

Lattice QCD for Cosmology

Sz. Borsanyi¹, Z. Fodor^{1,2,3}, K.-H. Kampert¹, S. D. Katz^{3,4}, T. Kawanai², T. G. Kovacs⁵, S. W. Mages², A. Pasztor¹, F. Pittler^{3,4}, J. Redondo^{6,7}, A. Ringwald⁸, K. K. Szabo^{1,2}

¹ Department of Physics, University of Wuppertal, D-42119 Wuppertal, Germany

² Jülich Supercomputing Centre, Forschungszentrum Jülich, D-52428 Jülich, Germany

³ Institute for Theoretical Physics, Eötvös University, H-1117 Budapest, Hungary

⁴ MTA-ELTE Lendület Lattice Gauge Theory Research Group, H-1117 Budapest, Hungary

⁵ Institute for Nuclear Research of the Hungarian Academy of Sciences, H-4026 Debrecen, Hungary

⁶ University of Zaragoza, E-50009 Zaragoza, Spain

⁷ Max Planck Institut für Physik, D-80803, Germany

⁸ Deutsches Elektronen-Synchrotron DESY, D-22607 Hamburg, Germany

We present a full result for the equation of state (EoS) in 2+1+1 (up/down, strange and charm quarks are present) flavour lattice QCD. We extend this analysis and give the equation of state in 2+1+1+1 flavour QCD. In order to describe the evolution of the universe from temperatures several hundreds of GeV to the MeV scale we also include the known effects of the electroweak theory and give the effective degree of freedoms. As another application of lattice QCD we calculate the topological susceptibility (χ) up to the few GeV temperature region. These two results, EoS and χ , can be used to predict the dark matter axion's mass in the post-inflation scenario and/or give the relationship between the axion's mass and the universal axionic angle, which acts as a initial condition of our universe.

The well-established theories of the strong interaction (QCD) and the electroweak theory determine the evolution of the early universe. The Hubble rate and the relationship between the age of the universe and its temperature (T) are determined by the equation of state (EoS). Since QCD is highly non-perturbative, the calculation of the EoS is a particularly difficult task. The only systematic way to carry out this calculation is based on lattice QCD. Here we present complete results of such a calculation. QCD, unlike the rest of the Standard Model, is surprisingly symmetric under time reversal, leading to a serious fine tuning problem. The most attractive solution for this is a new particle, the axion –a promising dark matter candidate. Assuming that axions are the dominant component of dark matter we determine the axion mass. The key quantities of the calculation are the previously mentioned EoS and the temperature dependence of the topological susceptibility ($\chi(T)$) of QCD, a quantity notoriously difficult to calculate. Determining $\chi(T)$ was believed to be impossible in the most relevant high temperature region, however an understanding of the deeper structure of the vacuum by splitting it into different sectors and re-defining the fermionic determinants has led to its fully controlled calculation. Thus, our two-fold prediction helps most cosmological calculations to describe the evolution of the early universe by using the fully controlled EoS and may be decisive for guiding experiments looking for dark matter axions. In the next couple of years, it should be possible to confirm or rule out post-inflation axions experimentally [17] if the axion’s mass is or is not found to be as predicted here. Alternatively, in a pre-inflation scenario our calculation determines the universal axionic angle that corresponds to the initial condition of our universe.

The Euclidean Lagrangian for the theory of the strong interaction is $\mathcal{L}_{\text{QCD}} = 1/(2g^2)\text{Tr}F_{\mu\nu}F_{\mu\nu} + \sum_f \bar{\psi}_f[\gamma_\mu(\partial_\mu + iA_\mu) + m_f]\psi_f$, where g is the QCD coupling constant, m_f represent the quark masses and f runs over the four quark flavors. In QCD $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]$, where A_μ is a Hermitian 3×3 matrix field denoting the gauge fields. The ψ_f are Dirac-spinor fields representing the quarks and carry a “colour” index, which runs from 1 to 3. The form of the action is (almost) unambiguously defined by the condition of gauge invariance and renormalizability. For many quantities, the only systematic way to solve this theory is to discretize it on a Euclidean space-time lattice, determine the partition function stochastically and using smaller and smaller lattice spacings to extrapolate to the continuum limit (the limit with vanishing lattice spacing). In this paper, we use this lattice formulation to determine the EoS of QCD and the topological susceptibility for low temperatures up to very high temperatures. Finally, as an application, we combine them to present our results on the dark matter axion mass.

Our most important qualitative knowledge about the QCD transition is that it is an analytic crossover [3], thus no cosmological relics are expected. Outside the narrow temperature range of the transition we know that the Hubble rate and the relationship between temperature and the age of the early universe can be described by a radiation-dominated EoS. The calculation of the EoS is a challenging task [17], the determination of the continuum limit at large temperatures is particularly difficult.

In our lattice QCD setup we used 2+1 or 2+1+1 flavours of staggered fermions with four steps of stout-smearing. For the gluonic sector tree-level improved gauge fields were used. The quark masses are set to their physical values, however we use degenerate up and down quark masses and the small effect of isospin breaking is included analytically. The continuum limit is taken using three, four or five lattice spacings with temporal lattice extensions of $N_T=6, 8, 10, 12$ and 16. In addition to dynamical staggered simulations we also used dynamical overlap simulations with 2+1 flavours down to physical masses. The inclusion of an odd number of flavours was a non-trivial task, however this setup was required for the determination of $\chi(T)$ at large temperatures in the several GeV region [17].

Charm quarks start to contribute to the equation of state above 300 MeV. Therefore up to 250 MeV we used 2+1 flavours of dynamical quarks. Connecting the 2+1 and the 2+1+1 flavour results at 250 MeV can be done smoothly. For large temperatures the step-scaling method for the equation of state of Ref. [9] was applied. We determined the EoS with complete control over all sources of systematics all the way to the GeV scale.

Two different methods were used to set the overall scale in order to determine the equation of state.

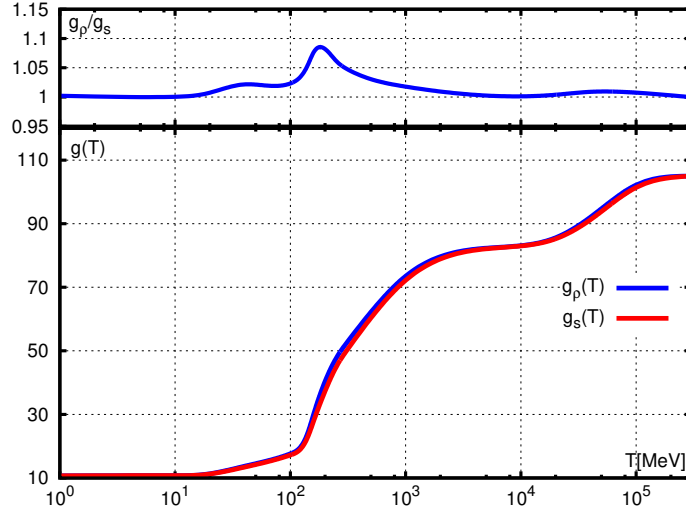


Figure 1: The effective degrees of freedom for the energy density (g_ρ) and for the entropy density (g_s). The line width is chosen to be the same as our error bars at the vicinity of the QCD transition where we have the largest uncertainties. At temperatures $T < 1$ MeV the equilibrium equation of state becomes irrelevant for cosmology, because of neutrino decoupling. The EoS comes from our calculation up to $T = 100$ GeV. At higher temperatures the electroweak transition becomes relevant and we use the results of Ref. [13]. Note that for temperatures around the QCD scale non-perturbative QCD effects reduce g_ρ and g_s by 10-15% compared to the ideal gas limit, an approximation which is often used in cosmology. For useful parametrizations for the QCD regime or for the whole temperature range see [17].

One of them took the pion decay constant the other applied the w_0 scale [10]. 32 different analyses (e.g. the two different scale setting procedures, different interpolations, keeping or omitting the coarsest lattice) entered our histogram method [12, 8] to estimate systematic errors. We also calculated the goodness of the fit Q and weights based on the Akaike information criterion AICc [8] and we looked at the unweighted or weighted results. This provided the systematic errors on our findings. In the low temperature region we compared our results with the prediction of the Hadron Resonance Gas (HRG) approximation and found perfect agreement. This HRG approach is used to parametrize the equation of state for small temperatures. In addition, we used the hard thermal loop approach [1] to extend the EoS to high temperatures.

In order to have a complete description of the thermal evolution of the early universe we supplement our QCD calculation for the EoS by including the rest of the Standard Model particles (leptons, bottom and top quarks, W , Z , Higgs bosons) and results on the electroweak transition [17]. As a consequence, the final result on the EoS covers four orders of magnitude in temperature from MeV to several hundred GeV.

Figure 1 shows the result for the effective numbers of degrees of freedom as a function of temperature. The widths of the lines represent the uncertainties. The tabulated data are also presented in [17]. Both the figure and the data can be used (similarly to Figure 22.3 of Ref. [14]) to describe the Hubble rate and the relationship between temperature and the age of the universe in a very broad temperature range.

