Physics potential of the International Axion Observatory (IAXO)

E. Armengaud, a D. Attie, a S. Basso, b P. Brun, a N. Bykovskiy, c
J. M. Carmona, d J. F. Castel, d S. Cebrián, d M. Cicoli, e,f
M. Civitani, b C. Cogollos, g J. P. Conlon, h D. Costa, g T. Dafni, d
R. Daido, i A.V. Derbin, j M. A. Descalle, k K. Desch, l
I.S. Dratchnev, j B. Döbrich, c A. Dudarev, c E. Ferrer-Ribas, a
I. Fleck, m J. Galán, d G. Galanti, b L. Garrido, g D. Gascon, g
L. Gastaldo, n C. Germani, g G. Ghisellini, b M. Giannotti, o
I. Giomataris, a S. Gninenko, p N. Golubev, p R. Graciani, g
I. G. Irastorza, d,1 K. Jakovčić, q J. Kaminski, l M. Krčmar, q
C. Krieger, b B. Lakić, q T. Lasserre, a P. Laurent, a O. Limousin, a
A. Lindner, n I. Lomskaya, j B. Lubsandorzhiev, p G. Luzón, d
M. C. D. Marsh, s C. Mergalejo, d F. Mescia, g M. Meyer, l
J. Miralda-Escudé, g,u H. Mirallas, v V. N. Muratova, j X. F. Navick, a
C. Nones, a A. Notari, a A. Nozik, p,v A. Ortiz de Solórzano, d
V. Pantuev, p T. Papaevangelou, a G. Pareschi, b K. Perez, w
E. Picatoste, g M. J. Pivovaroff, k J. Redondo, d,x A. Ringwald, r
M. Roncadelli, k E. Ruiz-Chóli, d J. Ruz, k K. Saiwaka, a J. Salvador, g
M. P. Samperiz, d T. Schiffer, l S. Schmidt, l U. Schneekloth, r
M. Schott, z H. Silva, c G. Tagliaferri, b F. Takahashi, j,aa
F. Tavecchio, b H. ten Kate, c I. Tkachev, p S. Troitsky, p
E. Unzhakov, j P. Vedrine, a J. K. Vogel, k C. Weinsheimer, z
A. Weltman, ab W. Yin, ac,ad

aIRFU, CEA, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France
bINAF - Osservatorio astronomico di Brera, Via E. Bianchi 46, Merate (LC), I-23807, Italy
cEuropean Laboratory for Particle Physics (CERN), Geneva, Switzerland
dGrupo de Física Nuclear y Astropartículas, Departamento de Física Teórica,
Universidad de Zaragoza C/ P. Cerbuna 12 50009, Zaragoza, Spain
eDipartimento di Física e Astronomia, Università di Bologna,
via Irnerio 46, 40126 Bologna, Italy
fINFN, Sezione di Bologna, viale Berti Pichat 6/2, 40127 Bologna, Italy
1Corresponding author
Abstract. We review the physics potential of a next generation search for solar axions: the International Axion Observatory (IAXO). Endowed with a sensitivity to discover axion-like particles (ALPs) with a coupling to photons as small as \( g_{a\gamma} \sim 10^{-12} \text{ GeV}^{-1} \), or to electrons \( g_{ae} \sim 10^{-13} \), IAXO has the potential to find the QCD axion in the 1 meV \text{–} 1 eV mass range where it solves the strong CP problem, can account for the cold dark matter of the Universe and be responsible for the anomalous cooling observed in a number of stellar systems. At the same time, IAXO will have enough sensitivity to detect lower mass axions invoked to explain: 1) the origin of the anomalous “transparency” of the Universe to gamma-rays, 2) the observed soft X-ray excess from galaxy clusters or 3) some inflationary models. In addition, we review
string theory axions with parameters accessible by IAXO and discuss their potential role in cosmology as Dark Matter and Dark Radiation as well as their connections to the above mentioned conundrums.
1 Introduction

Axions and axion-like particles (ALPs), as well as other more generic categories of particles (weakly interacting sub-eV particles, WISPs) at the low-mass frontier [1–3] are acquiring a strong interest as a portal for new physics, candidates to the dark Universe, or as solutions of poorly understood astrophysical phenomena. The detection of these particles in terrestrial experiments is currently pursued by a number of experimental techniques, see [4, 5] for recent reviews. In this experimental landscape, the International Axion Observatory (IAXO) [6, 7] stands out as one of the most ambitious projects under consideration. It is our purpose here to review the theoretical, cosmological and astrophysical motivation to carry out the primary goal of IAXO, i.e. the search for solar axions with sensitivity much beyond previous similar searches and well into unexplored parameter space. We will focus on more recent developments affecting regions of parameter space at reach of IAXO, as well as those highlighting its novelty and complementarity within the larger set of axion experimental efforts.

The QCD axion is a hypothetical 0− particle predicted in the Peccei-Quinn mechanism to solve the strong CP problem [8] of the standard model (SM) of particle physics. In the pure SM, the parity (P) and time-reversal (T) violation effects observed so far can be attributed to the phase of the CKM matrix, which is relatively large \( \delta \sim 30^\circ \) and has its origin in the Yukawa couplings of the Higgs to fermionic fields. The problem is that theory predicts the existence of another P,T violating phase, \( \bar{\theta} \), which appears in the Lagrangian density multiplying the topological-charge density of QCD,

\[
L_{\bar{\theta}} = -\frac{\alpha_s}{8\pi} G_{\mu\nu}^a \tilde{G}^{a\mu\nu} \bar{\theta}. \tag{1.1}
\]

This term produces CP violating effects like electric-dipole moments (EDMs) for hadrons, which have never been observed. The strongest upper limit on strong CP violation comes from the neutron EDM, \( |d_n| < 3.0 \times 10^{-13} \) e fm [9]. Given the calculation \( d_n = (2.4 \pm 1.0) \bar{\theta} \times 10^{-3} \) e fm [10], one finds an extremely strong upper bound,

\[
|\bar{\theta}| < 1.3 \times 10^{-10}. \tag{1.2}
\]

Indeed, \( \bar{\theta} \) arises in the SM as the sum of two contributions: the \( \theta \)-angle defining a gauge-invariant QCD vacuum and a common phase of the quark-mass matrix. The latter has an origin similar to the CKM phase and the former has no clear a-priori relation with them so it is extremely suspicious that these two will cancel so precisely as (1.2).

The solution proposed by Peccei and Quinn [11, 12] to alleviate the “strong CP issue” is based on the observation that the QCD vacuum energy, \( V_{\text{QCD}}(\bar{\theta}) \), has an absolute minimum\(^1\) at \( \bar{\theta} = 0 \). They proposed the existence of an extra global U(1) symmetry, spontaneously broken and color anomalous. Due to the color anomaly, the concomitant Goldstone-boson field, dubbed “axion” [13] by Wilczek and denoted here by \( A \), develops in the effective low-energy theory an anomalous coupling to gluons,

\[
L_{Ag} = -\frac{\alpha_s}{8\pi} G_{\mu\nu}^a \tilde{G}^{a\mu\nu} \frac{A}{f_A}, \tag{1.3}
\]

where \( f_A \) is the so-called axion decay constant, a new energy scale related to the scale of PQ symmetry breaking. Effectively, \( \bar{\theta} \) becomes replaced by \( \bar{\theta} + \langle A \rangle / f_A \), where \( \langle A \rangle \) is the

\(^1\)In a strict sense this is true if \( \bar{\theta} \) is the only source of CP violation in the SM. One expects a tiny shift of the minimum due to the non-zero CKM and from any other source of CP violation that can be communicated radiatively to the gluonic sector.
vacuum-expectation-value (VEV) of the axion. Since the axion has no other potential energy terms in the Lagrangian (as long as PQ is a symmetry at the classical level) it will adjust itself to minimise $V_{QCD}(\bar{\theta} + A/f_A)$ by taking the VEV $\langle A \rangle = -\theta f_A$, which cancels all the CP violating effects of $\bar{\theta}$! Note that the axion field effectively becomes a dynamical version (it is a space-time dependent field) of the $\bar{\theta}$ angle, which is a mere constant.

The QCD potential gives the axion field a mass,

$$m_A = \frac{\sqrt{\chi}}{f_A} \approx 5.7 \text{ meV} \frac{10^9 \text{ GeV}}{f_A},$$

(1.4)

(where $\chi = \partial^2 V_{QCD}/\partial \bar{\theta}^2 |_{\bar{\theta} = 0}$ is the topological susceptibility of QCD) and induces mixing of the QCD axion field with the $\eta'$ and the rest of the $0^-$ mesons. By virtue of this mixing, the QCD axion develops model-independently couplings to hadrons and, most importantly, a coupling to two photons,

$$L_{A\gamma} = -\frac{\alpha C_{A\gamma}}{8\pi f_A} F_{\mu\nu} \tilde{F}^{\mu\nu} A = -\frac{g_{A\gamma}}{4} F_{\mu\nu} \tilde{F}^{\mu\nu} A = g_{A\gamma} \vec{E} \cdot \vec{B} A.$$  

(1.5)

Here $F_{\mu\nu}$ is the electromagnetic field-strength, $\tilde{F}^{\mu\nu}$ its dual and $\vec{E}, \vec{B}$ are the electric and magnetic fields, respectively. The coupling constant $g_{A\gamma}$ has units of an inverse energy scale, which is $1/f_A$ except for an electromagnetic loop factor required for photon emission, $\alpha/2\pi$, and the mixing coefficient $C_{A\gamma} = -1.92$ [14]. The parameter $C_{A\gamma}$ can receive additional model-dependent additive contributions. These are usually of the order of 1 in simple models [15–18] but can be much larger in some engineered cases [19, 20]. Note that both the mass (1.4) and couplings, like (1.5), are inversely proportional to $f_A$. Ratios of coupling/mass are independent of $f_A$ and only depend on QCD physics and a few hopefully O(1) model-dependent parameters so the properties of the QCD axion are reasonably constrained.

The PQ solution to the strong CP problem is a minimal one, and is appealing for two reasons. The first is that the predicted axion particle can be experimentally confirmed or rejected (as opposed to some alternatives of the strong CP problem, e.g. [10]). So far QCD axions with $f_a \gtrsim 10^6$ GeV have already been robustly excluded by solar axion searches, stellar evolution, cosmology and laboratory experiments. The second is that the QCD axion with a large decay constant $f_a \gg 10^6$ GeV is an excellent cold dark matter (CDM) candidate. The combination of the strong CP problem and the resounding evidence of CDM in the Universe motivate very strongly the search for a low mass QCD axion.

Indeed, the popularity of the QCD axion has made generations of theorists to employ the word “axion” for other, equally-hypothetical, particles that are not related to QCD, the strong CP problem or the PQ mechanism simply because they have certain similarities to the QCD axion. The similarities often exploited are: being a low-mass 0$^-$ particle, a pseudo-Goldstone boson of an “axial” symmetry, featuring anomalous couplings to topological-charge densities like (1.3) for general non-abelian gauge-bosons, featuring anomalous couplings to two-photons, being low-mass and couple to two photons, to name a few. We will give more details below, but here we just note that the abuse is so overwhelming that we must bow to it and introduce here a disclaimer about nomenclature. Henceforth we will denote the true QCD “axion” as “QCD axion” or “the axion”, particles with similar properties as axion-like particles (ALPs) and the whole family as simply “axions”. The QCD axion will be $A$ and a generic ALP will be denoted as $a$. Couplings like (1.5) will be generic for ALPs but the mass relation is exclusive of the QCD axion. It is extraordinary that axions appear so profusely
in extensions of the SM. They are often encountered as pseudo-Nambu-Goldstone bosons of new global symmetries (spontaneously broken at a high energy scales) and in holomorphic theories of dynamical couplings, string theory being a prime example [21–24]. Because of the similar properties of the QCD axion and ALPs, axion experiments have the potential to discover more than the solution of the strong CP problem and the dark matter of the Universe. An example is the search for solar axions and the proposed International Axion Observatory (IAXO), the focus of this monograph.

The two-photon coupling allows the production of axions in the collision of photons and charged particles via the Primakoff-effect \( \gamma + q \rightarrow a + q \). The corresponding solar axion flux can be calculated to be \( 1.6 \times 10^{21} C_{a\gamma}^2 (10^8 \text{GeV}/f_a)^2 / (\text{m}^2 \text{year}) \) with a mean axion energy \( \langle E_a \rangle = 4.2 \text{ keV} \). The axion coupling to electrons can be responsible for a copious flux \( \sim 2 \times 10^{24} C_{ae}^2 (10^8 \text{GeV}/f_a)^2 / (\text{m}^2 \text{year}) \) with \( \langle E_a \rangle = 2.3 \text{ keV} \) through the ABC processes [25]. Note that although \( C_{ae} \) is of order unity in some important models like DFSZ [26, 27] or general Grand Unification Theory axions (GUTs) [28], it is absent at tree-level in other models, and the loop-induced contribution is very small. Fortunately, one can always resort to the former Primakoff flux. The solar axion flux is indeed copious even for very small values of \( f_a \) because the Sun is huge, but the chances of detecting solar axions on Earth-based experiments are severely hampered by a small detection probability. Naive estimates of the natural value of detection cross sections of order \( \sigma \sim g_{a\gamma}^2 \sim 5 \times 10^{-54} C_{a\gamma}^2 (10^8 \text{GeV}/f_a)^2 \) were initially discouraging. However, P. Sikivie presented in 1982 an experimental concept that uses the low mass of axions to boost the detection probability by making it coherent over macroscopic magnetic fields, the axion “helioscope” [29, 30].

Axions travelling through a transversely polarised \( B \)-field can convert into photons at any point along the magnetic region. The photon polarisation is aligned to the external \( B \)-field due to the \( \vec{E} \cdot \vec{B} \) coupling. The conversion probability can be understood as axion-photon oscillations [31] with oscillation length \( \lambda_a = 8\pi E / \sqrt{m_a^4 + (2g_{a\gamma}BE)^2} \) (\( E \) is the axion energy), but also as the square of the coherent sum of the conversion amplitudes at each point along the line of sight [32]. After a homogeneous magnetic length \( L \), the probability of conversion is

\[
P(a \rightarrow \gamma) = \frac{(2g_{a\gamma}BE)^2}{m_a^4 + (2g_{a\gamma}BE)^2} \sin^2 \frac{\sqrt{m_a^4 + (2g_{a\gamma}BE)^2} L}{4E},
\]

which is coherent (\( \propto L^2 \)) if the magnetic length is smaller than the oscillation length \( L \lesssim \lambda_a \).

For small \( m_a \) this can be very large and so can be the enhancement! For instance, for the parameters proposed for IAXO, this conversion probability is

\[
P(a \rightarrow \gamma) \simeq 10^{-19} C_{a\gamma}^2 \left( \frac{10^8 \text{GeV}}{f_a} \right)^2 \left( \frac{B}{3 \text{T}} \right)^2 \left( \frac{L}{20 \text{ m}} \right)^2,
\]

and stays in the coherent regime for \( m_a \lesssim 16 \text{ meV} \), making the IAXO search for solar axions realistic. Insisting on an incoherent detection scheme, one would need hundreds of kilometres of low-background instrumented detectors to obtain a similar conversion probability with \( \sigma \sim g_{a\gamma}^2 \). This is far too long compared to the most ambitious direct DM experiments looking for WIMP recoils, which are also parasitically used to search for solar axions. Their virtue is that their detection does not rely on macroscopic coherence and therefore they are relatively insensitive to the axion mass. Unfortunately, they will not be in the foreseeable future as sensitive as indirect stellar evolution constraints, or as current helioscopes, see e.g. [33, 34].
For helioscopes, the smallness of the axion mass is a virtue, as the coherent length can be macroscopic and the use of long and intense magnets can greatly enhance the axion detection probability with respect to incoherent detection. This so-called coherent inverse Primakoff-process is at the core of the power of the Helioscope technique to find solar axions with IAXO but it can lead to other very interesting phenomena.

Indeed, the most mature experimental efforts to detect axions in lab experiments [4, 5] rely on their conversion to photons in macroscopic magnetic fields [29, 31]. This approach includes the search for galactic axion dark matter [29] and the photon regeneration experiments (“shining light through a wall”) [30, 32, 35]. There is a large complementarity between the different strategies. While relic axion searches enjoy high sensitivity in terms of $g_{a\gamma}$, they typically accomplish this only for a narrow $m_a$ range and rely on the assumption that the DM is mostly composed of axions. Laboratory experiments are free from cosmological or astrophysical assumptions on the production of ALP, but they have a comparatively reduced sensitivity. Helioscopes enjoy an attractive compromise between axion source dependency (the solar flux is a robust prediction from any axion model) and competitive sensitivity, as will be shown. Indeed, a next generation helioscope is the only demonstrated technique that can discover QCD axions in the meV mass range [36]. The CERN Axion Solar Telescope (CAST) [37–40], in operation at CERN for more than a decade, represents the current state-of-the-art of the axion helioscope technique. The International Axion Observatory (IAXO) has been recently proposed as a follow-up of CAST, scaling the helioscope concept to the largest size realistically allowed. IAXO will be largely based on experience and concepts developed by CAST.

In this paper we provide an updated description of the physics case of IAXO. We organise our paper by reviewing the contexts in which the discovery of the QCD axion or an ALP in the parameter space accessible by IAXO could have strong implications in other contexts of theoretical particle physics, cosmology and astrophysics. We set the stage in section 2 with a detailed theoretical background for axions in extensions of the SM, reviewing recent advances in axion model building and axion couplings, and placing particular emphasis on stringy axions.

Section 3 is devoted to axions as dark matter. Indeed, a great deal of the appeal of axions is that they can be produced in the early Universe via non-thermal processes (realignment mechanism and decays of cosmic strings and domain walls [41, 42]), which makes them well motivated cold dark matter candidates. The relic density depends on $f_a, m_a$ but also on the initial conditions and the cosmological history until today. For the pure QCD axion, one can distinguish between two broad classes:

- If the PQ symmetry spontaneously breaks before inflation and is not restored afterwards, the axion field becomes homogeneous during inflation at some a-priori-unknown VEV. A broad range of axion masses can produce the correct relic density just adjusting to the adequate initial conditions, including the $m_A \sim$ meV values that could be discovered with IAXO.

- If the symmetry is restored after inflation, initial conditions are reset and the QCD axion becomes coherent only at very small scales compared to today’s horizon. An average over disconnected patches removes the uncertainty of initial conditions but the quantity of axions radiated from cosmic strings is still uncertain and those radiated from domain wall (DW) decay is model dependent. Models with short-lived DWs ($N_{DW} = 1$) can give the totality of the DM for $m_A \gtrsim 26 \mu$eV, the uncertainty in the
calculation encompassing the \( m_A \sim \text{meV} \) values accessible by IAXO. In models with long-lived DWs, the DM yield can be greatly increased, which points to \( m_A \sim \text{meV} \) as the most favoured mass range to account for all the DM in the Universe.

We discuss the status of the meV QCD axion as a DM candidate in detail in section 3. General ALP models have ample freedom to become DM candidates in a broad range of values of \( f_a \) and \( m_a \) [43].

Axions can be also produced from thermal processes or decays of heavy particles, contributing to the radiation energy density as effective relativistic degrees of freedom in the early Universe and/or today. Current cosmological observations present some tension that could suggest the presence of this Dark Radiation, in addition to the known neutrino species. For some values of the relevant parameters, this axionic Dark Radiation could give observable signals in helioscopes like IAXO. This scenario is described in section 4.

The fact that ALPs have masses protected by large radiative corrections is very suggestive to use them as candidates for the inflaton. A recent scenario showed that, for an adequate potential and coupling to photons (for reheating purposes), an axion accessible to IAXO might indeed be responsible for cosmic inflation. This interplay between ALPs and inflation is reviewed in section 5.

The existence of axions can have very important consequences in astrophysics. Indeed, the properties of well-known stellar systems have been providing the strongest constraints on axion properties for a long time [44]. More intriguing axion effects may account for unexplained astrophysical observations. Two of these cases deserve special consideration: the anomalous cooling rate observed in a number of stellar systems and the excessive transparency of the intergalactic medium to very high energy (VHE) photons. In both contexts, the existence of very light axions with properties at reach of IAXO has been repeatedly invoked as an explanation, although claims will be taken with caution, as more conservative explanations cannot be excluded. The observational situation as well as the potential interpretation in terms or axions are reviewed for both cases in sections 6 and 7, respectively.

To conclude the physics case update, we briefly review the conceptual design of IAXO [7] in section 8, present updated sensitivity projections of the experiment and conclude in section 9.

2 Axions from high-energy physics

2.1 Stringy axions

General Relativity (GR) is a perturbatively unitary theory only up to the Planck scale. Whether or not GR is non-perturbatively unitary is still an open question but the requirement of a Wilsonian (perturbative) unitarization of GR has a very promising venue in string theory where point particles are replaced by extended objects: strings. For theoretical consistency, string theory requires additional spatial dimensions beyond the known three. Experiments and observations constrain these additional dimensions to be compact and very small. The compactification schemes determine many of the properties of the low-energy four-dimensional effective description.

The field content in the higher dimensional theory includes many anti-symmetric tensor fields (\( p \)-forms) which may be integrated over any non-trivial cycle in the compactification geometry. In four-dimensions, the values of such integrals appear as dynamical scalar fields, and the gauge symmetry of the \( p \)-forms translates into a global shift symmetry making
them candidate axions [21–24]. These are commonly known as ‘closed string’ axions. In addition, D-branes (places where open strings can end) that fill four-dimensional spacetime can support additional ‘open string’ axions, which can realise the PQ-mechanism [45] or appear as independent ALPs.

In some string models, the mass spectrum and couplings of these scalars to ordinary matter can be explicitly computed [46, 47]. This serves as the starting point for several interesting applications to cosmology, astrophysics and particle physics. In particular, one linear combination of these scalars will play the rôle of the standard QCD axion if coupled to the QCD sector realised on D-branes, i.e. if leading to the low-energy coupling (1.3). All other scalars are instead ALPs, which might behave very similarly to the QCD axion but whose mass and decay constant are not related by QCD scales but by some other detail of the model. Effectively, mass and couplings become independent parameters. Hence the corresponding parameter space is much wider, featuring for instance regions where ALPs can successfully: (i) drive inflation with the production of observable primordial gravity waves [48], (ii) account for the observed dark matter content of the Universe [43, 49], (iii) contribute to dark radiation by behaving as extra neutrino-like degrees of freedom [50, 51], or (iv) become the longitudinal component of extra $U(1)$ gauge bosons with mass well below the string scale which could belong to either the visible or a hidden sector [52–54].

It is worth mentioning that some of these potential axions might acquire large masses [47]. For instance, they could be simply excluded in the compactifications leading to the SM gauge group [55]. Moreover, they could be eaten up by anomalous $U(1)$’s with masses of order the string scale in the Green-Schwarz mechanism of anomaly cancellation [56]). Finally, they could acquire a very large mass if they are stabilised in a supersymmetric way since they would become as massive as the corresponding supersymmetric partners, the so-called saxions, which have to be heavier than $\sim 50$ TeV to decay before primordial nucleosynthesis and avoid altering the successful Big Bang predictions [57–59].

