RICE UNIVERSITY

Cosmological and astrophysical probes of axionlike particles

by

Ray Mitchell Hagimoto

A Thesis Submitted
in Partial Fulfillment of the
Requirements for the Degree

Doctor of Philosophy

Approved, Thesis Committee:

Andrew J. Long, Co-chair
Assistant Professor of Physics and Astronomy

Mustafa A. Amin, Co-chair
Associate Professor of Physics and Astronomy

Arlei Da Silva
Assistant Professor of Computer Science

Houston, Texas

December, 2024

ABSTRACT

Cosmological and astrophysical probes of axionlike particles

by

Ray Mitchell Hagimoto

Axionlike particles (ALPs), pseudo Nambu-Goldstone bosons generated by the spontaneous breaking of global $U(1)_{\rm PQ}$ symmetries arise in solutions to open issues in fundamental physics and are ubiquitous in string theory compactifications. ALPs have a rich phenomenology which provides numerous ways to search for evidence of their existence. This work explores two potential discovery channels for ALPs. The first considers the possibility that hyperlight ALPs ($m_a \lesssim 10^{-28}\,\text{eV}$) with a Chern-Simons coupling to electromagnetism formed a cosmic string network in the early Universe that survives beyond recombination. In this scenario, CMB photons passing through string loops in the network experience a rotation in their plane of polarization, an effect known as CMB birefringence. Detecting CMB birefringence from ALP strings may be in reach of future CMB probes. I use existing CMB birefringence power spectrum data to constrain axion string network parameters. Next, I discuss non-Gaussian features of axion string induced CMB birefringence maps. Finally, I explore how one could use a neural network to estimate axion string network parameters from CMB birefringence maps. The second potential discovery channel considers how ALPs with lepton flavor violating couplings and masses $m_a \lesssim 1\,\text{MeV}$ affect the cooling rates of neutron stars. Through these studies, I develop tools that would assist in identifying signatures of ALPs in cosmological and astrophysical observations.

Acknowledgments

I decided to sign up for IB higher level physics in 10th grade because my brother, Ken, had taken it and told me it was difficult. This was very much in line with my tradition of wanting to copy almost everything he did. For example, when I was in infant school ($\sim$ elementary school) I wanted to be in the recorder club because Ken had been when he was my age. Unfortunately, I signed up too late so there were no remaining spots. That news brought me to tears. However, following in Ken's footsteps wasn't always a bad thing. Taking HL physics was one of the best decisions I ever made. It is the reason I found something I am passionate about and is why I met so many incredible people in the last 8 years. Between that, and buying me a Minecraft account I have much to thank you for.

My high school physics class was taught by Mr. Cesar Reyes. His fantastic, organised classes gave me a strong foundation as well as a love and appreciation for physics. That experience was made even more enjoyable due to the close camaraderie I had with my classmates. Vash Raman, Ruslan Kozmenko, Anirudd Roy, Raghu Sharma, Vasuki Kaushik were particularly close. (Vasuki, I still have the copy of Newton's Principia you gifted me in 2015).

After high school, I studied physics at the University of Texas at San Antonio where many faculty members left deep impressions on me. Professors Reza Aghayan, Eduardo Dueñez, Nicolas Large, Kathryn Mayer, and Rafael Lopez-Mobilila stand out.

My interactions with Rafael are especially treasured memories. At the end of my first year as an undergrad I cut my Summer vacation short to fly back to San Antonio from Singapore so I could audit Rafael's GR course. Looking back it's funny to think

that things like curvilinear coordinates and the geodesic equation were beyond the scope of my comprehension but I still picked up a lot of useful conceptual lessons by showing up to the lectures and attending office hours. That Summer, we would sometimes talk for as long as 3 hours. We discussed things like muon lifetimes and the twin paradox, but it was also in these conversations that Rafael introduced me to my favourite topic – Lagrangian mechanics. Learning about the principle of least action and how conservation laws are a consequence of continuous symmetries made me so excited to take classical mechanics the following year. When I TA'd freshman honours mechanics a few years ago I made sure to share Noether's theorem with my students with the hope that they would feel the same way I did when I learned it from you. Throughout my time at UTSA, Rafael shared many of his insights on physics with me. Whether it was funny stories about famous physicists, or technical explanations of topics in QFT and philosophy of physics, I always walked away from Rafael's office more knowledgeable and eager to learn.

My next two Summers as an undergraduate student were similarly influential. I participated in NSF REUs and was lucky enough to have incredible mentors. In Summer 2018 I studied the quantum dynamics of a qubit coupled to a harmonic oscillator under the supervision of Professors Jean-Francois Van Huele and Manuel Berrondo at BYU. Their enthusiastic lectures about time evolution in quantum mechanics, the Schrödinger picture versus the Heisenberg picture, and differential equations were very stimulating and enjoyable. In Summer 2019 I visited UChicago where Professors Stephan S. Meyer and Craig Hogan introduced me to cosmology. Together, Craig, Steve, me, and another REU student, Collin Lewin, investigated a peculiar feature of the CMB temperature two-point correlation function which culminated in my first publication in an academic journal. These experiences showed me how exciting research could be and were invaluable preparation for grad school.

A few months after I left Chicago, Craig and Steve invited me and Collin to attend the "Cosmic Controversies" conference. Steve kindly let me and Collin stay at

his house. At Cosmic Controversies I met many notable cosmologists and saw several outstanding presentations. I was thoroughly impressed by one in particular about oscillons after inflation by a professor from Rice University. As a result, I looked for an excuse to visit Rice, which is how I found out about the Gulf Coast Undergraduate Research Symposium (GCURS). At GCURS I gave a talk about my UChicago REU project where I met my (future) PhD advisors Mustafa Amin ("Mr. Oscillon") and Andrew Long for the first time.

Mustafa and Andrew are gifted educators and researchers with an inspiring work ethic. I will forever be indebted to them for their immense patience, superior organisation skills, and for teaching me almost everything I know about doing productive research. Upon completion of my first paper, Andrew gifted me a copy of Daniel Baumann's cosmology textbook which is one of the best presents I have ever received. On nights where I stayed late at the office Mustafa's Trader Joe's meals kept me sane. Beyond their role as advisors, Mustafa and Andrew are wonderful people.

I had many stimulating discussions with my group members Hong Yi Zhang, Siyang Ling, Mudit Jain, Dorian Amaral, Moira Venegas, Dashon Jones, Wisha Wanichwecharungruang, and Jonathan Thomas for which I am grateful. Also, my friends in the XENON experiment, Ivy Li, Luis Sanchez, Shixiao Liang, Juehang Qin.

I am especially indebted to Siyang for teaching me so much about mathematics, computer science, and more. There was one time I couldn't figure out how to interpret a formula in a review article about CMB non-Gaussianity. I discussed it with Siyang and he said the formula was a direct consequence of the definition of the Wigner 3-j symbol. I still didn't get it. The next day Siyang told me to come to the library with him and proceeded to lecture me about group theory and representations until I had all the prerequisite knowledge to be able to say, with a deep level of understanding, "Ah, yes. The formula is a direct consequence of the definition." Even the knowledge you shared with me about data structures and algorithms helped me land a Summer

internship. My time as a PhD student was significantly enriched because of you.

During my time at Rice I made many great friends, too many to list comprehensively. Aashish Kafle taught me how to make chai, Kevin Allen coached my bench press from 115 lbs to 185 lbs (in particle physics units that's around 4.6×10^{28} GeV), and Eva Jin inducted me to the church of Central Market. Umaima Memon shared lots of homemade macarons with me, including my favourite early grey flavour and introduced me to K-dramas. Claire, my best memories are hanging out with you and Miwa. Many of the friends I made before moving to Houston continued to be there for me. Thank you Keanu Dias, Arjun Bhalla, Alexander Wallis, Marielle Gaspar, Javier Rufino, Diego Ramos, Rudin Kraja, Aiden Harbick, Daniel Ruan, Dan Swartz, and Jon Kutasov. I love you all.

Towards the end of my PhD I also had a chance to meet some incredibly talented students in Philadelphia. Especially Georgy Gaitsgori, Akshat Bhardwaj, Kevin Sackel, Joshua Kazdan, Sajjad Nezhadi, Seungjae Son, Nicholas Tsao, Ansh Nagwekar, and Kai Wong. Learning about financial markets with you all was awesome. Our exciting conversations about life, trading, and aspirations were refreshing and encouraging. I am proud that I had an opportunity to work alongside you.

My family has selflessly supported me in pursuing my PhD. My Mum and Dad always encouraged me to follow my interests, and made sure I never had to worry about anything but school. My relatives George, Misty, Carolyn, and David made Texas feel like home. What you have given me is priceless and I wish I could pay it back. You would tell me to pay it forward instead.

Contents

Illustrations

Tables

Chapter 1

Introduction

1.1 Overview

This thesis explores the phenomenology of axionlike particles (ALPs) with a focus on a phenomenological consequence of ALP field configurations featuring topological defects known as cosmic strings. In chapter 1 (this chapter) I provide historical background and motivation for the works shown in chapters 2-5. More precisely, in sec. 1.2 I review the historical context for axions and ALPs, and in sec. 1.3 I introduce string defects and an effect called cosmic birefringence.

In chapter 2 I use existing measurements of CMB birefringence power spectra to constrain ALP-string parameters in a phenomenological model called the loop-crossing model (LCM). In chapter 3 I study the non-Gaussian signatures of LCM birefringence maps via their kurtosis and bispectrum. In chapter 4 I explore how neural networks can estimate LCM parameters from CMB birefringence maps. Finally, in chapter 5, I consider ALPs with lepton-flavour violating (LFV) couplings produced in neutron stars and use neutron star cooling data to put bounds on the LFV coupling strength.

1.2 Axions and axionlike particles

In this section I will provide historical context and motivation for the study of axions and axionlike particles. The Standard Model (SM) of particle physics is one of, if not, the most successful accomplishments in humanity's attempt to understand the fundamental laws of nature. The most striking example of the SM's success is its

prediction of the electron's anomalous magnetic moment, which has been verified to a staggering 12 significant figures (1) – a monumental agreement between theory and experiment. In spite of its successes, there are several areas in which the SM is limited. For example, the SM fails to explain the origin of neutrino masses (2), the nature of dark matter and dark energy (3), and the observed lack of Charge-Parity (CP) violation in the strong sector (4). The problem that is most relevant in the context of axions is the latter, known as "the strong CP problem", which will be explored in more detail below.

1.2.1 The Strong CP problem and the axion

The strong force is fundamentally described by the theory of quantum chromodynamics (QCD). The symmetries of the SM permit the inclusion of a CP-violating term in the QCD Lagrangian. It has the form (5)

$$\mathcal{L} \supset \frac{\theta_{\mathrm{QCD}}}{32\pi^2}\, G^a_{\mu\nu}\tilde{G}^{a,\mu\nu} \ , \tag{1.1}$$

where $G_{\mu\nu}$ is the gluon field strength tensor, $\tilde{G}^{\mu\nu} = \epsilon^{\mu\nu\alpha\beta}G_{\alpha\beta}/2$ is its dual, and the latin index a runs over the 8 generators of the QCD gauge group $SU(3)_c$. One implication of (1.1) is that it generates a electric dipole moment (EDM) for the neutron with a magnitude directly proportional to θ_{QCD} given by (6),

$$d_n = \frac{g_{\pi NN}\overline{g}_{\pi NN}\ln\left(M_N/m_\pi\right)}{4\pi^2 M_N} \approx 3.6\times 10^{-16}\,\theta_{\mathrm{QCD}}\,e\,\mathrm{cm} \ . \tag{1.2}$$

The latest experimental measurements of the neutron's EDM constrain $d_n < 1.8 \times 10^{-26}\,e\,\mathrm{cm}$ at 90% confidence level (7). In conjunction with eq. (1.2) this imposes a limit of $\theta_{\mathrm{QCD}} \lesssim 10^{-10}$, which is staggeringly small. There is no theoretical reason why the term in eq. (1.1) should be omitted, or why it should have such a small value, which hints that there may physics beyond the SM at play.

1.2.2 The Axion Solution

One of the most compelling solutions to the strong CP problem is to promote $\theta_{\rm QCD}$ from a parameter of the model to a dynamical CP-conserving field $\theta_{\rm QCD} \to \phi(x)$ in such a way that ϕ settles to 0 via its equations of motion. This idea was first introduced by Roberto Peccei and Helen Quinn in 1977 (8) and is known as the Peccei-Quinn (PQ) mechanism. While there are many variations on the PQ mechanism the basic ingredients are as follows (5). A new complex scalar field φ is added to the theory such that the full theory has a global $U(1)_{\rm PQ}$ symmetry. The φ field has a potential

$$V(\varphi) = \lambda\left(|\varphi|^2 - \frac{f_a^2}{2}\right)^2 , \tag{1.3}$$

where f_a is the axion decay constant, also known as the PQ scale. Moreover, φ couples to fermions in the theory via Yukawa terms. For energies below the PQ scale the ground state of the theory spontaneously breaks the global $U(1)_{\rm PQ}$ symmetry, giving masses to the fermions it is coupled to. The resultant (pseudo) Nambu-Goldstone boson ϕ is the angular degree of freedom of φ. The "pseudo-ness" of ϕ arises from instanton effects which generate a mass term for ϕ that explicitly breaks the PQ symmetry. In addition to generating the axion mass, instanton effects have two important consequences. The first is that they induce a "Chern-Simons" coupling between axions and gluons, adding the topological term

$$(\mathcal{C}\phi/f_a)G\tilde{G} \tag{1.4}$$

to the Lagrangian, where $\mathcal{C}$ is the quantum anomaly coefficient (see e.g. ref. (9)). Note that eq. (1.4) is of the same form as eq. (1.1) so there is now an effective term $(\theta_{\rm QCD} + \mathcal{C}\phi/f_a)G\tilde{G}$. But since the axion has a shift symmetry $\phi \to \phi + \text{const}$, the constant $\theta_{\rm QCD}$ can be absorbed into the definition of ϕ. This accomplishes the goal of replacing $\theta_{\rm QCD}$ with the dynamical field ϕ. The second important role played by

instantons is they generate a potential for the axion field of the form

$$V(\phi) \propto 1 - \cos(\mathcal{C}\phi/f_a) \ . \tag{1.5}$$

The equations of motion cause the axion to dynamically settle to a value which minimizes its potential energy eq. (1.5), which implies $\phi \to 0$. Hence the CP violating term vanishes and the strong CP problem is resolved.

1.2.3 Axionlike particles

In sec. 1.2 we introduced axions in the context of solving the strong CP problem. However, pseudoscalar fields with axion-like properties, known as axionlike particles (ALPs), arise naturally in other areas. Most notably, string theory compactifications of extra dimensions give rise to ALPs as Kaluza-Klein zero modes of antisymmetric tensor fields (10). In many compactifications there are a plethora of ALPs – hundreds, even hundreds of thousands – with masses populating many decades of energy scales down to the present Hubble scale $H_0 \sim 10^{-33}$ eV (11). See ref. (12) for an example in type IIB string theory. Throughout this thesis I will be primarily interested in ALPs with masses in the range * 10^{-33} eV $m_a \lesssim 10^{-28}$ eV.

Another general property of string theory ALPs is that they have Chern-Simons couplings to gauge fields. In my discussion of the PQ solution to the strong CP problem (sec. 1.2.1) instanton effects generated a coupling between axions and gluons as in eq. (1.4). However, it is also possible for axions to couple to photons in an analogous way as (9; 11; 15)

$$\mathcal{L}_{\phi\gamma\gamma} = \frac{\mathcal{A}\alpha_{\rm em}}{4\pi f_a}\,\phi\, F_{\mu\nu}\tilde{F}^{\mu\nu} \ , \tag{1.6}$$

*Note that ALPs with masses less than 10^{-19} eV are excluded from being a dominant component of the dark matter from observations of small-scale structures (13; 14). Therefore, the ALPs we consider in chapters 2-4 can not dominate the dark matter content.

where $F_{\mu\nu}$ is the electromagnetic field strength tensor, $\tilde{F}^{\mu\nu} = \epsilon^{\mu\nu\alpha\beta}F_{\alpha\beta}/2$ is its dual, $\alpha_{\rm em} \approx 1/137$ is the fine structure constant, and $\mathcal{A}$ is the electromagnetic anomaly coefficient. The axion-photon coupling in eq. (1.6) will play a central role in chapters 2, 3, and 4.

1.3 Cosmic axion strings and CMB birefringence

In this section I introduce one-dimensional topological defects called called strings. I briefly review how networks of strings can form in the ALP field in the early Universe via the Kibble mechanism and explore how ALP strings rather than particles provide a fascinating way to probe high energy physics through the topological ALP-photon interaction eq. (1.6). The review here will be brief, for more details the reader is directed to refs. (16; 17).

1.3.1 String defects

The potential for the complex scalar φ introduced in eq. (1.3) has an interesting consequence for certain field configurations wherein the angular field $\phi/f_a \equiv \arg\varphi$ (the axion) wraps around the circular minima at least once on a closed path in physical space. An example of such a solution is illustrated in fig. 1.1 where the colored plane shows the axion field wrapping around from $\phi = 0$ to $\phi = 2\pi f_a$. At the center of the plane is a 'vortex', a single point where ϕ must apparently wrap around from 0 to $2\pi f_a$ in order for the field to be smooth. Since this isn't possible ϕ is not well defined at that point. Instead, what must happen is that the radial component of φ goes to 0 at the vortex's center. This costs energy as it requires φ to climb up the potential's hill as seen in fig. 1.2.

Immediately next to this vortex must be another vortex where the field winds around in the same direction but possibly in a slightly different plane. By following the centers of the vortices one can trace out a one-dimensional structure in space called a string. This is depicted in fig. 1.1 by a black curve. Cosmic strings are

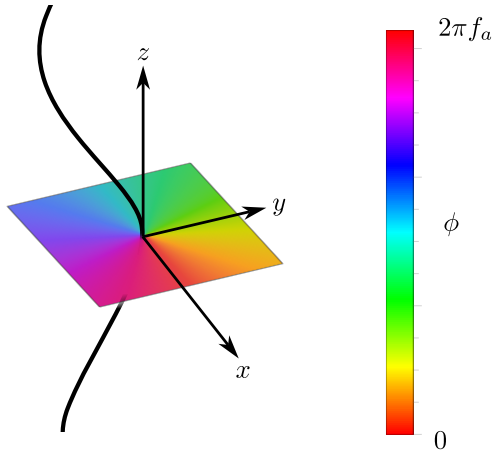


Figure 1.1 : Illustration of a string solution for $\phi(x, y, z)$ in physical space. The string core is represented by a black curve. A colored plane intersecting the string depicts the value of the axion field ϕ in the region around the string core. The axion winds from $\phi = 0$ on the x-axis to $\phi = 2\pi f_a$ for closed paths going counter-clockwise.

strings that form from cosmological phase transitions, a mechanism I review briefly in sec. 1.3.2.

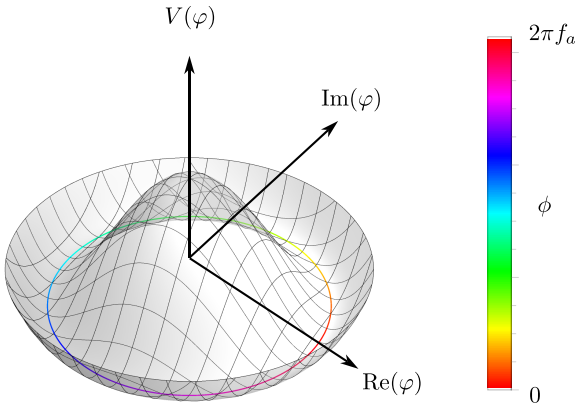


Figure 1.2 : Plot of the potential eq. (1.3) in field space. The field is parameterized by its real and imaginary components, $\mathrm{Re}(\varphi)$ and $\mathrm{Im}(\varphi)$, with the potential $V(\varphi) = \lambda\big(|\varphi|^2 - f_a^2/2\big)^2$. The value of the axion field in the minimum of the potential is depicted as a colored circle.

1.3.2 The Kibble mechanism and cosmic string formation

In this section I provide an overview of how cosmic strings form from cosmological phase transitions. A mechanism for the formation of topological defect networks due to cosmological phase transitions was first introduced in ref. (16) and it is known as the Kibble mechanism. My discussion is largely based on the presentation in ref. (17). In short, due to the Universe's expansion and state of thermal equilibrium, the potential eq. (1.3) has a temperature dependence. Above a critical temperature T_c the ground state of the theory is symmetric but below T_c it spontaneously breaks the global $U(1)_{\rm PQ}$ symmetry. Hence, in a Universe with an initial temperature above T_c, crossing below T_c results in a second order phase transition which can give rise to a network of topological defects such as strings. Below, I discuss the temperature-dependent potential in more detail and provide a simple example.

Observations of the CMB provide compelling evidence that the early Universe was in thermal equilibrium (18). We must therefore consider the effect that the presence of a thermal background has on potentials like eq. (1.3). The *effective potential* for weakly interacting fields at high temperature was derived perturbatively in (19–22), with the lowest order corrections being of the form

$$V_{\rm eff}(\varphi) = V(\varphi) - \mathcal{N}\frac{\pi^2}{90}\,T^4 + \frac{1}{24}\mathcal{M}^2(\varphi)\,T^2 + \mathcal{O}(T)\ , \tag{1.7}$$

where $V(\varphi)$ is the zero-temperature potential eq. (1.3), $\mathcal{N}$ controls the number of particles species, and $\mathcal{M}^2(\varphi)$ is related to the particle masses and is a function of φ since some of the masses are generated by the scalar φ. Importantly, in many models $\mathcal{M}^2(\varphi)$ contains a quadratic term. The significance of a quadratic term is that it shifts the position of the minima of $V(\varphi)$. To see this, we can expand eq. (1.3) to obtain

$$V(\varphi) = -\lambda f_a^2\,|\varphi|^2 + \lambda\,|\varphi|^4 + \text{constants}\ . \tag{1.8}$$

If one added a term of the form $c(T)\,|\varphi|^2$ to this expression where $c(T)$ is a temperature-dependent coefficient, then for $c(T) > \lambda f_a^2$ the coefficient of the $|\varphi|^2$ term becomes

positive and the 'hilltop' shape goes away. As a concrete example, consider the Goldstone model

$$\mathcal{L} = (\partial_\mu \varphi)^\dagger (\partial^\mu \varphi) - V(\varphi) \ , \tag{1.9}$$

where $V(\varphi)$ is given by eq. (1.3). In this model $\mathcal{M}^2(\varphi)$ is given by (17)

$$\mathcal{M}^2(\varphi) = 8\lambda \, |\varphi|^2 \ , \tag{1.10}$$

so the effective potential, according to eq. (1.7) is,

$$V_{\rm eff}(\varphi, T) = \frac{\lambda}{3}\left(T^2 - 3f_a^2\right)|\varphi|^2 + \lambda|\varphi|^4 \ , \tag{1.11}$$

where we have dropped φ-independent terms. Above a temperature $T_c = \sqrt{3f_a^2}$ the minimum of $V_{\rm eff}$ is at $\varphi = 0$ and there is no spontaneous symmetry breaking. In a hot big bang-like scenario this would imply symmetry restoration at higher temperatures. Then, as the Universe expands and T drops below T_c the symmetric state $\varphi = 0$ becomes unstable and the radial mode will incur a nonzero value $|\varphi| \neq 0$. On the other hand, in this picture there is no preferred direction for the angular mode ϕ. Thus, different causally disconnected regions of physical space will pick different values of ϕ at random. Necessarily, there will be regions of space where vortices form by chance, giving rise to a network of strings.

1.3.3 Birefringence from strings

In this section I discuss how cosmic axion strings can rotate the plane of polarization of light, an effect known as birefringence. Theories with a Chern-Simons term given by eq. (1.6) couple axions and photons, resulting in modified equations of electrodynamics. One of the consequences of this interaction is that if a linearly polarized electromagnetic wave propagates through a region of space with an inhomogeneous axion field it experiences a rotation of its plane of polarization by an amount (23–26),

$$\Delta\alpha = \frac{\mathcal{A}\alpha_{\rm em}}{2\pi f_a} \int_C \partial_\mu \phi \, {\rm d}X^\mu \tag{1.12}$$

where the integral is performed over the photon's worldline C. If the axion has a trivial topology then the integral in eq. (1.12) is simply the difference between the value of ϕ at the start and end points of C so that we have $\Delta\alpha = \mathcal{A}\alpha_{\rm em}\,\Delta\phi/2\pi f_a$.

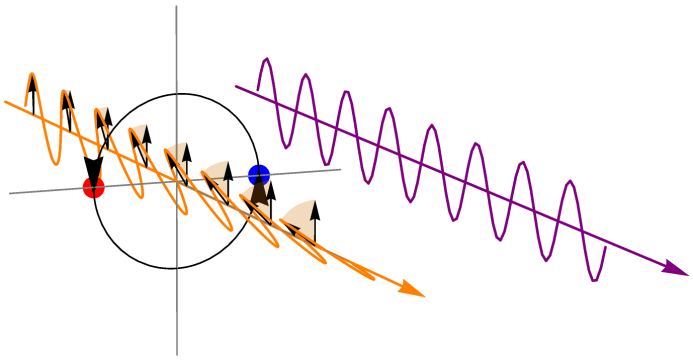
If there are strings, the axion field winds around from 0 to $2\pi f_a$ in closed paths around a single string core. Ref. (27) showed that for photons propagating along an infinitely long path through the center of a circular planar string loop the birefringence angle is

$$\Delta\alpha = \pm\mathcal{A}\alpha_{\rm em}\ , \tag{1.13}$$

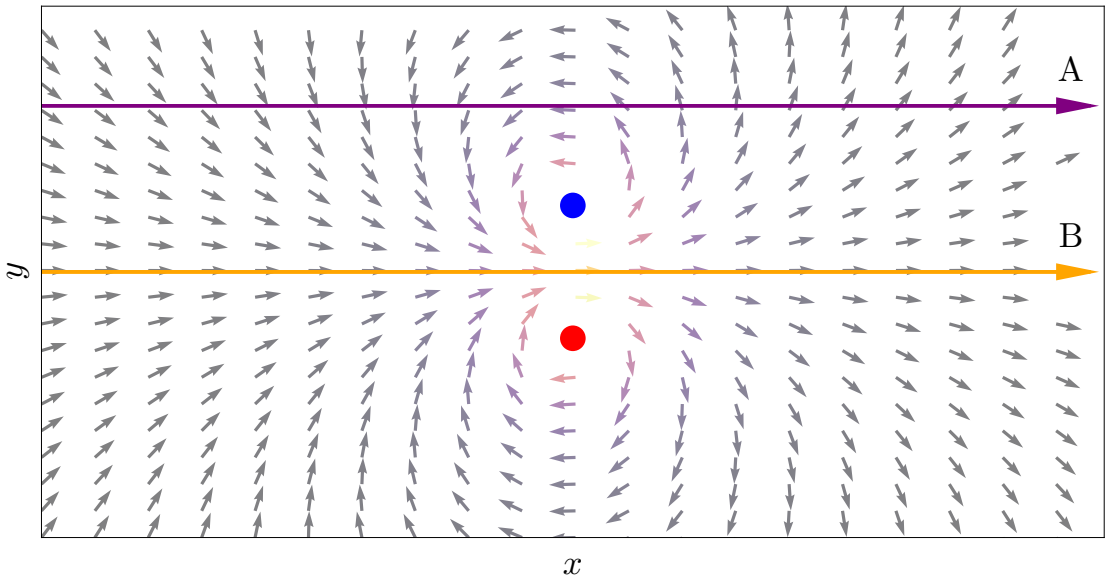
which is independent of f_a. Meanwhile, for photons that do not pass through the interior of a string loop, the incurred birefringence angle is 0. Fig. 1.3a illustrates this effect. A linearly polarized electromagnetic wave (photon) whose electric field is depicted as an orange sinusoid (the magnetic field is not shown for visual clarity). The photon propagates from left to right along an axis perpendicular to the plane of a circular planar string loop (black circle). The loop's clockwise winding number is illustrated with two arrowheads. As the photon propagates its plane of polarization is rotated by an amount dictated by the gradient of the axion field along its trajectory in accordance with eq. (1.12).

To help illustrate why the birefringence is accumulated only for photons passing through the interior of the loop fig. 1.3b shows a 2D cross-section of the planar loop and the gradient of ϕ as seen in that plane. To calculate $\nabla\phi$ the string loop was approximated as two infinitely long strings with opposite winding numbers. The trajectory of the photon shown in fig. 1.3a is shown as a orange arrow labeled B. Since the $\nabla\phi$ points along the direction of the photon's trajectory, the line integral along this path is nonzero and the photon experiences a rotation in its plane of polarization. On the other hand, a photon passing outside the loop exemplified by path A (purple arrow) experiences no net rotation since contributions to the integral of $\nabla\phi$ along this path cancel each other out.

For a string network that survives beyond recombination, linearly polarized CMB



(a) Illustration of birefringence experienced by an electromagnetic wave (orange) propagating through the center of a string loop (black circle). The string's winding orientation is depicted by black arrowheads. Electromagnetic waves that do not cross a string loop experience no rotation (purple).



(b) Illustration of $\nabla\phi$ for an axion field in the xy-plane of fig. 1.3a. Blue and red dots depict where the string loop crosses the plane and are consistent with the blue and red dots of fig. 1.3a. Arrows depicting the vector field are normalized to be the same length. Grey arrows have a smaller magnitude whereas orange and yellow arrows have a larger magnitude. The purple (orange) arrow labeled A (B) depicts an example of a photon trajectories passing outside (inside) the loop. The line integral along A is 0, so the photon experiences no birefringence. The line integral along B is nonzero since $\boldsymbol{\nabla}\phi$ is always pointing in the same direction as B, so there is a nonzero birefringence.

photons will propagate through the string network and experience a net birefringence equal to the sum of the birefringence from each string loop. From the perspective of an observer this means that a photon hitting a detector along the line of sight $\hat{\boldsymbol{n}}$ will have a birefringence angle of $\alpha(\hat{\boldsymbol{n}}) \approx \sum_{\text{loops crossed}} \pm \mathcal{A}\alpha_{\text{em}}$. Fig. 1.4 illustrates how CMB photons accumulate birefringence when propagating through a network of strings. The upper half of fig. 1.4 depicts strings as circular planar loops, drawn on orange shells that represent snapshots in time progressing from the past (right) to the present (left). Each time a photon crosses a loop its plane of polarization is rotated by $\pm\mathcal{A}\alpha_{\text{em}}$. In the lower half of fig. 1.4 the accumulated birefringence angle in each direction $\hat{\boldsymbol{n}}$ up to the indicated redshift z is shown in a Mollweide projection.

The possibility that ALP-strings may have left an imprint in CMB birefringence maps is intriguing because measuring birefringence is a science driver for current and future CMB probes (28–30). To date, studies analyzing CMB polarization measurements claim a nonzero isotropic birefringence angle of $0.342^{+0.094^\circ}_{-0.091^\circ}$ (31), whereas experimental efforts to measure the anisotropic component are consistent with zero (32–37). Future CMB probes are expected to improve on these measurements by orders of magnitude (29; 38; 39), and may therefore be sensitive enough to detect the anisotropic component. Through chapters (2-4) I obtain constraints of axion string network parameters from existing data, study the non-Gaussian signatures of axion-string induced birefringence maps, and show how neural networks can be used to extract axion-string parameters from CMB birefringence maps.

1.3.4 Axion-matter coupling

In sec. 1.3.3 I discussed how the axion-photon interaction term eq. (1.6) could imprint an observable signature on CMB radiation. Another possibly observable effect can occur if axions couple to matter. In particular, we consider an interaction that violates lepton flavour conservation, allowing for processes that, for example, permit

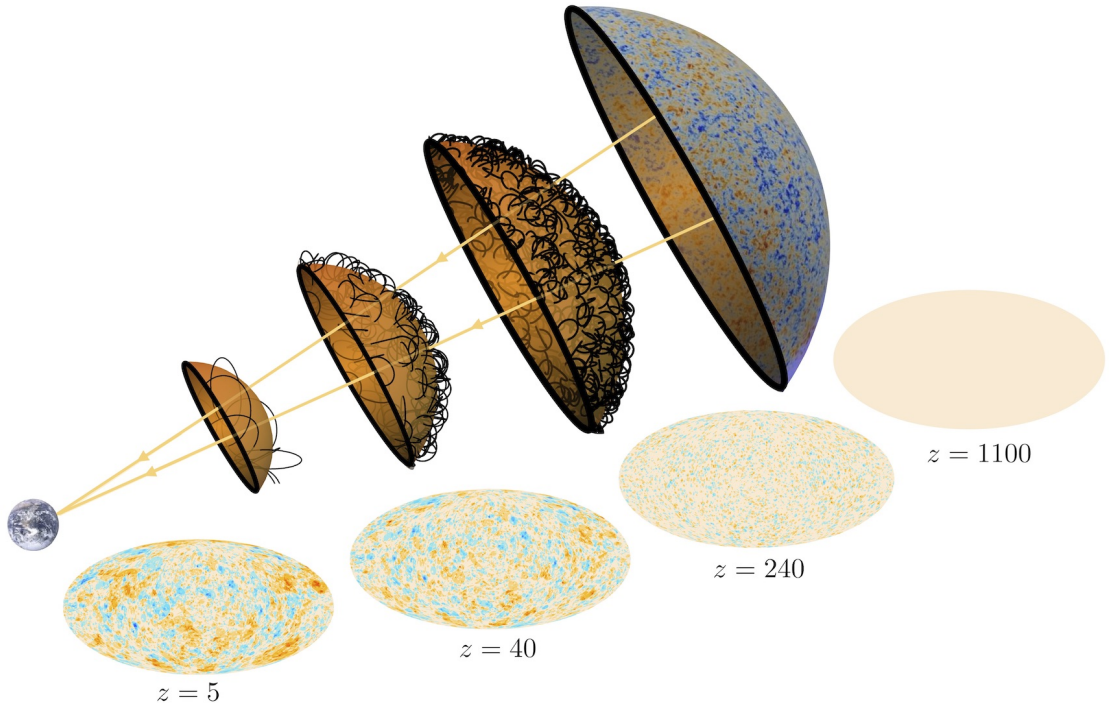


Figure 1.4 : Graphical illustration of the loop-crossing model, which we use to generate birefringence maps. The upper half of the figure depicts CMB photons propagating through a network of strings modeled as circular planar loops. The lower half depicts the birefringence angle $\alpha(\hat{\boldsymbol{n}})$ for photons along each line of sight $\hat{\boldsymbol{n}}$ as seen by an observer on earth. Red/orange indicates a counter-clockwise rotation and blue indicates a clockwise rotation.

the conversion of a muon to an electron through the emission of an axion:

$$\mathcal{L}_{\rm LFV} = \frac{g_{\phi e\mu}}{m_e + m_\mu} \overline{\Psi}_e \gamma^\rho \gamma_5 \Psi_\mu \, \partial_\rho \phi \; + \; \text{h.c.} \; . \tag{1.14}$$

In the environment of a neutron star the interaction eq. (1.14) would give rise to additional energy loss channels (via axion emission), which modify the neutron star's cooling rate.

There are multiple reasons why one would consider lepton flavour violating (LFV) interactions like eq. (1.14). For one, there is no reason to expect the full UV theory of

the SM + ALP to conserve lepton flavour since we already know that lepton flavour conservation is an accidental symmetry of the SM, which is broken by the fact that neutrinos are massive. Moreover, even if the UV theory does conserve lepton flavour, LFV effects can arise from radiative corrections (40–43).

In chapter 5, axion emissivity from neutron stars through the LFV interaction given in eq. (1.14) is computed and used to derive a limit on the coupling constant $g_{\phi e\mu}$.

Bibliography

[1] Fan, X. and Myers, T. G. and Sukra, B. A. D. and Gabrielse, G., *Measurement of the Electron Magnetic Moment*, *Phys. Rev. Lett.* **130** (2023) 071801, [2209.13084].

[2] Gonzalez-Garcia, M. C. and Nir, Yosef, *Neutrino Masses and Mixing: Evidence and Implications*, *Rev. Mod. Phys.* **75** (2003) 345–402, [hep-ph/0202058].

[3] Cirelli, Marco and Strumia, Alessandro and Zupan, Jure, *Dark Matter*, 2406.01705.

[4] Dine, Michael, *TASI lectures on the strong CP problem*, in *Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 2000): Flavor Physics for the Millennium*, pp. 349–369, 6, 2000. hep-ph/0011376.

[5] Marsh, David J. E., *Axion Cosmology*, *Phys. Rept.* **643** (2016) 1–79, [1510.07633].

[6] Crewther, R. J. and Di Vecchia, P. and Veneziano, G. and Witten, Edward, *Chiral Estimate of the Electric Dipole Moment of the Neutron in Quantum Chromodynamics*, *Phys. Lett. B* **88** (1979) 123.

[7] Abel, C. and others, *Measurement of the Permanent Electric Dipole Moment of the Neutron*, *Phys. Rev. Lett.* **124** (2020) 081803, [2001.11966].

[8] Peccei, R. D. and Quinn, Helen R., *CP Conservation in the Presence of Instantons*, *Phys. Rev. Lett.* **38** (1977) 1440–1443.

[9] Srednicki, Mark, *Axion Couplings to Matter. 1. CP Conserving Parts*, *Nucl. Phys. B* **260** (1985) 689–700.

[10] Arvanitaki, Asimina and Dimopoulos, Savas and Dubovsky, Sergei and Kaloper, Nemanja and March-Russell, John, *String Axiverse*, *Phys. Rev. D* **81** (2010) 123530, [0905.4720].

[11] Gendler, Naomi and Marsh, David J. E. and McAllister, Liam and Moritz, Jakob, *Glimmers from the axiverse*, *JCAP* **09** (2024) 071, [2309.13145].

[12] Mehta, Viraf M. and Demirtas, Mehmet and Long, Cody and Marsh, David J. E. and McAllister, Liam and Stott, Matthew J., *Superradiance in string theory*, *JCAP* **07** (2021) 033, [2103.06812].

[13] Dalal, Neal and Kravtsov, Andrey, *Excluding fuzzy dark matter with sizes and stellar kinematics of ultrafaint dwarf galaxies*, *Phys. Rev. D* **106** (2022) 063517, [2203.05750].

[14] Amin, Mustafa A. and Mirbabayi, Mehrdad, *A Lower Bound on Dark Matter Mass*, *Phys. Rev. Lett.* **132** (2024) 221004, [2211.09775].

[15] Agrawal, Prateek and Nee, Michael and Reig, Mario, *Axion Couplings in Heterotic String Theory*, 2410.03820.

[16] Kibble, T. W. B., *Some Implications of a Cosmological Phase Transition*, *Phys. Rept.* **67** (1980) 183.

[17] Vilenkin, A. and Shellard, E. P. S., *Cosmic Strings and Other Topological Defects*. Cambridge University Press, 7, 2000.

[18] Fixsen, D. J. and Cheng, E. S. and Gales, J. M. and Mather, John C. and Shafer, R. A. and Wright, E. L., *The Cosmic Microwave Background spectrum from the full COBE FIRAS data set*, *Astrophys. J.* **473** (1996) 576, [astro-ph/9605054].

[19] Kirzhnits, D. A. and Linde, Andrei D., *Macroscopic Consequences of the Weinberg Model*, *Phys. Lett. B* **42** (1972) 471–474.

[20] Weinberg, Steven, *Gauge and Global Symmetries at High Temperature*, *Phys. Rev. D* **9** (1974) 3357–3378.

[21] Bernard, Claude W., *Feynman Rules for Gauge Theories at Finite Temperature*, *Phys. Rev. D* **9** (1974) 3312.

[22] Dolan, L. and Jackiw, R., *Symmetry Behavior at Finite Temperature*, *Phys. Rev. D* **9** (1974) 3320–3341.

[23] Carroll, Sean M. and Field, George B. and Jackiw, Roman, *Limits on a Lorentz and Parity Violating Modification of Electrodynamics*, *Phys. Rev. D* **41** (1990) 1231.

[24] Carroll, Sean M. and Field, George B., *The Einstein equivalence principle and the polarization of radio galaxies*, *Phys. Rev. D* **43** (1991) 3789.

[25] Fedderke, Michael A. and Graham, Peter W. and Rajendran, Surjeet, *Axion Dark Matter Detection with CMB Polarization*, *Phys. Rev. D* **100** (2019) 015040, [1903.02666].

[26] Jain, Mudit and Long, Andrew J. and Amin, Mustafa A., *CMB birefringence from ultralight-axion string networks*, *JCAP* **05** (2021) 055, [2103.10962].

[27] Agrawal, Prateek and Hook, Anson and Huang, Junwu, *A CMB Millikan experiment with cosmic axiverse strings*, *JHEP* **07** (2020) 138, [1912.02823].

[28] CMB-S4 collaboration, Abazajian, Kevork N. and others, *CMB-S4 Science Book, First Edition*, 1610.02743.

[29] CMB-HD collaboration, Aiola, Simone and others, *Snowmass2021 CMB-HD White Paper*, 2203.05728.

[30] Chang, Clarence L. and others, *Snowmass2021 Cosmic Frontier: Cosmic Microwave Background Measurements White Paper*, 2203.07638.

[31] Eskilt, Johannes R. and Komatsu, Eiichiro, *Improved constraints on cosmic birefringence from the WMAP and Planck cosmic microwave background polarization data*, *Phys. Rev. D* **106** (2022) 063503, [2205.13962].

[32] POLARBEAR collaboration, Ade, Peter A. R. and others, *POLARBEAR Constraints on Cosmic Birefringence and Primordial Magnetic Fields*, *Phys. Rev. D* **92** (2015) 123509, [1509.02461].

[33] BICEP2, Keck Arrary collaboration, Ade, P. A. R. and others, *BICEP2 / Keck Array IX: New bounds on anisotropies of CMB polarization rotation and implications for axionlike particles and primordial magnetic fields*, *Phys. Rev. D* **96** (2017) 102003, [1705.02523].

[34] Contreras, Dagoberto and Boubel, Paula and Scott, Douglas, *Constraints on direction-dependent cosmic birefringence from Planck polarization data*, *JCAP* **12** (2017) 046, [1705.06387].

[35] Namikawa, Toshiya and others, *Atacama Cosmology Telescope: Constraints on cosmic birefringence*, *Phys. Rev. D* **101** (2020) 083527, [2001.10465].

[36] Gruppuso, Alessandro and Molinari, Diego and Natoli, Paolo and Pagano, Luca, *Planck 2018 constraints on anisotropic birefringence and its cross-correlation with CMB anisotropy*, *JCAP* **11** (2020) 066, [2008.10334].

[37] Bortolami, Marco and Billi, Matteo and Gruppuso, Alessandro and Natoli, Paolo and Pagano, Luca, *Planck constraints on cross-correlations between anisotropic cosmic birefringence and CMB polarization*, *JCAP* **09** (2022) 075, [2206.01635].

[38] Pogosian, Levon and Shimon, Meir and Mewes, Matthew and Keating, Brian, *Future CMB constraints on cosmic birefringence and implications for fundamental physics*, *Phys. Rev. D* **100** (2019) 023507, [1904.07855].

[39] BICEP/Keck collaboration, Ade, P. A. R. and others, *BICEP/Keck XVIII: Measurement of BICEP3 polarization angles and consequences for constraining cosmic birefringence and inflation*, 2410.12089.

[40] Choi, Kiwoon and Im, Sang Hui and Park, Chan Beom and Yun, Seokhoon, *Minimal Flavor Violation with Axion-like Particles*, *JHEP* **11** (2017) 070, [1708.00021].

[41] Chala, Mikael and Guedes, Guilherme and Ramos, Maria and Santiago, Jose, *Running in the ALPs*, *Eur. Phys. J. C* **81** (2021) 181, [2012.09017].

[42] Bauer, Martin and Neubert, Matthias and Renner, Sophie and Schnubel, Marvin and Thamm, Andrea, *The Low-Energy Effective Theory of Axions and ALPs*, *JHEP* **04** (2021) 063, [2012.12272].

[43] Bonilla, J. and Brivio, I. and Gavela, M. B. and Sanz, V., *One-loop corrections to ALP couplings*, *JHEP* **11** (2021) 168, [2107.11392].

Chapter 2

Searching for axion-like particles through CMB birefringence from string-wall networks

Abstract

Axion-like particles (ALPs) can form a network of cosmic strings and domain walls that survives after recombination and leads to anisotropic birefringence of the cosmic microwave background (CMB). In addition to studying cosmic strings, we clarify and emphasize how the formation of ALP-field domain walls impacts the cosmic birefringence signal; these observations provide a unique way of probing ALPs with masses in the range $3H_0 \lesssim m_a \lesssim 3H_{\rm cmb}$. Using measurements of CMB birefringence from several telescopes, we find no evidence for axion-defect-induced anisotropic birefringence of the CMB. We extract constraints on the model parameters that include the ALP mass m_a, ALP-photon coupling $\mathcal{A} \propto g_{a\gamma\gamma} f_a$, the domain wall number $N_{\rm dw}$, and parameters characterizing the abundance and size of defects in the string-wall network. Considering also recent evidence for isotropic CMB birefringence, we find it difficult to accommodate this with the non-detection of anisotropic birefringence under the assumption that the signal is generated by an ALP defect network.

Notes about this project: This chapter is from a paper I wrote with Mustafa A. Amin (MA), Andrew J. Long (AL), and Mudit Jain (MJ) that was published in the Journal of Cosmology and Astroparticle Physics in 2022 (1). It is a continuation of earlier work by MA, AL, and MJ who studied anisotropic CMB birefringence from hyperlight ($m_a \lesssim 10^{-28}$ eV axion string networks (2). Ref. (2) introduced a phenomenological model for birefringence all-sky maps from axion strings referred to as the "loop-crossing model" (LCM) and computed the birefringence power spectrum

analytically. A natural next step was to take the predicted birefringence power spectra and compare them with existing birefringence power spectra measurements to see what axion-string parameters are consistent with observations.

I officially joined the group in April 2021 and developed a Python code to perform LCM simulations which enabled us to visualise birefringence maps and provide further validation of the analytical results of ref. (2). Then I developed a framework for doing Bayesian inference on LCM parameters which was used to produce the main results in this paper.

2.1 Introduction

Exquisite measurements of cosmic microwave background (CMB) temperature and polarization anisotropies carried out over the past few decades have revolutionized our understanding of cosmology. The absence of B-mode polarization on large angular scales in the CMB has already provided important insights about cosmological initial conditions (3). Building on these measurements, more subtle analyses of achromatic CMB polarization rotation ("CMB birefringence"), has been a focus of a growing number of recent studies.

CMB birefringence provides an exciting window into physics beyond the Standard Model. For example, hypothetical axion-like particles (ALPs) coupled to photons in the following manner

$$\mathscr{L}_{\rm int} = -\frac{1}{4}\, g_{a\gamma\gamma}\, a\, F_{\mu\nu}\tilde{F}^{\mu\nu} \tag{2.1}$$

can induce a birefringence signal. A photon propagating through a classical ALP field $a(x)$ is expected to experience a frequency-independent birefringence (4–7) as its plane of polarization is rotated by an angle

$$\alpha = -\frac{g_{a\gamma\gamma}}{2} \int_C \mathrm{d}X^\mu\, \partial_\mu a(X)\ . \tag{2.2}$$

The integral runs over the photon's worldline X^μ from the point of emission to detection. Isotropic birefringence has been studied in the context of an approximately

homogeneous ALP field that may constitute the dark matter or dark energy (8–12), yielding information about the ALP-photon coupling ($g_{a\gamma\gamma}$) and the mass of the ALP (m_a).

ALPs may also form a network of topological defects in the Universe, and leave a distinctive imprint on the CMB polarization via cosmological birefringence (2; 13–16). Such a defect network of strings and walls in the ALP field can exist after recombination, depending on the value of m_a, the number of degenerate vacua $N_{\rm dw}$, and the symmetry breaking scale f_a. Even if the defect network is subdominant in energy density, it can lead to a potentially detectable, anisotropic, and frequency-independent birefringence signal, which depends on m_a, $N_{\rm dw}$, and the anomaly coefficient $\mathcal{A} \propto g_{a\gamma\gamma} f_a$.

Getting a handle on m_a, $N_{\rm dw}$, f_a, and $\mathcal{A}$ would be invaluable from a high energy physics point of view. For example, a global shift symmetry in the axion field would require a vanishing ALP mass $m_a = 0$, whereas a nonzero mass would signal that this symmetry is broken. The general expectation is that all global symmetries are explicitly broken due to quantum gravitational effects in string theory (17–21). On the other hand, some alternative constructions of quantum gravity such as asymptotic safe gravity, may allow global symmetries to be preserved (22). Therefore, probing the mass of ALPs and the vacuum structure of their effective potential would constitute a test of the underlying nature of quantum gravity (23). Birefringence from the defect network probes new physics at the scale f_a since the charges of particles at this scale determine the anomaly coefficient $\mathcal{A}$. Even if $f_a \gg$ TeV and these particles cannot be probed directly at colliders, measurements of CMB birefringence could provide valuable insight into new high energy physics.

For most of this work, we focus on exceptionally light ALPs, with masses that are comparable to the Hubble parameter between recombination and today (though much higher and lower masses are also discussed). This regime is relevant for the types of string-wall networks that can be present after recombination (see fig. 2.1). Such

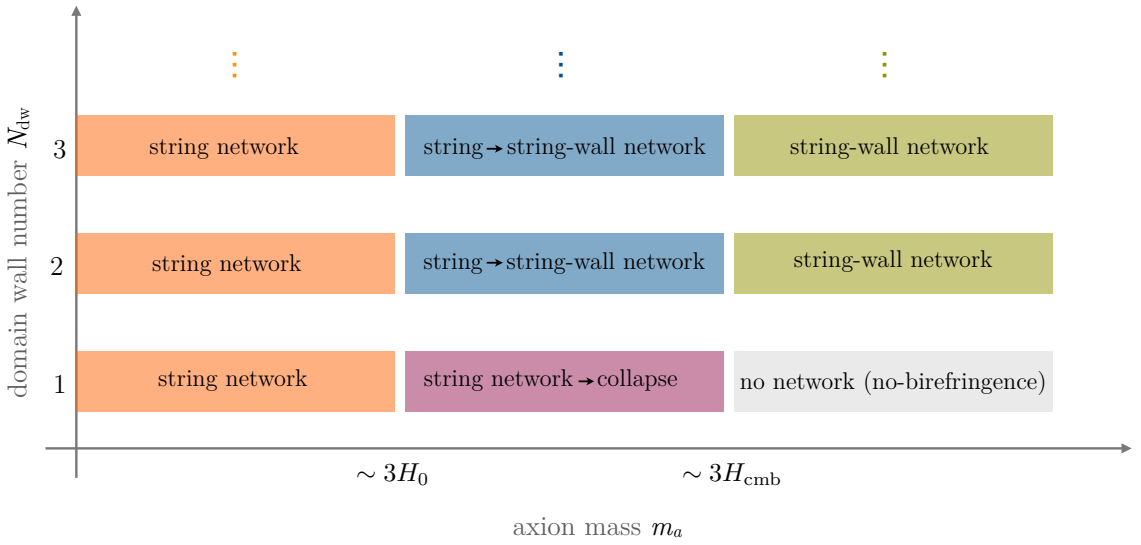


Figure 2.1 : The types of string-wall networks seen by CMB photons travelling from the surface of last-scattering to us. The possible types of networks seen by CMB photons depend on two parameters: the domain wall number $N_{\rm dw}$ and the mass of the axion m_a.

light masses arise naturally in many string theory constructions via non-perturbative effects (24–29).