We now turn to the determination of another cosmologically important quantity, $\chi(T)$. In general the Lagrangian of QCD should have a term proportional to $\mathcal{L}_Q = 1/(32\pi^2)\epsilon_{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}$, the four-dimensional integral of which is called the topological charge. This term violates the combined charge-conjugation and parity symmetry (CP). The surprising experimental observation is that the proportionality factor of this term θ is unnaturally small. This is known as the strong CP problem. A particularly attractive solution to this fundamental problem is the so-called Peccei-Quinn mechanism [15]. One introduces an additional (pseudo-)scalar U(1) symmetric field. The underlying Peccei-Quinn U(1) symmetry is spontaneously broken –which can happen pre-inflation or post-inflation– and an axion field A acts as a (pseudo-)Goldstone boson of the broken symmetry [19, 20]. Due to the chiral anomaly the axion also couples to \mathcal{L}_Q . As a

consequence, the original potential of the axion field with its $U(1)$ symmetry breaking gets tilted and has its minimum where $(\theta + A/f_A) = 0$. This sets the coefficient in front of \mathcal{L}_Q to zero and solves the strong CP problem. Furthermore, the free parameter f_A determines the mass of the axion at vanishing or non-vanishing temperatures by $m_A^2 = \chi/f_A^2$. Here χ is the susceptibility of the topological charge. We determined its value at $T = 0$, which [17] turned out to be $\chi(T = 0) = 0.0245(24)(12)/\text{fm}^4$ in the isospin symmetric case, where the first error is statistical, the second is systematic. For the error budget see [17]. Isospin breaking results in a small, 12% correction [17], thus the physical value is $\chi(T = 0) = 0.0216(21)(11)/\text{fm}^4 = [75.6(1.8)(0.9)\text{MeV}]^4$.

In an earlier study of ours [7] we looked at $\chi(T)$ in the quenched approximation. We provided a result within the quenched framework and reached a temperature about half to one third of the necessary temperatures for axion cosmology (a similar study with somewhat less control over the systematics is [5]). To obtain a complete result one should use dynamical quarks with physical masses. Dynamical configuration production is, however, about three orders of magnitude more expensive and the $\chi(T)$ values are several orders of magnitude smaller than in the quenched case. Due to cutoff effects [17] the continuum limit is far more difficult to carry out in dynamical QCD than in the pure gauge theory [7]. All in all we estimate that the brute-force approach to provide a complete result on $\chi(T)$ in the relevant temperature region would be at least ten orders of magnitude more expensive than the result of [7]. Ref. [6] used this approach, however it turned out to be quite difficult. As we will show our result for $\chi(T)$ is many orders of magnitude smaller than that of Ref. [6] in the cosmologically relevant temperature region.

Whilst writing up our results, a paper [16] appeared with findings similar to ours, for which two remarks are in order. A common set of criteria for assessing the reliability of lattice results was set by FLAG and published in Ref. [2]. They introduced for some quantities a measure indicating how far the continuum extrapolated result is from the lattice data. Firstly, in order to control the continuum extrapolation they demand to have a maximum difference of a few percent between the finest lattice data and the continuum extrapolated result. (In other words, no extrapolation with large cutoff effects are allowed.) In Ref. [16] the cutoff effects on the direct topological susceptibility $\chi(T)$ measurements are orders of magnitude larger than the FLAG requirement (even for secondary quantities such as $\chi(T)^{1/4}$ or for the exponent they are large), whereas here we show how to keep these cutoff effects on $\chi(T)$ to the $\mathcal{O}(10\%)$ level. Secondly, FLAG requires three or more lattice spacings for controlled continuum extrapolation results. When attempting to determine $\chi(T)$ for approximately $T=3$ GeV Ref. [16] fulfills this condition for the continuum extrapolation in a narrow temperature region between 250 to 330 MeV. However, no direct measurements are carried out at large temperatures, only an extrapolation using the behaviour in this narrow range, whereas we show how to determine $\chi(T)$ all the way into the physically relevant 3 GeV region.

Since cutoff effects are huge and the CPU demand is tremendous one faces significant physics challenges and needs dramatic algorithmic developments in a large-scale simulation setup in order to provide a complete answer for $\chi(T)$, thus resolve the tension between e.g. the two References [6] and [16] (or other analytic techniques).

The huge computational demand and the physics issue behind the determination of $\chi(T)$ has two main sources. a.) The tiny topological susceptibility needs extremely long simulation threads to observe enough changes of the topological sectors and b.) In high temperature lattice QCD the most widely used actions are based on staggered quarks. When dealing with topological observables staggered quarks have very large cutoff effects.

We solve both problems and determine the continuum result for $\chi(T)$ for the entire temperature range of interest.

ad a. We summarize our solution to problem a, which we call “fixed Q integration” (for a detailed discussion see [17]). As a first step one takes a starting-point in the quark mass-temperature space, at which it is possible to determine χ using conventional methods. One such point could be e.g. the quenched theory at some low temperature (below the phase transition), or alternatively relatively large quark masses at not too high temperatures could also be used in practice. Next we determine, at this point, the weights of the various topological sectors in a given volume. This is done by taking derivatives

with respect to the gauge coupling and m_f and then by integrating these parameters one can calculate the change in the weights of the various topological sectors all the way to the new point. This provides the weights of the topological sectors and therefore χ in a new point. The technique is similar to the so-called integral method used to determine the equation of state [17].

The CPU costs of the conventional technique scale as T^8 , whereas the new “fixed Q integration” method scales as T^0 . The gain in CPU time is tremendous. This efficient technique is used to obtain the final result for $\chi(T)$. Since we work with continuum extrapolated quantities both for the ratios in the starting-point as well as for their changes, one can in principle use any action in the procedure, we will use here overlap [17] and/or staggered actions.

ad b. We summarize our solution to problem b, which we call “eigenvalue reweighting”. According to the index theorem the continuum Dirac operator has exact zero modes with definite chirality on $Q \neq 0$ gauge configurations. Staggered fermions lack these exact zero modes, which introduces large to very large lattice artefacts when determining $\chi(T)$. We developed a reweighting technique, which resolves this problem. The method is based on the determination of the lowest eigenvalues of the Dirac operator and using the ratios of the eigenvalues and the mass in a reweighting procedure. The basic idea is to substitute the smallest eigenvalue (the would be zero modes) for a given configuration of a given topological charge with its continuum value and reweight according to it. Note that if we have a zero mode, the smallest eigenvalue should be m_f in the continuum with massive fermions, whereas for staggered quarks it can be significantly higher. We correct for this by reweighting the configurations with the ratio of the lowest continuum and the corresponding staggered eigenvalue. The topological charge is measured by the Wilson-flow method. The details and a discussion on the validity of the technique is presented in [17]. Overlap fermions do not suffer from this problem and have exact zero modes. Even though they are computationally too expensive to be used for the entire project (e.g. the integration in the gauge coupling requires tens of millions of trajectories [17]) the mass integration down to the physical point has been done with overlap quarks.

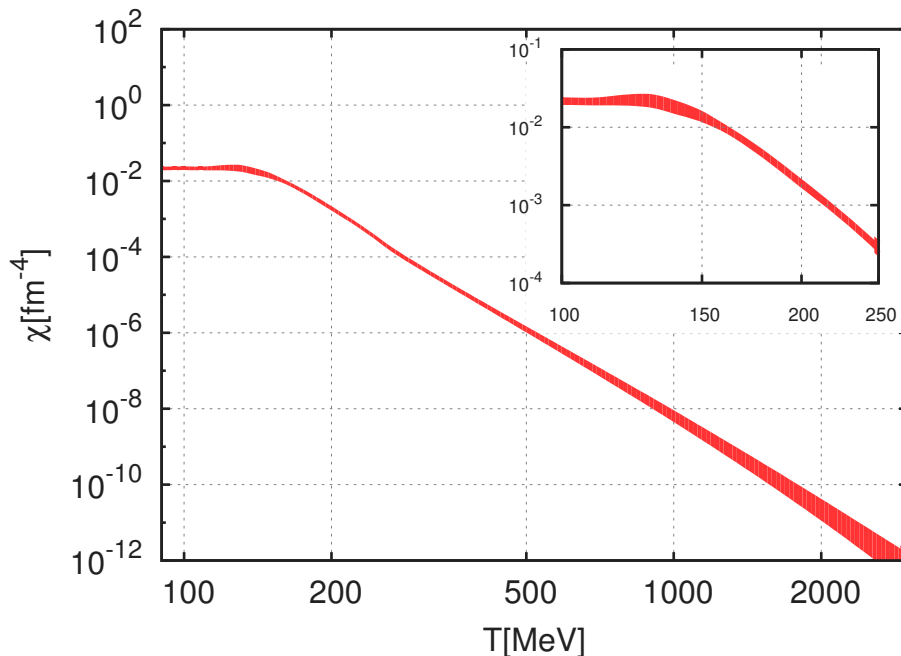


Figure 2: Continuum limit of $\chi(T)$. The insert shows the behaviour around the transition temperature. The width of the line represents the combined statistical and systematic errors. The dilute instanton gas approximation (DIGA) predicts a power behaviour of T^{-b} with $b=8.16$, which is confirmed by the lattice result for temperatures above ~ 1 GeV.