For each of the above discussed ‘closed string’ axions, there is generically also a ‘modulus’ scalar field that parametrises the size or shape of the compactification manifold. In contrast to the ‘closed string’ axions, moduli fields do not enjoy a shift symmetry, are not in general protected from large quantum corrections to their mass and therefore might be easily too heavy to be considered axions. However, the imaginary part $a$ of a moduli field $\Phi = \phi + i a$ is exactly shift symmetric and thus is a natural candidate for an axion.

The real part, the saxion $\phi$, parametrises either the size or the shape of the extra dimensions and comes from the dimensional reduction of the ten-dimensional metric. Hence the saxion is a gauge singlet with only gravitational couplings to ordinary matter. Moreover, as already mentioned, $\phi$ does not enjoy a shift symmetry, and so can be lifted and become massive by any kind of perturbative effect. Therefore it is crucial to study moduli stabilisation in order to work out the axion mass spectrum for phenomenological applications. There are two benchmark scenarios [22, 47]:

- The moduli are fixed by non-perturbative effects. Since the same mechanism gives mass to both $\phi$ and $a$, they will naturally tend to acquire masses of the same order of magnitude, which is constrained by the saxion phenomenology to $m_a \sim m_\phi \gtrsim 50$ TeV, too massive for our purposes.
- The saxion masses are fixed by perturbative effects. Axions will remain massless because they are protected by their perturbative shift symmetry, except for the non-perturbative effects which are now unconstrained by saxion phenomenology. In this
last case, the actual value of the axion mass is model-dependent. However, since the non-perturbative effects depend on exponentials of volume factors, the axions $a$ can be exponentially lighter than saxions and thus naturally very light.

Let us stress that fixing all the moduli via non-perturbative effects is highly non-generic because of various technical issues that can forbid non-perturbative effects [60]. Therefore we conclude that the low-energy four-dimensional limit of string compactifications generically gives rise to the presence of multiple light axions which can include the QCD axion plus one or more axions whose mass spectrum and couplings depend on the microscopic details of the physics and the geometry of the extra dimensions [47]. The axions which acquire small masses are of course of phenomenological interest.

The decay constants of stringy axions depend on their microscopic origin as closed or open string modes. In particular, the size of $f_a$ is determined by the geometry of the compactified dimensions. The closed string axion coupling $f_a$ can be either of order the compactification scale (the Kaluza-Klein scale $M_{KK}$) for strings associated with internal cycles in the extra dimensions bulk, or it can be of order the string scale $M_s$, for strings associated with resolutions of local singularities [22, 47].

Given that $M_s \sim g_s^{1/4} M_{Pl}/\sqrt{\mathcal{V}}$ and $M_{KK} \sim M_s/\sqrt{\mathcal{V}}^{1/6}$, where $M_{Pl} = 2.435 \cdot 10^{18}$ GeV denotes the (reduced) Planck mass, $g_s$ the string coupling, and $\mathcal{V}$ the internal volume in string units, compactifications at large volume and weak string coupling —both preferable properties of string models under perturbative control— predict $f_a \ll M_{Pl}$.

In models where all the moduli are safely heavier than 50 TeV, $M_s$ and $M_{KK}$ tend to be around the GUT scale which is therefore the natural value to expect for the decay constant of closed string axions: $f_{a}^{\text{closed}} \sim 10^{16}$ GeV. Notice also that in the closed string axions the effective PQ symmetry is always broken in the 4D effective field theory below the Kaluza-Klein scale. However, for sufficiently large $\mathcal{V}$, the decay constant of closed string axions can fall within the intermediate scale window [22] and thus the reach of IAXO. However, in these models with relatively low string and Kaluza-Klein scale, some moduli tend to get lighter than 50 TeV and cause cosmological problems.

Let us turn now to open string axions. They arise as the phases of complex fields charged under anomalous $U(1)$s that could live on a stack of D-branes [45]. In the process of anomaly cancellation, the $U(1)$ acquires a mass of order $M_s$ by eating up a closed string axion [61]. Therefore, the effective field theory below the string scale contains an effective global $U(1)$ PQ-like symmetry. The axion decay constant is then set by the magnitude of the radial part of the charged open string mode. In turn, in supersymmetric theories this radial part is set by a model-dependent Fayet-Iliopolous term. For models with D-branes at singularities where the SM sector is sequestered from the sources of supersymmetry breaking in the bulk [62, 63], the decay constant of open string axions can be suppressed with respect to $M_s$ due to either sequestering effects or a mixing of sequestering and $U(1)$ kinetic mixing, leading to an intermediate scale $f_{a}^{\text{open}}$. In some of these models the resulting decay constant can be as small as $f_{a}^{\text{open}} \sim 10^8$ GeV [47, 64], in an interesting region testable by IAXO.

In sum, axions are ubiquitous in string theory, and while most axions may be rendered heavy, some may remain light in the low-energy theory and contribute to a variety of physical processes that can be accessible to axion experiments like IAXO.

---

2A significantly lower $f_{a}^{\text{closed}}$ could be achieved if the cycle supporting the axion is located in a highly warped region of the compactification, where the effective action is however under less control.
2.2 The low-energy Lagrangian and field-theoretic models

The low-energy description of the properties and interactions of an axion is strongly determined by the requirement of possessing a global shift symmetry,

\[ a \rightarrow a + \epsilon f_a. \]  

(2.1)

If the symmetry is exact, the axion will be massless, as a mass term \( L_a \propto m_a^2 a^2 \) does not respect the shift symmetry. Low mass axions can exist if the symmetry is very weakly broken. There are two conceptually different ways of achieving this breaking. On the one hand, the shift symmetry can simply be explicitly broken by some term in the Lagrangian that is extremely small. In these cases there is a set of couplings or parameters \( \{ p \} \) which, when taken to zero, leave the Lagrangian density invariant under the shift symmetry. In this case, the axion mass can receive radiative corrections from the high-energy sector of the theory, hinting to a possible hierarchy problem. However, they will always be proportional to these coefficients. If those are very small, the theory can have low mass axions and still be technically natural. On the other hand, the shift symmetry might be violated only by anomalous couplings like (1.3). These soft-breaking terms can and do generate non-trivial potentials for axions but do not generate UV-sensitive radiative corrections to their mass. The resulting low mass axion is therefore naturally light. In practice it is often assumed that global symmetries are not respected by quantum gravity effects because sufficiently classical black-holes have no hair and thus can swallow arbitrary global charges. Therefore, unless the global symmetry is a remnant from a gauge symmetry or has some other means of protection, one assumes that quantum gravity will generate some kind of explicit symmetry breaking and thus a small potential for axions (and thus a mass). This kind of reasoning will play a role in Sec. 3.

In either case, the Lagrangian of a low-mass axion can be divided into a shift-symmetric and a shift-breaking part. The second must be subdominant to consider our axion naturally light. We assume that our axion is a periodic angular field defined in \( a \in (-\pi, \pi) v_a \) where \( v_a \) is a scale related to the spontaneous symmetry breaking of the shift symmetry. Below \( v_a \) and the electroweak scales, but above the QCD confining scale the Lagrangian of an axion can be written as

\[
L_a = \frac{1}{2} (\partial_\mu a) (\partial^\mu a) + \sum_{ij} \frac{c_{aij}}{2v_a} (\bar{\psi}_i \gamma^\mu \gamma^5 \psi_j) \partial_\mu a - \frac{\mathcal{E}_\alpha}{2\pi v_a} \frac{F_{\mu\nu} \tilde{F}^{\mu\nu}}{4} a - \frac{\mathcal{N}_\alpha}{2\pi v_a} \frac{\tilde{G}_{\mu\nu} \tilde{G}^{\mu\nu}}{4} a + \ldots
\]  

(2.2)

where \( c_{aij}'s \) are dimensionless coupling coefficients, \( \psi_i \) are the SM quarks and leptons, \( \mathcal{E} \) and \( \mathcal{N} \) are the electromagnetic and colour anomalies of the shift symmetry and the ellipsis stands for additional explicit symmetry breaking terms. If \( \mathcal{N} \neq 0 \) and additional explicit breaking terms are sufficiently small, the axion becomes the QCD axion. If we compare with (1.3), we find that we should define \( f_A = v_A/\mathcal{N} \). In this case, it is important to note that \( V_{\text{QCD}}(\theta) \) is 2\(\pi\) periodic in \( \theta \) and so it will in \( A/f_A \). Therefore, there are \( \mathcal{N} \) physically different values of the QCD axion VEV that minimise the QCD potential and cancel CP violation, \( \theta - \langle A \rangle / f_A = 0, 1, 2, \ldots N - 1 \). Once again, this will be crucial for our section on axion CDM, Sec. 3.

Below the QCD confining scale, \( \Lambda_{\text{QCD}} \), quarks and gluons are no longer adequate degrees of freedom and any axion coupling to \( GG \) will mix with the neutral mesons and acquire a
non-trivial potential. We define the low-energy Lagrangian for a generic axion as

$$L_a = \frac{1}{2} (\partial_\mu a)(\partial^\mu a) + \sum_{ij} \frac{C_{aij}}{2f_a} (\bar{\psi}_i \gamma^\mu \gamma^5 \psi_j) \partial_\mu a - \frac{C_{a\gamma} a}{2\pi f_a} F_{\mu\nu} \tilde{F}^{\mu\nu} a$$

(2.4)

where the ellipsis stands now for shift-symmetry couplings to mesons and explicit symmetry breaking terms. We also define the phenomenological coupling constants

$$g_{aff'} = C_{aff'} \frac{m_f + m_{f'}}{2f_a}, \quad g_{a\gamma} = \frac{\alpha}{2\pi f_a} C_{a\gamma},$$

(2.5)

where $m_f$ are SM fermion masses.

The distinction between $f_a$ and $v_a$ is only relevant for the QCD axion in our context so we will use $f_a = v_a$ for ALPs (we remove $V_{QCD}$ but we allow for other sources of ALP mass in the ellipsis). The coupling constant $C_{a\gamma}$ has a model-dependent contribution $\propto E$ and a model-independent contribution from meson-mixing just for the QCD axion [14],

$$C_{A\gamma} = -1.92(3) + \frac{\xi}{N_c} \quad \text{(QCD axion)},$$

$$C_{a\gamma} = \frac{\xi}{N_c} \quad \text{(ALP)}.$$

(2.6)

The couplings to fermions are very similar. They are inherited from the high-energy coefficients $c_{aij}$ except for a model-independent part which arises for the QCD axion coupling to hadrons. The relevant couplings for the low-energy phenomenology discussed here are to protons, neutrons and electrons. The case of an ALP is quite unconstrained so we quote here those of the QCD axion [14],

$$C_{Ap} = -0.47(3) + \Delta C_{App}, \quad C_{An} = -0.02(3) + \Delta C_{Ann}$$

(2.7)

$$C_{Ae} = \frac{\alpha^2 C_{A\gamma}}{3\pi} \log \left( \frac{\Lambda_{QCD}}{m_e} \right) + c_{aee}$$

(2.8)

where $\Delta C_{App}, \Delta C_{Ann}$ are the model-dependent parts that arise from the model-dependent QCD axion couplings to SM quarks. The model-dependent axion-electron coupling is zero at tree-level so we have included the loop correction arising from the axion-photon coupling [65], although it is typically very small. Similar loop effects appear for the proton and neutron couplings but are not very relevant.

If SM fermions do not transform under the PQ symmetry, QCD axions do not couple with them at tree level ($c's$ are all zero). These are called “hadronic axions”, of which the Kim-Shifman-Vainshtein-Zakharov (KSVZ) [66, 67] model is an often quoted example. A recent work [18, 68] has classified a wealth of these models constrained by their cosmological viability. Simple models involve one extra charged singlet under the SM gauge group and a new heavy coloured fermion and populate the range $0.25 < |C_{A\gamma}| < 12.75$. The range broadens when multiple coloured new fermions are allowed [18] but only very special charge assignments and field contents lead to significantly different values of $C_{A\gamma}$. Therefore, it turns out that this study provides a very nice bracketing of the photon coupling in axion models, encompassing predictions beyond the KSVZ-type constructions. Therefore, henceforth we will use the above-mentioned range when we plot the predictions for axion models for the axion-photon coupling.
Let us now consider the cases where SM fermions transform under the PQ symmetry. The simplest flavour preserving models involve two-Higgs-doublet models: the Peccei-Quinn model itself where $f_a \sim v_{EW}$—ruled out long ago— and the Dine-Fischler-Srednicki-Zhitnitsky (DFSZ) [26, 27], where an additional SM singlet is responsible for the very large value of $f_a$. Two existing DSFZ variants have $E/N = 8/3$ (type-I) and 2/3 (type-II). Their coupling to electrons is $C_{Ae} = \cos^2 2\beta / 3, \sin^2 2\beta / 3$ respectively, depending of the ratio of the two Higgs particle VEVs $\tan \beta = v_2 / v_1$. These potentially large couplings to electrons make these models especially interesting for explaining the stellar cooling anomalies and to be detected by IAXO.

Introducing new SM Higgs doublets/singlets, new PQ fermions and/or non-flavour diagonal couplings gives rise to axions featuring a broad range of couplings to SM particles, including flavour violating ones [69], which are strongly constrained by rare decays. Reference [70] demonstrated that, by adjusting the couplings to quarks and leptons, it is possible to arrange the model-dependent couplings to cancel the model-independent contributions to the proton, neutron and electron couplings to the QCD axion, the so-called astrophobic axion. This model and those accidentally similar are not severely constrained by stellar evolution (indeed supernova, neutron star, white dwarf and red-giant constraints become essentially irrelevant) and might be most advantageously discovered by IAXO itself. It is worth mentioning that other recently proposed models/constructions like the minimal flavour violating axion [71], the axi-flavon [72] and flaxion [73] do not seem to be particularly astrophobic in general, although a recent study focused on minimality found two interesting exceptions [74].

Another interesting QCD axion model is SMASH [75, 76], together with other axi-Majoron models [77]. SMASH was born as a minimal self-consistent model solving the most pressing issues of particle physics and cosmology: neutrino masses, inflation, dark matter, baryogenesis, the strong CP problem and the stability of the Higgs potential. SMASH is a KSVZ-type QCD axion model where the PQ symmetry mixes with lepton number and its spontaneous breaking gives a majorana mass to right-handed neutrinos. Note that the QCD axion here plays the role of the majoron as well. The SMASH axion as well as other majoron models does not have tree-level coupling to electrons, but they feature a potentially large radiative correction with right-handed neutrinos in the loop. Another interesting connection between the QCD axion and neutrinos is the Ma-axion [78].

At the end of the day, concerning solar axion detection, the main important model dependency is whether axions have couplings to electrons with $C_{Ae} \sim O(1)$ like DFSZ, or the couplings to electrons are loop-suppressed like in KSVZ. In the event of a discovery, IAXO could measure the QCD axion mass as well as the photon and electron couplings independently. Therefore it can potentially distinguish between generic possibilities and pinpoint very conspicuous models in optimistic conditions. We review this possibility in Sec. 8.1.

### 3 A meV mass QCD axion as dark matter candidate

The QCD axion is regarded as one of the best motivated candidates to be the DM of the Universe. Its couplings to SM particles are suppressed due to a large decay constant $f_A$, which ensures the stability and collisionless properties of DM. Furthermore, the axions are produced non-thermally in the early Universe, and hence they are “cold” in the sense that their velocity dispersion is small enough to fit the observed large scale structure.
The cosmological production mechanism of axion DM, called the vacuum realignment mechanism, was first discussed in [79–81]. In the early Universe, the axion field may have an expectation value which is different from the vacuum expectation value (VEV) at the minimum of the effective potential at the present Universe. The effective potential of the axion field arises due to instantons: non-perturbative and topologically non-trivial configurations of the gluon fields in QCD, which become increasingly relevant as QCD becomes non-perturbative, i.e. as the temperature of the Universe approaches the confining phase-transition around $\mathcal{O}(0.16)\text{GeV}$. When the QCD potential becomes sizeable, the axion field starts to oscillate around the minimum of the effective potential, and such a coherent oscillation of the classical axion field contributes to the matter energy density of the Universe. Taking into account this production mechanism, it was concluded in [79–81] that the QCD axion DM would overclose the Universe ($\Omega_A > 1$) if the decay constant takes values above $f_A \approx \mathcal{O}(10^{12})\text{GeV}$ corresponding to the axion mass of $m_A \approx \mathcal{O}(10^{-6})\text{eV}$.

Although the early discussion described above strongly motivated us to consider the axion as a DM candidate, the estimate of the relevant mass range was somewhat simplistic and should be refined in light of recent developments of theoretical and observational cosmology. First of all, the precise estimation of cosmological parameters revealed that DM constitutes only a fraction of the total energy density of the Universe. According to the recent results of the Planck collaboration, the matter density parameter is determined as $\Omega_c h^2 = 0.1206 \pm 0.0021$ (Planck TT+lowE) [82], where $h$ represents the value of the Hubble parameter today, $H_0 = 100h\text{km sec}^{-1}\text{Mpc}^{-1}$. An order of magnitude improvement in the accuracy of $\Omega_c$ leads to a tighter constraint on the decay constant $f_A \lesssim \mathcal{O}(10^{11})\text{GeV}$ for axion DM produced by the realignment mechanism [83]. Furthermore, the axion DM abundance estimate is not so straightforward if we follow the evolution of the axion field in the context of inflationary cosmology. If the Peccei-Quinn (PQ) symmetry is restored after inflation, topological defects such as strings and domain walls are formed, and they could also produce significant amounts of cold axions [84, 85]. On the other hand, if the PQ symmetry is never restored after inflation, such defect contributions can be neglected since their density is significantly diluted due to the exponential expansion during the inflationary epoch.

The difference between the two possible scenarios is sketched in Fig. 1. Here we show how the spatial distribution of an angular field $\theta(x)$ evolves over time, where we introduce the dimensionless axion field

$$\theta(x) = \frac{A(x)}{f_A},$$

which ranges from $-\pi$ to $\pi$. Let us assume that the PQ symmetry has been broken at a very early stage, and that the $\theta$ field takes a certain initial value $\theta_i$ within the Hubble radius $\sim H^{-1}_i$ at that time. If inflation occurs at a later time, the physical scale at which $\theta$ takes a uniform value $\theta_i$ exponentially grows, while the Hubble radius remains almost constant. After inflation, the Hubble radius grows with time, $H^{-1} \sim t$. As shown in the upper panel of Fig. 1, if the PQ symmetry is never restored after inflation, the size of the patch of the Universe in which $\theta$ takes the value $\theta_i$ can be much larger than the Hubble radius even at the present time. In this case, we can assume that the axion field has the same initial value $\theta_i$ throughout the observable Universe at the onset of its coherent oscillation. We call this scenario the pre-inflationary PQ symmetry breaking scenario.

The situation is drastically different if the PQ symmetry is restored after inflation, which is shown in the lower panel of Fig. 1. If the reheating temperature after inflation is large enough, the $\theta$ field has large fluctuations after inflation, taking random values from $-\pi$ to $\pi$. 

- 12 -
Figure 1. A schematic view of the time evolution of the spatial distribution of $\theta$ field in the pre-inflationary PQ symmetry breaking scenario (top) and the post-inflationary PQ symmetry breaking scenario (bottom).
on microscopic scales and restoring the PQ symmetry. After the Universe expands and cools enough, the PQ symmetry is spontaneously broken and $\theta$ takes a certain value $\theta_1$ within the Hubble radius at that time. Note that, however, $\theta$ takes different values ($\theta_2$, $\theta_3$, ... ) outside the Hubble radius, since such regions are causally disconnected. The Hubble radius continues to expand as time goes on, and the field relaxes to the uniform value within the new Hubble radius in order to minimize the gradient energy. Such behaviour continues until the time of the QCD phase transition, at which the mass energy of the axion field cannot be neglected and it starts to oscillate around the minimum of the potential. The typical scale of spatial variation of the axion field at the onset of the oscillation is given by the Hubble radius at the time of the QCD phase transition $\sim H_{\text{QCD}}^{-1}$. Such a scale (typically a few comoving mpc if the Universe is radiation dominated at the time) becomes shorter than the Hubble radius at later times, and we expect that our present observable Universe contains many different patches which had different values of $\theta$ at the time of the QCD phase transition. Furthermore, we also expect that topological defects such as strings and domain walls are formed around the boundaries of different patches. Therefore, the relic axion density should be estimated by summing over all possible field configurations, which include various initial values for $\theta$ at the onset of the coherent oscillation and those produced by the collapse of topological defects. We call this scenario the post-inflationary PQ symmetry breaking scenario.

Since there is a conceptual difference in the spatial distribution of the axion field between the pre-inflationary and post-inflationary PQ symmetry breaking scenarios, estimates of the axion DM abundance will be different accordingly. In the following subsections, we will discuss these two scenarios separately.

3.1 The pre-inflationary Peccei-Quinn symmetry breaking scenario

First, let us consider the pre-inflationary PQ symmetry breaking scenario. In this case, we can ignore the contribution from topological defects, and the relic axion DM abundance can be estimated by simply following the evolution of the homogeneous axion field $\theta$, which can be described by the standard realignment mechanism. As mentioned earlier, the axion field has a certain initial value $\theta_i$ throughout the observable Universe, and the relic axion abundance depends on this initial value as well as on $f_A$.

The axion potential arises from QCD non-perturbative effects and it is thus extremely suppressed at high temperatures due to the asymptotic freedom of the QCD running coupling. The axion is therefore effectively massless at high temperatures. As the temperature of the Universe approaches the QCD confinement scale, the axion mass becomes significant and starts influencing the axion field evolution. This phenomenon can be modelled by the following $T$–dependent potential,

$$V_{\text{QCD}}(A, T) = \chi(T) \left[ 1 - \cos \left( \frac{A}{f_A} \right) \right],$$  \hspace{1cm} (3.2)

where we emphasise that $\chi = \chi(T)$ depends on $T$. Note that we have redefined $A/f_A \to A/f_A - \theta$ for simplicity. Recently, lattice calculations of the topological susceptibility in full QCD became available [86], whose behavior at high temperatures is close to the power law $\chi(T) \propto T^{-n}$ with $n = 8.16$ predicted by the dilute instanton gas approximation [87–89]. Below the QCD phase transition, $\chi$ becomes a constant $\chi_0 = (75.5(5)\text{MeV})^4$ [14]. In what follows, we estimate the axion DM abundance by using the latest lattice QCD results [86].