The birefringence signature of an ALP defect network was first studied in ref. (13). Building on that work, some of us developed semi-analytic models to calculate the expected power spectrum of axion-defect-induced birefringence for different ranges of ALP masses and network dynamics (2). Recently, ref. (15) used *Planck* (2015) data to investigate some of these models (particularly those with stable string networks, orange blocks in fig. 2.1) and derive constraints on their parameters. Taking advantage of these prior studies, our goals for the present paper are as follows:

- Provide the first constraints on collapsing ALP string-wall networks imposed by measurements of anisotropic birefringence from CMB data.

- Clarify the role played by walls in string-wall networks in the context of bire-

fringence.

- Test for axion-defect-induced birefringence using measurements of anisotropic birefringence derived from data taken by various telescopes: *Planck*, ACTpol, SPTpol, BICEP2/*Keck Array*, and Polarbear.

- Validate the results of our earlier semi-analytic work on anisotropic birefringence (2) by ray-tracing through statistical ensembles of defect networks and calculating the corresponding birefringence maps and power spectra.

- Assess the compatibility of recent measurements of isotropic birefringence with limits on anisotropic birefringence assuming the source is an axion string-wall network.

In contrast with ref. (15), we also investigate scenarios with $m_a \gtrsim 3H_0$ in this paper. Since walls form when $m_a \sim 3H$, this necessitates including effects due to collapsing string-wall networks or stable string-wall networks (blue, purple and green regions in fig. 2.1). We find this higher mass window particularly intriguing since the associated phenomenology has the potential to provide a measurement of m_a in the $N_{\rm dw} = 1$ case. By contrast, astrophysical observations such as probes of exotic stellar emission, are sensitive to arbitrarily light ALPs, but such observations cannot constrain their masses.

The structure of the paper is as follows. In sec. 2.2 we review the loop-crossing model and discuss how each of the different string-wall network models are described by this framework. In sec. 2.3 we briefly summarize the current status of CMB birefringence measurements. In sec. 2.4 we report on the main results of our work: the non-observation of birefringence implies constraints on a hyperlight axion-like particle. In sec. 2.5 we discuss the implications of the measurements of isotropic birefringence for our models and analysis. Finally, we summarize and conclude in sec. 4.6. The article is extended by three appendices. App. 2.A contains a detailed outline of the procedure that we have used to simulate birefringence sky maps. App. 2.B presents

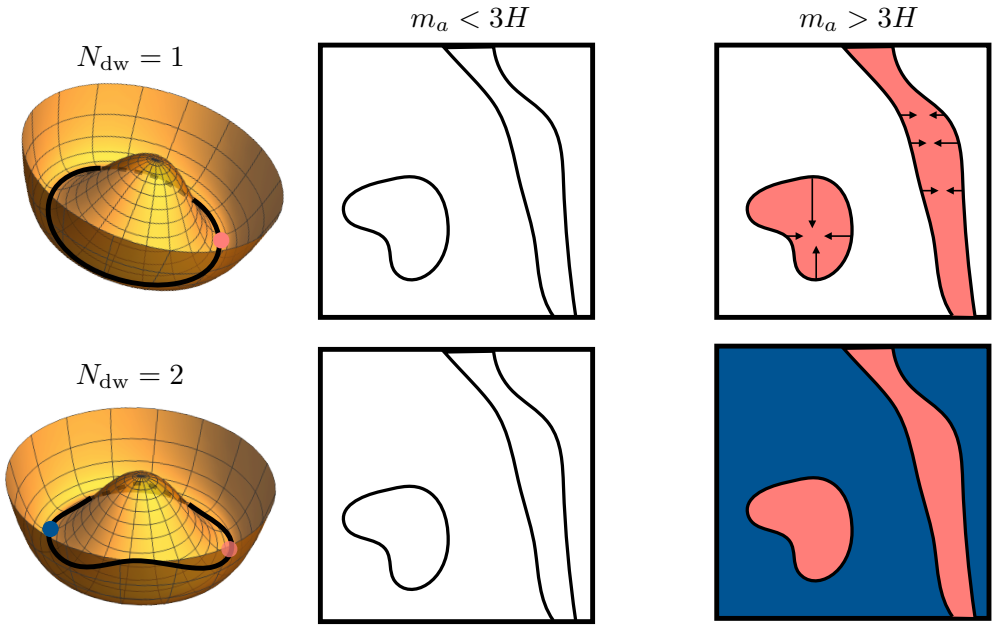


Figure 2.2 : An illustration of the axion string-wall network dynamics – black lines represent strings and colored regions represent walls. *Top row*: For $N_{\rm dw} = 1$, the string network survives from formation until $m_a \simeq 3H(t)$. Thereafter, field gradients in the space between strings realign to form domain walls, which pull on the strings causing the network to collapse in a few Hubble times. *Bottom row*: For $N_{\rm dw} = 2$, each string attaches to two domain walls, and the balance of forces from different walls prevents collapse and allows the network to survive after $m_a \simeq 3H$.

a statistical estimator that is often used to extract birefringence measurements from CMB polarization data. Finally, app. 2.C reports on searches for axion-defect-induced birefringence in additional data sets.

2.2 CMB birefringence from an axion string-wall network

The ALP field can form a topological defect network consisting of cosmic strings and domain walls (30; 31).* In the early Universe, if the ALP field's global symmetry is broken after inflation, then the associated phase transition fills the Universe with a network of cosmic strings (33; 34). The string network exhibits rich dynamics, such as the oscillation of curved string segments under the influence of their own tension, the formation of string loops from the crossing and reconnection of string segments, and the evaporation of string loops by the emission of ALPs (35–38). If the Hubble parameter drops below the axion mass scale, $3H(t) \simeq m_a$, the ALP field in the space between strings is released from Hubble drag, and the strings become bounded by domain walls; see fig. 2.2 for an illustration. The number $N_{\rm dw}$ of domain walls attached to each string is a parameter of the theory, associated with explicit symmetry breaking. For models with $N_{\rm dw} = 1$ the string-wall network collapses into a bath of ALPs within a few Hubble times, but for $N_{\rm dw} \geq 2$ the network is stable due to the balance of forces.

To assess the implications of an ALP string-wall network for CMB birefringence, it is useful to break up the parameter subspace $(m_a, N_{\rm dw})$ into four regions as illustrated in fig. 2.1. For $m_a \lesssim 3H_0$, the domain walls have not yet formed by today, which makes $N_{\rm dw}$ irrelevant, and the birefringence signal arises from axion strings alone. For $3H_0 \lesssim m_a \lesssim 3H_{\rm cmb}$ and $N_{\rm dw} = 1$, the domain walls form and cause the network to collapse between recombination and today. This shuts off the accumulation of birefringence at the time when $3H(t) \simeq m_a$. For $3H_0 \lesssim m_a \lesssim 3H_{\rm cmb}$ and $N_{\rm dw} \geq 2$, the formation of domain walls converts the string network into a stable string-wall network between recombination and today, whereas for $3H_{\rm cmb} \lesssim m_a$ and $N_{\rm dw} \geq 2$, this conversion occurs before recombination.

When coupled to electromagnetism (2.1), ALP strings and walls induce a frequency-independent birefringence signal. This signal is insensitive to the symmetry breaking

*See ref. (32) for an explicit discussion of cosmic strings arising in the String-Axiverse.

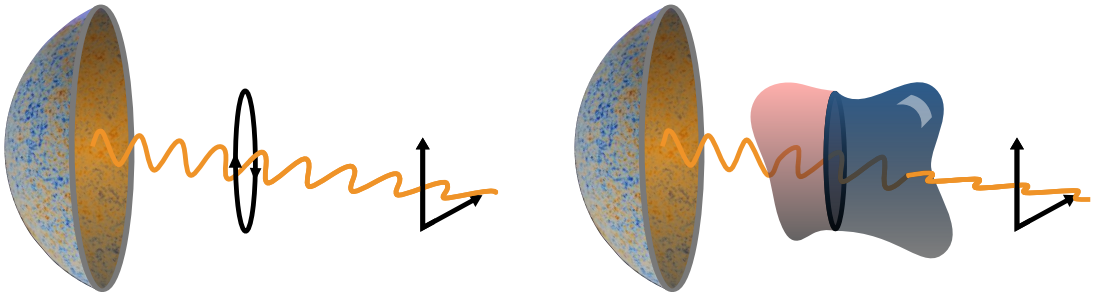


Figure 2.3 : An illustration of axion-defect-induced birefringence. *Left:* A photon's plane of polarization rotates gradually as it approaches and passes through a string loop without domain walls. *Right:* For a string loop bounded by two domain walls, the polarization axis rotates 'abruptly' upon crossing each wall. In both scenarios, with and without domain walls, the net effect is the same rotation angle $\alpha = \pm\mathcal{A}\alpha_{\rm em}$.

scale f_a but directly probes the anomaly coefficient $\mathcal{A} = -\pi f_a g_{a\gamma\gamma}/\alpha_{\rm em}$ (13). For instance, if walls have not yet formed, then a photon crossing through a string loop 'sees' the axion field pass through a full cycle $\Delta a = \int_C \mathrm{d}X^\mu\, \partial_\mu a(X) \to \pm 2\pi f_a$, and the resultant birefringence angle (2.2) is $\alpha \to \pm\mathcal{A}\alpha_{\rm em}$. The $\pm$ factor is the string's winding number,[†] and the arrow indicates the limiting value as the end points of C are taken far away from the loop; see fig. 2.3 for an illustration. If domain walls are present in the network, the birefringence effect is $\alpha = \pm\mathcal{A}\alpha_{\rm em}/N_{\rm dw}$ at each wall crossing, since the axion field changes "abruptly" by $\Delta a = \pm 2\pi f_a/N_{\rm dw}$. However, since each string connects to $N_{\rm dw}$ walls, the net effect is insensitive to $N_{\rm dw}$; we discuss this point further in sec. 2.2.3.

To calculate the birefringence signal from an axion string-wall network, we employ the 'loop crossing model' developed in refs. (2; 13). The loop crossing model

[†]Although strings can have winding numbers $\pm 1, \pm 2, ...$, only the ± 1 strings are the most stable. Throughout our work, we only consider $+1$ or -1 winding numbers.

captures features of the network's rich structure and dynamics that are particularly relevant for birefringence. In this framework the network is treated as a collection of circular, planar string loops that are uniformly distributed throughout space and isotropically oriented. The density of loops and its time dependence are controlled by the model parameters. The birefringence is calculated by considering photons propagating through the network and associating an angle $\alpha = \pm\mathcal{A}\alpha_{\rm em}$ to a photon that traverses the disk bounded by a loop. As the photon crosses through multiple loops with random winding numbers ± 1, the accumulated phase-shift grows like a random walk. For two points on the sky $\hat{\gamma}_1$ and $\hat{\gamma}_2$ separated by an opening angle $\theta_o = \arccos(\hat{\gamma}_1 \cdot \hat{\gamma}_2)$, the correlation between the accumulated birefringence of CMB photons from these points is taken to be (13)

$$\langle \alpha(\hat{\gamma}_1)\, \alpha(\hat{\gamma}_2) \rangle = (\mathcal{A}\alpha_{\rm em})^2\, N_{\rm both}(\theta_o)\ , \tag{2.3}$$

where $N_{\rm both}(\theta_o)$ is the average number of loops that both photons traverse. The associated angular power spectrum $C_\ell^{\alpha\alpha}$ is given by

$$C_\ell^{\alpha\alpha} = 2\pi \int_{-1}^{1} {\rm d}(\cos\theta_o)\, P_\ell(\cos\theta_o)\, \langle \alpha(\hat{\gamma}_1)\alpha(\hat{\gamma}_2) \rangle\ , \tag{2.4}$$

and $\ell(\ell+1)C_\ell^{\alpha\alpha}/2\pi$ would be constant for scale-invariant anisotropic birefringence.

The model dependence enters through $N_{\rm both}$, which knows about the density of loops in the network and their length distribution. It can be analytically approximated as (2)

$$N_{\rm both}(\theta_o) \approx \int_0^\infty {\rm d}\zeta \int_0^{\tilde{z}_*(\zeta,\theta_o)} {\rm d}z\, Q(\zeta, z, 0)\, \chi(\zeta, z)\ , \tag{2.5}$$

where ζ is a dimensionless measure of loop length, z is redshift, $Q(\zeta, z, 0)$ is a kernel function, and $\chi(\zeta, z)$ contains the model dependence. Expressions for $Q(\zeta, z, 0)$ and $\tilde{z}_*(\zeta, \theta_o)$ are available in eqs. (3.25) and (3.34) of ref. (2). For the networks that we study in the following subsections, the defects' size tracks the growing Hubble scale. Consequently birefringence on small angular scales is imprinted by small loops at early times, whereas large-scale features are imprinted at later times; see fig. 2.4

for an illustration. The model function $\chi(\zeta, z)$ is normalized such that it becomes time-independent for a string network in scaling $\chi(\zeta, z) = \chi(\zeta)$, and then the average string length per Hubble volume is ξ_0/H with constant $\xi_0 = \int_0^\infty \mathrm{d}\zeta\ \chi(\zeta)$.

Recent numerical simulations have sparked some debate as to whether global string networks (such as the ones we consider here) exhibit scaling (39–46) with constant ξ_0 or whether they deviate from scaling (47–53) with a slowly-growing ξ_0. In our work, the CMB birefringence signal only depends upon the string network evolution between recombination and today, so a logarithmic change in ξ_0 would induce a $1 - (\log f_a/H_{\rm cmb})/(\log f_a/H_0) \sim \mathcal{O}(10\%)$ effect on the birefringence signal, which can be neglected. If the network maintains scaling we expect $\xi_0 \sim \mathcal{O}(\text{few})$ and if there is a logarithmic growth during the long time interval between formation and recombination, we expect $\xi_0 \sim \log f_a/H_{\rm cmb} \sim \mathcal{O}(10)$; our analysis captures both scenarios by setting a wide prior on ξ_0.

2.2.1 Stable string network

For sufficiently small ALP masses $m_a < 3H_0$ and arbitrary $N_{\rm dw}$, domain walls have not yet formed today, and the defect network is a stable string network. We consider the 'uniform loop size' model of ref. (2); the model function is

$$\chi(\zeta) = \xi_0\, \delta(\zeta - \zeta_0)\ , \tag{2.6}$$

such that all loops in the network at time t have the same radius $\zeta_0/H(t)$, and ξ_0 is the average number of loops per Hubble volume. Recent simulations (47–53) of axion string networks identify a dominant population of large loops and infinite strings, which motivates parameters: $\xi_0 = \mathcal{O}(1-10)$ and $\zeta_0 = \mathcal{O}(0.1-1)$. The birefringence two-point correlation function is approximately (2)

$$\langle \alpha(\hat{\gamma}_1)\, \alpha(\hat{\gamma}_2)\rangle \approx \mathcal{A}^2 \xi_0\, \alpha_{\rm em}^2 \begin{cases} \dfrac{\zeta_0}{4}\left(\log\big(1+\tilde{z}_*(\zeta_0,\theta_o)\big) - \dfrac{\zeta_0}{3}\right) & \theta_o < \theta_t \\ \dfrac{1}{3\zeta_0}\ \log^3\big(1+\tilde{z}_*(\zeta_0,\theta_o)\big) & \theta_o > \theta_t \end{cases}\ , \tag{2.7}$$

where $\theta_t \approx 1$. Note that there is a degeneracy between ξ_0 and $\mathcal{A}$; only the combination $\mathcal{A}^2\xi_0$ appears in the correlation function, and it controls the amplitude of the signal.

The expected birefringence signal is shown in fig. 2.5 for $\zeta_0 = 1$ and $\mathcal{A}^2\xi_0 = 1$. The left panel shows a simulated realization of the birefringence angle over the sky; see app. 2.A for details of the simulation. Note the loop-like features that span a wide range of angular scales. The right panel shows the corresponding angular power spectrum (2.4). For each realization we calculate a $C_\ell^{\alpha\alpha}$; the blue curve shows their mean and the blue band shows the 68% containment region. This band broadens toward low ℓ due to cosmic variance. Note that the spectrum is almost scale invariant for $\ell \lesssim 100$, which follows from the assumed scale invariance of the string network. Exact scale invariance is broken by the angular size of the string loops at recombination, which sets a minimal angular scale for the birefringence anisotropies that corresponds to the peak at $\ell_p \sim 0.1\pi/(\zeta_0\,\theta_{\rm cmb}) \approx 40/\zeta_0$. The right panel also shows the analytic approximation in eq. (2.7) as the gray-dashed curve. Note that the approximation agrees exceptionally well with the direct simulation, which partly validates our use of eq. (2.7) for data fitting and parameter constraints in sec. 2.4.

2.2.2 Collapsing string-wall network

For ALP mass in the range $3H_0 < m_a < 3H_{\rm cmb}$ with $N_{\rm dw} = 1$, the string network develops domain walls and collapses between recombination and today. We consider the 'string network collapse' model of (2), which has the model function

$$\chi(\zeta, z) = \xi_0\ \delta(\zeta - \zeta_0)\ \Theta(z - z_c)\ . \tag{2.8}$$

The dimensionless parameters ξ_0 and ζ_0 have the same interpretation as in sec. 2.2.1, and z_c is the redshift when $3H(t) = m_a$, which is given by

$$z_c = \left[\left(\frac{(m_a/3H_0)^2 - \Omega_\Lambda}{\Omega_m}\right)^{1/3} - 1\right]\ , \tag{2.9}$$

in ΛCDM cosmology. The step function $\Theta(z - z_c)$ models a rapid formation of domain walls at $z = z_c$ and an abrupt collapse of the string-wall network, which shuts off

any further accumulation of birefringence.[‡] The birefringence two-point correlation function is approximately given as (2)

$$\langle\alpha(\hat{\gamma}_1)\,\alpha(\hat{\gamma}_2)\rangle \approx \xi_0\big(\mathcal{A}\,\alpha_{\rm em}\big)^2 \begin{cases} \frac{\zeta_0}{4}\log\left(\frac{1+(z_c^{3/2}+\tilde{z}_*^{3/2})^{2/3}}{1+z_c}\right) & \theta_o,\theta_c<\theta_t \\ \frac{\zeta_0}{4}\log\left(1+(z_c^{3/2}+\tilde{z}_*^{3/2})^{2/3}\right) & \\ \quad -\frac{1}{3\zeta_0}\log^3(1+z_c)-\frac{\zeta_0^2}{12} & \theta_o<\theta_t<\theta_c \\ \frac{1}{3\zeta_0}\Big(\log^3(1+(z_c^{3/2}+\tilde{z}_*^{3/2})^{2/3}) & \\ \quad -\log^3(1+z_c)\Big) & \theta_t<\theta_o,\theta_c \end{cases} , \quad (2.10)$$

where θ_c corresponds to the effective angular size of loops at the time of collapse.

We show the expected birefringence signal in fig. 2.6 for $\mathcal{A}^2\xi_0 = 1$, $\zeta_0 = 1$, and three choices of the ALP mass m_a. Note that $3H_0 \approx 4.5\times10^{-33}$ eV (for $h = 0.7$) and $H_{\rm cmb} \approx 1.0\times10^{-29}$ eV (for $z_{\rm cmb} = 1100$, $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $\Omega_r = 9\times10^{-5}$). Raising the ALP mass causes the network to collapse earlier, which suppresses power at large angular scales, since larger loops would have formed later. The power spectrum displays a strong scale dependence $\ell(\ell+1)C_\ell^{\alpha\alpha} \propto \ell^2$ for $\ell \lesssim \ell_c \sim \pi/\theta_c(m_a)$, where $\theta_c(m_a)$ corresponds to the angular size of loops at the time of network collapse (2). The power spectrum calculations also show good agreement between the direct numerical simulation and the analytic approximation in eq. (2.10). An $\mathcal{O}(1)$ discrepancy develops at large m_a for high $\ell \gtrsim 100$; for the purpose of data analysis, we neglect this mismatch and use the analytic calculation (dashed curves).

2.2.3 Stable string-wall network

For $N_{\rm dw} \geq 2$ the formation of domain walls at the time when $3H(t) \simeq m_a$ leads to a stable string-wall defect network. The resultant CMB birefringence signal is expected

[‡]After the string-wall network collapses, its energy is transferred to a population of non-relativistic ALPs (a subdominant component of the dark matter), which continue to induce birefringence. However, this contribution to the total birefringence is suppressed at low ℓ by $\sim 10^{-3}H_0/m_a$ for $m_a \gtrsim 100H_0$, making it negligible (2).

to be qualitatively unchanged from the stable string network without walls (13), which we discussed already in sec. 2.2.1. We argue this point in two steps: first, we argue that the realignment of smooth axion field gradients around strings into sharp gradients across walls does not change the integrated gradient 'seen' by a photon propagating through the network; and second, we argue that the abundance of walls in the network follows the same scaling as the abundance of string loops in the wall-free network.

Since the birefringence signal is proportional to the integrated axion field gradient (2.2), it is necessary to understand how this quantity differs whether or not walls are present in the network. For a network of axion strings without domain walls, the axion field's gradient varies smoothly throughout space. We illustrate this behavior in fig. 2.7 by showing an axion field configuration for a collection of parallel long strings (vortices in two dimensions). For instance, along the path from point A to point B the axion field passes through a full cycle and the integrated field gradient is $\Delta a = 2\pi f_a$. If the same collection of strings were each connected to $N_{\rm dw} = 3$ domain walls, then the field gradients would be localized in space in order to minimize the energy of the configuration. Nevertheless, on the path from A to B the integrated field gradient would remain equal to $\Delta a = 2\pi f_a$; each wall contributes only $2\pi f_a/3$, but there are 3 walls along the path. More generally, the presence or absence of walls for a given collection of vortices will not impact the integrated phase gradient modulo sub-2π variation in the field value at the endpoints.

The birefringence signal also depends upon the abundance of domain walls in the string-wall network. Here we argue that for a network in scaling, the number density of domain walls tracks the number density of string loops, which evolves in the same way whether or not there are walls. Since every domain wall ends on a string and each string has $N_{\rm dw}$ domain walls then the number densities of walls and strings are related by $n_{\rm dw} \approx N_{\rm dw} n_s$.[§] Numerical simulations (54–56) of axion string-wall

[§]Walls may also close on themselves forming bubbles. A photon passing through one of these configurations does not experience a net birefringence since the contributions from the two wall

networks indicate that the strings reach a scaling regime in which the number density $n_s(t)$ evolves as if the walls were absent. Thus we conclude that the density of walls in the string-wall network tracks the density of strings in the wall-free network.

It is worth noting that these arguments apply equally well for models with $m_a > 3H_{\rm cmb}$ that form walls before recombination and for those with $3H_0 < m_a < 3H_{\rm cmb}$ that form walls after recombination. In terms of the parameter space shown in fig. 2.1, the CMB birefringence signal will be qualitatively unchanged for "string network," "string→string-wall network," and "string-wall network."

The energy density in the stable string-wall network is stored mostly in the rest mass of the domain walls, which have tension $\sigma \approx 8 m_a f_a^2 / N_{\rm dw}^2$. Since the wall's energy redshifts more slowly than strings, the stable string-wall network could present a problem for cosmological observables. If $\xi_{\rm dw}$ is the average number of walls per Hubble volume today, then the string-wall network's energy density is approximately $\xi_{\rm dw} \sigma H_0$. A weak cosmological constraint is obtained by requiring this energy to be small compared to the critical density today $3 m_{\rm pl}^2 H_0^2$, which implies:

$$\frac{\xi_{\rm dw} \sigma H_0}{3 m_{\rm pl}^2 H_0^2} \sim \left(\frac{\xi_{\rm dw}}{N_{\rm dw}/2}\right)\left(\frac{N_{\rm dw}}{2}\right)^{-1}\left(\frac{m_a}{10^{-20}\,{\rm eV}}\right)\left(\frac{f_a}{10^{12}\ {\rm GeV}}\right)^2 \ll 1\ . \tag{2.11}$$

It is worth emphasizing that a sufficiently small decay constant f_a allows a larger axion mass m_a while still satisfying the overclosure condition and observational constraints on the axion-photon coupling (non-minimal models such as clockwork axion (57; 58) may help to further open this parameter space). For instance, if $f_a \sim 10^{10}$ GeV (allowed in type IIB String Theory constructions in the large volume scenario (59; 60)), then the upper bound on m_a can even be as large as $\sim 10^{-16}$ eV with $\xi_{\rm dw}/N_{\rm dw} \sim$

---

crossings cancel. On the other hand, some component of the birefringence signal arises from photons emitted within a bubble and detected at a point outside (or vice versa) (16). Assuming that bubbles are not nested, this component of the birefringence signal may be as large as $|\alpha| = \mathcal{A}\alpha_{\rm em}/N_{\rm dw}$. However, for the string-wall networks that we consider, the effect of multiple wall crossings (loop crossing) allows $|\alpha|$ to accumulate to values that are larger than $\mathcal{A}\alpha_{\rm em}$, which is the dominant component.

$\mathcal{O}(1)$. Axion masses above $m_a \gtrsim 10^{-21}$ eV are especially interesting, since they avoid the Lyman-α bound on ultralight dark matter (61; 62). This scenario in which the ultralight ALP makes up a significant fraction of dark matter and leads to a birefringence signal in the CMB merits closer study, which we do not pursue further here.

2.3 Measurements of cosmological birefringence with CMB data

Cosmological birefringence is expected to leave a distinctive imprint on the polarization anisotropies in the cosmic microwave background radiation. At a given point on the sky, a rotation in the CMB's plane of polarization cannot be measured directly, since the initial orientation on the surface of last scattering is not known. Nevertheless, the polarization pattern across the sky carries information about cosmological birefringence. CMB polarization maps may be decomposed into parity-even E-mode and parity-odd B-mode type polarization patterns (63–65). Thompson scattering at the surface of last scattering generates E-mode polarization, whereas B-mode polarization requires parity-violating sources, such as gravitational wave radiation (66). Cosmological birefringence partially converts E-mode polarization into B-mode polarization (and vice versa). This induces a B-mode signal if none was present otherwise and leads to correlations among the temperature and polarization patterns (67–69).

In order to extract information about birefringence from CMB temperature and polarization data, it is customary to work with a set of statistical quantities that provide unbiased estimators of the birefringence (67–69). These α-estimators are constructed from pairs of CMB power spectra (possibly also correlating across different frequency bins). The power spectrum of the α-estimators are equal to the birefringence angular power spectrum $C_\ell^{\alpha\alpha}$ up to a noise term. In this way, the measured EE, BB, EB, TE, and TB power spectra are used to reconstruct the birefringence power spectrum. For pedagogical purposes, in app. 2.B we define the α-estimators

and demonstrate the reconstruction procedure for mock polarization data.

The data from various CMB telescopes have been analyzed to search for evidence of cosmological birefringence. Assuming that the birefringence effect is isotropic (same rotation angle α at every point on the sky), several studies have recently reported evidence for nonzero birefringence (70–73), including ref. (73) that reports an angle $0.342^{\circ}{}^{+0.094^{\circ}}_{-0.091^{\circ}}$ using data from *Planck* and *WMAP*. On the other hand, for anisotropic birefringence that is statistically isotropic, data from several of the current-generation CMB telescopes has been used to extract a measurement of the birefringence power spectrum. The results of these studies are summarized in fig. 2.8. In particular, note that *Planck* data has been analyzed by two different groups using different α-estimators, which partly explains the scatter in their results. These various measurements in fig. 2.8 indicate an absence of evidence for anisotropic cosmological birefringence at the level of $\sim 0.1^{\circ}$ or greater on large angular scales. Next-generation surveys, such as COrE (74), LiteBIRD (75), Simons Observatory (76), CMB Stage IV (77), and PICO (78), expect to deliver measurements of CMB polarization with unprecedented precision. These observations will prove to be a powerful probe of cosmological birefringence, improving constraints by 2 to 3 additional orders of magnitude (79), and potentially uncovering evidence for birefringence from axion string-wall networks. For the time being, the data shown in fig. 2.8 imposes constraints on the models, which we quantify in the next section.

2.4 Constraints from anisotropic CMB birefringence measurements

In this section we outline the Bayesian inference method have used to search for evidence of birefringence in CMB data. For each dataset we approximate the likelihood as a Gaussian and assume vanishing covariance across multipoles (data is available for each multipole up to $\ell = 30$, whereas data above $\ell = 30$ is binned). We take the

log-likelihood to be

$$\ln \mathcal{L}\big(C_\ell^{\rm obs}|\theta\big) = \sum_\ell -\frac{1}{2\sigma_\ell^2}\Big[C_\ell^{\rm obs} - C_\ell^{\rm th}(\theta)\Big]^2\,, \tag{2.12}$$

where θ represents the model parameters of the theory (listed in tab. 2.1 along with the assumed priors), and σ_ℓ are the uncertainty in the observed values $C_\ell^{\rm obs}$. We take them to be the error bars in the published birefringence power spectrum plots, which are reproduced in fig. 2.8. To obtain the posterior for our likelihood and priors, we perform a Markov Chain Monte Carlo (MCMC) simulation using the Metropolis algorithm implemented in the Python package `PyMC` (87). For all data sets, we ran 10 chains in parallel for at least $5,000$ steps (with some up to even $50,000$ steps depending upon the data set).[¶] We assess their convergence by manually inspecting their trace plots for good mixing, and also ensuring that the Gelman-Rubin statistic $\hat{R}$ (88), for each model parameter, is close to 1. For each parameter $X \in \theta$ (e.g. $X = \mathcal{A}^2\xi_0$), our Markov chains satisfy $|\hat{R}_X - 1| < 0.04$.

2.4.1 Stable string network

We first study a network of stable hyperlight axion strings. The birefringence power spectrum $C_\ell^{\alpha\alpha}$ is calculated using the procedure described in sec. 2.2.1. It is a function of the amplitude parameter $\mathcal{A}^2\xi_0$ and the loop length parameter ζ_0, which have the priors shown in tab. 2.1. We allow the amplitude parameter $\mathcal{A}^2\xi_0$ to take unphysical, negative values in order to assess the presence of systematic bias in the data; we only use positive values to derive constraints.

The result of our MCMC sampling is summarized in fig. 2.9, which shows the posterior probability distribution over the amplitude parameter $\mathcal{A}^2\xi_0$ and the loop length parameter ζ_0. This figure illustrates the constraints from *Planck* (2018) ($1 \leq$

[¶] `PyMC` requires the user to specify the number of tuning steps, used to optimize the sampling algorithm. We used anywhere between $1000 - 2500$ tuning steps (for different data sets), which were eventually discarded to get all of our final results.

	stable strings	collapsing string-wall
amplitude:	$\mathcal{A}^2\xi_0 \sim \mathrm{U}(-\infty, \infty)$	$\mathcal{A}^2\xi_0 \sim \mathrm{U}(-\infty, \infty)$
loop length:	$\zeta_0 \sim \mathrm{U}(0.1, 1.0)$	$\zeta_0 \sim \mathrm{U}(0.1, 1.0)$
axion mass:	N/A	$\log_{10}\left(\frac{m_a}{\mathrm{eV}}\right) \sim \mathrm{U}(-32.4, -28.0)$
DW number:	N/A	$N_{\mathrm{dw}} = 1$

Table 2.1 : The two string-wall network models that we study, their model parameters, and the prior ranges used for MCMC sampling. The function $\mathrm{U}(a, b)$ denotes a uniform probability density on the interval from a to b and a vanishing probability outside this interval.

$\ell \leq 24$) and SPTpol ($75 \leq \ell \leq 525$); constraints from other data sets can be found in app. 2.C. The joint posterior distribution shows a degeneracy direction where $\mathcal{A}^2\xi_0 = 0$, since $C_\ell^{\alpha\alpha}$ becomes independent of ζ_0 when $\mathcal{A}^2\xi_0 = 0$. Similarly, the joint posterior broadens toward smaller ζ_0, since $\zeta_0 = 0$ is another degeneracy direction; our prior enforces $0.1 \leq \zeta_0$ (see discussion in sec. 2.2.1), and the degeneracy at $\zeta_0 = 0$ is not seen on the plot. The SPTpol data has a very slight preference for $\mathcal{A}^2\xi_0 < 0$ due to a pair of downward fluctuations in the data at $\ell = 120$ and 160, whereas the *Planck* (2018) data has a wider tail toward $\mathcal{A}^2\xi_0 > 0$ due to a few upward fluctuations at $\ell = 2$, 5, and 6. Note that the *Planck* (2018) and SPTpol measurements have comparable constraining power, even though the SPTpol measurements are almost two orders of magnitude more precise. The signal $C_\ell^{\alpha\alpha}$ falls off with increasing ℓ, while the precision of the data improves at a comparable rate up to about $\ell \sim 200$. Thus, similar limits are obtained from *Planck* at low ℓ and SPTpol at high ℓ (and also other data sets, with the exception of Polarbear; see app. 2.C). Furthermore, since the signal spectrum drops off more quickly than the measurements' precision beyond $\ell \gtrsim 200$, we do not expect such high ℓ data points to contribute significantly towards our results.

Using the marginalized posterior distribution over the amplitude parameter $\mathcal{A}^2\xi_0$, we derive 95% confidence level upper limits for both data sets. In doing so, we discard the unphysical parameter space with $\mathcal{A}^2\xi_0 < 0$, and we enforce $\mathcal{A}^2\xi_0 \geq 0$. *Planck* (2018) gives $\mathcal{A}^2\xi_0 < 13$ (95% CL) and SPTpol gives $\mathcal{A}^2\xi_0 < 3.7$ (95% CL). The SPTpol limit is tighter, partly because the posterior distribution is slightly skewed toward negative amplitudes, and we take only $\mathcal{A}^2\xi_0 > 0$ to derive the limits. Since we expect $\mathcal{A} = \mathcal{O}(1)$ from UV model building and $\xi_0 = \mathcal{O}(1-10)$ from string network simulations, these limits are already strongly constraining.

These results are in good agreement with a previous study (15) that calculated the posterior probability distribution over $\mathcal{A}^2\xi_0$ and ζ_0 and derived constraints on the amplitude $\mathcal{A}^2\xi_0$ using *Planck* (2015) data (82). See app. 2.C for our analysis of the *Planck* (2015) data. Ref. (15) also presents results for a string network model with a range of loops sizes (2), which we do not repeat here. Instead, we provide birefringence constraints on a collapsing string-wall network below.

2.4.2 Collapsing string-wall network

We now study an axion string-wall network that collapses between recombination and today. We calculate the birefringence power spectrum $C_\ell^{\alpha\alpha}$ by following the procedure described in sec. 2.2.2. Tab. 2.1 shows our priors on the three model parameters: the amplitude parameter $\mathcal{A}^2\xi_0$, the loop length parameter ζ_0, and the axion mass parameter m_a that controls when the network collapses; we fix $N_{\rm dw} = 1$. For $m_a \lesssim 10^{-32.4}$ eV the string network has not yet collapsed in the universe today, and we revert back to the analysis of sec. 2.4.1, whereas raising m_a causes the network to collapse earlier.

Our results are summarized in fig. 2.10, which shows the marginalized posterior probability distribution over the model parameters. The degeneracy direction at $\mathcal{A}^2\xi_0 = 0$ is consistent with the discussion in sec. 2.4.1. The data prefers larger values of the axion mass m_a, and the marginalized posterior is peaked at the cutoff imposed

by the prior $m_a < 10^{-28}$ eV. This is because the data is consistent with the absence of cosmological birefringence, and raising m_a suppresses power at large angular scales, as seen in fig. 2.6.

Using the marginalized posterior distributions we calculate the 95% CL upper limits on the amplitude parameter, which are found to be $\mathcal{A}^2\xi_0 < 55{,}000$ for *Planck* (2018) and $\mathcal{A}^2\xi_0 < 390$ for SPTpol. In comparison with our study of the stable string network from sec. 2.4.1, we see that the amplitude limits are weaker here. This is partially because raising m_a suppresses power at low ℓ and accommodates larger $\mathcal{A}^2\xi_0$. We also note that the *Planck* (2018) limit is weaker than the SPTpol limit here by a factor of ~ 100. The strongly scale-dependent power spectrum has $C_\ell^{\alpha\alpha} \propto \ell^0$ at large angular scales, which suppresses the signal in the range of multipoles from $1 \leq \ell \leq 24$ at which the *Planck* (2018) birefringence measurement is available, leading to a weaker limit on the amplitude $\mathcal{A}^2\xi_0$ as compared with SPTpol. This observation emphasizes the complementarity between all-sky and ground-based measurements of CMB polarization as probes of cosmological birefringence. A detection of anisotropic birefringence on small angular scales ($\ell \sim 100$) without a detection on large angular scales ($\ell \lesssim 10$) would point to a strongly scale-dependent source, and provide evidence for cosmological birefringence from a collapsing axion string-wall network.

2.5 Compatibility with isotropic birefringence measurements

Aside from searches for anisotropic birefringence, various groups (70–73) have recently analyzed all-sky polarization data to search for evidence of isotropic birefringence. In particular, the authors of ref. (73) report a measurement of $\alpha = -0.342^\circ{}^{+0.094^\circ}_{-0.091^\circ}$ (68% CL) using data from *Planck* and *WMAP*.[‖] These analyses provide strong evidence for isotropic birefringence in the CMB at more than 99.9% confidence. Here we

[‖] We have adopted the sign convention where a positive birefringence angle induces a counter clockwise rotation in the plane of polarization. Our convention is opposite to the one used in ref. (73), and we have added a minus sign in reporting their measurement.

address the implications of this measurement for birefringence from axion string-wall networks.

The axion string-wall network produces an anisotropic birefringence signal that is statistically isotropic. A general birefringence map can be decomposed onto spherical harmonics as $\alpha(\hat{\mathbf{n}}) = \sum_{\ell,m} \alpha_{\ell m} Y_{\ell m}(\hat{\mathbf{n}})$, and we are interested in the monopole α_{00}, which corresponds to isotropic birefringence. Averaging α_{00} over an ensemble of universes gives zero, since positive and negative fluctuations are equally likely. Nevertheless, every individual Universe has a nonzero α_{00}. To illustrate this point, we simulate 1000 realizations of the birefringence map for the stable string network model, and plot the distribution over the monopole in fig. 2.11. As expected, the mean is close to zero and the variance is approximately $C_0^{\alpha\alpha}$ that we calculate from theory. The distribution is clearly non-Gaussian, displaying a tighter central distribution, and moreover the circular features in the sky map imply correlations across modes. Nevertheless, a normal distribution with zero mean provides a good approximation. We intend to investigate the non-Gaussian behavior further in future work.

We evaluate the likelihood for isotropic birefringence as follows. We treat $\alpha_{00} = S + N$ as the sum of uncorrelated signal and noise terms. The signal is modeled as a Gaussian random variable with zero mean and variance $C_0^{\alpha\alpha}$ that we calculate in the loop crossing model. The slightly non-Gaussian nature of the α_{00} distribution is neglected for this analysis. The noise is modeled as a Gaussian random variable with zero mean and standard deviation $\sigma_0 = \sqrt{4\pi} \times 0.0925°$; this corresponds to the average uncertainty in the isotropic birefringence measurement from ref. (73), and the factor of $\sqrt{4\pi}$ accounts for the normalization of the spherical harmonics.** We extend our log-likelihood (2.12) to include the $\ell = 0$ mode in this way, assuming it is uncorrelated with the other multipoles, and we evaluate it at $\alpha_{00} = \sqrt{4\pi} \times (-0.342°)$.

**If $\alpha(\hat{\mathbf{n}}) = \sum_{\ell,m} \alpha_{\ell m} Y_{\ell m}(\hat{\mathbf{n}})$ with $Y_{00} = 1/\sqrt{4\pi}$ then the sky-averaged isotropic birefringence angle is $\overline{\alpha} = \int \mathrm{d}^2\hat{\mathbf{n}}\, \alpha(\hat{\mathbf{n}})/4\pi = \alpha_{00}/\sqrt{4\pi}$. We are grateful to Eiichiro Komatsu for pointing out this distinction.

Note that the sign of α_{00} does not provide any constraint on the model; the monopole α_{00} follows a symmetric distribution with zero mean.

We repeat the MCMC analysis and present the results in fig. 2.12. This figure shows the marginalized posterior distribution over the amplitude parameter $\mathcal{A}^2\xi_0$ for the stable string network model. First, taking only the anisotropic birefringence measurements from SPTpol, this data is consistent with the absence of an ALP string network, implying $\mathcal{A}^2\xi_0 < 3.7$ (95% CL), as we also discuss in sec. 2.4.1. Second, the isotropic birefringence measurement alone strongly favors the presence of an ALP string network; the posterior is broad, peaking at $\mathcal{A}^2\xi_0 \approx 40 - 50$ and extending to much larger values (not shown). Third, we show the fit to the joint likelihood, which leads to the measurement $\mathcal{A}^2\xi_0 = 0.5 \pm 1.0$ (68% CL).

The anisotropic and isotropic birefringence measurements are difficult to reconcile in the context of axion-defect-induced birefringence; this is one of the key results of our work. The isotropic birefringence measurement favors a large amplitude parameter $\mathcal{A}^2\xi_0$, which is in conflict with the anisotropic birefringence measurements. There is only a small overlap of the posterior distributions in the tail regions. Although we present results for SPTpol here, the same conclusions can be drawn from the *Planck* (2018) measurement of anisotropic birefringence (or other data sets) instead; however, the wider tail of the *Planck* posterior distribution (see fig. 2.9) leads to a smaller tension.

We have also performed a similar analysis for the collapsing string-wall network model. Since larger m_a suppresses the birefringence signal at low ℓ, an even larger amplitude $\mathcal{A}^2\xi_0$ is required to accommodate the isotropic birefringence measurement at $\ell = 0$. However, this large amplitude comes into sharper tension with the anisotropic birefringence measurements at $\ell > 0$.

2.6 Summary and conclusion

In this work we have studied models of axion-like particles that form a network of cosmic strings and domain walls. We distinguish four model classes in the parameter space spanned by the axion mass m_a and the domain wall number $N_{\rm dw}$: (1) a stable string network that survives in the universe today, (2) a string network that forms domain walls and collapses between recombination and today, (3) a string network that forms stable domain walls between recombination and today, and (4) a string network that forms stable domain walls before recombination.

We calculate the cosmological birefringence signal that these axion string-wall networks imprint on the polarization pattern of CMB radiation via the usual coupling of the axion-like particles to electromagnetism. Using measurements of anisotropic birefringence derived from polarization data taken by various CMB telescopes, we assess the extent to which they are compatible with axion-defect-induced birefringence. All of the measurements are consistent with the absence of birefringence from axion string-wall networks, and we derive constraints on the amplitude of the signal. Our main results are:

- For hyperlight ALP masses $m_a \lesssim 3H_0 \simeq 4 \times 10^{-33}$ eV, we find that SPTpol measurements constrain $\mathcal{A}^2\xi_0 < 3.7$ at 95% CL (assuming $\mathcal{A}^2\xi_0 \geq 0$) where $\mathcal{A} = -\pi f_a g_{a\gamma\gamma}/\alpha_{\rm em}$ parametrizes the strength of the axion-photon coupling, and ξ_0 parametrizes the average total string length in a Hubble volume in units of the Hubble length. In UV extensions of this effective theory, the parameter $\mathcal{A}$ corresponds to an anomaly coefficient, which is model dependent but typically equals an $\mathcal{O}(1)$ rational number. The precise expected value of ξ_0 is a matter of some debate, but broadly speaking $\xi_0 = \mathcal{O}(1-10)$. For instance if $\xi_0 = 30$ then the constraint implies $\mathcal{A} \lesssim 1/4$. Thus, we conclude that SPTpol measurements are already placing meaningful constraints on hyperlight axion-like particles and their UV embedding.

- For ALP masses in the range between $3H_0$ and $3H_{\rm cmb} \simeq 1 \times 10^{-28}$ eV and for $N_{\rm dw} = 1$, the anisotropic birefringence signal is predicted to be strongly scale dependent $\ell(\ell+1)C_\ell^{\alpha\alpha} \propto \ell^2$ for small ℓ (i.e. on large angular scales the birefringence angles are uncorrelated). This is because the ALP string network develops unstable domain walls when $3H \approx m_a$, causing the string-wall network to collapse and shutting off the source of large-scale birefringence. We find that current measurements of CMB polarization provide no evidence for this signal, which allows us to derive constraints on the axion mass m_a and signal amplitude $\mathcal{A}^2\xi_0$. Looking forward to future surveys, this distinctive scale-dependent signal provides a compelling target, since its detection would furnish a measurement of the axion mass scale in the range $3H_0 \lesssim m_a \lesssim 3H_{\rm cmb}$. It may also be accessible to redshift-dependent probes of birefringence (89).

- For larger domain wall numbers $N_{\rm dw} \geq 2$, the formation of domain walls when $3H \approx m_a$ leads to a stable string-wall defect network. We argue that the expected birefringence signal is qualitatively equivalent to the case of a stable string network with $m_a \lesssim 3H_0$. Thus, we do not perform a separate constraint analysis of these stable string-wall networks, but rather expect out constraints from the stable string network discussed in sec. 2.4.1 to carry over. It is interesting that the stable string-wall network can be consistent with overclosure constraints (depending on f_a and $N_{\rm dw}$) even for ALP masses as large as $m_a \sim 10^{-20}$ eV, which is the usual range of ultralight bosonic dark matter. This observation provides motivation to study the connections between ultralight ALP dark matter, astrophysical constraints, and cosmological signatures.

- For hyperlight ALP masses, we find that the measurements of anisotropic birefringence derived from ground-based telescopes such as ACTpol and SPTpol, currently provide the strongest constraints on axion string-wall networks. A plausible explanation is as follows. The birefringence power spectrum $\ell(\ell+$

1)$C_\ell^{\alpha\alpha}$ peaks around $\ell_p \sim 40/\zeta_0$ before decreasing again at larger multipoles. This translates to $C_{\ell_p}^{\alpha\alpha} \sim (\mathcal{A}^2\xi_0\,\zeta_0^2)\,(4\times10^{-5})$. Since the current ACTpol and SPTpol measurements have data points around this region with strongest precision ($\sigma_\ell \lesssim 10^{-4}$ deg^2), they end up providing the most stringent constraints.

- Among the various data sets that we consider, we find that anisotropic birefringence measurements derived from SPTpol data yield the strongest constraints on axion-defect-induced birefringence. This is partly because of two downward fluctuating data points at $\ell = 120$ and $\ell = 160$, which skew the amplitude distribution towards negative values, leading to a tighter 95% CL upper limit on $\mathcal{A}^2\xi_0$ assuming $\mathcal{A}^2\xi_0 \geq 0$.

- We assess the extent to which recent measurements of a nonzero isotropic birefringence are consistent with constraints on anisotropic birefringence in the context of an axion-defect-induced signal. We find that it is somewhat difficult for a stable string network (or stable string-wall network) to induce a birefringence signal that is compatible with the isotropic measurement and the lack of an anisotropic signal. The situation is further exacerbated for the collapsing string-wall network models due to additional reduced power on large angular scales.

Finally, let us remark on potential directions for extending the computational framework in which our results have been derived. The rich dynamics of a topological defect network present a challenge toward deriving phenomenological observables. In the work presented here, we have used the loop-crossing model (see sec. 2.2) to reduce the complex network down to a manageable number of degrees of freedom with which we can calculate a birefringence signal. The loop crossing model does not capture certain features that an axion-string wall network is expected to exhibit: the finite duration of domain wall formation and network collapse around $3H \sim m_a$ (for $N_{\rm dw} = 1$), or transition into a new scaling solution (for $N_{\rm dw} \geq 2$); and the gradual

change in the axion field nearby to a string loop. This last feature is expected to impact the low-ℓ power spectrum, since the axion field does not change by the full asymptotic amount of $\sim 2\pi f_a$, leading to only $\alpha < \mathcal{A}\alpha_{\rm em}$ during the last few e-foldings. Or in the case of stable walls, there aren't $\sim N_{\rm dw}$ wall crossings in these last few e-foldings before the CMB light reaches us in the present. It is important to understand how each of these features affects the birefringence signal in order to derive robust limits on axion string-wall networks from next-generation surveys.

Acknowledgements

We are grateful to Anthony Challinor, Richard Battye, Steve Gratton, Wayne Hu, Toshiya Namikawa, Ippei Obata, and Blake Sherwin for valuable discussions. We would particularly like to thank Joel Meyers and Levon Pogosian for guidance in the early stages of this project, and Eiichiro Komatsu for his timely feedback as we were finalizing the project. R.H, M.J., and A.J.L. are supported in part by the National Science Foundation under Award No. PHY-2114024. MA is supported by a DOE grant DOE-0000250746.

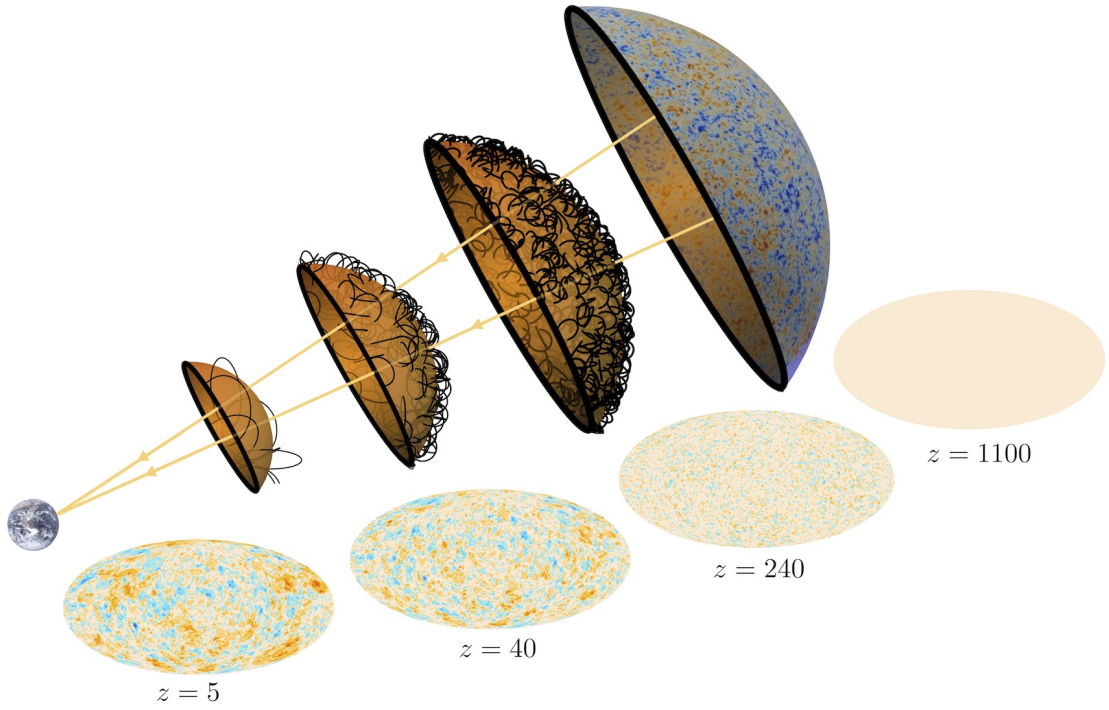


Figure 2.4 : An illustration of the loop-crossing model that we use to calculate the axion-defect-induced birefringence signal. The string-wall network is modeled as a collection of randomly oriented circular loops. The abundance and radius of the loops evolves in time, tracking the Hubble scale; parameter ξ_0 controls the number of loops per Hubble volume and parameter ζ_0 controls the loop size in Hubble units. Photons crossing through a loop experience birefringence $\alpha = \pm\mathcal{A}\,\alpha_{\rm em}$, and multiple loop crossings add incoherently like a random walk. The all-sky maps show the birefringence angle α that has accumulated up to the stated redshift based on a typical realization of the defect network.

