

**Axions, instantons, and the lattice**

Michael Dine, Patrick Draper, Laurel Stephenson-Haskins, and Di Xu

*Santa Cruz Institute for Particle Physics and Department of Physics,  
University of California at Santa Cruz, Santa Cruz, California 95064, USA  
and Amherst Center for Fundamental Interactions, Department of Physics,  
University of Massachusetts, Amherst, Massachusetts 01003, USA*

(Received 4 September 2017; published 3 November 2017)

If the QCD axion is a significant component of dark matter, and if the Universe was once hotter than a few hundred MeV, the axion relic abundance depends on the function  $\chi(T)$ , the temperature-dependent topological susceptibility. Uncertainties in this quantity induce uncertainties in the axion mass as a function of the relic density, or vice versa. At high temperatures, theoretical uncertainties enter through the dilute instanton gas computation, while in the intermediate and strong coupling regime, only lattice QCD can determine  $\chi(T)$  precisely. We reassess the uncertainty on the instanton contribution, arguing that it amounts to less than 20% in the effective action, or a factor of 20 in  $\chi$  at  $T = 1.5$  GeV. We then combine the instanton uncertainty with a range of models for  $\chi(T)$  at intermediate temperatures and determine the impact on the axion relic density. We find that for a given relic density and initial misalignment angle, the combined uncertainty amounts to a factor of 2–3 in the zero-temperature axion mass.

DOI: [10.1103/PhysRevD.96.095001](https://doi.org/10.1103/PhysRevD.96.095001)**I. INTRODUCTION**

The axion remains a promising candidate for dark matter [1–3], perhaps more so as the window for conventional weakly interacting massive particles shrinks. Searches for axion dark matter are underway [4], and there are proposals for future experiments which could conceivably widen the search window substantially [5].

In the early Universe, the axion begins to oscillate coherently when its thermal mass,  $m_a(T)$ , becomes comparable to the Hubble scale. The axion mass is related to the QCD topological susceptibility by

$$m_a^2(T) f_a^2 = \chi(T),$$

$$\chi = \int d^4x \langle F\tilde{F}(x)F\tilde{F}(0) \rangle_T = \partial_\theta^2 F(\theta, T), \quad (1.1)$$

where, in the last expression,  $F(\theta, T)$  is the  $\theta$ -dependent free energy.

At high temperatures,  $F(\theta, T)$  can be calculated by standard instanton methods [6]. At low temperatures,  $F(\theta, T)$  is known from chiral perturbation theory, and in fact converges rapidly to its  $T = 0$  limit below the confining phase transition [7,8]. However, for plausible cosmologies and a range of axion parameters, it is the case that the axion starts to oscillate at intermediate temperatures,  $T \sim \text{GeV}$ , where  $\alpha_s$  is approaching strong coupling and neither calculation applies. Instead, one can hope to extract  $\chi(T)$  from lattice QCD.

Recently, there have been a number of papers reporting lattice calculations of  $\chi(T)$  at temperatures above the critical temperature, both in pure gauge theory/quenched approximations [8–10] and in QCD [7,11–16]. In some

cases, discrepancies are found at high temperatures compared with the dilute instanton gas prediction. The free energies found in [12,13] differ by about an order of magnitude from the leading-order semiclassical result above the GeV scale, while the computation and extrapolation obtained in [7,14] (see also [17]) differs by many orders of magnitude (although systematic effects are not under control in the extrapolation, and thus the level of compatibility with the controlled high-temperature computations in [12,13] is unclear.) These results suggest a level of uncertainty in the microscopic parameters (the zero-temperature axion mass  $m_a$  and possibly the misalignment angle  $\theta_0$ ) required to achieve a given relic density  $\Omega$ .

Here we will assess the theoretical uncertainty on  $\Omega(m_a, \theta_0)$  in analytic computations and compare with lattice results. Our analysis has two prongs. First, in Sec. II, we reexamine the uncertainty on the leading-order semiclassical result for the free energy above the GeV scale. We consider a number of possible sources of error, the most important of which are likely to be corrections to the effective action that shift the IR cutoff on instanton sizes. However, in contrast to suggestions in the literature, we argue that infrared divergences that plague ordinary finite-temperature perturbation theory are not numerically relevant in the instanton effective action, and that the size of higher-order corrections can be reasonably estimated. As a result, the topological susceptibility has known low- and high-temperature asymptotics which appear compatible with [12,13]. Then, in Sec. III, we introduce a family of models for the topological susceptibility that interpolate through the region where neither analysis is reliable. In Sec. IV, we compute the relic density over this range of models, including the instanton uncertainty at the high-temperature

boundary and uncertainties in QCD parameters. In this way we determine the sensitivity of  $\Omega(m_a, \theta_0)$  to theoretical uncertainties. We find the sensitivity is limited, and overall the axion mass prediction from analytical methods appears robust at the level of a factor of 2–3. We comment on the implications of these results in Sec. V and conclude.

## II. THEORETICAL UNCERTAINTIES ON THE INSTANTON CONTRIBUTION TO THE FREE ENERGY

### A. The standard computation

At high temperatures, the  $\theta$  dependence of the free energy is controlled by instantons (Gross, Pisarski, and Yaffe [6], henceforth GPY). Classically, even at finite temperature, there are instantons of all scale sizes. But at one loop, there are two sources of scale invariance violation: the usual ultraviolet divergences familiar in the zero-temperature theory, and the finite temperature itself. Both correct the effective action, rendering finite the scale size integral both at small and large  $\rho$ .