\footnote{The state of the art calculations of the topological susceptibility $\chi(T)$ based on lattice simulations of full QCD with three flavours have reached a level of accuracy that allows for a precise determination of $\chi_0$.}$^3$
The zero momentum mode of the axion field obeys the following field equation,

\[ \dot{\theta} + 3H\dot{\theta} + \frac{m_A^2}{2}(T) \sin \theta = 0. \quad (3.3) \]

The realignment production of axions occurs when the Hubble friction becomes smaller than the potential force, i.e., \(3H \lesssim m_A\). At that time the axion field starts to oscillate around the minimum of the potential. The solution of Eq. (3.3) guarantees the conservation of the axion number in the comoving volume, from which we can obtain the present density of axion DM. Assuming radiation domination during the onset of oscillations, and the harmonic approximation \(\sin \theta \sim \theta\) one finds [75],

\[
\Omega_{A,\text{real}}h^2 \approx 0.35 \left( \frac{\theta_i}{0.001} \right)^2 \times \left( \frac{f_A}{3 \times 10^{17} \text{GeV}} \right)^{1.17} \quad \text{for} \quad f_A \lesssim 3 \times 10^{17} \text{GeV},
\]

\[
\text{for} \quad f_A \gtrsim 3 \times 10^{17} \text{GeV}. \quad (3.4)
\]

The case \(f_A \gtrsim 3 \times 10^{17} \text{GeV}\) corresponds the onset of axions oscillations happening when the axion mass has already reached its \(T = 0\) value (\(\chi = \chi_0\)). In the case of interest to Helioscope detection, \(f_a \ll 3 \times 10^{17} \text{GeV}\), the onset of axion oscillations happens before the confining phase transition, when \(\chi\) depends very strongly on the temperature. Unfortunately, Eq. (3.4) is based on an approximate solution, which is valid as long as \(|\theta_i| \ll \pi\) (corresponding to \(f_a \gtrsim O(10^{14}) \text{GeV}\)). If \(|\theta_i|\) takes a larger value, the factor \(\theta_i^2\) in Eq. (3.4) is replaced by a \(\theta_i\)-dependent correction term, which can be quantitatively estimated by solving the non-linear field equation [Eq. (3.3)] numerically. The anharmonic correction factor becomes very large in the limit \(|\theta_i| \to \pi\), and in such a case \(\Omega_{A,\text{real}}h^2\) can account for the total cold DM abundance for the lower values of the decay constant \(f_A\) accessible to IAXO. We have computed the required value of \(\theta_i\) to obtain 100\%, 33\%, 10\%, 3.3\% and 1\% of the observed DM abundance and plot them in Fig. 2. One can see that very high fine-tunings of \(\theta_i\) are required to have 100\% for \(f_A \sim 10^9 \text{GeV}\), already noticed by [83], but the values are quite reasonable for a few \%. It is precisely in these meV-mass scenarios, in which the direct axion DM detection is more difficult and one predicts typically a small DM abundance, where IAXO can be the only way of discovering the QCD axion.

In the pre-inflationary scenario, quantum fluctuations of an axion field during inflation lead to isocurvature fluctuations in its realignment dark matter density that are imprinted in the temperature fluctuations of the cosmic microwave background, whose amplitude is strongly constrained by observations [96, 97]. This constraint applies to any ALP DM candidate, not only to QCD axions. The upper limits on isocurvature fluctuations result in a constraint that relates the Hubble scale during inflation \(H_I\) (which gives the typical size of the fluctuations in the \(a\) field), the fraction of axion cold DM to the total, \(r_{aDM} = \Omega_{a,tot}h^2/0.12\), and the decay constant \(f_a\) (during inflation) that might depend on anharmonic effects if \(\theta_i\) is large. In the simplest cosmological scenario, the constraint is [98]

\[
r_{aDM} \frac{d\ln \Omega_{a,tot}}{d\theta_i} \frac{H_I}{2\pi f_a} \lesssim 10^{-5}. \quad (3.5)
\]

The case of the QCD axion is shown in Fig. 2 (right) as an upper bound on \(H_I\) for different values of \(r_{aDM}\). Note that the above graph uses our numerical calculations of \(\theta_i\) based on QCD result in some discrepancy among different numerical treatments [86, 90, 91]. In particular, the result of [90], which shows a much milder temperature dependence, leads to an order of magnitude smaller value for the axion mass explaining the total DM abundance compared with that obtained in [86]. Such a result might be interpreted as huge lattice artifacts caused by strong cutoff effects [86, 91].
Figure 2. Left: Value of the misalignment angle after inflation (initial conditions of the axion DM field) required to obtain today different fractions of the observed DM abundance. Right: Upper bound on the Hubble constant during inflation from Planck’s absence of isocurvature fluctuations for axion DM having such DM fractions. These bounds can be evaded in a number of scenarios [92–95]. Both calculations are performed with the topological susceptibility and equation of state of the SM from [86].

on [86], which is specific to the QCD axion. One can see that values of $f_A \sim 10^9$ GeV and a significant contribution of axion DM $r_{\text{ADM}} \sim O(1)$, require quite low $H_I$. Note that a bound on $H_I$ constrains the maximum possible reheating temperature to be $T_{\text{RH}} \sim \sqrt{H_I M_{\text{Pl}}}$ and this is constrained to be $T_{\text{RH}} \gtrsim 5$ MeV by Big Bang nucleosynthesis [99–102] and Cosmic Microwave Background (CMB) arguments [103, 104] but otherwise there is no phenomenological prejudice against small scale inflation. Indeed, a recent paper turned this argument around and showed that the quantum fluctuations during a low-$H_I$ period of inflation may also be responsible for generating a large enough misalignment angle to give rise to QCD axion or ALP DM [105, 106].

Future CMB polarisation probes like CMB-S4 and LiteBIRD can measure the CMB polarisation effect of gravitational waves from which we can obtain the value of $H_I$. However, due to the moderate sensitivity improvement with respect to the current limits the measured values will not be far from $H_I \sim 10^{14}$ GeV. Thus, if a future CMB polarisation experiment measures $H_I$, we will be forced to consider high scale inflation, which limits even tiny fractions of axion DM in this pre-inflation scenario. It is worth noting here that the isocurvature spectrum can be reduced in a number of non-minimal models [92–95], so pre-inflationary phenomenology is not completely eliminated in this scenario. However, we might say that the simplest axion models of high-scale inflation favour the scenario where the PQ symmetry is broken after inflation, which we discuss in the following subsection.

3.2 The post-inflationary Peccei-Quinn symmetry breaking scenario

The axion appears as a pseudo Nambu-Goldstone boson in the low energy effective theory, and such a description is only valid at energies below the scale of symmetry breaking $v_A$. Once the symmetry is restored, we cannot use the effective field theory description in terms of
the axion field $A(x)$. Instead, we can consider the evolution of a gauge singlet complex scalar field $\Phi(x)$ (the PQ field for QCD axions), which transforms as $\Phi \to \Phi e^{i\alpha}$ with $\alpha$ being a real constant parameter under the global $U(1)_{\text{PQ}}$ symmetry. The PQ symmetry is spontaneously broken when the PQ field acquires a VEV $|\langle \Phi \rangle| = v_{\text{PQ}}/\sqrt{2}$, and after that the axion field $A(x)$ is identified as an angular direction of the PQ field, i.e. $\Phi = (v_{\text{PQ}}/\sqrt{2}) \exp(iA(x)/v_{\text{PQ}})$. The VEV of $\Phi$, $v_{\text{PQ}}$ here plays the role of $v_A$. The Lagrangian for the PQ field is given by
\begin{equation}
(\sqrt{-g})^{-1}\mathcal{L} = \partial_\mu \Phi \partial^\mu \Phi^* - V_{\text{eff}}(\Phi, T),
\end{equation}
where $V_{\text{eff}}(\Phi, T)$ represents the finite-temperature effective potential for the PQ field,
\begin{equation}
V_{\text{eff}}(\Phi, T) = \lambda \left( |\Phi|^2 - \frac{v_{\text{PQ}}^2}{2} \right)^2 + \frac{\lambda}{3} T^2 |\Phi|^2.
\end{equation}
If we assume that the reheating temperature $T_R$ after inflation is sufficiently high, $T_R \gg v_{\text{PQ}}$, the PQ symmetry is restored ($\langle \Phi \rangle = 0$) at early times. Subsequently, it is broken when the temperature drops below $T \lesssim v_{\text{PQ}}$.\footnote{The PQ symmetry may also be restored non-thermally due to non-perturbative field dynamics after inflation $[75, 107-111]$ . The evolution of topological defects in this case is similar to that in the thermally restored case once they enter the scaling regime, and there is no significant difference in the prediction of DM abundance between the two cases. However, in the non-thermally restored case, relativistic axions can be abundantly produced, which tends to violate dark radiation constraints if such axions are not thermalized $[75]$.}

When the $U(1)_{\text{PQ}}$ symmetry is spontaneously broken, vortex-like objects called strings are formed due to the Kibble mechanism $[112]$. After their formation, the string network evolves according to an approximate scaling solution $[113-115]$, where the typical length scale of long strings is given by the cosmic time. In order to maintain the scaling property, the energy stored in strings is dissipated as radiation of massless axions, which leads to a further contribution to the present axion DM abundance $[84]$.

In addition to the dynamics of strings, we must take into account that of domain walls, which appear at the epoch of the QCD phase transition. The appearance of domain walls is understood through the effective potential of the axion field [Eq. (3.2)] written in terms of $v_{\text{PQ}}$,
\begin{equation}
V_{\text{QCD}}(A, T) = \chi(T) \left[ 1 - \cos \left( \frac{N_{\text{DW}} A}{v_{\text{PQ}}} \right) \right],
\end{equation}
where $N_{\text{DW}} = N$ is the colour-anomaly of the PQ symmetry, which turns out to be a positive integer and it is called the "domain wall number" in this context. The above potential explicitly breaks the $U(1)_{\text{PQ}}$ symmetry into its discrete subgroup $Z_{N_{\text{DW}}}$, in which the QCD axion field transforms as $A \to A + 2\pi v_{\text{PQ}} k/N_{\text{DW}}$ ($k = 0, 1, \ldots, N_{\text{DW}} - 1$). Such a symmetry breaking effect is irrelevant in early times since it vanishes at high temperature, $\chi(T) \to 0$, but it gradually arises as the Universe cools. This effect is schematically shown in Fig. 3. Once the Hubble friction becomes smaller than the potential force, the VEV field settles down into one of the $N_{\text{DW}}$ degenerate vacua $\langle A \rangle_k = 2\pi v_{\text{PQ}} k/N_{\text{DW}}$. Since the field value $\langle A(x) \rangle$ must be uncorrelated over distances larger than the Hubble radius at the QCD phase transition $\sim H_{\text{QCD}}^{-1}$, the QCD axion field relaxes into different vacua from one Hubble volume to another. Continuity of the VEV demands that there must be a sheet-like boundary between these regions in which the energy density is as high as $V_{\text{QCD}} \approx 2\chi(T)$. Such a non-trivial field configuration is called a domain wall, and its formation is an inevitable consequence of the QCD axion model $[116]$.
Figure 3. Effective potential of the PQ field (top) and that of the axion field (bottom). The left (right) panel shows the potential at the temperature much higher (lower) than the critical temperature of the QCD phase transition. The axion field corresponds to the direction along the bottom of the potential \( V(\Phi) \) (pink lines), and such a direction is exactly flat at high temperatures. At low temperatures, the periodic axion potential appears due to the presence of the topological susceptibility \( \chi(T) \), which gives rise to \( N_{\text{DW}} \) different minima. In these figures, we choose \( N_{\text{DW}} = 4 \).

Figure 4 illustrates a two-dimensional section of topological defects in the axion model with \( N_{\text{DW}} = 3 \). We note that in every case strings are attached by \( N_{\text{DW}} \) domain walls, since the value of the phase of the PQ field \( \langle A(x)/v_{\text{PQ}} \rangle \) must continuously change from \(-\pi\) to \(\pi\) around the string core. Therefore, we expect that hybrid networks of strings and domain walls, called string-wall systems, are formed at around the epoch of the QCD phase transition.

The domain wall number \( N_{\text{DW}} \) determines the number of degenerate vacua in the effective potential of the \( A \) field. We have already mentioned that in models where only one global \( U(1)_{\text{PQ}} \) symmetry exists, the value of \( N_{\text{DW}} \) is determined by the colour anomaly coefficient, \( \mathcal{N} \). In simple models this coincides with the number of new quark flavours that transform under the global \( U(1)_{\text{PQ}} \) symmetry [116]. For instance, in the KSVZ model there is one hypothetical heavy quark which transforms under \( U(1)_{\text{PQ}} \) and \( N_{\text{DW}} = \mathcal{N} = 1 \) while in DFSZ-I all of the standard model quarks transform under \( U(1)_{\text{PQ}} \) and we have \( N_{\text{DW}} = 6 \).

The subsequent evolution of the string-wall systems differs according to whether \( N_{\text{DW}} = 1 \) or \( N_{\text{DW}} > 1 \). If \( N_{\text{DW}} = 1 \), strings are pulled by one domain wall, which causes the disintegration into smaller pieces of a wall bounded by a string [117]. Therefore, these string-wall systems are unstable, and they collapse soon after their formation. On the other hand, if \( N_{\text{DW}} > 1 \), the tension force of domain walls acts on strings from \( N_{\text{DW}} \) different
Figure 4. A top view of string-wall systems with $N_{DW} = 3$. Gray circles represent the cores of strings, which are attached by three domain walls. Coloured lines correspond to the position of the centre of domain walls. Domain walls exist around boundary of three disconnected vacua, in which the axion field has a value $A/v_{PQ} = 0$, $2\pi/3$, and $-2\pi/3$. Around the centre of domain walls, it has a value $A/v_{PQ} = \pi/3$ (cyan), $\pi$ (magenta), and $-\pi/3$ (yellow).

directions, which makes the systems stable. Once such stable string-wall systems are formed, their evolution also obeys the scaling solution, in which the typical distance between two neighbouring walls is given by the Hubble radius. Since the energy density of such scaling domain wall networks decays slower than that of radiation or matter, it eventually overcloses the Universe, which conflicts with many observational results [116, 118]. Therefore, QCD axion models with $N_{DW} > 1$ suffer from the cosmological domain wall problem if the PQ symmetry is broken after inflation.

One possible solution to the domain wall problem is to consider models where the PQ symmetry is violated by something more than the colour anomaly. Indeed, a small explicit symmetry breaking term in the Lagrangian (2.4) leads to a small energy bias between degenerate vacua [116, 119, 120]. The small bias acts as a pressure force on the walls, squashing the false vacuum region and allowing the string-wall network to collapse. The magnitude of the bias must be large enough such that the collapse of domain walls occurs before they overclose the Universe, but must be sufficiently small because it also generically shifts the minimum of the overall effective potential field away from the CP conserving minimum, which spoils the solution to the strong CP problem.

If we consider the PQ symmetry to be an accidental symmetry at energies $\sim v_{PQ}$, the explicit symmetry breaking terms appearing from dynamics at much higher energy scales $\Lambda$ could be represented by higher dimensional PQ violating operators like

$$\Delta V = c_N \Lambda^4 \left( \frac{\Phi}{\Lambda} \right)^N + h.c.,$$

(3.9)
where $c_N$ is a dimensionless constant. With this parametrisation, even if $\Lambda$ is as high as the Planck scale, one needs to suppress the modulus of the coefficients $c_N$ with $N < 8$ or so to avoid fine-tuning the phases. This is naturally realised in scenarios where the PQ symmetry breaking operators are constrained to fulfill fundamental exact discrete symmetries [121, 122].

In summary, there are two distinct possibilities in the post-inflationary PQ symmetry breaking scenario according to the domain wall number $N_{\text{DW}}$. The string-wall systems are short-lived in the case with $N_{\text{DW}} = 1$, while they are long-lived in the case with $N_{\text{DW}} > 1$. In both cases the estimation of the relic QCD axion DM abundance is quantitatively different from the conventional case, since we must take account of the contribution from the collapse of string-wall systems. In the next subsection, we briefly review the recent results on the axion DM abundance and relevant mass ranges in these two scenarios.

### 3.3 Production of dark matter axions from topological defects

In the post-inflationary PQ symmetry breaking scenario, the present total QCD axion DM abundance can be somewhat artificially split into a sum of three contributions, 1) the contribution from the realignment mechanism, 2) that from global strings, and 3) that from the decay of string-wall systems,

$$
\Omega_{A,\text{tot}} h^2 = \Omega_{A,\text{real}} h^2 + \Omega_{A,\text{string}} h^2 + \Omega_{A,\text{dec}} h^2.
$$

(3.10)

We note now that ALPs can also develop string-wall networks and be produced in similar ways than QCD axions if they are endowed with periodic potentials with one or several minima. The QCD axion is peculiar in that we know its potential and its temperature dependence.

We summarize the timeline of the history of the early Universe in Fig. 5. Here, it is assumed that inflation has happened at sufficiently high energy scale, and that the subsequent reheating is so efficient that the PQ symmetry is restored after inflation. The PQ symmetry is spontaneously broken when the temperature of the Universe becomes $T \sim v_{\text{PQ}}$. At that time strings are formed, and they continue to produce massless axions until the epoch of the QCD phase transition.

The simple picture of the strings decaying into massless axions does not hold once the axion mass becomes non-negligible, which corresponds to the temperature of the Universe of $T \approx 1 \text{ GeV}$, and at that time axions are also produced from the realignment mechanism. The realignment contribution can be estimated in a similar way to Eq. (3.4), but in this case we must take an average over all possible values of the initial angle $\theta$, since the value of $\theta$ is different for each Hubble volume. After performing such averaging procedure, the contribution from the realignment mechanism becomes [75]

$$
\Omega_{A,\text{real}} h^2 \approx (3.8 \pm 0.6) \times 10^{-3} \times \left( \frac{f_A}{10^{10} \text{ GeV}} \right)^{1.165},
$$

(3.11)

where the uncertainty originates from the estimation of the topological susceptibility. This result assumes infinite distance from the strings and is thus probably an overestimate [75].

For temperatures below $T \lesssim 1 \text{ GeV}$, domain walls are formed due to the appearance of the QCD potential (3.8). The collapse of the string-wall systems gives an additional contribution to the cold DM abundance, which we denote by $\Omega_{A,\text{dec}}$. Since the contribution from the string-wall systems is different according to whether $N_{\text{DW}} = 1$ or $N_{\text{DW}} > 1$, in the following subsections we will discuss these two cases separately.
There has been a lot of controversy on the estimation of $\Omega_{A,\text{string}} h^2$ [123–130] and $\Omega_{A,\text{dec}} h^2$ [131, 132], which arises from the difficulty in understanding the energy loss process of topological defects and analyzing the spectrum of the axion produced from them in a quantitative way. A straightforward approach to this problem is to perform first principle field theory simulations of topological defects in the expanding Universe [129]. Along this line, the computational methods to estimate the energy spectrum of radiated axions have been developed in [115, 133–141]. In the following, we review these results and discuss their uncertainties.
### 3.3.1 Models with $N_{DW} = 1$

Let us first consider the case with $N_{DW} = 1$. In this case, the collapse of the string-wall systems occurs immediately after the formation, and the parameter dependence of the total DM abundance is similar to that of the realignment contribution (3.11).

The number density of axions produced from strings at time $t_s$ before the formation of domain walls can be estimated as [136]

$$n_A(t_s) \simeq \frac{\xi v_{PQ}^2}{e t_s} \left[ \ln \left( \frac{\sqrt{\lambda v_{PQ}} t_s}{\sqrt{\xi}} \right) - 3 \right],$$  \hspace{1cm} (3.12)

where the parameters $\xi$ and $\epsilon$ are defined in terms of the energy density of strings $\rho_{\text{string}}$ and the mean energy of axions produced from strings $\langle E_a \rangle$, respectively,

$$\xi = \rho_{\text{string}}(t_s) \left( \mu_{\text{string}}/t_s^2 \right), \quad \epsilon = \frac{\langle E_a \rangle}{(2\pi/t_s)},$$  \hspace{1cm} (3.13)

with $\mu_{\text{string}} = \pi v_{PQ}^2 \ln(\sqrt{\lambda v_{PQ}} t_s/\sqrt{\xi})$ being the tension of strings. The results of field theory simulations indicate that the network of strings evolves toward the scaling solution, in which $\xi$ takes an almost constant value of $O(1)$ [129, 133]. The results of the simulations also show that the mean energy of axions produced from strings is comparable to the Hubble scale at the production time, and $\epsilon$ takes a value of $O(1)$ [133, 136]. This fact implies that most of the axions produced from strings become non-relativistic during the radiation-dominated era, and they contribute to the cold DM abundance. We note that the estimate (3.12) implies that the axion abundance scales linearly with $\xi$. This dependence originates from the assumptions used to derive Eq. (3.12), that the energy density of strings obeys that of the scaling solution and that the total comoving energy of the system consisting of strings and axion radiations is conserved. Adopting the values $\xi = 1.0 \pm 0.5$ and $\epsilon = 4.02 \pm 0.70$ suggested in [136] and multiplying Eq. (3.12) by an appropriate dilution factor, we estimate the contribution from strings as

$$\Omega_{A, \text{string}} h^2 \approx 7.8^{+6.3}_{-4.5} \times 10^{-3} \times \left( \frac{f_A}{10^{10} \text{GeV}} \right)^{1.165}.$$  \hspace{1cm} (3.14)

In addition to the above contribution of axions produced from strings, there is a contribution from those produced from the collapse of string-wall systems. The results of numerical simulations in [136] indicate that the corresponding abundance is slightly smaller than (3.14),

$$\Omega_{A, \text{dec}} h^2 \approx 3.9^{+2.3}_{-2.1} \times 10^{-3} \times \left( \frac{f_A}{10^{10} \text{GeV}} \right)^{1.165} \left( N_{DW} = 1 \right).$$  \hspace{1cm} (3.15)

If we adopt the estimates shown in Eqs. (3.11), (3.14) and (3.15), the total axion abundance (3.10) reads

$$\Omega_{A, \text{tot}} h^2 \approx 1.6^{+1.0}_{-0.7} \times 10^{-2} \times \left( \frac{f_A}{10^{10} \text{GeV}} \right)^{1.165}.$$  \hspace{1cm} (3.16)

---

5Errors shown in Eqs. (3.14), (3.15), and (3.16) are not the standard uncertainty estimate (i.e. those from the propagation of uncertainty law), but maximum and minimum values obtained by using the results of numerical simulations.
Requiring that it explains the cold DM abundance observed today $\Omega_{\text{tot}} h^2 = \Omega_c h^2 \simeq 0.12$, we obtain predictions for the decay constant, $f_A \approx (3.8-9.9) \times 10^{10} \text{GeV}$, or the axion mass, $m_A \approx (0.6-1.5) \times 10^{-4} \text{eV}$.

Our estimate so far is based on the result of numerical simulations performed in [136], but there are some debates on the interpretation of the simulation results. In particular, the conventional simulation method [136] cannot realise a large logarithmic enhancement factor in the string tension $f_{\text{string}}$, and it was pointed out that the effects of such high string tension may further modify the estimate of the DM abundance [137, 138].

Recent studies have reported small (seemingly logarithmic) corrections to the scaling solution [115, 137, 141], which indicate a slow increase of the string length parameter $\xi$. The authors of [115] have studied the uncertainty due to the extrapolation of the simulation results to realistic values of the string tension and emphasize a much broader uncertainty. The extrapolation includes the uncertainty on the mean energy of radiated axions as well as the increase of the string length parameter. The main source of the uncertainty on the axion abundance is the ambiguity in the analysis of the spectrum of axions produced from strings.