Through combining these methods one can determine $\chi(T)$ by controlling all the systematic uncertainties (see Figure 2). Therefore, removing several thousand percent cutoff effects of previous approaches

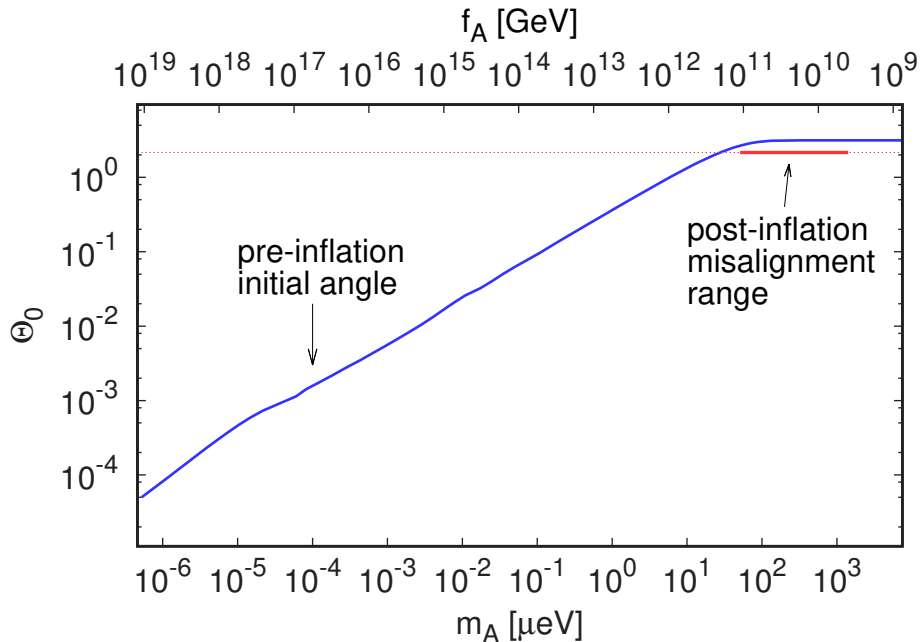


Figure 3: Relation between the axion's mass and the initial angle θ_0 in the pre-inflation scenario. The post-inflation scenario corresponds to $\theta_0 = 2.155$ with a strict lower bound on the axion's mass of $m_A = 28(2)\mu eV$. The thick red line shows our result on the axion's mass for the post-inflation case. E.g. $m_A = 50(4)\mu eV$ if one assumes that axions from the misalignment mechanism contributes 50% to dark matter. Our final estimate is $m_A = 50-1500\mu eV$ (the upper bound assumes that only 1% is the contribution of the misalignment mechanism the rest comes from other sources e.g. topological defects). For an experimental setup to detect post-inflationary axions see [17]. The slight bend around $m_A \sim 10^{-5}\mu eV$ corresponds to an oscillation temperature at the QCD transition [4, 11].

leaving a very mild $\mathcal{O}(10\%)$ continuum extrapolation to be performed. In addition, the direct determination of $\chi(T)$ all the way up to 3 GeV means that one does not have to rely on the dilute instanton gas approximation (DIGA). Note that *a posteriori* the exponent predicted by DIGA turned out to be compatible with our finding but its prefactor is off by an order of magnitude, similar to the quenched case.

As a possible application for these two cosmologically relevant lattice QCD results, we show how to calculate the amount of axionic dark matter and how it can be used to determine the mass of the axion. As we have seen, $\chi(T)$ is a rapidly decreasing function of the temperature. Thus, at high temperature m_A (which is proportional to $\chi(T)^{1/2}$) is small. In fact, much smaller than the Hubble expansion rate of the universe at that time or temperature ($H(T)$). The axion does not feel the tilt in the Peccei-Quinn Mexican hat type potential yet and it is effectively massless and frozen by the Hubble friction. As the Universe expands the temperature decreases, $\chi(T)$ increases and the axion mass also increases. In the meantime, the Hubble expansion rate –given by our equation of state– decreases. As the temperature decreases to T_{osc} the axion mass is of the same order as the Hubble constant (T_{osc} is defined by $3H(T_{\text{osc}}) = m_A(T_{\text{osc}})$). Around this time the axion field rolls down the potential, starts to oscillate around the tilted minimum and the axion number density increases to a nonzero value, thus axions as dark matter are produced. The details of this production mechanism, usually called misalignment, are quite well known (see e.g. [18, 17]).

In a post-inflationary scenario the initial value of the angle θ takes all values between $-\pi$ and π , whereas in the pre-inflationary scenario only one θ_0 angle contributes (all other values are inflated away). One should also mention that during the U(1) symmetry breaking topological strings appear which decay and also produce dark matter axions. In the pre-inflationary scenario they are inflated away. However, in the post-inflationary framework their role is more important. This sort of axion production mechanism is less well-understood and in our final results it is necessary to make some assumptions.

The possible consequences of our results on the predictions of the amount of axion dark matter can be seen in Figure 3. Here we also study cases, for which the dark matter axions are produced from the decay of unstable axionic strings (see the discussion in the figure’s caption). For the pre-inflationary Peccei-Quinn symmetry breaking scenario the axion mass determines the initial condition θ_0 of our universe.

Acknowledgments. We thank M. Dierigl, M. Giordano, S. Krieg, D. Nogradi and B. Toth for useful discussions. This project was funded by the DFG grant SFB/TR55, and by OTKA under grant OTKA-K-113034. The work of J.R. is supported by the Ramon y Cajal Fellowship 2012-10597 and FPA2015-65745-P (MINECO/FEDER). The computations were performed on JUQUEEN at Forschungszentrum Jülich (FZJ), on SuperMUC at Leibniz Supercomputing Centre in München, on Hazel Hen at the High Performance Computing Center in Stuttgart, on QPACE in Wuppertal and on GPU clusters in Wuppertal and Budapest.

Author contributions. SB and SM developed the fixed sector integral, TKG and KS the eigenvalue reweighting, FP the odd flavour overlap techniques, respectively. SB, SK, TKG, TK, SM, AP, FP and KS wrote the necessary codes, carried out the runs and determined the EoS and $\chi(T)$. JR, AP and AR calculated the DIGA prediction. K-HK and JR and AR worked out the experimental setup. ZF wrote the main paper and coordinated the project.

References

- [1] Jens O. Andersen et al. “NNLO hard-thermal-loop thermodynamics for QCD”. In: *Phys. Lett.* B696 (2011), pp. 468–472. arXiv:1009.4644 [hep-ph].
- [2] Sinya Aoki et al. “Review of lattice results concerning low-energy particle physics”. In: *Eur. Phys. J.* C74 (2014), p. 2890. arXiv:1310.8555 [hep-lat].
- [3] Y. Aoki et al. “The Order of the quantum chromodynamics transition predicted by the standard model of particle physics”. In: *Nature* 443 (2006), pp. 675–678. arXiv:hep-lat/0611014 [hep-lat].
- [4] Y. Aoki et al. “The QCD transition temperature: results with physical masses in the continuum limit II.” In: *JHEP* 06 (2009), p. 088. arXiv:0903.4155 [hep-lat].
- [5] Evan Berkowitz, Michael I. Buchoff, and Enrico Rinaldi. “Lattice QCD input for axion cosmology”. In: *Phys. Rev.* D92.3 (2015), p. 034507. arXiv:1505.07455 [hep-ph].
- [6] Claudio Bonati et al. “Axion phenomenology and θ -dependence from $N_f = 2 + 1$ lattice QCD”. In: *JHEP* 03 (2016), p. 155. arXiv:1512.06746 [hep-lat].
- [7] S. Borsanyi et al. “Axion cosmology, lattice QCD and the dilute instanton gas”. In: *Phys. Lett.* B752 (2016), pp. 175–181. arXiv:1508.06917 [hep-lat].
- [8] Sz. Borsanyi et al. “Ab initio calculation of the neutron-proton mass difference”. In: *Science* 347 (2015), pp. 1452–1455. arXiv:1406.4088 [hep-lat].
- [9] Sz. Borsanyi et al. “Precision SU(3) lattice thermodynamics for a large temperature range”. In: *JHEP* 07 (2012), p. 056. arXiv:1204.6184 [hep-lat].
- [10] Szabolcs Borsanyi et al. “High-precision scale setting in lattice QCD”. In: *JHEP* 09 (2012), p. 010. arXiv:1203.4469 [hep-lat].
- [11] Szabolcs Borsanyi et al. “Is there still any T_c mystery in lattice QCD? Results with physical masses in the continuum limit III”. In: *JHEP* 09 (2010), p. 073. arXiv:1005.3508 [hep-lat].
- [12] S. Durr et al. “Ab-Initio Determination of Light Hadron Masses”. In: *Science* 322 (2008), pp. 1224–1227. arXiv:0906.3599 [hep-lat].

- [13] M. Laine and M. Meyer. “Standard Model thermodynamics across the electroweak crossover”. In: *JCAP* 1507.07 (2015), p. 035. arXiv:1503.04935 [hep-ph].
- [14] K. A. Olive et al. “Review of Particle Physics”. In: *Chin. Phys.* C38 (2014), p. 090001.
- [15] R. D. Peccei and Helen R. Quinn. “CP Conservation in the Presence of Instantons”. In: *Phys. Rev. Lett.* 38 (1977), pp. 1440–1443.
- [16] Peter Petreczky, Hans-Peter Schadler, and Sayantan Sharma. “The topological susceptibility in finite temperature QCD and axion cosmology”. In: (2016). arXiv:1606.03145 [hep-lat].
- [17] *Supplementary Information.*
- [18] Olivier Wantz and E. P. S. Shellard. “Axion Cosmology Revisited”. In: *Phys. Rev.* D82 (2010), p. 123508. arXiv:0910.1066 [astro-ph.CO].
- [19] Steven Weinberg. “A New Light Boson?” In: *Phys. Rev. Lett.* 40 (1978), pp. 223–226.
- [20] Frank Wilczek. “Problem of Strong p and t Invariance in the Presence of Instantons”. In: *Phys. Rev. Lett.* 40 (1978), pp. 279–282.

Supplementary Information

Lattice QCD for Cosmology

Sz. Borsanyi¹, Z. Fodor^{1,2,3}, K.-H. Kampert¹, S. D. Katz^{3,4}, T. Kawanai², T. G. Kovacs⁵, S. W. Mages², A. Pasztor¹, F. Pittler^{3,4}, J. Redondo^{6,7}, A. Ringwald⁸, K. K. Szabo^{1,2}

¹ Department of Physics, University of Wuppertal, D-42119 Wuppertal, Germany

² Jülich Supercomputing Centre, Forschungszentrum Jülich, D-52428 Jülich, Germany

³ Institute for Theoretical Physics, Eötvös University, H-1117 Budapest, Hungary

⁴ MTA-ELTE Lendület Lattice Gauge Theory Research Group, H-1117 Budapest, Hungary

⁵ Institute for Nuclear Research of the Hungarian Academy of Sciences, H-4026 Debrecen, Hungary

⁶ University of Zaragoza, E-50009 Zaragoza, Spain

⁷ Max Planck Institut für Physik, D-80803, Germany

⁸ Deutsches Elektronen-Synchrotron DESY, D-22607 Hamburg, Germany

In the following sections we provide details of the work presented in the main paper. In Section S1 we summarize the simulation setup for our staggered lattice QCD calculations. This is the basis for the determination of the equation of state and one of the two key elements of the topological susceptibility calculations. Section S2 contains the technical details for the EoS focusing on the lattice QCD part, Section S3 presents the perturbative methods to determine the QCD equation of state. Section S4 lists the complete results for the equation of state both in the $n_f = 2 + 1 + 1$ and in the $n_f = 2 + 1 + 1 + 1$ frameworks. For cosmological applications we give the effective degrees of freedom for a wide temperature range, starting from 1 MeV all the way up to 500 GeV.