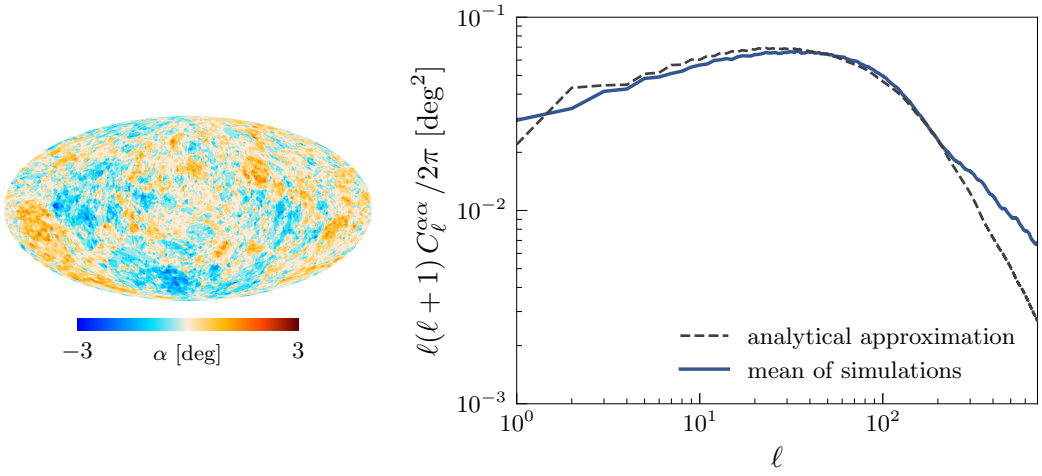


Figure 2.5 : The expected birefringence signal due to a string network that survives until today. We take $\mathcal{A}^2\xi_0 = 1$ and $\zeta_0 = 1$. *Left:* A sample sky map of the birefringence angle $\alpha(\hat{\mathbf{n}})$. *Right:* The angular power spectrum of the birefringence angle $C_\ell^{\alpha\alpha}$. The dashed black curve is our analytical approximation (2.7), while the blue curve is the mean of a suite of 1000 simulations of the loop crossing model. The shaded region shows the corresponding 68% central containment interval exhibiting cosmic variance.

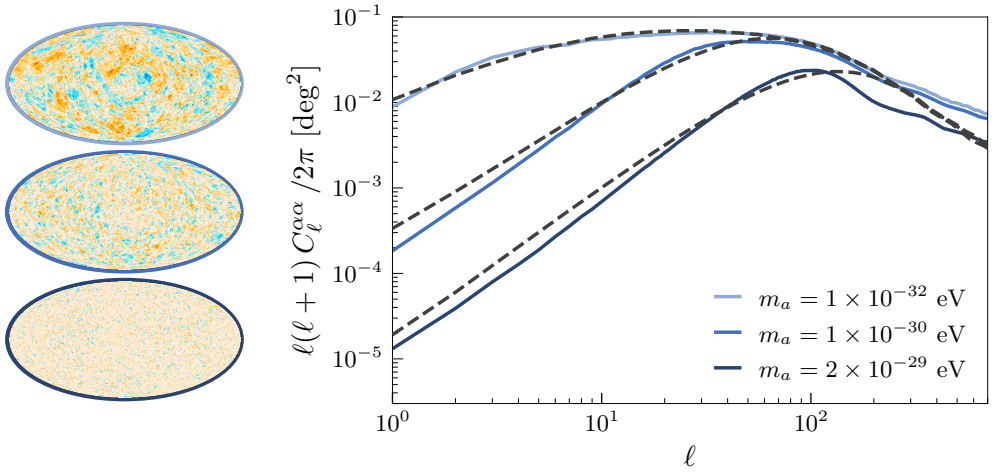


Figure 2.6 : Same as fig. 2.5 but for a collapsing string-wall network ($N_{\rm dw} = 1$). Several curves for different axion mass m_a are shown, corresponding to different collapse redshifts z_c given by eq. (2.9). For each m_a, we also show the corresponding all-sky birefringence map on the left. Increasing m_a causes the network to collapse earlier, and suppresses power at small ℓ (large angular scales).

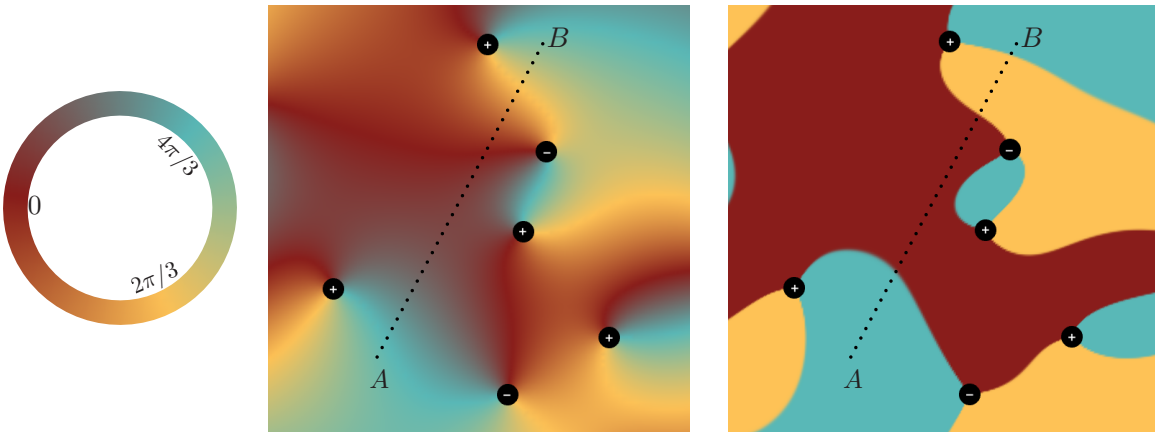


Figure 2.7 : An illustration of the axion field $a/f_a \in [0, 2\pi)$ in the vicinity of several parallel long strings (vortices) with and without walls. *Left panel:* Color legend indicating axion field values on a circle. *Middle panel:* The field around a long string in cylindrical coordinates obeys $\boldsymbol{\nabla} a = \pm\hat{\boldsymbol{\theta}}/\rho$, and several strings are superimposed to form the middle image. *Right panel:* Long strings are connected to $N_{\rm dw} = 3$ domain walls corresponding to sharp field gradients between minima at $a/f_a = 0, 2\pi/3$, and $4\pi/3$.

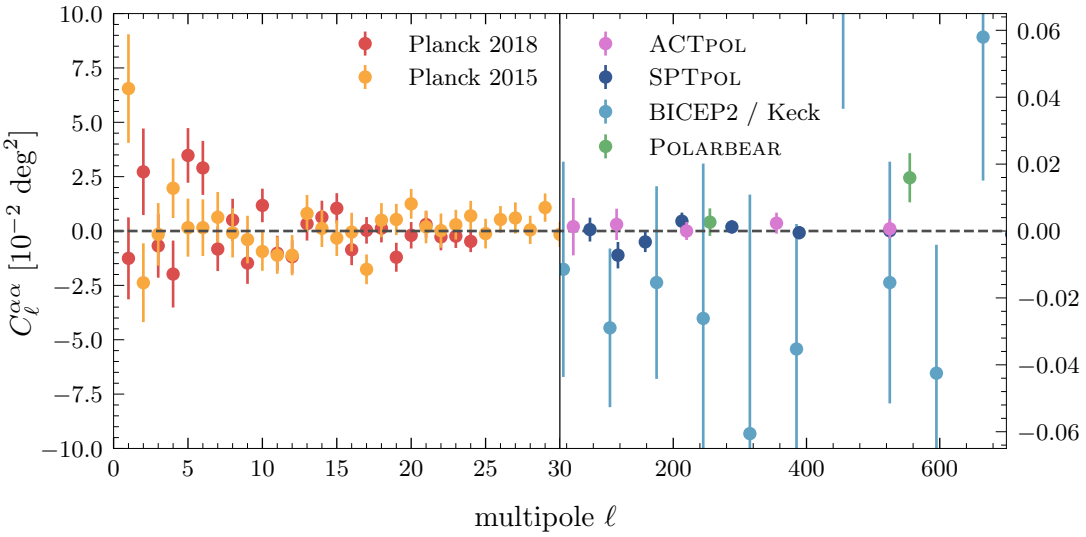


Figure 2.8 : Measurements of anisotropic cosmological birefringence with data from various CMB telescopes: *Planck* (2018) (80) (see also ref. (81)), *Planck* (2015) (82), ACTPOL (83), SPTPOL (84), BICEP2/*Keck Array* (85), and POLARBEAR (86). Note the different scales for $\ell < 30$ and $\ell > 30$. We do not show the measurements from *Planck* (2015) for $\ell > 30$ due to the large error bars. Additional data is available at higher multipoles, which is also not shown in this summary figure. These measurements are consistent with the absence of anisotropic cosmological birefringence.

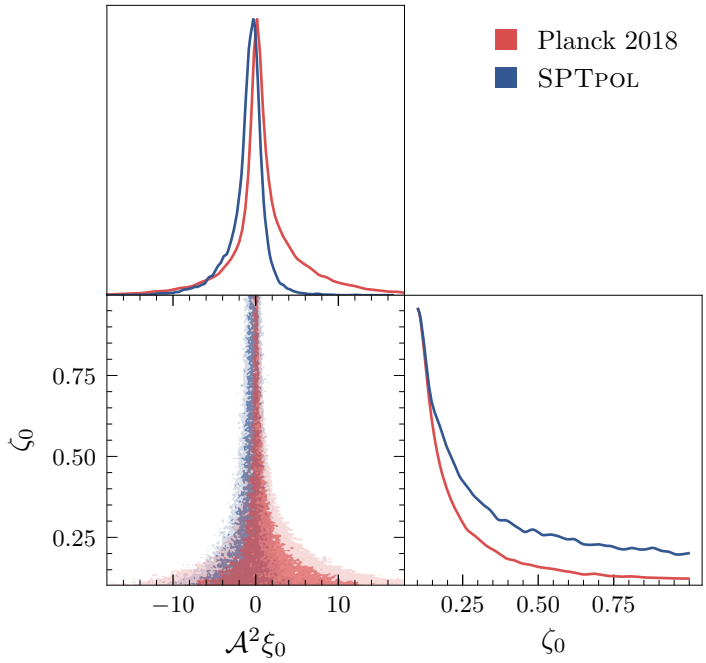


Figure 2.9 : For the stable string network we show joint posteriors obtained from our MCMC simulations using anisotropic birefringence measurements derived from *Planck* (2018) (80) and SPTpol (84) observations. Light and dark contours show 95% and 68% CL regions respectively.

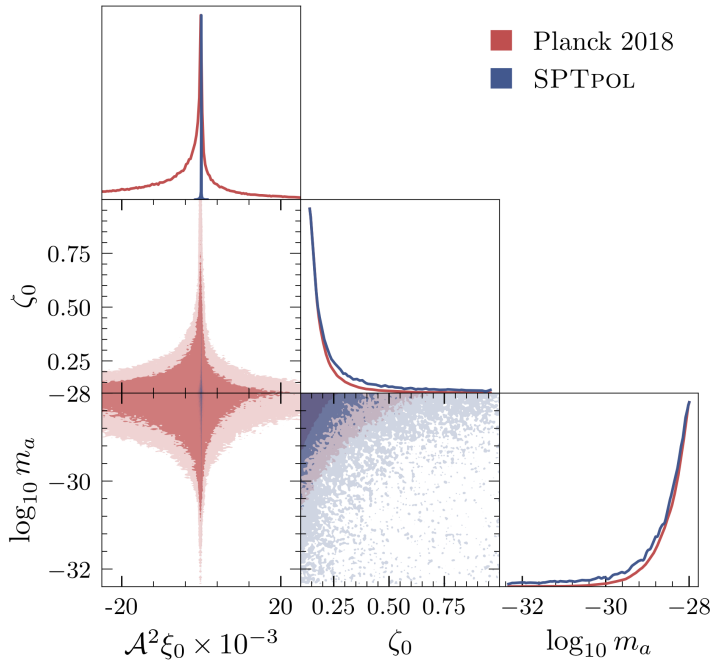


Figure 2.10 : Joint posteriors for the collapsing string-wall network. Data and contour shading is the same as fig. 2.9.

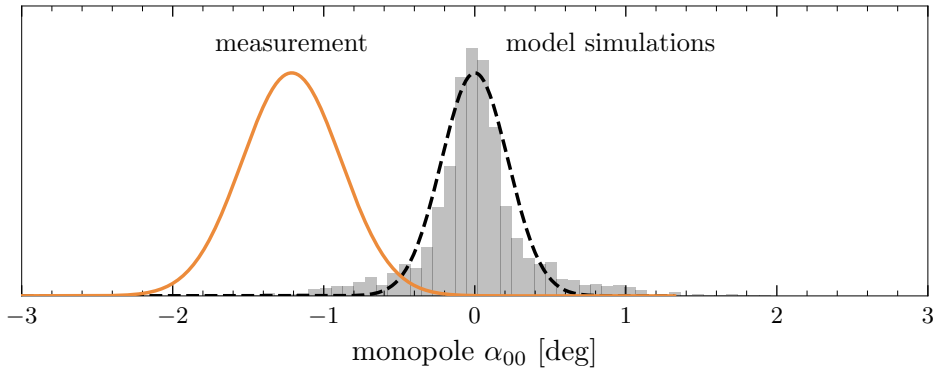


Figure 2.11 : Distribution over the monopole α_{00} of the birefringence map $\alpha(\hat{\mathbf{n}})$. To generate the gray histogram we simulate 1000 sky maps using the loop crossing model with stable strings for parameters $\mathcal{A} = 1$, $\xi_0 = 0.5$, $\mathcal{A}^2\xi_0 = 0.5$, and $\zeta_0 = 1$. These parameters are chosen to maximize the "Isotropic BF + SPTpol" distribution shown on fig. 2.12. We approximate the simulated distribution by a normal distribution (black-dashed curve) with zero mean and variance $C_0^{\alpha\alpha} = (0.23^\circ)^2$ calculated from the model. The orange curve shows the measurement of isotropic birefringence from ref. (73), which we model as a normal distribution with mean $\sqrt{4\pi} \times (-0.342^\circ) = -1.21^\circ$ and standard deviation $\sqrt{4\pi} \times (0.0925^\circ) = 0.328^\circ$.

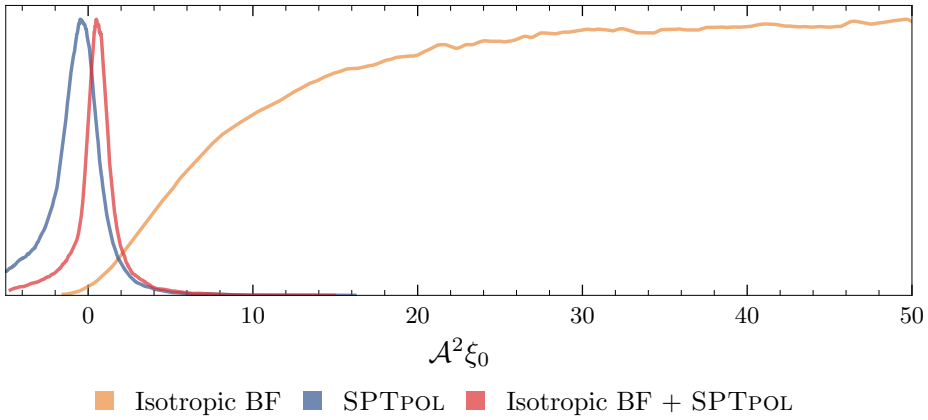


Figure 2.12 : Assessing the compatibility of isotropic and anisotropic birefringence measurements. We show the marginalized posterior over the amplitude parameter $\mathcal{A}^2\xi_0$ for the stable string network model. The isotropic birefringence measurement favors a nonzero amplitude to fit the monopole α_{00} (orange), whereas the anisotropic measurements using SPTpol data constrain the amplitude around zero (blue). The small overlap of the two distributions illustrates the difficulty in accommodating both measurements from axion-defect-induced birefringence. A joint likelihood combining both measurements (red) favors $\mathcal{A}^2\xi_0 = 0.5 \pm 1.0$ at 68% CL.

Appendix

2.A Simulating loop crossing model

Here we present a step-by-step procedure of how we simulate the loop crossing model to generate birefringence maps using HEALpix (90).

1. **Initialize HEALPix map.** We begin by making a HEALpix map with pixel parameter $N_{\rm side} = 2048$, and initialize a null array of length $N_{\rm pix} = 12N_{\rm side}^2$.

2. **Choose time steps.** For time evolution, we pick redshifts in logarithmic intervals from recombination $z_{\rm cmb} = 1100$ till the present $z = 0$, in the following manner

$$1 + z_n = (1 + z_{\rm cmb}) \left(\frac{1 + z_{\rm final}}{1 + z_{\rm cmb}} \right)^{(n-X)/N_{\rm steps}} . \tag{2.13}$$

   Here, the index n goes from 1 to $N_{\rm steps} = 28$, and X is a random variable sampled uniformly between $-1/2$ and $1/2$ at the start of the simulation (subsequent steps then use the same value). This ensures each simulation samples different redshift steps over many simulations. This reproduces a continuous evolution, when averaged over a large ensemble of simulations.

3. **Run simulation.** For each redshift step z_n:

   3.1. We populate the CMB light cone (from z_{n-1} to z_n) with circular loops, all having the same dimensionless radius ζ. The number of loops is determined by calculating the average number of loops N and randomly selecting each pixel to be the center of a loop with probability $N/N_{\rm pix}$. Assuming the

network is in scaling, the average number density is

$$n = \frac{\rho}{E} = \frac{\xi_0 H^3}{2\pi\zeta}\,. \tag{2.14}$$

Here $\rho = \xi_0 \mu H^2$ is the energy density of the scaling network, $E = \mu 2\pi\zeta_0 H^{-1}$ is the energy of a circular string loop with radius $\zeta_0 H^{-1}$ and tension μ. The average number of loops on the CMB light cone from z_{n-1} to z_n is therefore

$$\begin{aligned} N &= \int \mathrm{d}V_{\text{lightcone}}\,(\text{number density}) \\ &= \int_{z_n}^{z_{n-1}} \mathrm{d}z\, \frac{\xi_0 (aH)^3}{2\pi\zeta_0}\, 4\pi s(z)^2 H^{-1}(z)\,. \end{aligned} \tag{2.15}$$

Here $s(z)$ is the comoving distance from an observer on Earth to the centers of loops through which CMB photons crossed at redshift z.

3.2. Out of N_{pix}, we randomly select N_T pixels on the sphere for loop centers. Every loop is given a random orientation and assigned a random winding number w (equal to $+1$ or -1 with equal probability). For uniformly oriented loops, $\cos\Theta \sim \mathrm{U}(0,1)$ and $\Phi \sim \mathrm{U}(0, 2\pi)$, where (Θ, Φ) are polar and azimuthal angles measured relative to the normal vector of the loop's center pixel.

3.3. After populating the CMB light cone with N_T loops, for every loop we find the region/pixels bounded by its spherical projection onto the sphere. (The function `healpy.query_polygon()` was used for this purpose). All the pixels in the region are assigned the value $w\mathcal{A}\alpha_{\text{em}}$ where w is the winding number of the loop.

3.4. We repeat this procedure until $n = N_{\text{steps}}$, and perform many such simulations for ensemble averaging.

2.B Statistical estimator for anisotropic birefringence

This appendix includes supplementary material regarding: the effect of birefringence on the CMB, a statistical estimator that may be used to measure birefringence from CMB polarization data, and a demonstration of how the estimator works using simulated data.

The effect of birefringence on the CMB is most easily understood in position space where $\hat{\mathbf{n}}$ is a unit vector at some point on the sky. If $\tilde{T}(\hat{\mathbf{n}})$, $\tilde{Q}(\hat{\mathbf{n}})$, and $\tilde{U}(\hat{\mathbf{n}})$ denote the would-be CMB temperature and polarization sky maps in the absence of birefringence, then an anisotropic birefringence angle $\alpha(\hat{\mathbf{n}})$ mixes Q- and U-mode polarization (91), giving rise to the observable sky maps

$$T = \tilde{T} \tag{2.16a}$$

$$Q = \tilde{Q}\cos 2\alpha - \tilde{U}\sin 2\alpha \tag{2.16b}$$

$$U = \tilde{Q}\sin 2\alpha + \tilde{U}\cos 2\alpha \ . \tag{2.16c}$$

If $\tilde{T}$, $\tilde{Q}$, and $\tilde{U}$ were known, then eq. (2.16) would allow α to be extracted from measurements of T, Q, and U. Of course, the would-be temperature and polarization anisotropies of the CMB cannot be calculated (nor measured), but the *statistical properties* of these fields are calculable. This observation motivates one to define a birefringence estimator, which reproduces the true birefringence as a statistical average.

Several statistical estimators of anisotropic birefringence have been proposed in the literature. Here we discuss a particular set of estimators that have been used in recent CMB data analyses. These are similar to the quadratic estimators proposed by Hu and Okamoto (92) for studies of CMB weak lensing, building on which the authors of ref. (67–69) constructed another set of estimators for studies of CMB birefringence.

Following the notation of ref. (15), the birefringence estimators in the flat-sky

approximation (68) can be written as

$$\hat{\alpha}_{XY}(\mathbf{L}) = \lambda_{XY}(\mathbf{L}) \int \frac{\mathrm{d}^2 l_1}{(2\pi)^2} X(\mathbf{l}_1) Y^*(\mathbf{l}_2) F_{XY}(\mathbf{l}_1, \mathbf{l}_2)\big|_{\mathbf{l}_2 = \mathbf{l}_1 - \mathbf{L}} \tag{2.17}$$

where X and Y stand for temperature T, parity-even E-mode polarization, or parity-odd B-mode polarization. In fact there are five different birefringence estimators corresponding to the choice of $XY \in \{EE, BB, TE, TB, EB\}$ (since $XY = TT$ is trivial). In the flat-sky approximation, $\mathbf{L}, \mathbf{l}_1, \mathbf{l}_2 \in \mathbb{R}^2$ are the analogs of the spherical harmonic integer indices (ℓ, m), and eq. (2.17) assumes $\mathbf{L} \neq 0$. The estimator integrates over $X(\mathbf{l})$ and $Y(\mathbf{l})$, which represent the observed temperature or polarization maps, weighting them by the mode coupling coefficients

$$F_{XY}(\mathbf{l}_1, \mathbf{l}_2) = \begin{cases} \frac{f_{XY}(\mathbf{l}_1, \mathbf{l}_2)}{(1+\delta_{XY}) C_{l_1}^{XX} C_{l_2}^{YY}} & , \quad XY \neq TE \\ \frac{C_{l_1}^{YY} C_{l_2}^{XX} f_{XY}(\mathbf{l}_1, \mathbf{l}_2) - C_{l_1}^{XY} C_{l_2}^{XY} f_{XY}(\mathbf{l}_2, \mathbf{l}_1)}{C_{l_1}^{XX} C_{l_2}^{YY} C_{l_2}^{XX} C_{l_1}^{YY} - (C_{l_1}^{XY} C_{l_2}^{XY})^2} & , \quad XY = TE \end{cases} . \tag{2.18}$$

Here C_ℓ^{XY} are the predicted power spectra (or cross-correlation spectra if $X \neq Y$), δ_{XY} is the Kronecker delta, and the response functions $f_{XY}(\mathbf{l}_1, \mathbf{l}_2)$ are given in tab. 2.B.1. Finally the normalization coefficient

$$[\lambda_{XY}(\mathbf{L})]^{-1} = \int \frac{\mathrm{d}^2 l_1}{(2\pi)^2} f_{XY}(\mathbf{l}_1, \mathbf{l}_2) F_{XY}(\mathbf{l}_1, \mathbf{l}_2)\big|_{\mathbf{l}_2 = \mathbf{L} - \mathbf{l}_1} \tag{2.19}$$

ensures that $\hat{\alpha}_{XY}(\mathbf{L})$ is an unbiased estimator.

We seek to demonstrate how the estimator (2.17) reconstructs a known birefringence map $\alpha(\hat{\mathbf{n}})$ from an ensemble of simulated temperature and polarization maps. This is done with the following procedure.

1. We begin by calculating the angular power spectra $\tilde{C}_\ell^{TT}$, $\tilde{C}_\ell^{EE}$, and $\tilde{C}_\ell^{TE}$ using CAMB (93) (we take $\tilde{C}_\ell^{BB} = \tilde{C}_\ell^{TB} = \tilde{C}_\ell^{EB} = 0$). Assuming Gaussian fluctuations, we construct all-sky temperature and polarization maps, $\tilde{T}(\hat{\mathbf{n}})$, $\tilde{Q}(\hat{\mathbf{n}})$, and $\tilde{U}(\hat{\mathbf{n}})$ using `healpy` (the Python wrapper for `HEALPix`).

2. Next, we generate a birefringence map $\alpha(\hat{\mathbf{n}})$ using the procedure outlined in app. 2.A; since this is just a proof of principle demonstration, the parameters of the string network model are not particularly important.

XY	$f_{XY}(\mathbf{l}_1, \mathbf{l}_2)$
TT	0
TE	$-2\tilde{C}_{l_1}^{TE} \sin 2\varphi_{12}$
TB	$2\tilde{C}_{l_1}^{TE} \cos 2\varphi_{12}$
EE	$-2\big(\tilde{C}_{l_1}^{EE} - \tilde{C}_{l_2}^{EE}\big) \sin 2\varphi_{12}$
EB	$2\big(\tilde{C}_{l_1}^{EE} - \tilde{C}_{l_2}^{BB}\big) \cos 2\varphi_{12}$
BB	$-2\big(\tilde{C}_{l_1}^{BB} - \tilde{C}_{l_2}^{BB}\big) \sin 2\varphi_{12}$

Table 2.B.1 : Response functions. Here $\varphi_{ij} = \varphi_{\mathbf{l}_i} - \varphi_{\mathbf{l}_j}$ where $\cos\varphi_{\mathbf{l}} \equiv \hat{\mathbf{n}} \cdot \hat{\mathbf{l}}$. In `symlens` $\cos 2\varphi_{12}$ is represented symbolically by `symlens.cos2t12`.

3. Using the simulated birefringence map, we transform the temperature and polarization maps according to eq. (2.16), and convert Q and U maps to E and B maps.

4. In order to apply the flat-sky estimator, we isolate some small section of the all-sky maps and port it into the `pixell` (94) software.

5. With the chosen small section, we calculate the birefringence estimator $\hat{\alpha}_{EB}(\mathbf{L})$. We use `symlens` (95) to evaluate the integrals in eq. (2.17).

6. Finally we perform an inverse 2D Fourier transform on $\hat{\alpha}_{EB}(\mathbf{L})$ to obtain the reconstructed birefringence map $\hat{\alpha}_{EB}(\hat{\mathbf{n}})$.

Fig. 2.B.1 shows the result of the above procedure. The left panel shows the 'true' birefringence map $\alpha(\hat{\mathbf{n}})$, which acts on the CMB polarization anisotropies according to eq. (2.16). The middle and right panels show the reconstructed birefringence map $\hat{\alpha}_{EB}(\hat{\mathbf{n}})$ using only 1 and an average over 20 realizations of the CMB, respectively. Upon averaging over many CMB realizations, the estimator should converge to the true birefringence map (for an unbiased estimator). We note that even a single CMB

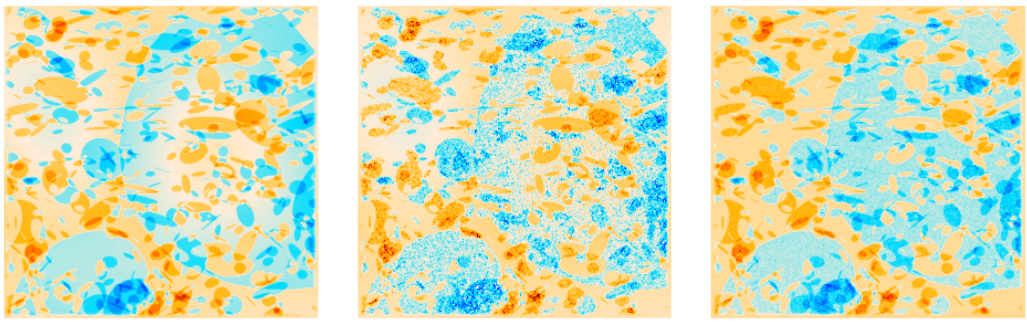


Figure 2.B.1 : A demonstration of how the statistical estimator $\hat{\alpha}_{EB}(\hat{\mathbf{n}})$ from eq. (2.17) reconstructs a birefringence map $\alpha(\hat{\mathbf{n}})$. *Left:* The 'true' birefringence map $\alpha(\hat{\mathbf{n}})$. *Middle:* The reconstructed birefringence map $\hat{\alpha}_{EB}(\hat{\mathbf{n}})$ obtained from a single realization of the CMB temperature and polarization maps. *Right:* Reconstructed birefringence map $\hat{\alpha}_{EB}(\hat{\mathbf{n}})$, averaged over a suite of 10 realizations. Our implementation of the estimator in this figure introduces a multiplicative bias (not perceptible here) that scales inversely with the map width.

realization leads to a reliable reconstruction that captures many qualitative properties of the true map, e.g. scale and shape of loop-like features.

2.C Alternative birefringence data

Measurements of anisotropic cosmological birefringence are available from the various CMB telescopes; see fig. 2.8. Using each of these data sets individually, we derive constraints on the axion string-wall network models. Our results are summarized in this appendix.

Fig. 2.C.1 shows the marginalized probability distribution over the amplitude parameter $\mathcal{A}^2\xi_0$ for the stable string network (left panel) and the collapsing string-wall network (right panel). Each curve corresponds to a measurement of anisotropic birefringence using data from a different CMB telescope; see fig. 2.8. The corresponding 95% CL upper limits on the amplitude parameter (assuming $0 < \mathcal{A}^2\xi_0$) are summa-

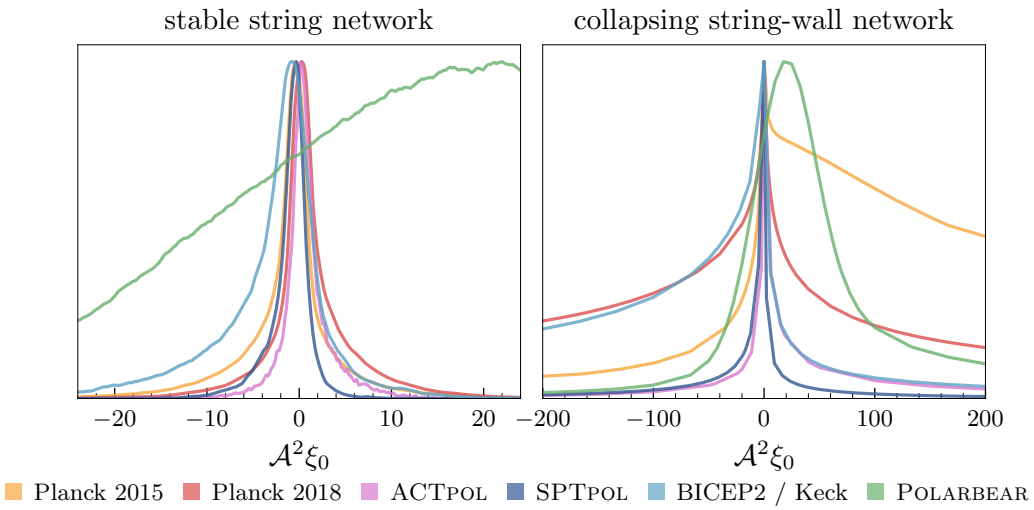


Figure 2.C.1 : Marginalized posterior on the amplitude parameter $\mathcal{A}^2\xi_0$ for a network of stables strings (left panel) and a collapsing string-wall network (right panel).

	stable strings	collapsing string-wall
Planck (2015):	$\mathcal{A}^2\xi_0 < 13$	$\mathcal{A}^2\xi_0 < 55,000$
Planck (2018):	$\mathcal{A}^2\xi_0 < 13$	$\mathcal{A}^2\xi_0 < 18,000$
ACTpol:	$\mathcal{A}^2\xi_0 < 7.1$	$\mathcal{A}^2\xi_0 < 1,100$
SPTpol:	$\mathcal{A}^2\xi_0 < 3.7$	$\mathcal{A}^2\xi_0 < 390$
BICEP2/*Keck*:	$\mathcal{A}^2\xi_0 < 11$	$\mathcal{A}^2\xi_0 < 3,200$
Polarbear:	$\mathcal{A}^2\xi_0 < 81$	$\mathcal{A}^2\xi_0 < 3,300$

Table 2.C.1 : Upper limits at 95% CL on the amplitude parameter $\mathcal{A}^2\xi_0$ derived from measurements of anisotropic birefringence using data from various CMB telescopes.

rized in tab. 2.C.1.

From these results, one can see that each of the data sets is consistent with $\mathcal{A}^2\xi_0 = 0$ at the 1σ level, corresponding to the absence of axion-defect-induced birefringence. SPTpol provides the strongest constraints on the amplitude of the signal, for both the

stable string network model and the collapsing string-wall network model. An upward fluctuation in the POLARBEAR data (86) leads to a $\sim 1\sigma$ preference for $\mathcal{A}^2\xi_0 > 0$, whereas several downward fluctuations in the BICEP2/*Keck Array* data (85) broaden the distribution toward negative amplitudes. For the collapsing string-wall network, the *Planck* data is less constraining, consistent with the discussion in sec. 2.4.

Using a birefringence measurement (82) derived from *Planck* (2015) data, we find that the network of stable strings is constrained by $\mathcal{A}^2\xi_0 < 13$ at 95% CL. This particular data set was also analyzed by another recent study (15), and our results are in good agreement. Note that ref. (15) presents a constraint $\mathcal{A}^2\xi_0 < 8.0$ at 95% CL, which is derived allowing $\mathcal{A}^2\xi_0 < 0$. To compare with our result, we digitize the marginalized posterior from fig. 5 of ref. (15) and calculate $\mathcal{A}^2\xi_0 < 12.6$ at 95% CL when imposing $\mathcal{A}^2\xi_0 > 0$, which is in excellent agreement with our result.

Bibliography

[1] Jain, Mudit and Hagimoto, Ray and Long, Andrew J. and Amin, Mustafa A., *Searching for axion-like particles through CMB birefringence from string-wall networks*, *JCAP* **10** (2022) 090, [2208.08391].

[2] Jain, Mudit and Long, Andrew J. and Amin, Mustafa A., *CMB birefringence from ultralight-axion string networks*, *JCAP* **05** (2021) 055, [2103.10962].

[3] PLANCK collaboration, Akrami, Y. and others, *Planck 2018 results. X. Constraints on inflation*, *Astron. Astrophys.* **641** (2020) A10, [1807.06211].

[4] Carroll, Sean M. and Field, George B. and Jackiw, Roman, *Limits on a Lorentz and Parity Violating Modification of Electrodynamics*, *Phys. Rev. D* **41** (1990) 1231.

[5] Carroll, Sean M. and Field, George B., *The Einstein equivalence principle and the polarization of radio galaxies*, *Phys. Rev. D* **43** (1991) 3789.

[6] Harari, Diego and Sikivie, Pierre, *Effects of a Nambu-Goldstone boson on the polarization of radio galaxies and the cosmic microwave background*, *Phys. Lett. B* **289** (1992) 67–72.

[7] Carroll, Sean M., *Quintessence and the rest of the world*, *AIP Conf. Proc.* **478** (1999) 291–294.

[8] Fedderke, Michael A. and Graham, Peter W. and Rajendran, Surjeet, *Axion Dark Matter Detection with CMB Polarization*, *Phys. Rev. D* **100** (2019) 015040, [1903.02666].

[9] Fujita, Tomohiro and Minami, Yuto and Murai, Kai and Nakatsuka, Hiromasa, *Probing axionlike particles via cosmic microwave background polarization*, *Phys. Rev. D* **103** (2021) 063508, [2008.02473].

[10] Obata, Ippei, *Implications of the Cosmic Birefringence Measurement for the Axion Dark Matter Search*, 2108.02150.

[11] Nakatsuka, Hiromasa and Namikawa, Toshiya and Komatsu, Eiichiro, *Is cosmic birefringence due to dark energy or dark matter? A tomographic approach*, *Phys. Rev. D* **105** (2022) 123509, [2203.08560].

[12] Komatsu, Eiichiro, *New physics from the polarized light of the cosmic microwave background*, *Nature Rev. Phys.* **4** (2022) 452–469, [2202.13919].

[13] Agrawal, Prateek and Hook, Anson and Huang, Junwu, *A CMB Millikan experiment with cosmic axiverse strings*, *JHEP* **07** (2020) 138, [1912.02823].

[14] Takahashi, Fuminobu and Yin, Wen, *Kilobyte Cosmic Birefringence from ALP Domain Walls*, 2012.11576.

[15] Yin, Weichen Winston and Dai, Liang and Ferraro, Simone, *Probing cosmic strings by reconstructing polarization rotation of the cosmic microwave background*, 2111.12741.

[16] Kitajima, Naoya and Kozai, Fumiaki and Takahashi, Fuminobu and Yin, Wen, *Power spectrum of domain-wall network and its implications for isotropic and anisotropic cosmic birefringence*, 2205.05083.

[17] Kallosh, Renata and Linde, Andrei D. and Linde, Dmitri A. and Susskind, Leonard, *Gravity and global symmetries*, *Phys. Rev. D* **52** (1995) 912–935, [hep-th/9502069].

[18] Vafa, Cumrun, *The String landscape and the swampland*, hep-th/0509212.

[19] Witten, Edward, *Symmetry and Emergence*, *Nature Phys.* **14** (2018) 116–119, [1710.01791].

[20] Brennan, T. Daniel and Carta, Federico and Vafa, Cumrun, *The String Landscape, the Swampland, and the Missing Corner*, *PoS* **TASI2017** (2017) 015, [1711.00864].

[21] Palti, Eran, *The Swampland: Introduction and Review*, *Fortsch. Phys.* **67** (2019) 1900037, [1903.06239].

[22] Eichhorn, Astrid and Pauly, Martin, *Constraining power of asymptotic safety for scalar fields*, *Phys. Rev. D* **103** (2021) 026006, [2009.13543].

[23] Alvey, James and Escudero Abenza, Miguel, *Constraints on global symmetry breaking in quantum gravity from cosmic birefringence measurements*, *Phys. Lett. B* **823** (2021) 136752, [2106.04226].

[24] Dine, Michael and Seiberg, N. and Wen, X. G. and Witten, Edward, *Nonperturbative Effects on the String World Sheet*, *Nucl. Phys. B* **278** (1986) 769–789.

[25] Gibbons, Gary W. and Green, Michael B. and Perry, Malcolm J., *Instantons and seven-branes in type IIB superstring theory*, *Phys. Lett. B* **370** (1996) 37–44, [hep-th/9511080].

[26] Becker, Katrin and Becker, Melanie and Strominger, Andrew, *Five-branes, membranes and nonperturbative string theory*, *Nucl. Phys. B* **456** (1995) 130–152, [hep-th/9507158].

[27] Svrcek, Peter and Witten, Edward, *Axions In String Theory*, *JHEP* **06** (2006) 051, [hep-th/0605206].

[28] Arvanitaki, Asimina and Dimopoulos, Savas and Dubovsky, Sergei and Kaloper, Nemanja and March-Russell, John, *String Axiverse*, *Phys. Rev. D* **81** (2010) 123530, [0905.4720].

[29] Alvey, James and Escudero, Miguel, *The axion quality problem: global symmetry breaking and wormholes*, *JHEP* **01** (2021) 032, [2009.03917].

[30] Vilenkin, A. and Everett, A. E., *Cosmic Strings and Domain Walls in Models with Goldstone and PseudoGoldstone Bosons*, *Phys. Rev. Lett.* **48** (1982) 1867–1870.

[31] Vilenkin, A. and Shellard, E. P. S., *Cosmic Strings and Other Topological Defects*. Cambridge University Press, 2000.

[32] March-Russell, John and Tillim, Hannah, *Axiverse Strings*, 2109.14637.

[33] Kibble, T. W. B., *Topology of Cosmic Domains and Strings*, *J. Phys. A* **9** (1976) 1387–1398.

[34] Kibble, T. W. B., *Some Implications of a Cosmological Phase Transition*, *Phys. Rept.* **67** (1980) 183.

[35] Sikivie, P., *Of Axions, Domain Walls and the Early Universe*, *Phys. Rev. Lett.* **48** (1982) 1156–1159.

[36] Harari, Diego and Sikivie, P., *On the Evolution of Global Strings in the Early Universe*, *Phys. Lett. B* **195** (1987) 361–365.

[37] Chang, Sanghyeon and Hagmann, C. and Sikivie, P., *Studies of the motion and decay of axion walls bounded by strings*, *Phys. Rev. D* **59** (1999) 023505, [hep-ph/9807374].

[38] Hagmann, C. and Chang, Sanghyeon and Sikivie, P., *Axion radiation from strings*, *Phys. Rev. D* **63** (2001) 125018, [hep-ph/0012361].

[39] Yamaguchi, Masahide and Kawasaki, M. and Yokoyama, Jun'ichi, *Evolution of axionic strings and spectrum of axions radiated from them*, *Phys. Rev. Lett.* **82** (1999) 4578–4581, [hep-ph/9811311].

[40] Yamaguchi, Masahide and Yokoyama, Jun'ichi, *Quantitative evolution of global strings from the Lagrangian view point*, *Phys. Rev. D* **67** (2003) 103514, [hep-ph/0210343].

[41] Hiramatsu, Takashi and Kawasaki, Masahiro and Sekiguchi, Toyokazu and Yamaguchi, Masahide and Yokoyama, Jun'ichi, *Improved estimation of radiated axions from cosmological axionic strings*, *Phys. Rev. D* **83** (2011) 123531, [1012.5502].

[42] Hiramatsu, Takashi and Kawasaki, Masahiro and Saikawa, Ken'ichi and Sekiguchi, Toyokazu, *Production of dark matter axions from collapse of string-wall systems*, *Phys. Rev. D* **85** (2012) 105020, [1202.5851].

[43] Kawasaki, Masahiro and Saikawa, Ken'ichi and Sekiguchi, Toyokazu, *Axion dark matter from topological defects*, *Phys. Rev. D* **91** (2015) 065014, [1412.0789].

[44] Lopez-Eiguren, Asier and Lizarraga, Joanes and Hindmarsh, Mark and Urrestilla, Jon, *Cosmic Microwave Background constraints for global strings and global monopoles*, *JCAP* **07** (2017) 026, [1705.04154].

[45] Hindmarsh, Mark and Lizarraga, Joanes and Lopez-Eiguren, Asier and

Urrestilla, Jon, *Scaling Density of Axion Strings*, *Phys. Rev. Lett.* **124** (2020) 021301, [1908.03522].

[46] Hindmarsh, Mark and Lizarraga, Joanes and Lopez-Eiguren, Asier and Urrestilla, Jon, *Approach to scaling in axion string networks*, 2102.07723.

[47] Gorghetto, Marco and Hardy, Edward and Villadoro, Giovanni, *Axions from Strings: the Attractive Solution*, *JHEP* **07** (2018) 151, [1806.04677].

[48] Vaquero, Alejandro and Redondo, Javier and Stadler, Julia, *Early seeds of axion miniclusters*, *JCAP* **04** (2019) 012, [1809.09241].

[49] Kawasaki, Masahiro and Sekiguchi, Toyokazu and Yamaguchi, Masahide and Yokoyama, Jun'ichi, *Long-term dynamics of cosmological axion strings*, *PTEP* **2018** (2018) 091E01, [1806.05566].

[50] Martins, C. J. A. P., *Scaling properties of cosmological axion strings*, *Phys. Lett. B* **788** (2019) 147–151, [1811.12678].

[51] Buschmann, Malte and Foster, Joshua W. and Safdi, Benjamin R., *Early-Universe Simulations of the Cosmological Axion*, *Phys. Rev. Lett.* **124** (2020) 161103, [1906.00967].

[52] Klaer, Vincent B. and Moore, Guy D., *Global cosmic string networks as a function of tension*, *JCAP* **06** (2020) 021, [1912.08058].

[53] Gorghetto, Marco and Hardy, Edward and Villadoro, Giovanni, *More Axions from Strings*, 2007.04990.

[54] Ryden, Barbara S. and Press, William H. and Spergel, David N., *THE EVOLUTION OF NETWORKS OF DOMAIN WALLS AND COSMIC STRINGS*, .

[55] Hiramatsu, Takashi and Kawasaki, Masahiro and Saikawa, Ken'ichi, *Evolution of String-Wall Networks and Axionic Domain Wall Problem*, *JCAP* **08** (2011) 030, [1012.4558].

[56] Hiramatsu, Takashi and Kawasaki, Masahiro and Saikawa, Ken'ichi and Sekiguchi, Toyokazu, *Axion cosmology with long-lived domain walls*, *JCAP* **01** (2013) 001, [1207.3166].

[57] Choi, Kiwoon and Im, Sang Hui, *Realizing the relaxion from multiple axions and its UV completion with high scale supersymmetry*, *JHEP* **01** (2016) 149, [1511.00132].

[58] Kaplan, David E. and Rattazzi, Riccardo, *Large field excursions and approximate discrete symmetries from a clockwork axion*, *Phys. Rev. D* **93** (2016) 085007, [1511.01827].

[59] Balasubramanian, Vijay and Berglund, Per and Conlon, Joseph P. and Quevedo, Fernando, *Systematics of moduli stabilisation in Calabi-Yau flux compactifications*, *JHEP* **03** (2005) 007, [hep-th/0502058].

[60] Conlon, Joseph P. and Quevedo, Fernando and Suruliz, Kerim, *Large-volume flux compactifications: Moduli spectrum and D3/D7 soft supersymmetry breaking*, *JHEP* **08** (2005) 007, [hep-th/0505076].

[61] Hui, Lam, *Recovery of the Shape of the Mass Power Spectrum from the Lyman Alpha Forest*, *Astrophys. J.* **516** (1999) 519–526, [astro-ph/9807068].

[62] Iršič, Vid and Viel, Matteo and Haehnelt, Martin G. and Bolton, James S. and Becker, George D., *First constraints on fuzzy dark matter from Lyman-α forest data and hydrodynamical simulations*, *Phys. Rev. Lett.* **119** (2017) 031302, [1703.04683].

[63] Kamionkowski, Marc and Kosowsky, Arthur and Stebbins, Albert, *A Probe of primordial gravity waves and vorticity*, *Phys. Rev. Lett.* **78** (1997) 2058–2061, [astro-ph/9609132].

[64] Seljak, Uros and Zaldarriaga, Matias, *Signature of gravity waves in polarization of the microwave background*, *Phys. Rev. Lett.* **78** (1997) 2054–2057, [astro-ph/9609169].

[65] Kamionkowski, Marc and Kosowsky, Arthur and Stebbins, Albert, *Statistics of cosmic microwave background polarization*, *Phys. Rev. D* **55** (1997) 7368–7388, [astro-ph/9611125].

[66] Abbott, L. F. and Wise, Mark B., *Constraints on Generalized Inflationary Cosmologies*, *Nucl. Phys. B* **244** (1984) 541–548.

[67] Kamionkowski, Marc, *How to De-Rotate the Cosmic Microwave Background Polarization*, *Phys. Rev. Lett.* **102** (2009) 111302, [0810.1286].

[68] Yadav, Amit P. S. and Biswas, Rahul and Su, Meng and Zaldarriaga, Matias, *Constraining a spatially dependent rotation of the Cosmic Microwave Background Polarization*, *Phys. Rev. D* **79** (2009) 123009, [0902.4466].

[69] Gluscevic, Vera and Kamionkowski, Marc and Cooray, Asantha, *De-Rotation of the Cosmic Microwave Background Polarization: Full-Sky Formalism*, *Phys. Rev. D* **80** (2009) 023510, [0905.1687].

[70] Minami, Yuto and Komatsu, Eiichiro, *New Extraction of the Cosmic Birefringence from the Planck 2018 Polarization Data*, *Phys. Rev. Lett.* **125** (2020) 221301, [2011.11254].

[71] Diego-Palazuelos, P. and others, *Cosmic Birefringence from the Planck Data Release 4*, *Phys. Rev. Lett.* **128** (2022) 091302, [2201.07682].

[72] Eskilt, J. R., *Frequency-Dependent Constraints on Cosmic Birefringence from the LFI and HFI Planck Data Release 4*, *Astron. Astrophys.* **662** (2022) A10, [2201.13347].

[73] Eskilt, Johannes R. and Komatsu, Eiichiro, *Improved Constraints on Cosmic Birefringence from the WMAP and Planck Cosmic Microwave Background Polarization Data*, 2205.13962.

[74] CORE collaboration, Bouchet, F. R. and others, *COrE (Cosmic Origins Explorer) A White Paper*, 1102.2181.

[75] Matsumura, T. and others, *Mission design of LiteBIRD*, *J. Low Temp. Phys.* **176** (2014) 733, [1311.2847].

[76] SIMONS OBSERVATORY collaboration, Ade, Peter and others, *The Simons Observatory: Science goals and forecasts*, *JCAP* **02** (2019) 056, [1808.07445].

[77] CMB-S4 collaboration, Abazajian, Kevork N. and others, *CMB-S4 Science Book, First Edition*, 1610.02743.

[78] NASA PICO collaboration, Hanany, Shaul and others, *PICO: Probe of Inflation and Cosmic Origins*, 1902.10541.

[79] Pogosian, Levon and Shimon, Meir and Mewes, Matthew and Keating, Brian, *Future CMB constraints on cosmic birefringence and implications for fundamental physics*, *Phys. Rev. D* **100** (2019) 023507, [1904.07855].

[80] Bortolami, Marco and Billi, Matteo and Gruppuso, Alessandro and Natoli, Paolo and Pagano, Luca, *Planck constraints on cross-correlations between anisotropic cosmic birefringence and CMB polarization*, 2206.01635.

[81] Gruppuso, Alessandro and Molinari, Diego and Natoli, Paolo and Pagano, Luca, *Planck 2018 constraints on anisotropic birefringence and its cross-correlation with CMB anisotropy*, *JCAP* **11** (2020) 066, [2008.10334].

[82] Contreras, Dagoberto and Boubel, Paula and Scott, Douglas, *Constraints on direction-dependent cosmic birefringence from Planck polarization data*, *JCAP* **12** (2017) 046, [1705.06387].

[83] Namikawa, Toshiya and others, *Atacama Cosmology Telescope: Constraints on cosmic birefringence*, *Phys. Rev. D* **101** (2020) 083527, [2001.10465].

[84] SPT collaboration, Bianchini, F. and others, *Searching for Anisotropic Cosmic Birefringence with Polarization Data from SPTpol*, *Phys. Rev. D* **102** (2020) 083504, [2006.08061].

[85] BICEP2, Keck Arrary collaboration, Ade, P. A. R. and others, *BICEP2 / Keck Array IX: New bounds on anisotropies of CMB polarization rotation and implications for axionlike particles and primordial magnetic fields*, *Phys. Rev. D* **96** (2017) 102003, [1705.02523].

[86] POLARBEAR collaboration, Ade, Peter A. R. and others, *POLARBEAR Constraints on Cosmic Birefringence and Primordial Magnetic Fields*, *Phys. Rev. D* **92** (2015) 123509, [1509.02461].

[87] Salvatier, John and Wiecki, Thomas V and Fonnesbeck, Christopher, *Probabilistic programming in python using pymc3*, *PeerJ Computer Science* **2** (2016) e55.

[88] Gelman, Andrew and Rubin, Donald B, *A single series from the gibbs sampler provides a false sense of security*, *Bayesian statistics* **4** (1992) 625–631.

[89] Sherwin, Blake D. and Namikawa, Toshiya, *Cosmic birefringence tomography and calibration-independence with reionization signals in the CMB*, 2108.09287.