Heuristically, the latter effect is associated with the effective mass of the  $A_4$  field,

$$m_D^2 = \frac{1}{3}(g^2 T^2) \left( N + \frac{N_f}{2} \right). \quad (2.1)$$

GPY note that a term in the effective action  $\frac{1}{2g^2} m_D^2 A_4^2$  gives rise to a correction to the instanton action, for  $\rho \gg T^{-1}$  [and  $\alpha_s(\rho^{-1}) \ll 1$ ], proportional to  $\rho^2$ ,

$$\int d^4x \frac{1}{2g^2} m_D^2 A_4^2 = \frac{\pi^2}{2g^2} m_D^2 \rho^2. \quad (2.2)$$

Note the  $g^{-2}$  in front of  $m_D^2$ , reflecting the  $1/g^2$  in front of the whole action, and the fact that the actual screening length is of order  $\frac{1}{gT}$ . If this were the complete result for the correction to the effective action, the  $\rho$  integration for the free energy would take the form, in the case of three flavors,

$$F(T) \propto m_u m_D m_s \int \frac{d\rho}{\rho^2} (\Lambda\rho)^9 e^{-3\pi^2 \rho^2 T^2}. \quad (2.3)$$

The integral is finite and dominated by  $\rho \sim (\pi T)^{-1}$ .

Since the dominant scale is of order  $T^{-1}$ , the effective action cannot be expanded in powers of  $\rho$ ; in a derivative expansion of the background field effective action, terms of the form

$$\frac{g^2}{T^{n-2}} A_4(\vec{x}) \partial_{i_1} \dots \partial_{i_n} A_4(\vec{x}) \quad (2.4)$$

are all of the same order,  $g^2 T^2$ , in the instanton background.

GPY indeed computed the full one-loop determinant [6]. At small  $\rho$ , in particular, the above expression for the action is modified as follows:

$$\delta S = \frac{1}{3} \pi^2 \rho^2 T^2 (2N + N_f) - \frac{1}{18} \pi^2 \rho^2 T^2 (N - N_f). \quad (2.5)$$

For  $N_f = 0$ , for example, this is not parametrically smaller than the Debye screening term, though it is numerically smaller. At one loop, the complete expression for the free energy in the presence of a single instanton is given by [6]

$$F(\theta, T) = - \int \frac{d\rho}{\rho^5} \left( \frac{4\pi^2}{g^2} \right)^{2N} e^{-\frac{8\pi^2}{g^2(\rho)} C_N} \times \prod_{i=1}^{N_f} (\xi \rho m_i) e^{-1/3 \lambda^2 (2N + N_f) - 12A(\lambda) [1 + \frac{1}{6}(N - N_f)] + i\theta} \quad (2.6)$$

where

$$A(\lambda) = -\frac{1}{12} \ln(1 + \lambda^2/3) + \alpha(1 + \gamma \lambda^{-2/3})^{-8} \quad (2.7)$$

$$\begin{aligned} \lambda &= \pi \rho T & C_N &= 0.097163; & \xi &= 1.3388 \\ \alpha &= .01290 & \gamma &= 0.1586 \end{aligned} \quad (2.8)$$

and  $N_f = 3$  in temperature regimes where three quarks are excited. At a temperature of  $T = 1.5$  GeV and using a renormalization scale  $\mu = T$ , we obtain

$$F_0(1.5) = -3.7 \times 10^{-14} \text{ GeV}^{-4} \quad (2.9)$$

where the subscript indicates  $\theta = 0$ . Here we have used the program RunDec [18–20] to obtain  $\alpha_s(1.5 \text{ GeV}) \simeq 0.345$  with three active flavors.

This computation of the free energy is subject to certain theoretical uncertainties, including higher-order corrections sensitive to the UV cutoff, parametric uncertainties on  $\alpha_s$ , effects of heavier quarks, and higher-order corrections that modify the infrared cutoff on  $\rho$ . In the next subsection, we estimate the uncertainties from the first three of these sources. In our view there has been some confusion in the literature about the uncertainty associated with corrections to the  $\rho$  cutoff, which is plausibly the dominant source of uncertainty. We therefore devote a separate subsection to this source.

### B. UV-sensitive corrections, heavy quarks, and parametric uncertainties

In the previous section, we evaluated the one-loop expression for  $F$  at  $T = 1.5$  GeV with  $\mu = T$  and three flavors. Let us comment on a few of the knobs we can turn

in this calculation to obtain estimates of theoretical uncertainty.

- (a) Because the dominant instanton size is of order  $(\pi T)^{-1}$ ,  $\mu = \pi T$  is another natural choice for the renormalization scale.
- (b) At  $T = 1.5$  GeV,  $\pi T$  is substantially above the charm threshold and near the bottom quark mass. We might therefore include at least the charm quark in the free energy.
- (c) A complete two-loop calculation of  $F$  is not available at present. In some places in the literature, UV-divergent two-loop corrections to the free energy are incorporated using renormalization group considerations, as discussed in [21]. These corrections are generally written as powers of  $\alpha_s(\rho^{-1})/\alpha_s(\mu)$  in the  $\rho$  integrand. However, for  $\mu$  of order  $T$  (or  $\pi T$ ), there is no justification for including some two-loop corrections and not others. Without performing an actual two-loop computation of the  $\theta$ -dependent part of the free energy, the only principled approach is to use the complete one-loop expression with  $\mu$  of order  $T$  (or  $\pi T$ ). However, the ‘‘UV two-loop’’ computation might instead be useful as an uncertainty estimator.