In Section S5 we discuss our simulations with overlap fermions using even and odd number of flavors. The line of constant physics is determined. In order to reduce cutoff effects for the topological susceptibility we introduced a eigenvalue reweighting method, which is presented in Section S6. This method significantly extends the quark-mass vs. temperature parameter space, which can be reached by lattice QCD. To reach even higher temperatures a second method was invented: the fixed sector integral method. The method is illustrated first in the quenched theory Section S7. In the next two Sections S8, S9 we apply the method for the real physical case using staggered and overlap fermions. It is interesting to mention that for the determination of the topological susceptibility overlap fermions turned out to be the less CPU-demanding fermion formulation. The non-perturbative lattice findings are compared with those of the dilute instanton gas approximation.

In Section S11 we use the equation of state and the topological susceptibility results to make predictions for axion cosmology. Both the pre-inflation and the post-inflation Peccei-Quinn symmetry breaking scenarios are discussed. In Section S12 we present experimental setups, which could explore axions in the predicted post-inflation mass range.

S1 Staggered simulations

For the majority of the results in this paper we use a four flavor staggered action with 4 levels of stout smearing. The action parameters are given in [S11]. The quark masses and the lattice spacing are functions of the gauge coupling:

$$m_s = m_s^{st}(\beta), \quad m_{ud} = R \cdot m_s^{st}(\beta), \quad m_c = C \cdot m_s^{st}(\beta), \quad a = a^{st}(\beta), \quad (\text{S1})$$

these sets of functions are called the Lines of Constant Physics (LCP), and are also given in [S11]. For the quark mass ratios we use $1/R = 27.63$ and $C = 11.85$, which are in good agreement with recent large scale lattice QCD simulations [S34, S35].

In addition to the algorithms mentioned in [S11] we now used a force gradient time integrator [S25, S77] to generate gauge configurations.

Throughout this paper the index f labels the quark flavors, $f = \{ud, s, c\}$, and N_t and N_s are the number of lattice points in the temporal and spatial directions. The temperature T is introduced in the fixed- N_t approach of thermodynamics, ie. $T = (aN_t)^{-1}$, which can be changed by the parameter β while N_t and N_s are fixed. The quark masses m_f are given in lattice units, if not indicated otherwise.

Two different sets of staggered ensembles are used in this paper: a physical $n_f = 2 + 1 + 1$ flavor simulation set and a three flavor symmetric $n_f = 3 + 1$ simulation set. In the following we describe them in more detail.

S1.1 Physical point $n_f = 2 + 1 + 1$ simulations

The lattice geometries and statistics of the $n_f = 2 + 1 + 1$ simulations at zero and finite temperature, are described in [S11]. The quark masses are set to their physical values. In a few cases we increased the statistics of the existing ensembles.

On these ensembles we evaluated the clover definition of the topological charge Q after applying a Wilson-flow [S64]. We used an adaptive step size integration scheme to reduce the computational time. For the finite temperature ensembles we chose a flow time of $(8T^2)^{-1}$, whereas at zero temperature $t = w_0^2$, where the w_0 scale is defined in [S19]. In the systematic error analyses we usually allow for a variation in the flow time. As was shown in our quenched study at finite temperature [S14] the susceptibility reaches a plateau for large flow times. The above choices are nicely in this plateau region even on coarse lattices. The so obtained charge is not necessarily an integer number. To evaluate the topological susceptibility we usually considered both definitions: with and without rounding the topological charge, the difference between the two is used to estimate systematic errors. When the $\langle Q^2 \rangle$ was large, or close to the continuum limit the rounding did not change the results significantly.

For some of the finite temperature ensembles we also calculated the eigenvalues and eigenvectors of the staggered Dirac-operator, D_{st} . For this purpose we used a variant of the symmetric Krylov-Schur algorithm described in [S50]. The chirality of these eigenmodes was determined using the staggered taste-singlet γ_5 -matrix, which is described in [S37].

S1.2 Topological susceptibility in the vacuum

Due to the index theorem, in the background of a gauge field with non-zero topological charge Q , the Dirac operator has at least $|Q|$ exact zero eigenvalues [S6]. These zero-modes have chirality $+1$ or -1 . This is, however, true only in the continuum theory. On the lattice, due to cut-off effects, a non-chiral Dirac operator, like the staggered operator, does not have exact zero eigenvalues, only close to zero small eigenvalues. Also the magnitude of the chirality of these would-be zero modes is smaller than unity. In the continuum limit of the lattice model the would-be zero modes become zero modes with chirality of unit magnitude. However, at any nonzero lattice spacing the lack of exact zero modes results in cut-off effects that can be unexpectedly large for some observables. This is shown for the topological susceptibility in

Figure S1, where χ is plotted as function of the lattice spacing squared. The result changes almost an order of magnitude by moving from the coarsest to the finest lattice spacing on the plot.

At zero temperature χ is proportional to the pion mass squared in the continuum. On the lattice with staggered fermions it is expected, that χ will be proportional to the mass squared of the taste singlet pion. For the staggered chiral perturbation theory analysis of χ , see [S9]. Thus it is natural to rescale χ with the pseudo-Goldstone mass squared over the taste-singlet pion mass squared. Since in our $n_f = 2 + 1 + 1$ simulations physical pion pseudo-Goldstone masses are used, in Figure S1 we plot χ multiplied by $(m_{\pi,phys}/m_{\pi,ts})^2$, where $m_{\pi,ts}$ is the taste-single pion mass [S48, S53]. The data shows much smaller cut-off effects, than without multiplication and a nice a^2 -scaling sets in starting from a lattice spacing of about 0.1 fm. The continuum extrapolated value is

$$\chi(T = 0) = 0.0245(24)_{\text{stat}}(03)_{\text{flow}}(12)_{\text{cont}}/\text{fm}^4, \quad (\text{S2})$$

the first error is statistical. The second, systematic error comes from varying the definition of the charge, i.e. the flow time at which the charge is measured. The third error comes from changing the upper limit of the lattice spacing range in the fit. According to leading order chiral perturbation theory

$$\Sigma/\chi = 2m_{ud}^{-1} + m_s^{-1} + \dots \quad (\text{S3})$$

where Σ is the condensate in the chiral limit and the ellipses stand for higher order terms. Using the values for quark masses and the condensate from [S34, S40] we obtain $\chi_{\text{LO}} = 0.0224(12)/\text{fm}^4$ in the isospin symmetric limit, which is in good agreement with Equation (S2). For isospin corrections see Section S10.

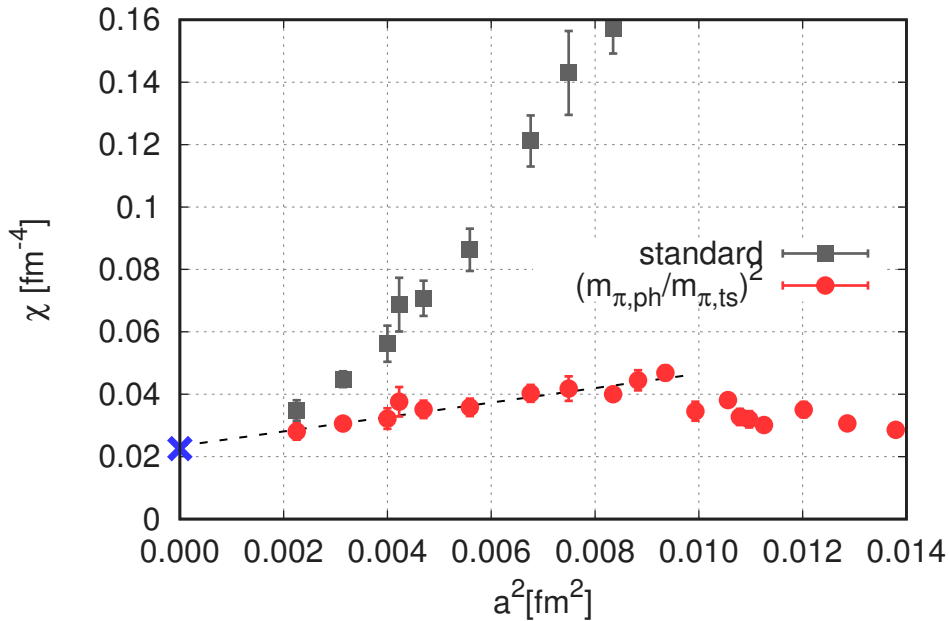


Figure S1: Lattice spacing dependence of the zero temperature topological susceptibility. The grey squares are obtained with the standard approach, the red circles after dividing by the taste singlet pion mass squared. The line is a linear fit. The blue cross corresponds to leading order chiral perturbation theory. The plot shows $n_f = 2 + 1 + 1$ flavor staggered simulations at zero temperature.

For the high temperature region the cut-off effects are not supposed to be described by chiral perturbation theory. Other techniques are required to get the large cut-off effects under control, such a technique is presented in Section S6.