[90] Górski, K. M. and Hivon, E. and Banday, A. J. and Wandelt, B. D. and Hansen, F. K. and Reinecke, M. and Bartelman, M., *HEALPix - A Framework*

for high resolution discretization, and fast analysis of data distributed on the sphere, *Astrophys. J.* **622** (2005) 759–771, [astro-ph/0409513].

[91] Zaldarriaga, Matias and Seljak, Uros, *An all sky analysis of polarization in the microwave background*, *Phys. Rev. D* **55** (1997) 1830–1840, [astro-ph/9609170].

[92] Hu, Wayne and Okamoto, Takemi, *Mass reconstruction with cmb polarization*, *Astrophys. J.* **574** (2002) 566–574, [astro-ph/0111606].

[93] Lewis, Antony and Challinor, Anthony and Lasenby, Anthony, *Efficient computation of CMB anisotropies in closed FRW models*, *Astrophys. J.* **538** (2000) 473–476, [astro-ph/9911177].

[94] Naess, Sigurd and Madhavacheril, Matthew and Hasselfield, Matthew and van Engelen, Alex and Hilton, Matt, "https://github.com/simonsobs/pixell."

[95] Madhavacheril, Matthew and Guan, Yilun, "https://github.com/simonsobs/symlens."

Chapter 3

Measures of non-Gaussianity in axion-string-induced CMB birefringence

In chapter 2 I presented LCM parameter inference results obtained by comparing LCM birefringence power spectra against published birefringence power spectrum measurements. Our analysis placed constraints on the LCM parameters $\mathcal{A}$, ζ_0 and ξ_0 (1). One of the main results of that paper is that we found the measured power spectra constrains the power spectrum amplitude to be $\mathcal{A}^2\xi_0 < 3.7$. The parameter $\mathcal{A}$ is a fundamental physics parameter related to the sum of electromagnetic charges of particles in the theory, unlike ξ_0 which is a phenomenological parameter. Since the birefringence power spectrum is directly proportional to $\mathcal{A}^2\xi_0$, our methodology in chapter 2 does not allow us to directly place constraints on $\mathcal{A}$. In order to make progress in this direction we identified that statistical properties of LCM birefringence other than the power spectrum needed to be studied.

However, random fields like all-sky birefringence maps $\alpha(\boldsymbol{n})$ contain more information than just their power spectrum unless they are Gaussian random fields. The maps produced by the LCM are distinctly non-Gaussian. The work shown in this chapter (chapter 3) characterise the nature of the non-Gaussianity using their kurtosis and bispectrum as measures.

Notes about this project: This chapter is from a paper I wrote with my Andrew J. Long that was published in the Journal of Cosmology and Astroparticle Physics in 2023 (2).

Abstract

The presence of axion strings in the Universe after recombination can leave an imprint on the polarization pattern of the cosmic microwave background radiation through the phenomenon of axion-string-induced birefringence via the hyperlight axion-like particle's coupling to electromagnetism. Across the sky, the polarization rotation angle is expected to display a patchwork of uniform regions with sharp boundaries that arise as the 'shadow' of axion string loops. The statistics of such a birefringence sky map are therefore necessarily non-Gaussian. In this article we quantify the non-Gaussianity in axion-string-induced birefringence using two techniques, kurtosis and bispectrum, which correspond to 4- and 3-point correlation functions. If anisotropic birefringence were detected in the future, a measurement of its non-Gaussian properties would facilitate a discrimination across different new physics sources generally, and in the context of axion strings specifically, it would help to break degeneracies between the axion-photon coupling and properties of the string network.

3.1 Introduction

Observations of the cosmic microwave background (CMB) temperature and polarization anisotropies have informed our understanding of the composition, structure, and evolution of the Universe. These precision measurements have also revealed some surprises, such as the mysterious dark matter and dark energy that permeate the Universe. Ongoing and future observations, with significantly higher precision, may uncover evidence for additional cosmological relics that are currently out of reach (3; 4) such as cosmic axion strings. In this work we seek to quantify the signatures of axion strings through their non-Gaussian cosmic birefringence.

Cosmic strings, one-dimensional topological defects formed from scalar fields (5), are predicted to arise in the early universe during phase transitions associated with as-yet undiscovered new physics. While the new particles and forces may be inaccessible, because they are too heavy to be produced at high-energy colliders or too feebly

coupled to be probed in the laboratory, the network of cosmic strings can leave a detectable imprint on the CMB radiation, which is both exquisitely measured and theoretically well understood. For example, searches for the gravitational influence of cosmic strings on the CMB anisotropies have already yielded an upper limit on the strings' tension (6; 7), which translates into a strong constraint on the scale of new physics. On the other hand, if the string-forming fields couple non-gravitationally to visible matter and radiation, novel channels for testing these theories become available. Cosmic strings formed from hyperlight axion-like particles (ALPs) that couple to electromagnetism provide an especially compelling target, since they are expected to induce a birefringence of CMB polarization (8).

The phenomenon of axion-induced birefringence has been a subject of great interest for many years (9–20). The important aspect of birefringence from axion strings (8) is that the typical axion field excursion is large $\Delta a \approx 2\pi f_a$, thereby evading a suppression factor that appears for other models, such as axion dark matter. Several recent studies (1; 21–27) have explored the signatures of axion-string-(and domain wall)-induced birefringence, calculated the angular power spectrum, and assessed compatibility with the various measurements of CMB birefringence (including a claimed detection of isotropic birefringence (28–33)). To summarize, these studies conclude that the current generation of CMB telescopes (Planck, SPTpol, ACTpol, BICEP2/*Keck Array*, Polarbear) are nearly sensitive enough to probe the most well-motivated parameter space, and next-generation telescopes will put these theories to the test.

Whereas most of the work on axion-string-induced birefringence has focused thus far on two-point statistics such as the angular power spectrum, the higher moments contain a wealth of valuable information that could help to discriminate across different sources of birefrigence (27) if a detection were made with next-generation surveys (34). We illustrate this point in fig. 3.1.1; the left panel shows a simulated map of the birefringence angle across a patch of the sky arising from a network of axion strings,

and the right panel shows a map that was simulated using Gaussian statistics with the same angular power spectrum. These two images can be distinguished easily: the map on the left displays disk-like structures, corresponding to the imprint of axion string loops. Since these birefringence maps have the same two-point correlations, the difference between them arises from higher-order correlations, which cannot reduce to two-point correlations for non-Gaussian statistics.

In this work we seek to quantify these non-Gaussian features in axion-string-induced birefringence using three-point correlations (bispectrum) and four-point correlations (kurtosis), which are familiar tools from studies of CMB non-Gaussianity (35). Similar techniques have been used in the past (36–39) to search for evidence of a cosmic string network's gravitational influence of the CMB anisotropies. Our approach is complementary to the one taken in ref. (27), which contains a related analysis of axion-string-induced birefringence using the scattering transform.

3.2 Kurtosis

We denote the birefringence map by $\hat{\alpha}(\boldsymbol{n})$ where $\boldsymbol{n}$ is a unit vector, indicating a direction on the sky, and $\hat{\alpha}$ is the birefringence angle, corresponding to the rotation of the polarization axis. We use hats to denote random variables and angled brackets to denote ensemble averaging. The birefringence map admits a multipole expansion:

$$\hat{\alpha}(\boldsymbol{n}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \hat{\alpha}_{\ell m} Y_{\ell m}(\boldsymbol{n}) \tag{3.1}$$

where $\hat{\alpha}_{\ell m}$ are called the multipole moment coefficients and $Y_{\ell m}(\boldsymbol{n})$ are the spherical harmonics; we use the standard normalization $\int \mathrm{d}^2\boldsymbol{n}\, |Y_{\ell m}(\boldsymbol{n})|^2 = 1$. Since the birefringence map is real $\hat{\alpha}(\boldsymbol{n})^* = \hat{\alpha}(\boldsymbol{n})$, the complex multipole moment coefficients obey $\hat{\alpha}^*_{\ell m} = (-1)^m\, \hat{\alpha}_{\ell -m}$.

Kurtosis is a convenient measure of non-Gaussianity that is both easy to calculate and intuitive to understand. The kurtosis of the (complex) multipole moment

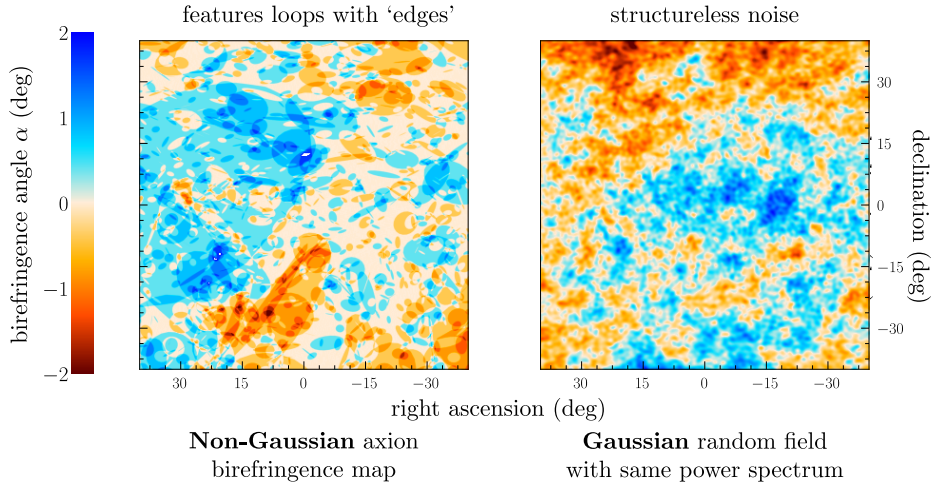


Figure 3.1.1 : *Left:* a simulated map of the birefringence angle $\alpha(\hat{\boldsymbol{n}})$ for an axion string network. *Right:* a simulated map assuming Gaussian statistics with the same power spectrum as the left map. The disk-like features on the left map are a manifestation of the non-Gaussian nature of the stochastic variable. The non-Gaussian map is generated using the loop-crossing model with $\zeta_0 = \xi_0 = \mathcal{A} = 1$; see sec. 3.2 for additional details.

coefficients is given by

$$\kappa_{\ell m} = \frac{\left\langle \left| \hat{\alpha}_{\ell m} - \langle \hat{\alpha}_{\ell m} \rangle \right|^4 \right\rangle}{\left\langle \left| \hat{\alpha}_{\ell m} - \langle \hat{\alpha}_{\ell m} \rangle \right|^2 \right\rangle^2} = \frac{\left\langle \left| \hat{\alpha}_{\ell m} \right|^4 \right\rangle}{\left\langle \left| \hat{\alpha}_{\ell m} \right|^2 \right\rangle^2} \,, \tag{3.2}$$

where the first equality is the general definition, and the second equality holds for axion-string-induced birefringence that has vanishing 1-point functions $\langle \hat{\alpha}_{\ell m} \rangle = 0$. If the real and imaginary parts of the multipole moment coefficients were i.i.d. Gaussian random variables, then Isserlis's theorem (Wick's theorem) would reduce the 4-point functions to products of 2-point functions. For modes with $m = 0$ the reality condition forces $\hat{\alpha}_{\ell 0}$ to be real implying $\kappa_{\ell m} = 3$, whereas for $m \neq 0$ the complex $\hat{\alpha}_{\ell m}$ would

have $\kappa_{\ell m} = 2$ instead.* We define the excess kurtosis

$$\Delta\kappa_{\ell m} = \begin{cases} \kappa_{\ell 0} - 3 & , \quad \text{for } m = 0 \\ \kappa_{\ell m} - 2 & , \quad \text{for } m \neq 0 \end{cases} , \tag{3.3}$$

which vanishes for Gaussian statistics. A positive excess kurtosis $\Delta\kappa_{\ell m} > 0$ corresponds to a distribution with a tighter center and broader tails than a Gaussian having the same mean and variance. In this way, kurtosis provides an intuitive measure of the departure from Gaussianity.

We seek to employ kurtosis as a measure of non-Gaussianity in axion-string-induced birefringence maps. To that end, we simulate birefringence maps using the loop-crossing model (LCM), as described in refs. (1; 23). The LCM is informed by simulations of axion string networks including refs. (40–56). In this model, the string network is approximated as a collection of circular planar loops with a statistically homogeneous distribution through space and a statistically isotropic orientation. On cosmological time scales, the number density of loops decreases and the length of loops grows so as to track the cosmological expansion. Specifically, the number density of loops at time t is $n(t) = \xi_0 H(t)^3/2\pi\zeta_0$ and the radius of loops at time t is $\zeta_0/H(t)$ where $H(t)$ is the Hubble parameter. The dimensionless coefficients, ξ_0 and ζ_0, are two model parameters, and string network simulations motivate values around $\xi_0 = 1\text{-}10$ and $\zeta_0 = 0.1\text{-}1$. As a photon propagates through the string network, from the CMB to a detector on Earth, birefringence accumulates each time the photon passes through the disk bounded by a string loop. The birefringence induced by each loop crossing is $\pm\mathcal{A}\alpha_{\rm em}$ where the dimensionless anomaly coefficient $\mathcal{A} = 0.1\text{-}1$ is another model parameter, $\alpha_{\rm em} \simeq 1/137$ is the electromagnetic fine structure constant, and the two equally-probable signs ± 1 depend on the relative orientation of the loop

*For a single Gaussian random variable $\hat{x}$ with $\langle\hat{x}\rangle = 0$, one finds $\langle\hat{x}^4\rangle = 3\langle\hat{x}^2\rangle^2$ and the kurtosis is $\langle\hat{x}^4\rangle/\langle\hat{x}^2\rangle^2 = 3$. For a complex random variable $\hat{X} = \hat{x} + i\hat{y}$ with statistically independent real and imaginary parts $\langle\hat{x}\hat{y}\rangle = 0$, one finds instead $\langle|\hat{X}|^4\rangle = \langle(\hat{x}^2 + \hat{y}^2)^2\rangle = 3\langle\hat{x}^2\rangle^2 + 2\langle\hat{x}^2\rangle\langle\hat{y}^2\rangle + 3\langle\hat{y}^2\rangle^2$ and $\langle|\hat{X}|^2\rangle^2 = \langle\hat{x}^2\rangle^2 + 2\langle\hat{x}^2\rangle\langle\hat{y}^2\rangle + \langle\hat{y}^2\rangle^2$, and the kurtosis is $\langle|\hat{X}|^4\rangle/\langle|\hat{X}|^2\rangle^2 = 2$ for $\langle\hat{x}^2\rangle = \langle\hat{y}^2\rangle$.

and the photon's propagation direction. The signal of axion-string-induced birefringence also depends upon the axion mass scale m_a, since the string network forms domain walls when the Hubble parameter is comparable to the axion mass scale, possibly suppressing the signal of axion-string-induced birefringence (1). In this work, we assume that the axion mass is smaller than the Hubble scale today $m_a \lesssim 3H_0$ and the string network survives until present times, allowing for an unsuppressed birefringence signal. We implement the loop-crossing model in a Python code that interfaces with HEALPix (57; 58) taking $N_{\rm side} = 128$ or 512 for different studies in this work. With a large number of simulated birefringence maps we calculate sample means to estimate ensemble averages and thereby evaluate the kurtosis of the multipole moment coefficients.

Using the loop-crossing model, we obtain more than 60,000 simulated realizations (up to 150,000) of the axion-string-induced birefringence map. For each map we extract the multipole moment coefficients $\hat{\alpha}_{\ell m}$. To assess the departure from Gaussianity, we show in fig. 3.2.1 the distributions over $\mathrm{Re}\,\hat{\alpha}_{\ell m}$ for the lowest several multipole moments. The distributions over $\mathrm{Im}\,\hat{\alpha}_{\ell m}$ (not shown) are similar. We only show multipole moments with $m > 0$ since the reality condition imposes $\mathrm{Re}\,\hat{\alpha}_{\ell m} = (-1)^m\ \mathrm{Re}\,\hat{\alpha}_{\ell -m}$. We give values of the sample mean $\hat{\mu}_{\ell m}$, sample standard deviation $\hat{\sigma}_{\ell m}$, and sample excess kurtosis $\Delta\hat{\kappa}_{\ell m}$ that were inferred from the suite of simulations. To highlight the departure of these distributions from Gaussianity, we show a normal distribution (dashed line) with the same mean and variance as each histogram. The histograms are approximately symmetric and centered close to zero, since each loop crossing shifts the birefringence by $\pm\mathcal{A}\alpha_{\rm em}$ with equal probability.

Fig. 3.2.1 displays a departure from Gaussianity for multipole moments with small index ℓ. For $\ell = 0$ and 2, the distinction between the histogram and the normal distribution is clearly evident. One can easily see that the histogram is tighter and taller around $\hat{\alpha}_{\ell m} = 0$, and close inspection reveals that it also has wider tails. In general such features correspond to a positive excess kurtosis. For the monopole we

find the excess kurtosis to be $\Delta\hat{\kappa}_{0,0} \approx 1.51$; for the quadrupole it is $\Delta\hat{\kappa}_{2,m} \approx 0.4$; and for $\ell = 8$ is it $\Delta\hat{\kappa}_{8,m} \approx 0.06$. For a given ℓ we find that each m has a similar distribution, which is consistent with the underlying statistical isotropy of the loop-crossing model. These examples illustrate that the excess kurtosis decreases as the multipole index ℓ increases.

We are interested in how the kurtosis varies across angular scales, and specifically how quickly the excess kurtosis decreases for higher multipole moments. Since the loop-crossing model generates a statistically isotropic birefringence map, we expect that $\Delta\hat{\kappa}_{\ell m}$ should only depend on the index ℓ. This observation motivates us to define the 'angle-averaged' excess kurtosis

$$\Delta\hat{\kappa}_\ell \equiv \frac{1}{\ell}\sum_{m=1}^{\ell} \Delta\hat{\kappa}_{\ell m} \ , \tag{3.4}$$

for $\ell > 0$. In fig. 3.2.2 we show the average excess kurtosis across a range of angular scales corresponding to multipole moment indices $\ell = 1$ to 100. We obtain these numerical results from simulated birefringence maps obtained through the loop-crossing model with four values of the dimensionless loop-length parameter: $\zeta_0 = 10$ corresponding to ten-times Hubble scale loops, $\zeta_0 = 1$ corresponding to Hubble scale loops, $\zeta_0 = 10^{-0.5} \approx 0.316$, and $\zeta_0 = 0.1$ corresponding to loops that are a tenth of the Hubble scale. Kurtosis is independent of the parameter $\mathcal{A}$, since $\hat{\alpha}_{\ell m} \propto \mathcal{A}$ and this factor cancels when calculating kurtosis as a ratio of multipole moment coefficients through eq. (3.2).

Fig. 3.2.2 exhibits several notable features. We observe that: (1) the excess kurtosis is positive across this range of multipoles and for this set of model parameters; (2) the excess kurtosis is $\approx$ 0.1-10 at low multipoles, and its value goes inversely with ζ_0; and (3) the excess kurtosis decreases with increasing multipole index ℓ in a (statistically) monotonic way, and it approximately follows a broken power law scaling. These features can be loosely understood as follows. For $\zeta_0 = 1$ the string network contains only one or two loops at the largest angular scales (smallest ℓ), and the statistics of the

birefringence map should also be order one numbers implying $\hat{\kappa}_{\ell m} \approx 1$ and $\Delta\hat{\kappa}_\ell \approx 1$. Increasing ℓ corresponds to decreasing the angular scale of interest, and the network contains more loops at smaller scales on average. As the number of loops increases, their imprint on the birefringence map corresponds to many overlapping disks and ellipses that each contribute $\pm\mathcal{A}\alpha_{\rm em}$. Since each loop's contribution can be either positive or negative (with equal probability, associated with the random orientation of the loop), the net birefringence grows like a random walk with a random number of steps. By the central limit theorem, the statistics of this quantity converge to Gaussian in the limit of many loops. Consequently, one expects an increasingly Gaussian birefringence map on smaller angular scales, corresponding to an inverse relationship between ℓ and $\Delta\hat{\kappa}_\ell$, such as the one seen in fig. 3.2.2. Furthermore, one expects the excess kurtosis to be positive, because the non-Gaussianity is primarily driven by the fact that there are few large loops. These rare outliers boost the tails of the $\hat{\alpha}_{\ell m}$ distribution at values that are relatively large compared to the standard deviation; such features are characteristic of a distribution with positive excess kurtosis.

The preceding loose argument can be formulated more concretely for the monopole $\hat{\alpha}_{00}$, which is proportional to the sky-average birefringence angle. This analysis is presented in app. 3.A. We find that the excess kurtosis in the monopole is inversely proportional to the average number of loops, $\Delta\hat{\kappa}_0 = 1/\overline{N}_{\rm loops}$. Extending this scaling to the higher multipole moments suggests the relationship $\Delta\hat{\kappa}_\ell \sim 1/\overline{N}_\ell$, where $\overline{N}_\ell$ is the average number of loops at a given angular scale $\sim \pi/\ell$. By evaluating the average number of loops as a function of ℓ and the string network model parameters, ζ_0 and ξ_0, we obtain an analytical estimate of the excess kurtosis

$$\Delta\hat{\kappa}_\ell \sim \frac{\zeta_0}{8\xi_0}\left(1 + \frac{\pi}{\lambda\zeta_0\ell}\right)^2 . \tag{3.5}$$

Here $\lambda \approx 0.3$ is a constant numerical factor. See app. 3.A for the derivation of eq. (3.5).

The analytical formula in eq. (3.5) agrees well with the numerical results presented in fig. 3.2.2. Eq. (3.5) implies that $\Delta\hat{\kappa}_\ell$ should scale like ℓ^{-2} for $\ell \ll \pi/(\lambda\,\zeta_0) \approx 10/\zeta_0$

and like ℓ^0 for larger ℓ. Similarly, fig. 3.2.2 shows an ℓ^{-2} scaling for small values of ℓ, and a flattening (in the $\zeta_0 = 10$, 1, and 0.316 curves) for larger values of ℓ approaching $\ell = 100$. Additionally, the angular scale dividing these two regimes is well approximated by $10/\zeta_0$. For the $\zeta_0 = 0.1$ curve, the flattening is not seen, and this is compatible with the analytical model since the transition scale $10/\zeta_0 \approx 100$, and the full plot range from $\ell = 1$ to 100 is in the ℓ^{-2} regime. Eq. (3.5) also predicts a scaling with the model parameters (ζ_0, ξ_0, and $\mathcal{A}$) that agrees well with fig. 3.2.2. For low multipoles, the formula implies $\Delta\hat{\kappa}_\ell \propto 1/\zeta_0$, which is consistent with the numerical results in the figure insofar as lowering ζ_0 increases the excess kurtosis for $\ell \lesssim 30$. For high multipoles, the formula implies $\Delta\hat{\kappa}_\ell \propto \zeta_0$, indicating a reversal of the scaling with ζ_0. The same reversal is seen on the figure, although the linear $\propto \zeta_0$ scaling is not observed. This is possibly because we only show multipoles up to $\ell = 100$, whereas larger values of ℓ are required to exhibit the linear scaling. Additionally, eq. (3.5) implies the relation $\Delta\hat{\kappa}_\ell \propto \xi_0^{-1}$, which we have also verified with numerical simulations taking $\mathcal{A} = \zeta_0 = 1$ and $\xi_0 = 0.1$, 1, and 10 (results not shown here). Eq. (3.5) implies that $\Delta\hat{\kappa}_\ell$ is independent of $\mathcal{A}$, and this is because $\mathcal{A}$ does not impact the average number of loops $\overline{N}_\ell$; more generally, $\mathcal{A}$ cancels from the kurtosis calculation entirely.

To conclude, let us address the issues of observability and cosmic variance. For a single realization of the CMB sky, one can measure the excess kurtosis using an unbiased kurtosis estimator. We consider a simple excess kurtosis estimator defined by

$$\Delta\hat{\kappa}_\ell^{(1)} = \frac{1}{\ell}\sum_{m=1}^{\ell}\frac{|\hat{\alpha}_{\ell m}|^4}{(C_\ell^{\alpha\alpha})^2} - 2 \ , \tag{3.6}$$

which is motivated by the assumption that the birefringence power spectrum is measured well enough that the true power spectrum $C_\ell^{\alpha\alpha}$ is approximately well known. One can apply $\Delta\hat{\kappa}_\ell^{(1)}$ to a measurement of anisotropic CMB birefringence to estimate the excess kurtosis. If the moments $\hat{\alpha}_{\ell m}$ were a Gaussian random field, then the mean

of this estimator would vanish (Gaussian variables have zero kurtosis), and the standard deviation would be $\mathrm{StDev}\,\Delta\hat{\kappa}_\ell^{(1)} = \sqrt{20/\ell}$. This spread in the estimator, even for Gaussian statistics, is a form of cosmic variance. To assess whether the excess kurtosis would be observable for a given model, we can compare the predicted excess kurtosis from eq. (3.5) with the typical variation $\sqrt{20/\ell}$. For the parameters shown in fig. 3.2.2, the predicted excess kurtosis typically falls below the cosmic variance across a wide range of multipoles. On the other hand, in models with small values of ζ_0 and ξ_0, the predicted kurtosis can be larger, especially at low multipoles.

3.3 Bispectrum

A widely-used measure of non-Gaussianity in studies of CMB temperature and polarization anisotropies is the bispectrum, and here we turn our attention to the birefringence bispectrum. We denote the first few moments of the multipole moment coefficients $\hat{\alpha}_{\ell m}$ as

$$\overline{\alpha}_{\ell_1 m_1} = \langle \hat{\alpha}_{\ell_1 m_1} \rangle \tag{3.7a}$$

$$P_{\ell_1 m_1 \ell_2 m_2} = \langle \hat{\alpha}_{\ell_1 m_1} \hat{\alpha}_{\ell_2 m_2} \rangle \tag{3.7b}$$

$$B_{\ell_1 m_1 \ell_2 m_2 \ell_3 m_3} = \langle \hat{\alpha}_{\ell_1 m_1} \hat{\alpha}_{\ell_2 m_2} \hat{\alpha}_{\ell_3 m_3} \rangle \ . \tag{3.7c}$$

For axion-string-induced birefringence, the 1-point functions vanish $\overline{\alpha}_{\ell m} = 0$. If the map is statistically isotropic and parity invariant, the 2-point and 3-point functions can be written in terms of the angular power spectrum C_ℓ and the reduced bispectrum $b_{\ell_1 \ell_2 \ell_3}$ through the relations (59; 60)

$$P_{\ell_1 m_1 \ell_2 m_2} = (-1)^{-m_2} \delta_{\ell_1 \ell_2} \delta_{m_1 - m_2} C_{\ell_1} \tag{3.8a}$$

$$B_{\ell_1 m_1 \ell_2 m_2 \ell_3 m_3} = h_{\ell_1 \ell_2 \ell_3} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} b_{\ell_1 \ell_2 \ell_3} \ , \tag{3.8b}$$

where $h_{\ell_1\ell_2\ell_3}$ is a geometrical factor given by

$$h_{\ell_1\ell_2\ell_3} = \sqrt{\frac{(2\ell_1+1)(2\ell_2+1)(2\ell_3+1)}{4\pi}} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix} , \tag{3.9}$$

and where the second factor is a Wigner 3-j symbol. The 3-j symbols vanish unless the multipole moment indices obey the triangle inequality $|\ell_1 - \ell_2| \leq \ell_3 \leq \ell_1 + \ell_2$ (and similarly for the other two index permutations), implying that one can think of ℓ_1, ℓ_2, and ℓ_3 as the lengths of the legs of a triangle. Additionally parity invariance requires the bispectrum to vanish unless $\ell_1 + \ell_2 + \ell_3$ is an even integer, and this parity condition is enforced by the geometrical factor $h_{\ell_1\ell_2\ell_3}$. It is useful to define the random variables (59):

$$\hat{C}_\ell = (2\ell+1)^{-1} \sum_{m=-\ell}^{\ell} \hat{\alpha}_{\ell m} \hat{\alpha}^*_{\ell m} \tag{3.10a}$$

$$\hat{b}_{\ell_1\ell_2\ell_3} = h^{-1}_{\ell_1\ell_2\ell_3} \sum_{m_1=-\ell_1}^{\ell_1} \sum_{m_2=-\ell_2}^{\ell_2} \sum_{m_3=-\ell_3}^{\ell_3} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \hat{\alpha}_{\ell_1 m_1} \hat{\alpha}_{\ell_2 m_2} \hat{\alpha}_{\ell_3 m_3} , \tag{3.10b}$$

which are unbiased estimators of the angular power spectrum and reduced bispectrum in the sense that $\langle \hat{C}_\ell \rangle = C_\ell$ and $\langle \hat{b}_{\ell_1\ell_2\ell_3} \rangle = b_{\ell_1\ell_2\ell_3}$.

The bispectrum is a measure of the non-Gaussianity in the birefringence map. This can be understood as follows. If the $\hat{\alpha}_{\ell m}$ were independent Gaussian random variables, then higher-point functions could be reduced to 1- and 2-point functions by applying Isserlis's theorem (Wick's theorem). Since the 1-point functions vanish, one would expect the 3-point functions to vanish as well implying $b_{\ell_1\ell_2\ell_3} = 0$ for a Gaussian birefringence map. Conversely, the presence of non-Gaussianity allows the bispectrum to be nonzero, $b_{\ell_1\ell_2\ell_3} \neq 0$. However, this need not be the case, and it is possible for a non-Gaussian birefringence map to have a vanishing bispectrum $b_{\ell_1\ell_2\ell_3} = 0$, and the non-Gaussianity only manifests itself in higher order moments such as the 4-point functions (trispectrum, kurtosis). In particular, although axion-string-induced birefringence is non-Gaussian, we nevertheless expect the bispectrum to vanish. This is because any configuration of loops that would give rise to a nonzero

3-point function has an equiprobable 'opposite' with all loop orientations reversed, which cancels this contribution in the ensemble average. However, it's important to bear in mind that although the bispectrum may vanish as an ensemble average $b_{\ell_1\ell_2\ell_3} = 0$, its estimator must be nonzero for any given realization $\hat{b}_{\ell_1\ell_2\ell_3} \neq 0$. Here we are primarily interested in evaluating the typical size of the bispectrum estimator, quantified through its standard deviation $\mathrm{StDev}[\hat{b}_{\ell_1\ell_2\ell_3}]$.

To assess the typical bispectrum arising from axion-string-induced birefringence, we have used the loop-crossing model to simulate a single realization of the birefingence map and calculate the bispectrum estimator $\hat{b}_{\ell_1\ell_2\ell_3}$. These results are presented in fig. 3.3.1. On the left we show a visualization of $\hat{b}_{\ell_1\ell_2\ell_3}$ where the multipole moment indices (ℓ_1, ℓ_2, ℓ_3) are mapped to points in a three-dimensional volume. Colored dots indicate the value of the bispectrum estimator on a log scale, and smaller values are rendered as semi-transparent to enhance visibility. The tetrahedral shape is a consequence of the triangle inequalities ($|\ell_1 - \ell_2| \leq \ell_3 \leq \ell_1 + \ell_2$ and permutations), since the bispectrum estimator vanishes outside of this region due to geometrical constraints imposed by the 3-j symbols. Additionally the parity condition requires $\ell_1 + \ell_2 + \ell_3$ to be an even integer, which further causes many $\hat{b}_{\ell_1\ell_2\ell_3}$ to vanish. The right panel plots the bispectrum estimator along two rays through the tetrahedron. These rays correspond to (1) the main diagonal of the tetrahedron along which $\ell_1 = \ell_2 = \ell_3$, corresponding to the equilateral triangle form; and (2) the edge of the tetrahedron along which $\ell_1 = \ell_2$ and $\ell_3 = 0$, corresponding to the squeezed triangle form. Due to the symmetry properties of the 3-j symbols, the values of $\hat{b}_{\ell_1\ell_2\ell_3}$ along the three tetrahedral edges are identical.

Several qualitative features of fig. 3.3.1 are easily understood. Since the bispectrum $b_{\ell_1\ell_2\ell_3}$ is expected to vanish for axion-string-induced birefringence, it is not surprising to see that the bispectrum estimator $\hat{b}_{\ell_1\ell_2\ell_3}$ evaluates to a scatter of positive and negative values. For $\ell_1 = \ell_2 = \ell_3 = 0$ the bispectrum estimator is simply the cube of the monopole multipole moment coefficient $\hat{b}_{000} = \sqrt{4\pi}(\hat{\alpha}_{00})^3$, and using

$\hat{\alpha}_{00} \sim 0.5\,\mathrm{deg}$ from fig. 3.2.1 (same simulation parameters) gives $\hat{b}_{000} \sim 0.4\,\mathrm{deg}^3$, which is compatible with the figure. Moving to larger ℓ, the bispectrum estimator tends to decrease in magnitude for higher multipoles, and we quantify and discuss this behavior further below. For this realization the bispectrum estimator is positive along the three tetrahedral edges, corresponding to the squeezed triangle form, but for other realizations they may be negative. The sign of $\hat{b}_{\ell_1\ell_2\ell_3}$ along these rays are correlated with the random sign of the monopole $\hat{\alpha}_{00}$. One can prove this using identities of the Wigner 3-j symbols, but heuristically the relation is $\hat{b}_{\ell\ell 0} \sim \langle |\hat{\alpha}_{\ell m}|^2 \hat{\alpha}_{00} \rangle$.

Repeating these simulations 5,000 times with the same LCM model parameters ($\zeta_0 = \xi_0 = \mathcal{A} = 1$), we evalaute the bispectrum estimator for each realization and present a sample of these results in fig. 3.3.2. We show histograms over the bispectrum estimator for $\ell_1 = \ell_2 = \ell_3 = 0$, 2, and 100, which are normalized so that their integral equals 1. These distributions appear to be centered at $\hat{b}_{\ell_1\ell_2\ell_3} = 0$, and they are approximately symmetric. Moreover, we have verified that the sample mean falls like $1/\sqrt{N_{\mathrm{sims}}}$, as one expects for a random variable with vanishing mean. The distributions in fig. 3.3.2 appear visibly non-Gaussian for $\ell = 0$ and 2, but this is not evidence of non-Gaussianity, since the product $\hat{b}_{\ell_1\ell_2\ell_3} \sim \hat{\alpha}^3_{\ell_1\ell_2\ell_3}$ would be non-Gaussian even if the individual factors $\hat{\alpha}_{\ell_1\ell_2\ell_3}$ were Gaussian. For $\ell = 100$ the distribution appears Gaussian, and this can be understood from the central limit theorem: since the bispectrum estimator is a sum over many terms $\hat{b}_{\ell_1\ell_2\ell_3} \sim \sum \hat{\alpha}^3$, see eq. (3.10), we expect that $\hat{b}_{\ell\ell\ell}$ should be approximately normally distributed at high ℓ since $\hat{b}_{\ell\ell\ell}$ is a linear combination of many i.i.d. random variables. The width of the histogram decreases for increasing multipole moment index ℓ, which is compatible with the trend seen already in fig. 3.3.1.

Although the bispectrum vanishes upon ensemble averaging, it is nonzero for each realization. Such fluctuations could still impact CMB polarization data, where only one realization is available. This observation motivates us to evaluate the standard deviation of the bispectrum estimator $\mathrm{StDev}[\hat{b}_{\ell_1\ell_2\ell_3}] = [\langle \hat{b}^2_{\ell_1\ell_2\ell_3} \rangle - \langle \hat{b}_{\ell_1\ell_2\ell_3} \rangle^2]^{1/2}$. If the

birefringence map were Gaussian, the 6-point function $\langle \hat{b}^2_{\ell_1\ell_2\ell_3} \rangle \sim \langle \hat{\alpha}^6_{\ell m} \rangle$ could be reduced to products of 2-point functions using Isserlis's theorem. By doing so we find

$$\begin{aligned} \mathrm{StDev}\big[\hat{b}_{\ell_1\ell_2\ell_3}\big]_{\text{if } \hat{\alpha}_{\ell m} \text{ are Gaussian}} &= |h_{\ell_1\ell_2\ell_3}|^{-1} \sqrt{C_{\ell_1} C_{\ell_2} C_{\ell_3}} \\ &\times \Big[1 + 2\delta_{\ell_1\ell_2}\delta_{\ell_2\ell_3} + \delta_{\ell_2\ell_3} + \delta_{\ell_1\ell_2} + \delta_{\ell_3\ell_1} + 6\,\delta_{\ell_10}\,\delta_{\ell_20}\,\delta_{\ell_30} \\ &\quad + (2\ell_1+1)\,\delta_{\ell_1\ell_2}\,\delta_{\ell_30} + (2\ell_2+1)\,\delta_{\ell_2\ell_3}\,\delta_{\ell_10} + (2\ell_3+1)\,\delta_{\ell_3\ell_1}\,\delta_{\ell_30} \Big]^{1/2}, \end{aligned} \tag{3.11}$$

where C_ℓ is the angular power spectrum, and we assumed that the multipole indices obey the triangle inequality and parity condition; variations of this formula (bispectrum covariance) appear in refs. (61–64). For a scale-invariant power spectrum $\ell(\ell+1)C_\ell$ is independent of ℓ, and one expects to find $\mathrm{StDev}[\hat{b}_{\ell_1\ell_2\ell_3}] \propto \ell^{-7/2}$ in the equilateral configuration and a larger ℓ^{-2} in the squeezed configuration. We are interested in whether departures from this scaling can arise from the inherent non-Gaussianity of axion-string-induced birefringence.

In fig. 3.3.3 we show the sample standard deviation of the bispectrum estimator $\mathrm{StDev}[\hat{b}_{\ell_1\ell_2\ell_3}]$, calculated using the same loop-crossing model parameters as in the previous figure, $\zeta_0 = \xi_0 = \mathcal{A} = 1$. The purple crosses and boxes correspond to axion-string-induced birefringence, and they were calculated using our simulations; the black dots correspond to Gaussian birefringence, and they were calculated using eq. (3.11). To evaluate C_ℓ in (3.11) we performed 1,000 LCM simulations and averaged over the power spectrum estimator $\hat{C}_\ell$ for each realization, which is approximately scale invariant for $\ell \lesssim 100$. From the figure, it can be seen that the standard deviation of the bispectrum tends to track closely to the expectation for Gaussian birefringence, particularly at higher multipoles with $\ell \gtrsim 5$. This explains why the bispectrum tends to be larger for the squeezed configuration as compared with the equilateral configuration, and why they decrease toward larger ℓ while approximately tracking power laws. For low multipoles $\ell \lesssim 4$ the difference between the bispectrum standard deviation and the Gaussian expectation can be significant, reaching a maximum fractional difference of approximately 80% for $\ell_1 = \ell_2 = 1$ and $\ell_3 = 0$. Since the bispectrum tends to be larger than the Gaussian expectation across a range

of low multipoles, correlated measurements could be used to search for evidence of non-Gaussian axion-string-induced birefringence.

3.4 Conclusion

If a network of axion strings is present in the Universe after recombination, then a coupling of the axion-like particles to electromagnetism will induce anisotropic cosmic birefringence. The birefringence angle will vary across the sky tracing the 'shadow' of the cosmic string network with sharp edges and loop-like features. The statistics of this birefringence map are therefore non-Gaussian, since a Gaussian random field would resemble featureless noise. In this work we have used two familiar measures of non-Gaussianity, kurtosis and bispectrum, to quantify the departure from Gaussian statistics.

Kurtosis is calculated from the fourth moment of the birefringence rotation angle; roughly $\kappa \sim \langle\hat{\alpha}^4\rangle/\langle\hat{\alpha}^2\rangle^2$. For Gaussian statistics, four-point functions are equal to products of two-point functions, and we define the excess kurtosis $\Delta\hat{\kappa}_\ell$ to measure the deviation from Gaussian statistics. We use a combination of numerical simulation, in a phenomenological framework called the loop-crossing model, and analytical approximation to evaluate the excess kurtosis across a range of angular scales (with multiple index $\ell = 0$ to 100) and for a range of string network model parameters (ζ_0, ξ_0, and $\mathcal{A}$). We find that excess kurtosis tends to be positive, order 0.1 to 10 at the largest angular scales (depending on model parameters), and decreasing toward smaller angular scales. To understand how the excess kurtosis varies with ℓ and depends on the model parameters, we have developed a simplified analytical model that leads to the approximation for $\Delta\hat{\kappa}_\ell$ provided in eq. (3.5). This formula agrees remarkably well with the scaling relations inferred from simulations. To assess observability, we have calculated the cosmic variance of an excess kurtosis estimator assuming $\hat{\alpha}$ to be a Gaussian random field and perfect knowledge of the power spectrum $C_\ell^{\alpha\alpha}$. For small values of ζ_0 and ξ_0, the excess kurtosis arising from axion-string induced birefringence

can be larger, on average, than the uncertainty from cosmic variance. These estimates indicate that the signal is detectable in principle, but likely challenging in practice.

The bispectrum is defined as the third moment of the birefringence rotation angle at different angular scales; roughly $b \sim \langle \hat{\alpha}_1 \hat{\alpha}_2 \hat{\alpha}_3 \rangle$. For axion-string-induced birefringence we expect the bispectrum to vanish as an ensemble average, but it must be nonzero in any given realization, and our analysis focuses on calculating its standard deviation. Using numerical simulations of the loop-crossing model, we evaluate the reduced bispectrum $\hat{b}_{\ell_1 \ell_2 \ell_3}$ for a range of angular scales from $\ell_i = 0$ to 100. We find that the bispectrum tends to be largest for the 'squeezed' triangle form ($\ell_1 = 0$, $\ell_2 = \ell_3$ and permutations) and relatively smaller in the 'equilateral' triangle form ($\ell_1 = \ell_2 = \ell_3$). For both cases the typical bispectrum decreases toward larger ℓ_i, approximately tracking a power law. We discuss how these general trends would arise even if the birefringence rotation angle followed Gaussian statistics. For the model parameters that we explored numerically here ($\zeta_0 = \xi_0 = \mathcal{A} = 1$), the typical bispectrum tracks the Gaussian expectation, and the largest difference occurs for $\ell_1 = \ell_2 = 1$ and $\ell_3 = 0$ (and permutations) where the fractional difference is approximately 80%. This deviation suggests that an anomalously large bispectrum would be consistent with axion-strings, although additional information such as a measurement of the power spectrum would be needed to claim evidence of axion strings from CMB polarization.

The work presented here serves to better characterize the cosmological signatures of an axion string network present in the Universe after recombination. If evidence for anisotropic birefringence is detected in CMB polarization measurements using two-point statistics, such as EB cross-correlation, the higher-moment statistics studied here will prove valuable to discriminate across different possible new physics sources of birefringence. For instance, at the level of the power spectrum the parameters of axion-string-induced birefringence exhibit a degeneracy; the signal is proportional to $\mathcal{A}^2 \xi_0$ where the anomaly coefficient $\mathcal{A}$ quantifies the strength of the axion-photon coupling, and the loop density parameter ξ_0 controls the number of axion string loops

per Hubble volume. A detection of anisotropic birefringence and a measurement of its power spectrum would not provide sufficient information to discriminate between $\mathcal{A}$ and ξ_0. However, in general this degeneracy can be broken by higher-point statistics (27). For example, eq. (3.5) reveals that the excess kurtosis $\Delta\hat{\kappa}_\ell$ is insensitive to $\mathcal{A}$ and goes inversely with ξ_0. Consequently, with sufficient information it becomes possible to independently determine the properties of the axion string network, parametrized here by ζ_0 and ξ_0, and the fundamental parameters of the new physics, parametrized by the anomaly coefficient $\mathcal{A}$ as well as the axion mass m_a.

Acknowledgments

We are grateful to Mustafa A. Amin, Mudit Jain, and Joel Meyers for discussions and comments on the draft. We thank Winston Yin for pointing out an incorrect reference in the first version of this paper. R.H. is grateful to Siyang Ling for many insightful conversations. R.H. and A.J.L. are supported in part by the National Science Foundation under Award No. PHY-2114024. Some of the results in this paper have been derived using the healpy and HEALPix package.

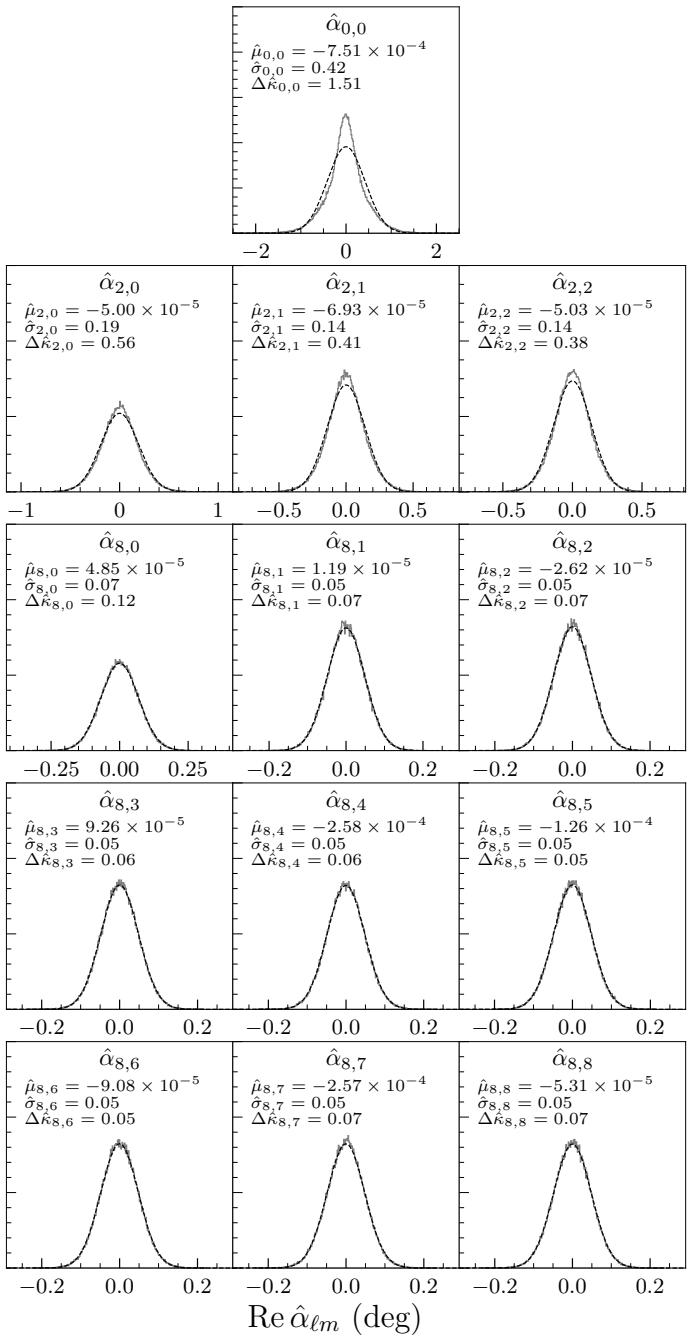


Figure 3.2.1 : Histogram showing distributions over the real part of the multipole moment coefficients $\hat{\alpha}_{\ell m}$ for axion-string-induced birefringence. These distributions were generated using 150,000 realizations of birefringence maps simulated in the loop-crossing model with $\zeta_0 = 1$, $\xi_0 = 1$, and $\mathcal{A} = 1$. In each panel the x-axis is the value of $\mathrm{Re}\,\hat{\alpha}_{lm}$ in degrees, $\hat{\mu}_{\ell m}$ and $\hat{\sigma}_{\ell m}$ are the sample mean and standard deviation of $\mathrm{Re}\,\hat{\alpha}_{\ell m}$, and $\Delta\hat{\kappa}_{\ell m}$ is the excess kurtosis of $\hat{\alpha}_{\ell m}$ through eqs. (3.2) and (3.3). Black dashed curves show a Gaussian distribution with the same mean and variance as the histogram.

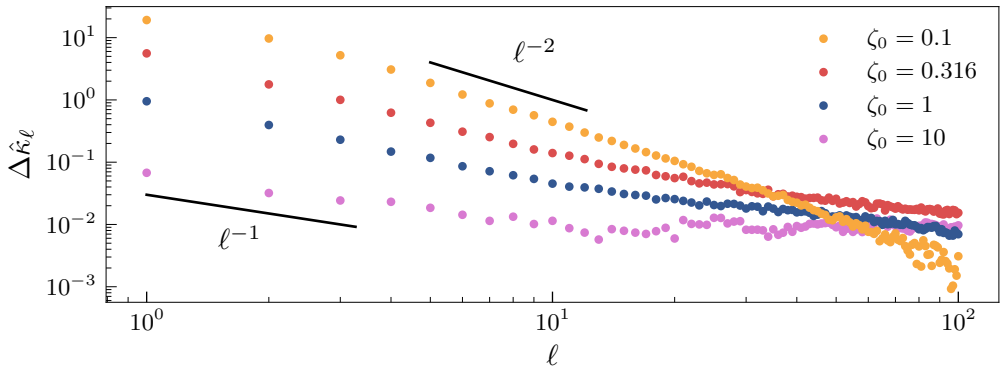


Figure 3.2.2 : Excess kurtosis of axion-string-induced birefringence for a range of multipole moments. We show the average excess kurtosis $\Delta\hat{\kappa}_\ell$ for multipole moments with index ℓ ranging from 1 to 100. The excess kurtosis is calculated from simulated birefringence maps that were created using the loop-crossing model with $\xi_0 = 1$, $\mathcal{A} = 1$, and three values of the loop-length parameter ζ_0. The number of realizations is 150,000 for $\zeta_0 = 1$, 40,000 for $\zeta_0 = 0.316$, 62,000 for $\zeta_0 = 0.1$, and 150,000 for $\zeta_0 = 10$. The curves approximately follow broken power law scalings for small and large multipoles.

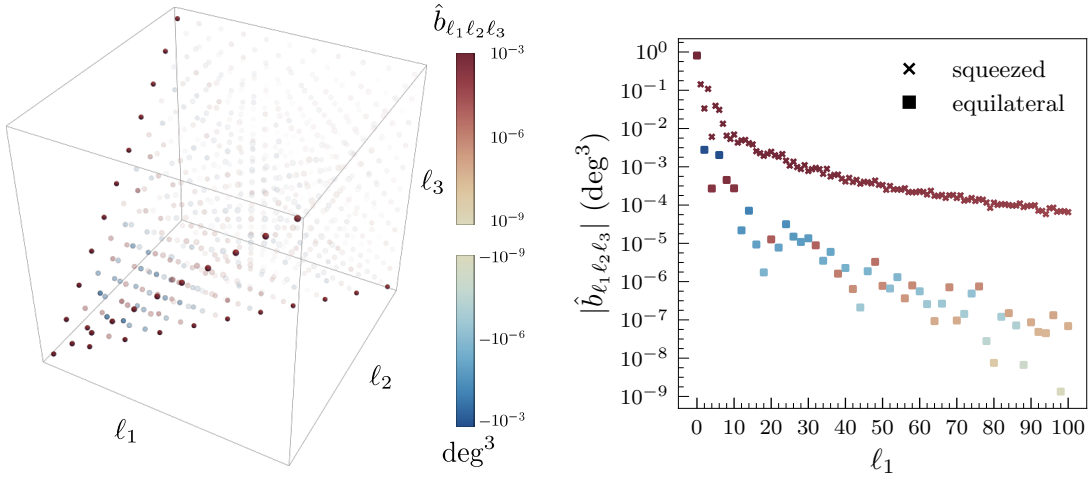


Figure 3.3.1 : A single realization of the bispectrum estimator $\hat{b}_{\ell_1\ell_2\ell_3}$ calculated from a simulated birefringence map using the loop-crossing model with parameters $\zeta_0 = \xi_0 = \mathcal{A} = 1$. For other values of $\mathcal{A}$ the bispectrum estimator would scale as $\propto \mathcal{A}^3$. *Left:* Colored dots indicate values of the bispectrum estimator for multipole moment indices ℓ_1, ℓ_2, ℓ_3 ranging from 0 to 100 in steps of 10. *Right:* Values of the bispectrum estimator along the edge of the tetrahedron where $\ell_1 = \ell_2$ and $\ell_3 = 0$ corresponding to a 'squeezed' triangle (cross markers) and along the main diagonal where $\ell_1 = \ell_2 = \ell_3$ corresponding to an 'equilateral' triangle (square markers).