We therefore recompute the free energy using  $\mu = \{1, \sqrt{\pi}, \pi\} \times T$ , three or four active flavors, and including or not including the UV-divergent two-loop corrections. For the change in renormalization scale, we again use RunDec to accurately determine  $\alpha_s(\mu)$  with different numbers of active flavors. The two-loop corrections are incorporated by running  $\alpha_s$  in the exponent to  $\mu = \rho^{-1}$  at two-loop order and running the quark masses and the coupling in the prefactor to  $\mu = \rho^{-1}$  at one-loop order. For example,

$$e^{-\frac{2\pi}{\alpha_s(\mu)}} \rightarrow e^{-\frac{2\pi}{\alpha_s(\mu)}(\mu\rho)^{b_0} \left(\frac{\alpha_s(\rho^{-1})}{\alpha_s(\mu)}\right)^{2b_1/b_0}} \quad (2.10)$$

where  $b_0 = 9$  and  $2b_1/b_0 = 32/9$  for  $N_F = 3$ , and  $\alpha_s(\rho^{-1})$  is determined from  $\alpha_s(\mu)$  at one-loop order.

Results are reported in Table I. The largest value for the free energy is obtained in the three-flavor scheme with  $\mu = T$ , adding the partial two-loop terms. This is not a surprise, since the included two-loop terms correspond entirely to running from  $T$  to  $\rho^{-1} \sim \pi T$ . However, it is likely an overestimate of the correction; if we use  $\mu = \pi T$  in the same computation, the partial two-loop result is smaller and much closer to the complete one-loop result. In reality, the complete two-loop result is likely to involve a mixture of scales, motivating the choice  $\mu = \sqrt{\pi}T$ . We observe that (excluding the  $3F, 2L, \mu = T$  result), the envelope of the values is contained within the  $\mu = \sqrt{\pi}T$  calculations, corresponding to an  $\mathcal{O}(1)$  uncertainty in the free energy,

$$\frac{\Delta F_0(1.5)}{F_0(1.5)} \simeq 1. \quad (2.11)$$

TABLE I. The instanton-induced free energy in units of  $-10^{-14}$  GeV $^{-4}$  at  $\theta = 0$  and  $T = 1.5$ . Rows correspond to a variety of computations: (3F, 4F) = three or four light flavors; (1L, 2L) = one-loop complete or partial two-loop; (T,  $\sqrt{\pi}T$ ,  $\pi T$ ) = renormalization scale.

3F, 1L, T	3.6
3F, 2L, T	10
3F, 1L, $\sqrt{\pi}T$	4.9
3F, 2L, $\sqrt{\pi}T$	7.2
4F, 1L, $\sqrt{\pi}T$	3.2
4F, 2L, $\sqrt{\pi}T$	5.2
3F, 1L, $\pi T$	6.0
3F, 2L, $\pi T$	5.5
4F, 1L, $\pi T$	4.0
4F, 2L, $\pi T$	3.8

We can also estimate a ‘‘parametric’’ uncertainty stemming from experimental uncertainty in  $\alpha_s$ . Using the 1-sigma error bar on  $\alpha_s(m_Z)$ , running down to  $\mu = \pi \times 1.5$  GeV and converting to the three-flavor scheme with RunDec [18–20], we obtain less than 2% uncertainty in  $\alpha_s$ .<sup>1</sup> This results in an uncertainty in  $F_0(1.5)$  of about a factor of 2, similar to the uncertainty from UV-sensitive corrections.

The lattice result for the topological susceptibility obtained in Ref. [13] corresponds to  $F_0(1.5) \approx -4 \times 10^{-13}$  GeV $^{-4}$ , which lies outside the uncertainty range that we estimate from these sources. Similar conclusions were drawn in the lattice studies [12,16]. In [12,17], it was suggested that the uncertainty in the 1-loop instanton computation arising from higher-order terms could actually be much larger, associated with infrared divergences in QCD perturbation theory at finite temperatures and with large shifts in the Debye screening length. We now turn to corrections of this type.

### C. Corrections to the IR cutoff on $\rho$ and infrared sensitivity

As discussed previously, the IR cutoff on the instanton size can be *qualitatively* associated with the Debye mass term in the effective action. In the perturbative vacuum, the Debye mass does receive large corrections beyond leading order [22,23]. A simple, heuristic understanding of these corrections can be obtained by considering  $\Pi_{44}$  as a function of (spatial) momentum,  $\vec{q}$ , for small  $\vec{q}$ . There is a variety of effects, but already at one loop, for example, there is a contribution to  $\partial\Pi/\partial q^2$  that diverges linearly as  $m_D \rightarrow 0$  at  $q = 0$ . The linear IR divergence is cut off by the leading-order  $m_D$ , leaving a weaker logarithmic IR divergence cut off by the nonperturbative magnetic mass. The next-to-leading-order Debye mass has the following form [22,23]:

<sup>1</sup>Uncertainty from higher-order corrections to the running of  $\alpha_s$  are extremely subdominant.

$$m_D^2 = (m_D)_0^2 + \frac{2Ng^2}{4\pi} T(m_D)_0 \ln(m_D/g^2 T) + \dots \quad (2.12)$$

The next-to-leading-order correction is of order  $g^3$ , signaling a breakdown of the perturbation expansion. It has been suggested [12,17] that the uncertainty on  $\chi(\theta, T)$  might be much larger than estimated in the previous section, due to the presence of such IR divergences.<sup>2</sup>

In the instanton computation, there is both a question of principle and a question of numerics. We have seen that it is not low spatial momenta that are relevant in the instanton background, but momenta of order  $k \sim 1/\rho \sim T$ . Consequently, for the dominant semiclassical configurations with  $\rho \sim 1/T \ll \Lambda$ , the IR divergences in  $\Pi_{44}$  are cut off at  $T$  in the instanton effective action  $S_{\text{eff}}$ . The corrections to  $S_{\text{eff}}$  from individual diagrams are then well-behaved and proportional to  $g(T)^2 T^2$ . Thus, as a matter of principle, IR-divergent corrections to  $m_D$  in the perturbative vacuum do not indicate a loss of perturbative control or a significant source of uncertainty in the instanton computation of  $F(\theta, T)$ .

