^r + 72L_7^r + 51L_8^r\right)\right)\lambda_K \\ &+12\left(7L_4^r + 3L_5^r - 2\left(7L_6^r + 9L_7^r + 6L_8^r\right)\right)\lambda_0\Bigg] w^2 + \Bigg[\frac{2}{9}\left(48L_4^r + 13L_5^r - 96L_6^r - 60L_7^r - 46L_8^r\right)\lambda_\eta \\ &+ \left(48L_4^r + 15L_5^r - 6\left(16L_6^r + 12L_7^r + 9L_8^r\right)\right)\lambda_K + 6\left(8L_4^r + 3L_5^r - 16L_6^r - 6L_8^r\right)\lambda_0\Bigg] w^3\end{aligned} \tag{120}$$

$$+\left[\left(2L_4^r+\frac{4}{9}\left(L_5^r-9L_6^r-2L_8^r\right)\right)\lambda_\eta+\left(10L_4^r+3L_5^r-20L_6^r-6L_8^r\right)\lambda_K+6\left(L_4^r-2L_6^r\right)\lambda_0\right]w^4\Bigg\}\ ,$$

$$\begin{aligned}\epsilon^\chi_{2,L}=\frac{8\kappa}{3}\Bigg\{&\frac{1}{3}\left[22L_4^r+\frac{35L_5^r}{3}-44L_6^r-8L_7^r-26L_8^r\right]+\left[48L_4^r+\frac{35L_5^r}{3}-96L_6^r-\frac{70L_8^r}{3}\right]w\\&+\left[74L_4^r+29L_5^r-148L_6^r+8L_7^r-\frac{166L_8^r}{3}\right]w^2+\frac{1}{3}\left[44L_4^r+\frac{31L_5^r}{3}-88L_6^r-16L_7^r-26L_8^r\right]w^3\Bigg\}\ ,\end{aligned}\tag{121}$$

$$\epsilon^\chi_{2,L\times L}=\frac{1024}{2+w}\left(3L_7^r+L_8^r\right)^2\left\{-w+2w^2-w^3\right\}\ ,\tag{122}$$

$$\begin{aligned}\epsilon^\chi_{2,\text{finite}}=\kappa^2\Bigg\{&\left[-6G\left(\frac{2w}{w+1}\right)-\frac{2}{9}G\left(1\right)+\frac{2}{3}G\left(\frac{2+w}{3w}\right)-\frac{4}{9}G\left(\frac{4}{3}\frac{1+w/2}{1+w}\right)+\frac{80}{9}\right]w\\&+\left[-3G\left(\frac{2w}{w+1}\right)-\frac{1}{9}G(1)-\frac{11}{3}G\left(\frac{2+w}{3w}\right)-\frac{5}{9}G\left(\frac{4}{3}\frac{1+w/2}{1+w}\right)+\frac{160}{9}\right]w^2\Bigg\}\ .\end{aligned}\tag{123}$$

In this last equation we defined

$$G(x)\equiv\frac{1}{\sigma}\left[4\mathrm{Li}_2\left(\frac{\sigma-1}{\sigma+1}\right)+\log^2\left(\frac{1-\sigma}{1+\sigma}\right)+\frac{\pi^2}{3}\right],\qquad\sigma=\sqrt{1-\frac{4}{x}}\,,\tag{124}$$

which arises in the evaluation of a two-mass scale sunset integral at vanishing external momentum [122].

B Extraction of the NNLO LECs and input parameters

As mentioned in sec. 2.4.2, the $\mathcal{O}(p^6)$ LECs c_i^r of the $N_f = 2$ chiral Lagrangian [116,117] are still poorly known from experimental data (see [121] for a review). To estimate the value of the c_i^r's appearing in eq. (50), we combined the results from recent lattice QCD simulations [104] with the information from the matching of pion mass, decay constant and topological susceptibility for $N_f = 2$ and $N_f = 3$ chiral perturbation theory of app. A and the estimates for the $N_f = 3$ LECs C_i^r provided in [121].

- The SU(2) partially quenched simulation of ref. [104] provide fits of 8 combinations of SU(2) LECs. In this analysis we consider the 450 MeV cut-fit for such combinations, reported in the last column of tab. 6 of [104]. While this estimate is less conservative than the one extracted from the 370 MeV cut-fit, the two are compatible for all reported combinations of LECs. In tab. 4 we quote the results expressed in terms of c_i^r using the relations between SU(N) and SU(2) LECs of [116].

- Ref. [121] provides estimates of 34 combinations of $\mathcal{O}(p^6)$ three-flavor LECs, C_i^r. The information of C_i^r can be translated into a value for three combinations of c_i^r by equating in the large m_s/m limit (and for $\alpha_{\text{em}} = 0$ and $m \equiv m_u = m_d$) the $N_f = 2$ and $N_f = 3$ formulas for the pion mass in eqs. (99) and (110), the pion decay constant in eqs. (100) and (111) and the topological susceptibility in eqs. (106) and (114). Such a matching leads respectively to the three following relations:

$$2c_6^r+c_7^r+2c_8^r+c_9^r-3c_{10}^r-2\left(3c_{11}^r+c_{17}^r+2c_{18}^r\right)=\tag{125}$$

$c_6^r - c_{17}^r$	$-5.33(77) \cdot 10^{-6}$
$2c_6^r - 12c_{10}^r + 18c_{11}^r - c_{18}^r$	$14.5(3.9) \cdot 10^{-6}$
c_7^r	$-3.9(2.3) \cdot 10^{-6}$
c_8^r	$0.0(1.8) \cdot 10^{-6}$
$2c_7^r + 4c_8^r$	$6.2(3.2) \cdot 10^{-6}$
c_9^r	$-0.2(1.2) \cdot 10^{-6}$
c_{10}^r	$-1.0(1.1) \cdot 10^{-6}$
$19c_{11}^r - 12c_{10}^r$	$10.1(3.1) \cdot 10^{-6}$

Table 4: *Numerical values of the combinations of SU(2) LECs at the scale $\overline{\mu} = 770$ MeV extracted from the 450 MeV cut-fit of the partially quenched simulations in [121] (tab. 6).*

$$
\begin{aligned}
&= \frac{\kappa f_\pi^2}{1152 m_\pi^2}\frac{m}{m_s} + \left[2C_{12}^r + 4C_{13}^r + C_{14}^r + 2C_{15}^r + 2C_{16}^r + C_{17}^r - 3C_{19}^r - 2\left(3C_{20}^r + 6C_{21}^r + C_{31}^r + 2C_{32}^r\right)\right] \\
&+ \frac{\kappa^2}{384}\left[\frac{1}{9}\lambda_\eta^2 + \lambda_\eta\lambda_K + \frac{13}{6}\lambda_K^2\right] + \frac{\kappa^2}{589824}\left[\frac{20975}{3}\lambda_\eta + 44123\lambda_K\right] + \kappa\Bigg[\frac{1}{36}\left(4L_1^r + L_2^r + L_3^r - 10L_4^r\right. \\
&\left. -3L_5^r + 12L_6^r + 12L_7^r + 10L_8^r\right)\lambda_\eta + \left(L_1^r + \frac{L_2^r}{4} + \frac{5L_3^r}{16} - L_4^r + L_6^r + \frac{L_8^r}{2} - \frac{L_5^r}{4}\right)\lambda_K\Bigg] + \\
&+ \frac{\kappa}{3}\left[5L_1^r + \frac{5L_2^r}{6} + \frac{205L_3^r}{144} + 6L_6^r + \frac{11L_7^r}{24} + \frac{73L_8^r}{24} - \frac{17L_4^r}{3} - \frac{13L_5^r}{9}\right] \\
&- 4\left[\left(2L_4^r + L_5^r\right)\left(2L_4^r + L_5^r - 2\left(2L_6^r + L_8^r\right)\right)\right] + \frac{\kappa^2}{589824}\left[\frac{1373}{4}G\left(\frac{4}{3}\right) + \frac{219836}{3}\right],
\end{aligned}
$$

$$
\begin{aligned}
&c_7^r + 2c_8^r + c_9^r = \qquad\qquad (126) \\
&= -\frac{\kappa f_\pi^2}{64 m_\pi^2}\frac{m}{m_s} + \left[C_{14}^r + 2C_{15}^r + 2C_{16}^r + C_{17}^r\right] - \frac{\kappa^2}{384}\left[\frac{7}{3}\lambda_\eta\lambda_K + \lambda_K^2\right] + \frac{\kappa^2}{589824}\left[44549\lambda_K - 10245\lambda_\eta\right] \\
&+ \kappa\left[\frac{1}{36}\left(4L_1^r + L_2^r + L_3^r - 2L_4^r - L_5^r\right)\lambda_\eta + \left(L_1^r + \frac{L_2^r}{4} + \frac{5L_3^r}{16} + \frac{L_4^r}{4} - L_6^r - \frac{L_8^r}{4}\right)\lambda_K\right] + \\
&+ \frac{\kappa}{3}\left[5L_1^r + \frac{5L_2^r}{6} + \frac{205L_3^r}{144} + \frac{5L_4^r}{4} + \frac{23L_5^r}{48} - 6L_6^r - \frac{15L_8^r}{8}\right] - \left[2L_4^r + L_5^r\right]^2 + \frac{\kappa^2}{49152}\left[\frac{5825}{16}G\left(\frac{4}{3}\right) + 5225\right],
\end{aligned}
$$

$$
\begin{aligned}
&3c_{10}^r + 6c_{11}^r + 2c_{17}^r + 4c_{18}^r + 4c_{19}^r = \qquad\qquad (127) \\
&= \frac{f_\pi^2}{2m_\pi^2}\frac{m}{m_s}\left[\frac{\kappa}{32}\left(\frac{\lambda_\eta}{3} + \frac{\lambda_K}{2} - \frac{23}{18}\right) - \left(\frac{3L_6^r}{2} + 9L_7^r + 3L_8^r\right)\right] + \Bigg[\frac{3C_{19}^r}{2} + \frac{21C_{20}^r}{4} + \frac{27C_{21}^r}{4} + C_{31}^r \\
&+ 3C_{32}^r + 3C_{33}^r\Bigg] - \frac{\kappa^2}{27}\left[\lambda_\eta^2 + \frac{3\lambda_K}{256}\left(26\lambda_\eta + 113\lambda_K\right)\right] - \frac{\kappa^2}{768}\left[\frac{1346\lambda_\eta}{27} + 95\lambda_K\right] + \frac{\kappa}{4}\Bigg[\Bigg(\frac{4L_4^r}{3} - 3L_7^r - L_8^r \\
&- \frac{10L_6^r}{3}\Bigg)\lambda_\eta + \frac{3}{8}\left(12L_4^r + 3L_5^r - 24L_6^r - 24L_7^r - 14L_8^r\right)\lambda_K\Bigg] + \frac{\kappa}{24}\Bigg[\frac{317L_4^r}{6} + \frac{217L_5^r}{12} - 127L_7^r - \frac{157L_8^r}{2} \\
&- \frac{317L_6^r}{3}\Bigg] + 4\left[-3\left(L_6^r + 6L_7^r + 2L_8^r\right)L_4^r + 6\left(L_6^r\right)^2 + 7\left(3L_7^r + L_8^r\right)^2 + 12L_6^r\left(3L_7^r + L_8^r\right)\right] + \frac{\kappa^2}{576}\Bigg[G(1) \\
&+ \frac{1}{2}G\left(\frac{4}{3}\right) - 7\pi^2 - \frac{2275}{9}\Bigg].
\end{aligned}
$$

The notation in eqs. (125-127) is as in app. A.2 (except that m/m_π^2, f_π and $M_{K,\eta}^2$ in $\lambda_{K,\eta}$ are

L_1^r	L_2^r	L_3^r	L_4^r	L_5^r	L_6^r	L_7^r	L_8^r	
0.5(2)	0.8(3)	−3.1(1.0)	0.09(34)	1.19(40)	0.16(20)	−0.34(11)	0.55(18)	$\times 10^{-3}$

C_{12}^r	C_{13}^r	C_{14}^r	C_{15}^r	C_{16}^r	C_{17}^r	C_{19}^r	C_{20}^r	C_{21}^r	C_{31}^r	C_{32}^r	C_{33}^r	
−2.8	1.5	−1.0	−3.0	3.2	−1.0	−4.0	1.0	−0.48	2.0	1.7	0.82	$\times 10^{-6}$

Table 5: *Numerical value of the NLO and NNLO couplings* L_i^r *and* C_i^r *at the scale* $\overline{\mu} = 770$ *MeV. We associated* 100% *uncertainty to the* C_i^r.

computed at $m_u = m_d = m = 0$), and the contributions on the r.h.s. have been ordered as in eq. (116). The numerical value of the NLO couplings L_i^r is reported in tab. 5: in particular, L_4^r, L_5^r and L_6^r are taken from lattice QCD studies [103], while the others from ref. [121]. Moreover, to all L_i^r a 30% intrinsic uncertainty from higher order 3-flavor corrections has been added (this is not present for 2-flavor where higher order corrections are much smaller). The value of the NNLO couplings C_i^r appearing in the r.h.s. of eqs. (125-127) taken from tab. 4 of ref. [121] is also reported in tab. 5. Since ref. [121] did not provide uncertainties for the C_i^r coefficients we assume that they reproduce at least the right orders of magnitude and conservatively assign to them a 100% uncertainty.

Eqs. (125-127) then lead to:

$$\begin{aligned} 2c_6^r + c_7^r + 2c_8^r + c_9^r - 3c_{10}^r - 2\left(3c_{11}^r + c_{17}^r + 2c_{18}^r\right) &= -\,3.5(22.0)\cdot 10^{-6}\,, \\ c_7^r + 2c_8^r + c_9^r &= 4.7(9.2)\cdot 10^{-6}\,, \\ 3c_{10}^r + 6c_{11}^r + 2c_{17}^r + 4c_{18}^r + 4c_{19}^r &= 0.3(25.5)\cdot 10^{-6}\,. \end{aligned} \tag{128}$$

The final value of the 9 couplings c_i^r in tab. 3 has been extracted by combining the lattice results in tab. 4 with the 2-3 flavor matching result in eq. (128) through a χ^2 fit, whose quality ($\chi^2 \sim 3$) turns out to be good. In principle, an estimate of c_6^r could be directly extracted from the pion scalar form factor, as in ref. [121] where $c_6^r \approx -1.9 \times 10^{-5}$. However, since there is still a factor ~ 3 uncertainty on how to theoretically model this last quantity [118–120], we chose not to use this estimate of c_6^r in our numerical analysis. In any case, the NNLO corrections to $\chi_{\rm top}$ in eq. (50) taking into account also $c_6^r = -1.9(1.9) \times 10^{-5}$ result in $\delta_2 = -0.006(3)$, still compatible with eq. (51), but with an overall lower quality fit of the c_i^r.

Finally, for convenience in tab. 6 we summarize the values of the parameters used in this work, which should be considered together with the LECs in tabs. 1, 2, 3 and 5. When uncertainties are not quoted it means that their effect was negligible and they have not been used.