S1.3 Three flavor symmetric $n_f = 3 + 1$ simulations

As it will be described in later Sections as an intermediate step we used results from simulations at the three flavor symmetric point, i.e. where all the light-quark masses are set to the physical strange quark mass $m_s^{st}(\beta)$ and the ratio of the charm and the 3 degenerate flavour masses is $C = 11.85$.

In this theory for each gauge coupling β one can measure the pseudoscalar mass m_π and the Wilson-flow based w_0 -scale. The dimensionless product $m_\pi w_0$ as well as the w_0 in physical units, i.e. $w_0 a^{st}(\beta)$, have well defined continuum limits, since the $n_f = 3 + 1$ and $n_f = 2 + 1 + 1$ theories differ only in the masses of quarks, that do not play a role in the ultraviolet behaviour of those theories. So we end up with the continuum value of the three flavor pseudoscalar mass $m_\pi^{(3)}$, and that of the w_0 -scale $w_0^{(3)}$. We performed $n_f = 3 + 1$ flavor zero temperature simulations in 64×32^3 volumes for seven lattice spacings ranging between $a = 0.15$ and 0.06 fm. We measured w_0 and $m_\pi w_0$ and performed a continuum extrapolation. This is shown in Figure S2. We obtain the continuum values

$$w_0^{(3)} = 0.153(1) \text{ fm} \quad \text{and} \quad m_\pi^{(3)} = 712(5) \text{ MeV}. \quad (\text{S4})$$

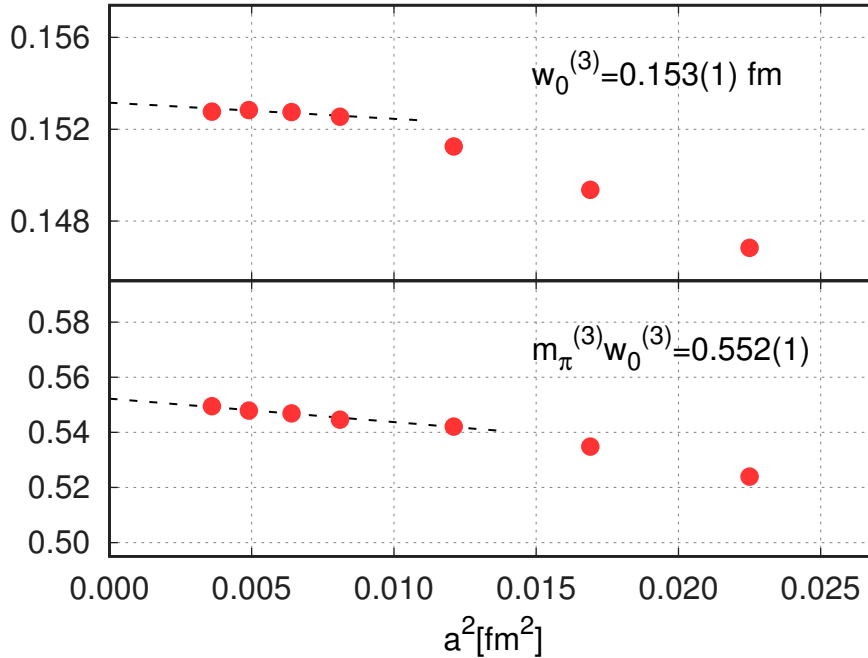


Figure S2: Continuum extrapolations of the w_0 -scale in lattice units (up) and $m_\pi w_0$ (down). The plots show $n_f = 3 + 1$ flavor staggered simulations at zero temperature.

In the $n_f = 3 + 1$ theory we performed finite temperature simulations and measured the same observables as in the $n_f = 2 + 1 + 1$ case. Additionally, for the new integral method described in Section S8, we generated configurations at fixed topology. This was achieved by measuring the topological charge after each Hybrid Monte-Carlo trajectory and rejecting the configuration in case of topology change. The typical acceptance rates were 40% on the coarsest lattices and higher on the finer ones. The finite temperature ensembles, unconstrained and fixed topology, are listed in the analysis section, Section S10.

S2 Lattice methods for the equation of state

For 2+1 dynamical flavors with physical quark masses we calculated the equation of state in Refs. [S22, S18]. This result was later confirmed in Ref. [S8]. The additional contribution of the charm quark has since been estimated by several authors using perturbation theory [S61], partially quenched lattice simulations [S24, S63] and from simulations with four non-degenerate quarks, but unphysical masses [S23, S10]. However, the final results must come from a dynamical simulation where all quark masses assume their physical values [S21].

We meet this challenge by using the 2+1+1 flavor staggered action (with 4 levels of stout smearing) that we introduced in Section S1. The action parameters, the bare quark masses and the tuning procedure, as well as the lattice geometries and statistics have been given in detail in Ref. [S11].

We calculate the EoS, ie. the temperature dependence of the pressure p , energy density ρ and entropy density s , from the trace anomaly $I(T)$. This latter is defined as $I = \rho - 3p$, and on the lattice it is given by the following formula:

$$\frac{I(T)}{T^4} = \frac{\rho - 3p}{T^4} = N_t^4 R_\beta \left[\frac{\partial}{\partial \beta} + \sum_f R_f m_f \frac{\partial}{\partial m_f} \right] \frac{\log Z[m_u, m_d, m_s, \dots]}{N_t N_s^3} \quad (\text{S5})$$

with

$$R_\beta = -\frac{d\beta}{d \log a} \quad \text{and} \quad R_f = \frac{d \log m_f}{d\beta}, \quad f = u, d, s, \dots \quad (\text{S6})$$

The derivatives of $\log Z$ with respect to the gauge coupling β and the quark masses m_f are easily accessible observables on the lattice: they are the gauge action S_g and the chiral condensate, respectively.

The R_β and R_f functions describe the running of the coupling and the mass. They can be obtained by differentiating the LCP in Eq. (S1). Since R_β appears as a factor in front of the final result, the systematics of the determination of the LCP directly distorts the trace anomaly. To estimate this uncertainty we use two different LCP's, one based on the w_0 -scale and another other on the pion decay constant f_π , which are supposed to give the same continuum limit, but can differ by lattice artefacts. We calculate R_β both from the w_0 and the f_π -based LCP and include the difference in the systematic error estimate. Let us mention, that the R_β and R_f functions are universal at low orders of perturbation theory: e.g. for QCD with n_f flavors we have $R_\beta = 12\beta_0 + 72\beta_1/\beta + \mathcal{O}(\beta^{-2})$ with $\beta_0 = (33 - 2n_f)/48\pi^2$ and $\beta_1 = (306 - 38n_f)/768\pi^2$.

There is an additive, temperature independent divergence in the trace anomaly. In the standard procedure, that we also followed in Ref. [S22], each finite temperature simulation is accompanied by a zero temperature ensemble. The trace anomaly is then calculated on both sets of configurations, their difference is the physical result. This defines a renormalization scheme where the zero temperature pressure and energy density both vanish. Using the configurations already introduced in Ref. [S11] and applying the standard method that we also described in Ref. [S18] we calculated the trace anomaly, as shown in the left panel of Fig. S3.

This strategy is bound to fail at high temperatures. Since the temperature on the lattice is given by $T = (aN_t)^{-1}$, high temperatures can only be reached with very fine lattices. As the lattice spacing is reduced, the autocorrelation times for zero temperature simulations rise, and the costs of these simulations explode beyond feasibility. Notice, however, that the renormalization constant is accessible not only from zero temperature simulations, but from any finite temperature data point that was taken using the same gauge coupling and quark masses.

In our approach we generate a renormalization ensemble for each finite temperature ensemble at exactly half of its temperature with the same physical volume and matching bare parameters. The trace anomaly difference is then divergence-free. We have tested this idea in the quark-less theory in Refs. [S43, S16].

Thus we determine the quantity $[I(T) - I(T/2)]/T^4$ in the temperature range 250...1000 MeV for four resolutions $N_t = 6, 8, 10$ and 12. The volumes are selected such that $LT_c \gtrsim 2$. In Fig. S3 we show this subtracted trace anomaly and its continuum limit.

We now turn to our systematic uncertainties. We use the histogram method to estimate the systematic

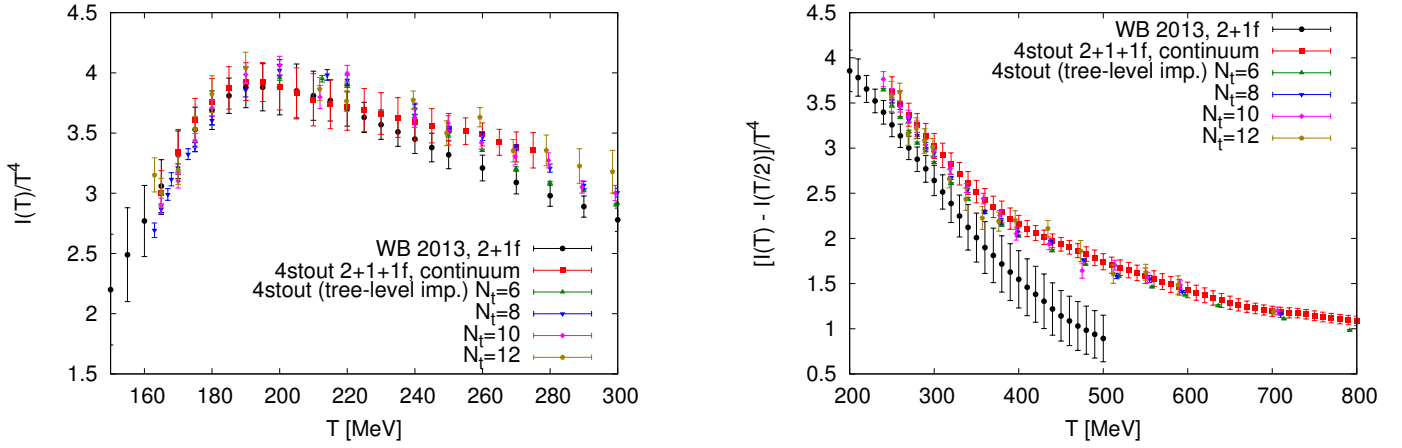


Figure S3: The trace anomaly renormalized with zero temperature simulations (left panel) and the subtracted trace anomaly (right panel) in the 2+1+1 and 2+1 flavor theories. For $T < 300$ MeV the two results agree within our uncertainty. We also show the individual lattice resolutions ($N_t = 6 \dots 12$) that contribute to the continuum limit.

errors, this means that we analyze our data in various plausible ways and form a histogram of the results. The median gives a mean, the central 68% defines the systematic error [S33]. Here we use uniform weights dropping the continuum results where the fit quality of the continuum limit was below 0.1. For the details of the analysis we largely follow our earlier work in Ref. [S11]: we interpolated the lattice data in temperature and then we performed a continuum extrapolation in $1/N_t^2$ temperature by temperature. The error bars in Fig. S3 combine the statistical and systematic errors, the latter we estimate by varying the scale setting prescription (w_0 -based or f_π -based scale setting), the observable (subtracted trace anomaly or its reciprocal), the interpolation, and whether or not we use tree-level improvement prior to the continuum extrapolation [S22].