However, until  $T$  is extremely large,  $g(T)$  is  $\mathcal{O}(1)$  in QCD, and there is no parametric separation between  $m_D$  and  $T$ . Therefore, Eq. (2.12), valid in the perturbative vacuum, might still be used as an estimate for the typical size of corrections to the effective action in the instanton background. Numerically, it gives rise to

$$\frac{(m_D)_1}{(m_D)_0} \simeq 0.6 \quad (2.13)$$

at  $T = 1.5$  GeV. The instanton-induced free energy scales approximately as the 7<sup>th</sup> – 8<sup>th</sup> power of the infrared cutoff on  $\rho$ , so from Eq. (2.13) we are led to associate an uncertainty in the free energy due to two-loop finite temperature corrections,

$$\frac{\Delta F_0(1.5)}{F_0(1.5)} \simeq 20. \quad (2.14)$$

We emphasize that a correction to the free energy of this size does *not* reflect a breakdown of the semiclassical analysis at this order. Organizing the instanton effective action as

$$S_{\text{inst}} = S_0 + S_1 + S_2 + \dots, \quad (2.15)$$

at leading order, the action is

<sup>2</sup>It should be noted that the existence of such corrections is not connected with the presence of light fermions. In particular, fermions do not introduce infrared divergences at high temperature. Therefore, even lattice studies focusing on the size of corrections in the pure gauge theory are of interest.

$$S_0 = \frac{8\pi^2}{g^2} \simeq 17 \quad (2.16)$$

at  $T = 1.5$  GeV. A shift in the free energy of order Eq. (2.14) corresponds to

$$\frac{S_2^{\text{Debye}}}{S_0} \lesssim 0.2, \quad (2.17)$$

a controlled correction to the effective action. More generally, we could estimate terms in the series by the three-dimensional loop factor, which is of order

$$\lambda = \frac{Ng^2(T)}{(4\pi)^{3/2}}. \quad (2.18)$$

At  $T = 1.5$  GeV,  $\lambda = 0.3$ . Two-loop corrections to the effective action would then be expected to be of order

$$\frac{S_2}{S_0} \simeq \lambda^2 = 0.1 \quad (2.19)$$

consistent with the Debye estimate. In other words, there is no reason to expect *arbitrarily* large corrections. The action is exponentiated in the free energy, leading to the order-of-magnitude uncertainty estimate in Eq. (2.14).

There is also the question of actual infrared divergent contributions to the instanton action. These are associated with low-momentum  $\vec{A}$  fields and corrections to the effective action involving no background fields. In the zero-instanton sector, such infrared divergences arise in the free energy first at four-loop order. It is believed that they are cut off at a scale of order  $g^2 T$ , the presumed mass gap of the three-dimensional gauge theory. The typical diagram involves six vertices connected by propagators, and the divergence arises when all vertices are well-separated. A computation at high order in the instanton background is complex, but for  $\rho$  of order  $T^{-1}$ , the infrared divergence should be similar. At zero temperature, the propagators are known [24,25], and at distances large compared to  $\rho$ , they are close to free-field propagators. At finite temperatures, when all coordinates except  $x_4$  are large compared to  $T^{-1}$  and  $\rho$ , we expect something similar, leading to an infrared divergent correction at the same order as at zero temperature. At 1.5 GeV, this suggests a perturbatively incalculable correction to the instanton action at the 1% level.

In summary, IR divergences do not appear relevant to the instanton computation, and semiclassical analysis is under sufficient theoretical control to admit uncertainty estimates. Absent a complete two-loop computation, we will take the perturbative Debye mass correction, Eq. (2.14), as a *conservative* estimate of the uncertainty in the  $\theta$ -dependent free energy. UV cutoff-sensitive corrections and uncertainties in  $\alpha_s$  are expected to be subdominant to the finite- $T$

corrections to the effective action, and Eq. (2.16) indicates that dilute gas corrections are expected to be negligible.

We therefore know with some confidence the range of possible behaviors for the axion potential both at temperatures below the critical temperature ( $\sim 150$  MeV for  $N_f = 3$ ) and at temperatures a few GeV and above. These boundary properties constrain the behavior in the intermediate range of temperatures, which happen to lie where the axion begins to oscillate in conventional scenarios. For this reason lattice computations (that successfully reproduce the high-temperature behavior) can be of value. On the other hand, as we will describe below, if we simply assume a smooth interpolation between the two regimes, the axion relic density is not very sensitive either to the form of the interpolation or the uncertainty in the high energy semiclassical computation.

In closing this section, we note that there have been arguments that the behavior of  $\chi$  is drastically different at high temperatures than the semiclassical result, even turning off exponentially rapidly with temperature [26–29]. We will not address this possibility further here, but it is certainly true that in such a circumstance substantially different axion relic densities can be obtained [10,28,29].

### III. $\chi(T)$ AT INTERMEDIATE TEMPERATURES

We have argued that we know the high-temperature behavior of  $\chi$  to about an order of magnitude. At scales below 1 GeV, the coupling rapidly becomes strong, and other methods are needed to determine the axion mass.