C Details of the Simulation

C.1 Evolution of the Field Equations

For the purpose of implementing numerical simulations, it is convenient to rewrite the equations of motion, given in eqs. (70) and (73), in terms of the rescaled field $\psi = R(t)\phi/f_a$, and the conformal time τ, which is defined as

$$\tau(t) = \int_0^t \frac{dt'}{R(t')} \propto t^{1/2}\ . \tag{129}$$

z	0.472(11)	eq. (32)	w^{-1}	27	[103]
f_π	92.3	[96]	$\bar{\ell}_3$	2.81(49)	[104]
m_{π^0}	134.98	[96]	$\bar{\ell}_4$	4.02(25)	[104]
m_{π^+}	139.57	[96]	$h_1^r - h_3^r - \ell_4^r$	-0.0049(12)	eq. (33)
m_K	495	[96]	ℓ_7	0.0065(38)	[104]
m_ρ	775	[96]	α_{em}^{-1}	137	[96]
m_μ	105.658	[96]	$\Gamma_{\pi^+\to\mu\nu(\gamma)}$	$2.5281 \cdot 10^{-14}$	[96]
G_F	$1.16638 \cdot 10^{-11}$	[96]	V_{ud}	0.9742	[96]

Table 6: *Numerical input values used in the computations. Dimensionful quantities are given in MeV.*

The time-dependent mass in the fat string scenario, eq. (73), is $m_r(\tau) = (\tau_i/\tau)m_i$, where $\tau_i \equiv \tau(t_i)$ is the time at which $m_r = m_i$. In this way, the equations of motion simplify to

$$\psi'' - \nabla_x^2\psi + u(\tau)\psi\left(|\psi|^2 - \frac{R^2}{2}\right) = 0 \ , \tag{130}$$

where ψ' and $\nabla_x\psi$ are derivatives with respect to the dimensionless time and distance variables $m_r\tau$ and $m_r x$ respectively (or $m_i\tau$ and $m_i x$ for the fat string case). In the physical scenario $u(\tau) = 1$, while in the fat string case $u(\tau) = \tau_i/\tau$.

We then solve eq. (130) numerically on a cubic lattice with periodic boundary conditions.[42] Space is discretised in a box of comoving side length L_c containing $N^3 = 1250^3$ uniformly distributed grid points, where the upper value of N is limited by our memory budget. Consequently the space-step between grid points in comoving coordinates is $a_c = L_c/N$, which is constant in time. It is again convenient to work in terms of the dimensionless comoving space-step $m_r a_c$ in the physical case, and $m_i a_c$ for the fat string system.

The physical length of the box is $L(t) = L_c R(t)$ and the physical space-step between grid points $a(t) = L/N = a_c R(t)$ grows $\sim t^{1/2}$. In the physical string scenario, the string core size m_r^{-1} is constant, and therefore the number of grid points in a core $(m_r a(t))^{-1} \sim 1/t^{1/2}$ decreases. Meanwhile, in the fat string scenario the string core size increases $\sim t^{1/2}$, and as a result the number of grid points in a string core $(m_r(t)a(t))^{-1} = (m_i a_c)^{-1}$ is constant. When considering systematic errors from the space steps in the fat string system we use the notation $m_r(t)a(t)$ for the size of the space step due to its direct physical interpretation (although this is entirely equivalent to $m_i a_c$).

The equations of motion are discretised following a standard central-difference Leapfrog algorithm for wave-like PDEs (see e.g. [146]). The system is evolved in fixed steps of conformal time a_τ, and we work in terms of the dimensionless time-step $m_r a_\tau$ in the physical case, and $m_i a_\tau$ in the fat string case. The derivatives are expanded to fourth order in the space-step and second order in the time-step.[43] In Appendix D we extensively study the systematics from the discretisation of space and time, as well as from finite volume effects.

As discussed in Section 4.3, we set the initial conditions in two ways, the second of which is used to produce a cleaner initial configuration with a fixed number of strings per Hubble volume.

(a) Random initial conditions: ϕ and $\dot{\phi}$ are both generated through the anti-Fourier transform of

---

[42] We parallelize the algorithm to step forward in time, and run on a cluster with 2 × 24-cores.

[43] The fourth order discretisation of space is probably not required, since additional tests show using a (less precise) second order discretisation does not lead to significant differences in the results, at least for the string length.

Fourier coefficients $\tilde{\phi}(\vec{k})$ and $\tilde{\dot{\phi}}(\vec{k})$, randomly chosen in the interval $[-\frac{f_a}{\sqrt{2}}, \frac{f_a}{\sqrt{2}}]$ for $|\vec{k}| \leq k_{max}$ and zero for $|\vec{k}| > k_{max}$ for a fixed choice of $k_{max} \in [0, m_r]$ in each set of simulations. Larger values of k_{max} lead to initial fluctuations with smaller wavelength, of order $2\pi/k_{max}$, and more initial strings. As a result, the initial string density is controlled by the parameter k_{max}/m_r. Even though this method produces configurations in which the axion field winds the fundamental domain $[0, 2\pi f_a]$ nontrivially, it does not lead to a clean string configuration at the initial time, because $|\phi|$ does not typically resemble the profile function of string-like solutions. Instead, the system takes some time to relax to a string network solution and in doing so releases a large quantity of energy, which produces extra contamination to the axion spectrum measured at later times. To solve this issue, we employ the method below that also allows us to construct axion field configurations with a predetermined string length, or equivalently with predetermined initial ξ.

(b) Fixed string number: This approach simply consists of evolving a field configuration produced by method (a) with $k_{max} = m_r$ until the required string length inside the box is obtained, and then using that configuration as initial condition for the actual simulation, with a different initial value of the Hubble parameter. Since this involves resetting the Hubble parameter, the strings produced do not in general have the right core-size, but we have checked that they quickly relax to their expected thickness long before the scaling regime is reached.

C.2 String Identification and String Length

To identify strings and measure their length, we first identify grid points that are likely to be close to a string core. We have carried this out in two ways, and have verified that the results obtained are extremely similar. In our main approach, we flag points such that, as a loop that surrounds it is travelled, there is at least one change in the axion field Δa between consecutive lattice points encountered that satisfy $|\Delta a|/f_a > \pi/2$. In particular, the loop is taken to be a square of side length 2 grid points. In order to capture strings with all possible orientations, at each point in the grid we consider loops in three orthogonal planes, and we flag a point if a loop in any of these satisfies the condition. We have checked that the results obtained are extremely similar for any reasonable choice of the threshold value for the change between adjacent grid points.

Having identified points close to the string core, we then combine these into strings. In particular, we cluster together flagged points that are adjacent in the x-y plane into an individual string point located at the mean of the flagged points. If such a cluster has a non-zero overlap with a cluster at the next level up or down in the z direction, these are connected into a string segment. As expected, the reconstructed strings form loops, and the length of each individual loop (which is required to analyse the loop distribution) as well as the total string length is recorded.

We have validated our string identification algorithm by comparing to the results obtained following the procedure adopted in [135]. In this, a lattice plaquette is identified as containing a string if the minimum axion field range that includes the field values on the four surrounding vertices satisfies $|\Delta a|/f_a > \pi$. It can be seen that this leads to the correct results for the prototype string solution, and indeed for any string solution for which the 2π field change incurred in traversing an enclosing path is distributed sufficiently homogeneously around the loop.

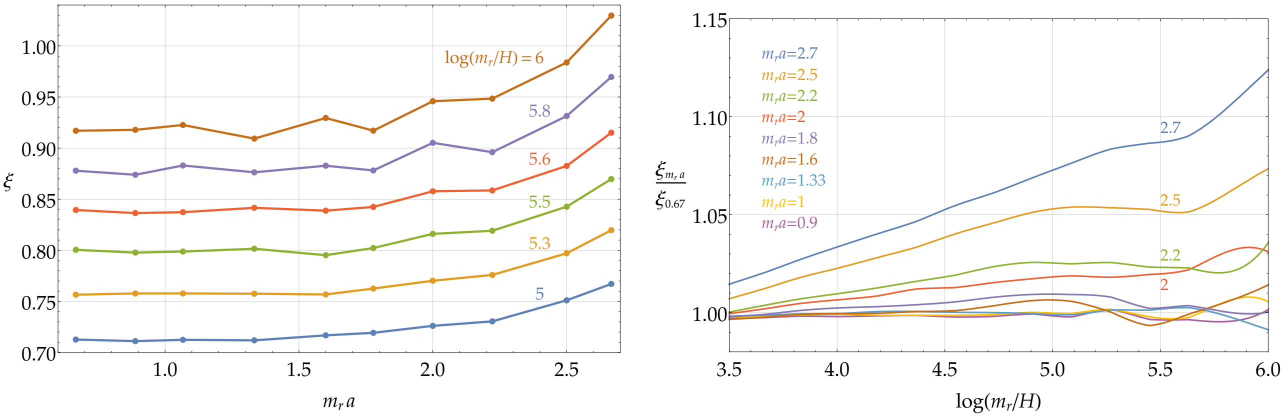

Figure 18: Left: The continuum extrapolation of $\xi(t)$ in the step-size, $m_r a \to 0$. The values obtained with $m_r a = 1.33$ appear to have already converged to the continuum limit, typically with better than percent level accuracy. Right: Results for $\xi(t)$ for different choices of the step-size $m_r a$, normalised to the same quantities calculated with the finest lattice spacing $m_r a = 0.67$. The other unphysical parameters are fixed to $a_\tau / a_c = 1/3$ and $H_f L_f = 2$.

D Analysis of Systematic Errors

To check that our results are free from numerical artifacts, we study how the key observables depend on the unphysical parameters in our simulations. Of particular importance are: (1) the lattice spacing a, (2) the time spacing a_τ, (3) the number of Hubble patches HL contained in the box, and (4) the way that the string cores are screened when evaluating the axion spectrum and energy. In this section we study the first three of these, postponing (4) until Section E. The results that we present here are all obtained from the fat string scenario, however we have tested that the conclusions we reach are also valid for the physical case.

D.1 Lattice Spacing

We study the dependence of $\xi(t)$ and of the axion spectrum on the discretisation parameter $m_r a$ to find the largest value compatible with the continuum limit $m_r a = 0$. As mentioned, $m_r a$ is constant and is equal to the number of grid points per string core at all times in fat string simulations (in contrast, for the physical system the number of points per string core decreases with time, and this source of systematic errors only becomes relevant towards the end of simulations). If $m_r a \gg 1$, the number of gridpoints per string core is much smaller than 1 and the evolution of strings is not correctly captured. Conversely, making the space-step unnecessarily small would restrict the maximum scale separations that could be analysed.

We have performed sets of simulations averaging over 20 samples, keeping all parameters fixed except $m_r a$, with identical initial conditions.

In Figure 18 we plot the continuum extrapolation for ξ and its dependence on $\log(m_r/H)$ for the different values of the space-step, normalised to those from the finest lattice spacing tested ($m_r a = 0.67$). For all $m_r a \leq 1.33$, $\xi(t)$ differs from the most precise result by less than 1%. However, for larger values of $m_r a$, it is systematically larger, especially at later times. In the main text, for our results of $\xi(t)$ and the loop distribution in the fat string scenario, we used the rather conservative

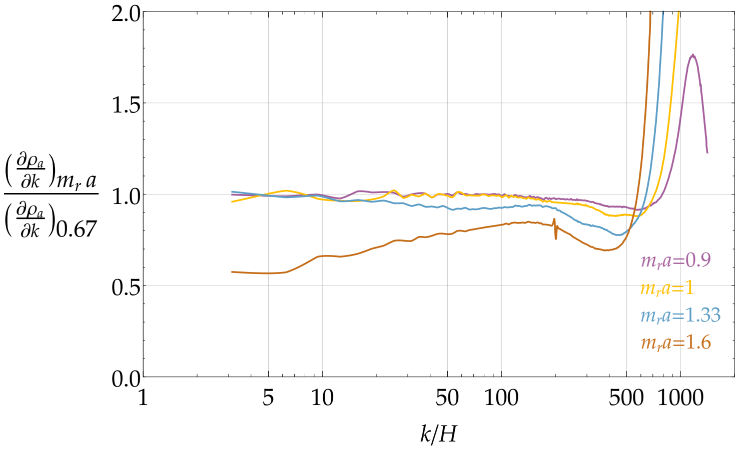

Figure 19: Axion spectrum at $\log(m_r/H_f) = 6$ for different choices of the step-size $m_r a$, normalised to the same quantities calculated with the finest lattice spacing $m_r a = 0.67$. The other unphysical parameters are fixed to $a_\tau/a_c = 1/3$ and $H_f L_f = 2$.

choice $m_r a = 1.33$. In the physical scenario the number of grid points inside a string core decreases as the simulation progresses. We use parameters such that, at the final time, we match the resolution in the fat string case $m_r a = 1.33$ (which means that at early times there are many more grid points inside a string core).

In Figure 19 we plot the axion spectrum at the time $\log(m_r/H) = 6$, for different space steps, again normalised to the results with $m_r a = 0.67$. This is slightly more sensitive to the value of $m_r a$ than $\xi(t)$ is, and we see that discretisation effects increase the production of UV states, reducing the energy emitted in the form of low momentum axions. In particular, only simulations with $m_r a \leq 1$ seem to have converged to within few percent of the continuum limit, while $m_r a = 1.33$ and 1.6 result in a spectrum that is systematically smaller in the IR. In the main text we have chosen $m_r a = 1$ in our analysis of the spectrum in the fat string system; this value is sufficiently small that we are confident that the UV dominated spectrum obtained is not an artifact of the finite space-step. Similarly, in the physical case we use parameters such that $m_r a = 1$ when the final time shot used in calculating the spectrum is taken.

D.2 Time Spacing

We study the dependence of our results on the time-step similarly. Given the form of eq. (130), the general theory of numerical solutions of PDEs tells us that the relevant quantity to which $m_i a_\tau$ should be compared is the comoving space-step $m_i a_c$ [146].[44] Consequently, we perform sets of simulations that differ only in the value of a_τ/a_c, averaging over 10 samples. In Figure D.2 we plot the results for $\xi(t)$ and the spectrum. For all values of a_τ/a_c, $\xi(t)$ is within 0.5% of the continuum limit, but for $a_\tau/a_c \leq 1/3$ the difference is less that 0.1%. The spectrum is also affected at less than percent level for $a_\tau/a_c < 1/3$, and we use $a_\tau/a_c = 1/3$ for our simulations in the main text in both the fat string and the physical scenarios.

---

[44] The Courant condition for the stability of the finite-difference algorithm sets an upper bound on the ratio $a_\tau/a_c < 1$ for the algorithm to converge, but further analysis is still required to quantify the systematic uncertainty.