It was pointed out in [115] that a different interpretation on the shape of the spectrum would alter the prediction for the axion DM abundance by a few orders of magnitude when extrapolated to realistic values of the string tension. In particular, if the mean energy is indeed comparable to the Hubble scale as claimed in [84, 124, 128, 129, 133, 136], the axion DM abundance can be further enhanced compared to the estimate shown above due to the increase of the string length parameter. Although such infrared (IR) dominated spectrum is incompatible with recent simulation results [115, 137, 141] obtained based on a string tension whose value is smaller than realistic ones, we cannot exclude the possibility that the energy of radiated axion is dominated by IR modes in realistic cases. Intriguingly, simulations performed in [115] show some indication that the spectrum slightly changes towards an IR dominated shape with increasing the string tension, and more careful studies on the axion spectrum in simulations with larger dynamical ranges are warranted to confirm such a trend.

We emphasize that even a small change of the shape of the spectrum could drastically enhance the axion DM abundance when extrapolated to realistic parameter values, and typically the axion DM mass becomes as large as $\mathcal{O}(\text{meV})$ in such scenarios, which is accessible by IAXO.

In order to quantify the potentially large uncertainty on the string contribution, let us recast it to the following form,

$$
\Omega_{A,\text{string}} h^2 \approx 2.6 \times 10^{-4} \times K \left( \frac{f_A}{10^{10} \text{GeV}} \right)^{1.165},
$$

where $K$ represents the axion production efficiency from strings, which gives the number of axions produced until the radiation from strings is terminated at a time $t_e$ around the epoch of the QCD phase transition, $i.e.$ $n_A(t_e) = K H(t_e) v_{\text{PQ}}^2$, and we have used the condition $m_A = 3H$ to estimate $t_e$. The value of $K$ depends on the logarithmic correction to the string tension $\ln(\sqrt{\lambda} v_{\text{PQ}} t_e / \sqrt{\xi})$, which requires extrapolation over an enormous separation range between $\sqrt{\lambda} v_{\text{PQ}} t_e$ and $t_e^{-1}$. The estimate shown in Eq. (3.14) corresponds to $K \approx 13-54$, while it was claimed in [115] that its value can differ by many orders of magnitude according to different interpretations of simulation results. The extrapolation with an extremely IR dominated spectrum leads to a value of $K$ as large as $\lesssim 5 \times 10^3$ [115] which corresponds to the axion DM mass of $m_A \lesssim 4.4 \text{meV}$.

On the other hand, the authors of [139, 140] introduced an alternative technique to perform direct simulations of string-wall networks with high effective string tension. This
study concludes that axions emitted from strings and walls are largely irrelevant, obtaining a DM QCD axion mass on the lower side, $m_A \approx (0.262 \pm 0.034) \times 10^{-4} \text{eV}$ [140], that corresponds to the production efficiency $K \approx 13$ in Eq. (3.17). The results of these simulations also show that the string networks become denser for larger values of the string tension, which leads to a larger value of $\xi \sim 4$ [139, 140]. Reinterpreting these results in light of energy conservation, we expect that the large contribution from the energy density of strings should be compensated by producing more energetic axions in order to realise smaller axion production efficiency, which leads to a smaller DM axion mass. Indeed, if we adopt the value $\xi \sim 4$ in Eq. (3.12) together with the axion production efficiency obtained in [140], we obtain $\epsilon \sim 30–40$ for the mean energy of radiated axions, which is an order of magnitude larger than the value $\epsilon = 4.02 \pm 0.70$ obtained in the conventional field theory simulations [136]. This implies that physics at smaller scales can be relevant to the determination of the axion DM abundance. We note that the new simulation method introduced in [139] is based on an effective theory that breaks down at some smaller distance scales, and hence it is still not straightforward to figure out how the physics of small scale strings affects the axion production efficiency. Further studies on the dynamics of string-wall networks are required to include precise modeling of physics at smaller distance scales.

Regarding the fact that there remains the large uncertainty on the estimation of the axion abundance produced from strings, here we treat $K \approx 13$ corresponding to the result of [140] as a lower limit and $K \lesssim 5 \times 10^3$ obtained in [115] as an upper limit on the axion production efficiency. This corresponds to the following range of the axion DM mass in the models with $N_{\text{DW}} = 1$,

$$2.6 \times 10^{-5} \text{eV} \lesssim m_A \lesssim 4.4 \times 10^{-3} \text{eV} \quad (N_{\text{DW}} = 1),$$

or the axion decay constant,

$$1.3 \times 10^9 \text{GeV} \lesssim f_A \lesssim 2.2 \times 10^{11} \text{GeV} \quad (N_{\text{DW}} = 1).$$

We again emphasize that the latest simulation results [115] show a trend that the IR modes are getting more important for larger values of the string tension, and that the extrapolation with such a feature shows a preference for a higher axion DM mass close to the upper limit on $m_A$ shown above. Note that this mass range is derived based on the assumption that axions produced from strings account for the total cold DM abundance. In other words, IAXO will be able to probe the parameter space where axions can constitute 100% of the observed cold DM abundance within the range of uncertainty.

### 3.3.2 Models with $N_{\text{DW}} > 1$

If $N_{\text{DW}} > 1$, the string-wall systems live longer than those in the case with $N_{\text{DW}} = 1$. They eventually collapse due to the effect of the explicit symmetry breaking term. The present abundance of axions produced from these string-wall systems can be estimated as

$$\Omega_{A,\text{dec}} h^2 \approx 1.75 \times \frac{C_d^{1/2} A^{3/2}}{\epsilon a N_{\text{DW}}^2} \left( \frac{\Xi}{10^{-52}} \right)^{-1/2} \left( \frac{f_A}{10^9 \text{GeV}} \right)^{-1/2} \quad (N_{\text{DW}} > 1),$$

where

$$\Xi \equiv \frac{\Delta V}{2 v_{\text{Q}}^4}$$

$$\Delta V$$

– 24 –
parameterises the magnitude of the explicit symmetry breaking term $\Delta V$ that induces the energy bias between different domains.\footnote{The result of numerical simulations for axionic domain walls with $N_{DW} > 1$ shows some deviation from the exact scaling solution, which leads to additional uncertainties of $\Omega_{a,dec} h^2$ \cite{136}.} The dimensionless parameters $C_d$, $A$, and $\tilde{\epsilon}_a$ appearing in Eq. (3.20) represent the decay time of the string-wall systems, the area of domain walls, and the mean energy of axions radiated from them, whose values can be estimated from numerical simulations \cite{136}. They generically take values of $\mathcal{O}(1)$ but slightly depend on $N_{DW}$.

If the lifetime of the string-wall systems is not sufficiently long, one must take account of the realignment and string contributions in addition to $\Omega_{A,dec} h^2$. For the string contribution, we extend Eq. (3.17) to the case with $N_{DW} > 1$,

$$\Omega_{A,\text{string}} h^2 \approx 2.6 \times 10^{-4} \times K N_{DW}^2 \left( \frac{f_A}{10^{10} \text{ GeV}} \right)^{1.165},$$

(3.22)

and assume that $K$ has a similar uncertainty as in the models with $N_{DW} = 1$. The factor of $N_{DW}^2$ just comes from the fact that the string tension is proportional to $v_{PQ}^2$, which becomes $N_{DW}^2 f_A^2$ when we write everything in terms of the axion decay constant.

If we assume that the PQ symmetry breaking is protected by a discrete $Z_N$ symmetry, and Planck-suppressed operators with $\Lambda = M_{Pl}$, the parameter $\Xi$ in Eq. (3.21) is given by (3.9). In this case, we have the following parameterisation $\cite{122}$,

$$\Xi = \frac{|c_N| N_{DW}^{N-4}}{(\sqrt{2})^N} \left( \frac{f_A}{M_{Pl}} \right)^{N-4}. \quad (3.23)$$

It turns out that $N = 9$ or 10 is favoured if we require that the axions produced by the string-wall systems should not overclose the Universe and that the operator $\Delta V$ does not shift the minimum of the QCD potential (3.2) away from $\theta = 0$ even for $\mathcal{O}(1)$ values of the phase and/or modulus of $c_N$.

Since the magnitude of the energy bias (3.23) becomes more suppressed for smaller values of $f_A$, the lifetime of the string-wall systems becomes longer in the small $f_A$ range. For such long-lived defects, the dominant contribution to the DM abundance is given by Eq. (3.20), which places a lower bound on $f_A$ from the requirement that the axion abundance should not exceed the observed cold DM abundance. On the other hand, if $f_A$ takes a larger value, the lifetime of the defects becomes shorter, and the DM abundance is determined by the realignment and string contributions, which leads to an upper bound on $f_A$.\footnote{As mentioned before, the realignment contribution shown in Eq. (3.11) can be an overestimate. Instead of using it, here we adopt Eq. (3.22) with a lower value of $K \approx 13$ suggested in Sec. 3.3.1 to obtain upper limits on $f_A$ and lower limits on $m_A$ shown in Eqs. (3.24)-(3.27).} Therefore, the value of $f_A$ is constrained to a finite range in the post-inflationary PQ symmetry breaking scenario with $N_{DW} > 1$. For instance, the allowed value of $f_A$ in the models with $N_{DW} = 6$ and $N = 9$ reads

$$4.4 \times 10^7 \text{ GeV} \lesssim f_A \lesssim 9.9 \times 10^9 \text{ GeV} \quad (N_{DW} = 6 \text{ and } N = 9),$$

(3.24)

which corresponds to the mass range

$$5.8 \times 10^{-4} \text{ eV} \lesssim m_A \lesssim 1.3 \times 10^{-1} \text{ eV} \quad (N_{DW} = 6 \text{ and } N = 9).$$

(3.25)
Similarly, for the models with $N_{\text{DW}} = 6$ and $N = 10$, we have
\[
1.3 \times 10^9 \text{GeV} \lesssim f_A \lesssim 9.9 \times 10^9 \text{GeV} \quad (N_{\text{DW}} = 6 \text{ and } N = 10),
\]
which corresponds to the mass range
\[
5.8 \times 10^{-4} \text{eV} \lesssim m_A \lesssim 4.5 \times 10^{-3} \text{eV} \quad (N_{\text{DW}} = 6 \text{ and } N = 10).
\]

Since the relic QCD axion abundance depends on the coefficient $|c_N|$ in (3.23) in addition to $f_A$, axions can explain the total cold DM abundance in the whole mass ranges described above, up to the tuning of the coupling parameter $c_N$ and the uncertainty on the production efficiency $K$ from strings. In particular, axions can account for the whole DM in the parameter range $f_A \approx \mathcal{O}(10^8-10^9)$ GeV and $m_A \approx \mathcal{O}(10^{-3}-10^{-2})$ eV with a mild tuning of the phase of the parameter $c_N$ [122, 142]. Interestingly, such a parameter range agrees with that indicated by the anomalous cooling of stars in various evolutionary stages, which will be discussed in the next section.

We plot the parameter range where QCD axions can account for the observed DM abundance for the models with $N_{\text{DW}} = 1$ and for the models with $N_{\text{DW}} = 6$ in Fig. 6. We see that the predicted value for $m_A$ can be much larger than the conventional estimation $m_A \approx \mathcal{O}(10^{-6})$ eV due to the contribution from strings and string-wall systems. In particular, the axion can be DM in the meV mass range both for the models with $N_{\text{DW}} = 1$ and those with $N_{\text{DW}} > 1$, and such a parameter range can be decisively probed by IAXO.

### 4 Axions as dark radiation

According to the standard cosmological model, ΛCDM, the energy budget of the Universe today consists of (in order of decreasing magnitude): dark energy, dark matter, ordinary baryonic matter, and relic neutrino and photon radiation. However, the dark sector may be significantly richer than this phenomenological model suggests. Extensions of the Standard Model commonly include light particles in the dark sector, e.g. axions and hidden photons. Beyond their potential role as cold DM, these light particles may have been produced in the hot Big Bang from thermal processes or decays of heavy particles and would then contribute to the energy density of the Universe as a ‘dark radiation’ component. The observational success of the ΛCDM model indicates that dark radiation, if it exists, must contribute to the total energy density today by an even smaller amount than the CMB photons and the $C\nu B$ neutrinos. By convention, the dark radiation is parametrised phenomenologically as the ‘excess number of neutrino species’, $\Delta N_{\text{eff}}$:

\[
\rho_{d,x} = \rho_{\text{rad}} - \rho_{\text{CMB}} - \rho_{C\nu B} = \frac{7}{8} \left( \frac{4}{11} \right)^{4/3} \Delta N_{\text{eff}} \rho_{\text{CMB}},
\]

where $\rho_{\text{rad}}$ denotes the total relativistic energy density. The CMB data from Planck (when combined with other experiments) is consistent with $\Delta N_{\text{eff}} = 0$, but a significant dark radiation energy density is still allowed within the 68% error bars of $\sigma(\Delta N_{\text{eff}}) = 0.23$ [143]. Dark radiation is also capable of reducing the persistent tensions between local measurements of $H_0$ [144] and the value inferred using ΛCDM and CMB data. Dark radiation present at the time of Big Bang nucleosynthesis (BBN) would increase both the expansion rate and the freeze-out abundance of neutrons, and consequently affect the light element abundances. Observational determinations of the primordial deuterium abundance give no evidence for
Figure 6. The predicted mass ranges in which QCD axions can account for the observed cold DM abundance in the post-inflationary PQ symmetry breaking scenario. The generic prediction of QCD axion models is plotted in the yellow region. The sensitivity prospects of IAXO are also plotted. The prediction of the post-inflationary PQ symmetry breaking scenario differs according to the value of $N_{DW}$ and the structure of the explicit symmetry breaking terms in the models with $N_{DW} > 1$. Dashed regions correspond to the predictions (up to uncertainties in the estimation of the relic axion abundance) of the models with $N_{DW} = 1$ (red), $N_{DW} = 6$ and $N = 9$ (gray), and $N_{DW} = 6$ and $N = 10$ (blue).

dark radiation with error bars comparable to Planck, $\sigma(\Delta N_{\text{eff}}) = 0.28$ [145]. Future CMB experiments will be significantly more sensitive to dark radiation: the ground based ‘Stage-3’ and ‘Stage-4’ CMB polarisation programmes state a projected sensitivity of $\sigma(\Delta N_{\text{eff}}) = 0.06$ and $\sigma(\Delta N_{\text{eff}}) = 0.02$, respectively.

4.1 Thermally produced ALP dark radiation

The definition of $\Delta N_{\text{eff}}$ is chosen so that an additional neutrino species that initially was in thermal equilibrium with the photon-baryon plasma and decoupled simultaneously with the Standard Model neutrinos gives $\Delta N_{\text{eff}} = 1$. More generally, any thermally produced, light, feebly interacting particles that decouple at time $T_d$ contribute to the dark radiation by:

$$\Delta N_{\text{eff}} = \left( \frac{g_\star(T_\nu)}{g_\star(T_d)} \right)^{4/3} \times \left\{ \begin{array}{ll} 1 & \text{Majorana fermion,} \\ \frac{4}{7} & \text{Scalar.} \end{array} \right.$$ (4.2)
Here $g_\star = g_\star(T)$ denotes the effective number of relativistic degrees of freedom in thermal equilibrium at temperature $T$. Assuming that the dark radiation component decoupled earlier than the regular neutrinos, the Standard Model contribution to $g_\star(t)$ ranges from $g_\star(T_\nu) = 10.75$ to $g_\star(T \gtrsim 3m_{\text{top}}) = 106.75$ (see [86]) the latter figure being applicable at very early times when even the top quark was in thermal equilibrium. Beyond the Standard Model, $g_\star(T)$ may have received contributions from additional particles, and the Standard Model contribution in general only gives a lower bound on $g_\star(T)$. Hence, a thermally produced ALP that decouples before $T \sim 3m_{\text{top}}$ contributes to the dark radiation by:

$$\Delta N_{\text{eff}} \leq 0.027,$$

which saturates if $g_\star$ never received any contributions other than those of the Standard Model and the ALP. Figure 7 shows the value of $\Delta N_{\text{eff}}$ as a function of the decoupling temperature under the SM-only assumption.

A thermal population of QCD axions can be always established through interactions with gluons at sufficiently high temperatures, which happens provided the reheating temperature satisfies $T_{\text{RH}} > T_d$ where [146] (see also [147–149]):

$$T_d \approx 9.6 \cdot 10^6 \text{ GeV} \left( \frac{f_A}{10^{10} \text{ GeV}} \right)^{2.246}.$$  (4.4)

Moreover, the efficiency of the thermal production is enhanced by about 3 orders of magnitude if there is a direct coupling between axion and top quarks [148], and so in this case the required $T_{\text{RH}}$ goes down by the same amount.

From analysis of Planck data and BBN constraints, the reheating temperature of the hot Big Bang plasma is only bounded from below, $T_{\text{RH}} \geq 4.7 \cdot 10^{-3} \text{ GeV}$ [104], but it is not known whether $T_{\text{RH}}$ far exceeded this bound.\(^8\)

\(^8\)Note that a determination of the energy scale of inflation would not model-independently determine the
Even if the reheating temperature is too low to fully thermalise the ALPs, small contributions to the energy density may arise from (non-equilibrium) ALP interactions with the thermal plasma, generically giving $\Delta N_{\text{eff}} \ll 0.027$ \[146\].

If axions have directly couplings to quarks and leptons, some of which are poorly constrained observationally, they can thermalize at temperatures between about 0.1 GeV and 100 GeV, for $f$ below $10^9$ GeV, leading to larger values of $\Delta N_{\text{eff}}$, up to about 0.05 for quarks and even up to 0.5 for leptons \[149–151\], as shown in Fig. 8. Such values are not upper bounds, but predictions, and should be visible by future CMB ‘Stage-4’ experiments, allowing for a very interesting interplay between CMB experiments and IAXO. Large values of $\Delta N_{\text{eff}}$ might even mitigate the present cosmological tension on Hubble constant measurements \[151\], at about 3.5 $\sigma$ between CMB and Supernovae measurements \[82, 152\], and claimed at 3.7 $\sigma$ \[153\], and even 4.4 $\sigma$ \[154\] more recently.

4.2 Non-thermal ALP dark radiation from modulus decay

Non-thermally produced axions may potentially give much larger contributions to the dark radiation energy density of the Universe. In fact, substantial levels of axionic dark radiation (corresponding to $\Delta N_{\text{eff}} \sim O(0.1–10)$) is a generic consequence of string theory models of inflation and its aftermath. The reason is very simple: when the reheating temperature is insufficient to thermalise them, direct decay of the field driving reheating into ALPs will produce a non-thermal cosmic background of relativistic axions, and hence dark radiation. While not specific to string theory, this scenario is a generic consequence of reheating scenarios involving weakly coupled moduli fields, as we will now discuss.

Compactifications of string theory include geometric deformation moduli and axions in general in the low-energy theory (cf. section 2). During inflation, the compactification geometry becomes slightly distorted due to the energy density of the inflaton, and moduli become displaced from their post-inflationary minima. After the end of inflation, the geometry will relax back to its final vacuum configuration, and will undergo oscillations as itsettles down. Specifically, when the Hubble parameter $H$ has decreased to the same order of the mass of one of the moduli fields, $m_\phi$, the modulus starts oscillating around its minimum with an initial amplitude set by the displacement acquired during inflation. The energy density of such a coherently oscillating field red-shifts only like matter, $\rho_\phi \sim R^{-3}$ ($R$ the scale factor of the Friedman-Robertson-Walker metric), and if the modulus is sufficiently long-lived, its energy density comes to dominate over any previous radiation component which red-shifts like $\rho_r \sim R^{-4}$. This way, the most long-lived modulus field comes to dominate the energy density of the Universe, and the reheating temperature of the hot Big Bang plasma will be set at its decay. Moduli arise from higher-dimensional components of the metric tensor and generically couple with gravitational strength interactions, and their decay rate is typically given by,

$$\Gamma \sim \frac{1}{4\pi} \frac{m_\phi^3}{M_{\text{Pl}}^2},$$

up to an $O(1)$ constant. Clearly, parametrically lighter moduli are more long-lived than heavier moduli. Assuming instantaneous reheating, $T_{\text{RH}}$ is given by (for the more precise
Figure 8. Top: $\Delta N_{\text{eff}}$ as a function of $f/c_i$. The green, red and blue label correspond to the predictions for the charm, bottom and top particle, respectively. The orange bands represent the 1$\sigma$ and 2$\sigma$ forecasted contours of CMBS4, plus a more futuristic 1$\sigma$ band, according to [155]. See also [156] for similar forecasts combining CMB-S4 and large scale structure. Bottom: Contribution to $\Delta N_{\text{eff}}$ from muon (blue) and tau (red) scattering as a function of $c_\ell/f$. Decays are possible for off-diagonal couplings; we show only results for tau decays (magenta), which are the only allowed ones in the above range for $c_{\ell\ell}/f$. Each process is shown as a band, parametrizing the uncertainty in the number of relativistic degrees of freedom: the straight line $g_*$ is taken from [157] and the dashed line from [86]. We also show the analytical expectation, $\Delta N_{\text{eff}} \propto f^{-8/3}$, for non-thermalized axions. The orange bands represent the forecasted sensitivities for future CMB experiments [155].
equation, see e.g. [158]):

\[ T_{\text{RH}} \sim \sqrt{H M_{\text{Pl}}} \approx \sqrt{\Gamma M_{\text{Pl}}} \approx \frac{m_{\phi}^{\frac{3}{2}}}{M_{\text{Pl}}^{\frac{1}{2}}} = 20 \text{ MeV} \left( \frac{m_{\phi}}{100 \text{ TeV}} \right)^{\frac{3}{2}}. \]  

(4.6)

Hence, the observational requirement of thermalisation before BBN and neutrino decoupling \( T_{\text{RH}} \geq 4.7 \cdot 10^{-3} \text{ GeV} \) [104], translates into a generic bound on the mass of the lightest modulus: \( m_{\phi} \gtrsim 50 \text{ TeV} \). The typical values of \( m_{\phi} \) from string compactifications is not known in general, but some scenarios that also predict supersymmetry breaking ‘soft’ terms at the TeV scale give \( m_{\phi} \sim 10^6 - 10^7 \text{ GeV} \) [62, 63, 159, 160] which gives from (4.6) \( T_{\text{RH}} \sim 1 – 10 \text{ GeV} \).

Larger minimum moduli masses (say, \( m_{\phi} \gtrsim 10^{11} \text{ GeV} \) for the QCD axion) may produce a thermal axion dark radiation, leading, in effect, to a suppression of \( \Delta N_{\text{eff}} \).