At very low temperatures, the  $\theta$  dependence of the vacuum energy is known reliably from current algebra,

$$F(\theta, 0) = -3.6 \times 10^{-5} \text{ GeV}^4 \cos(\theta). \quad (3.1)$$

Finite temperature lattice computations indicate that the topological susceptibility,

$$\chi(T) = \left. \frac{\partial^2 V(T)}{\partial \theta^2} \right|_{\theta=0}, \quad (3.2)$$

is near its zero-temperature chiral perturbation theory (ChPT) value at temperatures of order 100 MeV, and remains approximately equal to it until at least the chiral phase transition near 150 MeV [8]. Beyond this scale, only lattice computations can accurately determine  $\chi(T)$ , and at present there are varied results in the literature.

However, for the purposes of computing the axion relic density, it turns out to be sufficient to consider simple models that interpolate between the ChPT and instanton regimes. We will adopt the following class of models for  $F(\theta, T)$ :

$$F(\theta, T) = \begin{cases} -\chi(0) \cos \theta, & 0 < T < T_2 \\ -\chi(T_0) \left(\frac{T_0}{T}\right)^n \cos \theta, & T_2 < T < T_0 \\ -\chi(T_0) \left(\frac{T_0}{T}\right)^8 \cos \theta, & T > T_0 \end{cases} \quad (3.3)$$

Here  $T_0$  is the ‘‘anchor point’’ for the instanton regime. The results of [8] suggest that the slope of  $\chi$  is instantonlike down to temperatures a few times  $T_c$ ; however, to maintain a minimal uncertainty in the semiclassical computation, we fix  $T_0 = 1.5$  GeV. As discussed below, our modeling still includes the possibility of instantonlike slopes at lower  $T$ . We will vary  $\chi(T_0)$  within the uncertainty on the instanton computation.  $T_2$ , the anchor point for the ChPT regime, is related to  $T_0$  and the slope of the power law in the model by

$$T_2^n = T_0^n \times \frac{\chi(T_0)}{\chi(0)}. \quad (3.4)$$

We vary  $n$  such that  $T_2$  varies between 100 and 500 MeV. Given  $T_0 = 1.5$  GeV, values of  $T_2$  of order 150 in fact correspond to  $n \approx 8$ , equivalent to assuming instantonlike behavior persists significantly below  $T_0$ . Larger values for  $T_2$  above the critical temperature  $T_c$  are not based on physical considerations, but instead are included to partially accommodate the lattice results of [7,14], which found very *shallow* falloff of  $\chi(T)$  above the chiral phase transition. This behavior is then approximated in the models of Eq. (3.3) for larger  $T_2$  by zero falloff until  $T_2$ . However, Refs. [7,14,17] extrapolated the shallow power law behavior up to high temperatures, leading to values for  $F_0$  many orders of magnitude different from the semiclassical result. Our insistence on reaching instanton behavior by 1.5 GeV [within the uncertainty (2.14)] requires an even *steeper* power law to set in above  $T_2$  when  $T_2 \gg T_c$ . Numerically,  $n$  will fall in the range 7–20, with the lower values corresponding to lower values of  $T_2$ .

### IV. AXION RELIC DENSITY FROM MISALIGNMENT

We can now assess the sensitivity of the axion relic density to uncertainties in  $\chi(T)$ , including both the uncertainties in the instanton computation and the range of models for the behavior at intermediate temperatures. In the figures below we will numerically integrate the equation of motion,

$$\ddot{a} + 3H\dot{a} + V'(a) = 0. \quad (4.1)$$

However, for qualitative purposes, a good approximation is obtained by treating the axion as frozen until a temperature  $T_{\text{osc}}$  [30,31]:

$$m_a(T_{\text{osc}}) = 3H(T_{\text{osc}}). \quad (4.2)$$

At this point, the axion begins to oscillate with a time-(temperature-) dependent mass. Approximating the energy density by

$$\rho(t) = \frac{1}{2} \dot{a}^2 + \frac{1}{2} m_a^2(T) a^2, \quad (4.3)$$

TABLE II. Axion relic density as a function of model parameters as computed with the approximate formula (4.5). Here  $m_a(0) = 30 \mu\text{eV}$ ,  $\chi_0$  is given in units of  $3.7 \times 10^{-14} \text{ GeV}^4$ , and the misalignment angle is set to the value appropriate for post-inflationary breaking of the Peccei-Quinn symmetry,  $\theta_0 = 2.16$ .

$\Omega$			
$n$	$\chi_0 = 1/10$	$\chi_0 = 1$	$\chi_0 = 10$
8	0.22	0.18	0.15
14	0.18	0.16	0.14
20	0.17	0.15	0.14

one can show that it evolves with temperature as

$$\rho(T) = \rho(T_{\text{osc}}) \left( \frac{R^3(T_{\text{osc}})}{R^3(T)} \right) \frac{m_a(T)}{m_a(T_{\text{osc}})}. \quad (4.4)$$

Within the range of  $m_a(T)$  that we consider,  $T_{\text{osc}}$  is always less than the instanton anchor point  $T_0$ . Therefore, Eq. (4.2) can be solved by substituting the intermediate power-law behaviors for  $m_a(T)$ ; the instanton asymptotics constrain the range of intermediate power laws considered.