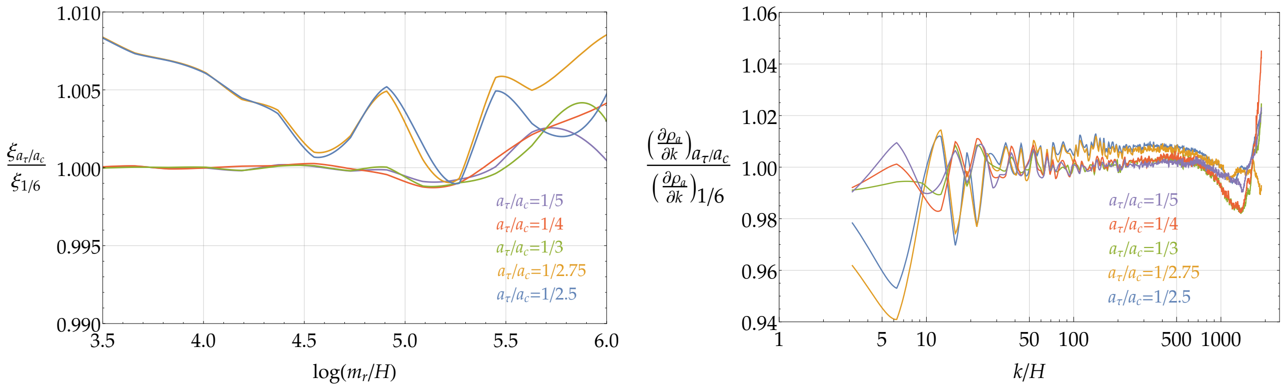

Figure 20: Results for $\xi(t)$ (left) and the axion spectrum at $\log(m_r/H_f) = 6$ (right), for different choices of the time-step a_τ. This is measured relative to the comoving space-step a_c, which is fixed to $m_i a_c = m_r a = 0.67$, and the final box size is fixed to $H_f L_f = 2$. The results are normalized to the shortest time step tested: $a_\tau/a_c = 1/6$.

D.3 Finite Volume

Finite volume effects are a particularly delicate issue. These are controlled by the parameter HL, which counts the number of Hubble lengths per box side length, with the physical limit for a spatially flat Universe corresponding to $(HL)^{-1} \to 0$.

Since HL decreases over the course of a simulation, we want to determine the smallest value that it can take at the final time, $H_f L_f$, such that the results obtained match those in the physical limit. Although we use periodic boundary conditions, if $H_f L_f > 2$ every point is causally disconnected from itself from the beginning to the end of the simulations. Nevertheless, even choosing $H_f L_f > 2$, finite volume effects might still affect some observables such as the loop distribution (see Section 4.3.3).

We performed a set of simulations keeping the values of all the parameters except for $H_f L_f$ fixed (all ending at a final scale separation $\log(m_r/H_f) = 6$). In Figure 21 we show the convergence of ξ to the infinite volume limit for $HL > 2$ for different values of $\log(m_r/H)$ and as a function of $\log(m_r/H)$ (where we indicate with vertical lines the times at which $HL = 2$ for each simulation). Finite volume effects do not affect ξ while $HL > 2$, and they remain small even slightly later. However, for $HL \lesssim 1$ they result in a dramatic change, since the the whole network is in causal contact and starts to be destroyed.

We also study how the loop distribution is affected by finite volume effects. In Figure 22 we show the ratio ξ_ℓ/ξ_∞, defined in Section 4.3.3, for different values of ℓ, for a set of simulations with $H_f L_f = 2$ and $\log(m_r/H_f) = 6.7$. Vertical lines correspond to the times at which $HL = 5, 4, 3$. As mentioned in Section 4.3.3, the constant value of ξ_ℓ/ξ_∞ in time is a strong indication that the system is in the scaling regime. However, the finite box size limits the maximum loop radius that can be contained. As a result, finite volume effects lead to ξ_ℓ/ξ_∞ growing at late times once larger loops are no longer possible, and the larger ℓ is the earlier this occurs. For our analysis of ξ and the loop distribution we have chosen $H_f L_f = 2$ for both the fat string and the physical systems. This is rather conservative for ξ, but slightly sub-optimal for the study of the distribution of relatively long loops with $\ell H/\pi \gtrsim 1/2$ (although earlier time shots are also plotted in Figure 8, so the effect of the finite volume is clear).

In Figure 23 (left) we plot the axion spectrum at $\log(m_r/H_f) = 6$ for simulations with different

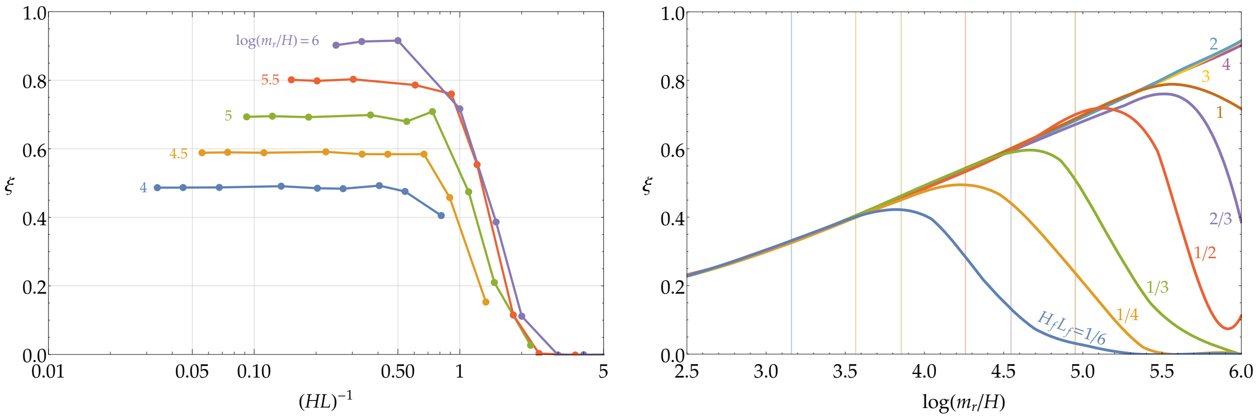

Figure 21: Left: The extrapolation of $\xi(t)$ to the infinite box size limit, $(HL)^{-1} \to 0$. The value $(HL)^{-1} = 1/2$ seems to have already converged, with better than percent level accuracy at most times. Right: $\xi(t)$ calculated in simulations with different box sizes (measured by $H_f L_f$); vertical lines indicate the times at which $HL = 2$ in each simulation.

values of $H_f L_f$. The peak at the Hubble scale is well reproduced for $H_f L_f \geq 2$, even if its position is dangerously close to the IR cut-off (indeed, both are of order Hubble). On the other hand, for $H_f L_f \lesssim 1$, not only is the Hubble peak not present, but there is also a significant overproduction of UV modes, related to the shrinking of a significant fraction of the string network, which towards the end of the simulation is made up of loops with radius smaller than the Hubble distance.

In Figure 23 (right) we plot the axion number density n_a as a function of $\log(m_r/H)$ for simulations with different values of $H_f L_f$; with vertical lines corresponding to times at which $HL = 3$ for each simulation. n_a has been normalised to the values obtained in the simulation with $H_f L_f = 4$, which is the least affected by finite volume effects. The number density seems to be slightly more sensitive to the finite box size than ξ. While $HL > 3$ the effect on n_a is less than one percent, for $HL < 3$ finite volume effects result in a systematic underestimation of a few percent. This is reasonable, since n_a is strongly dependent on the IR of the axion spectrum, which is the part most sensitive to finite volume effects. For this reason in the main text we chose $H_f L_f = 3$ when studying both the spectrum and the number density, for the fat string and also the physical system.

We have also studied the dependence on $H_f L_f$ of the effective string tension $\mu_{\rm eff}$. This depends on the distribution of string length in loops of different shapes and sizes, since e.g. the logarithmic divergence in the string tension is expected to be cut off at a smaller scale for small loops than long strings. Consequently, $\mu_{\rm eff}$ could in principle be sensitive to finite volume effects. However, we do not observe any change in the string tension for $H_f L_f \geq 2$.

E String Screening, Energies and the Axion Spectrum

E.1 Components of the energy

In this section we discuss in more detail the way that the total energy density of the system $\rho_{\rm tot}$ splits into the contributions defined in eq. (69), and demonstrate that our method of screening the strings when extracting the energy in free axions and radial modes is consistent and does not introduce

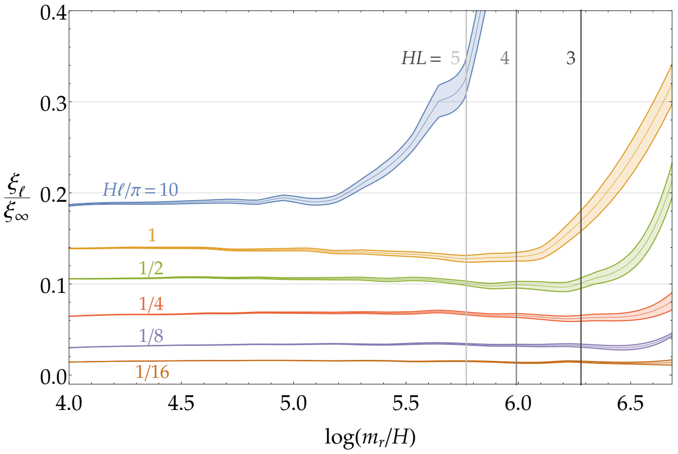

Figure 22: ξ_ℓ/ξ_∞ for the set of simulations with $H_f L_f = 2$, for different values of ℓ. The shaded bands represent statistical errors in the average over different simulations. The other parameters are fixed as $m_r a = 1.33$ and $a_\tau/a_c = 1/3$.

systematic errors. We also study the energy contributions in the absence of strings, and the effect of the interaction term between axion and radial modes. Finally, we describe the algorithm that we use to calculate the axion spectrum.

As summarised in Section 4.4.1, the total energy density is written as the sum $\rho_{\rm tot} = \rho_s + \rho_a + \rho_r$. The energy density in free axions ρ_a has been calculated as the spatial average $\rho_a = \langle \dot{a}^2 \rangle$ far away from string cores, and in these regions axion and radial modes are to a good approximation decoupled at sufficiently late times. Similarly, the radial energy density is computed from $\rho_r = \langle \frac{1}{2}\dot{r}^2 + \frac{1}{2}\left|\nabla r\right|^2 + V(r) \rangle$, again away from the string cores. We define the parameter d_s such that the minimum distance from the strings' centre of the regions that are included in the average is $d_s m_r^{-1}$, so that, at all times, d_s is the screening distance in units of the core size for both the fat string and physical scenarios.

There is actually an alternative approach to calculating the energy density of axions, which does not require screening. This consists of including the interaction terms in the axion energy density, i.e. evaluating $\langle (1 + r/f_a)^2 \dot{a}^2 \rangle$ over the whole space. On the string centre the factor $(1 + r/f_a)^2$ vanishes, and the contribution from the core regions is thus suppressed. Calculated in this way, the axion energy density will include the contribution from the axion-radial interaction energy. When we discuss the energy contributions to $\rho_{\rm tot}$, we will show that energy densities in axions calculated in both ways, i.e. with $\langle \dot{a}^2 \rangle$ screened and with $\langle (1 + r/f_a)^2 \dot{a}^2 \rangle$, are in close agreement.

In Figure 24 we plot the ratio of different components of the energy density to $\rho_{\rm tot}$ as a function of $m_r t/2\pi$ for a single simulation of the fat string system carried out on a small grid, with a space-step $m_r a = 1$. The goal here is not to understand how energies behave in the scaling regime (which has not been reached in the simulation analysed here), but rather to check that our approach to calculating the energy in axion and radial modes is consistent. All of the energy contributions plotted are calculated as the spatial average over regions a distance of at least $d_s = 1$ away from string centres, except for $\frac{1}{2}\langle (1 + r/f_a)^2 \dot{a}^2 \rangle$, which is averaged over the whole space. The simulation has been run until the box only contains 1/8 of a Hubble volume, i.e. $H_f L_f = 1/2$. As a result the entire system is in causal contact, and all the strings are destroyed at about $m_r t/2\pi \approx 19$, before the end of the simulation.

Many key features concerning the way that the total energy is split up can be understood from

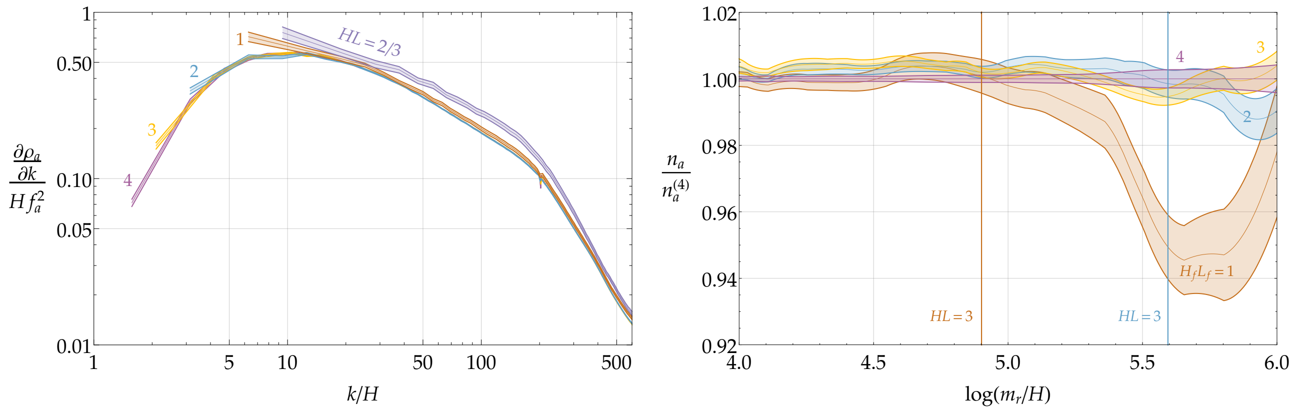

Figure 23: The spectrum at $\log(m_r/H) = 6$ (left) and axion number density (normalised to the number density calculated in a simulation with the largest box size tested: $H_f L_f = 4$) as a function of $\log(m_r/H)$ (right) for different final numbers of Hubble patches at the end of the simulation. Error bars represent statistical errors in the average over different samples.

Figure 24. First, the kinetic part $\frac{1}{2}\langle \dot{a}^2 \rangle$ of the axion energy outside the string cores is systematically smaller than the gradient part $\frac{1}{2}\langle |\nabla a|^2 \rangle$ outside the string cores while strings are present. This is expected since the former only gets contributions from free axions (in the approximation that the motions of strings does not have a significant effect on $\dot{a}$ far from string cores), while the latter contains energy from both axion waves and a large fraction of the tension of strings. However, once the string network starts shrinking $\frac{1}{2}\langle |\nabla a|^2 \rangle$ decreases, and both $\frac{1}{2}\langle \dot{a}^2 \rangle$ and the radial energy $\langle \frac{1}{2}\dot{r}^2 + \frac{1}{2}|\nabla r|^2 + V(r) \rangle$ increase. This indicates that as the strings are destroyed the energy stored in their tension is transferred to free axion and radial modes.

As expected for free classical waves, after the network shrinks the axion energy $\langle \frac{1}{2}\dot{a}^2 + \frac{1}{2}|\nabla a|^2 \rangle$ is repeatedly interchanged between its kinetic and gradient parts, which on average are equal. Similarly, the total energy in radial modes $\langle \frac{1}{2}\dot{r}^2 + \frac{1}{2}|\nabla r|^2 + V(r) \rangle$ behaves as expected (for free states with a mass decreasing adiabatically in time, due to the fat string trick), with $\langle \frac{1}{2}\dot{r}^2 \rangle$ and the sum $\langle \frac{1}{2}|\nabla r|^2 + V(r) \rangle$ giving equal contributions. After the string network disappears, the ratios of axion and radial energy to the total energy stay approximately constant. This is a sign that axions and radial models (as well as the total energy) redshift at the same rate, i.e. as massless radiation, or equivalently as massive radiation with a mass decreasing with time as $\sim 1/R(t)$.