Then we use the trace anomaly data with zero temperature renormalization from the left panel of Fig. S3 and extend it towards lower temperatures from our already established 2+1 flavor equation of state result. Then we can calculate $I(T)/T^4$ from the continuum limit of $[I(T) - I(T/2)]/T^4$ using the formula:

$$\frac{I(T)}{T^4} = \sum_{k=0}^{n-1} 2^{-4k} \frac{I(T/2^k) - I(T/2^{k+1})}{(T/2^k)^4} + 2^{-4n} \frac{I(T/2^n)}{(T/2^n)^4}, \quad (S7)$$

where n is the smallest non-negative integer with $T/2^n < 250$ MeV.

The temperature integral of $I(T)/T^5$ gives the normalized pressure $p(T)/T^4$. The energy density and entropy density are then obtained from the thermodynamic relations: $\rho = I + 3p$ and $sT = \rho + p$, respectively.

S3 The QCD equation of state in the perturbative regime

S3.1 Massless perturbation theory

Recent progress in Hard Thermal Loop (HTL) perturbation theory has provided for a next-to-next-to-leading-order (NNLO) result for the free energy, which is in fair agreement with our data both for the 2+1 theory [S1] and also for the 2+1+1 flavor theory for high enough temperatures. The trace anomaly and the pressure are shown in Fig. S4 for both cases.

We also show a comparison to the results of conventional perturbation theory with four massless quarks in Fig. S5. Here $g = \sqrt{4\pi\alpha_s}$ is the QCD coupling constant. The completely known fifth order [S78] result is in good agreement with the lattice data. Note, that whereas the perturbative result treats even the charm quark massless, the lattice EoS includes the mass effects correctly.

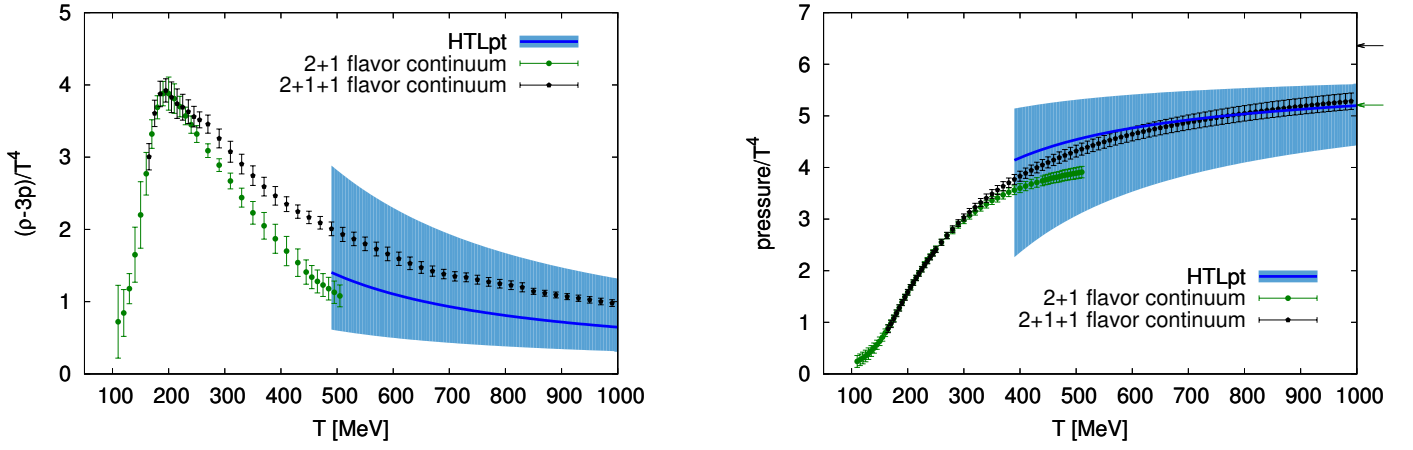


Figure S4: The QCD trace anomaly and pressure in the 2+1+1 and 2+1 flavor theories. We also give the four flavor NNLO HTL result at high temperatures [S1].

S3.2 Charm mass threshold in the QCD equation of state

Thanks to the lattice data that we have generated, we can present non-perturbative results for the charm quark contribution. It is instructive to study the inclusion of the charm quark in detail. This way we can design an analytical technique for the inclusion of the bottom quark, for which the standard formulation of lattice QCD is computationally not feasible.

The quark mass threshold for the charm quark entering the EoS has already been estimated in Ref. [S61]. There, the effect of a heavy quark was calculated to a low order of perturbation theory. This effect was expressed as a pressure ratio between QCD with three light and one heavy flavor and QCD with only three light flavors. When that paper was completed the lattice result for the QCD equation of state was not yet available, but the perturbative methods were already in an advanced state.

Despite the known difficulties of perturbation theory the estimate of Ref. [S61] is very close to our lattice result if we plot the ratio of the pressure with and without the charm quark included. We show our lattice data together with the perturbative estimate in Fig. S6.

Though the individual values for the 2+1+1 and 2+1 flavor pressures of [S61] are not very accurate, their ratio describes well the lattice result. This is true both for the leading and for the next-to-leading order results (See Fig. S6).

The tree-level charm correction is given by

$$\frac{p^{(2+1+1)}(T)}{p^{(2+1)}(T)} = \frac{SB(3) + F_Q(m_c/T)}{SB(3)} \quad (S8)$$

where $SB(n_f)$ is the Stefan Boltzmann limit of the n_f flavor theory, and $F_Q(m/T)T^4$ is the free energy density of a free quark field with mass m . In this paper we used the \overline{MS} mass $m_c(m_c) = 1.29$ GeV [S68].

Order g^2 in the ratio of Fig. S6 starts to be important correction below a temperature of about $2 - 3T_c^{QCD}$ temperature. Near $2T_c$ the difference between the two approximations is 3%. The difference reduces to 0.2% at 1 GeV up to which point we have lattice data.

S3.3 Bottom mass threshold in the QCD equation of state

In the previous discussion we saw that even the tree-level quark mass threshold gives a correct estimate for the equation of state. This allows us to introduce the bottom threshold along the same lines.

First, we remark that one can write the charm threshold relative to the 2 + 1 + 1 flavor theory:

$$\frac{p^{(2+1+1)}(T)}{p^{(2+1+1)}(T)|_{m_c=0}} = \frac{SB(3) + F_Q(m_c/T)}{SB(4)}. \quad (S9)$$

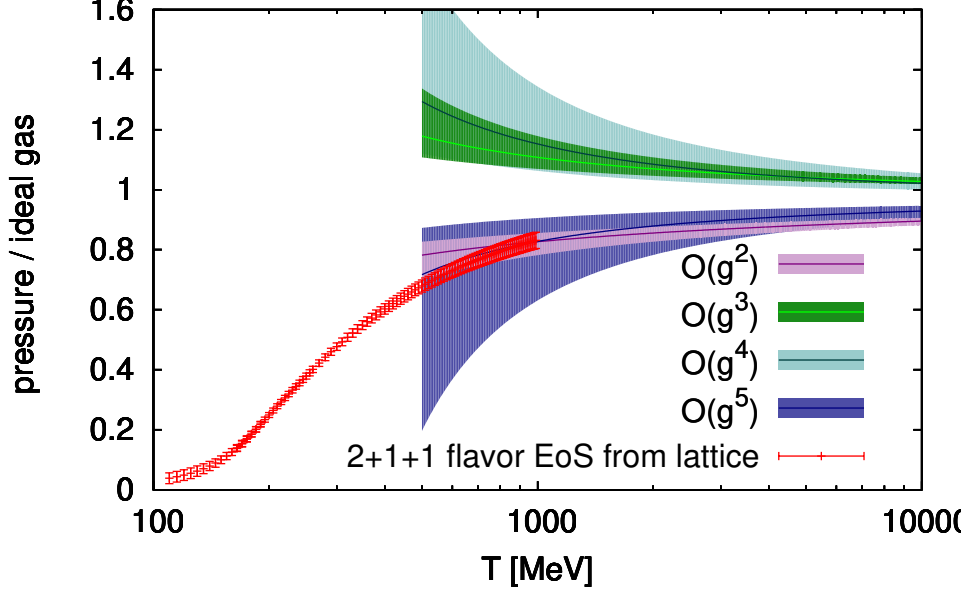


Figure S5: The QCD pressure for 2+1+1 flavors together with various orders of conventional perturbation theory. The renormalization scale is varied in the range $\mu = (1 \dots 4)\pi T$. The middle lines correspond to $\mu = 2\pi T$.