The moduli field \( \phi \) generically has open decay channels into any sufficiently light degree of freedom. In particular, \( \phi \) can decay into axions: \( \phi \rightarrow a + a \). Importantly, weakly coupled axions do not thermalise, and so are more energetic than the average particle in the hot Big Bang plasma,

\[ E_{a,\text{initial}} = \frac{m_{\phi}}{2} \gg m_{\phi} \sqrt{\frac{m_{\phi}}{M_{\text{Pl}}}}. \]  

(4.7)

As the Universe expands, the axion energy redshifts with the inverse scale factor. However, the ratio of the axion energy to the plasma/CMB temperature remains large until the present day [158]:

\[ \left( \frac{E_a}{T} \right)_{\text{now}} = \left( \frac{4}{11} \right)^{1/3} \left( \frac{g_\ast(t_{\nu})}{g_\ast(t_{\text{RH}})} \right)^{1/3} \left( \frac{E_a}{T} \right)_{\text{initial}}. \]  

(4.8)

The levels of axions dark radiation produced directly from modulus decay is then determined by the branching ratio \( B_a \) of \( \phi \) decaying into axions, \( B_a = \Gamma_{\phi \rightarrow aa}/\Gamma_{\phi \rightarrow \text{anything}} \). The relevant value of \( \Delta N_{\text{eff}} \) is given by [50, 51]:

\[ \Delta N_{\text{eff}} = 43 \left( \frac{B_a}{7} \right) \left( \frac{g_\ast(T_{\nu})}{g_\ast(T_{\text{RH}})} \right)^{1/3}. \]  

(4.9)

The exceptionally sensitive CMB Stage-4 experiments will probe sub-percent branching ratios into axions: in the Standard Model (i.e. for \( g_\ast(T_{\text{RH}}) \) in the range 10.75–106.75), a constraint on \( \Delta N_{\text{eff}} < 0.02 \) corresponds to bounds on the branching ratio into axions \( B_a < O(0.3–0.7\%) \). Hence, precision constraints on dark radiation will provide an incredibly sensitive probe of the high-energy physics driving reheating.

The spectrum of the axions produced from modulus decay differs from the thermal Boltzmann distribution. At decay, momentum conservation ensures that all axions have \( E_a = m_{\phi}/2 \), however, the decay of the modulus field is not instantaneous, and, when observed today, axions created early will be slightly more red-shifted than axions created at a later time. This results in a ‘quasi-thermal’ shape [158]. For moduli masses of \( O(10^6) \text{ GeV} \), the characteristic energy of this relativistic ‘Cosmic axion Background’ is around \( O(0.2) \text{ keV} \).

Particular realisations of this scenario have been considered for various string compactifications [50, 51, 161–164], where also the branching ratio into axions can be estimated. Commonly, \( O(\Delta N_{\text{eff}}) \sim 0.1–10 \), and scenarios with many light axions (cf. the ‘axiverse’ [24]) tend to produce levels of dark radiation significantly above current observational limits (see [164, 165] for suggestions on how to ameliorate this problem). Non-thermal axionic dark radiation can also be produced by other mechanisms than freeze-in and moduli decay, see...
e.g. [166]. In sum, dark radiation is a very sensitive probe of the spectrum and dynamics relevant in the early Universe. The next generation of CMB experiments will detect the imprints of a Cosmic axion Background or rule out large classes of string compactifications.

4.3 Additional observational hints of axionic dark radiation

Relativistic ALPs can produce a number of additional signals beyond their effect on the energy density of the Universe. These arise through either direct scattering of the axions off matter, or through axion-photon conversion in background magnetic fields, as explained in section 7.

- **Scattering of ALPs off the thermal plasma**

  Direct scattering of axions off the thermal plasma is a telltale feature of non-thermally produced axions which are naturally much more energetic than the plasma temperature, cf. (4.8). These high-energy axions can access dynamical processes with a centre-of-mass energy $E_{\text{CoM}} \gg T$, which in the thermal plasma would be suppressed by $\exp(-E_{\text{CoM}}/T) \ll 1$. As these axions are non-thermal, the scattering rate is always $\Gamma \ll H$, yet rare scattering events can cause significant deviations from standard cosmology. For example, scattering off photons during BBN is constrained by the observationally inferred primordial helium abundance [158]. However, scattering off the pre-recombination thermal plasma in the red-shift range $z \sim 1100 - 2 \cdot 10^6$ will produce very small spectral distortions of the CMB that will be hard to detect [167]. Highly energetic ALP scattering may furthermore produce dark matter [158].

- **ALP-photon conversion in primordial magnetic fields**

  ALP dark radiation may also be detected through ALP-photon conversion in astrophysical magnetic fields [168]. Particularly interesting are hypothetical primordial cosmic magnetic fields spanning Mpc distances, and galaxy cluster magnetic fields, which tend to be coherent over kpc scales.

  The magnitude of primordial intergalactic fields is unknown, but lies between $10^{-9}$ and $10^{-16}$ G [169]. In the presence of a substantial primordial, cosmic magnetic field, a highly relativistic Cosmic ALP Background can convert into energetic photons that will contribute to the reionization of the Universe. Such dark radiation can be constrained by its contribution to the optical depth to recombination, $\tau$. Conservative estimates using Planck constraints on $\tau$ and $\Delta N_{\text{eff}}$ indicate that this could lead to a combined constraint of $g_{a\gamma}B \leq 10^{-18}$ GeV$^{-1}$ nG [170], where $B$ denotes the cosmic magnetic field, which is observationally constrained to $B \leq nG$ [171]. Hence, if a Cosmic ALP Background is detected by other means, it will provide strong constraints on the strength of cosmological magnetic fields, and vice-versa.

- **Soft X-ray excess from galaxy clusters**

  The magnetic fields of galaxy clusters are even more interesting, as they are well constrained observationally. The fields of galaxy clusters are typically $O(1-10)\mu G$ in strength (e.g. see [172]). These magnetic fields are measured via observations of Faraday rotation of radio sources located in or behind clusters. On passing through a magnetic field, the polarisation angle $\phi$ rotates with wavelength, $\lambda$, as,

  $$\phi = \phi_0 + RM\lambda^2,$$

  (4.10)

-
where the rotation measure is given by:

\[
RM = 812 \int \left( \frac{n_e}{1 \text{ cm}^{-3}} \right) \left( \frac{B_{\parallel}}{1 \mu \text{G}} \right) \left( \frac{dl}{\text{kpc}} \right)
\]  

along the line of sight. As the electron density in a cluster \(n_e(r)\) is well-determined from X-ray measurements, radio observations of the Rotation Measure lead to a statistical characterisation of the magnetic field strength (although as the magnetic field reverses direction many times along the line of sight, it is not possible to know the exact 3-dimensional field configuration). The rate of variation in the Rotation Measure also allows the typical coherence lengths of the magnetic field to be estimated as 1 - 10 kpc, with the field extending throughout the ~1Mpc extent of a galaxy cluster.

The typical electron density \(n_e\) inside a galaxy cluster is between \(10^{-3}\) and \(10^{-2}\) cm\(^{-3}\) and the typical extent of a galaxy cluster is around 1 Mpc. For ALP masses less than around \(10^{-12}\) eV (the plasma mass within galaxy clusters), this implies that at X-ray energies galaxy clusters provide ideal environments for ALP-photon conversion. Indeed, at X-ray energies the ALP-photon interconversion probability on passing through a cluster is \(O(1)\) when \(g_{a\gamma} \sim 10^{-11} \text{GeV}^{-1}\). The conversion probability is energy-dependent, and maximises at \(E \gtrsim \text{keV}\) energies in the galaxy cluster environment.\(^9\) Detailed simulations of ALP-photon conversion passing through the Coma cluster at various energies can be found in [173]. These both confirm these general results and also show the presence of a threshold energy for conversion to occur efficiently (for the Coma cluster, this is around 50 eV).

Non-thermally produced ALP dark radiation arising from the decay of a modulus with \(m_\phi = 10^6\) GeV would currently have a mean energy of around 200 eV [158]. For \(g_{a\gamma}\) in the range \(10^{-13}\)GeV\(^{-1}\) to \(10^{-11}\)GeV\(^{-1}\) this would lead via ALP-photon conversion to an additional contribution to cluster X-ray spectra in the \(O(0.1-1\text{ keV})\) range [168].

In fact, an unidentified excess luminosity above the ICM spectrum at soft X-ray energies, \(O(\lesssim 0.4\text{ keV})\), has been observed by a number of X-ray satellites – and in particular EUVE and ROSAT – from several clusters [174]. This excess is especially strong and well established in the Coma cluster, where it extends from the centre of the cluster out to several megaparsecs away. The soft excess is a somewhat unusual phenomenon, as older satellites such as ROSAT are far more sensitive to it than more recent ones such as XMM-Newton and Chandra. This arises because ROSAT was designed for an all-sky survey, and so has a larger field of view and a much lower internal background than modern satellites which are designed to observe point sources. This gave ROSAT a far greater sensitivity to weak diffuse emission (such as the soft excess) than for more recent satellites.

It was shown in [168, 173], taking the nearby Coma cluster as a particular example for which well-developed models of the magnetic field exists, that the morphology and amplitude of the excess could be explained by ALP-photon conversion of a Cosmic ALP Background. Consistent results were subsequently found for data from the outskirts of the Coma cluster [175], and from several other galaxy clusters for which the magnetic field is relatively well-constrained from Faraday rotation measurements [176]. Hence,\(^9\) Significant conversion probabilities can also be obtained by ‘resonant’ conversion of lower-energy ALPs into photons if the ALP mass is very close to the local plasma frequency.
Figure 9. IAXO sensitivity to the axion/ALP two-photon coupling of cosmological dark radiation (DR). Lines correspond to 1 event/year for $\Delta N_{\text{eff}} = 0.1, 0.5, 1.0$ from top to bottom.

ALP dark radiation may explain this galaxy cluster soft X-ray excess in the parameter space region with $m_a < 10^{-12}$ eV and $g_{a\gamma} \sim 10^{-11}$ GeV$^{-1}$ which is within the reach of detectability of IAXO.

- **Direct detection with IAXO**

  For obtaining the sensitivity of IAXO to dark radiation we parametrize the flux produced by the modulus $\phi$ partially decaying to ALPs as

  \[
  \frac{d\Phi}{dE} = 2 \times 10^6 \times \Delta N_{\text{eff}} \times \frac{E}{E_*} e^{-(E/E_*)^2} \left[ \frac{1}{\text{cm}^2\text{s}\text{keV}} \right],
  \]

  where $E_*$ is the average ALP energy. Under the assumption that negligible background is achievable down to a threshold of 0.2 keV, the sensitivity of IAXO to dark radiation axions/ALPS would be the one shown in Fig. 9, as a function of $E_*$. More realistic prospects depend on the experimental parameters (background and threshold) actually achieved with the technologies of choice to extend IAXO energy window to lower values, and will be studied in the near future.

5 **Axions and inflation**

The precise measurements of the CMB temperature and polarization anisotropies provide strong support for the slow-roll inflation paradigm [177]. Successful inflation requires an extremely flat inflaton potential for minimally coupled scalars, which is naturally realised if the inflaton is an axion.
The simplest axion inflation is the natural inflation with the following potential \[178, 179\],

\[ V_{\text{inf}}(a) = \Lambda^4 \left( 1 - \cos \left( \frac{a}{f_a} \right) \right), \]  

(5.1)

where the inflation scale is set by the energy scale \( \Lambda \). The above potential leads to the large-field inflation where the inflaton field excursion exceeds the Planck mass. In fact, the decay constant \( f_a \) is constrained to be \( f_a \gtrsim 5 M_{\text{Pl}} \) \[177\]. Such a super-Planckian decay constant has been questioned from the perspective of a string theory \[180–184\] and quantum gravity in general\(^{10}\). Even if such a large decay constant is justified, the predicted spectral index \( n_s \) and tensor-to-scalar ratio \( r \) are not favoured by the current CMB observations.

From the effective field theory point of view, there are two possible simple ways to overcome both the trans-Planckian values of the decaying constant and the incompatibilities with current CMB data. The first one invokes the presence of multiple non-abelian gauge fields coupled to the axion in a shift invariant way. \[185–188\]. Here we assume that the axion potential consists of multiple cosine terms with different height and period.

In the minimal case, the potential consists of two cosine terms \[185\]:

\[ V_{\text{inf}}(a) = \Lambda^4 \left( \cos \left( \frac{a}{f_a} + \vartheta \right) - \frac{\kappa}{n^2} \cos \left( n \frac{a}{f_a} \right) \right) + \text{const.}, \]  

(5.2)

where \( n(> 1) \) is a rational number, \( \kappa \) is a numerical coefficient, \( \vartheta \) is a relative phase, and the last term is a constant to realize the tiny cosmological constant at present. The quartic hilltop inflation is realized for \( \vartheta \approx 0 \) and \( \kappa \approx 1 \). The inflation takes place in the vicinity of the origin where the curvature, or the effective mass, is much smaller than the Hubble parameter during inflation. After inflation the inflaton oscillates about one of the potential minima, \( a_{\text{min}} \). Interestingly, if \( n \) is an odd integer, the axion mass at the potential maximum and minimum is equal in magnitude but has an opposite sign, i.e., the axion has a flat-top and flat-bottomed potential. As a result, the axion is very light and cosmologically stable in the present vacuum, and it may also account for DM. This opens up a possibility to unify inflation and DM in terms of an axion which can be searched for by experiments. In the following subsections we briefly summarize implications of the axionic unification of inflation and DM \[189\].

The second possibility is to change the axion propagator by non-minimally coupling to gravity. The only gravity-axion interaction conserving the tree-level shift invariance is Einstein tensor coupling \[190\]

\[ - \frac{G^{\alpha\beta}}{M^2} \partial_{\mu} a \partial_{\nu} a, \]  

(5.3)

where \( M \) is a new mass scale\(^{11}\).

### 5.1 Axion hilltop inflation

In the limit of \( \vartheta = 0 \) and \( \kappa = 1 \), the inflation model with (5.2) reduces to the quartic hilltop inflation, where the inflaton is massless both at the potential maximum and minimum. This case is ruled out by current observations implying a spectral index of order \( n_s = 0.968 \pm 0.006 \) \[177\].

\(^{10}\) The cosmological constant cannot be larger than \( M_{\text{Pl}} \).

\(^{11}\) Note that this is not the unitarity violation scale \[190\].
The predicted spectral index can be increased to give a better fit to the observation by allowing a non-zero $\vartheta$ in the small field scenario [185, 191]. Interestingly, this implies a small axion mass,

$$m_a^2 \equiv V''(a_{\text{min}}) \simeq \left( \frac{9(n^2 - 1)}{2} \right)^{\frac{1}{2}} \vartheta^{1/2} \Lambda^2 \frac{f}{f},$$

(5.4)

where we have approximated $|\vartheta| \ll 1$. Then the observed spectral index fixes the inflaton mass as

$$\frac{m_a^2}{H_i^2} \sim \frac{|V''(a_*)|}{H_i^2} = \mathcal{O}(0.01),$$

(5.5)

where $H_i$ is the Hubble parameter during inflation, and $a_*$ is the axion field value at the horizon exit of cosmological scales. We emphasize that the second equality of (5.5) holds generally in small-field inflation. A similar equality holds when $n_s$ is increased also by allowing $\kappa \neq 1$. Thus, the Planck normalization of curvature perturbation and the observed spectral index determine the relation between the decay constant and the axion mass [189, 192]

$$f_a \simeq 5 \times 10^7 \text{GeV} \left( \frac{n_3}{3} \right)^{\frac{1}{2}} \left( \frac{m_a}{1 \text{ eV}} \right)^{\frac{1}{2}}.$$

(5.6)

It seems that the QCD axion cannot behave as described here without changing radically the behaviour of the QCD potential and compromising the solution to the strong CP problem, however an ALP can very well realize this model. Moreover, if we couple the inflaton to photons we can obtain successful reheating. The coupling to photons is thus given as a function of the ALP mass as

$$g_{a\gamma} \simeq 3 \times 10^{-11} C_{a\gamma} \left( \frac{n_3}{3} \right)^{-\frac{1}{2}} \left( \frac{m_a}{1 \text{ eV}} \right)^{\frac{1}{2}} \text{GeV}^{-1}.$$

(5.7)

### 5.2 Reheating of the ALP inflaton into photons

After inflation, the ALP oscillates about the potential minimum, and the Universe is dominated by coherent oscillations of the ALP. Since the ALP mass is much smaller than the typical curvature $\sim \Lambda^2 / f$, the potential is well approximated by a quartic one. As the Universe expands, the averaged oscillation amplitude, $a_0(t)$, gradually decreases inversely proportional to the scale factor.

The ALP decays and dissipates into plasma through its coupling to photons. First the ALP decays into two photons with the rate,

$$\Gamma(a \rightarrow \gamma\gamma) = \frac{g_{a\gamma}^2}{64\pi} m_{\text{eff}}(t)^3,$$

(5.8)

where $m_{\text{eff}}(t) \equiv |V''(a_0(t))|^{1/2}$ is the effective mass of the ALP when the oscillation amplitude is large. The decay proceeds until it is kinematically blocked by the thermal mass of photons. Afterwards, the ALP gradually dissipates into plasma through scatterings of photons and electrons off the ALP condensate. The dissipation rate is suppressed by the effective ALP mass [193], and given by,

$$\Gamma_{\text{dis,}\gamma} = C_{\text{dis,}\gamma} \frac{g_{a\gamma}^2 T^3}{8} \left( \frac{m_{\text{eff}}^2}{e^4 T^2} \right).$$

(5.9)
where $C_{\text{dis}, \gamma}$ is a numerical factor of $\mathcal{O}(1 - 10)$ that represents uncertainties of the order-of-magnitude estimate as well as the effects of tachyonic preheating and the subsequent self-resonance [194–197]. When the plasma temperature exceeds the weak scale, one needs to take into account the dissipation through similar couplings to the weak gauge bosons.

The effective ALP mass decreases as the averaged oscillation amplitude becomes smaller. At a certain point, the dissipation becomes ineffective, and the reheating ends. To quantify the efficiency of the reheating, we define $\xi$ as the ratio of the energy density of the remnant ALP to the total energy density just after the reheating. For instance, if $\xi = 0.1$, 90% of the initial ALP energy dissipates into plasma, while the rest becomes the remnant.

Once the decay and dissipation rates are given, one can easily solve the Boltzmann equation to evaluate $\xi$. In order for the ALP to dissipate most of its energy, say, $\xi \lesssim \mathcal{O}(0.01)$, it turns out that $g_{\alpha \gamma}$ must be larger than $\mathcal{O}(10^{-11}) \text{GeV}^{-1}$, taking account of the above mentioned uncertainties. In this case, the reheating ends in several Hubble times after inflation, and the reheating temperature is estimated to be

$$T_{\text{RH}} \simeq 40 \left( \frac{g_{\ast}(T_{\text{RH}})}{106.75} \right)^{-\frac{1}{4}} \left( \frac{m_{a}}{1 \text{ eV}} \right)^{\frac{1}{2}} \text{TeV}.$$  \hspace{1cm} (5.10)

In addition, ALPs are also thermalized in plasma, and its abundance is given by

$$\Delta N_{\text{eff}} \simeq 0.03,$$  \hspace{1cm} (5.11)

where it decouples at a temperature slightly below the top quark mass. Thermalized ALPs behave as dark radiation during nucleosynthesis, and they become hot DM at a later time, suppressing the matter power spectrum at small scales. Both effects can be searched for by the future CMB and large-scale structure observations.

### 5.3 The ALP miracle: axionic unification of inflaton and DM

After the reheating becomes ineffective, a small amount of the ALP is left over. The remnant ALP is stable on cosmological time scales, and so, it contributes to DM. The energy density of the ALP remnant, $\rho_a$, first decreases like radiation since the potential is well approximated to be a quartic potential. Then, $\rho_a$ starts to decrease like matter when its oscillation amplitude becomes so small that the ALP mass becomes non-negligible. Assuming that the ALP remnant accounts for the observed DM density, the transition temperature is given by

$$T_c \simeq 0.6 \xi^{-1} \left( \frac{g_{\ast}(T_{\text{RH}})}{g_{\ast}(T_c)} \right)^{\frac{3}{2}} \text{eV},$$  \hspace{1cm} (5.12)

where $T_c$ is the temperature at the transition. Since the ALP condensate behaves like dark radiation at $T > T_c$, it suppresses the small-scale matter power spectrum, which is constrained by the SDSS and/or Lyman-α data. In order to be consistent with the observed data, the transition should take place no later than the redshift $z_c = \mathcal{O}(10^5)$ [198]. Then the ratio $\xi$ is bounded above:

$$\xi \lesssim 0.02 \left( \frac{5 \times 10^5}{1 + z_c} \right),$$  \hspace{1cm} (5.13)

where we substituted $g_{\ast}(T_{\text{RH}}) = 106.75$ and $g_{\ast}(T_c) \simeq 3.909$. Therefore, the small-scale matter power spectrum sets the lower bound on the ALP-photon coupling,

$$g_{\alpha \gamma} \gtrsim \mathcal{O}(10^{-11}) \text{GeV}^{-1}.$$  \hspace{1cm} (5.14)
<table>
<thead>
<tr>
<th>Observable</th>
<th>Stellar System</th>
<th>Proposed Solution(s)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate of period change</td>
<td>WD Variables</td>
<td>axion or ALP coupled to electrons; neutrino magnetic moment;</td>
<td>[202–206]</td>
</tr>
<tr>
<td>Shape of WDLF</td>
<td>WDs</td>
<td>axion or ALP coupled to electrons;</td>
<td>[207, 208]</td>
</tr>
<tr>
<td>Luminosity of the RGB tip</td>
<td>Globular Clusters (M5, ω-Centauri)</td>
<td>axion or ALP coupled to electrons; neutrino magnetic moment;</td>
<td>[209–211]</td>
</tr>
<tr>
<td>R-parameter</td>
<td>Globular Clusters</td>
<td>axion or ALP coupled to photons;</td>
<td>[212, 213]</td>
</tr>
<tr>
<td>B/R</td>
<td>Open Clusters</td>
<td>axion or ALP coupled to photons;</td>
<td>[214–217]</td>
</tr>
<tr>
<td>Neutron stars</td>
<td>CAS A</td>
<td>axion or ALP coupled to neutrons.</td>
<td>[218]</td>
</tr>
</tbody>
</table>

Table 1. Summary of anomalous cooling observations [219]. All the anomalies have, individually, about 1-2σ statistical significance, except for the last two for which the significance has not been quantified.

After the transition, the ALP remnant behaves like CDM. The observed DM abundance is explained if the ALP mass is given by

\[ m_a \sim 0.1 x^{-1} \left( \frac{\xi}{0.01} \right)^{-\frac{3}{4}} \text{eV}, \]  

(5.15)

where \( x \) is a numerical factor of order unity which parametrizes the typical oscillation amplitude of the ALP remnant as \( a_0^{(\text{rem})} = x \xi^{1/4} f \).

Combining (5.7), (5.13), (5.14), and (5.15), we arrive at the unique parameter region,

\[ 0.01 \text{eV} \lesssim m_a \lesssim 1 \text{eV}, \]

(5.16)

\[ g_{a\gamma} = \mathcal{O}(10^{-11}) \text{GeV}^{-1}, \]

(5.17)

where both inflation and DM are simultaneously explained by the ALP. It is highly non-trivial that such a viable region exists without running afoul of the current experimental and observational limits. We refer to this coincidence as the ALP miracle [189]. The region of the ALP parameter space consistent with the CMB observation has been more carefully evaluated in a recent work [192], and it is shown in Fig. 6. Interestingly, the predicted sweet spot of the ALP miracle significantly overlaps with the sensitivity reach of IAXO.