The two ways of computing the axion energy density: evaluating $\langle \dot{a}^2 \rangle$ outside strings or $\langle (1 + r/f_a)^2 \dot{a}^2 \rangle$ over the whole space, are compatible within few percent while strings are present. The small discrepancy arises because the factor $(1 + r/f_a)$ is not exactly a step function at the edge of the string cores. As we will see in the calculation of the spectrum, the difference is stored in UV axion modes (as expected) and therefore do not affect the calculated axion number density. Once the strings disappear the two measurements of the energy density give almost identical results since the difference is due to the non-vanishing interaction energy $\frac{1}{2}(r^2/f_a^2 + 2r/f_a)\dot{a}^2$, which goes to zero as the Universe expands.

The sum $\langle \frac{1}{2}\dot{a}^2 + \frac{1}{2}|\nabla a|^2 + \frac{1}{2}\dot{r}^2 + \frac{1}{2}|\nabla r|^2 + V(r) \rangle$ converges to the total energy $\rho_{\rm tot}$ only once all the strings have been destroyed. Prior to this, a significant proportion of the energy is stored in string cores, and this is largest at small values of the log. The difference corresponds to part of the string

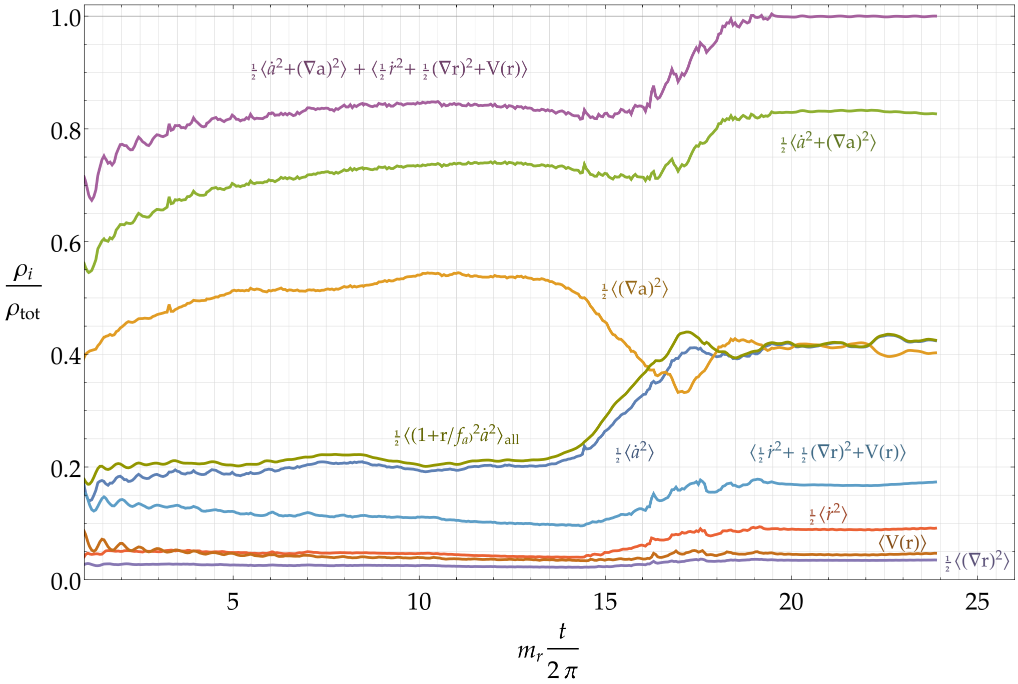

Figure 24: Different contributions to the total energy density $\rho_{\rm tot}$, each one normalised to $\rho_{\rm tot}$, as a function of time for fat string system. All contributions are calculated as the spatial average at a distance $d_s = 1$ from the strings' center, except for $\frac{1}{2}\langle(1+r/f_a)^2\dot{a}^2\rangle$, which is calculated over the whole space. The string network begins to be destroyed by the finite box size at $m_r t/2\pi \sim 19$, corresponding to the purple line going to 1.

energy, with the remainder coming from the difference $\langle\frac{1}{2}\left|\nabla a\right|^2 - \frac{1}{2}\dot{a}^2\rangle$ away from the string cores. The small mismatch between $\langle\frac{1}{2}\dot{a}^2 + \frac{1}{2}\left|\nabla a\right|^2 + \frac{1}{2}\dot{r}^2 + \frac{1}{2}\left|\nabla r\right|^2 + V(r)\rangle$ and $\rho_{\rm tot}$ in the absence of strings is due to the non-vanishing axion-radial interaction energy, which however decreases as the Universe expands.

We also note that the energy densities in axions and radial modes have small oscillations with frequency $\mathcal{O}(m_r)$. The amplitude of the oscillations decreases with time due to redshifting, and as a result for a fixed value of the log they are larger in the physical case compared to the fat string scenario. Calculating the instantaneous axion emission spectrum involves taking the difference between the axion spectra (appropriately redshifted) at two times, which are separated by more than m_r^{-1}. Consequently, the result obtained for momenta around m_r is sensitive to the phase of the oscillations at the two times. This is the origin of the fluctuations of the instantaneous spectrum in Figure 15, which as expected are more significant in the physical case. To reduce this effect in the physical case we averaged over simulations starting at slightly different times, with relative differences of order $2\pi/m_r$.

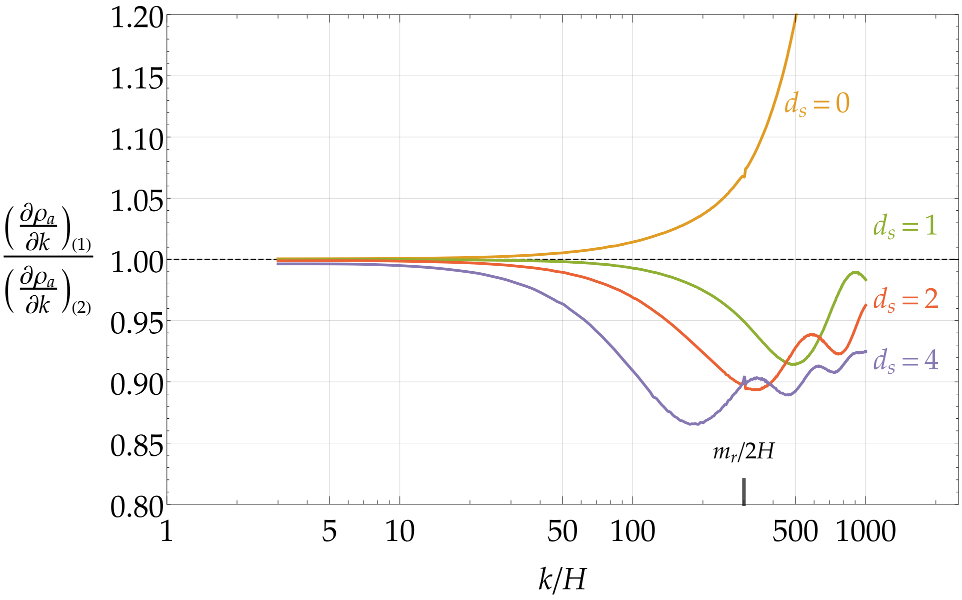

Figure 25: The spectrum calculated with method (1), described in the text, for different screening distances d_s, normalised to the spectrum calculated with method (2).

E.2 The Axion Spectrum

We now describe the details of the calculation of the differential axion spectrum $\partial\rho_a/\partial k$. In the absence of strings, the axion energy density $\rho_a = \langle \dot{a}^2 \rangle$ is

$$\rho_a = \frac{1}{L^3}\int d^3x_p\ \dot{a}^2(x_p) = \frac{1}{L^3}\int \frac{d^3k}{(2\pi)^3}|\tilde{\dot{a}}(k)|^2\ , \tag{131}$$

where $x_p = R(t)x$ are physical coordinates, and $\tilde{\dot{a}}(k)$ is the Fourier transform of $\dot{a}(x_p)$. The axion spectrum $\partial\rho_a/\partial|k|$ is defined by $\int d|k|\ \partial\rho_a/\partial|k| = \rho_a$, and is therefore given by

$$\frac{\partial\rho_a}{\partial|k|} = \frac{|k|^2}{(2\pi L)^3}\int d\Omega_k|\tilde{\dot{a}}(k)|^2\ . \tag{132}$$

In the following, as in the main text, we use the notation $\partial\rho_a/\partial k \equiv \partial\rho_a/\partial|k|$.

In performing the integral $\int d\Omega_k$ in eq. (132), which converts the three-dimensional spectrum to the one-dimensional spectrum, the three-dimensional momenta $\vec{k} \equiv \frac{2\pi\vec{n}}{L}$ with $|\vec{n}| \leq \frac{N}{2}$ and $|\vec{n}| \in \left]m-\frac{1}{2}, m+\frac{1}{2}\right]$, $m \in \mathbb{N}$, are grouped into the same one-dimensional momentum bin labeled $\frac{2\pi m}{L}$. Momenta with $\frac{N}{2} < |\vec{n}| < \sqrt{3}\frac{N}{2}$ have not been included since they lie far above the string core scale and do not matter for either the spectrum or the number density.

In the presence of strings, eqs. (131) and (132) are no longer valid because of the contribution of the string cores to $\dot{a}(x_p)$ and so to $\tilde{\dot{a}}(k)$. We adopted two methods to mask strings out of the calculation of $\partial\rho_a/\partial k$:

1. The first is the Pseudo Power Spectrum Estimator (PPSE) introduced in the analysis of cosmic microwave background data [147], and first used in the context of cosmic strings in ref. [135].

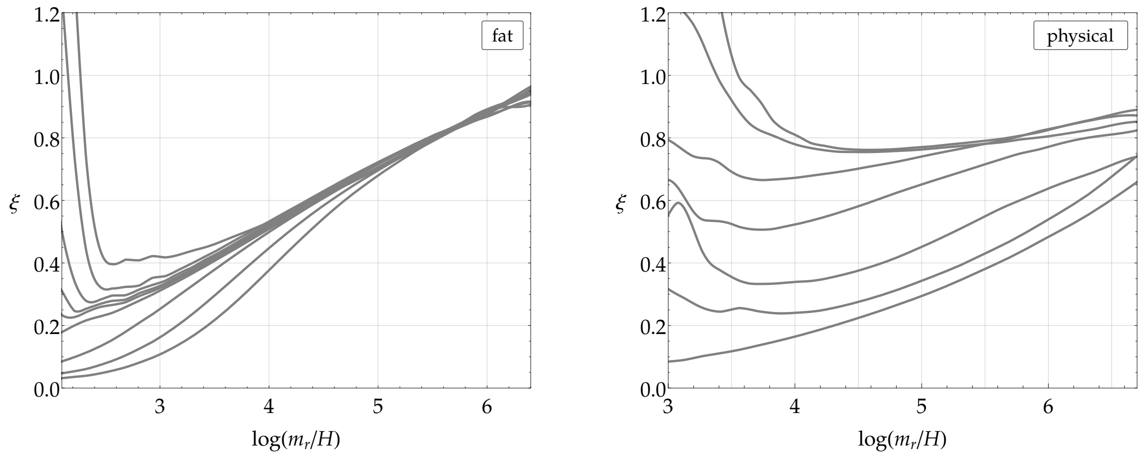

Figure 26: Convergence of $\xi(t)$ to the attractor solution starting from random initial conditions with different values of k_{max}/m_r, between 1/2 and 1/50, for the fat string (left) and physical (right) scenarios.

This method involves calculating the spectrum $\partial\rho_a/\partial k$ after removing the regions of space that are less than $d_s m_r^{-1}$ from the strings' center. Briefly, it is based on eq. (132) with $\tilde{\dot{a}}(k)$ taken to be the Fourier transform of $\theta(x_p)\dot{a}(x_p)$, where $\theta(x_p) = 1$ for x_p a distance of more than $d_s m_r^{-1}$ from the strings' center, and $\theta(x_p) = 0$ otherwise. $\partial\rho_a/\partial k$ is then appropriately corrected to account for the bias introduced by the window function $\theta(x_p)$ (more details may be found in [147]).

2. The second approach takes advantage of the automatic screening of string cores provided by the factor $(1 + r/f_a)$, and consists of using eq (132) with $\tilde{\dot{a}}(k)$ taken to be the Fourier transform of $(1 + r(x_p)/f_a)\dot{a}(x_p)$. This is similar to the previous method, except that the string cores do not have to be identified, and the degree of masking varies smoothly as the core is approached.

We have tested that the computation of the axion spectrum $\partial\rho_a/\partial k$ is independent of the method used, and in particular we studied the dependence of the first method on the screening distance d_s. In Figure 25 we plot the ratio of spectra at $\log = 6$ computed using the two methods for different values of the screening parameter d_s, including the case $d_s = 0$, where the strings are not screened. We see that the difference between the spectrum from method 1 with $d_s = 1$ and the one from method 2 is 5% at $k = m_r/2$ and rapidly decreases for smaller momenta. This is expected since, given the form of the string profile function for the radial mode, method 2 is roughly equivalent to masking distances up to m_r^{-1}. Meanwhile, increasing d_s suppresses the spectrum at momenta of order $2\pi/(d_s m_r^{-1})$, but leaves IR momenta unchanged. Again this is not surprising, since masking distances of order $d_s m_r^{-1}$ only loses information about momenta larger than $2\pi/(d_s m_r^{-1})$. The spectrum for $d_s = 0$, which is affected by the presence of strings, is much more UV dominated than the others, but at momenta lower than $m_r/2$ still does not differ by more than 10%.

In the main text we calculated ρ_a and $\partial\rho_a/\partial k$ using method 1 with $d_s = 1$. The resulting systematic uncertainty from the masking algorithm at $k \sim m_r/4$ is less than 1%, while for $k \sim m_r/2$ it can be estimated to be of order $5 \div 10\%$. As explained in Appendix F, we extract the power law of the spectrum between the IR and the UV peaks, q, by considering momenta smaller than $m_r/6$. As a result the systematic uncertainty on q introduced by the masking procedure is negligible. The fact

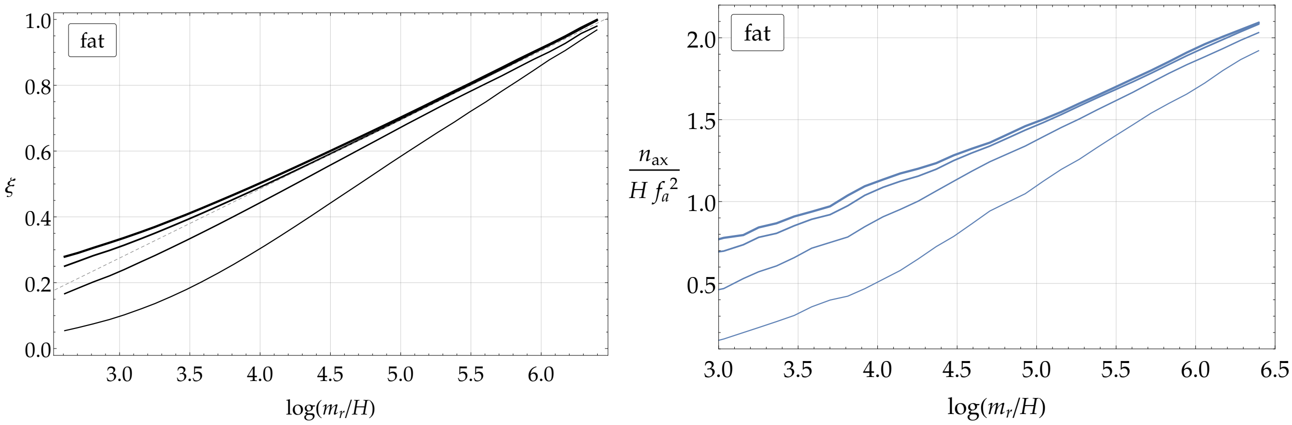

Figure 27: Convergence to the attractor for $\xi(t)$ (left) and n_a (right) for fat strings. Denser initial conditions are represented by thicker lines, and the results in both cases are obtained from the same sets of initial conditions, with fixed string densities.

that ρ_a and $\partial\rho_a/\partial k$ are stable under the change of algorithm, and that ρ_a redshifts as radiation, gives us further confidence that, even in presence of strings, our results only include free axion modes, as required.