The relic density  $\Omega_a$  can then be expressed as a function of the parameters  $\chi(T_0)$  and  $n$  (or  $T_2$ ). The result is:

$$\Omega_{\text{axion}} = 0.13 \times (7.3)^{\frac{2}{4+n}} \left( \frac{m_a}{30 \mu\text{eV}} \right)^{-\frac{6+n}{4+n}} \times \left( \frac{\chi_0(1.5)}{3.7 \times 10^{-14} \text{ GeV}^4} \right)^{-\frac{1}{4+n}} \left( \frac{\theta_0}{2.155} \right)^2 \quad (4.5)$$

where  $\theta_0$  is the initial misalignment angle and  $m_a$  is the zero-temperature axion mass. Taking, for example,  $m_a = 30 \mu\text{eV}$  and a few values for  $\chi_0(1.5)$  and  $n$  gives the results in Table II. Alternatively, for fixed  $\Omega_{\text{axion}} = \Omega_{\text{DM}}$ , we obtain

$$\left( \frac{m_a}{30 \mu\text{eV}} \right) = 0.51 \times (2.4)^{\frac{6}{6+n}} \left( \frac{\chi_0(1.5)}{3.7 \times 10^{-14} \text{ GeV}^4} \right)^{-\frac{1}{6+n}} \times \left( \frac{\theta_0}{2.155} \right)^{\frac{8+2n}{6+n}} \left( \frac{\Omega_{\text{DM}}}{0.25} \right)^{-\frac{4+n}{6+n}}. \quad (4.6)$$

Eq. (4.5) indicates that the relic density is substantially insensitive to the magnitude of the free energy at high temperatures; for  $T_2 = 150 \text{ MeV}$  ( $n \sim 8$ ),  $\chi_0$  enters to the  $-\frac{1}{12}$  power. Therefore, sizable uncertainties in  $\chi_0$  translate into modest uncertainties in  $m_a$ , also observed in [8]. Similarly the dependence on  $n$  ( $T_2$ ) is mild. The lattice results of [8,13], for example, differ from our estimate of  $\chi_0(1.5)$  by a factor of about 10, and exhibit power-law behavior corresponding to  $n \approx 8$ . If the Peccei-Quinn phase transition occurs after inflation, this factor of 10 leads to

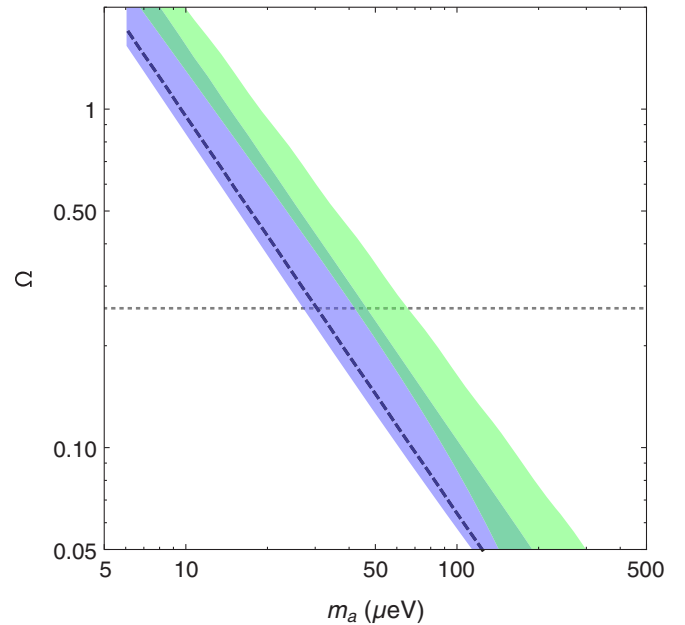


FIG. 1. Axion relic density from misalignment in the post-inflationary scenario. Colors correspond to different models for the temperature-dependent free energy between the dilute gas at high temperatures and chiral perturbation theory at low temperatures. Specifically, the blue (green) band sets the anchor point for ChPT at  $T_2 = 100(500) \text{ MeV}$ . The width of each band reflects the uncertainty in the instanton computation of the free energy used as an anchor at  $T = 1.5 \text{ GeV}$ ,  $F_0 \rightarrow (1/20, 20) \times F_0$ , c.f. Eqs. (2.9) and (2.14). The dashed line corresponds to the value of  $F_0(1.5)$  obtained in the lattice calculation of Ref. [13].

about a 15% decrease in the value of the axion mass required to account for the observed dark matter density.

To obtain a more accurate result for the late-time relic density, we solve the full axion equation of motion numerically through the time where it starts to oscillate. Figure 1 shows the relic density obtained in this way for two values of  $T_2$  and a range of  $\chi_0$ , in the postinflationary Peccei-Quinn-breaking scenario ( $\theta_0 = 2.155$ ). Compared to the analytic estimate (4.5), the full numerical solution yields marginally higher  $\Omega$  for fixed  $m_a$ . Even with the factor of 5 variation in  $T_2$  and the factor of  $20^2$  variation in  $\chi_0$ , we find that the axion mass required to account for all of dark matter varies by only a factor of 2–3.

Additional sources of axion production (cosmic strings) can force a larger axion mass. These masses are, indeed, at the edge of capability of cavity experiments like ADMX, and are the focus of much future planning. However, these sources of energy density, as well as the constraint on  $\theta_0$ , are not necessarily present in the early Universe.

In fact, there is not necessarily a Peccei-Quinn transition at all [32]. The approximate Peccei-Quinn symmetry, if it exists, is almost certainly an accident. This accident may not occur at the high temperatures or high curvatures that characterize the early Universe. In this case, the initial value of  $\theta$  is a fixed number, or possibly one of a set of discrete