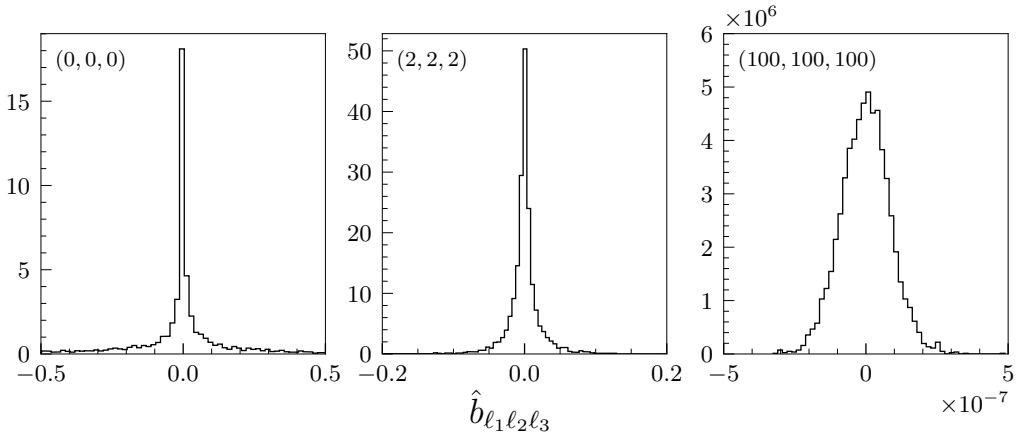


Figure 3.3.2 : Distributions of bispectrum estimators for $\ell_1 = \ell_2 = \ell_3 = 0$ (left), 2 (middle), and 100 (right). We have used 5,000 simulations of the loop-crossing model with parameters $\zeta_0 = \xi_0 = \mathcal{A} = 1$. For other values of $\mathcal{A}$ the bispectrum estimator would scale as $\propto \mathcal{A}^3$.

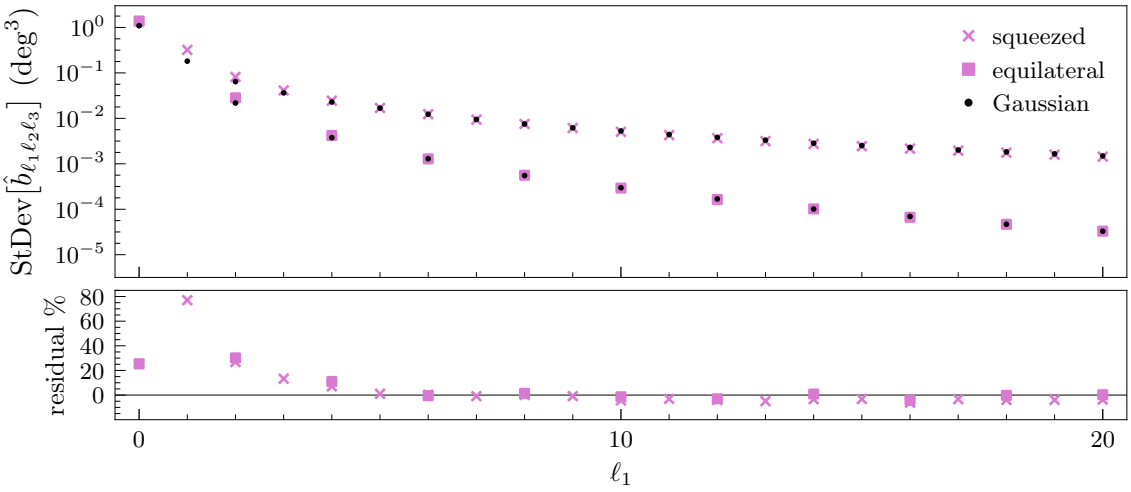


Figure 3.3.3 : *Top:* Standard deviation of the birefringence bispectrum estimator StDev$[\hat{b}_{\ell_1\ell_2\ell_3}]$. We show all values that are nonzero for $\ell \leq 20$; no binning was performed. Purple markers denote results from axion-string-induced birefringence, calculated as the sample average of 5,000 LCM simulations for the model with parameters $\zeta_0 = \xi_0 = \mathcal{A} = 1$. Crossed markers correspond to the squeezed triangle form with $\ell_1 = \ell_2,\ \ell_3 = 0$, and square markers correspond to the equilateral triangle with $\ell_1 = \ell_2 = \ell_3$. Black dots indicate the expected bispectrum standard deviation for Gaussian birefringence, given by eq. (3.11). *Bottom:* The fractional difference between the bispectrum standard deviation and the expectation for Gaussian statistics.

Appendix

3.A Analytical analysis for kurtosis

To develop an analytical understanding of the kurtosis arising in axion-string-induced birefringence, we provide here a simplified description that is analytically tractable. We first consider the monopole $\hat{\alpha}_{00}$ and then extend this analysis to higher multipoles with $\ell > 0$.

3.A.1 Monopole

Consider the monopole of the birefringence map

$$\hat{\alpha}_{00} = \int \mathrm{d}^2 n \, Y_{00}^*(\boldsymbol{n}) \, \hat{\alpha}(\boldsymbol{n}) = \frac{1}{\sqrt{4\pi}} \int \mathrm{d}^2 n \, \hat{\alpha}(\boldsymbol{n}) \, . \tag{3.12}$$

In the loop-crossing model, the birefringence map $\hat{\alpha}(\boldsymbol{n})$ is built up from random overlapping string loops of different sizes and orientations, distributed isotropically across the sky. Photons propagating through the disk encircled by a loop experience a random birefringence, which accumulates with multiple loop crossings. For simplicity we suppose here that every loop crossing leads to a statistically equivalent shift in the monopole, $\Delta\hat{\alpha}_{00} = +C$ or $-C$ with equal probability. More realistically in the loop-crossing model, larger loops contribute more and smaller loops less, and the loop's orientation affects the solid angle it spans on the sky, but these effects are ignored for this simplified analysis. Note that the location of the loops on the sky is irrelevant for the monopole. We also suppose that the number of loops giving this contribution, denoted as $\hat{N}_{\mathrm{loops}}$ is random and Poisson distributed with intensity parameter $\overline{N}_{\mathrm{loops}}$. The quantity $\overline{N}_{\mathrm{loops}}$ is calculable within the loop-crossing model in terms of the properties of the string network. These simplifications allow the

monopole to be written as

$$\begin{aligned} \hat{\alpha}_{00} &= C \sum_{i=1}^{\hat{N}_{\text{loops}}} \hat{W}_i \\ \hat{W}_i &\sim -1 \text{ or } 1 \text{ with equal probability} \\ \hat{N}_{\text{loops}} &\sim \text{Poisson}(\overline{N}_{\text{loops}}) \ . \end{aligned} \tag{3.13}$$

This is an example of a hierarchical random model, where the number of random variables (loop crossings) is itself a random variable.

We are interested in the moments of $\hat{\alpha}_{00}$, which give the kurtosis. It is useful to recognize that $(\hat{W}_i + 1)/2$ is a Bernoulli(1/2) random variable, taking values 0 and 1 with equal probability. The sum over a sequence of n i.i.d. Bernoulli(1/2) random variables is a binomial(n, 1/2) random variable. This motivates us to define $\hat{Y}_n = \sum_{i=1}^{n} (\hat{W}_i + 1)/2$ and write the monopole as

$$\hat{\alpha}_{00} = C\Big(2\hat{Y}_{\hat{N}_{\text{loops}}} - \hat{N}_{\text{loops}}\Big) \ . \tag{3.14}$$

This expression can be used to calculate the kurtosis of $\hat{\alpha}_{00}$ analytically using the fact that for any two random variables $\hat{x}$ and $\hat{y}$, expectation values can be calculated as (65)

$$\mathrm{E}(\hat{x}) = \mathrm{E}\big[\mathrm{E}(\hat{x}|\hat{y})\big] \tag{3.15}$$

so long as the expectation values exist. For example, the first moment is calculated as follows:

$$\begin{aligned} \frac{1}{C}\,\mathrm{E}(\hat{\alpha}_{00}) &= 2\,\mathrm{E}\big[\hat{Y}_{\hat{N}_{\text{loops}}}\big] - \mathrm{E}(\hat{N}_{\text{loops}}) \\ &= 2\,\mathrm{E}\big[\mathrm{E}(\hat{Y}_{\hat{N}_{\text{loops}}}|\hat{N}_{\text{loops}})\big] - \overline{N}_{\text{loops}} \\ &= 2\,\mathrm{E}\big[\hat{N}_{\text{loops}}/2\big] - \overline{N}_{\text{loops}} \\ &= 0 \ . \end{aligned} \tag{3.16}$$

Repeating this procedure for $\mathrm{E}(\hat{\alpha}_{00}^2)$, and $E(\hat{\alpha}_{00}^4)$ we find

$$\mathrm{E}(\hat{\alpha}_{00}^2) = C^2\,\overline{N}_{\text{loops}} \tag{3.17}$$

$$\mathrm{E}(\hat{\alpha}_{00}^4) = C^4\left[3\overline{N}_{\text{loops}}^2 + \overline{N}_{\text{loops}}\right] \ . \tag{3.18}$$

The corresponding excess kurtosis is

$$\Delta\kappa_0 = \kappa_{00} - 3 = \frac{\mathrm{E}(\hat{a}_{00}^4)}{\mathrm{E}(\hat{a}_{00}^2)^2} - 3 = 1/\overline{N}_{\mathrm{loops}} \,, \tag{3.19}$$

which is the result quoted in the main text.

3.A.2 Higher multipoles

We suppose that the monopole relation in eq. (3.19) extends to higher multipoles as

$$\Delta\kappa_\ell \sim 1/\overline{N}_\ell \,, \tag{3.20}$$

where $\overline{N}_\ell$ denotes the average number of loops on an angular scale $\sim \pi/\ell$. By calculating $\overline{N}_\ell$ in the loop-crossing model, we obtain an expression for the angle-averaged excess kurtosis $\Delta\kappa_\ell$ in terms of the multipole index ℓ and the string network model parameters.

First, in the loop-crossing model, the typical length of loops in the network grows with time to track the growing Hubble scale. At redshift z, the typical angular scale of the loops is (23)

$$\delta\theta \sim \pi/\ell \sim \frac{2\lambda\zeta_0}{a(z)H(z)s(z)} \,, \tag{3.21}$$

where $\lambda = 0.3$ accounts for the random orientation of the loops, and in a matter-dominated cosmology: $a(z) \propto (1+z)^{-1}$ is the scale factor, $H(z) \propto (1+z)^{-3/2}$ is the Hubble parameter, and $s(z) = \int_0^z \mathrm{d}z'/a_0 H(z')$ is the comoving distance to redshift z. Solving this relation for z gives

$$z_\ell \sim \frac{\lambda\zeta_0\ell(2\pi + \lambda\zeta_0\ell)}{\pi^2} \,, \tag{3.22}$$

which represents the redshift at which loops with angular scale π/ℓ were present in the network. In the loop-crossing model, the average comoving number density of loops in the network at redshift z is taken to be

$$\overline{n}(z) = \frac{\xi_0 a(z)^3 H(z)^3}{2\pi\zeta_0} \,. \tag{3.23}$$

Integrating over a spherical shell of redshifts $z_\ell < z < z_\ell + \Delta z$ gives

$$\begin{aligned}\overline{N}_\ell &= \int_{z_\ell}^{z_\ell+\Delta z} \mathrm{d}z\, 4\pi s^2(z) \frac{\mathrm{d}s}{\mathrm{d}z}\, \overline{n}(z) \\ &\approx \Delta z\, 4\pi s^2(z_\ell) \frac{1}{a_0 H(z_\ell)} \frac{\xi_0 a(z_\ell)^3 H(z_\ell)^3}{2\pi\zeta_0} \\ &\sim \frac{8\lambda^2 \zeta_0 \xi_0 \ell^2 \Delta z}{(\pi + \lambda\zeta_0\ell)^2}\,,\end{aligned} \tag{3.24}$$

which represents the average number of loops with angular extent π/ℓ. If the excess angle-averaged kurtosis can be estimated as $\Delta\kappa_\ell \sim 1/\overline{N}_\ell$, then we have

$$\Delta\kappa_\ell \sim \frac{\zeta_0}{8\,\xi_0} \left(1 + \frac{\pi}{\lambda\zeta_0\ell}\right)^2 \tag{3.25}$$

where we have taken $\Delta z = 1$. This expression matches the result quoted in the main text.

Bibliography

[1] Jain, Mudit and Hagimoto, Ray and Long, Andrew J. and Amin, Mustafa A., *Searching for axion-like particles through CMB birefringence from string-wall networks*, *JCAP* **10** (2022) 090, [2208.08391].

[2] Hagimoto, Ray and Long, Andrew J., *Measures of non-Gaussianity in axion-string-induced CMB birefringence*, *JCAP* **09** (2023) 024, [2306.07351].

[3] CMB-HD collaboration, Aiola, Simone and others, *Snowmass2021 CMB-HD White Paper*, 2203.05728.

[4] CMB-S4 collaboration, Abazajian, Kevork and others, *Snowmass 2021 CMB-S4 White Paper*, 2203.08024.

[5] Vilenkin, A. and Shellard, E. P. S., *Cosmic Strings and Other Topological Defects*. Cambridge University Press, 2000.

[6] Charnock, Tom and Avgoustidis, Anastasios and Copeland, Edmund J. and Moss, Adam, *CMB constraints on cosmic strings and superstrings*, *Phys. Rev. D* **93** (2016) 123503, [1603.01275].

[7] Lizarraga, Joanes and Urrestilla, Jon and Daverio, David and Hindmarsh, Mark and Kunz, Martin, *New CMB constraints for Abelian Higgs cosmic strings*, *JCAP* **10** (2016) 042, [1609.03386].

[8] Agrawal, Prateek and Hook, Anson and Huang, Junwu, *A CMB Millikan experiment with cosmic axiverse strings*, *JHEP* **07** (2020) 138, [1912.02823].

[9] Harvey, Jeffrey A. and Naculich, Stephen G., *Cosmic Strings From Pseudoanomalous U(1)s*, *Phys. Lett. B* **217** (1989) 231–237.

[10] Carroll, Sean M. and Field, George B. and Jackiw, Roman, *Limits on a Lorentz and Parity Violating Modification of Electrodynamics*, *Phys. Rev. D* **41** (1990) 1231.

[11] Carroll, Sean M. and Field, George B., *The Einstein equivalence principle and the polarization of radio galaxies*, *Phys. Rev. D* **43** (1991) 3789.

[12] Harari, Diego and Sikivie, Pierre, *Effects of a Nambu-Goldstone boson on the polarization of radio galaxies and the cosmic microwave background*, *Phys. Lett. B* **289** (1992) 67–72.

[13] Carroll, Sean M., *Quintessence and the rest of the world*, *AIP Conf. Proc.* **478** (1999) 291–294.

[14] Fedderke, Michael A. and Graham, Peter W. and Rajendran, Surjeet, *Axion Dark Matter Detection with CMB Polarization*, *Phys. Rev. D* **100** (2019) 015040, [1903.02666].

[15] Greco, Alessandro and Bartolo, Nicola and Gruppuso, Alessandro, *Cosmic birefrigence: cross-spectra and cross-bispectra with CMB anisotropies*, *JCAP* **03** (2022) 050, [2202.04584].

[16] Murai, Kai and Naokawa, Fumihiro and Namikawa, Toshiya and Komatsu, Eiichiro, *Isotropic cosmic birefringence from early dark energy*, *Phys. Rev. D* **107** (2023) L041302, [2209.07804].

[17] Cai, Hongbo and Guan, Yilun and Namikawa, Toshiya and Kosowsky, Arthur, *Impact of anisotropic birefringence on measuring cosmic microwave background lensing*, *Phys. Rev. D* **107** (2023) 043513, [2209.08749].

[18] Greco, Alessandro and Bartolo, Nicola and Gruppuso, Alessandro, *Probing Axions through Tomography of Anisotropic Cosmic Birefringence*, *JCAP* **05** (2023) 026, [2211.06380].

[19] Yin, Lu and Kochappan, Joby and Ghosh, Tuhin and Lee, Bum-Hoon, *Is Cosmic Birefringence model-dependent?*, 2305.07937.

[20] Naokawa, Fumihiro and Namikawa, Toshiya, *Gravitational lensing effect on cosmic birefringence*, 2305.13976.

[21] Agrawal, Prateek and Hook, Anson and Huang, Junwu and Marques-Tavares, Gustavo, *Axion string signatures: a cosmological plasma collider*, *JHEP* **01** (2022) 103, [2010.15848].

[22] Takahashi, Fuminobu and Yin, Wen, *Kilobyte Cosmic Birefringence from ALP Domain Walls*, *JCAP* **04** (2021) 007, [2012.11576].

[23] Jain, Mudit and Long, Andrew J. and Amin, Mustafa A., *CMB birefringence from ultralight-axion string networks*, *JCAP* **05** (2021) 055, [2103.10962].

[24] Yin, Weichen Winston and Dai, Liang and Ferraro, Simone, *Probing cosmic strings by reconstructing polarization rotation of the cosmic microwave background*, *JCAP* **06** (2022) 033, [2111.12741].

[25] Kitajima, Naoya and Kozai, Fumiaki and Takahashi, Fuminobu and Yin, Wen, *Power spectrum of domain-wall network, and its implications for isotropic and anisotropic cosmic birefringence*, *JCAP* **10** (2022) 043, [2205.05083].

[26] Gonzalez, Diego and Kitajima, Naoya and Takahashi, Fuminobu and Yin, Wen, *Stability of domain wall network with initial inflationary fluctuations and its implications for cosmic birefringence*, *Phys. Lett. B* **843** (2023) 137990, [2211.06849].

[27] Yin, Weichen Winston and Dai, Liang and Ferraro, Simone, *Testing charge quantization with axion string-induced cosmic birefringence*, 2305.02318.

[28] Minami, Yuto and Komatsu, Eiichiro, *New Extraction of the Cosmic Birefringence from the Planck 2018 Polarization Data*, *Phys. Rev. Lett.* **125** (2020) 221301, [2011.11254].

[29] Diego-Palazuelos, P. and others, *Cosmic Birefringence from the Planck Data Release 4*, *Phys. Rev. Lett.* **128** (2022) 091302, [2201.07682].

[30] Eskilt, J. R., *Frequency-dependent constraints on cosmic birefringence from the LFI and HFI Planck Data Release 4*, *Astron. Astrophys.* **662** (2022) A10, [2201.13347].

[31] Eskilt, Johannes R. and Komatsu, Eiichiro, *Improved constraints on cosmic birefringence from the WMAP and Planck cosmic microwave background polarization data*, *Phys. Rev. D* **106** (2022) 063503, [2205.13962].

[32] Diego-Palazuelos, P. and others, *Robustness of cosmic birefringence*

measurement against Galactic foreground emission and instrumental systematics, *JCAP* **01** (2023) 044, [2210.07655].

[33] Eskilt, J. R. and others, *Cosmoglobe DR1 results. II. Constraints on isotropic cosmic birefringence from reprocessed WMAP and Planck LFI data*, 2305.02268.

[34] Guzman, Eric and Meyers, Joel, *Reconstructing cosmic polarization rotation with ResUNet-CMB*, *JCAP* **01** (2022) 030, [2109.09715].

[35] Meerburg, P. Daniel and others, *Primordial Non-Gaussianity*, 1903.04409.

[36] Gangui, A., *On the CMB kurtosis from cosmic strings.*, *Astrophysical Letters and Communications* **33** (Feb., 1996) 25–31.

[37] Landriau, M. and Shellard, E. P. S., *Cosmic String Induced CMB Maps*, *Phys. Rev. D* **83** (2011) 043516, [1004.2885].

[38] Planck collaboration, Ade, P. A. R. and others, *Planck 2013 results. XXV. Searches for cosmic strings and other topological defects*, *Astron. Astrophys.* **571** (2014) A25, [1303.5085].

[39] Hergt, Lukas and Amara, Adam and Brandenberger, Robert and Kacprzak, Tomasz and Refregier, Alexandre, *Searching for Cosmic Strings in CMB Anisotropy Maps using Wavelets and Curvelets*, *JCAP* **06** (2017) 004, [1608.00004].

[40] Yamaguchi, Masahide and Kawasaki, M. and Yokoyama, Jun'ichi, *Evolution of axionic strings and spectrum of axions radiated from them*, *Phys. Rev. Lett.* **82** (1999) 4578–4581, [hep-ph/9811311].

[41] Yamaguchi, Masahide and Yokoyama, Jun'ichi, *Quantitative evolution of global strings from the Lagrangian view point*, *Phys. Rev. D* **67** (2003) 103514, [hep-ph/0210343].

[42] Hiramatsu, Takashi and Kawasaki, Masahiro and Sekiguchi, Toyokazu and Yamaguchi, Masahide and Yokoyama, Jun'ichi, *Improved estimation of radiated axions from cosmological axionic strings*, *Phys. Rev. D* **83** (2011) 123531, [1012.5502].

[43] Hiramatsu, Takashi and Kawasaki, Masahiro and Saikawa, Ken'ichi and Sekiguchi, Toyokazu, *Production of dark matter axions from collapse of string-wall systems*, *Phys. Rev. D* **85** (2012) 105020, [1202.5851].

[44] Kawasaki, Masahiro and Saikawa, Ken'ichi and Sekiguchi, Toyokazu, *Axion dark matter from topological defects*, *Phys. Rev. D* **91** (2015) 065014, [1412.0789].

[45] Lopez-Eiguren, Asier and Lizarraga, Joanes and Hindmarsh, Mark and Urrestilla, Jon, *Cosmic Microwave Background constraints for global strings and global monopoles*, *JCAP* **07** (2017) 026, [1705.04154].

[46] Gorghetto, Marco and Hardy, Edward and Villadoro, Giovanni, *Axions from Strings: the Attractive Solution*, *JHEP* **07** (2018) 151, [1806.04677].

[47] Kawasaki, Masahiro and Sekiguchi, Toyokazu and Yamaguchi, Masahide and Yokoyama, Jun'ichi, *Long-term dynamics of cosmological axion strings*, *PTEP* **2018** (2018) 091E01, [1806.05566].

[48] Vaquero, Alejandro and Redondo, Javier and Stadler, Julia, *Early seeds of axion miniclusters*, *JCAP* **04** (2019) 012, [1809.09241].

[49] Martins, C. J. A. P., *Scaling properties of cosmological axion strings*, *Phys. Lett. B* **788** (2019) 147–151, [1811.12678].

[50] Buschmann, Malte and Foster, Joshua W. and Safdi, Benjamin R., *Early-Universe Simulations of the Cosmological Axion*, *Phys. Rev. Lett.* **124** (2020) 161103, [1906.00967].

[51] Hindmarsh, Mark and Lizarraga, Joanes and Lopez-Eiguren, Asier and Urrestilla, Jon, *Scaling Density of Axion Strings*, *Phys. Rev. Lett.* **124** (2020) 021301, [1908.03522].

[52] Klaer, Vincent B. and Moore, Guy D., *Global cosmic string networks as a function of tension*, *JCAP* **06** (2020) 021, [1912.08058].

[53] Gorghetto, Marco and Hardy, Edward and Villadoro, Giovanni, *More axions from strings*, *SciPost Phys.* **10** (2021) 050, [2007.04990].

[54] Gorghetto, Marco and Hardy, Edward and Nicolaescu, Horia, *Observing invisible axions with gravitational waves*, *JCAP* **06** (2021) 034, [2101.11007].

[55] Hindmarsh, Mark and Lizarraga, Joanes and Lopez-Eiguren, Asier and Urrestilla, Jon, *Approach to scaling in axion string networks*, *Phys. Rev. D* **103** (2021) 103534, [2102.07723].

[56] Buschmann, Malte and Foster, Joshua W. and Hook, Anson and Peterson, Adam and Willcox, Don E. and Zhang, Weiqun and Safdi, Benjamin R., *Dark matter from axion strings with adaptive mesh refinement*, *Nature Commun.* **13** (2022) 1049, [2108.05368].

[57] Górski, K. M. and Hivon, E. and Banday, A. J. and Wandelt, B. D. and Hansen, F. K. and Reinecke, M. and Bartelmann, M., *HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere*, *The Astrophysical Journal* **622** (Apr., 2005) 759–771, [arXiv:astro-ph/0409513].

[58] Andrea Zonca and Leo Singer and Daniel Lenz and Martin Reinecke and Cyrille Rosset and Eric Hivon and Krzysztof Gorski, *healpy: equal area pixelization and spherical harmonics transforms for data on the sphere in python*, *Journal of Open Source Software* **4** (Mar., 2019) 1298.

[59] Komatsu, Eiichiro, *The pursuit of non-gaussian fluctuations in the cosmic microwave background*. PhD thesis, Tohoku U., 2001. astro-ph/0206039.

[60] Planck collaboration, Ade, P. A. R. and others, *Planck 2013 Results. XXIV. Constraints on primordial non-Gaussianity*, *Astron. Astrophys.* **571** (2014) A24, [1303.5084].

[61] Luo, Xiao-chun, *The Angular bispectrum of the cosmic microwave background*, *Astrophys. J. Lett.* **427** (1994) L71, [astro-ph/9312004].

[62] Spergel, David N. and Goldberg, David M., *Microwave background bispectrum. 1. Basic formalism*, *Phys. Rev. D* **59** (1999) 103001, [astro-ph/9811252].

[63] Gangui, Alejandro and Martin, Jerome, *Cosmic microwave background bispectrum and slow roll inflation*, *Mon. Not. Roy. Astron. Soc.* **313** (2000) 323, [astro-ph/9908009].

[64] Gangui, Alejandro and Martin, Jerome, *Best unbiased estimators for the three point correlators of the cosmic microwave background radiation*, *Phys. Rev. D* **62** (2000) 103004, [astro-ph/0001361].

[65] Casella, George and Berger, Roger, *Statistical Inference*. Duxbury Resource Center, 2001.

Chapter 4

Extracting Axion String Network Parameters from Simulated CMB Birefringence Maps using Convolutional Neural Networks

Abstract

Axion-like particles may form a network of cosmic strings in the Universe today that can rotate the plane of polarization of cosmic microwave background (CMB) photons. Future CMB observations with improved sensitivity might detect this axion-string-induced birefringence effect, thereby revealing an as-yet unseen constituent of the Universe and offering a new probe of particles and forces that are beyond the Standard Model of Elementary Particle Physics. In this work, we explore how spherical convolutional neural networks (SCNNs) may be used to extract information about the axion string network from simulated birefringence maps. We construct a pipeline to simulate the anisotropic birefringence that would arise from an axion string network, and we train SCNNs to estimate three parameters related to the cosmic string length, the cosmic string abundance, and the axion-photon coupling. Our results demonstrate that neural networks are able to extract information from a birefringence map that is inaccessible with two-point statistics alone (i.e., the angular power spectrum). We also assess the impact of noise on the accuracy of our SCNN estimators, demonstrating that noise at the level anticipated for Stage IV (CMB-S4) measurements would significantly bias parameter estimation for SCNNs trained on noiseless simulated data, and necessitate modeling the noise in the training data.

Notes about this project: This chapter is from a paper I wrote with Mustafa A. Amin and Andrew J. Long that is available on the arXiv preprint server at

2411.05002 [astro-ph.CO] (1). In chapter 3 I studied certain non-Gaussian statistical properties of LCM birefringence maps, namely their kurtosis and bispectrum. These statistics are conceptually easy to interpret, but may not be optimal at encoding information about LCM parameters. We were interested to know if a neural network could be trained to do a better job estimating LCM parameters by learning complicated statistical features in LCM birefringence maps.

4.1 Introduction

Precision measurements of the cosmic microwave background (CMB) radiation yield a wealth of data with which cosmologists are able to infer the constituents of the cosmos (2). The CMB's statistical properties provide compelling evidence for the presence of new physics, beyond the Standard Model of Particle Physics, such as dark matter and dark energy. The next generation of CMB telescopes is expected to reach unprecedented levels of precision, particularly in regard to polarization measurements (3). These measurements will provide an exciting opportunity to probe signatures of additional beyond the Standard Model physics that are inaccessible with current sensitivity levels (4).

In this work, we are interested in the cosmological signatures of hypothetical axion-like particles (ALPs) coupled to electromagnetism. We assume the standard Chern-Simons interaction, which takes the form

$$\mathcal{L}_{\text{int}} = -\frac{1}{4} g_{a\gamma\gamma} a F_{\mu\nu} \tilde{F}^{\mu\nu} \ , \tag{4.1}$$

where $g_{a\gamma\gamma}$ is the axion-photon coupling parameter, $a(x)$ is the pseudoscalar axion field, $F_{\mu\nu}(x)$ is the electromagnetic field strength tensor, and $\tilde{F}^{\mu\nu}(x)$ is the dual tensor. The coupling depends on the fine structure constant $\alpha_{\text{em}} \approx 1/137$, the axion decay constant f_a, and the electromagnetic anomaly coefficient $\mathcal{A}$ as $g_{a\gamma\gamma} = \mathcal{A}\alpha_{\text{em}}/(4\pi f_a)$. ALPs arise naturally in string theory as a consequence of the additional compact spatial dimensions (5; 6). In these theories, instanton effects lift the ALP potential

and their exponential sensitivity leads to a vast range of ALP masses spanning from nearly the Planck scale to far below the current Hubble scale $H_0 \approx 10^{-33}$ eV (7; 8). Such ALPs are also expected to interact with electromagnetism at a strength that can possibly be probed with astrophysical and cosmological observations (8).

One of the most well-studied cosmological signatures of ALPs is cosmic birefringence. As a linearly-polarized electromagnetic wave propagates through a varying ALP field, the plane of polarization is rotated by an angle $\alpha \propto g_{a\gamma\gamma}\Delta a$ that depends on the Chern-Simons coupling and the change in the ALP field (9–14). Measuring cosmic birefringence is an important science driver for current and future CMB experiments (3; 4; 15). Measurements of isotropic and anisotropic birefringence in upcoming CMB surveys are expected to improve by at least an order of magnitude (3; 16; 17). An exciting development in recent years is that a measurement of isotropic birefringence in CMB data has been reported with $\approx 3\sigma$ statistical significance (18–21). See ref. (22) a review of recent developments in the measurement of isotropic birefringence.

Various studies have explored the implications of axion-induced birefringence for cosmologically distance sources of polarized light like the CMB. While a homogeneous ALP field could induce isotropic cosmic birefringence (10; 12), there is theoretical motivation to consider configurations such as string and domain wall networks that can form from phase transitions in the early Universe (23). A remarkable feature of birefringence from string loops is that it does not directly depend on string tension and arises even if the string network is a subdominant component of the the total energy budget of the universe (24). Moreover, the anisotropic birefringence signal from a network, if detected, is likely to be more robust against calibration errors that might affect isotropic birefringence measurements.

Dynamics of topological axion defects and the birefringence induced by them have been studied in refs. (25; 26). CMB birefringence power spectra and non-Gaussian signatures from axion strings have been computed in refs. (24; 27; 28). Constraints

on axion-string parameters using published birefringence power spectra have been derived in refs. (29; 30). The potential of using radio emissions from spiral galaxies to probe birefringence caused by axion strings has been explored in ref. (31). A tomographic constraint on anisotropic birefringence generated at reionization was provided in ref. (32; 33). The birefringence from axion domain walls has been studied in refs. (34–37) and axion dark energy in ref. (38).

Recent advances in statistical learning and computing have made it possible to harness the power of neural networks in cosmology (39–47). In particular, spherical convolutional neural networks (SCNNs) designed for data with a spherical topology have been developed, such as those implemented in the Python package DeepSphere (48; 49). These networks have demonstrated capability in distinguishing cosmological models using simulated weak lensing all sky maps (50), and have been used for cosmological parameter inference in KiDS-1000 weak lensing maps (51).

In this work we explore how SCNNs can be used to estimate axion string network parameters from simulated all sky maps of CMB birefringence. To this end, we train SCNNs to estimate the parameters of a phenomenological model known as the loop-crossing model (LCM) (27), given simulated noiseless CMB birefringence maps. In previous works, measured birefringence power spectra have been compared against the power spectrum predicted from the LCM (29; 30), with the tightest constraint yielding $\mathcal{A}^2\xi_0 \lesssim 0.93$ at 95% confidence level. However, a fundamental limitation of this approach is that some model parameters are degenerate at the level of the power spectrum since in the LCM, the birefringence power spectrum is directly proportional to $\mathcal{A}^2\xi_0$. The LCM parameter ξ_0 is a phenomenological parameter that describes the energy density of the string network in a Hubble volume. This means that inference using only the power spectrum is fundamentally unable to independently measure $\mathcal{A}$ and ξ_0. To address this issue, the use of higher order statistics such as the trispectrum and wavelet scattering transform have been explored and were shown to break the degeneracy between $\mathcal{A}$ and ξ_0 (28; 52). The ability for SCNNs to learn statis-

tical properties of image data motivates the exploration of this tool for axion string parameter estimation.

Our strategy in this work is to train three spherical convolutional neural networks to estimate the LCM parameters ζ_0, $\mathcal{A}$, and ξ_0 using realizations of axion-string induced birefringences maps as training data. We then assess their performance by using approximate Bayesian computation to sample the posterior distributions. This pipeline is illustrated in fig. 4.1.1. Our results demonstrate that neural networks are a powerful tool that can be used to look for evidence of cosmic axion strings in future CMB polarization measurements. In addition to considering noiseless simulations, we explore the effect that adding noise to the maps has on the estimates produced by these neural networks.

4.2 Mock birefringence data simulation procedure

In this section we discuss how we generated mock data of the anisotropic birefringence arising from an axion string network by employing Loop Crossing Model (LCM) simulations. The LCM (53) treats all axion strings in the network to be circular planar loops whose positions are statistically homogeneous, whose orientations are statistically isotropic, and whose mean abundance and typical size evolve to scale with the cosmological expansion. The LCM is informed by numerical 3D lattice simulations of axion string network dynamics (54–63), which reveal that the network scales with the cosmological expansion (up to a possible logarithmic correction that remains under debate).

In our implementation, the LCM has four parameters: a size parameter ζ_0 related to the radius of string loops, an abundance parameter ξ_0 related to the number of string loops, the mass parameter m_a related to the string network collapse, and an intensity parameter $\mathcal{A}$ related to the amplitude of birefringence. At time t we assume that all loops in the network have the same (physical) radius $r(t) = \zeta_0\, d_H(t)$, which grows with time to scale with the increasing Hubble distance $d_H(t) = 1/H(t)$. The

Parameter	**Prior**
$\log_{10}\zeta_0$	$U\big(\log_{10}(0.3), \log_{10}(3)\big)$
$\log_{10}\mathcal{A}$	$U\big(\log_{10}(0.1), \log_{10}(1)\big)$
$\log_{10}\xi_0$	$U\big(\log_{10}(3), \log_{10}(30)\big)$

Table 4.2.1 : Priors used in inference and mock-data generation. The logarithms are drawn uniformly from the ranges shown.

average number density of loops also decreases as $n(t) = \xi_0 d_H^{-3}(t)/(2\pi\zeta_0)$ to maintain scaling. A logarithmic deviation from scaling would correspond to a growth in ξ_0 by a negligible factor between recombination and today, which we neglect. The axion mass m_a controls the time when the string network develops domain walls and collapses, through the relation $m_a \sim 3H(t)$ (27). In this work we restrict ourselves to masses $m_a \lesssim 4 \times 10^{-33}$ eV so that the string networks survive at least until today. Photons passing through the disk enclosed by a loop develop a birefringence rotation angle of $\Delta\alpha = \pm\mathcal{A}\alpha_{\rm em}$ where $\mathcal{A}$ is the electromagnetic anomaly coefficient, $\alpha_{\rm em} \approx 1/137$ is the electromagnetic fine structure constant, and the sign depends on the loop's winding number and orientation (24). As a photon propagates though multiple loops, its birefringence accumulates $\alpha = \sum \Delta\alpha$. Fig. 4.2.1 is a graphical illustration of the LCM and induced CMB birefringence; we indicate photons propagating through a network of circular planar string loops, shown in three redshift slices. In the bottom half of fig. 4.2.1 we show mollweide projections of the cumulative birefringence map from $z = 1100$ to the indicated redshift.

The procedure that we employ to create mock data is the following.

1. Draw a set of LCM model parameters ζ_0, $\mathcal{A}$, and ξ_0 from the prior distributions in tab. 4.2.1. We set $m_a = 0$. These priors are informed by theoretical expectations for the values these parameters may take. For example $\mathcal{A}$ is a sum over the squared electromagnetic charge of particles in the theory, so $\mathcal{A}$ is not

expected to be much smaller than an $\mathcal{O}(1)$ number. This partially motivates us to take our prior to have support in the region $\mathcal{A} \in [0.1, 1]$.

2. Select a HEALPix resolution parameter. For all of the work present in this article, we take $N_{\rm side} = 128$ corresponding to $N_{\rm pix} = 196,608$ pixels. This corresponds to an angular scale of approximately 0.5 degrees.

3. Calculate the average number of loops in the network $\langle N_{\rm loops} \rangle$ by using numerical methods to evaluate the integral (27)

$$\langle N_{\rm loops} \rangle = 2 \frac{\xi_0}{\zeta_0} \int_0^{z_{\rm cmb}} {\rm d}z \, H^2(z)(1+z)^{-3} \, s^2(z) \; . \tag{4.2}$$

Here $z_{\rm cmb} = 1100$ is the fiducial redshift at recombination, $s(z) \equiv \int_0^z H^{-1}(z') \, {\rm d}z'$ is the comoving distance from the observer (at $z = 0$) to redshift z, and $H(z) \equiv H_0 \sqrt{\Omega_r \, (1+z)^4 + \Omega_m \, (1+z)^3 + \Omega_\Lambda}$ is the Hubble parameter at redshift z. We assume an ΛCDM cosmology with $\Omega_r = 9 \times 10^{-5}$, $\Omega_m = 0.3$, and $\Omega_\Lambda = 0.7$.

4. Choose the number of loops in this realization by sampling $\hat{N}_{\rm loops} \sim {\rm Poisson}(\langle N_{\rm loops} \rangle)$.

5. To populate the network with loops, for each loop we generate a random position drawn uniformly from a 2-sphere, a random orientation drawn uniformly from a 2-sphere, a random winding number drawn uniformly from ± 1, and a random redshift z drawn from the probability density (see app. 4.A)

$$p(z) = \frac{H^2(z)(1+z)^{-3} \, s^2(z)}{\int_0^{z_{\rm cmb}} {\rm d}z' \, H^2(z')(1+z')^{-3} s^2(z')} \tag{4.3}$$

with support only on the interval $z \in (0, z_{\rm cmb})$. A derivation of eq. (4.3) can be found in app. 4.A. Once a random z is chosen, the circular loop's comoving radius is taken to be $r_{\rm co} = \zeta_0 (1+z)/H(z)$.

6. Using a HEALPix discretization scheme (64), for each loop find the pixels whose line of sight vectors pass through the interior of the loop. At these pixels increment their values by $\Delta\alpha = \pm \mathcal{A} \alpha_{\rm em}$.

The result of this procedure is a spatially-discretized birefringence map $\alpha_i = \alpha(\boldsymbol{n}_i)$ that takes values on each of the pixels $\boldsymbol{n}_i$.

4.3 Neural network architecture and training

Our mock data takes the form of an all-sky birefringence map, having a spherical topology. We therefore choose to use a neural network architecture which appropriately accounts for the geometry of the data. To this end, we use the DeepSpherePython package, which is a library for creating SCNNs (48; 49). Previous work has applied DeepSphereto cosmological mock data with HEALPix (50); and to KiDS-1000 weak lensing maps for parameter inference (51).

We train three neural networks, such that each of them is an estimator for one of the three LCM model parameters. Rather than directly learning ζ_0, ξ_0, and $\mathcal{A}$, we find that it is advantageous for the networks to instead learn

$$Z \equiv \log_{10}(\zeta_0) \ , \quad A \equiv \log_{10}(\mathcal{A}^2\xi_0) \ , \quad \text{and} \quad X \equiv \log_{10}(\xi_0^2/\mathcal{A}) \ . \tag{4.4}$$

This is the case for two reasons. First, the birefringence angular power spectrum in the LCM is directly proportional to $\mathcal{A}^2\xi_0$ and its shape is controlled by ζ_0 (53). This means that power spectrum information can be used to infer these parameters. On the other hand, $\xi_0^2/\mathcal{A}$ is orthogonal to $\mathcal{A}^2\xi_0$ in the $(\log_{10}\xi_0, \log_{10}\mathcal{A})$ plane so inferring this combination of parameters requires information beyond the power spectrum. Hence, training a neural network to estimate $\xi_0^2/\mathcal{A}$ will allow us to assess whether or not information beyond the power spectrum is being recovered. Second, since $\mathcal{A}^2\xi_0$ and $\xi_0^2/\mathcal{A}$ vary over several orders of magnitude, and because we care more about the neural networks' ability to provide accurate estimates to within an order of magnitude, training them to learn the base-10 logarithms of these parameters is a direct way to enforce this.

We use the architectures shown in tables 4.B.1, 4.B.2, and 4.B.3 which can be found in app. 4.B. We use 3 convolutional layers for ζ_0 with a total of 9,881 param-

eters since we found that this was sufficient for good performance on the range of parameters allowed by our priors (see tab. 4.2.1). For the A and X networks we used deeper networks with 6 to 8 convolutional layers and 20,113 to 3,566,273 trainable parameters respectively. The convolutions themselves are implemented using DeepSphere's `ChebyshevConv` layers which approximate the convolution kernels as a Chebyshev polynomial expansion in terms of the discrete Laplacian operator as explained in refs. (49; 65). These convolutions are approximately equivariant under rotations of the input map, i.e., a rotation of the map followed by a convolution is the same as a convolution followed by a rotation. After all the convolutions we use global average pooling (66) to ensure that the outputs of the neural networks are invariant under rotations of the input map.

For our mock data set we generate 20,000 axion-string-induced birefringence maps (with HEALPix pixelization) by performing repeated LCM simulations using the parameters ζ_0, ξ_0, and $\mathcal{A}$ sampled from the priors in tab. 4.2.1 and by following the procedure described in sec. 4.2. These data are then split into training and validation sets with a ratio of 80:20. Finally, we train three neural networks to provide estimators for Z, A, or X using a mean-squared-error loss (67). Training was performed using the Adam optimizer (68) with a learning rate schedule starting at 0.005, which decays by 4% every epoch. Training is stopped when the validation loss does not improve by at least 1×10^{-5} over 8 consecutive epochs. We define the trained neural network as the set of network parameters with the lowest validation loss.

To illustrate the performance of the three networks, we show an example in fig. 4.3.1. For this example, we take the LCM model parameters to be $\zeta_0 = 1$, $\mathcal{A} = 0.316$, and $\xi_0 = 10$, which correspond to $Z = 0$, $A \approx 0$, and $X \approx 2.5$. We generate four pixelated birefringence maps $\alpha(\boldsymbol{n}_i)$, which are random realizations of the LCM simulation procedure. These maps are then passed to the three SCNNs, which have already been trained, yielding estimates $\hat{Z}$, $\hat{A}$, and $\hat{X}$. In general we will use hatted variables to denote the outputs of the SCNNs, and unhatted variables to

denote the LCM model parameters. We emphasize that the neural network is deterministic, such that the same map $\alpha(\boldsymbol{n}_i)$ always yields the same estimate, $\hat{Z}$ for example. However, the procedure of generating the mock data via the LCM simulation is stochastic, so a single set of LCM model parameters generates many possible realizations $\alpha(\boldsymbol{n}_i)$. In this example, one can see that two of the SCNNs are performing very well; the $\hat{Z}$ and $\hat{A}$ estimates are close to the input model parameters Z and A. The third SCNN performs moderately well; the estimate $\hat{X}$ differs from the input parameter $X = \log_{10}(\xi_0^2/\mathcal{A}) = 2.5$ by as much as $\Delta X = 0.44$, which corresponds to a factor of 2.8 in $\xi_0^2/\mathcal{A}$. We discuss the SCNN performance further in the next section, where we also provide quantitative measures of success.

4.4 Validation of neural network performance

In order to assess the performance of the neural networks to provide accurate and precise estimates of the LCM model parameters, we perform the following two tests. To quantify the networks' accuracy, we directly compare SCNN parameter estimators with true LCM model parameters across a 2D slice of the parameter space. To quantify the networks' precision, we employ approximate Bayesian computation to sample the posterior distribution and infer the spread in the network output. In the following subsections, we discuss both approaches.

4.4.1 Estimator performance on known inputs

To get a sense of the reliability of the parameter estimates from our neural networks we want to quantify deviations from known inputs. This can be achieved by scanning the parameter space and calculating the average error at each point. To do this calculation we perform 992,600 draws of the LCM model parameters ζ_0, ξ_0, and $\mathcal{A}$ from the priors in tab. 4.2.1, and calculate the corresponding Z, A, and X using eq. (4.4). For each draw we do an LCM simulation and pass the mock data to the neural networks which return parameter estimates $\hat{Z}$, $\hat{A}$, and $\hat{X}$. We then divide the

(A, X) plane into $50 \times 50 = 2500$ bins. In each bin we compute the error magnitude defined as

$$\text{error magnitude} = \text{average of } \sqrt{(\hat{A} - A)^2 + (\hat{X} - X)^2}\,, \tag{4.5}$$

where the average is with respect to the samples in the bin. Some bins have no samples, because they are excluded by our priors; other bins contain between 65 and 809 samples.

The results of this analysis are shown in fig. 4.4.1. The color heat map shows the error magnitude in each bin. The white regions have no samples, because they are outside of our prior range, as indicated by the labeled values of $\mathcal{A}$ and ξ_0. The arrows indicate the displacement from the true value to the average estimated value. For example, an arrow pointing upward means that the neural networks tend to overestimate the value of X but are relatively accurate for A.

This figure offers several indications of the performance of the SCNNs. The error magnitude is typically smaller than 0.4 across most of the (A, X) parameter space, corresponding to a factor of $10^{0.4} \approx 2.5$ in the parameters $\mathcal{A}^2\xi_0$ and $\xi_0^2/\mathcal{A}$. However, the error predominantly arises from $\xi_0^2/\mathcal{A}$, as indicated by the mostly-vertical arrows. This poorer performance in the $X = \log_{10}(\xi_0^2/\mathcal{A})$ estimator was expected, and we discuss it further in sec. 4.4.3.

Notice that the region of parameter space in which X is large tends to have larger error magnitude (i.e., warmer colors). For larger ξ_0 the error is expected to grow, because the network contains a greater number of loops. The birefringence map generated by many overlapping loops appears increasingly like a Gaussian random field (28), and obscures information about the string network. This leads to a larger error magnitude.

The performance of the network to accurately estimate X is particularly poor at the upper-left boundary. This is indicated by the bright orange cells on the heat map and the long downward arrows. Here the error magnitude reaches 0.8 corresponding to underestimating $\xi_0^2/\mathcal{A}$ by a factor of $10^{0.8} \approx 6.3$.

The X neural network tends to overpredict when the input value is small (i.e., upward arrows near the bottom of the figure) and underpredict when the input value is large (i.e., downward arrows near the top). Consequently, there is a region of parameter space where the predictions are particularly accurate – this is the very dark region of the plot. If the network were trained again using a new set of randomly-generated training data, we anticipate that the general trends seen in fig. 4.4.1 would persist, while the particular values of the error magnitude would change.

4.4.2 Approximate Bayesian computation

We designed the SCNNs to provide a single estimator of the LCM model parameters (rather than a probability density). In this section, we discuss how approximate Bayesian computation (ABC) (69) can be employed to ascribe an uncertainty to those estimates. More formally, the uncertainty of the network is quantified by the posterior distribution $p(\theta|\hat{\theta})$, where $\theta \equiv (Z, A, X)$ are the parameters and $\hat{\theta} \equiv (\hat{Z}, \hat{A}, \hat{X})$ are the estimates of the parameters. To obtain samples from $p(\theta|\hat{\theta})$ we use ABC, following the procedure outlined below.

1. Define a distance measure between parameter estimates $\hat{\theta}$ and a target value $\hat{\theta}_{\text{target}}$ as

$$\rho(\hat{\theta}, \hat{\theta}_{\text{target}}) = \sqrt{(\hat{Z} - \hat{Z}_{\text{target}})^2 + (\hat{A} - \hat{A}_{\text{target}})^2 + (\hat{X} - \hat{X}_{\text{target}})^2} \, .$$

2. Pick a tolerance ϵ. We use $\epsilon = 0.07$, which was found to be sufficient for convergence.

3. Sample LCM model parameters ζ_0, $\mathcal{A}$, and ξ_0 from the priors in tab. 4.2.1, and compute the corresponding $\theta = (Z, A, X)$ using eq. (4.4).

4. Generate a pixelated LCM birefringence map $\alpha(\boldsymbol{n}_i)$ following the procedure in sec. 4.2.

5. Pass the map through each neural network to obtain estimates $\hat{\theta} = (\hat{Z}, \hat{A}, \hat{X})$.

6. Keep the sampled θ if $\rho(\hat{\theta}, \hat{\theta}_{\text{target}}) < \epsilon$.

7. Repeat steps $3 - 6$ until a desired number of accepted samples is reached.

Fig. 4.4.2 shows the outcome of this procedure for a representative parameter point. We show samples from the posterior distribution for $\hat{\theta}_{\text{target}} = (0, 0, 2.5)$, which corresponds to $\hat{\zeta}_0 = 1$, $\hat{\xi}_0 = 10$, and $\hat{\mathcal{A}} = 0.316$. The three lower subplots show the 2D marginal posteriors for each pair of parameters, and the three upper subplots show the 1D marginal posterior (black histogram) and prior (gray dashed histogram) of each individual parameter. The red cross indicates the target value $\hat{\theta}_{\text{target}}$.

From the 2D posteriors, we can assess possible correlations between estimator errors. No significant correlations are observed. Since the networks are trained independently we expect uncorrelated errors.

From the 1D posteriors, we can assess the SCNNs' uncertainties. For each of the three estimators, the posteriors are approximately centered at the target value, which is an indication of the SCNNs' accuracy. The posteriors have standard deviations of $\sigma_Z = 0.03$, $\sigma_A = 0.014$, and $\sigma_X = 0.3$. This means that the ζ_0 parameter is typically within a factor of $10^{0.03} \approx 1.07$ of the target value; the $\mathcal{A}^2\xi_0$ parameter is typically within a factor of $10^{0.014} \approx 1.03$ of the target value; and the $\xi_0^2/\mathcal{A}$ parameter is typically within a factor of $10^{0.3} \approx 2$ of the target value. For ζ_0 and $\mathcal{A}^2\xi_0$ the SCNNs are quite precise with uncertainties below 10% (for this example). For $\xi_0^2/\mathcal{A}$ the uncertainty is much larger. Nevertheless, all three 1D posteriors, even $\xi_0^2/\mathcal{A}$, have standard deviations that are smaller than the prior distribution's as evidenced by the black histograms being narrower than the grey lines in fig. 4.4.2.

4.4.3 Interpretation of results

The results shown in fig. 4.4.1 and fig. 4.4.2 indicate that the neural networks have learned ways to extract information to infer these parameters. This is especially interesting for $X = \log_{10}(\xi_0^2/\mathcal{A})$ since part of our motivation for this work was to

see if SCNNs can break the parameter degeneracy between $\mathcal{A}$ and ξ_0 that is present in power spectrum-only analyses (28; 52). The stronger predictive power of the Z and A neural networks over the X network can be understood in the following way: unlike for Z and A, there is no information contained in the power spectrum which allow one to discern between different values of X. Therefore, the only way for the X estimator to learn useful information is to extract information beyond the power spectrum, which is harder because it requires the network to learn more complex patterns.