F Convergence from Different Initial Conditions

In Section 4.3.1 we demonstrated the existence of the attractor solution by considering the convergence of the string density $\xi(t)$, starting from initial conditions with fixed values of ξ. Here we present further evidence for the attractor solution. First, rather than fixing initial conditions with a predetermined initial string density by the method (b) described in Section 4.3.1 and C.1, we use random initial conditions given by the method (a) varying the maximum coefficients k_{max}/m_r. Second, we show how other key properties of the string network, in particular the axion spectrum and axion number density, also converge to the same late time values starting with different string densities.

In Figure 26 we plot $\xi(t)$ starting from random field initial conditions for different choices of k_{max}/m_r at the initial time $H_i = m_r$. Although the results obtained are slightly less regular than those starting with fixed numbers of strings, shown in Figure 7, the convergence to the attractor, and the logarithmic increase, is clear.

To analyse the convergence of other properties of the string network, we present results starting from fixed string number densities, which lead to the values of $\xi(t)$ plotted in Figure 27 (left). We have also confirmed that the quantities that we study converge to the same attractor solution starting from random initial conditions. In Figure 27 (right) we show the results for the axion number density obtained from the same set of simulations as $\xi(t)$ in the left panel.

Notice that while at $\log = 3$ the values for ξ and n_a are spread respectively by a factor of 3 and 15, at $\log = 6$ the spread reduces to around 10%.

The convergence of the axion spectrum $\partial\rho_a/\partial k$ to the attractor, shown in Figure 28 (left), is also revealing. At $\log(m_r/H) = 3$, the axion spectrum is highly suppressed in simulations with a lower initial string density than in the others. However, by $\log(m_r/H) = 6$ the spectra are very similar. This is especially the case for the modes around Hubble, which have nearly the same amplitude for all

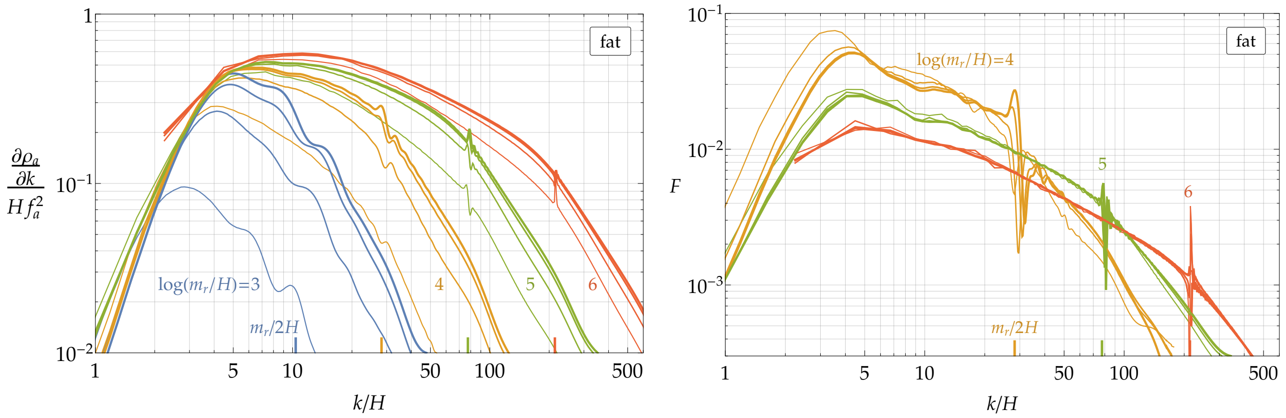

Figure 28: The convergence of the total axion spectrum (left) and the instantaneous emission spectrum (right), starting from the same set of initial conditions as Figure 27, in the fat string system. Simulations with more strings initially are represented by thicker lines, and spectra at the same time instant are represented by the same colour.

initial conditions. More dramatically, the instantaneous emission shown in Figure 28 (right) converges extremely fast at late times. This is because, unlike the overall spectrum, it only depends on the properties of the network at a fixed time instant, which are practically the same for all initial conditions by the end of the simulations. We see that underdense networks tend to have more instantaneous emission in the IR at early times, which is expected since at these times the typical distance between strings is larger in this case.

Taken all together, our results provide convincing evidence that the properties of the string scaling solution at late enough times are sufficiently insensitive to the initial conditions chosen.

G Extraction of the Scaling Parameters

In this appendix we describe how we extract the slope α in the fit of $\xi(t)$, eq. (78), and the power law q of the instantaneous spectrum, and estimate the associated uncertainties. We also describe how we obtain the extrapolations of the number density in Figure 17.

Given the convergence shown in Figure 7, initial conditions that result in a constant value of $d\xi/d\log(m_r/H) \equiv d\xi/d\log$ for the longest time correspond to string networks that are the closest to the scaling regime. Therefore, we estimate α by calculating $d\xi/d\log$ for initial conditions with different fixed initial string densities, and then restrict to those such that $d\xi/d\log$ is approximately constant at late times.[45] The constant common value that $d\xi/d\log$ reaches in such simulations is a good estimate of α, and the spread indicates the uncertainty.

In Figure 29 we plot $d\xi/d\log$ for different initial string densities, for both the fat string and the physical case. Slopes of different simulations tend to converge asymptotically at to a common constant value. This is evident in the fat case, for which the constant approached is $d\xi/d\log = 0.22 \pm 0.02$, where the error has been estimated from the spread of the results from simulations that have constant

---

[45] Since the derivative $d\xi/d\log$ is more sensitive to local fluctuation we smooth it by convoluting with a Gaussian $g(x)$ with $\sigma = 1/4$, i.e. $\int dy\ g(x-y)d\xi(y)/d\log$.

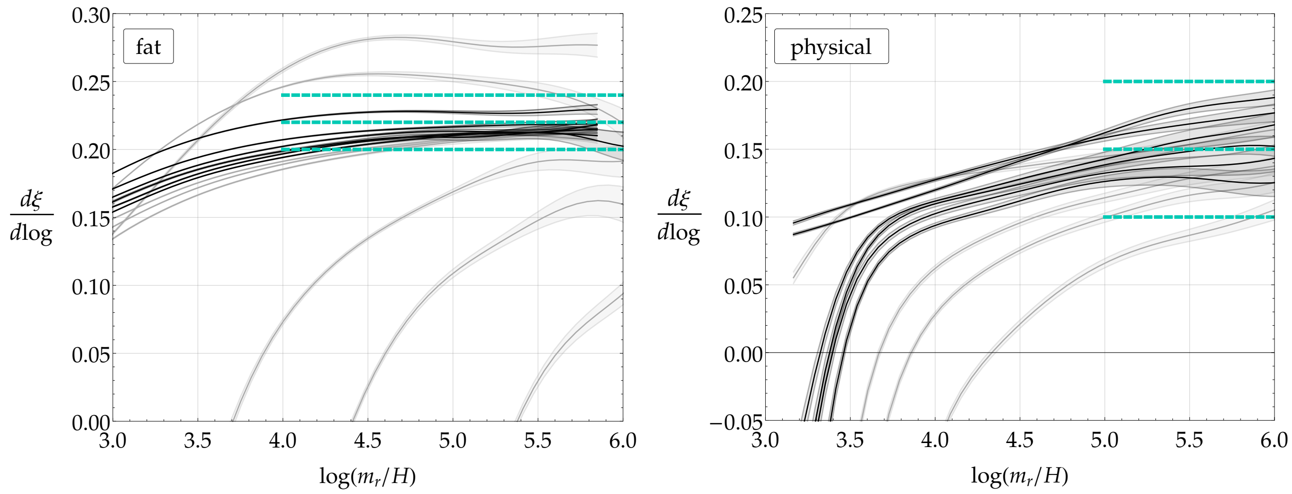

Figure 29: The derivative $d\xi/d\log(m_r/H)$ for different initial conditions. The black curves are approximately constant for the largest amount of time and therefore the closest to scaling, and their spread is used to estimate the error on α. Blue dashed lines represent the estimate of α, and are plotted at times when the system is reasonably close to scaling. Error bands represent statistical errors in the average over different samples.

$d\xi/d\log$ for $\log(m_r/H) \gtrsim 4$ (plotted in black). In the physical case, the scaling regime is reached at larger values of the log and $d\xi/d\log$ is changing over most of the simulated range. As a result, the uncertainty on $d\xi/d\log = 0.15 \pm 0.05$, estimated as the spread of $d\xi/d\log$ over the simulations for which $d\xi/d\log$ is constant for $\log(m_r/H) \gtrsim 5$, is larger.

We now turn to the extraction of the power law of the spectrum q. For this purpose, we consider simulations starting from the initial conditions that lead to string networks that are the closest to scaling (corresponding to results for $\xi(t)$ plotted in black in Figure 7).[46] As shown in Figure 15, for $\log(m_r/H) \gtrsim 5$ the instantaneous emission spectrum, parametrized by $F(k/H, m_r/H)$, has an approximate power law behavior $1/k^q$ for momenta that are large enough with respect to the peak at around the Hubble scale (at $k/H \sim 5 \div 10$) and are small compared to the UV cutoff at $k = m_r/2$.

The fact that q is less than 1 can be more clearly seen from Figure 30, in which we plot $xF(x, y)$. In these plots $q < 1$ corresponds to the increase between the IR and UV peaks, which is evident both in the fat string and physical cases. Moreover, most of the area under the curves in Figure 30 is at UV momenta $k \sim m_r/2$, which shows that the energy density is dominantly emitted in the form of very high momentum axions, although the axion number density is dominated by low momentum states since $q > 0$. In Figure 31 we plot the total energy spectrum of Figure 14 on a log scale, i.e. $\partial\rho_a/\partial \log k$. The positive gradients show that the energy density in axions is dominated by UV modes at all times, for both the fat string and the physical case.

Due to the challenges in the physical case, discussed in Section 4.4.3, we only attempt a detailed analysis of q in the fat string scenario. To do this, we consider F at late times, $\log(m_r/H) \gtrsim 5$, and focus on the region with momenta a factor of 3 larger than the Hubble peak and a factor of 3 smaller than the core peak, i.e. $30H \lesssim k \lesssim m_r/6$, which is sufficiently uncontaminated by the cutoffs.

---

[46] As shown in Appendix F, the results obtained starting from different initial conditions are very similar, well within the uncertainties on q that we quote.

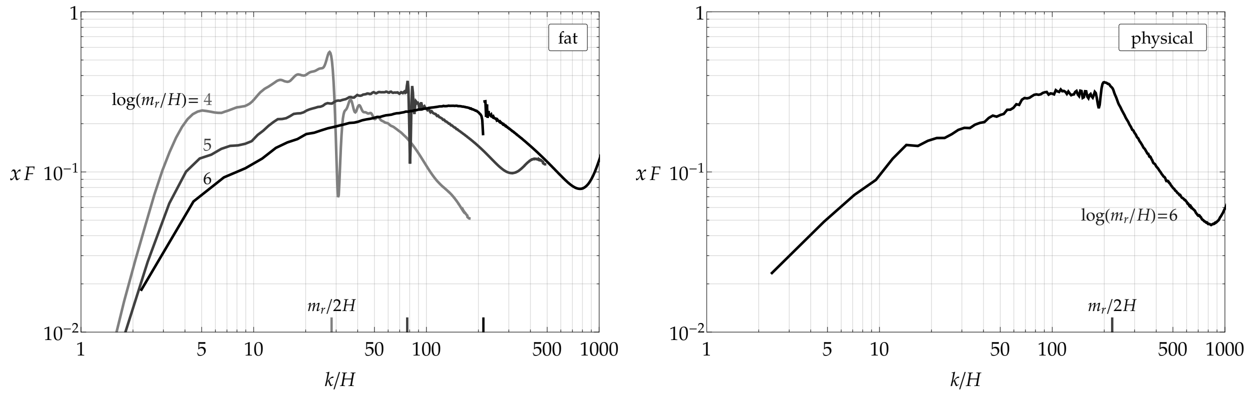

Figure 30: Plot of $xF(x, y)$ with $x = k/H$ for the fat string (left) and the physical (right) cases, for different values of $y = m_r/H$. The growth of $xF(x, y)$ for x between the IR and UV cutoffs corresponds to a UV dominated spectrum, i.e. $q < 1$.

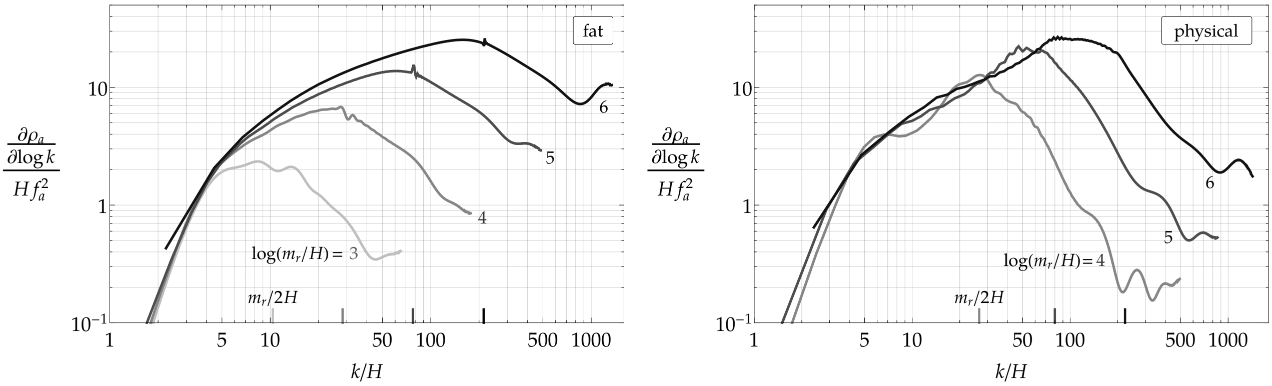

Figure 31: The axion energy density spectrum $\partial\rho_a/\partial \log k$ for the fat string (left) and the physical (right) cases (the data is the same as in Figure 14 but represented on a log scale).