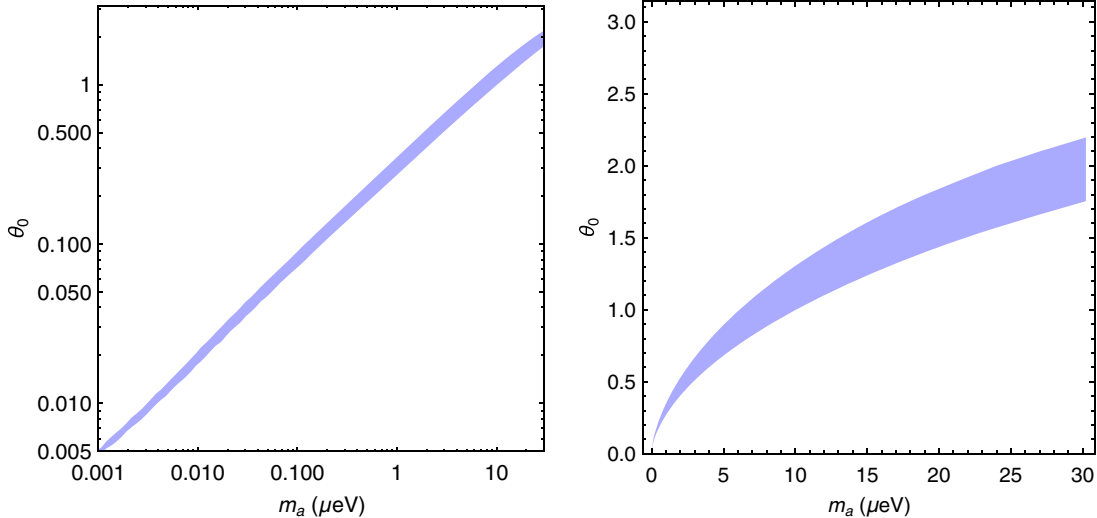


FIG. 2. Axion relic density from misalignment in the preinflationary scenario. The curve shows the misalignment angle needed to obtain  $\Omega = 0.258$ . The band reflects the uncertainty in the instanton computation of the free energy, Eq. (2.14), used as an anchor at  $T = 1.5$  GeV, and the anchor point for ChPT has been fixed to  $T_2 = 140$  MeV. Left panel: log-log axes over a broad range of axion masses. Right panel: linear axes over a range of axion masses in reach of current and next-generation ADMX [33].

numbers. This number might well be small, or might be  $\mathcal{O}(1)$ . Either has significant implications for the final dark matter density.

Alternatively, there may be an approximate symmetry both for low and high temperature (or curvature). The question of whether the symmetry is broken during or after inflation then depends, for example, on the coupling of the inflaton to the field responsible for Peccei-Quinn symmetry breaking. For example, there might be an effective mass term for this field, of either sign. There seems to be no particular reason to believe that one or the other outcome is favored.

These different possibilities have been extensively studied in the literature. If the symmetry breaking occurs after inflation, one has to average over random initial misalignment angles, which fixes the parameter  $\theta_0$  as above. In the case of symmetry breaking before inflation,  $\theta_0$  is a free parameter. In Fig. 2 we show the sensitivity of the misalignment angle required to saturate the relic density to the uncertainty in  $\chi_0$ . As in the postinflationary case studied above, we find that the theoretical uncertainties have essentially no qualitative impact on the required parameters. Furthermore, for a wide range of  $\mathcal{O}(1)$  values for  $\theta_0$ , the relevant axion masses are compatible with current and next-generation cavity experiments [4,33].

## V. CONCLUSIONS

Within the conventional picture of axion cosmology, we have found that the standard computation of the axion relic density is relatively robust against theoretical uncertainties stemming from the dilute gas computation of the QCD free energy at high temperatures and the behavior of

the free energy at strong coupling. In particular, we have argued that the instanton computation is under sufficient control at temperatures of 1–2 GeV to allow a reasonable assessment of uncertainties due to higher-order corrections. These corrections cannot amount to much more than an order of magnitude in the free energy without an unexplained breakdown in the semiclassical analysis. In particular, we have argued that infrared divergences in the Debye mass in the perturbative vacuum are not relevant in the instanton background and cannot inject arbitrarily large corrections to  $F(\theta, T)$ . Thus, while an improved determination of the finite-temperature topological susceptibility would lead to improvement in the precision of the (relic density, axion mass) relation, it is not expected to lead to qualitative (order-of-magnitude) changes, and modern cavity experiments retain significant discovery potential.

However, in closing, we note that it has long been recognized that the underlying cosmological assumptions of the standard calculation may not hold, and that there is good theoretical motivation to consider lighter axions with larger decay constants. Within conventional effective field theory, for example, it is hard to account for the requisite quality of the Peccei-Quinn symmetry without invoking large discrete symmetries. String theory points to a different picture, in which the Peccei-Quinn symmetry appears more natural [34]. The assumption that the underlying mass scale is of the order of the Planck or unification scale is suggestive of larger decay constants. It could also be that four-dimensional effective field theory is not useful at scales orders of magnitude below the Planck scale, as in large or warped extra dimension scenarios, and early Universe cosmology might be substantially modified.

## ACKNOWLEDGMENTS

This work was supported in part by the U.S. Department of Energy Grant No. DE-FG02-04ER41286. We thank N. Arkani-Hamed and R. Kitano for valuable comments,

and Leslie Rosenberg for bringing the work of [13] to our attention. M. D. thanks the Institute for Advanced Study for its hospitality while much of this work was performed.