4.5 Estimator degradation due to noise

We trained our neural networks to provide parameter estimates on noiseless maps. However, real birefringence maps are reconstructed from CMB polarization data. This means that birefringence maps will have noise sourced from the Q and U measurements as well as the statistical estimators used to reconstruct the birefringence. A popular technique for birefringence reconstruction is to use quadratic estimators (70–72). For example, POLARBEAR, ACT, and SPT have all used quadratic estimators in their analysis of anisotropic birefringence (73–75). For a more general discussion of CMB birefringence quadratic estimators and their statistical reconstruction noise see (29). In this section we explore how our neural networks, which were trained on noiseless maps, can perform on noisy data.

We model the noisy pixelated birefringence maps as $\alpha_S(\boldsymbol{n}_i) + \alpha_N(\boldsymbol{n}_i)$ where the signal $\alpha_S(\boldsymbol{n}_i)$ is generated from the LCM simulation. The noise $\alpha_N(\boldsymbol{n}_i)$ is assumed to be a Gaussian random field drawn from the angular power spectrum N_ℓ, uncorrelated with the signal. We assume $N_\ell \propto \ell^0$ is constant, corresponding to white noise. This choice is motivated by the noise being dominated by reconstruction noise (29). We study different noise levels by multiplying the expected CMB-S4 noise level $N_\ell^{\mathrm{CMB-S4}} = 1.5 \times 10^{-5}$ deg^2, which we obtained from figure 1 of ref. (29), by a multiplicative factor. We use the `synfast` method of `healpy` to generate $\alpha_N(\boldsymbol{n}_i)$.

An example of the signal and noise maps appears in fig. 4.5.1.

The histograms in fig. 4.5.2 show the distribution of estimators for each neural network given LCM parameters $\zeta_0 = 1$, $\mathcal{A} = 0.316$, and $\xi_0 = 10$, which correspond to $Z = 0$, $A = 0$, and $X = 2.5$. The networks were trained on noiseless maps and tested on birefringence maps with various noise levels. The results for each estimator are shown across three panels. In each panel we depict the input LCM parameter value with a vertical black line. In general the estimators tend to give biased predictions with comparable variances when provided with noisy data.

In order from top to bottom the panels show our results for the Z, A, and X estimators. The Z estimator exhibits a bias toward smaller values than the true parameter at $Z = 0$. This is shown by the fact that the centers of the distributions move to smaller values with increasing noise. For example, the prediction for $0.5\,N_\ell^{\mathrm{CMB-S4}}$ is centered at $Z \approx -0.45$, whereas the prediction for $N_\ell^{\mathrm{CMB-S4}}$ is centered at around $Z \approx -0.65$. This can be understood as the neural network interpreting the noise as part of the signal. Since the noise power spectrum is approximately a power law with a positive index over the range of multipoles that can be probed with the map resolution of $N_{\mathrm{side}} = 128$, the noise looks like many small loops about the size of a pixel. In a similar way, the A estimator biases its predictions to larger values than the true value. Again, this is to be expected since the neural network interprets the noise as part of the signal. At the level of the power spectrum the noise is additive: $C_\ell^{\mathrm{tot}} = C_\ell^{\mathrm{signal}} + C_\ell^{\mathrm{noise}}$, which can partially account for the trend in the prediction bias. The X estimator is the most affected by the addition of noise. Like the A estimator, its predictions are biased above the true value. However, we see that while the bias for the Z and A estimators is less than an order of magnitude for all noise levels, the X estimator's bias is larger than the other estimators' bias at every noise level. This upward bias is expected since X is proportional to ξ_0^2, which controls the number density of loops. If the noise is interpreted as signal it will have the appearance of adding many loops on small scales, leading to an upward bias.

In order for all three estimators to be biased by no more than an order of magnitude, one would need an experiment with about 10 times less noise than CMB-S4 as shown by the yellow histograms in fig. 4.5.2. This analysis reveals that our noiseless estimators prove inadequate to provide reliable estimates at the noise level of a CMB-S4-like experiment. Therefore methods for improving the estimators are required. A discussion of some of these can be found in sec. 4.6.

4.6 Summary and conclusion

In this work, we have explored how spherical convolutional neural networks (SCNNs) can be used to perform parameter inference on simulated axion-string-induced birefringence maps. Below is a summary of the main contributions and results of this work:

- We developed a pipeline for generating HEALPix maps of anisotropic birefringence based on a simulation of the loop-crossing model (LCM), which has phenomenological parameters ζ_0, ξ_0, and $\mathcal{A}$, and which control the radius of string loops, the number of string loops, and the birefringence accumulated by a photon after passing through a single loop, respectively. The LCM parameterization is informed by string network dynamics.

- We trained three independent SCNNs to estimate the parameters $Z \equiv \log_{10} \zeta_0$, $A \equiv \log_{10} \mathcal{A}^2 \xi_0$, and $X \equiv \log_{10} \xi_0^2/\mathcal{A}$ from these maps. The choice of these parameters is driven by the following considerations. First, the birefringence power spectrum from axion strings is proportional to $\mathcal{A}^2\xi_0$, so it is natural to train our SCNNs to learn this combination of parameters and another, $\xi_0^2/\mathcal{A}$ for which the power spectrum provides no information. Second, some parameters can take values over several orders of magnitude. This motivates us to train the SCNNs to infer the base-10 log of the parameters.

- We evaluated the SCNN estimators' performance on birefringence maps generated from known LCM model parameters (see fig. 4.4.1), finding that the networks performed very well for Z and A, and moderately well for X. For example, the Z and A estimators are typically biased by no more than 0.04 units, which corresponds to a factor of $10^{0.04} \approx 1.10$ deviation from the true value of ζ_0 or $\mathcal{A}^2\xi_0$. On the other hand, depending on the input value the X estimator is typically biased by less than 0.3 units corresponding to a factor of $10^{0.3} \approx 2$ error in $\xi_0^2/\mathcal{A}$ – although the error can be as high as a factor of 6 for some values of $\xi_0^2/\mathcal{A}$.

- We used approximate Bayesian computation to sample the posterior distribution for the case when the estimators yield $\hat{Z} = 0$, $\hat{A} = 0$, and $\hat{X} = 2.5$, which allowed us to quantify the statistical uncertainty of the predictions (see fig. 4.4.2). The estimates for Z and A had standard deviations of $\sigma_Z = 0.03$ and $\sigma_A = 0.014$, respectively. The X estimator had a larger uncertainty with $\sigma_X = 0.3$.

- By simulating CMB birefringence maps with various noise levels we demonstrated that the accuracy of the neural networks is degraded in a qualitatively predictable way.

Our work was motivated in part by the question: can a neural network learn information about a birefringence map that is inaccessible with its power spectrum alone? In particular, in the context of the loop crossing model, the birefringence angular power spectrum only depends upon the intensity and abundance parameters, $\mathcal{A}$ and ξ_0, through the combination $\mathcal{A}^2\xi_0$, but not the combination $\xi_0^2/\mathcal{A}$. Based on several tests of their accuracy and precision, we conclude that the SCNNs that learned $\mathcal{A}^2\xi_0$ and $\xi_0^2/\mathcal{A}$ are able to furnish reliable estimators of these parameters when provided with noiseless birefringence maps generated from LCM simulations. Of course the uncertainty in $\xi_0^2/\mathcal{A}$ is much larger than in $\mathcal{A}^2\xi_0$, approximately a factor of 2 versus 3% for the sample parameter point illustrated in fig. 4.4.2. This

was expected, since information about $\xi_0^2/\mathcal{A}$ is not encoded in the power spectrum, but rather stored in higher-point correlations. Nevertheless, our priors based on UV considerations and string network simulations allow the parameters $\mathcal{A}$ and ξ_0 to each vary by about an order of magnitude, and even the factor of 2 uncertainty in $\xi_0^2/\mathcal{A}$ is informative, relatively speaking.

Our work demonstrates that even with an unoptimized architecture and inference pipeline, SCNNs are capable of extracting information beyond the power spectrum. We anticipate that better performance (particularly in extracting $\xi_0^2/\mathcal{A}$) could be achieved with an alternative architecture and/or extended inference pipeline. For example, we have trained estimators using mean-squared-error loss which may allow for biases, but the method presented in ref. (76) uses deep summaries to construct minimally-biased point estimators.

These results are a step toward a pipeline that can be used to do inference on real data, but there are areas available for improvement. In sec. 4.5 we found that noise introduces additional biases in the predictions, particularly for X, which exhibited biases as large as two orders of magnitude for CMB-S4 noise levels. We found that to yield parameter estimates within an order of magnitude from the true value, one would need an experiment with 10% of the noise level compared to CMB-S4. To address noise more effectively, one could explore two potential strategies. First, neural networks could (and should!) be trained on maps that already include noise specific to the experimental setup. This would allow the model to learn both the signal and the noise features, leading to more robust predictions. Additionally, one could implement techniques from (46), which demonstrated improvements by using a combination of max/average pooling and random permutations in the deeper layers of the network. These regularization and data augmentation techniques help the network generalize better in noisy environments by preventing overfitting to large-scale correlations that might include noise. We leave the exploration of these techniques to future work.

This material is based upon work supported (in part: R.H. and A.J.L.) by the

National Science Foundation under Grant Nos. PHY-2114024 and PHY-2412797. M.A.A. is supported by a DOE grant DOE-SC0021619. Some of the results in this paper have been derived using the healpy and HEALPix packages. A.J.L. thanks Bhuvnesh Jain and Matthew Johnson for illuminating discussions of machine learning methods. R.H. would like to thank Siyang Ling for his invaluable contributions to the LCM simulation code, and Juehang Qin, Dorian Amaral, and Ivy Li for useful conversations about statistical learning.

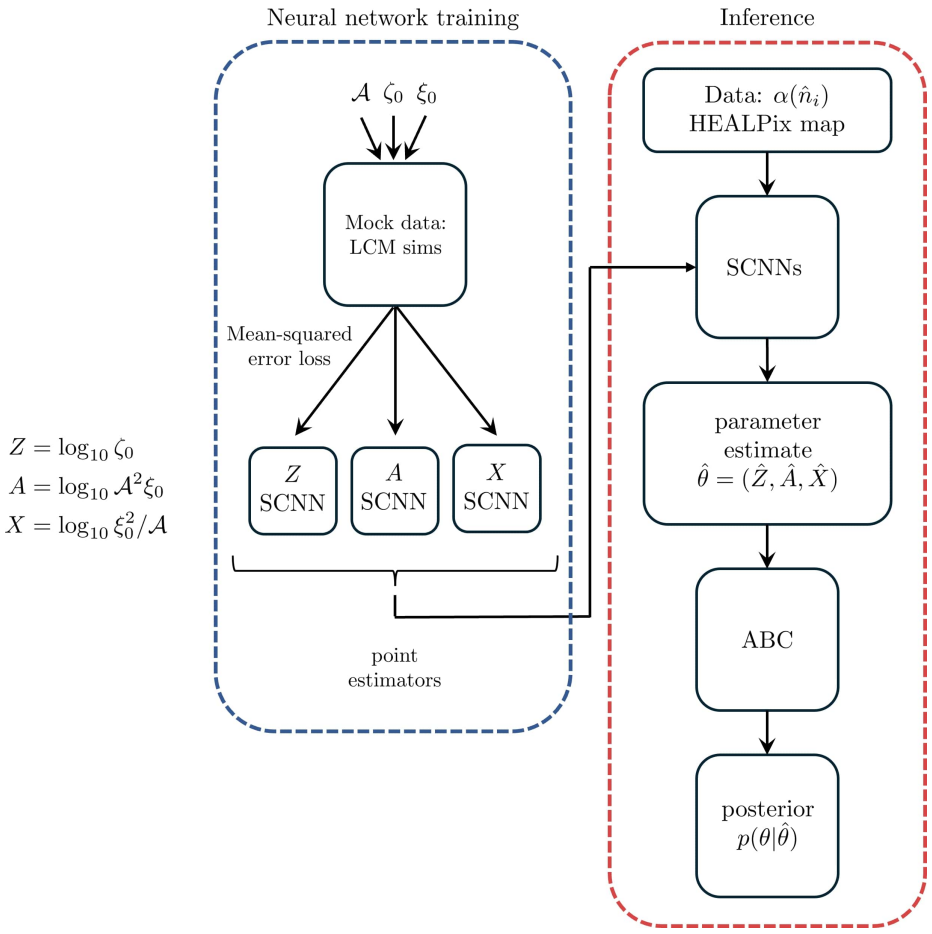


Figure 4.1.1 : Illustration of our neutral network training and inference pipelines. Mock data is generated by performing simulations of the loop crossing model (LCM), which has parameters $\mathcal{A}$, ζ_0, and ξ_0. Data takes the form of a pixelated map of birefringence rotation angles $\alpha(\boldsymbol{n}_i)$ in HEALPix format. Mock data is used to train three spherical convolutional neural networks (SCNNs), which return point estimates of $Z = \log_{10}(\zeta_0)$, $A = \log_{10}(\mathcal{A}^2\xi_0)$, and $X = \log_{10}(\xi_0^2/\mathcal{A})$. We validate the training of the SCNNs using two checks of their parameter inference. One method is approximate Bayesian computation (ABC), which calculates the posterior over the LCM model parameters.

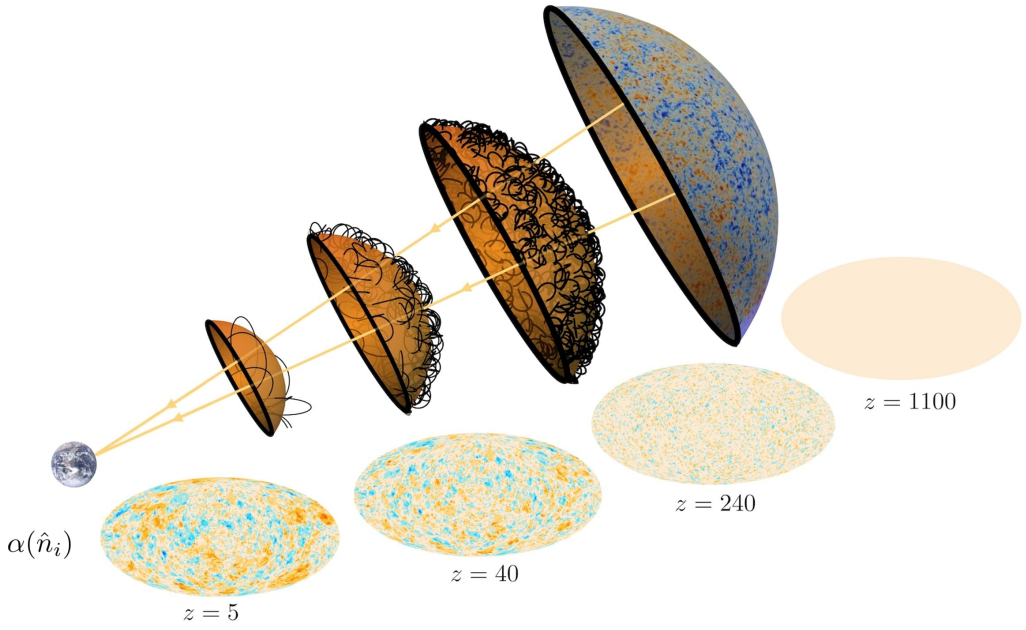


Figure 4.2.1 : Illustration of birefringence accumulation in the loop crossing model (LCM). CMB photons (yellow arrowed lines) propagate from the surface of last scattering ($z = 1100$) to Earth ($z = 0$). The intervening space is filled with an LCM string network consisting of circular planar loops with statistically homogeneous positions and statistically isotropic orientations. Each time a photon passes through a loop its plane of polarization incurs a rotation of $\Delta\alpha \pm \mathcal{A}\alpha_{\rm em}$ depending on the orientation of the loop. Birefringence accumulates over time with multiple loop crossings. Three redshift slices ($z = 240$, $z = 40$, and $z = 5$) are illustrated, showing a possible realization of the string network (orange shells with black circles) and the accumulated birefringence (mollweide projections). This graphic is figure 4 of ref. (30), and we reproduce it here with permission from the authors.

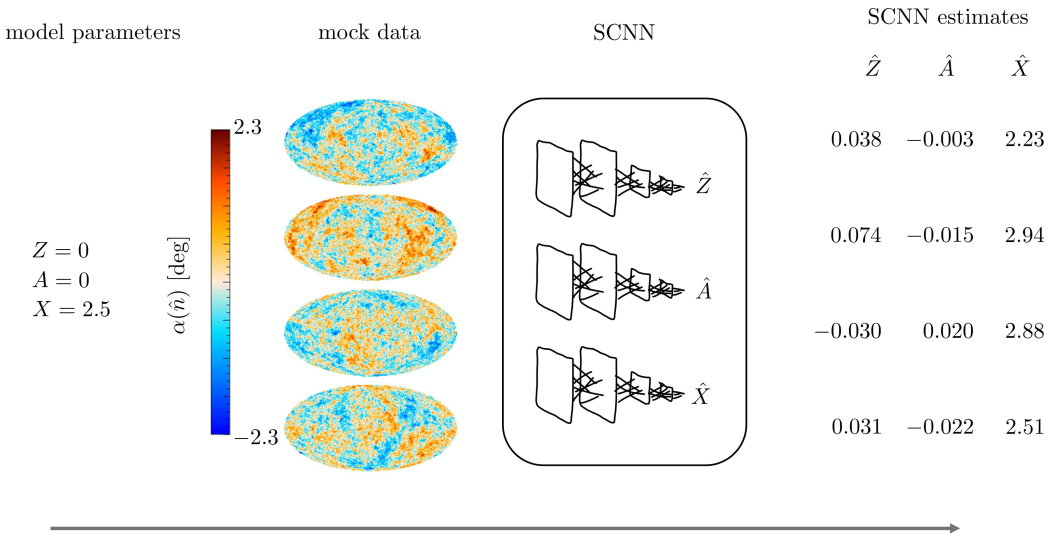


Figure 4.3.1 : Illustration of LCM parameter estimation using SCNNs on simulated birefringence maps. From left to right we show, (1) a set of LCM model parameters $Z = 0$, $A = 0$, and $X = 2.5$, corresponding to $\zeta_0 = 1$, $\mathcal{A} = 0.316$, and $\xi_0 = 10$; (2) four realizations of birefringence mock data generated using LCM simulation; (3) graphical depiction of our three SCNNs; and (4) parameter estimates furnished by each of the three SCNNs for each of the four maps.

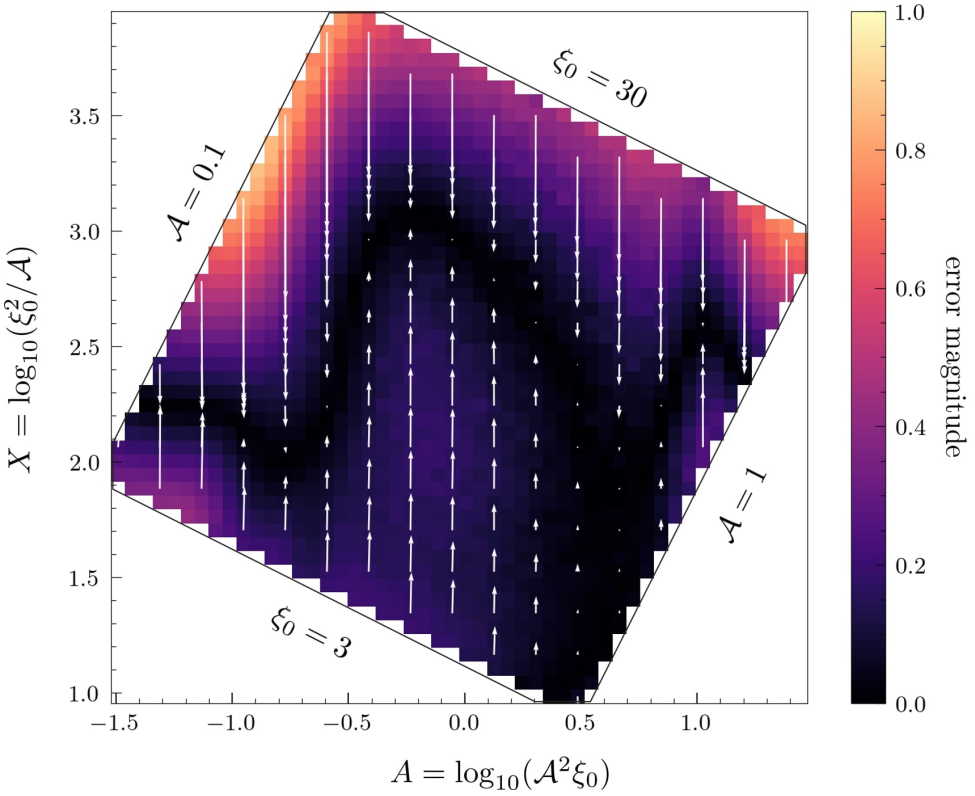


Figure 4.4.1 : An illustration of the performance of our trained SCNNs. We show the error magnitude (4.5) as a colored heatmap where cooler/darker colors indicate better performance. We also show the displacement from input parameter pair to the average output of the SCNNs as a white arrow. There are no samples in the white regions, which are outside of our prior range, as indicated by the diagonal labels; see also tab. 4.2.1.

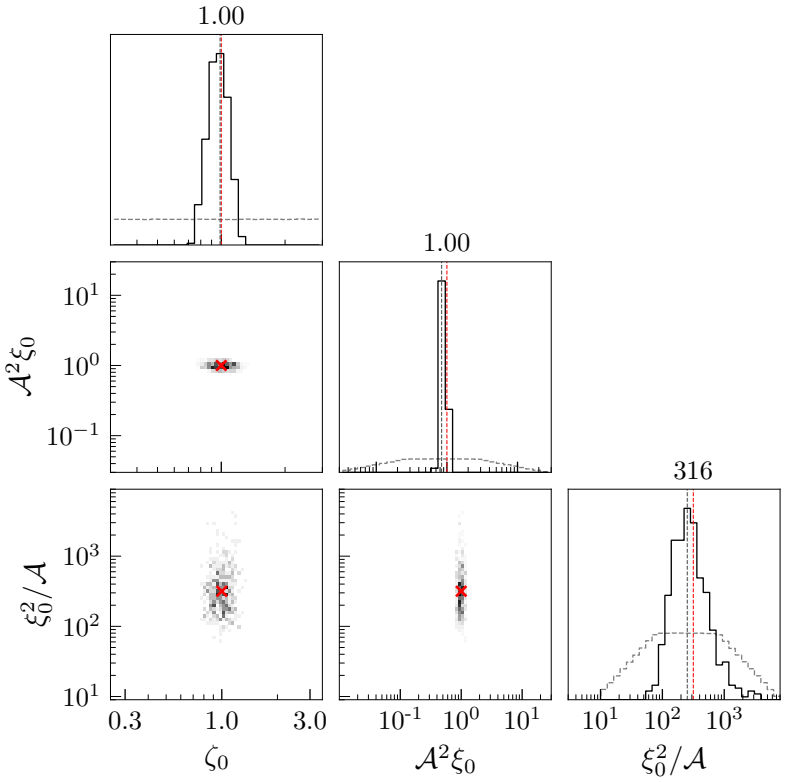


Figure 4.4.2 : Posteriors on the parameters $Z = \log_{10}(\zeta_0)$, $A = \log_{10}(\mathcal{A}^2\xi_0)$, and $X = \log_{10}(\xi_0^2/\mathcal{A})$ obtained using ABC sampling when $\hat{\theta}_{\text{target}} = (\hat{Z},\ \hat{A},\ \hat{X}) = (0,\ 0,\ 2.5)$. The diagonal subplots show the 1D marginal posteriors. The gray dashed curves in the diagonal subplots depict the prior for each parameter. Gray vertical lines depict the parameter with the highest posterior density. Red vertical lines depict $\hat{\theta}_{\text{target}}$. Lower-left subplots show a histogram of 2D marginal posteriors. Red crosses depict the projection of $\hat{\theta}_{\text{target}}$ in each plane.

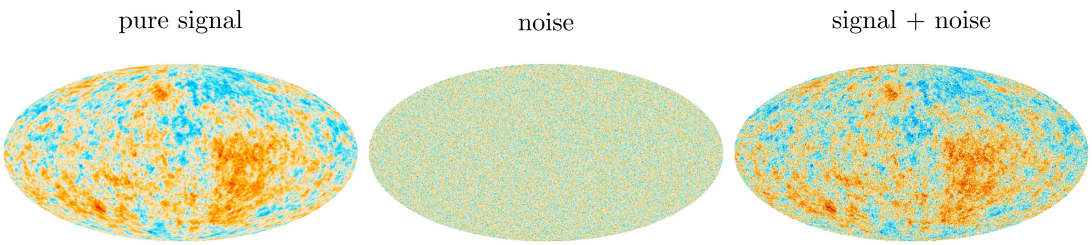


Figure 4.5.1 : *Left:* a realization of a CMB birefringence map generated with LCM parameters $\zeta_0 = 1.0$, $\mathcal{A} = 0.316$, and $\xi_0 = 10.0$. *Middle:* a realization of Gaussian random noise for a CMB-S4-like experiment. *Right:* combined signal and noise. Since the noise power spectrum follows a white noise profile, the noise level increases toward smaller angular scale. As a result, the granularity of the noise in the map shown is determined by the resolution of the map. A higher resolution would give the appearance of smaller scale fluctuations.

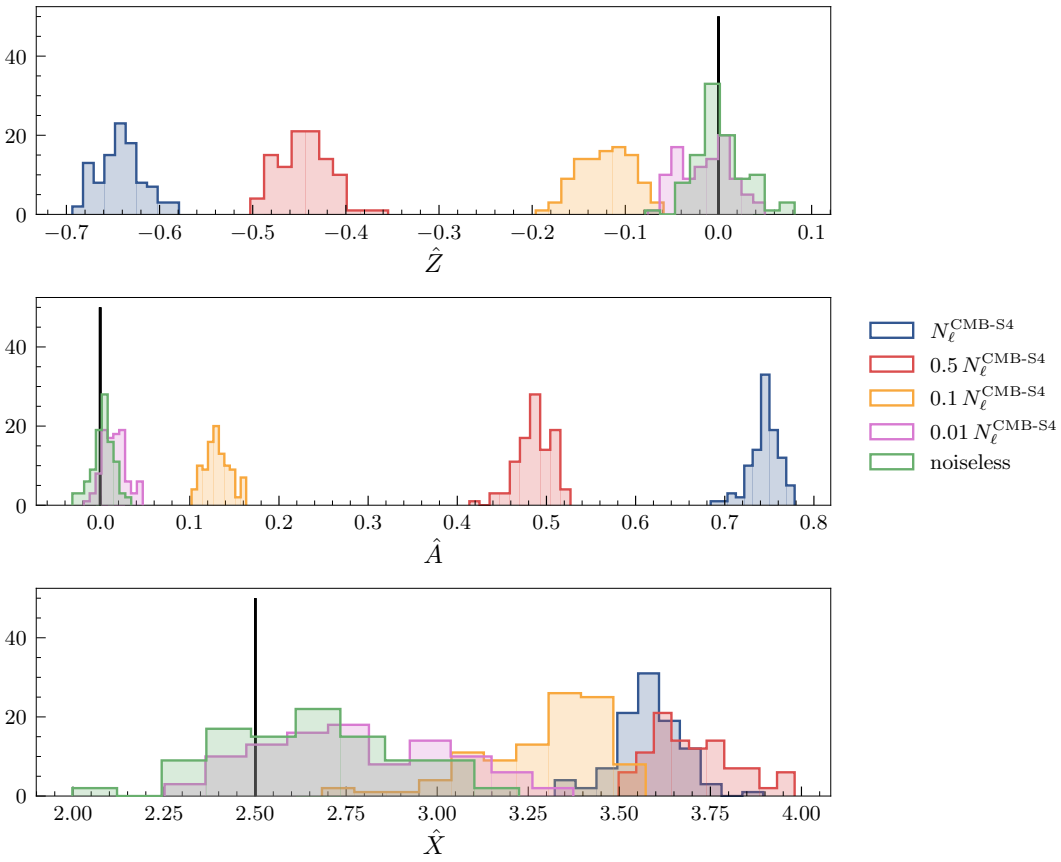


Figure 4.5.2 : Performance of our SCNNs on noisy mock data. For each noise level (colored histograms) we show the sample distribution over the SCNN estimators $\hat{Z}$ (top), $\hat{A}$ (middle), and $\hat{X}$ (bottom). To generate each histogram, we perform 100 LCM simulations with $\zeta_0 = 1$, $\xi_0 = 10$, and $\mathcal{A} = 0.316$ (corresponding to $Z = 0$, $A = 0$, and $X = 2.5$, indicated by the vertical black bar), add white noise up to the level of a CMB-S4-like experiment, and pass these noisy birefringence maps to our SCNNs, which were trained on noiseless mock data.

Appendix

4.A Loop redshift probability density

This appendix provides a derivation of eq. (4.3), which gives the probability density $p(z)$ to find a string loop at redshift z. In the LCM the average number of string loops with comoving radius between r and $r + \mathrm{d}r$, with comoving position between $\vec{s}$ and $\vec{s} + \mathrm{d}\vec{s}$ on the observer's past light cone, and with orientation (normal to the plane of the loop) between $\hat{k}$ and $\hat{k} + \mathrm{d}\hat{k}$ is given by (27) $\Big[\mathrm{d}N = \nu(r,z)\,\mathrm{d}r\,\mathrm{d}^3\vec{s}\,\frac{\mathrm{d}^2\hat{k}}{4\pi}\,.\Big]\; The assumption that loops are oriented isotropically implies that \nu$ is independent of $\hat{k}$, and the assumption that the loops are distributed homogeneously throughout space at a given time implies that ν only depends on $\vec{s}$ through $s = |\vec{s}|$, which is a proxy for time (or redshift) on the past light cone. In spherical polar form we have $\mathrm{d}^3\vec{s} = s^2\,\mathrm{d}^2\boldsymbol{n}\,\mathrm{d}s$, where $\boldsymbol{n}$ is the unit vector pointing in the direction of $\vec{s}$. To write $\mathrm{d}^3\vec{s}$ in terms of z note that $s(z) = \int_0^z H^{-1}(z')\mathrm{d}z'$. Hence, $\mathrm{d}^3\vec{s} = s^2(z)\,H^{-1}(z)\,\mathrm{d}z\,\mathrm{d}^2\boldsymbol{n}$.

In order to more conveniently parameterize the scaling property of string loop networks we introduce a kernel function $\chi(\zeta, z)$ as in ref. (27) such that,

$$\nu(r,z) = \int_0^\infty \mathrm{d}\zeta\,\chi(\zeta,z)\;\frac{H(z)^2(1+z)^{-2}}{2\pi r}\,\delta\big[r - \zeta(1+z)/H(z)\big]\;. \tag{4.6}$$

To implement the assumption that all string loops have the same radius at a given time, we take $\chi(\zeta, z) = \xi_0\,\delta(\zeta - \zeta_0)$. Integrating over the loop radius and orientation gives the average number of loops with comoving position between $\vec{s}$ and $\vec{s} + \mathrm{d}\vec{s}$ to be

$$\mathrm{d}N \;=\; \int_0^\infty \mathrm{d}r \int_{4\pi} \frac{\mathrm{d}^2\hat{k}}{4\pi}\,\nu(r,z)\,\mathrm{d}^3\vec{s} \;=\; \frac{\xi_0}{2\pi\zeta_0}\Big(H(z)^2(1+z)^{-3}\,s^2(z)\,\mathrm{d}z\Big)\mathrm{d}^2\boldsymbol{n}\;. \tag{4.7}$$

The expression on the right implicitly defines a probability density over z which is obtained by normalizing it over the range of possible redshift values. Hence, taking

$z_{\rm cmb} = 1100$, we have

$$p(z) = \frac{H^2(z)(1+z)^{-3}\, s^2(z)}{\int_0^{z_{\rm cmb}} H^2(z)(1+z)^{-3}\, s^2(z)\, \mathrm{d}z} \,,$$

which appears in eq. (4.3).

4.B Neural network architectures

This appendix contains the architectures used for our SCNNs in tables 4.B.1-4.B.3. To construct our neural networks we use the Python package DeepSphere (48; 49). This package provides implementations of layers designed for use on HEALPix formatted maps. These include the `ChebyshevConv` and `MaxPool` layers which perform convolutions and pooling.

Bibliography

[1] Hagimoto, Ray and Long, Andrew J. and Amin, Mustafa A., *Extracting Axion String Network Parameters from Simulated CMB Birefringence Maps using Convolutional Neural Networks*, 2411.05002.

[2] Planck collaboration, Aghanim, N. and others, *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6, [1807.06209].

[3] CMB-HD collaboration, Aiola, Simone and others, *Snowmass2021 CMB-HD White Paper*, 2203.05728.

[4] Chang, Clarence L. and others, *Snowmass2021 Cosmic Frontier: Cosmic Microwave Background Measurements White Paper*, 2203.07638.

[5] Svrcek, Peter and Witten, Edward, *Axions In String Theory*, *JHEP* **06** (2006) 051, [hep-th/0605206].

Layer type	**Output shape**	**Parameters**
`Input`	$(N_b, N_{\text{pix}}, 1)$	0
`ChebyshevConv (K=32, Fout=40)`	$(N_b, N_{\text{pix}}, 40)$	840
`MaxPool (p=2)`	$(N_b, 12288, 40)$	0
`ChebyshevConv (K=10, Fout=20)`	$(N_b, 12288, 20)$	8020
`MaxPool (p=1)`	$(N_b, 768, 20)$	0
`ChebyshevConv (K=5, Fout=10)`	$(N_b, 768, 10)$	1010
`GlobalAvgPool`	$(N_b, 10)$	0
`Flatten`	$(N_b, 10)$	0
`Dense`	$(N_b, 1)$	11

Table 4.B.1 : Neural network architecture used for the $Z = \log_{10}(\zeta_0)$ estimator. The batch size is $N_b = 32$ and the HEALPix resolution parameter is $N_{\text{side}} = 128$, so the number of pixels is $N_{\text{pix}} = 196,608$. All layers use ReLU activations and batch normalization is disabled (`use_bn=False`).

Layer type	**Output shape**	**Parameters**
`Input`	$(N_b, N_{\text{pix}}, 1)$	0
`ChebyshevConv (K=5, Fout=8)`	$(N_b, N_{\text{pix}}, 8)$	48
`MaxPool (p=2)`	$(N_b, 12288, 8)$	0
`ChebyshevConv (K=5, Fout=16)`	$(N_b, 12288, 16)$	656
`ChebyshevConv (K=5, Fout=16)`	$(N_b, 12288, 16)$	1296
`MaxPool (p=1)`	$(N_b, 3072, 16)$	0
`ChebyshevConv (K=5, Fout=32)`	$(N_b, 3072, 32)$	2592
`ChebyshevConv (K=5, Fout=32)`	$(N_b, 3072, 32)$	5152
`MaxPool (p=1)`	$(N_b, 768, 32)$	0
`ChebyshevConv (K=5, Fout=64)`	$(N_b, 768, 64)$	10304
`GlobalAvgPool`	$(N_b, 64)$	0
`Flatten`	$(N_b, 64)$	0
`Dense`	$(N_b, 1)$	65

Table 4.B.2 : Neural network architecture used for the $A = \log_{10}(\mathcal{A}^2\xi_0)$ estimator. The batch size is $N_b = 32$ and the HEALPix resolution parameter is $N_{\text{side}} = 128$, so the number of pixels is $N_{\text{pix}} = 196,608$. All layers use ReLU activations and batch normalization is disabled (`use_bn=False`).

Layer type	**Output shape**	**Parameters**
`Input`	$(N_b, N_{\rm pix}, 1)$	0
`ChebyshevConv (K=5, Fout=16)`	$(N_b, N_{\rm pix}, 16)$	96
`ChebyshevConv (K=5, Fout=32)`	$(N_b, N_{\rm pix}, 32)$	2592
`MaxPool (p=2)`	$(N_b, 12288, 32)$	0
`ChebyshevConv (K=5, Fout=64)`	$(N_b, 12288, 64)$	10432
`ChebyshevConv (K=5, Fout=128)`	$(N_b, 12288, 128)$	41344
`MaxPool (p=1)`	$(N_b, 3072, 128)$	0
`ChebyshevConv (K=5, Fout=256)`	$(N_b, 3072, 256)$	164608
`ChebyshevConv (K=5, Fout=512)`	$(N_b, 3072, 512)$	656896
`MaxPool (p=1)`	$(N_b, 768, 512)$	0
`ChebyshevConv (K=5, Fout=512)`	$(N_b, 768, 512)$	1312256
`ChebyshevConv (K=5, Fout=512)`	$(N_b, 768, 512)$	1312256
`GlobalAvgPool`	$(N_b, 512)$	0
`Flatten`	$(N_b, 512)$	0
`Dense`	$(N_b, 128)$	65664
`Dense`	$(N_b, 1)$	129

Table 4.B.3 : Neural network architecture used for the $X = \log_{10}(\xi_0^2/\mathcal{A})$ estimator. The batch size is $N_b = 8$ and the HEALPix resolution parameter is $N_{\rm side} = 128$, so the number of pixels is $N_{\rm pix} = 196,608$. All layers use ReLU activations and batch normalization is disabled (`use_bn=False`).

[6] Arvanitaki, Asimina and Dimopoulos, Savas and Dubovsky, Sergei and Kaloper, Nemanja and March-Russell, John, *String Axiverse*, *Phys. Rev. D* **81** (2010) 123530, [0905.4720].

[7] Hui, Lam and Ostriker, Jeremiah P. and Tremaine, Scott and Witten, Edward, *Ultralight scalars as cosmological dark matter*, *Phys. Rev. D* **95** (2017) 043541, [1610.08297].

[8] Gendler, Naomi and Marsh, David J. E. and McAllister, Liam and Moritz, Jakob, *Glimmers from the axiverse*, *JCAP* **09** (2024) 071, [2309.13145].

[9] Carroll, Sean M. and Field, George B., *The Einstein equivalence principle and the polarization of radio galaxies*, *Phys. Rev. D* **43** (1991) 3789.

[10] Harari, Diego and Sikivie, Pierre, *Effects of a Nambu-Goldstone boson on the polarization of radio galaxies and the cosmic microwave background*, *Phys. Lett. B* **289** (1992) 67–72.

[11] Carroll, Sean M., *Quintessence and the rest of the world*, *AIP Conf. Proc.* **478** (1999) 291–294.

[12] Lue, Arthur and Wang, Li-Min and Kamionkowski, Marc, *Cosmological signature of new parity violating interactions*, *Phys. Rev. Lett.* **83** (1999) 1506–1509, [astro-ph/9812088].

[13] Fedderke, Michael A. and Graham, Peter W. and Rajendran, Surjeet, *Axion Dark Matter Detection with CMB Polarization*, *Phys. Rev. D* **100** (2019) 015040, [1903.02666].

[14] Fujita, Tomohiro and Murai, Kai and Nakatsuka, Hiromasa and Tsujikawa, Shinji, *Detection of isotropic cosmic birefringence and its implications for axion-like particles including dark energy*, *Phys. Rev. D* **103** (2021) 043509, [2011.11894].

[15] CMB-S4 collaboration, Abazajian, Kevork N. and others, *CMB-S4 Science Book, First Edition*, 1610.02743.

[16] Pogosian, Levon and Shimon, Meir and Mewes, Matthew and Keating, Brian, *Future CMB constraints on cosmic birefringence and implications for fundamental physics*, *Phys. Rev. D* **100** (2019) 023507, [1904.07855].

[17] BICEP/Keck collaboration, Ade, P. A. R. and others, *BICEP/Keck XVIII: Measurement of BICEP3 polarization angles and consequences for constraining cosmic birefringence and inflation*, 2410.12089.

[18] Minami, Yuto and Komatsu, Eiichiro, *New Extraction of the Cosmic Birefringence from the Planck 2018 Polarization Data*, *Phys. Rev. Lett.* **125** (2020) 221301, [2011.11254].

[19] Diego-Palazuelos, P. and others, *Cosmic Birefringence from the Planck Data Release 4*, *Phys. Rev. Lett.* **128** (2022) 091302, [2201.07682].

[20] Eskilt, J. R., *Frequency-Dependent Constraints on Cosmic Birefringence from the LFI and HFI Planck Data Release 4*, *Astron. Astrophys.* **662** (2022) A10, [2201.13347].

[21] Eskilt, Johannes R. and Komatsu, Eiichiro, *Improved Constraints on Cosmic Birefringence from the WMAP and Planck Cosmic Microwave Background Polarization Data*, 2205.13962.

[22] Komatsu, Eiichiro, *New physics from the polarized light of the cosmic microwave background*, *Nature Rev. Phys.* **4** (2022) 452–469, [2202.13919].

[23] Kibble, T. W. B., *Topology of Cosmic Domains and Strings*, *J. Phys. A* **9** (1976) 1387–1398.

[24] Agrawal, Prateek and Hook, Anson and Huang, Junwu, *A CMB Millikan experiment with cosmic axiverse strings*, *JHEP* **07** (2020) 138, [1912.02823].

[25] Huang, M. C. and Sikivie, P., *The Structure of Axionic Domain Walls*, *Phys. Rev. D* **32** (1985) 1560.

[26] Harvey, Jeffrey A. and Naculich, Stephen G., *Cosmic Strings From Pseudoanomalous U(1)s*, *Phys. Lett. B* **217** (1989) 231–237.

[27] Jain, Mudit and Long, Andrew J. and Amin, Mustafa A., *CMB birefringence from ultralight-axion string networks*, *JCAP* **05** (2021) 055, [2103.10962].

[28] Hagimoto, Ray and Long, Andrew J., *Measures of non-Gaussianity in axion-string-induced CMB birefringence*, *JCAP* **09** (2023) 024, [2306.07351].

[29] Yin, Weichen Winston and Dai, Liang and Ferraro, Simone, *Probing cosmic strings by reconstructing polarization rotation of the cosmic microwave background*, 2111.12741.

[30] Jain, Mudit and Hagimoto, Ray and Long, Andrew J. and Amin, Mustafa A., *Searching for axion-like particles through CMB birefringence from string-wall networks*, *JCAP* **10** (2022) 090, [2208.08391].

[31] Yin, Weichen Winston and Dai, Liang and Huang, Junwu and Ji, Lingyuan and Ferraro, Simone, *A New Probe of Cosmic Birefringence Using Galaxy Polarization and Shapes*, 2402.18568.

[32] Sherwin, Blake D. and Namikawa, Toshiya, *Cosmic birefringence tomography and calibration independence with reionization signals in the CMB*, *Mon. Not. Roy. Astron. Soc.* **520** (2023) 3298–3304, [2108.09287].

[33] Namikawa, Toshiya, *Tomographic constraint on anisotropic cosmic birefringence*, 2410.05149.

[34] Takahashi, Fuminobu and Yin, Wen, *Kilobyte Cosmic Birefringence from ALP Domain Walls*, 2012.11576.

[35] Gonzalez, Diego and Kitajima, Naoya and Takahashi, Fuminobu and Yin, Wen, *Stability of domain wall network with initial inflationary fluctuations and its implications for cosmic birefringence*, *Phys. Lett. B* **843** (2023) 137990, [2211.06849].

[36] Kitajima, Naoya and Kozai, Fumiaki and Takahashi, Fuminobu and Yin, Wen, *Power spectrum of domain-wall network and its implications for isotropic and anisotropic cosmic birefringence*, 2205.05083.

[37] Ferreira, Ricardo Z. and Gasparotto, Silvia and Hiramatsu, Takashi and Obata, Ippei and Pujolas, Oriol, *Axionic defects in the CMB: birefringence and gravitational waves*, *JCAP* **05** (2024) 066, [2312.14104].

[38] Obata, Ippei, *Implications of the Cosmic Birefringence Measurement for the Axion Dark Matter Search*, 2108.02150.

[39] Schmelzle, Jorit and Lucchi, Aurelien and Kacprzak, Tomasz and Amara, Adam and Sgier, Raphael and Réfrégier, Alexandre and Hofmann, Thomas, *Cosmological model discrimination with Deep Learning*, 1707.05167.

[40] Peel, Austin and Lalande, Florian and Starck, Jean-Luc and Pettorino, Valeria and Merten, Julian and Giocoli, Carlo and Meneghetti, Massimo and Baldi, Marco, *Distinguishing standard and modified gravity cosmologies with machine learning*, *Phys. Rev. D* **100** (2019) 023508, [1810.11030].

[41] Ribli, Dezső and Pataki, Bálint Ármin and Zorrilla Matilla, José Manuel and Hsu, Daniel and Haiman, Zoltán and Csabai, István, *Weak lensing cosmology with convolutional neural networks on noisy data*, *Mon. Not. Roy. Astron. Soc.* **490** (2019) 1843–1860, [1902.03663].

[42] Fluri, Janis and Kacprzak, Tomasz and Lucchi, Aurelien and Refregier, Alexandre and Amara, Adam and Hofmann, Thomas and Schneider, Aurel,

Cosmological constraints with deep learning from KiDS-450 weak lensing maps, *Phys. Rev. D* **100** (2019) 063514, [1906.03156].

[43] Hortua, Hector J. and Volpi, Riccardo and Marinelli, Dimitri and Malagò, Luigi, *Parameter estimation for the cosmic microwave background with Bayesian neural networks*, *Phys. Rev. D* **102** (2020) 103509, [1911.08508].

[44] Guzman, Eric and Meyers, Joel, *Reconstructing cosmic polarization rotation with ResUNet-CMB*, *JCAP* **01** (2022) 030, [2109.09715].

[45] Taylor, Peter L. and Craigie, Matthew and Ting, Yuan-Sen, *Unsupervised searches for cosmological parity violation: An investigation with convolutional neural networks*, *Phys. Rev. D* **109** (2024) 083518, [2312.09287].

[46] Zhong, Kunhao and Gatti, Marco and Jain, Bhuvnesh, *Improving convolutional neural networks for cosmological fields with random permutation*, *Phys. Rev. D* **110** (2024) 043535, [2403.01368].

[47] DES collaboration, Jeffrey, N. and others, *Dark Energy Survey Year 3 results: likelihood-free, simulation-based wCDM inference with neural compression of weak-lensing map statistics*, 2403.02314.

[48] Defferrard, Michaël and Milani, Martino and Gusset, Frédérick and Perraudin, Nathanaël, *DeepSphere: a graph-based spherical CNN*, in *International Conference on Learning Representations (ICLR)*, 2020.

[49] Defferrard, Michaël and Perraudin, Nathanaël and Kacprzak, Tomasz and Sgier, Raphael, *DeepSphere: towards an equivariant graph-based spherical CNN*, in *ICLR Workshop on Representation Learning on Graphs and Manifolds*, 2019. 1904.05146.

[50] Perraudin, Nathanaël and Defferrard, Michaël and Kacprzak, Tomasz and Sgier, Raphael, *DeepSphere: Efficient spherical convolutional neural network with*

HEALPix sampling for cosmological applications, *Astronomy and Computing* **27** (Apr., 2019) 130–146, [1810.12186].

[51] Fluri, Janis and Kacprzak, Tomasz and Lucchi, Aurelien and Schneider, Aurel and Refregier, Alexandre and Hofmann, Thomas, *Full wCDM analysis of KiDS-1000 weak lensing maps using deep learning*, *Phys. Rev. D* **105** (2022) 083518, [2201.07771].

[52] Yin, Weichen Winston and Dai, Liang and Ferraro, Simone, *Testing charge quantization with axion string-induced cosmic birefringence*, *JCAP* **07** (2023) 052, [2305.02318].

[53] Jain, Mudit and Vilenkin, Alexander, *Clustering of cosmic string loops*, *JCAP* **09** (2020) 043, [2006.15358].

[54] Yamaguchi, Masahide and Kawasaki, M. and Yokoyama, Jun'ichi, *Evolution of axionic strings and spectrum of axions radiated from them*, *Phys. Rev. Lett.* **82** (1999) 4578–4581, [hep-ph/9811311].

[55] Yamaguchi, Masahide and Yokoyama, Jun'ichi, *Quantitative evolution of global strings from the Lagrangian view point*, *Phys. Rev. D* **67** (2003) 103514, [hep-ph/0210343].

[56] Hiramatsu, Takashi and Kawasaki, Masahiro and Sekiguchi, Toyokazu and Yamaguchi, Masahide and Yokoyama, Jun'ichi, *Improved estimation of radiated axions from cosmological axionic strings*, *Phys. Rev. D* **83** (2011) 123531, [1012.5502].

[57] Hiramatsu, Takashi and Kawasaki, Masahiro and Saikawa, Ken'ichi and Sekiguchi, Toyokazu, *Production of dark matter axions from collapse of string-wall systems*, *Phys. Rev. D* **85** (2012) 105020, [1202.5851].

[58] Kawasaki, Masahiro and Saikawa, Ken'ichi and Sekiguchi, Toyokazu, *Axion dark matter from topological defects*, *Phys. Rev. D* **91** (2015) 065014, [1412.0789].

[59] Lopez-Eiguren, Asier and Lizarraga, Joanes and Hindmarsh, Mark and Urrestilla, Jon, *Cosmic Microwave Background constraints for global strings and global monopoles*, *JCAP* **07** (2017) 026, [1705.04154].

[60] Gorghetto, Marco and Hardy, Edward and Villadoro, Giovanni, *Axions from Strings: the Attractive Solution*, *JHEP* **07** (2018) 151, [1806.04677].

[61] Hindmarsh, Mark and Lizarraga, Joanes and Lopez-Eiguren, Asier and Urrestilla, Jon, *Scaling Density of Axion Strings*, *Phys. Rev. Lett.* **124** (2020) 021301, [1908.03522].

[62] Gorghetto, Marco and Hardy, Edward and Villadoro, Giovanni, *More Axions from Strings*, 2007.04990.

[63] Hindmarsh, Mark and Lizarraga, Joanes and Lopez-Eiguren, Asier and Urrestilla, Jon, *Approach to scaling in axion string networks*, 2102.07723.

[64] Górski, K. M. and Hivon, E. and Banday, A. J. and Wandelt, B. D. and Hansen, F. K. and Reinecke, M. and Bartelman, M., *HEALPix - A Framework for high resolution discretization, and fast analysis of data distributed on the sphere*, *Astrophys. J.* **622** (2005) 759–771, [astro-ph/0409513].

[65] Shuman, David and Narang, Sunil and Frossard, Pascal and Ortega, Antonio and Vandergheynst, Pierre, *The Emerging Field of Signal Processing on Graphs: Extending High-Dimensional Data Analysis to Networks and Other Irregular Domains*, *IEEE Signal Processing Magazine* **30** (May, 2013) 83–98, [1211.0053].

[66] Lin, Min and Chen, Qiang and Yan, Shuicheng, *Network In Network*, *arXiv e-prints* (Dec., 2013) arXiv:1312.4400, [1312.4400].

[67] James, Gareth and Witten, Daniela and Hastie, Trevor and Tibshirani, Robert and others, *An introduction to statistical learning*, vol. 112. Springer, 2013.

[68] Kingma, Diederik P, *Adam: A method for stochastic optimization*, *arXiv preprint arXiv:1412.6980* (2014) .

[69] Y. Fan and S. A. Sisson, *Abc samplers*, 2018.

[70] Kamionkowski, Marc, *How to De-Rotate the Cosmic Microwave Background Polarization*, *Phys. Rev. Lett.* **102** (2009) 111302, [0810.1286].