The error of not using the g^2 order is about 0.2% of the total QCD contribution at 1 GeV.

From the lattice data we have $p^{(2+1+1)}(T)$. Using Eq. (S9) we can calculate the QCD pressure for the theory with four light quarks. Perturbation theory can provide just that, at least, if the temperature is high enough. In the following we demonstrate that it is possible to use a perturbative formula that matches our lattice-based $p^{(2+1+1)}(T)|_{m_c=0}$ already from approx. 500 MeV, i.e. below the bottom threshold.

The perturbative results have a mild dependence on the choice of the Λ parameter, here we use the standard $\Lambda_{\overline{MS}} = 290$ MeV value for $n_f = 4$ [S68].

The highest fully known order for the QCD pressure is $g^6 \log g$ [S56]. The coefficient of the g^6 order is not known analytically, but the missing term, q_c (following the notation of Ref. [S56]) can be fitted against lattice data. We fix the renormalization scale to $\mu = 2\pi T$. This fitting method has already been applied for the Yang-Mills theory [S56, S16]. The order g^6 result describes our pressure data at 1 GeV within errors if $-3400 < q_c < -2600$, for the trace anomaly we have $-3200 < q_c < -2800$. We propagate this uncertainty into the perturbative result, keeping the range $2700 < -q_c < 3200$. We show the fitted curves for the central choice, $q_c = -3000$, in Fig. S7. The result in the plot has already been converted to the case of a massive charm. Notice, that both for the trace anomaly and for the pressure the $\mathcal{O}(g^6)$ perturbative result follows the lattice data already from 500 MeV.

Having a pressure and trace anomaly function that is valid for the 2+1+1 flavor theory, that agrees with the lattice data below 1 GeV and provide a perturbatively correct continuation towards high temperatures we can proceed to include the effect of the bottom quark. The tree-level correction for the bottom quark reads

$$p^{(2+1+1+1)}(T) = p^{(2+1+1)}(T) \frac{SB(4) + F_Q(m_b/T)}{SB(4)} \quad (\text{S10})$$

where $m_b(m_b) = 4.18$ GeV is the bottom mass [S68]. Eq. (S10) works beyond the bottom threshold, too, since the ratio of the perturbative massless 4 and 5 flavor pressure is, to a very good approximation, equal to the ratio of the Stefan-Boltzmann limits. Comparing the free energy up to order g^5 this statement holds to 0.3% accuracy in the entire relevant temperature range. (For earlier formulations of this idea see Refs. [S57, S51].)

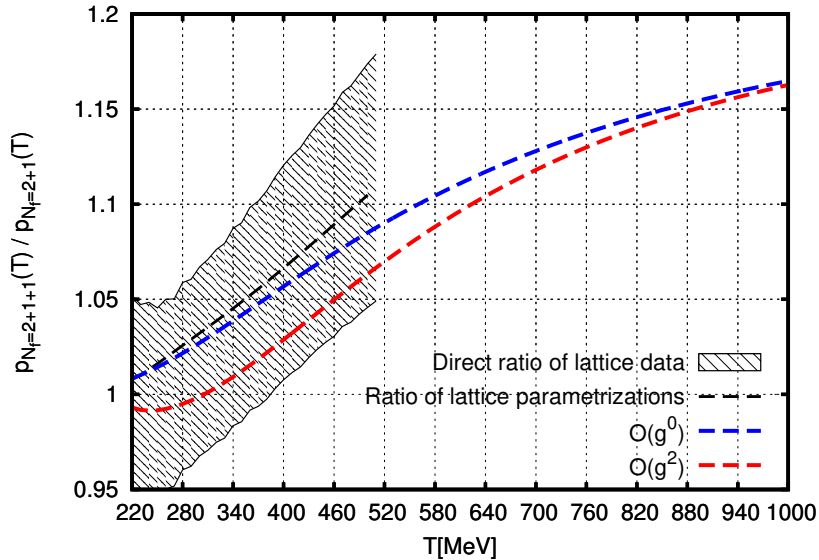


Figure S6: The ratio of the pressure between 2+1+1 flavor and 2+1 flavor theories from our lattice simulations. Note that the parametrization smoothly approaches 1 as we further decrease the temperature (see Eq. (S11)). We also show two perturbative estimates for the ratio of the pressure functions. The tree-level ($\mathcal{O}(g^0)$) estimate (see text), and the $\mathcal{O}(g^2)$ estimate of Ref. [S61]. Both of these estimates show agreement with the lattice data within our accuracy.

Since the massless four flavor perturbative result is used as a starting point for both heavy quarks we have a fair approximation between 500 MeV and 10 GeV with one analytical formula.

S4 Equation of state for 2+1+1 flavor and 2+1+1+1 flavor QCD and for the whole Standard Model

It has been a longstanding challenge to determine the pressure, energy density, and the number of effective degrees of freedom as a function of the temperature from first principles. This is the equation of state of the universe. Cosmology requires this information over a temperature range of many orders of magnitude, ranging from beyond the electroweak scale down to the MeV scale [S74, S51].

As the Universe covers this broad temperature range it passes through several epochs, each with a different dominant interaction. We restrict our study here to the Standard Model of particle physics. At the high end of the temperature range of interest there is the electroweak phase transition at a temperature of about 160 GeV [S31]. The equation of state of the electroweak transition has been worked out in Refs. [S61, S60, S59] with the intent to provide a description for the entire Standard Model. For this reason the contribution from all other degrees of freedom (i.e. up to the bottom quark) had to be estimated. While the photon, neutrinos and leptons can easily be described as practically free particles, the QCD part requires a non-perturbative approach. This was not available when Refs. [S61, S60] were published.

In this paper we add the last missing piece to the cosmological equation of state: the QCD contribution. In this section we give the results of our efforts for the 2+1+1 and 2+1+1+1 flavor theories separately. Finally, we combine all the elements of the Standard Model and present the number of effective degrees of freedom from the energy density and entropy ($g_\rho(T)$ and $g_s(T)$) in the full temperature range.

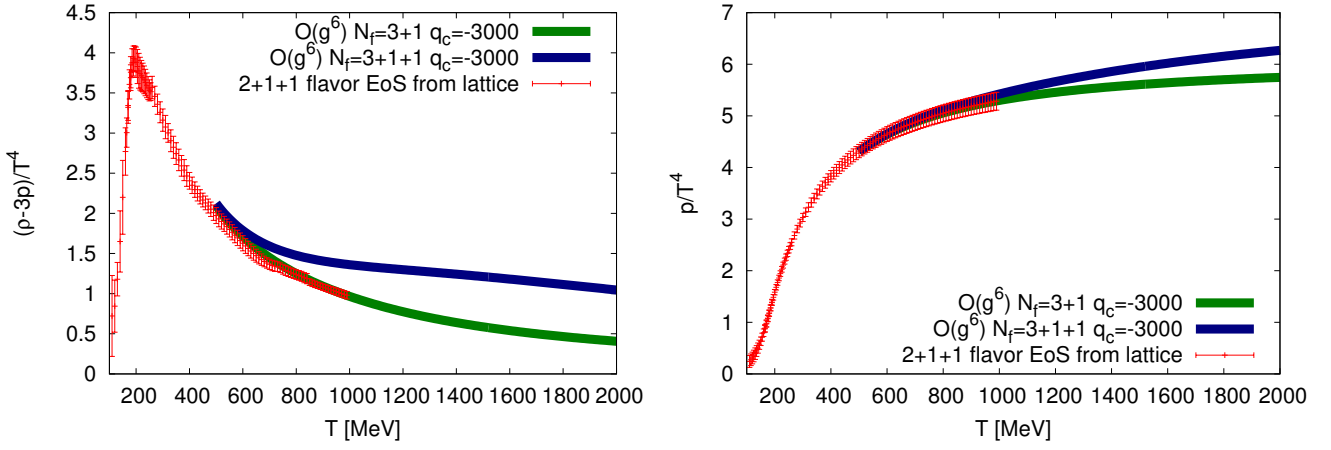


Figure S7: The lattice result for the 2+1+1 flavor QCD pressure together with the fitted value of the g^6 order. We included the charm mass at tree-level. The perturbative result agrees with the data from about 500 MeV temperature. Using the same fitted coefficient we also calculated the effect of the bottom quark with the same method. The blue curve shows the EoS including the bottom contribution.

S4.1 The 2+1+1 flavor QCD equation of state

Now we show the complete result obtained from $n_f = 2 + 1 + 1$ lattice QCD. Figure S8 depicts the trace anomaly (left panel) and pressure (right panel). For comparison the 2+1 flavor results are also shown.

Plotting p/T^4 (which is the normalized free energy density), we can compare our result to other approaches. At low temperatures the Hadron Resonance Gas model (using the 2014 PDG spectrum) gives a good description of the lattice data. This was already observed in Ref. [S18].

In Ref. [S18] we gave a simple parametrization for the 2+1 flavor equation of state. Here we update the 2+1 flavor parameters and provide a parametrization that covers the 100-1000 MeV temperature range and describes the 2+1+1 lattice data, i.e. including the effect of the charm quark. As before, the parametrizing formula reads

$$\frac{I(T)}{T^4} = \exp(-h_1/t - h_2/t^2) \cdot \left(h_0 + f_0 \frac{\tanh(f_1 \cdot t + f_2) + 1}{1 + g_1 \cdot t + g_2 \cdot t^2} \right), \quad (\text{S11})$$

with $t = T/200$ MeV. The parameters are given in Table. S1, the resulting curves are shown in Fig. S8. For completeness the $n_f = 2 + 1$ parametrization is also shown.