The power q is then given by $-d\log F(x, y)/d\log(x)$.[47] In Figure 32 we plot $d\log F(x, y)/d\log(x)$ for different values of $y = m_r/H$, corresponding to values of $\log(m_r/H) > 5$. The sections of the curves plotted in black indicate momenta in the range $30H < k < m_r/6$, which are safe from contamination from the IR and UV peaks.

$d\log F/d\log(x)$ crosses zero at about $k \sim 5H$, signalling the position of the Hubble scale peak, and it is smaller than -1 at large momenta, meaning that after the UV cutoff the instantaneous spectrum falls off steeply. Moreover, in the intermediate region of interest $-d\log F/d\log(x)$ changes slowly as a function of momentum and across different time shots, ranging from about 0.7 to 0.8. We note that in this range $-d\log F/d\log(x)$ (and therefore q) shows a tendency to increase at later times. However, given the present uncertainties this apparent change is not significant enough to draw any conclusion, in particular we are not able to assess whether it is due to a residual contamination from the nearby UV peak or it constitutes a genuine feature of the spectrum (similar analysis with larger grids and

---

[47] Because the derivative $d\log F(x, y)/d\log(x)$ is subject to more fluctuations we smooth it by convoluting with a Gaussian $g(x)$ with $\sigma = 1/6$, i.e. we consider $\int d\log(z)\ g(\log(x) - \log(z))\ d\log F(z, y)/d\log(z)$.

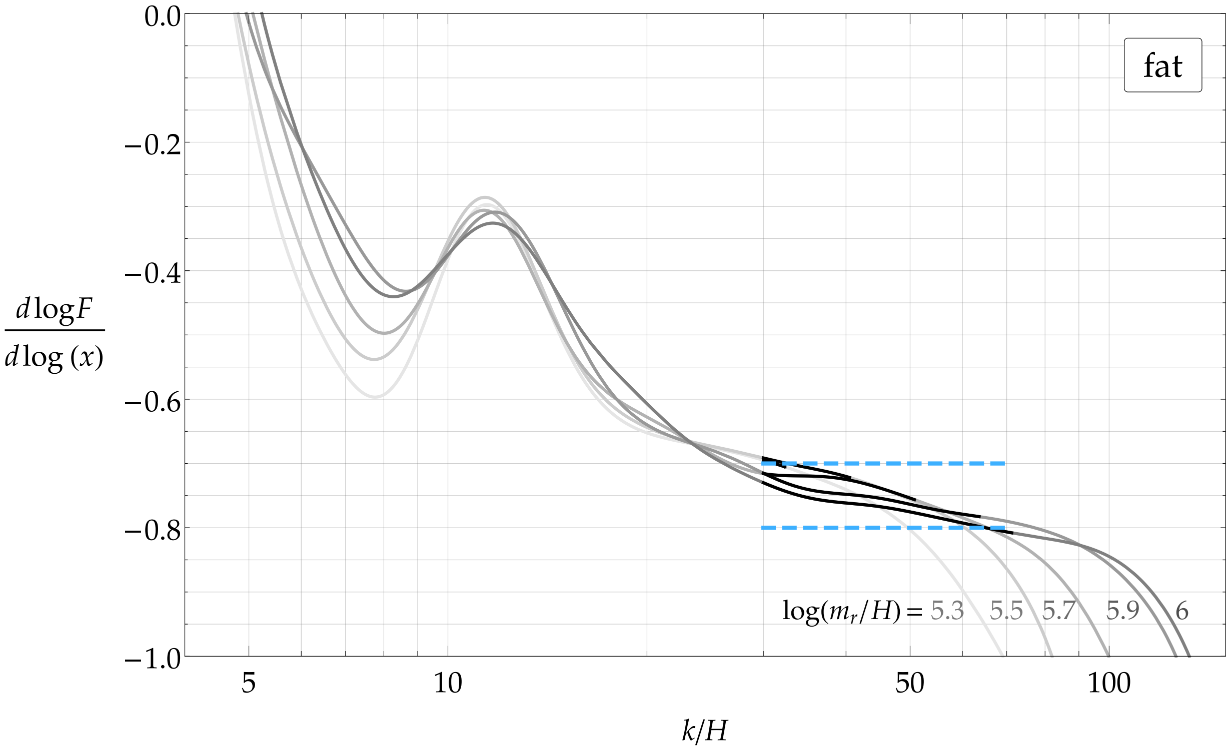

Figure 32: The value of $d\log F(x,y)/d\log x$ as a function of the rescaled momentum $x = k/H$, for different time shots, i.e. different values of $y = m_r/H$, for the fat string system in the scaling regime. The black sections corresponds to momenta $30H < k < m_r/6$, which are sufficiently far from the IR and UV peaks to be uncontaminated by effects at these scales. The roughly constant values in these regions correspond to the value of $-q$ in the approximate power law, and the blue dashed lines represent our estimated range for this.

more statistics could be beneficial to understand this point). In the regime $5 \lesssim \log(m_r/H) \lesssim 6$, we therefore estimate $q = 0.75 \pm 0.05$, where the uncertainly mostly comes from $d\log F/d\log(x)$ not being constant over the momentum range of interest.

In order to extrapolate the axion number density with different power laws, we used an analytic form of $F(x,y)$ that closely matches the three approximate power laws visible in Figure 15,

$$F(x,y) = N\frac{\left(\frac{x}{x_1}\right)^{q_1}\left[1+\theta(x-x_2)\left(\left(\frac{x_2}{x}\right)^{q_2-q}-1\right)\right]}{\left(\frac{x}{x_1}\right)^{q_1+q}+1} \propto \begin{cases} x^{q_1} & x \ll x_1 \\ \frac{1}{x^q} & x_1 \ll x \ll x_2 \\ \frac{1}{x^{q_2}} & x > x_2\,, \end{cases} \tag{133}$$

where $\theta(x)$ is the step function. In this expression x_1 and x_2 are the positions of the IR and UV cutoffs in units of Hubble, q_1, q_2 are the powers that suppress the spectrum in the IR and in the UV respectively, and q is the power between the two cutoffs. The normalization N is required so that $\int F(x,y)dx = 1$. A reasonable fit for the 5 free parameters is $x_1 \sim 3, x_2 = y/2$ and $q = 0.7 \div 0.8$, $q_1 \sim 3$, $q_2 \sim 2$ in the fat string case, and similarly in the physical case except that $q_2 \sim 3$. The extrapolations in Figure 17 are carried out by keeping all parameters fixed except for q. What matters most for the final extrapolation is the parameter q, which regulates the hardness of the spectrum, while the remaining parameters only lead to changes of order 1.

H Comparison with EFT Estimates

As mentioned in Section 4.4, the dynamics of global string loops can be equivalently described by an effective theory where the fundamental degrees of freedom are the string and the axion radiation [38], with an interaction governed by the Kalb–Ramond action [148]. This theory is valid in the regime where the string and the emitted radiation (with frequency $\omega \sim 1/R$) are not strongly coupled, which corresponds to $\log(m_r R) \gg 1$, where R is the loop radius (as shown in [38], the effective coupling of the emitted axion radiation to the string is proportional to $1/\log(m_r R)$). As a result, for loops with a large hierarchy between the radius and the core size, the emitted axion radiation should approximate the one predicted by the Nambu–Goto effective theory. Indeed, using such a theory it was shown [38] that a circular loop starting at rest with $\log(m_r R_0) = 100$ follows the cosine time-law for the Nambu–Goto strings with percent precision, at least for values of the loop radius such that $\log(m_r R) \gg 1$, where the theory is applicable.

We show now that the evolution of circular loops provided by the solutions of the field equations matches the one predicted by the effective theory of strings, for the values of $\log(m_r R_0)$ reachable in our field theoretic simulations. We solve eq. (70) in Minkowski space, i.e. with a time independent scale factor, $H = 0$, and initial conditions $\phi(x)$ and $\dot{\phi}(x)$ that approximately resemble a static circular loop with initial radius R_0.[48] In Fig. 33 we plot the time-law for the loop radius $R(t)$ normalized to the initial radius R_0 for $\log(m_r R_0) = 4$ and 5. We also plot the prediction given by the effective theory for $\log(m_r R_0) = 5$ and the free Nambu–Goto time law, $R_{\rm NG}(t) = R_0 \cos(t/R_0)$. The result of the simulation for $\log(m_r R_0) = 5$ matches very well the EFT prediction where this is valid. Moreover, as $\log(m_r R_0)$ increases, the circular loop time law gets closer to the free Nambu–Goto prediction, indeed indicating that at large $\log(m_r R_0)$ global string dynamics converge to that of free Nambu–Goto strings.

Although the convergence $R(t) \to R_{\rm NG}(t)$ is good when the loop radius is sufficiently larger than the core size m_r^{-1}, i.e. for $\log(m_r R) \gg 1$, there is a substantial deviation when the string becomes strongly coupled, $\log(m_r R) \lesssim 2$, which prevents the loop from bouncing many times. This however is not in conflict with the EFT prediction, which correctly reproduces the right time law before the loops collapses even for $\log(m_r R_0)$ as small as 5.

Given the limitations in performing direct field theoretic simulations at much larger values of $\log(m_r R_0)$ we cannot test whether the EFT expectation that loops will bounce many times (thus emitting an IR dominated spectrum) is correctly reproduced or whether core effects when the loops shrink keep inhibiting the rebounce.

---

[48] More precisely, $\dot{\phi}(x) = 0$, while $\phi(x)$ is chosen to be cylindrically symmetric around the z-axis and in the $y = 0$ plane is given by the field generated by the superposition of two point like charges with charge ± 1 in the position $(\pm R_0/2, 0, 0)$. The field generated by a point-like charge ± 1 in the origin is provided by eq. (71) (with phase $e^{\pm i\theta}$ resp.) and the superposition of fields is defined by their product.

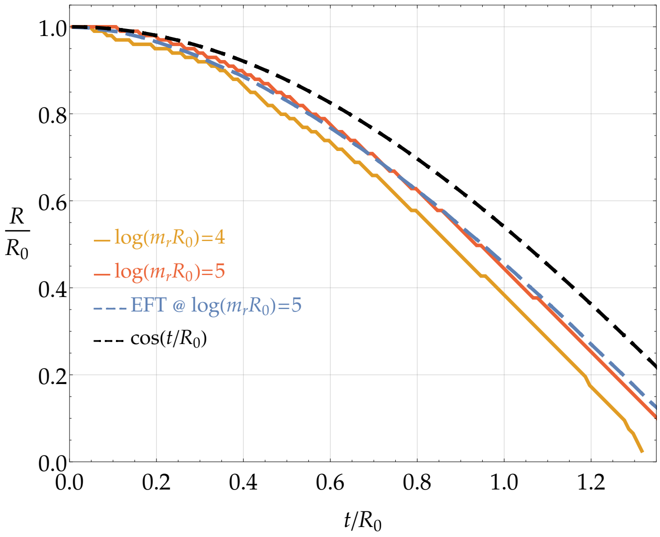

Figure 33: Circular loop radius time-law $R(t)$, normalised to the initial radius R_0. Green and orange dashed lines correspond to the solution of the field equations for $\log(m_r R_0) = 4, 5$. The solid blue line is the effective field theory prediction for $\log(m_r R_0) = 5$ and the solid black line is the free Nambu–Goto solution $R_{\rm NG}(t) = R_0 \cos(t/R_0)$.

References

[1] R. D. Peccei and H. R. Quinn, “CP Conservation in the Presence of Instantons,” Phys. Rev. Lett. **38** (1977) 1440.

[2] F. Wilczek, “Problem of Strong p and t Invariance in the Presence of Instantons,” Phys. Rev. Lett. **40** (1978) 279.

[3] S. Weinberg, “A New Light Boson?,” Phys. Rev. Lett. **40** (1978) 223.

[4] J. E. Kim, “Weak Interaction Singlet and Strong CP Invariance,” Phys. Rev. Lett. **43** (1979) 103.

[5] 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.

[6] A. R. Zhitnitsky, “On Possible Suppression of the Axion Hadron Interactions. (In Russian),” Sov. J. Nucl. Phys. **31** (1980) 260 [Yad. Fiz. **31** (1980) 497].

[7] M. Dine, W. Fischler and M. Srednicki, “A Simple Solution to the Strong CP Problem with a Harmless Axion,” Phys. Lett. B **104** (1981) 199.

[8] J. Preskill, M. B. Wise and F. Wilczek, “Cosmology of the Invisible Axion,” Phys. Lett. B **120** (1983) 127.

[9] L. F. Abbott and P. Sikivie, “A Cosmological Bound on the Invisible Axion,” Phys. Lett. B **120** (1983) 133.

[10] M. Dine and W. Fischler, "The Not So Harmless Axion," Phys. Lett. B **120** (1983) 137.

[11] P. W. Graham, I. G. Irastorza, S. K. Lamoreaux, A. Lindner and K. A. van Bibber, "Experimental Searches for the Axion and Axion-Like Particles," Ann. Rev. Nucl. Part. Sci. **65** (2015) 485 doi:10.1146/annurev-nucl-102014-022120 [arXiv:1602.00039 [hep-ex]].

[12] I. G. Irastorza and J. Redondo, "New experimental approaches in the search for axion-like particles," arXiv:1801.08127 [hep-ph].

[13] A. Arvanitaki, S. Dimopoulos, S. Dubovsky, N. Kaloper and J. March-Russell, "String Axiverse," Phys. Rev. D **81** (2010) 123530 doi:10.1103/PhysRevD.81.123530 [arXiv:0905.4720 [hep-th]].

[14] S. Weinberg, "Phenomenological Lagrangians," Physica A **96** (1979) no.1-2, 327. doi:10.1016/0378-4371(79)90223-1

[15] J. Gasser and H. Leutwyler, "Chiral Perturbation Theory to One Loop," Annals Phys. **158** (1984) 142. doi:10.1016/0003-4916(84)90242-2

[16] J. Gasser and H. Leutwyler, "Chiral Perturbation Theory: Expansions in the Mass of the Strange Quark," Nucl. Phys. B **250** (1985) 465. doi:10.1016/0550-3213(85)90492-4

[17] P. Sikivie, "Axion Cosmology," Lect. Notes Phys. **741** (2008) 19 [astro-ph/0610440].

[18] D. J. E. Marsh, "Axion Cosmology," Phys. Rept. **643** (2016) 1 doi:10.1016/j.physrep.2016.06.005 [arXiv:1510.07633 [astro-ph.CO]].

[19] D. H. Lyth, Phys. Lett. B **236** (1990) 408. doi:10.1016/0370-2693(90)90374-F

[20] M. S. Turner and F. Wilczek, "Inflationary axion cosmology," Phys. Rev. Lett. **66** (1991) 5. doi:10.1103/PhysRevLett.66.5

[21] A. D. Linde, "Axions in inflationary cosmology," Phys. Lett. B **259** (1991) 38. doi:10.1016/0370-2693(91)90130-I

[22] A. Vilenkin and A. E. Everett, "Cosmic Strings and Domain Walls in Models with Goldstone and PseudoGoldstone Bosons," Phys. Rev. Lett. **48** (1982) 1867.

[23] P. Sikivie, "Of Axions, Domain Walls and the early universe," Phys. Rev. Lett. **48** (1982) 1156.