[71] Gluscevic, Vera and Kamionkowski, Marc and Cooray, Asantha, *De-Rotation of the Cosmic Microwave Background Polarization: Full-Sky Formalism*, *Phys. Rev. D* **80** (2009) 023510, [0905.1687].

[72] Yadav, Amit P. S. and Biswas, Rahul and Su, Meng and Zaldarriaga, Matias, *Constraining a spatially dependent rotation of the Cosmic Microwave Background Polarization*, *Phys. Rev. D* **79** (2009) 123009, [0902.4466].

[73] POLARBEAR collaboration, Ade, Peter A. R. and others, *POLARBEAR Constraints on Cosmic Birefringence and Primordial Magnetic Fields*, *Phys. Rev. D* **92** (2015) 123509, [1509.02461].

[74] Namikawa, Toshiya and others, *Atacama Cosmology Telescope: Constraints on cosmic birefringence*, *Phys. Rev. D* **101** (2020) 083527, [2001.10465].

[75] SPT collaboration, Bianchini, F. and others, *Searching for Anisotropic Cosmic Birefringence with Polarization Data from SPTpol*, *Phys. Rev. D* **102** (2020) 083504, [2006.08061].

[76] Fluri, Janis and Lucchi, Aurelien and Kacprzak, Tomasz and Refregier, Alexandre and Hofmann, Thomas, *Cosmological parameter estimation and inference using deep summaries*, *Phys. Rev. D* **104** (2021) 123526, [2107.09002].

Chapter 5

Neutron star cooling with lepton-flavor-violating axions

Abstract

The cores of dense stars are a powerful laboratory for studying feebly coupled particles such as axions. Some of the strongest constraints on axionlike particles and their couplings to ordinary matter derive from considerations of stellar axion emission. In this work we study the radiation of axionlike particles from degenerate neutron star matter via a lepton-flavor-violating coupling that leads to muon-electron conversion when an axion is emitted. We calculate the axion emission rate per unit volume (emissivity) and by comparing with the rate of neutrino emission, we infer upper limits on the lepton-flavor-violating coupling that are at the level of $|g_{ae\mu}| \lesssim 10^{-6}$. For the hotter environment of a supernova, such as SN 1987A, the axion emission rate is enhanced and the limit is stronger, at the level of $|g_{ae\mu}| \lesssim 10^{-11}$, competitive with laboratory limits. Interestingly, our derivation of the axion emissivity reveals that axion emission via the lepton-flavor-violating coupling is suppressed relative to the familiar lepton-flavor-preserving channels by the square of the plasma temperature to muon mass ratio, which is responsible for the relatively weaker limits.

Notes about this project: This chapter is from a paper I wrote with Andrew J. Long and Hong-Yi Zhang that was published in the journal The Physical Review D in 2023 (1).

The goal of this project, similar to the previous projects was to study axion phenomenology and like chapter 2 it place constraints on axion parameters. Unlike in previous chapters, here we consider axion particles with masses on the order $m_a \sim$

1 MeV produced in neutron stars. The basic idea is that if axions have a lepton-flavour violating coupling to electrons and muons, then the cooling rates of neutron stars would have an additional channel by which they can lose energy which would modify the rate at which they cool. My primary contribution to this work was the numerical evaluation of the axion emissivity given by eq. (5.5). The Monte Carlo techniques I used are discussed in detail in app. 5.C.

5.1 Introduction

Axions are pseudo-Goldstone bosons associated with a spontaneously broken global symmetry that is anomalous to the standard model (SM) gauge couplings (2). Initially proposed as a natural solution to explain the absence of the neutron electric dipole moment (3–5), a QCD axion is characterized by its decay constant f_a (6–9) and its mass is determined by $m_a \approx 5.7\,\mu\text{eV}(10^{12}\,\text{GeV}/f_a)$ (10; 11). Apart from the QCD axion, axionlike particles have also been extensively studied in string theory (12–14) and dark matter physics (15–19). For recent reviews, refer to (20–23).

Due to their weak interactions with SM particles, detecting axions in terrestrial experiments is challenging. Therefore, it is motivated to search for evidence of axions in astrophysical systems where their feeble couplings are partially compensated by high temperatures and densities (24). For instance, probing axion emission from the white dwarf luminosity function (25–28) places a stringent limit on the axion-electron coupling at the level of $g_{aee} \lesssim 10^{-13}$. Additionally, the axion's interaction with nucleons is probed by neutron star (NS) cooling (29–31) and supernova neutrino emission (32–39), which imply tight upper limits at the level of $g_{aNN} \lesssim 10^{-10}$.

As an extension of the SM, there is no strong reason for the ultraviolet theory of axions to respect lepton flavor conservation since it is an accidental symmetry of the SM broken by tiny neutrino masses. The axions whose ultraviolet theory is responsible for the breaking of the flavor symmetry are known as flavons or familons (40–44), which can also explain the strong CP problem if they have a coupling to

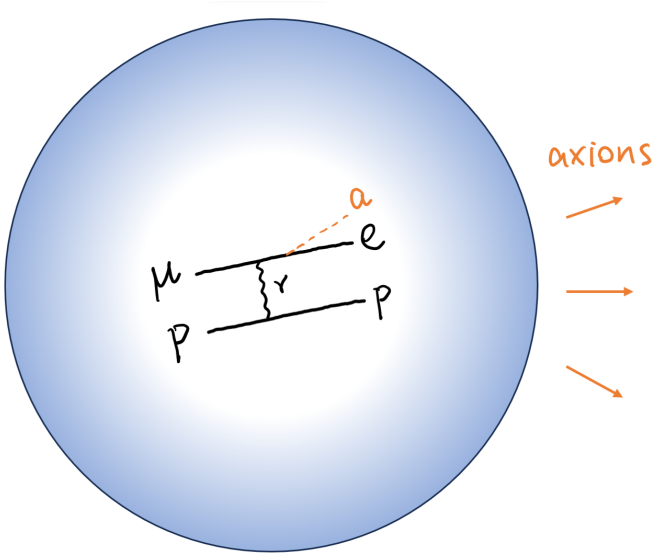


Figure 5.1.1 : If axions are produced in neutron star cores, they will carry energy out of the star and make the neutron star cool down more efficiently than expected.

gluons (45; 46). Even if the underlying theory preserves lepton flavor, lepton-flavor-violating (LFV) effects can arise from radiative corrections (47–50). It has been shown that LFV interactions can account for the production of dark matter through thermal freeze-in (51). Tests of lepton flavor conservation thus provide important information about new physics.

Laboratory tests of lepton-flavor violation serve as an indirect probe of the axion's LFV interactions. Notably, charged lepton flavor violation would lead to rare lepton decays (52). If the axion were heavier than the muon, an effective field theory approach could be used to study decays such as $\mu \to e\gamma$, $\mu \to 3e$ and $\mu - e$ conversion, being the best process to detect LFV in the $e\mu$ sector.* For lighter axions, $\mu \to ea$ could be the dominating channel and the current limit on $\mathrm{Br}(\mu \to ea)$ is of order 10^{-6}

*In the SM, LFV decays are suppressed by the neutrino mass-squared difference and $\mathrm{Br}(\mu \to e\gamma) \sim \mathrm{Br}(\mu \to 3e) \sim 10^{-54}$ (52–54), far below the current experimental limits $\mathrm{Br}(\mu \to e\gamma) < 4.2 \times 10^{-13}$ (55) and $\mathrm{Br}(\mu \to 3e) < 1.0 \times 10^{-12}$ (56).

(57) or 10^{-5} (58) depending on the axion mass and chirality of the interaction. The limit will be improved in the future experiments MEG II (59; 60) and Mu3e (61) by up to two orders of magnitude (62).

In this work, we aim to establish an astrophysical limit on the axion's LFV interactions based on NS cooling arguments, as a complement to current lab limits. The basic idea is illustrated in figure 5.1.1; if axions are produced in NS cores, they must not carry energy out of the star more efficiently than standard neutrino-mediated cooling channels (24). In a NS core, unlike nondegenerate stars or even white dwarf stars, the particle densities are so high that the electron Fermi energy exceeds the muon mass, and an appreciable population of muons is present (63). As such, NSs provide a unique opportunity to probe the axion's LFV coupling with muons and electrons.

5.2 Axions with LFV couplings

We consider a LFV coupling among the electron, muon, and axion, which is expressed as

$$\mathscr{L}_{\rm LFV} = \frac{g_{ae\mu}}{m_e + m_\mu} \overline{\Psi}_e \gamma^\rho \gamma_5 \Psi_\mu \, \partial_\rho a \;\; + \;\; \text{h.c.} \; , \tag{5.1}$$

where $\Psi_e(x)$ is the electron field, $\Psi_\mu(x)$ is the muon field, $a(x)$ is the axion field, $m_e \approx 0.511\,\text{MeV}$ is the electron mass, $m_\mu \approx 106\,\text{MeV}$ is the muon mass, and $g_{ae\mu}$ is the axion's LFV coupling. The coupling may also be written in terms of the axion decay constant f_a as $g_{ae\mu} = C_{ae\mu}(m_e + m_\mu)/(2f_a)$. This interaction can naturally arise, e.g., in the models of the LFV QCD axion (8; 9), the LFV axiflavon (45; 46; 64), the leptonic familon (65–67) and the majoron (68; 69) (also see (62) for a summary of constraints). Past studies of charged lepton flavor violation, from both terrestrial experiments and cosmological / astrophysical observations, furnish constraints on the axion LFV coupling $g_{ae\mu}$, which we summarize here.

The LFV interaction opens an exotic decay channel for the muon $\mu \to ea$, as long

as the axion mass is not too large $m_a < m_\mu - m_e$. The branching ratio is predicted to be (70)

$$\mathrm{Br}(\mu \to ea) \approx \frac{\Gamma(\mu \to ea)}{\Gamma(\mu \to e\nu\overline{\nu})} = 7.0 \times 10^{15} g_{ae\mu}^2 \ . \tag{5.2}$$

Initial searches for the two-body muon decay were performed by Derenzo using a magnetic spectrometer, resulting in an upper limit on the branching ratio of 2×10^{-4} for the mass range 98.1–103.5 MeV (71). Jodidio et al. constrained the branching ratio for a massless familon to be $< 2.6 \times 10^{-6}$, which was later extended to massive particles up to ~ 10 MeV (62). Bryman and Clifford analyzed data of muon and tauon decays obtained from NaI(Tl) and magnetic spectrometers, concluding an upper limit of 3×10^{-4} for masses less than 104 MeV (72). Bilger et al. studied muon decay in the mass range 103–105 MeV using a high purity germanium detector and established a limit of 5.7×10^{-4} (73), while the PIENU Collaboration improved the limit in the mass range 87.0–95.1 MeV (74). The TWIST experiment performed a broader search for masses up to ~ 80 MeV by accommodating nonzero anisotropies, resulting in an upper limit of 2.1×10^{-5} for massless axions (58). These constraints on $\mathrm{Br}(\mu \to ea)$ translate into upper limits on the LFV coupling $g_{ae\mu}$, and we summarize the current status in table 5.2.1.

Apart from terrestrial experiments, cosmological and astrophysical observations also constrain the axion's LFV interaction. If this interaction were too strong, relativistic axions would be produced thermally in the early universe; however, the presence of a dark radiation in the universe is incompatible with observations of the cosmic microwave background anisotropies. Constraints on dark radiation are typically expressed in terms of a parameter N_{eff} called the effective number of neutrino species. A recent study of flavor-violating axions in the early universe finds that current observational limits on N_{eff} require the LFV coupling to obey $|2f_a/C_{ae\mu}| > 2.5 \times 10^8$GeV (75). Astrophysical probes of the axion's LFV interaction have not been extensively explored. Calibbi et al. considered the bound on $\mathrm{Br}(\mu \to ea)$ from SN 1987A associated with the cooling of the proto-NS (62). Assuming that the dominant energy loss

$\lvert g_{ae\mu}\rvert$	$\frac{2f_a}{C_{ae\mu}}$ [GeV]	Br($\mu \to ea$)	m_a [MeV]	Experiment	Reference
$< 3.0 \times 10^{-6}$	$> 3.5 \times 10^{4}$	< 1.0	$\lesssim 1$	NS cooling	This work
$\lesssim 8 \times 10^{-10}$	$\gtrsim 1 \times 10^{8}$	$\lesssim 4 \times 10^{-3}$	$\lesssim 50$	SN 1987A, $\mu \to ea$	(62)
$< 4.2 \times 10^{-10}$	$> 2.5 \times 10^{8}$	$< 1.3 \times 10^{-3}$	$\lesssim 10^{-7}$	Cosmology, ΔN_{eff}	(75)
$< 2.9 \times 10^{-10}$	$> 3.7 \times 10^{8}$	$< 5.7 \times 10^{-4}$	$103 - 105$	Rare muon decay	(73)
$\lesssim 2 \times 10^{-10}$	$\gtrsim 5 \times 10^{8}$	$\lesssim 3 \times 10^{-4}$	< 104	Rare muon decay	(72)
$< 2 \times 10^{-10}$	$> 6 \times 10^{8}$	$< 2 \times 10^{-4}$	$98.1 - 103.5$	Rare muon decay	(71)
$< 1 \times 10^{-10}$	$> 9 \times 10^{8}$	$< 1 \times 10^{-4}$	$47.8 - 95.1$	Rare muon decay (PIENU)†	(74)
$< 5.5 \times 10^{-11}$	$> 1.9 \times 10^{9}$	$< 2.1 \times 10^{-5}$	< 13	Rare muon decay (TWIST)	(58)
$\lesssim 4 \times 10^{-11}$	$\gtrsim 3 \times 10^{9}$	$\lesssim 9 \times 10^{-6}$	$\lesssim 50$	SN 1987A, $lf \to l'fa$	This work
$< 1.9 \times 10^{-11}$	$> 5.5 \times 10^{9}$	$< 2.6 \times 10^{-6}$	$\lesssim 10$	Rare muon decay	(57; 62)

Table 5.2.1 : Summary of constraints on the axion's LFV coupling in the e-μ sector, where stronger constraints are presented at the bottom. See the main text for more detailed descriptions. For the NS cooling limit, we calculate the axion emissivity via $l + f \to l' + f + a$ and compare with the neutrino emissivity via Murca channels. For the SN 1987A limit, we compare with the upper bound on energy loss rate.

channel is free muon decay $\mu \to ea$, they derive an upper limit on the branching ratio at the level of 4×10^{-3}. We find that a stronger constraint is obtained from the 2-to-3 scattering channels, such as $\mu p \to epa$, and we discuss this result further below.

To provide a comprehensive overview, we also introduce the constraints on LFV couplings involving τ leptons. Currently, laboratory limits on the branching ratios of rare tauon decays are $\mathrm{Br}(\tau \to ea) < 2.7 \times 10^{-3}$ and $\mathrm{Br}(\tau \to \mu a) < 4.5 \times 10^{-3}$ (62; 76). Constraints from N_{eff} are more stringent, $\mathrm{Br}(\tau \to ea) \lesssim 3 \times 10^{-4}$ and $\mathrm{Br}(\tau \to \mu a) \lesssim 5 \times 10^{-4}$ (75). Each of these limits is expected to improve significantly, by up to three orders of magnitude, in the future Belle II (62; 77) and CMB-S4 experiment (75; 78; 79). However, it remains challenging to impose constraints on τ leptons from astrophysical systems due to their considerable mass of 1.8 GeV, which far exceeds stellar core temperatures.

5.3 Axion emission via LFV couplings

The emission of axions from NS matter via the LFV interaction can proceed through various channels. One might expect the dominant channel to be the decay of free muons $\mu \to ea$; however, since the electrons in NS matter are degenerate, this channel is Pauli blocked, and its rate is suppressed in comparison with scattering channels. Since NS matter consists of degenerate electrons, muons, protons, and neutrons, various scattering channels are available. We denote these collectively as[‡]

$$l + f \to l' + f + a \ , \tag{5.3}$$

where a lepton $l = e, \mu$ is converted to another $l' = \mu, e$ with the spectator particle $f = p, e, \mu$. We consider channels in which the NS's muon is present in the initial state, and channels in which muons are created thanks to the large electron Fermi momentum. The scattering is mediated by the electromagnetic interaction (photon

[‡]We neglect the Compton process for axions, since the number density of photons is low compared to other particles.

exchange), and channels involving neutrons are neglected. Assuming that all particles are degenerate, scattering predominantly happens for particles at the Fermi surface. These processes are kinematically allowed if $|p_{F,l}-p_{F,f}| < p_{F,l'}+p_{F,f}$ and $|p_{F,l'}-p_{F,f}| < p_{F,l}+p_{F,f}$, implying the existence of a threshold momentum of the spectator particle

$$p_{F,f} > (p_{F,e} - p_{F,\mu})/2 \ . \tag{5.4}$$

Here we have introduced the Fermi momentum $p_{F,i}$ of the particle species i.

The quantities of interest are the axion emissivities $\varepsilon_a^{(lf)}$, which corresponds to the energy released in axions per unit volume per unit time through the channel $lf \to l'fa$. We assign $(E_1, \boldsymbol{p}_1)$ and $(E_1', \boldsymbol{p}_1')$ for the initial and final four-momenta of the converting leptons l and l', $(E_2, \boldsymbol{p}_2)$ and $(E_2', \boldsymbol{p}_2')$ for the spectator f, and $(E_3', \boldsymbol{p}_3')$ for the axion. Then the axion emissivity is calculated as

$$\begin{aligned} \varepsilon_a^{(lf)} = \frac{(2\pi)^4}{S} \int \prod_{i=1}^{2} \widetilde{dp_i} \ \prod_{j=1}^{3} \widetilde{dp_j'} \ \sum_{\text{spin}} \left|\mathcal{M}^{(lf)}\right|^2 \\ \times \delta^{(4)}(p_1 + p_2 - p_1' - p_2' - p_3') \\ \times E_3' \, f_1 \, f_2 \, (1 - f_1') \, (1 - f_2') \ , \end{aligned} \tag{5.5}$$

where S is the symmetry factor accounting for identical initial and final state particles, $\mathcal{M}^{(lf)}$ is the Lorentz invariant matrix element, f_i and f_i' are the Fermi-Dirac distribution functions, the factor $(1 - f_i')$ takes into account the Pauli blocking due to particle degeneracy, and $\widetilde{dp} \equiv d^3p/[(2\pi)^3 2E]$ is the Lorentz-invariant differential phase space element. We do not include a factor of $(1 + f_3')$, since $f_3' \ll 1$ and there is no Bose enhancement for axion production since NSs are essentially transparent to axions for the currently allowed parameter space.

Calculating the emissivity (5.5) requires evaluating the 15 momentum integrals along with the 4 constraints from energy and momentum conservation. We evaluate all but 2 of these integrals analytically using the Fermi surface approximation, and we calculate the last 2 integrals using numerical techniques. The Fermi surface approximation assumes that the integrals are dominated by momenta near the Fermi surface

$|\boldsymbol{p}| \approx p_F$; smaller and larger momenta do not contribute because of Pauli blocking or Boltzmann suppression. We find the axion emissivity of the $lf \to l'fa$ channel to be

$$\varepsilon_a^{(lf)} = \frac{328\pi^2\alpha^2 g_{ae\mu}^2}{945 m_\mu^4} \frac{\beta_{F,l} E_{F,e}^3}{\beta_{F,f}^2 p_{F,f}^2} F^{(lf)} T^8 \ , \tag{5.6}$$

where $\alpha \approx 1/137$ is the electromagnetic fine-structure constant, $E_{F,i}$ is the Fermi energy, $\beta_{F,i} \equiv p_{F,i}/E_{F,i}$ is the Fermi velocity, T is the plasma temperature, and $F^{(lf)}$ is a factor depending on both the specific process and the Fermi velocity of the scattering particles. To derive (5.6), we have assumed that the axion mass is small compared to the NS temperature $m_a \ll T$, muons and electrons are in the beta equilibrium (i.e., $E_{F,e} \approx E_{F,\mu}$), electrons are ultra relativistic but muons are not (i.e., $p_{F,\mu} \lesssim m_\mu$), and $T \ll m_\mu^2/E_{F,e}$. Our derivation of (5.6) can be found in appendix 5.A. In addition, we evaluate the emissivity fully numerically using Monte Carlo integration methods to estimate the integrals in (5.5) without employing the Fermi surface approximation. In the regime of interest, the two methods agree very well. The impact of an axion mass $m_a \gtrsim T$ is discussed in appendix 5.C.

The temperature dependence of the axion emissivity (5.6) is especially interesting and important for understanding the limits from NS cooling. For comparison, note that axion bremstrahlung via lepton-flavor-preserving (LFP) interactions (such as $ep \to epa$ or $\mu p \to \mu pa$) goes as $\varepsilon_a \propto T^6$. In other words, the LFV interaction leads to an emissivity that's suppressed by an additional factor of $T^2 E_{F,e}^2/(m_\mu^2 - m_e^2)^2 \sim T^2/m_\mu^2$, which is of order $(100\,\mathrm{keV}/100\,\mathrm{MeV})^2 \sim 10^{-6}$ for $T \sim 10^9\,\mathrm{K}$. A detailed discussion appears in appendix 5.A, but the essential idea can be understood as follows. The phase-space integrals over momenta can be converted to energy integrals, and each integral for degenerate leptons and protons is restricted to the Fermi surface of thickness $\sim T$, giving a factor of T^4. The phase-space integral of axions (i.e., d^3p_3'/E_3') gives a factor of T^2. The axions are emitted thermally and have an energy $\sim T$. The energy conservation delta function gives T^{-1}. The squared matrix element has a temperature dependence T^2. Putting all these together, we see that the emissivity is proportional to T^8. In comparison, the squared matrix element

for the LFP interactions has no temperature dependence since one power of T from the coupling vertex is canceled by T^{-1} from the lepton propagator.

We numerically evaluate the axion emissivities (5.6) and present these results in figure 5.3.1 for the six channels $lf \to l'fa$, where the effective mass of protons is taken to be $0.8\, m_p$ (see (80) and references therein).[§] Using the strong degeneracy of particles and the beta equilibrium condition $E_{F,e} \approx E_{F,\mu}$, one can show that the emissivities are equal for the channels $ef \to \mu fa$ and $\mu f \to efa$. Thus the plot only shows three curves corresponding to in-states consisting of a muon and a spectator particle $f = p, e, \mu$. The channels with a spectator proton ($f = p$) have the largest emissivity across the range of muon Fermi momenta shown here; this is a consequence of the enhanced matrix element and the larger available phase space for these scatterings. For the channels with a spectator muon ($f = \mu$), the emissivity drops to zero below $\beta_{F,\mu} \approx 0.34$; this corresponds to a violation of the kinematic threshold in (5.4). For all channels, the emissivity decreases with decreasing muon Fermi velocity due to the reduced kinematically allowed phase space. On the other hand, for larger muon Fermi velocity, the channels with spectator electrons and muons coincide, since both particles can be regarded as massless. For the top axis in figure 5.1.1, we show the corresponding mass density of a NS assuming the $npe\mu$ model; see appendix 5.B for more details.

The total axion emissivity is obtained by summing over the six channels. For this estimate we set $\beta_{F,\mu} = 0.84$. We find the axion emissivity via LFV interactions to be

$$\varepsilon_a^{\rm LFV} \simeq 4.8 \times 10^{32} g_{ae\mu}^2 T_9^8 \ \text{erg cm}^{-3}\ \text{s}^{-1}\ , \tag{5.7}$$

where $T_9 \equiv T/(10^9\,\text{K})$ and $10^9\,\text{K} \approx 86.2\,\text{keV}$.

[§]Using electric charge neutrality and the beta equilibrium condition $E_{F,e} \approx E_{F,\mu}$, the emissivity is fully determined given the effective proton mass and $\beta_{F,\mu}$.

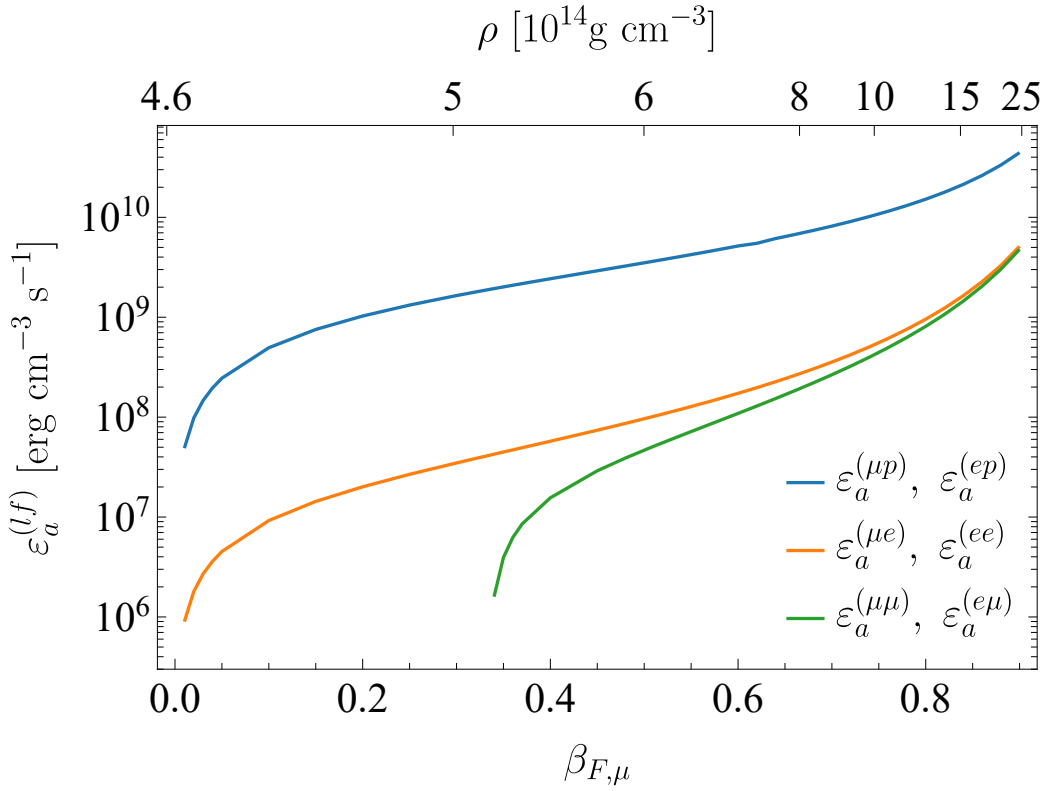


Figure 5.3.1 : Axion emissivities $\varepsilon_a^{(lf)}$ for the LFV process $l + f \to l' + f + a$, given by equation (5.6), as a function of the muon Fermi velocity $\beta_{F,\mu}$. The top axis, in a nonlinear scale, represents the corresponding mass density of a NS assuming the $npe\mu$ matter. Here we take $g_{ae\mu} = 10^{-11}$ and $T = 10^9$ K, and more generally $\varepsilon_a^{(lf)} \propto g_{ae\mu}^2 T^8$.

5.4 Implications for NS cooling

In low-mass NSs, slow cooling could occur via neutrino emission by the modified Urca (Murca) processes $nn \to npe\overline{\nu}$, $npe \to nn\nu$ or slightly less efficient processes such as the nucleon bremsstrahlung (81; 82). At the density $\rho = 6\rho_0$, where $\rho_0 = 2.5 \times 10^{14}\,\mathrm{g\ cm^{-3}}$ is the nuclear saturation density (83), and with the effective nucleon mass taken to be $0.8\,m_N$ (80), the emissivity of the Murca process is given by $\varepsilon_\nu = 4.4 \times 10^{21} T_9^8$ erg cm^{-3} s^{-1} (84). Comparing this rate with (5.7), one finds that the axion emission from LFV couplings dominates the neutrino emission unless

$$|g_{ae\mu}| \lesssim 3.0 \times 10^{-6}\ , \tag{5.8}$$

which is consistent with existing constraints. In heavier NSs, the LFV emission of axions tends to have a less significant impact. This is because fast neutrino emission could occur via the direct Urca processes (85). In the presence of superfluidity, the formation of Cooper pairs can dominate over the Murca process (86; 87), further diminishing the role of LFV axion emission. Medium effects for neutrino emission processes are discussed in (82; 88; 89).

Axions are predominantly produced in NSs through the nucleon bremsstrahlung process $nn \to nna$. At the same core conditions, its emissivity is given by $\varepsilon_a^{(nn)} \simeq 2.8 \times 10^{38} g_{ann}^2 T_9^6$ erg cm^{-3} s^{-1} (90; 91). The nucleon bremsstrahlung process dominate the LFV processes if

$$|g_{ae\mu}| \lesssim 7.6 \times 10^2 |g_{ann}| T_9^{-1}\ . \tag{5.9}$$

The current best constraint on the axion-neutron coupling is $|g_{ann}| \lesssim 2.8 \times 10^{-10}$ (30). Therefore, it is unlikely for the LFV couplings to play a significant role in NSs with an age $\gtrsim 1\,\mathrm{yr}$, where the temperature has cooled to $10^9\,\mathrm{K}$ (92).

These limits on the axion's LFV coupling are relatively weak, and this is a consequence of the $\varepsilon_a^{\mathrm{LFV}} \propto T^8$ scaling, which is suppressed compared to LFP channels by a factor of $(T/m_\mu)^2$, which is tiny in old NSs. However, in the proto-NS that

forms just after a supernova, this ratio can be order one, which suggests that stronger limits can be obtained by considering the effect of axion emission on supernova rather than NSs. Since our analysis has focused on NS environments, adapting our results to the more complex proto-NS system requires some extrapolation. We estimate the axion emissivity from a supernova by extrapolating (5.7) to high temperatures. By imposing the bound on the energy loss of SN 1987A, $\varepsilon_a/\rho \lesssim 10^{19}\,\mathrm{erg\ g^{-1}\ s^{-1}}$ (24), one finds that at a typical core condition $\rho \sim 8 \times 10^{14}\,\mathrm{g\ cm^{-3}}$,

$$|g_{ae\mu}| \lesssim 4 \times 10^{-11} \left(\frac{50\,\mathrm{MeV}}{T}\right)^4 , \tag{5.10}$$

which is to be evaluated at $T \sim (30 - 60)$ MeV. This constraint is more stringent than that obtained from considering $\mu \to ea$ in a supernova and is comparable to the current best terrestrial limit.

One should note that at typical core conditions of a proto-NS, nucleons and muons are at the borderline between degeneracy and nondegeneracy where electromagnetic field screening effects become significant. In appendix 5.C, we discuss the effect of electromagnetic field screening due to the presence of a degenerate plasma with charged constituents on the axion emissivity. We then account for this effect in our numerical code by introducing an effective mass for photon propagators of order the Thomas-Fermi wavenumber k_{TF}.[¶] Using Monte Carlo integration we evaluate the axion emissivity up to temperatures of 100 MeV and find that extrapolating the degenerate rate tends to overestimate the emissivity by a factor of ~ 10, leading to a weaker supernova constraint by a factor ~ 3.

[¶]While this methodology is not apt for strongly coupled plasmas like NSs and white dwarfs, it does furnish reasonably accurate estimates of the screening effect in axion bremsstrahlung processes within white dwarfs (24).

5.5 Discussion

In this article, we study the astrophysical signatures of an axionlike particle's LFV coupling with muons and electrons. We focus on axion emission from NS cores, where the electron Fermi energy is large enough to maintain a high abundance of muons. Our limits on the LFV coupling $g_{ae\mu}$ derive from comparing the axion emission rate with the energy loss rate due to neutrino emission, since excessively strong axion emission would conflict with the observations of old NSs and SN 1987A. The summary of current constraints is shown in figure 5.5.1.

Further research is needed to assess the impact of axion's LFV interactions on the entire cooling history of the star, including a careful treatment of equations of state and nuclear interactions. Stronger nuclear interactions would result in higher number densities of protons and muons at the same mass density, thereby enhancing the rate of the LFV interactions. Such an analysis is particularly motivated for axion emission from proto-NSs formed after type-II supernovae, where the transition from nondegenerate to degenerate matter and the creation of the muon population could impact axion emissivities. Our work highlights the importance of assessing both the free muon decay channel $\mu \to ea$ as well as scattering channels $lf \to l'fa$ in such studies.

Acknowledgments

We would like to thank Mustafa Amin for helpful discussions. H.Y.Z. is partly supported through DOE Grant No. DESC0021619. R.H. and A.J.L. are partly supported by the National Science Foundation under Award No. PHY-2114024.

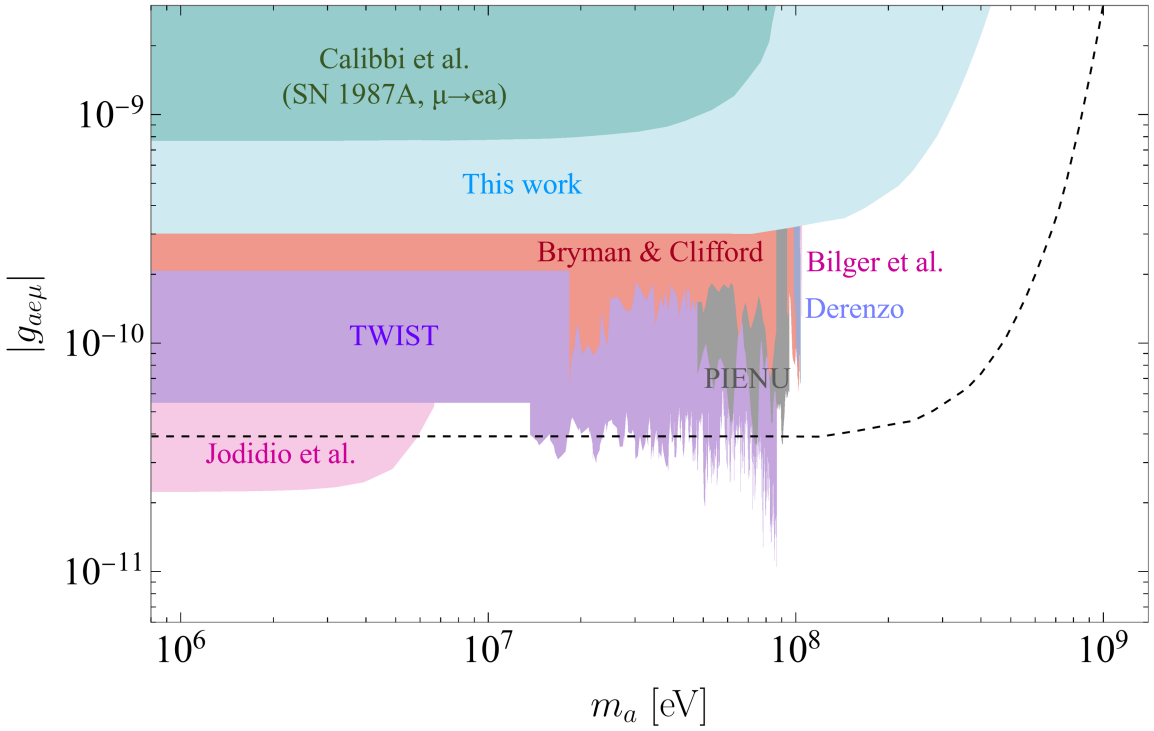


Figure 5.5.1 : Summary of constraints on the axion's LFV coupling in the e-μ sector. The constraints labeled with "Calibbi et al." and "This work" are astrophysical and the others are lab limits obtained by measuring rare muon decay rates. The weaker constraint we derive from NS cooling and the cosmological constraint inferred from the ΔN_{eff} observation, shown in table 5.2.1, do not appear on this part of parameter space. For the region labeled with "This work", we assume a supernova core temperature $T = 30\,\text{MeV}$ and a higher temperature $T = 50\,\text{MeV}$ would expand the exclusion region into that enclosed by the black dashed line.

Appendix

5.A Calculation of axion emissivity

In this section, we implement the Fermi surface approximation and evaluate the axion emissivity from the process $l + f \to l' + f + a$, where a lepton $l = e, \mu$ is converted to another $l' = \mu, e$ with the spectator particle f being one of p, e, μ. This approximation was also used in the calculation of neutrino emissivities (84; 93) and axion emissivities for the bremsstrahlung process by nucleons (90; 91; 94; 95). The metric signature is $(-,+,+,+)$.

The axion emissivity is calculated as

$$\begin{aligned}\varepsilon_a^{(lf)} = \frac{1}{S}\int \frac{d^3p_1}{(2\pi)^3}\frac{1}{2E_1}\frac{d^3p_2}{(2\pi)^3}\frac{1}{2E_2}\frac{d^3p_1'}{(2\pi)^3}\frac{1}{2E_1'}\frac{d^3p_2'}{(2\pi)^3}\frac{1}{2E_2'}\frac{d^3p_3'}{(2\pi)^3}\frac{1}{2E_3'}\sum_{\text{spin}}\left|\mathcal{M}^{(lf)}\right|^2 \\ \times (2\pi)\,\delta(E_1 + E_2 - E_1' - E_2' - E_3')\;(2\pi)^3\,\delta^{(3)}(\boldsymbol{p}_1 + \boldsymbol{p}_2 - \boldsymbol{p}_1' - \boldsymbol{p}_2' - \boldsymbol{p}_3') \\ \times E_3'\, f_1\, f_2\,(1 - f_1')\,(1 - f_2')\;,\end{aligned} \tag{5.11}$$

where $\mathcal{M}^{(lf)}$ is the Lorentz-invariant matrix element for the scattering $l+f \to l'+f+a$. The symmetry factor S is needed to avoid double counting of identical particles if l or $l' = f$. The energies E_i are determined by the on-shell conditions: $E_i = \sqrt{|\boldsymbol{p}_i|^2 + m_i^2}$ for $i = 1, 2, 1', 2', 3'$. The thermal factors $f_1\, f_2\,(1 - f_1')\,(1 - f_2')$ restrict the fermion particle energies (E_1, E_2, E_1', and E_2') to be near their respective Fermi energies $E_{F,i}$ within a narrow range of order temperature $T \ll E_{F,i}$. This observation motivates the Fermi surface approximation, by which the emissivity is factorized into angular integrals with momenta restricted to the Fermi surface and energy integrals. To implement the Fermi surface approximation we introduce Dirac delta functions that fix the magnitude of the fermion 3-momenta to equal their respective Fermi momenta,

and we promote the fermion energies to integration variables via the prescription:

$$d^3p \to d^3p \int \frac{E}{p_F}\delta(p - p_F)dE \ . \tag{5.12}$$

This approximation allows the emissivity to be written as

$$\varepsilon_a^{(lf)} = \frac{1}{2^5(2\pi)^{11}p_{F,1}p_{F,2}p_{F,1'}p_{F,2'}S}JA \ , \tag{5.13}$$

which splits the calculation into two parts: an angular integral A and an energy integral J, defined by

$$\begin{aligned} A \equiv & \int d^3p_1 d^3p_2 d^3p_1' d^3p_2' d^2\Omega_3' \delta(p_1 - p_{F,1})\delta(p_2 - p_{F,2})\delta(p_1' - p_{F,1'})\delta(p_2' - p_{F,2'}) \\ & \times \delta^3(\boldsymbol{p}_1 + \boldsymbol{p}_2 - \boldsymbol{p}_1' - \boldsymbol{p}_2')\frac{\sum_{\text{spin}} \left|\mathcal{M}^{(lf)}\right|^2_{\text{Fermi}}}{{E_3'}^n} \ , \end{aligned} \tag{5.14}$$

$$J \equiv \int dE_1 dE_2 dE_1' dE_2' dE_3' \delta(E_1 + E_2 - E_1' - E_2' - E_3') f_1 f_2 (1 - f_1')(1 - f_2'){E_3'}^{n+2} \ . \tag{5.15}$$

The matrix element $\left|\mathcal{M}^{(lf)}\right|_{\text{Fermi}}$ is evaluated with fermion 3-momenta and energies fixed to the respective Fermi momenta and Fermi energies. The exponent n is chosen such that ${E_3'}^{-n}\sum_{\text{spin}}\left|\mathcal{M}^{(lf)}\right|^2_{\text{Fermi}}$ is independent of E_3'. We have neglected the axion momentum in the momentum conservation delta function since $p_3' \sim T \ll p_{F,\mu}$. The mass dimension of J and A is $6 + n$ and $3 - n$, and that of $\left|\mathcal{M}^{(lf)}\right|^2$ is -2. For the LFV channels considered in this work, we note that $p_{F,2} = p_{F,2'}$, $n = 2$, and $S = 1$ for f being a proton and $S = 2$ otherwise.

5.A.1 Energy integral

The energy integral can be written as

$$\begin{aligned} J &\approx \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2 \int_{-\infty}^{\infty} dx_1' \int_{-\infty}^{\infty} dx_2' \int_0^{\infty} dz \frac{T^{6+n} z^{2+n} \delta\left(x_1 + x_2 + x_1' + x_2' - z\right)}{(e^{x_1} + 1)(e^{x_2} + 1)(e^{x_1'} + 1)(e^{x_2'} + 1)} \\ &= \frac{T^{6+n}}{6} \int_0^{\infty} dz \frac{z^{3+n}(z^2 + 4\pi^2)}{e^z - 1} \ , \end{aligned} \tag{5.16}$$

where $x_i \equiv (E_i - E_{F,i})/T$, $x_i' \equiv (E_{F,i}' - E_i')/T$, and $z \equiv E_3'/T$. The approximation symbols arise from extending the limits of integration to infinity. The second equality is derived using the technique in (96). For $n = 2$, we obtain

$$J = \frac{164\pi^8}{945} T^8 \ . \tag{5.17}$$

5.A.2 Angular integral

For the angular integral, we first integrate d^3p_2' with the momentum delta function and dp_1, dp_2, dp_1' with the Fermi surface delta function. It is convenient to align all angles with respect to $\boldsymbol{p}_1$, so $\int d^2\Omega_1$ simply gives 4π. The angular integral A becomes

$$\begin{aligned} A = {} & 4\pi p_{F,1}^2 p_{F,2}^2 p_{F,1'}^2 \int_{-1}^{1} dc_{12} \int_{-1}^{1} dc_{11'} \int_{-1}^{1} dc_{13'} \int_{0}^{2\pi} d\varphi_{12} \int_{0}^{2\pi} d\varphi_{11'} \int_{0}^{2\pi} d\varphi_{13'} \\ & \delta(p_2' - p_{F,2'}) E_3'^{-n} \sum_{\text{spin}} \left|\mathcal{M}^{(lf)}\right|^2_{\text{Fermi}} \ , \end{aligned} \tag{5.18}$$

$$= 32\pi^3 p_{F,1}^2 p_{F,2}^2 p_{F,1'}^2 \int_{-1}^{1} dc_{12} \int_{-1}^{1} dc_{11'} \int_{-1}^{1} dc_{13'} \int_{0}^{\pi} dv_\varphi \ \delta(p_2' - p_{F,2'}) \langle E_3'^{-n} \sum_{\text{spin}} \left|\mathcal{M}^{(lf)}\right|^2_{\text{Fermi}} \rangle_{\varphi_{13'}} \ , \tag{5.19}$$

where c_{ij} denotes the cosine of the angle between $\boldsymbol{p}_i$ and $\boldsymbol{p}_j$, $u_\varphi \equiv \varphi_{11'} + \varphi_{12}$, $v_\varphi \equiv \varphi_{11'} - \varphi_{12}$, and $\langle \cdots \rangle_{\varphi_{13'}}$ stands for an average over $\varphi_{13'}$. To obtain the second equality, we have assumed that $\langle E_3'^{-n} \sum_{\text{spin}} |\mathcal{M}|^2_{\text{Fermi}} \rangle_{\varphi_{13'}}$ and $\delta(p_2' - p_{F,2'})$ do not depend on u_φ, and may rely on v_φ only through $\cos v_\varphi$.

To simplify the expression further, we note that 2 and $2'$ represent identical particle species whereas 1 and $1'$ represent different particle species, and either $p_{F,2} \geq p_{F,1}, p_{F,1'}$ or $p_{F,2} < p_{F,1}, p_{F,1'}$. The delta function then becomes

$$\delta(p_2' - p_{F,2'}) = \frac{\delta(v_\varphi - v_{\varphi,0})}{p_{F,1'} \sqrt{(1 - c_{11'}^2)(1 - c_{12}^2)(1 - \cos^2 v_{\varphi,0})}} \ , \tag{5.20}$$

where

$$v_{\varphi,0} = \arccos \left[\frac{p_{F,1}^2 + p_{F,1'}^2 - 2p_{F,1}p_{F,1'}c_{11'} + 2p_{F,2}(p_{F,1} - p_{F,1'}c_{11'})c_{12}}{2p_{F,1'}p_{F,2}\sqrt{(1 - c_{11'}^2)(1 - c_{12}^2)}} \right] \ . \tag{5.21}$$

To have a real-valued $v_{\varphi,0}$ within the range from 0 to π, we must require $\cos^2 v_{\varphi,0} < 1$. This restricts the range of $dc_{11'}$ and dc_{12} integrals to be within

$$c_{11'}^- < c_{11'} < c_{11'}^+ \ , \quad c_{12}^- < c_{12} < c_{12}^+ \ , \tag{5.22}$$

where

$$c_{11'}^\pm = \frac{(p_{F,1} + p_{F,2}c_{12})(p_{F,1}^2 + p_{F,1'}^2 + 2p_{F,1}p_{F,2}c_{12})}{2p_{F,1'}(p_{F,1}^2 + p_{F,2}^2 + 2p_{F,1}p_{F,2}c_{12})} \tag{5.23}$$

$$\pm \frac{p_{F,2}\sqrt{(c_{12}^2-1)[(p_{F,1}^2 - p_{F,1'}^2 + 2p_{F,1}p_{F,2}c_{12})^2 - (2p_{F,2}p_{F,1'})^2]}}{2p_{F,1'}(p_{F,1}^2 + p_{F,2}^2 + 2p_{F,1}p_{F,2}c_{12})} \ , \tag{5.24}$$

and

$$c_{12}^+ = \min\left[1, \frac{p_{F,1'}^2 - p_{F,1}^2 + 2p_{F,2}p_{F,1'}}{2p_{F,1}p_{F,2}}\right] \ , \quad c_{12}^- = \max\left[-1, \frac{p_{F,1'}^2 - p_{F,1}^2 - 2p_{F,2}p_{F,1'}}{2p_{F,1}p_{F,2}}\right] \ . \tag{5.25}$$

Combining equations (5.18)-(5.25), we find

$$A = 32\pi^3 p_{F,1}^2 p_{F,2}^2 p_{F,1'} \int_{c_{12}^-}^{c_{12}^+} dc_{12} \int_{c_{11'}^-}^{c_{11'}^+} dc_{11'} \int_{-1}^{1} dc_{13'} \frac{\langle E_3'^{-n} \sum_{\text{spin}} \left|\mathcal{M}^{(lf)}\right|^2_{\text{Fermi}} \rangle_{\varphi_{13'}, v_\varphi = v_{\varphi,0}}}{\sqrt{(1-c_{11'}^2)(1-c_{12}^2)(1-\cos^2 v_{\varphi,0})}} \ . \tag{5.26}$$

We need to calculate the matrix element at the Fermi surface to evaluate this integral.

5.A.3 Matrix element

Now we evaluate the matrix element. It is convenient to use the LFV coupling

$$\mathcal{L}_{\text{LFV}} = -ig_{ae\mu}a(\overline{\Psi}_e\gamma_5\Psi_\mu + \overline{\Psi}_\mu\gamma_5\Psi_e) \ , \tag{5.27}$$

which is equivalent to the use of the pseudovector (derivative) form written in the main text if each fermion line is attached to at most one axion line (97). Given the two Feynman diagrams in figure 5.A.1, the matrix elements are

$$i\mathcal{M}^{(1)} = \pm e^2 g_{ae\mu}\left[\overline{u}_1'\gamma^\mu \frac{-\not{r} + m_1'}{r^2 + m_1'^2}\gamma_5 u_1\right] \frac{-g_{\mu\nu}}{k^2}[\overline{u}_2'\gamma^\nu u_2] \ , \tag{5.28}$$

$$i\mathcal{M}^{(2)} = \pm e^2 g_{ae\mu}\left[\overline{u}_1'\gamma_5 \frac{-\not{s} + m_1}{s^2 + m_1^2}\gamma^\mu u_1\right] \frac{-g_{\mu\nu}}{k^2}[\overline{u}_2'\gamma^\nu u_2] \ , \tag{5.29}$$

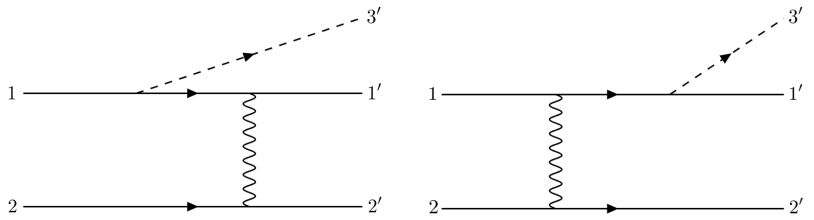


Figure 5.A.1 : Feynman diagrams for the LFV process $l + f \to l' + f + a$. If f is a lepton, there occur two more graphs which can be obtained by exchanging $(1 \leftrightarrow 2)$ for f being identical to l or $(1' \leftrightarrow 2')$ for f being identical to l'.

where $k \equiv p_2 - p_2'$, $r \equiv p_1 - p_3'$, $s \equiv p_1' + p_3'$ and $\pm$ refers to the sign of the spectator particle's electric charge. In NSs we have $|m_1^2 - {m_1'}^2| \approx m_\mu^2 \gg E_F E_3'$, thus $r^2 + {m_1'}^2 \approx -m_1^2 + {m_1'}^2$ and $s^2 + m_1^2 \approx -{m_1'}^2 + m_1^2$. The matrix element for exchange diagrams can be obtained by $(1 \leftrightarrow 2)$ or $(1' \leftrightarrow 2')$, with an additional factor of -1 included.

The spin-summed squared matrix element is

$$\sum_{\text{spin}} \left|\mathcal{M}^{(lp)}\right|^2 = -\frac{128 g_{ae\mu}^2 e^4}{(p_2 - p_2')^4} \frac{(p_1 \cdot p_1' + m_1 m_1')(p_2 \cdot p_3')(p_2' \cdot p_3')}{(m_1^2 - {m_1'}^2)^2} \ , \tag{5.30}$$

$$\sum_{\text{spin}} \left|\mathcal{M}^{(ll)}\right|^2 = \sum_{\text{spin}} \left|\mathcal{M}^{(lp)}\right|^2 + (1 \leftrightarrow 2) + \mathcal{T}^{(ll)} \ , \tag{5.31}$$

$$\sum_{\text{spin}} \left|\mathcal{M}^{(ll')}\right|^2 = \sum_{\text{spin}} \left|\mathcal{M}^{(lp)}\right|^2 + (1' \leftrightarrow 2') + \mathcal{T}^{(ll')} \ , \tag{5.32}$$

where $l = e, \mu$ and $l' = \mu, e$. The second term in (5.31) and (5.32) is the contribution solely from the exchange diagrams given by the first term but with $(1 \leftrightarrow 2)$. The third term in (5.31) is the interference between prototype and exchange diagrams given by

$$\begin{aligned}\mathcal{T}^{(ll)} = &\frac{64 g_{ae\mu}^2 e^4}{(p_1 - p_2')^2 (p_2 - p_2')^2} \frac{p_2' \cdot p_3'}{(m_1^2 - {m_1'}^2)^2} \\ &\times \left[(p_2 \cdot p_1' + m_1 m_1')(p_1 \cdot p_3') + (p_1 \cdot p_1' + m_1 m_1')(p_2 \cdot p_3') - (p_1 \cdot p_2 + m_1^2)(p_1' \cdot p_3')\right] \ ,\end{aligned} \tag{5.33}$$

and $T^{(ll')}$ in (5.32) by $\mathcal{T}^{(ll)}$ but with $(1 \leftrightarrow 1')$ and $(2 \leftrightarrow 2')$. Here we evaluate the

traces of products of gamma matrices and spinors with the help of the Mathematica package FeynCalc (98).