6 Hints for axions in the anomalous cooling of stellar objects

Independent observations of diverse stellar systems have shown deviations from the predicted behaviour, indicating in all cases an over-efficient cooling [142, 199–201]. These deviations, often referred to as cooling anomalies, have been observed in: 1) several pulsating white dwarfs (WDs), in which the cooling efficiency was extracted from the rate of the period change; 2) the WD luminosity function (WDLF), which describes the distribution of WD as a function of their brightness; 3) red giants branch (RGB) stars, in particular the luminosity of the tip of the branch; 4) horizontal branch stars (HB) or, more precisely, the R-parameter, that is the ratio of the number of HB over RGB stars; 5) helium burning supergiants, more specifically the ratio B/R of blue and red supergiants; and 6) neutron stars. Table 1 summarizes the results.
Furthermore, axions may have observable effects on the late evolutionary stages of massive stars, for example influencing the nucleosynthesis [220]. Axions with couplings at reach of IAXO would also modify the threshold initial mass for which carbon is ignited in the stellar core (Mup), ultimately shifting the required mass to explode as core collapse SN by 1-2 $M_\odot$, depending on the couplings [221, 222]. Finally, a recent study indicates that axions reduce the final (pre-SN) luminosity expected for a star of a given initial mass [223]. Presently, observations are sparse but they do seem to indicate a preference for the axion scenario, with couplings somewhat below the recent CAST bound and likely accessible to IAXO and possibly BabyIAXO.

It is quite remarkable that such different stellar systems show, systematically, an excessive amount of energy loss, indicating a lack of understanding in the current modeling of stellar cooling. Notice that the amount of anomalous energy loss is different in different objects. For example, main sequence stars, such as our Sun, do not show this anomalous behaviour and, in general, the amount of exotic cooling indicates some correlation with the interior temperature of the stellar object. The accumulation in the recent years of data from such different stellar objects offers a unique possibility to revise our understanding of the cooling mechanisms in stars.

An appealing explanation to these anomalous observations is to assume that the source of the exotic cooling is a new light, weakly interacting particle, produced in the stellar core and able to stream freely outside, carrying energy away, in much the same way as neutrinos do. Interestingly, the peculiar dependence of the observed additional cooling on temperature and density is extremely selective on the possible particles which could account for all the observations, and indicate a clear preference for a pseudoscalar (axion-like) particle coupled to photons and matter [200]. Indeed, given the large variations in temperature and density of the systems considered, it is quite remarkable that one particle alone can simultaneously address all the observed problems, a fact that, in our opinion, strengthens considerably the physics case for axions and ALPs.

6.1 Summary of cooling anomalies

6.1.1 Pulsating White Dwarfs

Original hints to cooling anomalies were derived from the observations, starting from Kepler et. al. [224] in 1991, that the rate of period change $\dot{P}/P$ of G117 - B15A, a pulsating WD, was larger than predicted by the standard pulsation theory [225].

After over 20 years of additional observations, the hint from G117 - B15A remains [202]. Additionally, the WD variables R548 [203], PG 1351+489 [206], L 19-2 (113) [205] and L 19-2 (192) [205] have all shown similar anomalies, with the observed rate of period change being always larger than expected. The current results are reported in Table 2.12

Since $\dot{P}/P$ is practically proportional to the cooling rate $\dot{T}/T$, the results seem to indicate that these WDs are cooling substantially faster than expected. The unaccounted energy loss could be due to a novel particle, efficiently produced in the dense core of a WD and freely escaping carrying energy away. Examples considered in the literature are axions

---

12The discrepancy has been calculated using the data for $\dot{P}/P$ in the quoted references and does not account for possible unknown systematics. In particular, the high significance of the discrepancy for G117 - B15A could be due to the hypothesis that the particular oscillating mode examined (with period about 215 s) is trapped in the envelope (see, e.g., [202]), an assumption which could be incorrect [208]. Relaxing this hypothesis would significantly reduce the discrepancy [202] (see also discussion in [200]).
produced through electron bremsstrahlung [202, 205] (see sec. 6.2) and neutrinos with an anomalous large magnetic moment produced through plasmon decay [204].

6.1.2 The White Dwarf Luminosity Function
An anomalous behaviour was also observed in the WD luminosity function (WDLF), which describes the number distribution of WDs in brightness intervals. The particular shape of this distribution depends on the lifetime of WDs in a specific luminosity bin and, consequently, on the efficiency of the cooling mechanisms. A number of studies (see, e.g., [208]) showed that additional cooling provided by axions/ALPs coupled to electrons could improve considerably the fit, while even a large neutrino magnetic moment would not have any substantial effect on the WDLF [207]. This result can be attributed to the too steep temperature dependence of the plasmon decay into neutrinos [200].

A more recent study of the hot part of the WDLF [226] did not confirm this anomalous behaviour. However, the hotter section of the WDLF has much larger observational errors and the ALP production would be almost completely hidden by standard neutrino cooling in the hottest WDs. Whatever the case, a decisive improvement in our understanding of the WDLF is expected in the next decade or so. Observations from the Gaia satellite have already increased the catalog of WDs by an order of magnitude with respect to SDSS, and LSST is expected to ultimately increase the census of the WDs to tens of millions [227].

6.1.3 Globular Clusters
Further hints to anomalous energy loss emerged in the recent analyses of Red Giant Branch (RGB) stars in the globular cluster M5 [209, 210] and, with less significance, in the globular cluster ω-Centauri [211].

The red giant is the evolutionary stage of low mass stars that follows the main sequence. Stars in this stage have a He core and burn H in a shell. During their evolution in the RGB, stars become brighter and brighter until they reach a tip in the color magnitude diagram at the time of the He-flash, corresponding to the He ignition in their core. After the He-flash, the luminosity decreases and the stars move into the Horizontal Branch (HB) stage.

Additional cooling of the core during the RGB stage would delay the He ignition allowing the star to become brighter before moving into the HB phase. The tip of the RGB is, therefore, a measure of the cooling efficiency during the RGB evolutionary stage.

The studies in [209–211] showed a brighter than expected tip of the RGB, indicating a somewhat over-efficient cooling during the evolutionary phase preceding the helium flash. In all cases, the significance is fairly low, ~ 1.2 σ in the M5 analysis and less than 1 σ in ω-Centauri. In spite of the low significance of these results, however, it is remarkable that RGB observations seem to confirm the need for additional cooling in agreement with the

<table>
<thead>
<tr>
<th>WD</th>
<th>class</th>
<th>$P$ [s]</th>
<th>$P_{\text{obs}}$ [s/s]</th>
<th>$P_{\text{th}}$ [s/s]</th>
<th>discrepancy</th>
</tr>
</thead>
<tbody>
<tr>
<td>G117 - B15A</td>
<td>DA</td>
<td>215</td>
<td>$(4.19 \pm 0.73) \times 10^{-15}$</td>
<td>$(1.25 \pm 0.09) \times 10^{-15}$</td>
<td>$4\sigma$</td>
</tr>
<tr>
<td>R548</td>
<td>DA</td>
<td>213</td>
<td>$(3.3 \pm 1.1) \times 10^{-15}$</td>
<td>$(1.1 \pm 0.09) \times 10^{-15}$</td>
<td>$2\sigma$</td>
</tr>
<tr>
<td>PG 1351+489</td>
<td>DB</td>
<td>489</td>
<td>$(2.0 \pm 0.9) \times 10^{-13}$</td>
<td>$(0.81 \pm 0.5) \times 10^{-13}$</td>
<td>$1.1\sigma$</td>
</tr>
<tr>
<td>L 19-2 (113)</td>
<td>DA</td>
<td>113</td>
<td>$(3.0 \pm 0.6) \times 10^{-15}$</td>
<td>$(1.42 \pm 0.85) \times 10^{-15}$</td>
<td>$1.5\sigma$</td>
</tr>
<tr>
<td>L 19-2 (192)</td>
<td>DA</td>
<td>192</td>
<td>$(3.0 \pm 0.6) \times 10^{-15}$</td>
<td>$(2.41 \pm 1.45) \times 10^{-15}$</td>
<td>$0.4\sigma$</td>
</tr>
</tbody>
</table>

Table 2. Results for $\dot{P}$ measured and expected in WD variables.
completely unrelated results from WDs. These results can be improved using multi-band photometry of multiple globular clusters [228]. A considerable reduction of the observational uncertainties, particularly those related to the clusters distances, are expected from the data of the Gaia satellite mission [229].

Moreover, an independent analysis [212] showed a disagreement between the observed and expected $R$-parameter, $R = N_{\text{HB}}/N_{\text{RGB}}$, which compares the numbers of stars in the HB ($N_{\text{HB}}$) and in the upper portion of the RGB ($N_{\text{RGB}}$). Assuming Gaussian errors, the discrepancy between the observed, $R = 1.39 \pm 0.03$, and the expected, $R = 1.47 \pm 0.03$, values is about 2σ. The result indicates a surplus of RGB with respect to the numerical prediction and can be interpreted as an anomalous cooling with a different degree of efficiency in the two evolutionary stages. This anomaly is known as the $R$-parameter or HB hint.

Assuming the result is due to new physics, the ideal candidate to explain the low number of HB would be an ALP coupled to photons and produced through the Primakoff mechanism in the stellar core. This process is fairly inefficient in the high density environment of the RG core and could have the effect of accelerating only the following HB stage (hence, HB hint), reducing therefore the number of HB versus RGB stars and explaining the result for $R$ [212, 213]. A cooling mechanism efficient during the RGB stage could, however, also explain the discrepancy in the predicted and measured $R$-parameter. In particular, an ALP coupled to electrons and produced through electron bremsstrahlung or Compton could produce a similar effect [200]. In general, we could expect a combination of the two mechanisms, as discussed in sec. 6.2.

### 6.1.4 He-burning Supergiants

An additional deviation from the standard cooling theory was observed in core He-burning stars of intermediate mass ($M \sim 10M_\odot$). The problem, in this case, is that numerical simulations predict a larger number ratio of blue (hot) over red (cold) supergiants (B/R), with respect to what is actually observed [214, 215]. The predicted number would be lowered (alleviating or, perhaps, solving the B/R problem) in the hypothesis of an additional cooling channel efficient in the stellar core but not in the H-burning shell [230], which could be provided by axions coupled to photons and produced through the Primakoff process in a way analogous to what discussed for HB stars [216, 217, 231]. An exact prediction of the required additional cooling is, however, presently unavailable.

### 6.1.5 Neutron Star in CAS A

Finally, X-ray observations of the surface temperature of a neutron star in Cassiopeia A [232–234] showed a cooling rate considerably faster than expected. The effect seems to indicate the need for an additional energy loss roughly equal to the standard one. This was interpreted in terms of an axion-neutron coupling [218]

$$g_{an} \approx 4 \times 10^{-10}.$$  \hspace{1cm} (6.1)

Although this result is compatible with well established limits from NS cooling [235, 236], new results based on different assumptions of the NS micro-physics have challenged this picture [237, 238]. Moreover, the observed anomaly may also have origin in the phase transition of the neutron condensate into a multicomponent state [239]. Given this controversy, and the general difficulty in the modelling of NS, we decided not to include this result in our global analysis of the cooling anomalies presented here.
6.2 Cooling anomalies, axions and IAXO

As discussed above, axions/ALPs are favourite candidates, among the various new physics options, to explain the cooling anomalies [200]. More recently, these have also been interpreted in terms of concrete QCD axions models [142] such as KSVZ, DFSZ, and Axi-Majoron (A/J). Here we present the regions of the ALP and axion parameter space hinted by the anomalous observations and the IAXO potential to probe these areas.

We consider first the case, often discussed in the literature, of an ALP interacting only with photons. In this case, the most relevant axion production mechanism is the Primakoff process,

$$\gamma + Ze \rightarrow Ze + a,$$  \hspace{1cm} (6.2)

which consists in the conversion of a photon into an ALP in the electric field of nuclei and electrons in the stellar core. This process depends strongly on the environment temperature and is suppressed at high density (e.g., those characterizing the core of WDs and RGB stars) by the plasma frequency and degeneracy effects (see, e.g., [240]). Thus, if we ignore the interaction with electrons, axions cannot provide a solution for the excessive cooling observed in WDs and RGB. Instead, their main effect on the evolution of low mass stars would be to accelerate the HB phase while leaving essentially unchanged the RGB stage, and therefore to reduce the expected R-parameter.

This property has been used to constrain the axion-photon coupling [212, 213, 240]. The current 2σ bound, $g_{a\gamma} < 0.65 \times 10^{-10}\text{GeV}^{-1}$ [212, 213], is shown in Fig. 10 (left) by the solid black line labeled “HB”.

The red-hashed region in the left panel of Fig. 10 is the hint from the $R$-parameter (HB-hint) [212, 213, 241] at 1σ confidence level: $g_{ae} = (0.29 \pm 0.18) \times 10^{-10}\text{GeV}^{-1}$. Notice that the bound and the hinted regions are obtained assuming a vanishing ALP-electron interaction. The more general case will be discussed below. As evident from the figure, IAXO is expected to have sufficient sensitivity to detect ALPs in this region, for masses below $m_a \sim 0.1$ eV.

Axions (or ALPs) interacting also with electrons have the potential to explain the additional hints from the WD pulsation, the WDLF, and RGB stars. Moreover, the axion-electron coupling opens up new axion production channels in the Sun, improving the discovery potential of IAXO.

The 1σ intervals on the axion/ALPs coupling with electrons and photons derived from the observations of individual stellar systems are shown in Tab. 3.\textsuperscript{13} The most relevant axion production mechanism at high density is the bremsstrahlung process,

$$e + Ze \rightarrow Ze + e + a,$$  \hspace{1cm} (6.3)

which induces an additional energy loss rate proportional to $T^4$. As shown in [200], this temperature dependence is optimal to fit the WDLF and provides a reasonably good explanation for the observed excess cooling in DA and DB WD variables, whose internal temperatures differ by a factor of a few. The combined analysis of all the observed WD variables gives a fairly good fit, $\chi^2_{\text{min}}$/d.o.f= 1.1, for $g_{ae} = 2.9 \times 10^{-13}$ and favours the axion solution at 2σ.

\textsuperscript{13}The confidence intervals for the WD variables shown in Tab. 3 have been derived from a likelihood analysis of the data in [203–205] and are not given in the original references. The hint from the WDLF is derived from the data in [208]. The RGB hint refers to the M5 data [210] and is calculated from a likelihood analysis combining the observational and computational 1σ errors in quadrature. It has a slightly smaller error than the one reported in [200], which was based on a conservative interpretation of Fig. 2 in [210].
**Figure 10.** Summary of astrophysical hints in the \((g_{a\gamma}, m_a)\) and \((\sqrt{g_{a\gamma}g_{ae}}, m_a)\) planes compared with the expected sensitivity of IAXO. The 2\(\sigma\) regions corresponding to five explicit QCD axion models (DFSZ I, DFSZ II and A/J with \(\mathcal{E}/N = 2/3, 5/3\) and 8/3) that can account for the WD/RBG/R anomalies following the work in \[142\] are shown. Only the fraction of the regions for which solar axion production through the \(g_{a\gamma}\) or \(g_{ae}\) couplings is dominant, is shown on the left and right plots respectively. The unitarity constraints have also been imposed. Left: The red-hashed region is the hint from HB stars \[212, 213, 241\], obtained in the assumption of interaction with photons only. The yellow region shows the customary band for QCD axion models, see text for details. The mentioned hinted regions are shown as red segments. Note that the DFSZ lines have been displaced 5\% upwards for visibility as they overlap with A/J models. The A/J model with \(\mathcal{E}/N = 5/3\) does not appear here because solar production via \(g_{ae}\) is dominant for all this region. Right: The IAXO sensitivity assumes only solar production through the electron coupling. The 2\(\sigma\) hinted regions appear as diamond-shaped regions in yellow (A/J \(\mathcal{E}/N = 5/3\)), red (A/J \(\mathcal{E}/N = 8/3\)), orange (A/J \(\mathcal{E}/N = 2/3\)), light green (DFSZ II) and dark green (DFSZ I). The wide yellow band encompasses all possible DFSZ models within unitarity constraints. As shown, the combined prospects of IAXO via the \(g_{a\gamma}\) and \(g_{ae}\) channels will probe most of the hinted region, with the only exception of the A/J \(\mathcal{E}/N = 5/3\) model, due to the extremely suppressed \(g_{a\gamma}\) coupling.

<table>
<thead>
<tr>
<th>observable</th>
<th>hint (1(\sigma))</th>
<th>observable</th>
<th>hint (1(\sigma))</th>
</tr>
</thead>
<tbody>
<tr>
<td>WD G117 - B15A</td>
<td>(\alpha_{26} = 1.89 \pm 0.47)</td>
<td>WD R548</td>
<td>(\alpha_{26} = 1.84 \pm 0.93)</td>
</tr>
<tr>
<td>WD PG 1351+489</td>
<td>(\alpha_{26} = 0.36 \pm 0.38)</td>
<td>WD L19-2 (113 s mode)</td>
<td>(\alpha_{26} = 2.08 \pm 1.35)</td>
</tr>
<tr>
<td>WD L19-2 (192 s mode)</td>
<td>(\alpha_{26} = 0.5 \pm 1.2)</td>
<td>WDLF</td>
<td>(\alpha_{26} = 0.16^{+0.19}_{-0.14})</td>
</tr>
<tr>
<td>luminosity of RGB tip</td>
<td>(\alpha_{26} = 0.28^{+0.47}_{-0.24})</td>
<td>R-parameter</td>
<td>(g_{10} = 0.29 \pm 0.18)</td>
</tr>
</tbody>
</table>

*Table 3.* Hints at 1 \(\sigma\) from stellar cooling anomalies (from ref. \[219\]). Here, \(\alpha_{26} = (g_{ae} \times 10^{13})^2/4\pi\) and \(g_{10} = g_{a\gamma} \times 10^{10}\)GeV. The hint on the R-parameter shown here assumes no ALP-electron interaction.

Moreover, the peculiar temperature dependence of the axion bremsstrahlung rate allows to account for the excessive cooling observed in RGB stars \[210\] (which have a considerably larger internal temperature than WDs) with a comparable axion-electron coupling. The combination of the hints from the WD pulsation, the WDLF and RGB stars gives \(g_{ae} = \)
Figure 11. Hinted regions for the axion-electron coupling, with $\alpha_{26} = (g_{ae} \times 10^{13})^2 / 4\pi$. The segments in the left side plot show the 1 σ intervals. On the right we show the fit of all the hints combined, excluding the G117-B15A hint, as discussed in [200].

$1.51^{+0.25}_{-0.29} \times 10^{-13}$ with $\chi^2_{\text{min}} / \text{d.o.f} = 1.1$ and favors the axion solution at about 3 σ, see Fig. 11.

A more recent independent analysis of the RGB luminosity tip in ω-Centauri [211] confirms the hint to additional cooling, though this study has so far been carried out considering only the neutrino magnetic moment as source of exotic cooling.

Finally, if we add also the analysis from the $R-$parameter anomaly, we find the results presented in Fig. 12, where the light brown region shows the 1 σ hinted area. Notice that the $R$-parameter depends, in general, on both the axion-electron and axion-photon couplings [200]. This explains the slight left bending shape of the $R-$parameter hinted region. In particular, the quantitative analysis with the data at hand shows that it is possible to explain all the observed cooling hints, including the $R$-parameter anomaly, even neglecting the axion-photon coupling, though there is a preference for non-vanishing couplings with both electrons and photons. The best fit values, $g_{ae} = 1.5 \times 10^{-13}$ and $g_{a\gamma} = 0.13 \times 10^{-10}$ GeV$^{-1}$, indicated in the figure with a red dot, are well within reach of IAXO though the 1 σ hinted region extends, in this case, to lower axion-photon couplings.

The sensitivity of IAXO to axions coupled to both electrons and photons is shown in Fig. 10 (right). Here it is assumed that production of solar axions proceeds through the axion-electron coupling and detection through the photon coupling. If production via the Primakoff process dominates the solar flux the relevant graphical reference is Fig. 10 (left). We can already note that the best fit value $g_{ae} = 1.5 \times 10^{-13}$ with $g_{a\gamma} = 0.13 \times 10^{-10}$ GeV$^{-1}$ would be detectable by IAXO up to ALP masses of 0.1 eV. However, we recall again that within the 1σ hinted region, the coupling to photons could be zero and this would prevent any possibility of detection by IAXO.

The situation is much better defined for QCD axion models, where we know that only fine-tuning can provide a small coupling to photons. Indeed, it is more generic to have a photon coupling than a sizeable electron coupling. Moreover, axions couple generically to protons and neutrons and this implies a strong constraint from the measured duration of the neutrino pulse of SN1987A (see, e.g., [242–244]). Reference [142] studied three classes of axion models, KSVZ, DFSZ and A/J, which differ on the origin of the axion coupling to electrons. In KSVZ, this arises from a loop involving the photon coupling and turns out to be very small. The interplay of the direct constraints on the photon coupling and the SN1987A bound does not allow a KSVZ solution to the cooling anomalies [142]. The situation is very different in DFSZ I and DFSZ II models, which are two Higgs doublet models extended
with an additional SM singlet scalar that allows a PQ symmetry spontaneously broken at high \( f_a \) (see for instance [77]). DFSZ II has a larger \( C_{a\gamma} = E/\mathcal{N} - 1.92 = 2/3 - 1.92 \) and will be thus easier to discover with IAXO than DFSZ I, for which \( E/\mathcal{N} = 8/3 \), implying a larger cancellation between the model dependent and model independent contributions to the photon coupling. The regions where these models can account for the WB/RGB/HB hints are depicted in Fig. 10 (right) and Fig. 13 (top panels). In Fig. 10, we are assuming production solely through the axion-electron coupling. This condition is relaxed in Fig. 13.

In Axi-Majoron models, the electron coupling arises through a loop involving right-handed neutrinos and can be large if the involved Yukawa couplings are large. The three different Axi-Majoron models considered consists in the addition of a new heavy quark and a SM scalar singlet and can be considered as variations of KSVZ where the PQ scalar field gives mass to RH neutrinos and breaks lepton number spontaneously. They differ in the SM charges of the heavy quark that give different couplings to photons. The hinted regions for these models are shown in the right panel of Fig. 10 and in Fig. 13 (bottom panels). Just like in the case of DFSZ axions, in Fig. 10 we are assuming production solely through the axion-electron coupling while we relaxed this conditions in Fig. 13.

As shown in [142], DFSZ and Axi-Majoron models can well explain the stellar hints even when the constraints from SN 1987A are accounted for. However, some tension does exist in these models between the hinted values for the axion coupling with electrons and photons (from WD and globular cluster stars), and the bound on the axion-nucleus coupling extracted from the observed neutrino signal of SN 1987A. Recent astrophobic models [70] relax further this tension and promise even better fits. An quantitative analysis of these models in the contest of the cooling anomalies is in preparation.