[24] A. Vilenkin, "Cosmic Strings and Domain Walls," Phys. Rept. **121** (1985) 263.

[25] R. L. Davis, "Goldstone Bosons in String Models of Galaxy Formation," Phys. Rev. D **32** (1985) 3172. doi:10.1103/PhysRevD.32.3172

[26] R. L. Davis, "Cosmic Axions from Cosmic Strings," Phys. Lett. B **180** (1986) 225.

[27] T. W. B. Kibble, "Topology of Cosmic Domains and Strings," J. Phys. A **9** (1976) 1387. doi:10.1088/0305-4470/9/8/029

[28] T. W. B. Kibble, "Some Implications of a Cosmological Phase Transition," Phys. Rept. **67** (1980) 183. doi:10.1016/0370-1573(80)90091-5

[29] A. Vilenkin, "Cosmic Strings," Phys. Rev. D **24** (1981) 2082. doi:10.1103/PhysRevD.24.2082

[30] A. Albrecht and N. Turok, “Evolution of Cosmic Strings,” Phys. Rev. Lett. **54** (1985) 1868. doi:10.1103/PhysRevLett.54.1868

[31] D. P. Bennett and F. R. Bouchet, “Evidence for a Scaling Solution in Cosmic String Evolution,” Phys. Rev. Lett. **60** (1988) 257.

[32] B. Allen and E. P. S. Shellard, “Cosmic string evolution: a numerical simulation,” Phys. Rev. Lett. **64** (1990) 119. doi:10.1103/PhysRevLett.64.119

[33] G. R. Vincent, M. Hindmarsh and M. Sakellariadou, “Scaling and small scale structure in cosmic string networks,” Phys. Rev. D **56** (1997) 637 [astro-ph/9612135].

[34] C. J. A. P. Martins and E. P. S. Shellard, “Extending the velocity dependent one scale string evolution model,” Phys. Rev. D **65** (2002) 043514 doi:10.1103/PhysRevD.65.043514 [hep-ph/0003298].

[35] A. Vilenkin and E. P. S. Shellard, “Cosmic Strings and Other Topological Defects,”

[36] A. Vilenkin and T. Vachaspati, Phys. Rev. D **35** (1987) 1138. doi:10.1103/PhysRevD.35.1138

[37] R. L. Davis and E. P. S. Shellard, “Do Axions Need Inflation?,” Nucl. Phys. B **324** (1989) 167. doi:10.1016/0550-3213(89)90187-9

[38] A. Dabholkar and J. M. Quashnock, “Pinning Down the Axion,” Nucl. Phys. B **333** (1990) 815. doi:10.1016/0550-3213(90)90140-9

[39] R. A. Battye and E. P. S. Shellard, “Global string radiation,” Nucl. Phys. B **423** (1994) 260 doi:10.1016/0550-3213(94)90573-8 [astro-ph/9311017].

[40] R. A. Battye and E. P. S. Shellard, “Axion string constraints,” Phys. Rev. Lett. **73** (1994) 2954 Erratum: [Phys. Rev. Lett. **76** (1996) 2203] doi:10.1103/PhysRevLett.73.2954, 10.1103/PhysRevLett.76.2203 [astro-ph/9403018].

[41] M. Yamaguchi, M. Kawasaki and J. Yokoyama, “Evolution of axionic strings and spectrum of axions radiated from them,” Phys. Rev. Lett. **82** (1999) 4578 doi:10.1103/PhysRevLett.82.4578 [hep-ph/9811311].

[42] M. Yamaguchi, “Scaling property of the global string in the radiation dominated universe,” Phys. Rev. D **60** (1999) 103511 doi:10.1103/PhysRevD.60.103511 [hep-ph/9907506].

[43] C. Hagmann, S. Chang and P. Sikivie, “Axion radiation from strings,” Phys. Rev. D **63** (2001) 125018 doi:10.1103/PhysRevD.63.125018 [hep-ph/0012361].

[44] H. Georgi and M. B. Wise, “Hiding the Invisible Axion,” Phys. Lett. **116B** (1982) 123. doi:10.1016/0370-2693(82)90989-3

[45] S. Chang, C. Hagmann and P. Sikivie, “Studies of the motion and decay of axion walls bounded by strings,” Phys. Rev. D **59** (1999) 023505 doi:10.1103/PhysRevD.59.023505 [hep-ph/9807374].

[46] Y. B. Zeldovich, I. Y. Kobzarev and L. B. Okun, “Cosmological Consequences of the Spontaneous Breakdown of Discrete Symmetry,” Zh. Eksp. Teor. Fiz. **67** (1974) 3 [Sov. Phys. JETP **40** (1974) 1].

[47] G. B. Gelmini, M. Gleiser and E. W. Kolb, “Cosmology of Biased Discrete Symmetry Breaking,” Phys. Rev. D **39** (1989) 1558. doi:10.1103/PhysRevD.39.1558

[48] S. E. Larsson, S. Sarkar and P. L. White, “Evading the cosmological domain wall problem,” Phys. Rev. D **55** (1997) 5129 doi:10.1103/PhysRevD.55.5129 [hep-ph/9608319].

[49] T. Hiramatsu, M. Kawasaki, K. Saikawa and T. Sekiguchi, “Axion cosmology with long-lived domain walls,” JCAP **1301** (2013) 001 doi:10.1088/1475-7516/2013/01/001 [arXiv:1207.3166 [hep-ph]].

[50] L. Di Luzio, E. Nardi and L. Ubaldi, “Accidental Peccei-Quinn symmetry protected to arbitrary order,” Phys. Rev. Lett. **119** (2017) no.1, 011801 doi:10.1103/PhysRevLett.119.011801 [arXiv:1704.01122 [hep-ph]].

[51] C. Hagmann and P. Sikivie, “Computer simulations of the motion and decay of global strings,” Nucl. Phys. B **363** (1991) 247. doi:10.1016/0550-3213(91)90243-Q

[52] W. A. Bardeen, “Anomalous Currents in Gauge Field Theories,” Nucl. Phys. B **75** (1974) 246. doi:10.1016/0550-3213(74)90546-X

[53] C. Vafa and E. Witten, “Parity Conservation in QCD,” Phys. Rev. Lett. **53** (1984) 535. doi:10.1103/PhysRevLett.53.535

[54] G. ’t Hooft, “Computation of the Quantum Effects Due to a Four-Dimensional Pseudoparticle,” Phys. Rev. D **14** (1976) 3432 Erratum: [Phys. Rev. D **18** (1978) 2199]. doi:10.1103/PhysRevD.18.2199.3, 10.1103/PhysRevD.14.3432

[55] G. ’t Hooft, “Symmetry Breaking Through Bell-Jackiw Anomalies,” Phys. Rev. Lett. **37** (1976) 8. doi:10.1103/PhysRevLett.37.8

[56] A. I. Vainshtein, V. I. Zakharov, V. A. Novikov and M. A. Shifman, “ABC’s of Instantons,” Sov. Phys. Usp. **25** (1982) 195 [Usp. Fiz. Nauk **136** (1982) 553]. doi:10.1070/PU1982v025n04ABEH004533

[57] V. Baluni, “CP Violating Effects in QCD,” Phys. Rev. D **19** (1979) 2227. doi:10.1103/PhysRevD.19.2227

[58] R. J. Crewther, P. Di Vecchia, G. Veneziano and E. Witten, “Chiral Estimate of the Electric Dipole Moment of the Neutron in Quantum Chromodynamics,” Phys. Lett. **88B** (1979) 123 Erratum: [Phys. Lett. **91B** (1980) 487]. doi:10.1016/0370-2693(80)91025-4, 10.1016/0370-2693(79)90128-X

[59] J. M. Pendlebury *et al.*, “Revised experimental upper limit on the electric dipole moment of the neutron,” Phys. Rev. D **92** (2015) no.9, 092003 doi:10.1103/PhysRevD.92.092003 [arXiv:1509.04411 [hep-ex]].

[60] J. R. Ellis and M. K. Gaillard, “Strong and Weak CP Violation,” Nucl. Phys. B **150** (1979) 141. doi:10.1016/0550-3213(79)90297-9

[61] I. B. Khriplovich and A. I. Vainshtein, “Infinite renormalization of Theta term and Jarlskog invariant for CP violation,” Nucl. Phys. B **414** (1994) 27 doi:10.1016/0550-3213(94)90419-7 [hep-ph/9308334].

[62] G. ‘t Hooft, “Recent Developments in Gauge Theories” (Plenum Press, 1980) 135

[63] J. E. Kim, “A Composite Invisible Axion,” Phys. Rev. D **31** (1985) 1733. doi:10.1103/PhysRevD.31.1733

[64] K. Choi and J. E. Kim, “Dynamical Axion,” Phys. Rev. D **32** (1985) 1828. doi:10.1103/PhysRevD.32.1828

[65] H. Georgi, D. B. Kaplan and L. Randall, “Manifesting the Invisible Axion at Low-energies,” Phys. Lett. **169B** (1986) 73. doi:10.1016/0370-2693(86)90688-X

[66] G. G. Raffelt, “Astrophysical axion bounds,” Lect. Notes Phys. **741** (2008) 51 [hep-ph/0611350].

[67] A. Arvanitaki and S. Dubovsky, “Exploring the String Axiverse with Precision Black Hole Physics,” Phys. Rev. D **83** (2011) 044026 doi:10.1103/PhysRevD.83.044026 [arXiv:1004.3558 [hep-th]].

[68] S. J. Asztalos *et al.* [ADMX Collaboration], “A SQUID-based microwave cavity search for dark-matter axions,” Phys. Rev. Lett. **104** (2010) 041301 doi:10.1103/PhysRevLett.104.041301 [arXiv:0910.5914 [astro-ph.CO]].

[69] R. D. Peccei and H. R. Quinn, “Constraints Imposed by CP Conservation in the Presence of Instantons,” Phys. Rev. D **16** (1977) 1791. doi:10.1103/PhysRevD.16.1791

[70] S. R. Coleman, J. Wess and B. Zumino, Phys. Rev. **177** (1969) 2239. doi:10.1103/PhysRev.177.2239

[71] C. G. Callan, Jr., S. R. Coleman, J. Wess and B. Zumino, “Structure of phenomenological Lagrangians. 2.,” Phys. Rev. **177** (1969) 2247. doi:10.1103/PhysRev.177.2247

[72] P. Di Vecchia and G. Veneziano, “Chiral Dynamics in the Large n Limit,” Nucl. Phys. B **171** (1980) 253. doi:10.1016/0550-3213(80)90370-3

[73] S. Chang and K. Choi, “Hadronic axion window and the big bang nucleosynthesis,” Phys. Lett. B **316** (1993) 51 doi:10.1016/0370-2693(93)90656-3 [hep-ph/9306216].

[74] J. Kodaira, “QCD Higher Order Effects in Polarized Electroproduction: Flavor Singlet Coefficient Functions,” Nucl. Phys. B **165** (1980) 129. doi:10.1016/0550-3213(80)90310-7

[75] S. A. Larin, “The Renormalization of the axial anomaly in dimensional regularization,” Phys. Lett. B **303** (1993) 113 doi:10.1016/0370-2693(93)90053-K [hep-ph/9302240].

[76] M. I. Buchoff *et al.*, “QCD chiral transition, U(1)A symmetry and the dirac spectrum using domain wall fermions,” Phys. Rev. D **89** (2014) no.5, 054514 doi:10.1103/PhysRevD.89.054514 [arXiv:1309.4149 [hep-lat]].

[77] A. Trunin, F. Burger, E. M. Ilgenfritz, M. P. Lombardo and M. Müller-Preussker, “Topological susceptibility from $N_f = 2 + 1 + 1$ lattice QCD at nonzero temperature,” J. Phys. Conf. Ser. **668** (2016) no.1, 012123 doi:10.1088/1742-6596/668/1/012123 [arXiv:1510.02265 [hep-lat]].

[78] D. J. Gross, R. D. Pisarski and L. G. Yaffe, “QCD and Instantons at Finite Temperature,” Rev. Mod. Phys. **53** (1981) 43. doi:10.1103/RevModPhys.53.43

[79] M. S. Turner, “Thermal Production of Not SO Invisible Axions in the Early Universe,” Phys. Rev. Lett. **59** (1987) 2489 Erratum: [Phys. Rev. Lett. **60** (1988) 1101]. doi:10.1103/PhysRevLett.59.2489, 10.1103/PhysRevLett.60.1101.3

[80] E. Masso, F. Rota and G. Zsembinszki, “On axion thermalization in the early universe,” Phys. Rev. D **66** (2002) 023004 doi:10.1103/PhysRevD.66.023004 [hep-ph/0203221].

[81] P. Graf and F. D. Steffen, “Thermal axion production in the primordial quark-gluon plasma,” Phys. Rev. D **83** (2011) 075011 doi:10.1103/PhysRevD.83.075011 [arXiv:1008.4528 [hep-ph]].

[82] M. Archidiacono, S. Hannestad, A. Mirizzi, G. Raffelt and Y. Y. Y. Wong, “Axion hot dark matter bounds after Planck,” JCAP **1310** (2013) 020 doi:10.1088/1475-7516/2013/10/020 [arXiv:1307.0615 [astro-ph.CO]].

[83] P. Arias, D. Cadamuro, M. Goodsell, J. Jaeckel, J. Redondo and A. Ringwald, “WISPy Cold Dark Matter,” JCAP **1206** (2012) 013 doi:10.1088/1475-7516/2012/06/013 [arXiv:1201.5902 [hep-ph]].

[84] P. A. R. Ade *et al.* [Planck Collaboration], “Planck 2015 results. XIII. Cosmological parameters,” Astron. Astrophys. **594** (2016) A13 doi:10.1051/0004-6361/201525830 [arXiv:1502.01589 [astro-ph.CO]].

[85] D. H. Lyth and E. D. Stewart, “Constraining the inflationary energy scale from axion cosmology,” Phys. Lett. B **283** (1992) 189. doi:10.1016/0370-2693(92)90006-P

[86] J. Hamann, S. Hannestad, G. G. Raffelt and Y. Y. Y. Wong, “Isocurvature forecast in the anthropic axion window,” JCAP **0906** (2009) 022 doi:10.1088/1475-7516/2009/06/022 [arXiv:0904.0647 [hep-ph]].

[87] S. Weinberg, “Gauge and Global Symmetries at High Temperature,” Phys. Rev. D **9** (1974) 3357. doi:10.1103/PhysRevD.9.3357

[88] R. N. Mohapatra and G. Senjanovic, “Soft CP Violation at High Temperature,” Phys. Rev. Lett. **42** (1979) 1651. doi:10.1103/PhysRevLett.42.1651

[89] D. A. Kirzhnits and A. D. Linde, “Macroscopic Consequences of the Weinberg Model,” Phys. Lett. **42B** (1972) 471. doi:10.1016/0370-2693(72)90109-8

[90] E. Berkowitz, M. I. Buchoff and E. Rinaldi, “Lattice QCD input for axion cosmology,” Phys. Rev. D **92** (2015) no.3, 034507 doi:10.1103/PhysRevD.92.034507 [arXiv:1505.07455 [hep-ph]].