At the Fermi surface, the spin-summed squared matrix element becomes

$$\sum_{\text{spin}} \left|\mathcal{M}^{(lf)}\right|^2_{\text{Fermi}} = \frac{32e^4 g_{ae\mu}^2 {E_3'}^2}{E_{F,1}^2 E_{F,2}^2 \beta_2^4 (\beta_1^2 - {\beta_1'}^2)^2} G^{(lf)} \ , \tag{5.34}$$

where $f = p, e, \mu$. The $G^{(lf)}$ factor is found to be

$$G^{(lp)} = \frac{(1-\beta_{F,2}c_{23'})(1-\beta_{F,2}c_{2'3'})(1-\beta_{F,1}\beta_{F,1'}c_{11'})}{(1-c_{22'})^2} \ , \tag{5.35}$$

$$G^{(ll)} = G^{(lp)} + (1 \leftrightarrow 2) + H^{(ll)} \ , \tag{5.36}$$

$$G^{(ll')} = G^{(lp)} + (1' \leftrightarrow 2') + H^{(ll')} \ , \tag{5.37}$$

where we have assumed that electrons are ultra relativistic so $\beta_{F,e} = 1$. The second term in (5.36) and (5.37) is the contribution solely from the exchange diagrams given by the first term but with $(1 \leftrightarrow 2)$. The third term in (5.36) is the interference between prototype and exchange diagrams given by

$$\begin{aligned} H^{(ll)} = \frac{(1-\beta_{F,1}c_{2'3'})}{2(1-c_{12'})(1-c_{22'})} [\beta_{F,1}\big(c_{13'} + c_{23'} + \beta_{F,1}(1-c_{12}) + \beta_{F,1'}(c_{11'} + c_{21'}) \\ + \beta_{F,1}\beta_{F,1'}(c_{12}c_{1'3'} - c_{11'}c_{23'} - c_{13'}c_{21'} - c_{1'3'})\big) - 2] \ , \end{aligned} \tag{5.38}$$

and $H^{(ll')}$ in (5.37) by $H^{(ll)}$ but with $(1 \leftrightarrow 1')$ and $(2 \leftrightarrow 2')$.

5.A.4 Axion emissivity

In summary, the axion emissivity is given by

$$\varepsilon_a^{(lf)} = \frac{328\pi^2\alpha^2 g_{ae\mu}^2}{945 m_\mu^4} \frac{\beta_{F,1} E_{F,1}^3}{\beta_{F,2}^2 p_{F,2}^2} F^{(lf)} T^8 \ , \tag{5.39}$$

$$F^{(lf)} \equiv \frac{1}{8S} \int_{c_{12}^-}^{c_{12}^+} dc_{12} \int_{c_{11'}^-}^{c_{11'}^+} dc_{11'} \int_{-1}^{1} dc_{13'} \frac{\langle G^{(lf)} \rangle_{\varphi_{13'}, v_\varphi = v_{\varphi,0}}}{\sqrt{(1-c_{11'}^2)(1-c_{12}^2)(1-\cos^2 v_{\varphi,0})}} \ . \tag{5.40}$$

The $dc_{13'}$ integral can be evaluated analytically. We calculate the other integrals using numerical techniques and present the result for $F^{(lf)}$ in figure 5.A.2. In the left

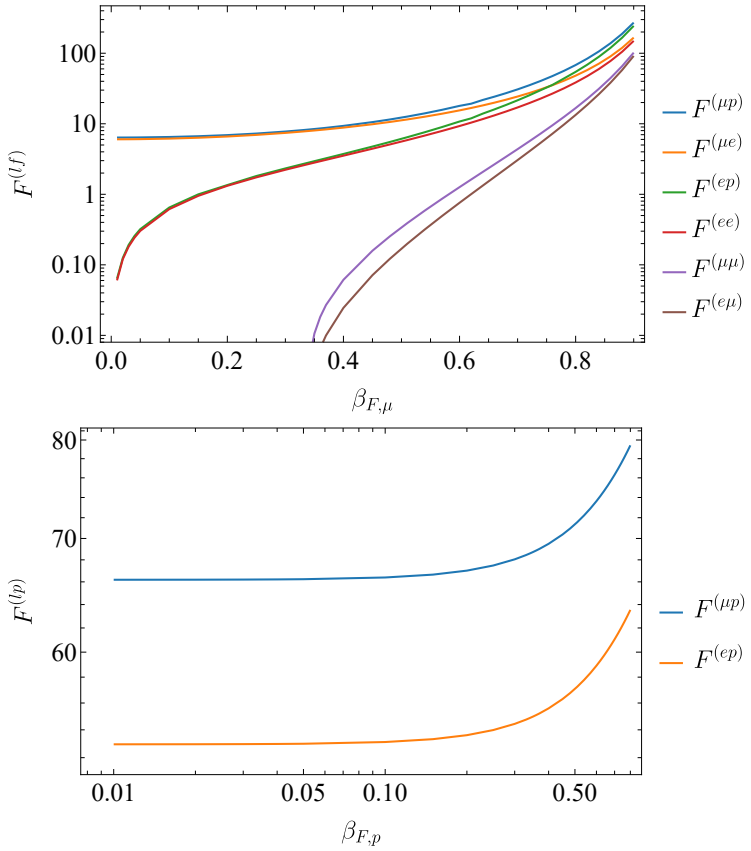


Figure 5.A.2 : The factor $F^{(lf)}$ as a function of the Fermi velocity of muons (left) and protons (right). Here we have set $\beta_{F,p} = 0.3$ and $\beta_{F,\mu} = 0.8$ for the left and right panels respectively for the $f = p$ processes.

panel we vary the muon Fermi velocity $\beta_{F,\mu} = p_{F,\mu}/E_{F,\mu}$. From the right panel we see that $F^{(lp)}$ is not sensitive to $\beta_{F,p}$ if protons are nonrelativistic, i.e., $\beta_{F,p} \lesssim 0.5$, which is expected in NSs. Therefore, we use the values of $F^{(lf)}$ shown in the left panel to calculate the emissivity shown in the main text.

5.A.5 Different temperature dependence from LFV and LFP interactions

In the main text we contrast the temperature dependence of the axion emissivity for LFV and LFP interactions. The LFP interaction leads to axion emission via

channels such as $l + f \to l + f + a$ with an emissivity that scales as $\varepsilon_a \propto T^6$ (similar for $nn \to nna$ (24)). By considering the LFV interaction here, we find that channels such as $l + f \to l' + f + a$ lead to an emissivity $\varepsilon_a \propto T^8$ instead. This different scaling may be understood by inspecting the form of the matrix element. Consider the Feynman diagram in the left panel of figure 5.A.1. The fermion propagator and the axion vertex contribute factors of

$$\frac{E_3'}{(p_1 - p_3')^2 + m_1'^2} = \frac{E_3'}{m_1'^2 - m_1^2 + 2E_3'(E_1 - 2|\boldsymbol{p}_1|c_{13'})} \ , \tag{5.41}$$

in the $(-,+,+,+)$ metric signature and neglecting the axion mass $E_3' = |\boldsymbol{p}_3'|$. The axion energy E_3' in the numerator arises from the derivative nature of the axion interaction. The temperature dependence enters via the typical axion energy, $E_3' \sim T$. For LFP channels such as $\mu p \to \mu p a$, we have $m_1' = m_1$, the $E_3' \sim T$ factor in the numerator is canceled by the factor in the denominator, and consequently the squared matrix element is insensitive to the temperature. On the other hand, for the LFV channels, the $m_1'^2 - m_1^2$ term dominates in the denominator. Consequently, the LFV axion emissivity is suppressed relative to the LFP calculation by a factor of order $T^2 E_{F,e}^2/(m_\mu^2 - m_e^2)^2 \sim T^2/m_\mu^2 \sim 7 \times 10^{-7} T_9^2$.

5.B The $npe\mu$ matter

At typical NS densities $\sim 10^{15}\,\mathrm{g\ cm^{-3}}$, the equilibrium composition involves neutrons, protons, electrons, muons and other exotic matter states such as hyperons. Neglecting the exotic matter, equations of state for a NS are relatively easy to calculate (99). Thermal equilibrium and conservation of the baryon number and electric charge impose (63)

$$E_{F,\mu} = E_{F,e} \ , \quad E_{F,n} = E_{F,p} + E_{F,e} \ , \quad n_p = n_e + n_\mu \ , \tag{5.42}$$

where we have approximated the chemical potential with the Fermi energy. We also have the Fermi energy $E_{F,i}^2 = m_i^2 + p_{F,i}^2$, the number density $n_i = p_{F,i}^3/3\pi^2$, and the

mass density $\rho = \sum_i m_i n_i$. If one of $\rho, n_n, n_p, n_e, n_\mu$ is fixed, the other quantities can be fully determined. For this work, we have taken $0.8\, m_N \approx 750\,\mathrm{MeV}$ for the mass of nucleons to account for their nuclear interactions. At $\rho = 6\rho_0 \approx 1.5 \times 10^{15}\,\mathrm{g\ cm^{-3}}$, we find

$$p_{F,n} \simeq 624\,\mathrm{MeV}\ , \quad p_{F,p} \simeq 226\,\mathrm{MeV}\ , \quad p_{F,e} \simeq 193\,\mathrm{MeV}\ , \quad p_{F,\mu} \simeq 162\,\mathrm{MeV}\ , \tag{5.43}$$

corresponding to $\beta_{F,p} \simeq 0.29$ and $\beta_{F,\mu} \simeq 0.84$.

5.C Numerical integration

5.C.1 Numerical integrator

In this section we discuss the numerical method used to evaluate (5.11). To prepare the integrand for numerical integration we simplify it by using the Dirac deltas to perform 4 integrals analytically. We use the momentum conserving Dirac delta to carry out the d^3p_2' integrals which enforces $\boldsymbol{p}_2' = \boldsymbol{p}_1 + \boldsymbol{p}_2 - \boldsymbol{p}_1' - \boldsymbol{p}_3'$. Next, we rewrite the momentum integrals in spherical coordinates by making the replacements $d^3p \to |\boldsymbol{p}|^2\, d|\boldsymbol{p}|\, d\cos\theta\, d\phi$ where θ and ϕ give the polar and azimuthal angles of $\boldsymbol{p}$ in the rest frame of the NS. The coordinate system is oriented so that the z-axis points in the same direction as $\boldsymbol{p}_3'$ so that the $d\cos\theta_3'\, d\phi_3'$ integral yields a trivial factor of 4π. We then change variables from momentum magnitudes $|\boldsymbol{p}|$ to energies by using the relation $E^2 = |\boldsymbol{p}|^2 + m^2$ to write $E\, dE = |\boldsymbol{p}|\, d|\boldsymbol{p}|$. Finally, the energy Dirac delta is used to fix $|\boldsymbol{p}_3'|$ so that, assuming the axion is massless ($m_3' = 0$),

$$\begin{aligned} &E_1 + E_2 - E_1' - E_2' - E_3' = E_1 + E_2 - E_1' - \sqrt{|\boldsymbol{p}_1 + \boldsymbol{p}_2 - \boldsymbol{p}_1' - \boldsymbol{p}_3'|^2 + m_2'^2} - |\boldsymbol{p}_3'| \\ &= E_1 + E_2 - E_1' - \sqrt{|\boldsymbol{P}|^2 - 2\, P_z\, |\boldsymbol{p}_3'| + |\boldsymbol{p}_3'|^2 + m_2'^2} - |\boldsymbol{p}_3'| = 0 \end{aligned} \tag{5.44}$$

where $\boldsymbol{P} \equiv \boldsymbol{p}_1 + \boldsymbol{p}_2 - \boldsymbol{p}_1' - \boldsymbol{p}_3'$. This adds a factor of $|1 + (|\boldsymbol{p}_3'| - P_z)/E_2|^{-1}$ to the integrand since $\delta[f(x)] = \delta(x - x_*)/|f'(x_*)|$ where x_* is the root of $f(x)$. In practice, (5.44) is enforced by using Newton-Raphson iteration to find the value of $|\boldsymbol{p}_3'|$ which is a root of

this equation when the integration variables E_1, E_2, E_1', $\cos\theta_1$, $\cos\theta_2$, $\cos\theta_1'$, ϕ_1, ϕ_2, and ϕ_1' are fixed. All together, this rewrites the integral (5.11) as

$$\varepsilon_a^{(lf)} = \frac{4\pi}{2^5(2\pi)^{11}}\frac{1}{S}\int dE_1\, d\cos\theta_1\, d\phi_1\, dE_2\, d\cos\theta_2\, d\phi_2\, dE_1'\, d\cos\theta_1'\, d\phi_1' \times \frac{|\boldsymbol{p}_1||\boldsymbol{p}_2||\boldsymbol{p}_1'||\boldsymbol{p}_3'|}{E_2'\left|1+(E_3'-P_z)/E_2'\right|}\sum_{\text{spin}}\left|\mathcal{M}^{(lf)}\right|^2 E_3'\, f_1\, f_2\,(1-f_1')\,(1-f_2')\ , \tag{5.45}$$

where the matrix element is given by (5.30 - 5.32). We evaluate the integral in this form using the Vegas package in Python which performs Monte Carlo integration using two adaptive strategies: importance sampling, and stratified sampling, to improve convergence (100). We choose to use this Monte Carlo integrator because of its flexibility and ease of use. The integral is evaluated by passing the integrand as an explicit function of the 9 integration variables $(E_1, E_2, E_1', \cos\theta_1, \cos\theta_2, \cos\theta_1', \phi_1, \phi_2, \phi_1')$ to an instance of the `vegas.Integrator` class. We split the calculation of the integral into two steps. First, we adapt the `vegas.Integrator` object to the integrand by calling it with the parameters `nitn` $= 10$, `neval` $= 5\times10^7$, and `alpha` $= 0.1$. These parameters control the number of iterations used to adapt the integrator; the number of points on the integration domain where the integrand is evaluated; and the sensitivity of the adaptation algorithms, respectively. We then discard the results obtained from the first run but keep the adapted integrator and call it again with the same parameter choices except with `alpha = 0` so that there is no further adaptation. The value of the integral and the errors we report below are taken as the `mean` and `sdev` attributes of the second run `vegas.Integrator` object. The mean is a weighted average of the results of each of the `nitn = 10` iterations of the Vegas algorithm, where the weights are the inverse variance in each iteration. The uncertainty, `sdev` is the square root of the variance of the weighted average assuming the sample average in each iteration is approximately normally distributed – this is a good approximation if `neval` is sufficiently large.

In principle the energy integrals over E_1, E_2, E_1' should be over the domain $E_i \in$

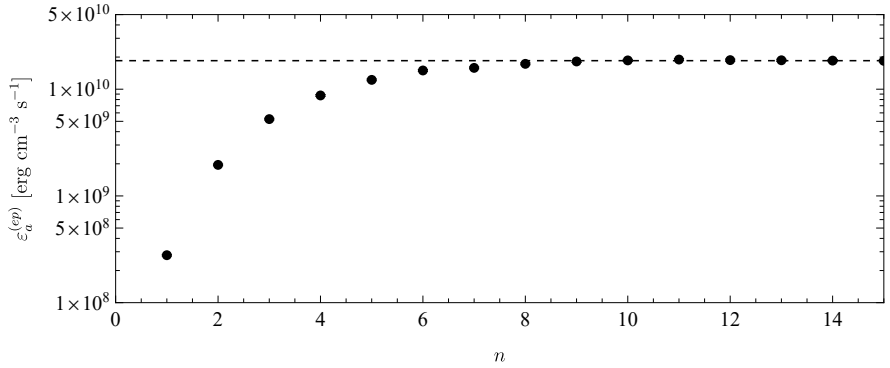


Figure 5.C.1 : Axion emissivity for the $ep \to \mu pa$ channel vs energy integration domain $E_i \in [E_{F,i} - nT,\ E_{F,i} + nT]$ parameterized by n. The gray dashed line is the constant $1.85 \times 10^{10}\,\mathrm{erg\,cm^{-3}\,s^{-1}}$, which is the value to which the integral converges. For these calculations we have fixed $\beta_{F,\mu} = 0.836788$, $g_{ae\mu} = 10^{-11}$, and $T = 10^9\,\mathrm{K}$. As n increases, the value of the emissivity integral converges to a constant value of $\approx 1.8 \times 10^{10}\,\mathrm{erg\,cm^{-3}\,s^{-1}}$.

$[m_i, \infty)$ but in practice we can only integrate over a finite window. The thermal factors in (5.45) provide support only in a window around the Fermi level $E_F = \sqrt{p_F^2 + m^2}$ whose width is of order $\sim T$. This motivates integrating E_1, E_2, E_1' over the finite window $E_i \in [\max(m_i, E_{F,i} - nT),\ E_{F,i} + nT]$ with a value of n sufficiently large that the integral is insensitive to its exact value. We find $n = 10$ to be large enough that the integral is independent of n, but small enough that Monte Carlo convergence is not too slow. The n-independence is demonstrated for the process $ep \to \mu pa$ for $\beta_{F,\mu} = 0.84$, $T = 10^9$ K and $m_a = 0$ in figure 5.C.1. Note how as n increases, the emissivity approaches a constant value of approximately 1.8×10^{10} erg cm^{-3} s^{-1}, which corresponds to the blue data point at $\beta_{F,\mu} \approx 0.84$ in figure 5.C.2.

5.C.2 Numerical validation of Fermi surface approximation

The results of our numerical evaluations of (5.45) for the various axion emission channels are shown in figure 5.C.2. The numerical results (dots and squares) agree

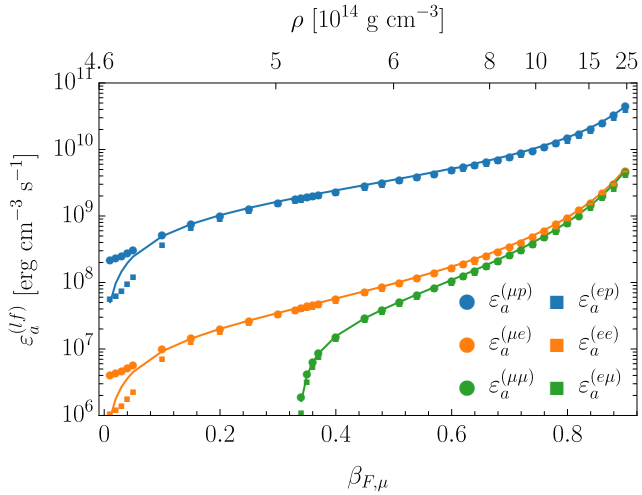


Figure 5.C.2 : Axion emissivity computed using the Monte Carlo integration method (dots and squares) vs. Fermi surface approximation (lines). The results agree well for $\beta_{F,\mu} \gtrsim 0.1$ and the agreement is good within about 10% at $\beta_{F,\mu} \approx 0.8$. At small $\beta_{F,\mu} \lesssim 0.1$, the Fermi surface approximation underestimates the emissivity for the μp, μe channels and overestimates it for the ep and ee channels. To make this plot, we choose $g_{ae\mu} = 10^{-11}$ and $T = 10^9$ K to be consistent with the parameters in figure 2 of the main text.

very well with the analytical results (lines) for a wide range of $\beta_{F,\mu}$. For small $\beta_{F,\mu}$ the numerical results tend to diverge from the analytical results, which is expected because in this regime the number density of muons is small, which means that the degenerate matter approximation breaks down. In addition, we observe that for $\beta_{F,\mu} \gtrsim 0.1$ the emissivities are paired by channel such that $\varepsilon_a^{(lf)} \approx \varepsilon_a^{(l'f)}$. This is a consequence of the strong particle degeneracy and the beta equilibrium condition $E_{F,e} \approx E_{F,\mu}$. We have verified this numerically and analytically by imposing the relation $E_{F,e} = E_{F,\mu} + \Delta$ and observing that the difference between emissivities for the channels $ef \to \mu fa$ and $\mu f \to efa$ grows with Δ but is only significant if $\Delta \gtrsim T$. For $\Delta > 0$ the electron's Fermi energy is larger than the muon's which allows for electrons with energies below the Fermi level to also convert into muons, enhancing the emissivity of this channel. Conversely, for muon to electron conversion, the emissivity is exponentially suppressed since the muons' energies are below the electrons' energies.

5.C.3 Effect of temperature on axion emissivity

In addition to verifying that numerically evaluating the axion emissivity at $T = 10^9$ K agrees with the analytical approximation, we also numerically computed the axion emissivity as a function of temperature while fixing $\beta_{F,\mu} = 0.836788$ and $m_a = 0$. We are motivated to do this for two reasons. The first is to confirm the T^8 scaling of the emissivity at low temperatures, i.e. equation (5.39). The second is to calculate the emissivity for larger temperatures such as $T \sim 50$ MeV, the scale of supernovae; allowing us to comment on constraints imposed on axion LFV interactions by supernovae observations.

In degenerate NS matter, there is a screening of electromagnetic fields due to the presence of a degenerate plasma with charged constituents. To estimate this effect, we replace the photon propagator k^{-2} in the matrix element by $(k^2 + k_{\rm TF}^2)^{-1}$ (101), where $k_{\rm TF}^2 = \sum_i 4\alpha p_{F,i} E_{F,i}/\pi$ is the Thomas-Fermi screening scale which receives contributions from electrons, muons and protons. Noting that $k^2 \sim (p_{F,e} - p_{F,\mu})^2 \sim$

$E_{F,e}^2(1-\beta_{F,\mu})^2$ at low temperatures, the screening effect is insignificant if $\beta_{F,\mu} \lesssim 1 - k_{\rm TF}/E_{F,e}$, which becomes $\beta_{F,\mu} \lesssim 0.75$ at the core condition given by (5.43). Therefore, for mildly relativistic muons with $\beta_{F,\mu} \sim 0.8$, the emissivity of LFV axions without including the screening effect is subject to $\mathcal{O}(1)$ corrections. On the other hand, incorporating the screening effect in axion emissivities is important at high temperatures since $k_{\rm TF}^2$ dominates over k^2, especially near the pole $k^2 = 0$.

The temperature dependence of the axion emissivity is presented in figure 5.C.3 for 10^{-3} MeV $\leq T \leq 100$ MeV. Since we expect the emissivity to scale as $\varepsilon^{(lf)} \propto T^8$ for low temperatures we normalize the emissivity by T^8 so that a T^8 scaling would be a constant line in this figure. The figure displays several interesting features. (1) At temperatures below $T \sim 10$ MeV, the emissivity is seen to scale like $\varepsilon^{(lf)} \propto T^8$ (up to $\mathcal{O}(1)$ factors), which confirms the prediction from the Fermi surface approximation. (2) The emissivity tends to decrease relative to T^8 for all six channels at temperatures $T \gtrsim 10\,\text{MeV}$. (3) For lower temperatures, the emissivities are paired by channel such that $\varepsilon_a^{(lf)} \approx \varepsilon_a^{(l'f)}$; however, at higher temperatures these relations do not hold. This is expected since the Fermi surface approximation, one of the assumptions needed to show that $\varepsilon_a^{(lf)}$ and $\varepsilon_a^{(l'f)}$ coincide, breaks down in this regime. The significance of $T = 10\,\text{MeV}$ can be understood as follows: at low temperatures the thermal factors lead to a strong suppression of the integrand away from the Fermi surface. As we lift the temperature the accessible phase space broadens and the pole becomes significant.

5.C.4 Effect of axion mass on emissivity

In previous results we assumed axions were massless. Here, we use our numerical integration method to explore the effect of raising the axion mass on the emissivity. To do this we must modify (5.44) to accommodate a massive axion by replacing $E_3' = |\boldsymbol{p}_3'|$ with $E_3' = \sqrt{|\boldsymbol{p}_3'|^2 + m_a^2}$ so that energy conservation imposes the following constraint on $|\boldsymbol{p}_3'|$,

$$E_1 + E_2 - E_1' - \sqrt{|\boldsymbol{P}|^2 - 2\,P_z\,|\boldsymbol{p}_3'| + |\boldsymbol{p}_3'|^2 + m_2'^2} - \sqrt{|\boldsymbol{p}_3'|^2 + m_a^2} = 0\ . \tag{5.46}$$

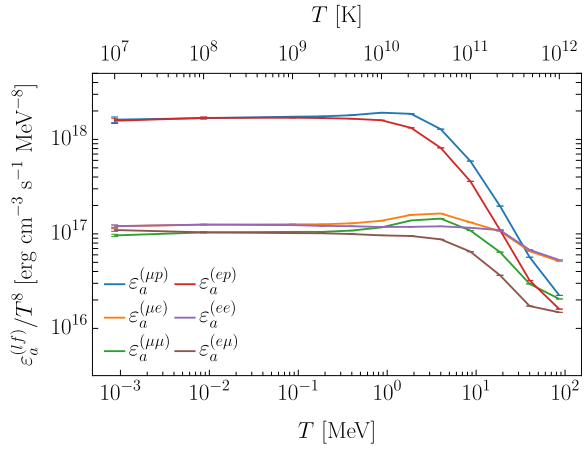


Figure 5.C.3 : Numerically evaluated axion emissivity vs. temperature, calculated using (5.45) with matrix elements given by (5.30 - 5.32). To generate these data we fixed $\beta_{F,\mu} = 0.836788$, $m_a = 0$, and $g_{ae\mu} = 10^{-11}$. The data presented here were computed with $\texttt{neval} = 5 \times 10^7$. The error bars are typically between 100 to 10,000 times smaller the mean values.

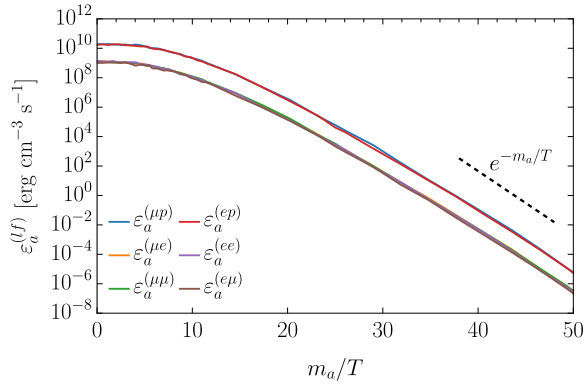


Figure 5.C.4 : Numerically evaluated axion emissivity vs. axion mass, calculated using (5.45) with matrix elements given by (5.30 - 5.32). To generate these data we fixed $\beta_{F,\mu} = 0.836788$, $T = 10^9$ K, and $g_{ae\mu} = 10^{-11}$. For large masses the emissivity falls off with an exponential tail (compare with black dashed line). The data presented here were computed with $\texttt{neval} = 10^6$.

In principle, we must also account for the axion's mass in the matrix element since (5.30)–(5.32) were derived assuming $m_a = 0$. However, we argue that the most important contribution of the mass to the emissivity is an exponential suppression arising from the thermal factors and therefore report results obtained using the 'massless' matrix element of (5.30 - 5.32). We set the temperature T to a fiducial value of 10^9 K and fix $\beta_{F,\mu} = 0.836788$ and calculate the emissivity for a range of masses satisfying $0 \leq m_a/T \leq 50$. The emissivities calculated are presented in figure 5.C.4. We find that the emissivity is approximately constant for $m_a/T \leq 10$, after which point the emissivity is exponentially suppressed.

Bibliography

[1] Zhang, Hong-Yi and Hagimoto, Ray and Long, Andrew J., *Neutron star cooling with lepton-flavor-violating axions*, *Phys. Rev. D* **109** (2024) 103005, [2309.03889].

[2] Kim, Jihn E., *Light Pseudoscalars, Particle Physics and Cosmology*, *Phys. Rept.* **150** (1987) 1–177.

[3] Peccei, R. D. and Quinn, Helen R., *CP Conservation in the Presence of Instantons*, *Phys. Rev. Lett.* **38** (1977) 1440–1443.

[4] Weinberg, Steven, *A New Light Boson?*, *Phys. Rev. Lett.* **40** (1978) 223–226.

[5] Wilczek, Frank, *Problem of Strong P and T Invariance in the Presence of Instantons*, *Phys. Rev. Lett.* **40** (1978) 279–282.

[6] Kim, Jihn E., *Weak Interaction Singlet and Strong CP Invariance*, *Phys. Rev. Lett.* **43** (1979) 103.

[7] Shifman, Mikhail A. and Vainshtein, A. I. and Zakharov, Valentin I., *Can Confinement Ensure Natural CP Invariance of Strong Interactions?*, *Nucl. Phys. B* **166** (1980) 493–506.

[8] Zhitnitsky, A. R., *On Possible Suppression of the Axion Hadron Interactions. (In Russian)*, *Sov. J. Nucl. Phys.* **31** (1980) 260.

[9] Dine, Michael and Fischler, Willy and Srednicki, Mark, *A Simple Solution to the Strong CP Problem with a Harmless Axion*, *Phys. Lett. B* **104** (1981) 199–202.

[10] Borsanyi, Sz. and others, *Calculation of the axion mass based on high-temperature lattice quantum chromodynamics*, *Nature* **539** (2016) 69–71, [1606.07494].

[11] Gorghetto, Marco and Villadoro, Giovanni, *Topological Susceptibility and QCD Axion Mass: QED and NNLO corrections*, *JHEP* **03** (2019) 033, [1812.01008].

[12] Svrcek, Peter and Witten, Edward, *Axions In String Theory*, *JHEP* **06** (2006) 051, [hep-th/0605206].

[13] Arvanitaki, Asimina and Dimopoulos, Savas and Dubovsky, Sergei and Kaloper, Nemanja and March-Russell, John, *String Axiverse*, *Phys. Rev. D* **81** (2010) 123530, [0905.4720].

[14] Ringwald, Andreas, *Searching for axions and ALPs from string theory*, *J. Phys. Conf. Ser.* **485** (2014) 012013, [1209.2299].

[15] Dine, Michael and Fischler, Willy, *The Not So Harmless Axion*, *Phys. Lett. B* **120** (1983) 137–141.

[16] Abbott, L. F. and Sikivie, P., *A Cosmological Bound on the Invisible Axion*, *Phys. Lett. B* **120** (1983) 133–136.

[17] Preskill, John and Wise, Mark B. and Wilczek, Frank, *Cosmology of the Invisible Axion*, *Phys. Lett. B* **120** (1983) 127–132.

[18] Ferreira, Elisa G. M., *Ultra-light dark matter*, *Astron. Astrophys. Rev.* **29** (2021) 7, [2005.03254].

[19] Hui, Lam, *Wave Dark Matter*, *Ann. Rev. Astron. Astrophys.* **59** (2021) 247–289, [2101.11735].

[20] Marsh, David J. E., *Axion Cosmology*, *Phys. Rept.* **643** (2016) 1–79, [1510.07633].

[21] Di Luzio, Luca and Giannotti, Maurizio and Nardi, Enrico and Visinelli, Luca, *The landscape of QCD axion models*, *Phys. Rept.* **870** (2020) 1–117, [2003.01100].

[22] Sikivie, Pierre, *Invisible Axion Search Methods*, *Rev. Mod. Phys.* **93** (2021) 015004, [2003.02206].

[23] Particle Data Group collaboration, Workman, R. L. and others, *Review of Particle Physics*, *PTEP* **2022** (2022) 083C01.

[24] Raffelt, Georg G., *Astrophysical methods to constrain axions and other novel particle phenomena*, *Phys. Rept.* **198** (1990) 1–113.

[25] Miller Bertolami, Marcelo M. and Melendez, Brenda E. and Althaus, Leandro G. and Isern, Jordi, *Revisiting the axion bounds from the Galactic white dwarf luminosity function*, *JCAP* **10** (2014) 069, [1406.7712].

[26] Giannotti, Maurizio and Irastorza, Igor G. and Redondo, Javier and Ringwald, Andreas and Saikawa, Ken'ichi, *Stellar Recipes for Axion Hunters*, *JCAP* **10** (2017) 010, [1708.02111].

[27] J Isern and E Garcí a-Berro and S Torres and R Cojocaru and S Catalán, *Axions and the luminosity function of white dwarfs: the thin and thick discs, and the halo*, *Monthly Notices of the Royal Astronomical Society* **478** (may, 2018) 2569–2575.

[28] Córsico, Alejandro H. and Althaus, Leandro G. and Miller Bertolami, Marcelo M. and Kepler, S. O., *Pulsating white dwarfs: new insights*, *Astron. Astrophys. Rev.* **27** (2019) 7, [1907.00115].

[29] Hamaguchi, Koichi and Nagata, Natsumi and Yanagi, Keisuke and Zheng, Jiaming, *Limit on the Axion Decay Constant from the Cooling Neutron Star in Cassiopeia A*, *Phys. Rev. D* **98** (2018) 103015, [1806.07151].

[30] Beznogov, Mikhail V. and Rrapaj, Ermal and Page, Dany and Reddy, Sanjay, *Constraints on Axion-like Particles and Nucleon Pairing in Dense Matter from the Hot Neutron Star in HESS J1731-347*, *Phys. Rev. C* **98** (2018) 035802, [1806.07991].

[31] Buschmann, Malte and Dessert, Christopher and Foster, Joshua W. and Long, Andrew J. and Safdi, Benjamin R., *Upper Limit on the QCD Axion Mass from Isolated Neutron Star Cooling*, *Phys. Rev. Lett.* **128** (2022) 091102, [2111.09892].

[32] Burrows, Adam and Turner, Michael S. and Brinkmann, R. P., *Axions and SN 1987a*, *Phys. Rev. D* **39** (1989) 1020.

[33] Burrows, Adam and Ressell, M. Ted and Turner, Michael S., *Axions and SN1987A: Axion trapping*, *Phys. Rev. D* **42** (1990) 3297–3309.

[34] Keil, Wolfgang and Janka, Hans-Thomas and Schramm, David N. and Sigl, Gunter and Turner, Michael S. and Ellis, John R., *A Fresh look at axions and SN-1987A*, *Phys. Rev. D* **56** (1997) 2419–2432, [astro-ph/9612222].

[35] Hanhart, Christoph and Phillips, Daniel R. and Reddy, Sanjay, *Neutrino and axion emissivities of neutron stars from nucleon-nucleon scattering data*, *Phys. Lett. B* **499** (2001) 9–15, [astro-ph/0003445].

[36] Fischer, Tobias and Chakraborty, Sovan and Giannotti, Maurizio and Mirizzi, Alessandro and Payez, Alexandre and Ringwald, Andreas, *Probing axions with the neutrino signal from the next galactic supernova*, *Phys. Rev. D* **94** (2016) 085012, [1605.08780].

[37] Carenza, Pierluca and Fischer, Tobias and Giannotti, Maurizio and Guo, Gang and Martínez-Pinedo, Gabriel and Mirizzi, Alessandro, *Improved axion emissivity from a supernova via nucleon-nucleon bremsstrahlung*, *JCAP* **10** (2019) 016, [1906.11844].

[38] Carenza, Pierluca and Fore, Bryce and Giannotti, Maurizio and Mirizzi, Alessandro and Reddy, Sanjay, *Enhanced Supernova Axion Emission and its Implications*, *Phys. Rev. Lett.* **126** (2021) 071102, [2010.02943].

[39] Lella, Alessandro and Carenza, Pierluca and Co', Giampaolo and Lucente, Giuseppe and Giannotti, Maurizio and Mirizzi, Alessandro and Rauscher, Thomas, *Getting the most on supernova axions*, *Phys. Rev. D* **109** (2024) 023001, [2306.01048].

[40] Davidson, Aharon and Wali, Kameshwar C., *MINIMAL FLAVOR UNIFICATION VIA MULTIGENERATIONAL PECCEI-QUINN SYMMETRY*, *Phys. Rev. Lett.* **48** (1982) 11.

[41] Wilczek, Frank, *Axions and Family Symmetry Breaking*, *Phys. Rev. Lett.* **49** (1982) 1549–1552.

[42] Anselm, A. A. and Uraltsev, N. G. and Khlopov, M. Yu., *mu —> e FAMILON DECAY*, *Sov. J. Nucl. Phys.* **41** (1985) 1060.

[43] Feng, Jonathan L. and Moroi, Takeo and Murayama, Hitoshi and Schnapka, Erhard, *Third generation familons, b factories, and neutrino cosmology*, *Phys. Rev. D* **57** (1998) 5875–5892, [hep-ph/9709411].

[44] Bauer, Martin and Schell, Torben and Plehn, Tilman, *Hunting the Flavon*, *Phys. Rev. D* **94** (2016) 056003, [1603.06950].

[45] Ema, Yohei and Hamaguchi, Koichi and Moroi, Takeo and Nakayama,

Kazunori, *Flaxion: a minimal extension to solve puzzles in the standard model*, *JHEP* **01** (2017) 096, [1612.05492].

[46] Calibbi, Lorenzo and Goertz, Florian and Redigolo, Diego and Ziegler, Robert and Zupan, Jure, *Minimal axion model from flavor*, *Phys. Rev. D* **95** (2017) 095009, [1612.08040].

[47] Choi, Kiwoon and Im, Sang Hui and Park, Chan Beom and Yun, Seokhoon, *Minimal Flavor Violation with Axion-like Particles*, *JHEP* **11** (2017) 070, [1708.00021].

[48] Chala, Mikael and Guedes, Guilherme and Ramos, Maria and Santiago, Jose, *Running in the ALPs*, *Eur. Phys. J. C* **81** (2021) 181, [2012.09017].

[49] Bauer, Martin and Neubert, Matthias and Renner, Sophie and Schnubel, Marvin and Thamm, Andrea, *The Low-Energy Effective Theory of Axions and ALPs*, *JHEP* **04** (2021) 063, [2012.12272].

[50] Bonilla, J. and Brivio, I. and Gavela, M. B. and Sanz, V., *One-loop corrections to ALP couplings*, *JHEP* **11** (2021) 168, [2107.11392].

[51] Panci, Paolo and Redigolo, Diego and Schwetz, Thomas and Ziegler, Robert, *Axion dark matter from lepton flavor-violating decays*, *Phys. Lett. B* **841** (2023) 137919, [2209.03371].

[52] Calibbi, Lorenzo and Signorelli, Giovanni, *Charged Lepton Flavour Violation: An Experimental and Theoretical Introduction*, *Riv. Nuovo Cim.* **41** (2018) 71–174, [1709.00294].

[53] Petcov, S. T., *The Processes $\mu \to e + \gamma, \mu \to e + \overline{e}, \nu' \to \nu + \gamma$ in the Weinberg-Salam Model with Neutrino Mixing*, *Sov. J. Nucl. Phys.* **25** (1977) 340.

[54] Hernández-Tomé, G. and López Castro, G. and Roig, P., *Flavor violating leptonic decays of τ and μ leptons in the Standard Model with massive neutrinos*, *Eur. Phys. J. C* **79** (2019) 84, [1807.06050].

[55] MEG collaboration, Baldini, A. M. and others, *Search for the lepton flavour violating decay $\mu^+ \to e^+\gamma$ with the full dataset of the MEG experiment*, *Eur. Phys. J. C* **76** (2016) 434, [1605.05081].

[56] SINDRUM collaboration, Bellgardt, U. and others, *Search for the Decay mu+ —> e+ e+ e-*, *Nucl. Phys. B* **299** (1988) 1–6.

[57] Jodidio, A. and others, *Search for Right-Handed Currents in Muon Decay*, *Phys. Rev. D* **34** (1986) 1967.

[58] TWIST collaboration, Bayes, R. and others, *Search for two body muon decay signals*, *Phys. Rev. D* **91** (2015) 052020, [1409.0638].

[59] MEG II collaboration, Baldini, A. M. and others, *The design of the MEG II experiment*, *Eur. Phys. J. C* **78** (2018) 380, [1801.04688].

[60] Jho, Yongsoo and Knapen, Simon and Redigolo, Diego, *Lepton-flavor violating axions at MEG II*, *JHEP* **10** (2022) 029, [2203.11222].

[61] Mu3e collaboration, Perrevoort, Ann-Kathrin, *The Rare and Forbidden: Testing Physics Beyond the Standard Model with Mu3e*, *SciPost Phys. Proc.* **1** (2019) 052, [1812.00741].

[62] Calibbi, Lorenzo and Redigolo, Diego and Ziegler, Robert and Zupan, Jure, *Looking forward to lepton-flavor-violating ALPs*, *JHEP* **09** (2021) 173, [2006.04795].

[63] Haensel, P. and Potekhin, A. Y. and Yakovlev, D. G., *Neutron stars 1: Equation of state and structure*, vol. 326. Springer, New York, USA, 2007, 10.1007/978-0-387-47301-7.

[64] Linster, Matthias and Ziegler, Robert, *A Realistic $U(2)$ Model of Flavor*, *JHEP* **08** (2018) 058, [1805.07341].

[65] Froggatt, C. D. and Nielsen, Holger Bech, *Hierarchy of Quark Masses, Cabibbo Angles and CP Violation*, *Nucl. Phys. B* **147** (1979) 277–298.

[66] Leurer, Miriam and Nir, Yosef and Seiberg, Nathan, *Mass matrix models*, *Nucl. Phys. B* **398** (1993) 319–342, [hep-ph/9212278].

[67] Leurer, Miriam and Nir, Yosef and Seiberg, Nathan, *Mass matrix models: The Sequel*, *Nucl. Phys. B* **420** (1994) 468–504, [hep-ph/9310320].

[68] Chikashige, Y. and Mohapatra, Rabindra N. and Peccei, R. D., *Are There Real Goldstone Bosons Associated with Broken Lepton Number?*, *Phys. Lett. B* **98** (1981) 265–268.

[69] Schechter, J. and Valle, J. W. F., *Neutrino Decay and Spontaneous Violation of Lepton Number*, *Phys. Rev. D* **25** (1982) 774.

[70] Björkeroth, Fredrik and Chun, Eung Jin and King, Stephen F., *Flavourful Axion Phenomenology*, *JHEP* **08** (2018) 117, [1806.00660].

[71] Derenzo, S. E., *Measurement of the low-energy end of the mu-plus decay spectrum*, *Phys. Rev.* **181** (1969) 1854–1866.

[72] Bryman, D. A. and Clifford, E. T. H., *EXOTIC MUON DECAY mu —> e x*, *Phys. Rev. Lett.* **57** (1986) 2787.

[73] Bilger, R. and Foehl, Klaus and Clement, H. and Croni, M. and Erhardt, A. and Meier, R. and Patzold, J. and Wagner, G. J., *Search for exotic muon decays*, *Phys. Lett. B* **446** (1999) 363–367, [hep-ph/9811333].

[74] PIENU collaboration, Aguilar-Arevalo, A. and others, *Improved search for two body muon decay $\mu^+ \to e^+ X_H$*, *Phys. Rev. D* **101** (2020) 052014, [2002.09170].

[75] D'Eramo, Francesco and Yun, Seokhoon, *Flavor violating axions in the early Universe*, *Phys. Rev. D* **105** (2022) 075002, [2111.12108].

[76] ARGUS collaboration, Albrecht, H. and others, *A Search for lepton flavor violating decays tau —-> e alpha, tau —> mu alpha*, *Z. Phys. C* **68** (1995) 25–28.

[77] Belle-II collaboration, Altmannshofer, W. and others, *The Belle II Physics Book*, *PTEP* **2019** (2019) 123C01, [1808.10567].

[78] CMB-S4 collaboration, Abazajian, Kevork N. and others, *CMB-S4 Science Book, First Edition*, 1610.02743.

[79] Abazajian, Kevork and others, *CMB-S4 Science Case, Reference Design, and Project Plan*, 1907.04473.

[80] Li, Bao-An and Cai, Bao-Jun and Chen, Lie-Wen and Xu, Jun, *Nucleon Effective Masses in Neutron-Rich Matter*, *Prog. Part. Nucl. Phys.* **99** (2018) 29–119, [1801.01213].

[81] Yakovlev, Dima G. and Pethick, C. J., *Neutron star cooling*, *Ann. Rev. Astron. Astrophys.* **42** (2004) 169–210, [astro-ph/0402143].

[82] Potekhin, A. Y. and Pons, J. A. and Page, Dany, *Neutron stars - cooling and transport*, *Space Sci. Rev.* **191** (2015) 239–291, [1507.06186].

[83] Horowitz, C. J. and Piekarewicz, J. and Reed, Brendan, *Insights into nuclear saturation density from parity violating electron scattering*, *Phys. Rev. C* **102** (2020) 044321, [2007.07117].

[84] Friman, B. L. and Maxwell, O. V., *Neutron Star Neutrino Emissivities*, *Astrophys. J.* **232** (1979) 541–557.

[85] Lattimer, J. M. and Prakash, M. and Pethick, C. J. and Haensel, P., *Direct URCA process in neutron stars*, *Phys. Rev. Lett.* **66** (1991) 2701–2704.

[86] Leinson, L. B., *Superfluid response and the neutrino emissivity of baryon matter. Fermi-liquid effects*, *Phys. Rev. C* **79** (2009) 045502, [0904.0320].

[87] Leinson, L. B., *Neutrino emission from triplet pairing of neutrons in neutron stars*, *Phys. Rev. C* **81** (2010) 025501, [0912.2164].

[88] Voskresensky, D. N. and Senatorov, A. V., *Emission of Neutrinos by Neutron Stars*, *Sov. Phys. JETP* **63** (1986) 885–897.

[89] Grigorian, H. and Voskresensky, D. N. and Blaschke, D., *Influence of the stiffness of the equation of state and in-medium effects on the cooling of compact stars*, *Eur. Phys. J. A* **52** (2016) 67, [1603.02634].

[90] Iwamoto, N., *Axion Emission from Neutron Stars*, *Phys. Rev. Lett.* **53** (1984) 1198–1201.

[91] Brinkmann, Ralf Peter and Turner, Michael S., *Numerical Rates for Nucleon-Nucleon Axion Bremsstrahlung*, *Phys. Rev. D* **38** (1988) 2338.

[92] Pethick, C. J., *Cooling of neutron stars*, *Rev. Mod. Phys.* **64** (1992) 1133–1140.

[93] Yakovlev, D. G. and Kaminker, A. D. and Gnedin, Oleg Y. and Haensel, P., *Neutrino emission from neutron stars*, *Phys. Rept.* **354** (2001) 1, [astro-ph/0012122].

[94] Iwamoto, Naoki, *Nucleon-nucleon bremsstrahlung of axions and pseudoscalar particles from neutron star matter*, *Phys. Rev. D* **64** (2001) 043002.

[95] Harris, Steven P. and Fortin, Jean-Francois and Sinha, Kuver and Alford, Mark G., *Axions in neutron star mergers*, *JCAP* **07** (2020) 023, [2003.09768].

[96] Baym, Gordon and Pethick, Christopher, *Landau Fermi-liquid theory: concepts and applications*. John Wiley & Sons, 2008.

[97] Raffelt, Georg and Seckel, David, *Bounds on Exotic Particle Interactions from SN 1987a*, *Phys. Rev. Lett.* **60** (1988) 1793.

[98] Shtabovenko, Vladyslav and Mertig, Rolf and Orellana, Frederik, *FeynCalc 9.3: New features and improvements*, *Comput. Phys. Commun.* **256** (2020) 107478, [2001.04407].

[99] Cohen, Jeffrey M and Langer, William D and Rosen, Leonard C and Cameron, AGW, *Neutron star models based on an improved equation of state*, *Astrophysics and Space Science* **6** (1970) 228–239.

[100] Lepage, G. Peter, *Adaptive multidimensional integration: VEGAS enhanced*, *J. Comput. Phys.* **439** (2021) 110386, [2009.05112].

[101] Raffelt, G. G., *Stars as laboratories for fundamental physics: The astrophysics of neutrinos, axions, and other weakly interacting particles*. 5, 1996.

Chapter 6

Conclusions

Through the work presented in this thesis I have developed and implemented pipelines that placed limits on axion parameters using measurements of CMB birefringence and neutron star cooling. In chapters 2-4 I considered the possibility that long-lived cosmic strings in ALP fields could impart a rotation of the plane of polarization of CMB photons, an effect known as CMB birefringence.

In chapter 2 I obtained limits on axion parameters in a phenomenological model called the loop-crossing model (LCM) by performing Bayesian inference on published CMB birefringence power spectrum data. I found that the birefringence power spectrum is constrained by current probes to be less $\mathcal{A}^2\xi_0 < 3.7$ (95% CL). This limit is intriguing because, for reasonable values of the parameters ($\mathcal{A} \sim \mathcal{O}(1)$, $\xi_0 = 0.3-30$), the amplitude can be around this value. With future CMB probes expected to improve over current polarization measurements by orders of magnitude, we have the exciting prospect that next generation CMB data can access a compelling parameter in cosmic axion string models.

In chapter 3 I explored how non-Gaussian statistics can be used to resolve a limitation of power spectrum-based analyses. Specifically, the power spectrum is directly proportional to the combination of parameters $\mathcal{A}^2\xi_0$, which means that the power spectrum does not contain information that could allow one to independently measure $\mathcal{A}$ and ξ_0. I showed that axion-string induced birefringence maps exhibit positive kurtosis in their spherical harmonic coefficients and derived an analytic formula which successfully describes the qualitative feature of the excess kurtosis as a function of multipole. One of the takeaways from this project was that kurtosis poses a challeng-

ing way to look for a signal of axion strings in real CMB birefringence maps. Therefore it is advantageous to look for statistics with more concentrated information.

Pursuing this direction of inquiry further, in chapter 4 I studied how convolutional neural networks could be trained to extract information about LCM parameters. In particular, I trained three spherical convolutional neural networks (SCNNs) to learn the combination of parameters $Z = \log_{10}\zeta_0$, $A = \log_{10}(\mathcal{A}^2\xi_0)$, and $X = \log_{10}(\xi_0^2/\mathcal{A})$. Based on several tests of their accuracy and precision, we conclude that the SCNNs that learned $\log_{10}(\mathcal{A}^2\xi_0)$ and $\log_{10}(\xi_0^2/\mathcal{A})$ are able to furnish reliable estimators of these parameters when provided with noiseless birefringence maps generated from LCM simulations.

In chapter 5 I studied how neutron star cooling curves can be used to look for evidence of axions by considering a theory where lepton-flavour violating LFV interactions enable electrons colliding with a nucleon to produce a muon and an axion. The emitted axions would carry energy away from the neutron star, slightly increasing the cooling rate. The absence of an anomalously fast cooling rate in neutron star observations allowed us to put a limit $|g_{ae\mu}| \lesssim 4\times10^{-11}$ on the LFV coupling constant $g_{ae\mu}$.

To conclude, there are several avenues for advancing the study of cosmic axion strings and their imprint on CMB birefringence. A straightforward next step is to extend the loop-crossing model (LCM) to incorporate additional features revealed by full 3D cosmic string simulations, such as the non-circular geometry of loops, the distribution of loop radii at a given redshift, and other nuances that are currently simplified. This would make the simulations more realistic and provide a more robust foundation for inference.

A critical step toward improving inference pipelines is addressing the limitations of neural networks trained on LCM-based simulations. While these networks effectively infer parameters as point estimators, there is a concern that they may have learned features specific to the LCM, such as the circular geometry of strings. To

mitigate this, neural networks could be trained on simulations that account for perturbations in string geometry or other features from 3D simulations, ensuring that the inferred parameters are robust to these complexities. Additionally, instead of directly inferring parameters, future pipelines could train networks to learn summary statistics that maximize the Fisher information about the parameters, which would allow for unbiased point estimators or simulation-based inference using these compressed datasets.

Ultimately, a key way to unify these efforts would be through full ray-tracing simulations of CMB photons propagating through realistic cosmic string networks. These simulations would capture a wide range of features, including birefringence, lensing, and non-trivial string geometries, providing training data that bridges the gap between simplified models and observational data. By integrating these improvements, we can build more robust inference pipelines, paving the way for reliable analysis of next-generation CMB data and maximizing our potential to discover axion strings in the cosmos.