- 
- [1] L. F. Abbott and P. Sikivie, A cosmological bound on the invisible axion, *Phys. Lett.* **120B**, 133 (1983).
- [2] J. Preskill, M. B. Wise, and F. Wilczek, Cosmology of the invisible axion, *Phys. Lett. B* **120B**, 127 (1983).
- [3] M. Dine and W. Fischler, The not so harmless axion, *Phys. Lett. B* **120B**, 137 (1983).
- [4] L. J. Rosenberg, Dark-matter QCD-axion searches, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 12278 (2015).
- [5] P. W. Graham, I. G. Irastorza, S. K. Lamoreaux, A. Lindner, and K. A. van Bibber, Experimental searches for the axion and axion-like particles, *Annu. Rev. Nucl. Part. Sci.* **65**, 485 (2015).
- [6] D. J. Gross, R. D. Pisarski, and L. G. Yaffe, QCD and instantons at finite temperature, *Rev. Mod. Phys.* **53**, 43 (1981).
- [7] C. Bonati, M. D’Elia, M. Mariti, G. Martinelli, M. Mesiti, F. Negro, F. Sanfilippo, and G. Villadoro, Axion phenomenology and  $\theta$ -dependence from  $N_f = 2 + 1$  lattice QCD, *J. High Energy Phys.* **03** (2016) 155.
- [8] S. Borsanyi, M. Dierigl, Z. Fodor, S. D. Katz, S. W. Mages, D. Nogradi, J. Redondo, A. Ringwald, and K. K. Szabo, Axion cosmology, lattice QCD and the dilute instanton gas, *Phys. Lett. B* **752**, 175 (2016).
- [9] E. Berkowitz, M. I. Buchoff, and E. Rinaldi, Lattice QCD input for axion cosmology, *Phys. Rev. D* **92**, 034507 (2015).
- [10] R. Kitano and N. Yamada, Topology in QCD and the axion abundance, *J. High Energy Phys.* **10** (2015) 136.
- [11] J. Frison, R. Kitano, H. Matsufuru, S. Mori, and N. Yamada, Topological susceptibility at high temperature on the lattice, *J. High Energy Phys.* **09** (2016) 021.
- [12] P. Petreczky, H.-P. Schadler, and S. Sharma, The topological susceptibility in finite temperature QCD and axion cosmology, *Phys. Lett. B* **762**, 498 (2016).
- [13] S. Borsanyi *et al.*, Calculation of the axion mass based on high-temperature lattice quantum chromodynamics, *Nature (London)* **539**, 69 (2016).
- [14] C. Bonati, M. D’Elia, M. Mariti, G. Martinelli, M. Mesiti, F. Negro, F. Sanfilippo, and G. Villadoro, Recent progress on QCD inputs for axion phenomenology, *Eur. Phys. J. Web Conf.* **137**, 08004 (2017).
- [15] A. Trunin, F. Burger, E.-M. Ilgenfritz, M. P. Lombardo, and M. Muller-Preussker, Topological susceptibility from  $N_f = 2 + 1 + 1$  lattice QCD at nonzero temperature, *J. Phys. Conf. Ser.* **668**, 012123 (2016).
- [16] F. Burger, E.-M. Ilgenfritz, M. P. Lombardo, M. Muller-Preussker, and A. Trunin, Topology (and axion’s properties) from lattice QCD with a dynamical charm, *Nucl. Phys.* **A967**, 880 (2017).
- [17] G. Grilli di Cortona, E. Hardy, J. Pardo Vega, and G. Villadoro, The QCD axion, precisely, *J. High Energy Phys.* **01** (2016) 034.
- [18] K. G. Chetyrkin, J. H. Kuhn, and M. Steinhauser, RunDec: A Mathematica package for running and decoupling of the strong coupling and quark masses, *Comput. Phys. Commun.* **133**, 43 (2000).
- [19] B. Schmidt and M. Steinhauser, CRunDec: a C++ package for running and decoupling of the strong coupling and quark masses, *Comput. Phys. Commun.* **183**, 1845 (2012).
- [20] F. Herren and M. Steinhauser, Version 3 of RunDec and CRunDec, [arXiv:1703.03751](https://arxiv.org/abs/1703.03751).
- [21] T. R. Morris, D. A. Ross, and C. T. Sachrajda, Higher order quantum corrections in the presence of an instanton background field, *Nucl. Phys.* **B255**, 115 (1985).
- [22] A. K. Rebhan, Non-Abelian Debye mass at next-to-leading order, *Phys. Rev. D* **48**, R3967 (1993).
- [23] P. B. Arnold and L. G. Yaffe, Non-Abelian Debye screening length beyond leading order, *Phys. Rev. D* **52**, 7208 (1995).
- [24] L. S. Brown, R. D. Carlitz, D. B. Creamer, and C.-k. Lee, Propagation functions in pseudoparticle fields, *Phys. Rev. D* **17**, 1583 (1978).
- [25] H. Levine and L. G. Yaffe, Higher order instanton effects, *Phys. Rev. D* **19**, 1225 (1979).
- [26] T. D. Cohen, High temperature phase of QCD and U(1)-A symmetry, *Phys. Rev. D* **54**, R1867 (1996).
- [27] S. Aoki, H. Fukaya, and Y. Taniguchi, Chiral symmetry restoration, eigenvalue density of Dirac operator and axial U(1) anomaly at finite temperature, *Phys. Rev. D* **86**, 114512 (2012).
- [28] V. Azcoiti, Topology in the SU(N<sub>f</sub>) chiral symmetry restored phase of unquenched QCD and axion cosmology, *Phys. Rev. D* **94**, 094505 (2016).
- [29] V. Azcoiti, Topology in the SU(N<sub>f</sub>) chiral symmetry restored phase of unquenched QCD and axion cosmology II, *Phys. Rev. D* **96**, 014505 (2017).
- [30] M. S. Turner, Windows on the axion, *Phys. Rep.* **197**, 67 (1990).
- [31] E. W. Kolb and M. S. Turner, The early Universe, *Front. Phys.* **69**, 1 (1990).
- [32] M. Dine and A. Anisimov, Is there a Peccei-Quinn phase transition?, *J. Cosmol. Astropart. Phys.* **07** (2005) 009.
- [33] <http://depts.washington.edu/admx/future.shtml>.
- [34] E. Witten, Some properties of O(32) superstrings, *Phys. Lett.* **149B**, 351 (1984).