In general, IAXO will be capable to find the hinted QCD axions in a sizable part of the parameter space, although it is with the upgraded IAXO configuration that most of the
At high mass there is room for a QCD axion solution, in a range of parameters still partially accessible to IAXO, while at lower masses they point to ALPs which could provide the required CDM [43]. Finally, at slightly lower masses there is an overlap with the parameters required to explain the transparency hints [245]. In all cases, IAXO shows a high potential to explore the relevant parameter space.

hinted parameter space will be covered. The A/J model $R_Q = (3, 2, +1/6)$ is an exception but it is already quite a tuned solution very close to the unitarity constraint [142]. In the region accessible by IAXO the solar axion emission happens mostly through the axion-electron coupling. This is particularly true for the DFSZ I model. In this case, the production rate induced by the axion-photon coupling would dominate for low values of $\tan \beta$, in a region almost entirely excluded by the unitarity constraints.

The regions hinted by the cooling anomalies, shown in Figs. 10 and 12, are phenomenologically quite interesting. They are largely accessible to IAXO and partially to ALPS II [35]. At high mass there is room for a QCD axion solution, in a range of parameters still partially accessible to IAXO, while at lower masses they point to ALPs which could provide the required CDM [43]. Finally, at slightly lower masses there is an overlap with the parameters required to explain the transparency hints [245]. In all cases, IAXO shows a high potential to explore the relevant parameter space.

\textbf{Figure 13.} Isocontours of the $\chi^2$ fit to the WD, RGB and HB anomalies in different axion models from [142]. In DFSZ $\tan \beta$ is the ratio of the 2 Higgs doublet VEVs, which is constrained by unitarity in the Yukawa couplings to be within the dashed lines. The parameter entering in the axion-electron coupling in the A/J models is the trace over RH Yukawa shown as $|\text{tr} - 2\kappa_{\chi e}|$ and it also has an upper bound from unitarity. The green regions are the sensitivities of IAXO and ARIADNE. See [142] for details. The lower axis gives the axion mass in units of eV.
7 Axion-like particles in the propagation of photons over astronomical distances

IAXO will be capable to probe the ALP parameter space which is currently only accessible to astrophysical observations. Astrophysical magnetic fields can range from values below $10^{-9}$ G in intergalactic space to $\gtrsim 10^{14}$ G on the surface of magnetars (highly magnetized neutron stars) on a large range of spatial scales. As the photon-ALP conversion probability scales with the product of the field strength $B$ and the spatial extend of the field $L$, in some astrophysical environments a large conversion probability could be possible [246–248]. For a homogeneous magnetic field, the conversion probability becomes maximal and independent of energy for energies $E_{\text{crit}} \lesssim E \lesssim E_{\text{max}}$, with [e.g. 246, 249] (neglecting dispersion due to, e.g., the CMB [250])

\[
E_{\text{crit}} \approx \frac{|m_a^2 - \omega_{\text{pl}}^2|}{2g_{\text{av}}B} \sim 2.5 \text{ GeV} |m_a^2 \text{peV} - 1.4 \times 10^{-3} \text{ neV}^{\text{cm}^{-3}}| g_{\text{av}}^{-1} B_{\mu G}^{-1},
\]

\[
E_{\text{max}} \approx \frac{90\pi B_{\text{cr}}^2 g_{\text{av}}}{7\alpha} \sim 2.12 \times 10^6 \text{ GeV} g_{\text{av}} B_{\mu G}^{-1}.
\]

In the above equations, we have introduced the notation $B_X = B/X$, $m_X = m_a/X$, and $g_X = g_{\text{av}} \times 10^X / \text{GeV}^{-1}$. The plasma frequency is connected to the electron density of the medium through $\omega_{\text{pl}} \sim 0.037 \sqrt{n_{\text{cm}^{-3}}} \text{ neV}$. Above $E_{\text{max}}$, the oscillations are damped due to the QED vacuum polarisation. The chosen units in the above equations already indicate that for magnetic fields with a strength of $\mu G$ and ALP masses around neV, a strong mixing can be expected at gamma-ray energies. Such magnetic fields are commonly observed in the intra-cluster medium of galaxy clusters, in galaxies themselves, and in the lobes of jets launched by active galactic nuclei (AGNs) [e.g. 251, 252]. For lower ALP masses $m_a \lesssim 10^{-12} \text{ eV}$, $E_{\text{crit}}$ shifts to X-ray energies. For X-rays or gamma rays produced in AGN, Tab. ?? displays the typical strength of the magnetic field $B$, its coherence length $L$, and the photon-ALP oscillation length $\lambda_a$ of magnetic fields traversed by the photons.

Below, we summarize current hints, constraints, and future sensitivities of gamma-ray and X-ray observations. In comparison to future IAXO observations, the X-ray and gamma-ray measurements suffer from the unknown exact conditions of the involved magnetic fields and photon fluxes. In this respect, IAXO offers the unique opportunity to test any hints found in astrophysical observations in a more controlled setting.

7.1 Conversion between gamma-ray photons and ALPs

For gamma rays, one expects two possible observables as a result of photon-ALP mixing. On the one hand, the conversions should lead to oscillatory features around $E_{\text{crit}}$ and $E_{\text{max}}$ as some photons will oscillate into ALPs thereby reducing the photon flux. The shape of these oscillations will depend on the morphology of the magnetic field, the plasma frequency, the ALP parameters, and the gamma-ray energy. Since astrophysical magnetic fields are often turbulent in nature, the spectral features are expected to show a complicated dependence on energy.

On the other hand, photon-ALP oscillations could lead to a boost in the photon flux due to an astrophysical version of the “light shining through a wall” experiment, in which photons convert to axions or ALPs and back on either side of an opaque wall. For gamma rays emitted by AGNs, the opacity is caused by the interaction of gamma rays with lower-energy photons from infrared (IR) to ultraviolet (UV), which produces an electron-positron
Table 4. Various astronomical objects with typical values of average magnetic field $B$, domain length $L$ and corresponding photon-ALP oscillation length $\lambda_a$ traversed by a photon beam originating from AGNs. IGM stands for intergalactic medium. The magnetic field in the AGN jet does not possess a domain-like structure [253]. The values of $\lambda_a$ are calculated in the strong mixing regime where the plasma and ALP mass effects at low energies, the QED vacuum polarization and CMB photon dispersion effects at high energies can be neglected. In the case where $\lambda_a \lesssim L$ the simple sharp edges model model to describe domain-like magnetic fields gives unphysical results because the photon/ALP beam becomes sensitive to the discontinuities of the magnetic field components: in this case, more sophisticated models must be used [e.g. 250, 254, 255]. *In the case of the jet QED effects are quite important and for energies $E \gtrsim O(100 \text{ GeV})$ they strongly reduce the reported value of $\lambda_a$. **A similar fact happens in the extragalactic space because of the ALP mass term for $E \lesssim O(100 \text{ GeV})$ and because of the CMB photon dispersion on the CMB for $E \gtrsim O(5 \text{ TeV})$ [256].

<table>
<thead>
<tr>
<th>Astronomical object</th>
<th>$B$</th>
<th>$L$</th>
<th>$\lambda_a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGN jet</td>
<td>$0.1 \rightarrow 5 \text{ G}$</td>
<td>none</td>
<td>$O(0.01 \rightarrow 0.5 \text{ pc})^*$</td>
</tr>
<tr>
<td>Radio lobes</td>
<td>$10 \mu\text{G}$</td>
<td>$10 \text{kpc}$</td>
<td>$O(10 \text{kpc})$</td>
</tr>
<tr>
<td>Spiral galaxies</td>
<td>$7 \mu\text{G}$</td>
<td>$10 \text{kpc}$</td>
<td>$O(10 \text{kpc})$</td>
</tr>
<tr>
<td>Starburst galaxies</td>
<td>$50 \mu\text{G}$</td>
<td>$10 \text{kpc}$</td>
<td>$O(1 \text{kpc})$</td>
</tr>
<tr>
<td>Elliptical galaxies</td>
<td>$5 \mu\text{G}$</td>
<td>$150 \text{pc}$</td>
<td>$O(10 \text{kpc})$</td>
</tr>
<tr>
<td>Cluster</td>
<td>$5 \mu\text{G}$</td>
<td>$10 \rightarrow 100 \text{kpc}$</td>
<td>$O(10 \text{kpc})$</td>
</tr>
<tr>
<td>IGM</td>
<td>$&lt; 1.7 \text{nG}$</td>
<td>$0.1 \rightarrow 10 \text{Mpc}$</td>
<td>$O(50 \text{Mpc})^\dagger$</td>
</tr>
<tr>
<td>Milky Way</td>
<td>$5 \mu\text{G}$</td>
<td>$10 \text{kpc}$</td>
<td>$O(10 \text{kpc})$</td>
</tr>
</tbody>
</table>

pairs [257, 258]. In intergalactic space, these lower-energy photons originate from the extragalactic background light (EBL), the background radiation field which encompasses the stellar emission integrated over the age of the Universe, and the emission absorbed and re-emitted by dust [see 259, 260, for reviews]. Due to strong foreground emission in the solar system in the EBL wavelength range, the EBL is extremely difficult to measure directly [261]. Within the AGN jet, the radiation fields from the dusty torus, the accretion disk, or optical emission from ionized clouds (the so-called broad line region, BLR) can cause the attenuation [e.g. 262]. The presence of the BLR is usually associated with a sub-class of AGNs, so called flat-spectrum radio quasars (FSRQs). Basically, because of photon-ALP oscillations, the photon acquires a “split personality”: sometimes it travels as an ordinary photon and gets absorbed by the EBL, but sometimes it travels as an ALP which does not interact with EBL photons. As a consequence, the effective optical depth, $\tau_{\text{eff}}$ is smaller than the optical depth as evaluated by conventional physics. Since the photon survival probability is given by $P_{\gamma\rightarrow\gamma}^{\text{ALP}} = \exp[-\tau_{\text{eff}}]$, even a small decrease of the effective optical depth can give rise to a large enhancement in photon flux.

The exponential attenuation of the initial gamma-ray flux is described with the optical depth $\tau$, which increases monotonically with the primary gamma-ray energy, the source distance (in the case of absorption on the EBL), and the photon density of the radiation field. ALPs are not absorbed during their propagation and if a sizeable amount of photons converts into ALPs, which then reconvert back to photons close to Earth, a boost of the expected gamma-ray flux is expected [e.g. 247, 256, 263, 264].
7.2 Hints for anomalies in the gamma-ray opacity?

Several authors have found evidence that state-of-the-art EBL models over-predict the attenuation of gamma rays using published data points of AGN spectra obtained with imaging air Cherenkov telescopes (IACTs), which measure gamma rays above energies of $\sim 50$ GeV. Such an over-prediction would manifest itself through a hardening of the AGN spectra. For example, if the observed spectrum at low energies can be described with a power law, $dN/dE \propto E^{-\Gamma_{\text{low}}}$ a spectral hardening would mean that at high energies $dN/dE \propto E^{-\Gamma_{\text{high}}}$ with $\Gamma_{\text{high}} < \Gamma_{\text{low}}$. Such a behavior is in general not expected from standard gamma-ray emission scenarios (although specific models can produce such spectra [e.g. 265]), especially if the hardening $\Delta \Gamma = \Gamma_{\text{low}} - \Gamma_{\text{high}}$ correlates with increasing optical depth for several sources. Such a correlation has indeed been found and ALPs have been proposed as a possible explanation [247, 266–273]. An over-predicted EBL attenuation should also lead to a correlation between fit residuals and the optical depth when smooth concave, i.e. non-hardening, functions are assumed for the emitted AGN spectra. A $4\sigma$ indication for this effect has been found [274, 275], and ALP parameters reducing this tension were derived [276] (see the “T-Hint” labelled region in Fig. 14). Furthermore, using recent EBL measurements with the CIBER experiment, which suggest a larger attenuation than current EBL models, the authors of Ref. [277] found that ALPs can improve the fits of IACT spectra when again concave intrinsic spectra are assumed (region on top of the “T-Hint” region in Fig. 14). Instead of ALPs, these evidences have also been interpreted as evidence for particle cascades initiated by ultra-high energy cosmic rays [278] or a correlation between AGN lines of sight with cosmic voids [279].

However, recent analyses could not confirm the above correlations for a spectral hardening. Extending the IACT data sample of Ref. [276], the authors of Ref. [280] did not find a correlation of fit residuals with the optical depth. Furthermore, when including systematic uncertainties such as the IACT energy resolution, no spectral hardening between IACT spectra and spectra measured at lower gamma-ray energies with the Large Area Telescope (LAT) on board the Fermi satellite could be found [281]. Using Fermi-LAT data alone, no spectral hardening as a function of redshift (or equivalently the optical depth) was found in a recent analysis [282].

It has also been hypothesized that photon-ALP conversions could be responsible for IACT observations of FSRQs above 100 GeV. If gamma rays are produced close to the central super massive black hole, their flux should be severely attenuated due to the interaction with the radiation fields within the jet as mentioned above. Yet, a number of FSRQs has indeed been observed [e.g. 283–285] and it has been shown that the inclusion of ALPs can reproduce the observed spectra for ALP parameters $g_{11} \sim 1$ and $m_{\text{neV}} \lesssim 10$ [264]. One has to keep in mind, though, that astrophysical mechanisms could also produce the gamma-ray emission beyond the BLR and thus circumvent the pair production and the attenuation [e.g. 286, 287]. But it was the dissatisfaction of the ad hoc nature of these astrophysical attempted explanations that led to the above ALP-based proposal. In addition for BL Lacs (a sub-class of AGNs), by combining all the magnetic environments crossed by the photon/ALP beam – namely BL Lac jet, host galaxy, extragalactic space and Milky Way – it is possible to infer important predictions for BL Lacs spectra which present peculiar observable features induced by photon-ALP oscillations: (i) the oscillatory behavior of the energy spectrum (ii) photon excess above 20 TeV [288].

\footnote{See also \url{http://tevcat.uchicago.edu/}.
As shown in Fig. 14, IAXO is sensitive to the entire parameter space where ALPs have been proposed to alter the gamma-ray transparency. Therefore, IAXO measurements could give the definite answer on this matter.

7.2.1 Statistical distributions of X-ray source luminosities

While this part does not deal specifically with opacity, we mention this effect here as another possible hint for a light ALP (∼ 10 neV) with the same coupling as what is discussed in this section (10^{-11} \text{ GeV}^{-1}). In the strong mixing regime, it is easy to show that the average luminosity of a photon beam is 2/3 that without mixing [289]. It means that the luminosity of X-ray sources would be actually higher than what we observe. As all source would be concerned, this effect based on average luminosities would not be observable. However, photon-ALP mixing would also affect higher moments of the luminosity distribution of the sources. By studying these higher moments, the authors of Ref. [290] showed an anomaly compatible with ALPs in the same parameter space region as mentioned in the previous subsection. It has later been noted that this effect could be due to outliers in the used source catalog [291] so it could be a selection-bias-induced fake signal.

7.3 Constraints from high-energy astronomy and remaining parameter space

Searches for spectral distortions in gamma-ray spectra have already constrained the parameter space where ALPs could explain a reduced transparency. Observations with the H.E.S.S. telescopes of one blazar, i.e. an AGN with its jet closely aligned to the line of sight, have led to the exclusions labelled “H.E.S.S.” in Fig. 14 under the assumption that the blazar is located in a Galaxy group that harbours a magnetic field of $B = 1 \mu\text{G}$ [292]. Further constraints were derived using Fermi-LAT observations of NGC 1275, the central AGN of the Perseus galaxy cluster [293]. This galaxy cluster could have a central magnetic field as large as 25 $\mu\text{G}$ [294]. The constraints are the strongest to date in the mass range between $0.5 \lesssim m_{\text{neV}} \lesssim 20$ (see Fig. 14). Similar analyses using X-ray observations of AGN in galaxy clusters have constrained lower mass ALPs (see below).

For lower ALP masses, strong constraints where derived from the non-observation of a gamma-ray burst from the core-collapse supernova (SN) SN1987A [295–297]. During the core collapse, gamma rays in the core could convert to ALPs in the electrostatic fields of ions and escape the explosion. If they convert back into gamma rays in the Galactic magnetic field, a gamma ray burst lasting tens of seconds could be observed in temporal co-incidence with the SN neutrino burst. Interestingly, if a Galactic SN occurred in the field of view of the Fermi LAT, a wide range of photon-ALP couplings could be probed for $m_{\text{neV}} \lesssim 100$ (see the red dashed line in Fig 14). These prospects for a detection rely on a simplified picture of the proton-neutron star nuclear medium [297] and encourage future refinements of the theoretically expected ALP flux from such objects [256, 298].

The next generation of gamma-ray observatories will allow deeper studies of ALP-induced effects and will have enough sensitivity to probe the whole parameter space relevant for the gamma ray transparency hints. In particular, analyses of Cherenkov Telescope Array (CTA) data will consist of searches for the spectral irregularities [see 299, for preliminary results], the reduced opacity [300], or spatial correlations between blazar spectra with the magnetic field of the Milky Way [301]. A similar correlation study could also be conducted with future gamma-ray observations of the HAWC or LHAASO observatories [302]. However, CTA will be fully operational within a decade, and the necessary accumulation of data could take years.
On the other hand, as shown on Fig. 14, IAXO will be sensitive to the relevant masses and coupling. In conclusion, light ALPs can have a strong impact on gamma-ray astronomy and IAXO will cover the whole relevant parameter space.

**Figure 14.** The IAXO potential in the low ALP mass part of the $g_{\alpha\gamma}$ vs. $m_\alpha$ plane. The hashed region indicates the R-parameter hint, discussed in section 6.1.3, in the case of ALPs interacting only with photons. In this case, this is known also as the HB-hint [212, 213, 219]. The region indicated with ”T-hint” is the transparency region, discussed in the section 7. ALPs with parameters in this region have been invoked to address the unexpected transparency of the Universe to very high energy photons [276, 277] and some anomalous redshift-dependence of AGN gamma-ray spectra [303]. Notice that the lower mass section of the hinted region has been excluded by the non-observation of gamma rays from SN 1987A [297] and, more recently, by the search for spectral irregularities in the gamma ray spectrum of NGC 1275 [293]. The region enclosed within the dashed red line, labelled Fermi SN prospects, shows the Fermi LAT potential to probe the ALP parameter space in case of a new nearby (galactic) SN explosion [304]. Finally, the figure shows the regions excluded by the analysis of the spectral distortions of X-ray point sources in galaxy clusters: the Chandra observations of the AGN in Hydra A [305], the Perseus cluster NGC1275 [306] (indicated in the figure as NGC1275 and to be distinguished from the region labelled Fermi NG1275), and the M87 AGN of the Virgo cluster [307]. Refer to the text for more details.
7.4 Conversion between X-ray photons and ALPs

Let us now describe further aspects of ALP physics that rely on the interconversion of ALPs and X-ray photons in the magnetic field of galaxy clusters.

- **Spectral distortions of X-ray point sources in galaxy clusters**
  
  Just as with gamma-ray observations, it is possible to search for ALPs by looking for oscillatory features in the spectra of arriving X-rays from AGN or quasars in or behind galaxy clusters – and conversely, the absence of such features can be used to place bounds on ALP parameter space.

  The main disadvantage of this method is that the actual magnetic field along the line of sight is unknown. For any one source, it is then never possible to exclude the possibility that the magnetic field configuration along the line of sight is particularly unfavourable for ALP-photon conversion. This motivates considering sufficiently broad energy ranges so that multiple features are expected, and observing multiple X-ray sources within distinct astrophysical environments.

  In X-rays, this method has been applied for the AGN in Hydra A [305], the central AGN of the Perseus cluster NGC1275 [306] (see above for gamma rays), the central M87 AGN of the Virgo cluster [307], and a variety of weaker quasars and AGNs in and behind clusters [308]. Of these sources, the two bright local AGNs NGC1275 and M87 are the most constraining, as both are bright AGNs with deep exposures at the heart of large cool-core galaxy clusters (which tend to have the highest magnetic fields). For reasonable and observationally supported values for the magnetic field structure within the Perseus and Virgo cluster, the absence of large spectral modulations was used in [306] and [307] to constrain $g_{a\gamma} \lesssim 1.5 \times 10^{-12}\text{GeV}^{-1}$ for ALP masses $m_a \lesssim 10^{-12}\text{eV}$. These bounds depend on the magnetic field model for the cluster. For magnetic fields weaker or stronger than assumed, the constraint scales inversely with the magnetic field.

- **Spectral distortions of the continuum thermal bremsstrahlung emission of galaxy clusters**

  The advantage of bright point sources is that they are sensitive to the magnetic field along a single line of sight, and so avoid effects of destructive interference when averaging over many different sightlines. However bright point sources are also rather rare.

  Another approach to searching for ALPs is to use instead the continuum emission from galaxy clusters. This arises as thermal bremsstrahlung from the intrachannel medium, the hot ($T \sim 2-8\text{keV}$) ionised plasma where around 90% of the baryons in a cluster are located. The thermal emission contains both continuum and line emission, and provides an excellent fit to the overall photon spectrum from clusters.

  As with point sources, the aim is to observe (or constrain) ALP-induced deviations from the thermal bremsstrahlung fit. In the presence of $g_{a\gamma} \sim 10^{11}\text{GeV}$, this thermal emission will experience significant spectral modulation along a single line of sight (as for point sources). Averaged over a large region of the cluster, the modulations from individual sightlines will average out. However, in the presence of significant ALP-photon conversion significant spectral distortions will be present on small scales [309].
• The 3.5 keV Line

One of the most interesting recent results in particle astrophysics has been the observation of an unexplained line at $E \sim 3.55$ keV [310, 311]. The line was found originally in observations of stacked samples of galaxy clusters and is not at an energy that corresponds to a known atomic line. There has been considerable interest around the possibility that this could arise from dark matter [312–320], while also various possible astrophysical explanations have been proposed.

If the line arises from dark matter, one difficulty with ‘standard’ interpretations (such as a sterile neutrino) is that the line is much stronger in clusters than in galaxies, and in particular is much stronger at the centre of the Perseus cluster than at any other location. This implies that the line is not sensitive only to the dark matter content, but also to some aspect of the astrophysical environment. One way this can arise is in models where the dark matter decays originally to a relativistic ALP with energy $E = 3.5$ keV, and the observed photon signal comes from conversion of this ALP in the magnetic field of the cluster [321]. Depending on the dark matter lifetime, the line signal can be reproduced for axion-photon couplings of order $10^{-15}$ GeV$^{-1} \lesssim g_{a\gamma} \lesssim 10^{-10}$ GeV$^{-1}$ for ALP masses $m_a < 10^{-12}$ eV. Interestingly, IAXO will be able to test a relatively large region of this parameter space.

The morphology of the 3.5 keV signal can be explained by ALP-photon conversion. The stronger signal in clusters would then arise from the larger and more extended magnetic fields present in clusters compared to galaxies. As Perseus is a close cool-core cluster, observations of Perseus only cover the central region with a large magnetic field (as $B \propto n_e(r)^{1/2}$, the magnetic field is significantly enhanced in the central high-$n_e$ cool core). The extremely strong signal in the centre of Perseus would then be a consequence of the efficiency of ALP-photon conversion in strong magnetic fields.