[91] S. Borsanyi *et al.*, “Axion cosmology, lattice QCD and the dilute instanton gas,” Phys. Lett. B **752** (2016) 175 doi:10.1016/j.physletb.2015.11.020 [arXiv:1508.06917 [hep-lat]].

[92] P. Petreczky, H. P. Schadler and S. Sharma, “The topological susceptibility in finite temperature QCD and axion cosmology,” Phys. Lett. B **762** (2016) 498 doi:10.1016/j.physletb.2016.09.063 [arXiv:1606.03145 [hep-lat]].

[93] S. Borsanyi *et al.*, “Calculation of the axion mass based on high-temperature lattice quantum chromodynamics,” Nature **539** (2016) no.7627, 69 doi:10.1038/nature20115 [arXiv:1606.07494 [hep-lat]].

[94] F. Burger, E. M. Ilgenfritz, M. P. Lombardo and A. Trunin, "Chiral observables and topology in hot QCD with two families of quarks," Phys. Rev. D **98** (2018) no.9, 094501 doi:10.1103/PhysRevD.98.094501 [arXiv:1805.06001 [hep-lat]].

[95] C. Bonati, M. D'Elia, G. Martinelli, F. Negro, F. Sanfilippo and A. Todaro, "Topology in full QCD at high temperature: a multicanonical approach," arXiv:1807.07954 [hep-lat].

[96] M. Tanabashi *et al.* [Particle Data Group], "Review of Particle Physics," Phys. Rev. D **98** (2018) no.3, 030001. doi:10.1103/PhysRevD.98.030001

[97] G. Grilli di Cortona, E. Hardy, J. Pardo Vega and G. Villadoro, "The QCD axion, precisely," JHEP **1601** (2016) 034 doi:10.1007/JHEP01(2016)034 [arXiv:1511.02867 [hep-ph]].

[98] C. Bonati, M. D'Elia, M. Mariti, G. Martinelli, M. Mesiti, F. Negro, F. Sanfilippo and G. Villadoro, "Axion phenomenology and θ-dependence from $N_f = 2 + 1$ lattice QCD," JHEP **1603** (2016) 155 doi:10.1007/JHEP03(2016)155 [arXiv:1512.06746 [hep-lat]].

[99] Z. Fodor *et al.*, "Up and down quark masses and corrections to Dashen's theorem from lattice QCD and quenched QED," Phys. Rev. Lett. **117** (2016) no.8, 082001 doi:10.1103/PhysRevLett.117.082001 [arXiv:1604.07112 [hep-lat]].

[100] D. Giusti, V. Lubicz, C. Tarantino, G. Martinelli, S. Sanfilippo, S. Simula and N. Tantalo, "Leading isospin-breaking corrections to pion, kaon and charmed-meson masses with Twisted-Mass fermions," Phys. Rev. D **95** (2017) no.11, 114504 doi:10.1103/PhysRevD.95.114504 [arXiv:1704.06561 [hep-lat]].

[101] S. Basak *et al.* [MILC Collaboration], "Lattice computation of the electromagnetic contributions to kaon and pion masses," arXiv:1807.05556 [hep-lat].

[102] N. Carrasco *et al.* [European Twisted Mass Collaboration], "Up, down, strange and charm quark masses with N_f = 2+1+1 twisted mass lattice QCD," Nucl. Phys. B **887** (2014) 19 doi:10.1016/j.nuclphysb.2014.07.025 [arXiv:1403.4504 [hep-lat]].

[103] S. Aoki *et al.*, "Review of lattice results concerning low-energy particle physics," Eur. Phys. J. C **77** (2017) no.2, 112 doi:10.1140/epjc/s10052-016-4509-7 [arXiv:1607.00299 [hep-lat]].

[104] P. A. Boyle *et al.*, "Low energy constants of SU(2) partially quenched chiral perturbation theory from N_f=2+1 domain wall QCD," Phys. Rev. D **93** (2016) no.5, 054502 doi:10.1103/PhysRevD.93.054502 [arXiv:1511.01950 [hep-lat]].

[105] J. Gasser and H. Leutwyler, "Quark Masses," Phys. Rept. **87** (1982) 77. doi:10.1016/0370-1573(82)90035-7

[106] M. Knecht and R. Urech, "Virtual photons in low-energy pi pi scattering," Nucl. Phys. B **519** (1998) 329 doi:10.1016/S0550-3213(98)00044-3 [hep-ph/9709348].

[107] C. Haefeli, M. A. Ivanov and M. Schmid, "Electromagnetic low-energy constants in chiPT," Eur. Phys. J. C **53** (2008) 549 doi:10.1140/epjc/s10052-007-0493-2 [arXiv:0710.5432 [hep-ph]].

[108] B. Ananthanarayan and B. Moussallam, "Four-point correlator constraints on electromagnetic chiral parameters and resonance effective Lagrangians," JHEP **0406** (2004) 047 doi:10.1088/1126-6708/2004/06/047 [hep-ph/0405206].

[109] J. Bijnens and J. Prades, "Electromagnetic corrections for pions and kaons: Masses and polarizabilities," Nucl. Phys. B **490** (1997) 239 doi:10.1016/S0550-3213(97)00107-7 [hep-ph/9610360].

[110] V. Cirigliano and I. Rosell, "pi/K → e anti-nu(e) branching ratios to O(e^2p^4) in Chiral Perturbation Theory," JHEP **0710** (2007) 005 doi:10.1088/1126-6708/2007/10/005 [arXiv:0707.4464 [hep-ph]].

[111] S. Descotes-Genon and B. Moussallam, "Radiative corrections in weak semi-leptonic processes at low energy: A Two-step matching determination," Eur. Phys. J. C **42** (2005) 403 doi:10.1140/epjc/s2005-02316-8 [hep-ph/0505077].

[112] R. Urech, "Virtual photons in chiral perturbation theory," Nucl. Phys. B **433** (1995) 234 doi:10.1016/0550-3213(95)90707-N [hep-ph/9405341].

[113] M. Knecht, H. Neufeld, H. Rupertsberger and P. Talavera, "Chiral perturbation theory with virtual photons and leptons," Eur. Phys. J. C **12** (2000) 469 doi:10.1007/s100529900265 [hep-ph/9909284].

[114] V. Lubicz, G. Martinelli, C. T. Sachrajda, F. Sanfilippo, S. Simula, N. Tantalo and C. Tarantino, "Electromagnetic corrections to the leptonic decay rates of charged pseudoscalar mesons: lattice results," PoS LATTICE **2016** (2016) 290 doi:10.22323/1.256.0290 [arXiv:1610.09668 [hep-lat]].

[115] D. Giusti, V. Lubicz, G. Martinelli, C. T. Sachrajda, F. Sanfilippo, S. Simula, N. Tantalo and C. Tarantino, "First lattice calculation of the QED corrections to leptonic decay rates," Phys. Rev. Lett. **120** (2018) no.7, 072001 doi:10.1103/PhysRevLett.120.072001 [arXiv:1711.06537 [hep-lat]].

[116] J. Bijnens, G. Colangelo and G. Ecker, "The Mesonic chiral Lagrangian of order p^6," JHEP **9902** (1999) 020 doi:10.1088/1126-6708/1999/02/020 [hep-ph/9902437].

[117] J. Bijnens, G. Colangelo and G. Ecker, "Renormalization of chiral perturbation theory to order p^6," Annals Phys. **280** (2000) 100 doi:10.1006/aphy.1999.5982 [hep-ph/9907333].

[118] J. Bijnens, G. Colangelo and P. Talavera, "The Vector and scalar form-factors of the pion to two loops," JHEP **9805** (1998) 014 doi:10.1088/1126-6708/1998/05/014 [hep-ph/9805389].

[119] J. F. Donoghue, J. Gasser and H. Leutwyler, "The Decay of a Light Higgs Boson," Nucl. Phys. B **343** (1990) 341. doi:10.1016/0550-3213(90)90474-R

[120] I. Caprini, "Model-independent constraint on the pion scalar form factor and light quark masses," Phys. Rev. D **98** (2018) no.5, 056008 doi:10.1103/PhysRevD.98.056008 [arXiv:1803.04150 [hep-ph]].

[121] J. Bijnens and G. Ecker, "Mesonic low-energy constants," Ann. Rev. Nucl. Part. Sci. **64** (2014) 149 doi:10.1146/annurev-nucl-102313-025528 [arXiv:1405.6488 [hep-ph]].

[122] B. Ananthanarayan, J. Bijnens and S. Ghosh, "An Analytic Analysis of the Pion Decay Constant in Three-Flavoured Chiral Perturbation Theory," Eur. Phys. J. C **77** (2017) no.7, 497 doi:10.1140/epjc/s10052-017-5019-y [arXiv:1703.00141 [hep-ph]].

[123] G. Amoros, J. Bijnens and P. Talavera, "Two point functions at two loops in three flavor chiral perturbation theory," Nucl. Phys. B **568** (2000) 319 doi:10.1016/S0550-3213(99)00674-4 [hep-ph/9907264].

[124] U. Burgi, “Charged pion pair production and pion polarizabilities to two loops,” Nucl. Phys. B **479** (1996) 392 doi:10.1016/0550-3213(96)00454-3 [hep-ph/9602429].

[125] J. Bijnens, G. Colangelo, G. Ecker, J. Gasser and M. E. Sainio, “Pion-pion scattering at low energy,” Nucl. Phys. B **508** (1997) 263 Erratum: [Nucl. Phys. B **517** (1998) 639] doi:10.1016/S0550-3213(97)80013-2, 10.1016/S0550-3213(97)00621-4, 10.1016/S0550-3213(98)00127-8 [hep-ph/9707291].

[126] G. Colangelo, J. Gasser and H. Leutwyler, “$\pi\pi$ scattering,” Nucl. Phys. B **603** (2001) 125 doi:10.1016/S0550-3213(01)00147-X [hep-ph/0103088].

[127] J. Bijnens and N. Hermansson Truedsson, “The Pion Mass and Decay Constant at Three Loops in Two-Flavour Chiral Perturbation Theory,” JHEP **1711** (2017) 181 doi:10.1007/JHEP11(2017)181 [arXiv:1710.01901 [hep-ph]].

[128] Y. Y. Mao *et al.* [TWQCD Collaboration], “Topological Susceptibility to the One-Loop Order in Chiral Perturbation Theory,” Phys. Rev. D **80** (2009) 034502 doi:10.1103/PhysRevD.80.034502 [arXiv:0903.2146 [hep-lat]].

[129] A. I. Davydychev and J. B. Tausk, “Two loop selfenergy diagrams with different masses and the momentum expansion,” Nucl. Phys. B **397** (1993) 123. doi:10.1016/0550-3213(93)90338-P

[130] R. Kaiser, “On the two-loop contributions to the pion mass,” JHEP **0709** (2007) 065 doi:10.1088/1126-6708/2007/09/065 [arXiv:0707.2277 [hep-ph]].

[131] O. V. Tarasov, “Connection between Feynman integrals having different values of the space-time dimension,” Phys. Rev. D **54** (1996) 6479 doi:10.1103/PhysRevD.54.6479 [hep-th/9606018].

[132] O. V. Tarasov, “Generalized recurrence relations for two loop propagator integrals with arbitrary masses,” Nucl. Phys. B **502** (1997) 455 doi:10.1016/S0550-3213(97)00376-3 [hep-ph/9703319].

[133] L. Visinelli and P. Gondolo, “Axion cold dark matter in view of BICEP2 results,” Phys. Rev. Lett. **113** (2014) 011802 doi:10.1103/PhysRevLett.113.011802 [arXiv:1403.4594 [hep-ph]].

[134] D. Harari and P. Sikivie, “On the Evolution of Global Strings in the Early Universe,” Phys. Lett. B **195** (1987) 361. doi:10.1016/0370-2693(87)90032-3

[135] T. Hiramatsu, M. Kawasaki, T. Sekiguchi, M. Yamaguchi and J. Yokoyama, “Improved estimation of radiated axions from cosmological axionic strings,” Phys. Rev. D **83** (2011) 123531 doi:10.1103/PhysRevD.83.123531 [arXiv:1012.5502 [hep-ph]].

[136] V. B. Klaer and G. D. Moore, “How to simulate global cosmic strings with large string tension,” JCAP **1710** (2017) 043 doi:10.1088/1475-7516/2017/10/043 [arXiv:1707.05566 [hep-ph]].

[137] M. Yamaguchi, J. Yokoyama and M. Kawasaki, “Numerical analysis of formation and evolution of global strings in (2+1)-dimensions,” Prog. Theor. Phys. **100** (1998) 535 doi:10.1143/PTP.100.535 [hep-ph/9808326].

[138] L. Fleury and G. D. Moore, “Axion dark matter: strings and their cores,” JCAP **1601** (2016) 004 doi:10.1088/1475-7516/2016/01/004 [arXiv:1509.00026 [hep-ph]].

[139] M. Hindmarsh, J. Lizarraga, J. Urrestilla, D. Daverio and M. Kunz, "Scaling from gauge and scalar radiation in Abelian Higgs string networks," Phys. Rev. D **96** (2017) no.2, 023525 doi:10.1103/PhysRevD.96.023525 [arXiv:1703.06696 [astro-ph.CO]].

[140] M. Hindmarsh, S. Stuckey and N. Bevis, "Abelian Higgs Cosmic Strings: Small Scale Structure and Loops," Phys. Rev. D **79** (2009) 123504 doi:10.1103/PhysRevD.79.123504 [arXiv:0812.1929 [hep-th]].

[141] C. Hagmann, S. Chang and P. Sikivie, "Axions from string decay," Nucl. Phys. Proc. Suppl. **72** (1999) 81 doi:10.1016/S0920-5632(98)00506-4 [hep-ph/9807428].

[142] V. B. Klaer and G. D. .Moore, "The dark-matter axion mass," JCAP **1711** (2017) no.11, 049 doi:10.1088/1475-7516/2017/11/049 [arXiv:1708.07521 [hep-ph]].

[143] M. Gorghetto, E. Hardy and G. Villadoro, in preparation.

[144] K. Clough, P. Figueras, H. Finkel, M. Kunesch, E. A. Lim and S. Tunyasuvunakool, "GRChombo : Numerical Relativity with Adaptive Mesh Refinement," Class. Quant. Grav. **32** (2015) no.24, 245011 [Class. Quant. Grav. **32** (2015) 24] doi:10.1088/0264-9381/32/24/245011 [arXiv:1503.03436 [gr-qc]].

[145] L. M. Fleury and G. D. Moore, "Axion String Dynamics I: 2+1D," JCAP **1605** (2016) no.05, 005 doi:10.1088/1475-7516/2016/05/005 [arXiv:1602.04818 [hep-ph]].

[146] R. J. LeVeque, "Finite Difference Methods for Ordinary and Partial Differential Equations," Cambridge University Press

[147] G. Hinshaw *et al.* [WMAP Collaboration], "First year Wilkinson Microwave Anisotropy Probe (WMAP) observations: The Angular power spectrum," Astrophys. J. Suppl. **148** (2003) 135 doi:10.1086/377225 [astro-ph/0302217].

[148] R. L. Davis and E. P. S. Shellard, "Antisymmetric Tensors and Spontaneous Symmetry Breaking," Phys. Lett. B **214**, 219 (1988). doi:10.1016/0370-2693(88)91472-4