If this line is found to arise from new physics, ALPs then offer a way to reproduce the unusual morphology.

8 Updated sensitivity prospects for IAXO

The International Axion Observatory (IAXO) is a next generation axion helioscope that aims at a substantial step forward, of more than one order of magnitude, in sensitivity to $g_{a\gamma}$ with respect to current best limits. The baseline layout of the experiment is based on the enhanced axion helioscope studied in [322] and sketched in Fig. 15. In this configuration the entire cross sectional area of the magnet is equipped with X-ray focusing optics to increase the signal-to-noise ratio. When the magnet is pointing to the Sun, solar axions are converted into photons, that are focused and detected by low background X-ray detectors placed at the focal point of the telescopes. In this way a larger magnet aperture $A$ translates directly into the figure of merit of the experiment, as a larger signal is expected while the detector background remains low. This opens the way for new large-volume magnet configurations, like the ones of superconducting detector magnets typically developed for high energy physics.

A useful figure of merit (FOM) was introduced in [322] to easily gauge the relative importance of the various experimental parameters affecting the sensitivity of a helioscope:

$$f \equiv f_M f_{DO} f_T$$ (8.1)
Solar axions are converted into photons by the transverse magnetic field inside the bore of a powerful magnet. The resulting quasi-parallel beam of photons of cross sectional area $A$ is concentrated by an appropriate X-ray optics onto a small spot area $a$ in a low background detector. The envisaged design for IAXO, shown in Fig. 16, includes eight such magnet bores, with their respective optics and detectors.

where we have factored the FOM to explicitly show the contributions from various experimental subsystems: magnet, detectors and optics, and tracking (effective exposure time of the experiment)

$$f_M = B^2 L^2 A \quad f_{DO} = \frac{\epsilon_d \epsilon_o \epsilon_t}{\sqrt{b a}} \quad f_T = \sqrt{\epsilon_t t}.$$

(8.2)

where $B$, $L$ and $A$ are the magnet field, length and cross sectional area, respectively. The efficiency $\epsilon = \epsilon_d \epsilon_o \epsilon_t$, being $\epsilon_d$ the detectors’ efficiency, $\epsilon_o$ the optics throughput or focusing efficiency (it is assumed that the optics covers the entire area $A$), and $\epsilon_t$ the data-taking efficiency, i.e. the fraction of time the magnet tracks the Sun (a parameter that depends on the extent of the platform movements). Finally, $b$ is the normalized (in area and time) background of the detector, $a$ the total focusing spot area and $t$ the duration of the data taking campaign. The expressions in (8.2) assume some simplifications, like that $B$ is constant in all the volume of the magnet, or that $b$ and $\epsilon$ are constant throughout the energy range of interest. Generalizations of the figure of merit for arbitrary distributions of the parameters are straightforward, and the simplified versions quoted here are in any case useful to see the main dependencies.

Following this metric, IAXO thus aims at a $f$ more than a factor $10^4$ larger that its predecessor CAST. The conceptual design report (CDR) of the experiment [6, 7] demonstrates the technical feasibility of this step in sensitivity, and its main parameters will be reviewed below. Previous implementations of axion helioscopes have relied on existing equipment that were originally built for other experimental purposes. On the contrary, IAXO subsystems (magnet, optics and detectors) are entirely conceived and optimized for solar axion detection. The design prescriptions have been to rely on state-of-the-art technologies scaled up within realistic limits, i.e. no R&D is needed to reach the stated experimental parameters. The baseline sensitivity projections studied below will refer to the experimental parameters anticipated in the CDR. More recently, the realization of an intermediate experimental
stage, BabyIAXO, featuring a scaled-down prototype version of the magnet, optics and detectors, is being considered. This new activity brings the opportunity to explore potential improvements over the CDR figure of merit of the full infrastructure, potentially resulting on an upgraded sensitivity scenario for IAXO. The sensitivity projections shown below and in the plots all throughout this paper refer to these three experimental scenarios: BabyIAXO, IAXO-baseline and IAXO-upgraded. The experimental parameters corresponding to each of them are listed in table 5, and the prescriptions to define them are discussed in the following.

The central component of IAXO is therefore a large superconducting magnet. Contrary to previous helioscopes, IAXO’s magnet will follow a toroidal multibore configuration [323], to efficiently produce an intense magnetic field over a large volume. The baseline layout of the IAXO magnet is a 25 m long and 5.2 m diameter toroid assembled from 8 coils, and generating effectively 2.5 T average (5 T maximum) in 8 bores of 600 mm diameter. The toroid’s stored energy is 500 MJ. The design is inspired by the ATLAS barrel and end-cap toroids [324, 325], the largest superconducting toroids built and presently in operation at CERN. The superconductor used is a NbTi/Cu based Rutherford cable co-extruded with Aluminum, a successful technology common to most modern detector magnets. Figure 16 shows the conceptual design of the overall infrastructure [7]. IAXO needs to track the Sun for the longest possible period. For the rotation around the two axes to happen, the 250 tons magnet is supported at the centre of mass by a system also used for very large telescopes. The necessary magnet services for vacuum, helium supply, current and controls are rotating along with the magnet.

Each of the eight magnet bores is equipped with X-ray telescopes that rely on the high X-ray reflectivity on multi-layer surfaces at grazing angles. By means of nesting, that is, placing concentric co-focal X-ray mirrors inside one another, large surface of high-throughput optics
can be built. The IAXO collaboration envisions using optics similar to those used on NASA’s NuSTAR [326], an X-ray astrophysics satellite with two focusing telescopes that operate in the 3 - 79 keV band. The NuSTAR’s optics, shown in Fig. 17, consists of thousands of thermally-formed glass substrates deposited with multilayer coatings to enhance the reflectivity above 10 keV. For IAXO, the mirror arrangement and coatings are designed to match the solar axion spectrum. The conceptual design of the IAXO telescopes [327] can be seen on the right of Fig. 17. As proven in [6, 7, 327], this technology can equip the aperture area of IAXO magnet bores with focusing efficiency of around 0.6 and focal spot areas of about 0.2 cm$^2$.

At the focal plane in each of the optics, IAXO will have low-background X-ray detectors. The baseline technology for these detectors are small gaseous chambers read by pixelated planes of micro-mesh gas structure (Micromegas) [328] manufactured with the microbulk technique [329]. These detectors have been successfully used and developed in CAST and other low background applications [330]. The latest CAST detectors have achieved background levels of $10^{-6}$ counts keV$^{-1}$ cm$^{-2}$ s$^{-1}$ with prospects for improvement down to $10^{-7}$ or even $10^{-8}$ counts keV$^{-1}$ cm$^{-2}$ s$^{-1}$ [331]. These background levels are achieved by the use of radiopure detector components, appropriate shielding, and offline discrimination-algorithms on the 3D event topology in the gas registered by the pixelised readout. A pathfinder system combining an X-ray optics of the same type as proposed for IAXO and a Micromegas detector has been operated in CAST during 2014 and 2015 with the expected performance [332]. Alternative or additional technologies are being considered to complement the capabilities of Micromegas detectors or to extent them in specific aspects. GridPix detectors, similar to Micromegas detectors but built on a small CMOS pixelized readout [333], enjoy very low energy threshold down to the tens of eV, and thus are of interest for the search of specific solar axion production channels lying at lower energies, like the ones mediated by the axion-electron coupling. Silicon Drift Detectors (SDD) offer better energy resolution with flexible and cost-effective implementations [334, 335]. Finally, bolometric detectors like Magnetic

![Figure 17](image.jpg)

**Figure 17.** Left: the NuSTAR X-ray telescope, with optics very similar to that proposed for IAXO. Right: conceptual design of the X-ray optics needed for IAXO.
Metallic Calorimeters (MMC) [336], or Transition Edge Sensors (TES) [337], enjoying much lower energy threshold and energy resolution, are also under consideration. An R&D activity is ongoing to assess the low-background capabilities of all this technologies, and their suitability as detectors for IAXO. Low energy threshold and resolution are also of interest to extract model parameters in case of a positive signal (see next section).

The values considered for the main experimental parameters are listed in table 5 for three different scenarios. The column labelled IAXO baseline shows the set of values anticipated in our CDR for IAXO, and are considered a realistic estimation within current state of the art. We refer to [7] for a careful justification of those values. However, we now envision, as a first step, the realization of an intermediate stage, called BabyIAXO, featuring a scaled-down prototype magnet, as well as prototype optics and detectors, all representative of the final systems. BabyIAXO will serve as a testbed for all technologies in the full IAXO, but at the same time will deliver competitive physics. The BabyIAXO magnet is conceived as a common-coil dipole magnet, with two bores placed in between the coils. The two superconducting coils are 10-m long, but otherwise they enjoy quite similar engineering parameters than the ones proposed for the final IAXO toroid. The BabyIAXO magnet bores will have a diameter of 70 cm and each one will be equipped with one full detection line, optics and detector, of similar dimensions than the final IAXO systems. In the baseline configuration of the experiment, one of the BabyIAXO lines will host a newly built IAXO optics prototype similar to the one described above (potentially extended to 70 cm diameter), while the second one is expected to host an existing XMM spare optics [338]. A technical description of the BabyIAXO system will be object of another dedicated publication. Figure 18 shows a conceptual design of the BabyIAXO setup, whose expected experimental parameters are listed in the first column of table 5.

In addition, the experience with BabyIAXO is expected to test enhanced design choices that eventually lead to improved IAXO FOM values, especially $f_M$, beyond the ones anticipated in our CDR. The column labelled IAXO upgraded represent this possible improved scenario. Collectively they constitute a factor $\sim$10 better $f$ (of which a factor 4 better $f_M$) than the baseline scenario. However, the extent to which those possible improvements may get eventually realized is tentative, and this scenario must be considered as a desirable target.
whose feasibility will be studied as part of the BabyIAXO stage.

The IAXO sensitivity projections shown in Fig. 19 as well as in all plots throughout this paper refer to the three scenarios of table 5. They have been computed by means of Monte Carlo simulation of the expected background counts in the optics spot area, computation of the likelihood function and subsequent derivation of the 95% upper limit on the $g_{a\gamma}$ assuming no detected signal. The calculation is repeated for a range of $m_a$ values in order to build full sensitivity lines in the $(g_{a\gamma}, m_a)$--plane. For the purpose of this analysis, and following similar prescriptions as in [6], detector background and efficiency are assumed flat with energy down to arbitrarily low energies. The axion-photon conversion in the magnet is approximated to the conversion in an homogeneous field, and the focusing effect is reduced to the equivalent effect of enhancing the signal-to-noise ratio due to the fact of confining the signal counts into the spot area. For all scenarios, an additional buffer gas data taking phase is considered, giving rise to the extended step-wise sensitivity line at high masses (0.01-0.25 eV). This data taking phase is composed by a number of overlapping gas density steps spanning the desired mass range. While in previous projections [322] an equal exposure time is assigned to each step, resulting in a exclusion line more or less horizontal in $g_{a\gamma}$, in this case the exposure time is different for every step and adjusted to obtain a sensitivity down to the DFSZ $g_{a\gamma}$ (KSVZ for the BabyIAXO case) for every $m_a$ value. Of course different prescriptions to distribute the total exposure time among the steps are possible, depending on the motivation, e.g. to go for larger $m_a$ or lower $g_{a\gamma}$. The total exposure of this second phase is the same as the vacuum phase ($t$ in Table 5).

### 8.1 Measuring axion parameters

In the case of a positive signal, and depending on the axion parameters, IAXO will be able to extract information on its mass $m_a$ and relative coupling with electrons and photons, potentially providing invaluable information on the underlying theoretical model. If the axion mass is above around 0.02 eV, axion-photon oscillations destroy the coherence of the conversion along the magnet length. This coherence can be restored if the conversion takes place in a buffer gas with density matching the axion mass, this being the rationale of the gas phase of both IAXO and BabyIAXO. If a positive detection happens during the gas scanning

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Units</th>
<th>BabyIAXO</th>
<th>IAXO baseline</th>
<th>IAXO upgraded</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B$</td>
<td>T</td>
<td>~2</td>
<td>~2.5</td>
<td>~3.5</td>
</tr>
<tr>
<td>$L$</td>
<td>m</td>
<td>10</td>
<td>20</td>
<td>22</td>
</tr>
<tr>
<td>$A$</td>
<td>m$^2$</td>
<td>0.77</td>
<td>2.3</td>
<td>3.9</td>
</tr>
<tr>
<td>$f_M$</td>
<td>T$^2$m$^4$</td>
<td>~230</td>
<td>~6000</td>
<td>~24000</td>
</tr>
<tr>
<td>$b$</td>
<td>keV$^{-1}$ cm$^{-2}$ s$^{-1}$</td>
<td>$1 \times 10^{-7}$</td>
<td>$10^{-8}$</td>
<td>$10^{-9}$</td>
</tr>
<tr>
<td>$\epsilon_d$</td>
<td></td>
<td>0.7</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>$\epsilon_o$</td>
<td></td>
<td>0.35</td>
<td>0.7</td>
<td>0.7</td>
</tr>
<tr>
<td>$a$</td>
<td>cm$^2$</td>
<td>$2 \times 0.3$</td>
<td>$8 \times 0.15$</td>
<td>$8 \times 0.15$</td>
</tr>
<tr>
<td>$\epsilon_t$</td>
<td></td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>$t$</td>
<td>year</td>
<td>1.5</td>
<td>3</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 5. Indicative values of the relevant experimental parameters representative of BabyIAXO as well as IAXO, both the baseline and upgraded scenarios, based on the considerations explained in the text.
phase, the gas density provides the axion mass. But also in the vacuum phase and for lower mass values the onset of these spectral oscillations can be observed and used to determine the axion mass. As studied in [339], provided the signal is measured with sufficient statistics, \( m_a \) values as low as \( 3 \times 10^{-3} \) eV could be determined by this method.

Moreover, if the axion signal is composed by significant fractions of Primakoff and ABC solar axions, the combined spectral fitting can provide independent estimations of \( g_{a\gamma} \) and \( g_{ae} \) [340]. This combined determination works in areas of parameter space particularly well motivated by the stellar cooling anomalies. To exploit these capabilities high-resolution and low-threshold detectors are preferred, because part of the ABC solar axion spectrum lies at lower (< 1 keV) energies, and also because it features several high-resolution peaks [25]. Devices like the bolometric detectors described above, under consideration for IAXO, could play a major role in a post-discovery high-precision measurement campaign.

8.2 Direct search for DM axions with IAXO and BabyIAXO

Although the focus of this paper has been the physics potential of IAXO in its baseline configuration as a helioscope, the (Baby)IAXO magnet constitutes a remarkable infrastructure to implement additional setups to search for axions/ALPs in alternative ways. A most ap-

![Figure 19. Sensitivity prospects of BabyIAXO and IAXO (semitransparent regions) in the overall context of other experimental and observational bounds. We refer to [5] for details on the latter.](image)
pealing option is to implement DM axion detectors that could exploit the particular features of the IAXO magnet. The basic idea of the axion-haloscope-technique proposed in [29] is to combine a high-$Q$ microwave cavity inside a magnetic field to trigger the conversion of axions of the DM halo into photons. A straightforward implementation of this concept in one bore of the IAXO magnet gives competitive sensitivity [341] even using relatively conservative values for detection parameters, thanks to the large $B^2V$ available, in the approximate mass range of $0.6 - 2 \mu eV$. Another recent concept proposes the use of a carefully placed pick-up coil to sense the small oscillating magnetic field that could be produced by the DM axion interacting in a large magnetic volume [342]. This technique is best suited for even lower $m_a$ and is better implemented in toroidal magnets [343, 344]. The large size and toroidal geometry of IAXO suggest that it could host a very competitive version of this detection technique.

However, as can be seen from Fig.6, there is a strong motivation to explore the mass range for QCD axion DM well above $10^{-5}$ eV. The difficulty in searching in the “high-mass” range can be understood from the fact that the figure of merit of scanning with haloscopes for axion DM scales with the square of the cavity volume times the $Q$ factor of the cavity. Most existing setups use solenoidal magnets and cylindrical cavities. The diameter of the cylinder sets the frequency scale of the resonance and thus the axion mass scale which the experiment is sensitive to. Thus going to high mass means going to small diameters. In addition, the cavity quality factor $Q$ typically decreases for smaller cavities.

To tackle higher masses, different strategies are being developed in the community [5]. For instance, to compensate the loss in $V$ one can go to very strong magnetic fields and/or use superconducting cavities to keep $Q$ large. Another avenue is to decouple $V$ from the resonant frequency and go for large $V$ structures resonating at high frequency. Several strategies in this direction are being explored. One of them, currently being tested in exploratory setups at the CAST experiment at CERN [345, 346], employ long rectangular cavities [347]. The particular advantage is that the volume can be kept very large (long cavity), while the resonance frequency can be rather high, through the usage of relatively thin cavities. Given the size of the (Baby)IAXO magnet, it could host a multitude of rectangular cavities. The implementation of one those concepts in IAXO is under consideration, but is out of the scope of the present paper. It clearly deserves serious study and will be the object of one or more future publications.

9 Discussion and conclusions

Axions and more generic ALPs are recently receiving an increasing attention from the experimental community. This is in part due to the lack of experimental confirmation of the WIMP dark matter paradigm in recent searches both at colliders and direct detection experiments. Axions are known to be attractive dark matter candidates, but despite continued experimental activity almost since their proposal more than 30 years ago, most of the allowed parameter space remains largely unexplored so far. Recently a plethora of ideas and new small-scale initiatives are being put forward [5], under the assumption that dark matter is entirely made of axions, and mostly addressed to the to low-mass range $m_a \sim 10^{-6} - 10^{-4}$ eV. However, as described in section 3, axion (or ALP) dark matter can be also realized in other mass ranges, in particular at higher mass values, for which dark matter experiments are increasingly difficult to realize. In addition, axions/ALPs may compose only a subdominant fraction of the galactic dark matter density, a possibility that may have a particular theoret-
ical interest [348]. In general, a signal in a dark matter experiment is proportional to $g_{a\gamma}^2 \rho_a$, being the local galactic axion density $\rho_a$ degenerate with $g_{a\gamma}$. Experiments not relying on the axion being the dark matter are needed to break such degeneracy. In any case, axions are strongly motivated by theory, without relying on the dark matter question. They constitute our only compelling solution to the strong-CP problem in the Standard Model. As reviewed in section 2, axions and ALPs are also very generic low energy signatures of high-energy completions of the SM in extra dimensions.

Experiments that invoke both the production and detection of ALPs entirely in the laboratory are the least model-dependent. Among these experiments, only the ALPS-II project will reach a sensitivity beyond current astrophysical and experimental bounds on $g_{a\gamma}$, for ALP masses below $m_a < 10^{-4}$ eV. In this context, the search for solar axions constitutes a good compromise between model-independency (as the emission of axions by the Sun is a robust prediction of any axion model), and intensity of the source axion flux. The IAXO project embodies the technical know-how accumulated in previous realizations of the axion helioscope concept, most in particular the CAST experiment at CERN, extended to a much larger scale. While CAST has been the first axion helioscope reaching a sensitivity to $g_{a\gamma}$ comparable to astrophysical bounds, IAXO will largely advance well beyond them. In particular, and as a summary of the material reviewed in previous sections:

- IAXO will cover a large fraction of unexplored ALP parameter. In particular, it will probe a large range of QCD axion models in the mass range $m_a \sim 1$ meV – 1 eV.

- Most of this region is not attainable by any other experimental technique, stressing the complementarity of IAXO in the wider axion experimental landscape.

- IAXO will comfortably cover the range of $g_{a\gamma}$ invoked as ALP solutions to the possible anomalies observed in the propagation of high energy photon over astronomical distances, fully testing this hypothesis.

- IAXO will cover a large fraction of the region of parameter space invoked as possible solutions to the anomalies observed in the cooling of several stellar systems, both involving $g_{a\gamma}$ and $g_{ae}$.

- The parameter space to be covered by IAXO could contain a viable dark matter candidate. In particular, post-inflation models with $N_{DM} > 1$ allows for QCD axion with $m_a$ in this range to account for the totality of the DM density. In addition, that very $(g_{a\gamma}, m_a)$ region is also suggested by ALP-miracle models recently proposed in which ALPs can account for both DM and inflation.

- Combined with a positive detection in a haloscope, it will break the $g_{a\gamma}^2 \rho_a$ degeneracy and determine the local density of the galactic axionic dark matter.

- Although not covered in this paper, the IAXO infrastructure could be used to host additional experimental setups, in particular microwave cavities or other devices able to directly search for DM axions.

IAXO will play a prominent role in the new generation of axion experiments currently under proposal or preparation. It is highly complementary with the other experimental frontiers (laboratory and relic axions). Collectively a relatively large fraction of the most motivated parameter space for axions will be explored, potentially leading to substantial progress in the low energy frontier in the coming years.
Acknowledgments

We acknowledge support from the the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme, grant agreement ERC-2017-AdG-788781 (IAXO+) as well as ERC-2018-StG-802836 (AxScale), and from the Spanish MINECO under grant FPA2016-76978-C3-1-P. Part of this work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. JR is supported by the Ramon y Cajal Fellowship 2012-10597, the grant FPA2015-65745-P (MINECO/FEDER), the EU through the ITN “Elusives” H2020-MSCA-ITN-2015/674896 and the Deutsche Forschungsgemeinschaft under grant SFB-1258 as a Mercator Fellow. We also acknowledge partial support of RFBR (grants 17-02-00305A, 16-29-13014 ofi-m).

References


[82] PLANCK collaboration, N. Aghanim et al., Planck 2018 results. VI. Cosmological parameters, [1807.06209].


[150] R. Z. Ferreira and A. Notari, *Observable windows for the QCD axion through \( N_{\text{eff}} \)*, 1801.06090.


[188] D. Croon and V. Sanz, Saving Natural Inflation, JCAP 1502 (2015) 008, [1411.7809].


[228] O. Straniero, I. Domínguez, M. Giannotti and A. Mirizzi, Axion-electron coupling from the RGB tip of Globular Clusters, in Proceedings, 13th Patras Workshop on Axions, WIMPs and


[241] M. Giannotti, ALP hints from cooling anomalies, in *Proceedings, 11th Patras Workshop on Axions, WIMPs and WISPs (Axion-WIMP 2015)*: Zaragoza, Spain, June 22-26, 2015, pp. 26–30, 2015, [1508.07576], DOI.


G. Galanti, F. Tavecchio, M. Roncadelli and C. Evoli, *Photon-ALP oscillations from a blazar to us up to 1000 TeV*, 1811.03548.


S. Aune et al., *X-ray detection with Micromegas with background levels below 10^{-6} keV^{-1} cm^{-2} s^{-1}*, JINST 8 (2013) C12042, [1312.4282].

F. Aznar et al., *A Micromegas-based low-background x-ray detector coupled to a slumped-glass telescope for axion research*, JCAP 1512 (2015) 008, [1509.06190].


M. Silva-Feaver et al., *Design Overview of the DM Radio Pathfinder Experiment*, 1610.09344.

A. I. Melcón et al., *Axion Searches with Microwave Filters: the RADES project*, JCAP 1805 (2018) 040, [1803.01243].