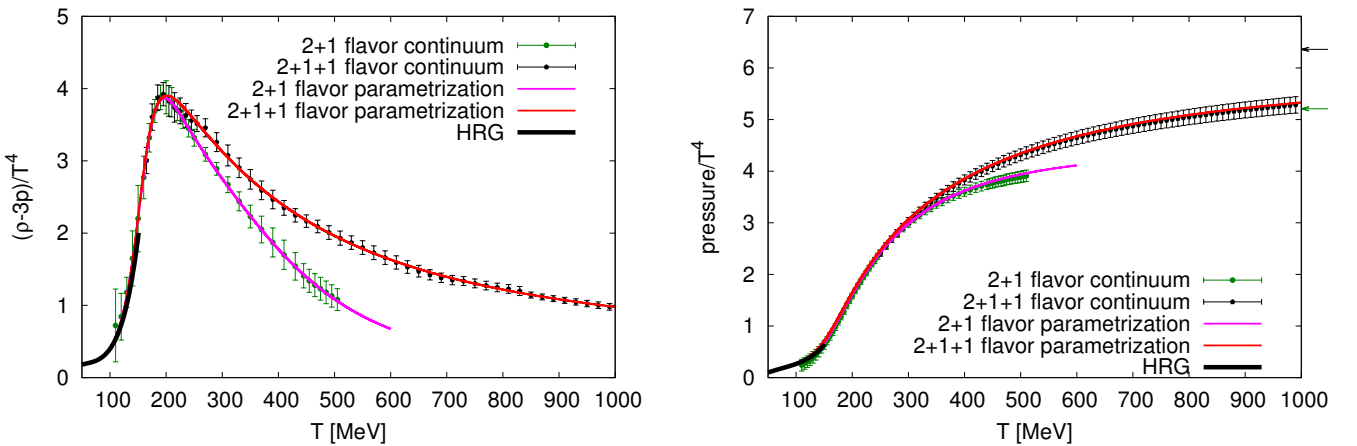


Figure S8: The QCD trace anomaly and pressure in the 2+1+1 and 2+1 flavor theories in our parametrization Eq. (S11). We also show the Hadron Resonance Gas model's prediction for comparison.

	h_0	h_1	h_2	f_0	f_1	f_2	g_1	g_2
2+1+1 flavors	0.353	-1.04	0.534	1.75	6.80	-5.18	0.525	0.160
2+1 flavors	-0.00433	-1.00	-0.288	0.293	6.10	-4.90	-0.787	0.289

Table S1: Constants for our parametrization of the trace anomaly in Eq. (S11).

S4.2 The 2+1+1+1 flavor QCD equation of state

Here we present our final result on 2+1+1+1 flavor QCD. The bottom threshold has been added as described in Sec. S3.3. We use the 2+1 flavor lattice results up to 250 MeV, 2+1+1 flavor data up to 500 MeV. In the range 500...1000 MeV we observed that our $\mathcal{O}(\alpha_s^3)$ order perturbative result agrees very well with the 2+1+1 flavor lattice data. This justifies the use of the $\mathcal{O}(\alpha_s^3)$ formula to include the effect of the bottom quark as described in Sec. S3.3. The effect of the bottom quark starts at a temperature of about 600 MeV. See Fig. S7.

Because of its large mass the top quark can only be included in the framework of the electroweak theory. Thus the calculation of the bottom quark's effect completes the discussion of the QCD contribution to the equation of state. The resulting thermodynamic functions are shown in Fig. S9.

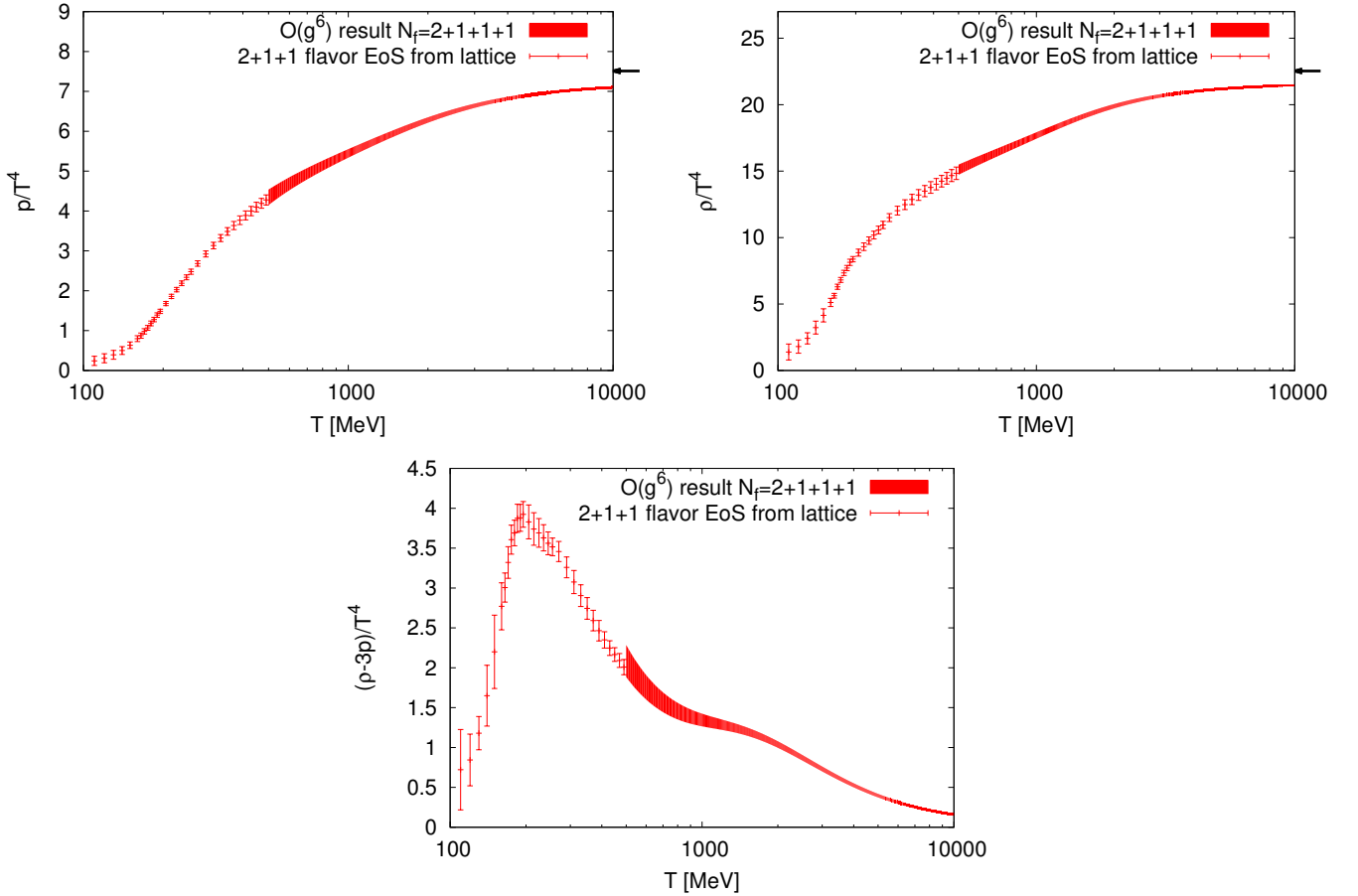


Figure S9: The QCD pressure, energy density and the trace anomaly in the 2+1+1+1 theory.

S4.3 The Standard Model

Now we are in the position to construct an equation of state that gives a good description over the entire temperature range of the Standard Model.

In particular, we compute the effective number of degrees of freedom. This is defined by the energy density or the entropy normalized by the Stefan-Boltzmann limit of a single scalar field:

$$g_\rho(T) = \rho(T) \frac{30}{\pi^2 T^4}, \quad g_s(T) = s(T) \frac{45}{2\pi^2 T^3}. \quad (\text{S12})$$

(These quantities are not to be confused with the strong coupling constant of the previous section.)

The final result for the pressure is the combination of the following contributions, the energy density and the entropy density can be obtained from the pressure using the thermodynamical identities

$$\rho(T) = 3p(T) + T^5 \frac{dp(T)/T^4}{dT}, \quad s(T)T = \rho(T) + p(T). \quad (\text{S13})$$

a) Photons and neutrinos

We treat these light particles in their Stefan-Boltzmann limit, assuming three generations of left-handed neutrinos. In this paper we work out the equilibrium equation of state, the neutrinos give a trivial contribution of $p_\nu/T^4 = \frac{7}{8} \cdot 2 \cdot 3 \cdot \pi^2/90$, for the photons we have $p_\gamma/T^4 = \pi^2/45$.

b) Charged leptons

We sum the free energy of the non-interacting leptons with the formula

$$p/T^4 = \frac{1}{2\pi^2} \sum_i g_i \left(\frac{m_i}{T}\right)^2 \sum_{k=1}^{\infty} \frac{(-1)^k}{k^2} K_2 \left(\frac{m_i k}{T}\right), \quad (\text{S14})$$

where K_2 is a modified Bessel function of the second kind, and g_i is the spin degeneracy factor, $g_i = 4$ for leptons, and m_i is its mass. The right hand side of Eq. (S14) with $g_i = 12$ gives the free quark contribution $F_Q(m_i/T)$ already introduced in Eq. (S8).

c) Light hadrons

In Refs. [S18, S11] we have tested the Hadron Gas Model's prediction in detail. Here we use this model's prediction up to a temperature of 120 MeV. From that point we switch over to the lattice result.

d) QCD

We have continuum extrapolated lattice results for the equation of state up to a temperature of 1 GeV. We have found these as a combination of 2+1 and 2+1+1 flavor simulations. The lattice data can be smoothly continued with α_s^3 order perturbative result, where one analytically unknown parameter was fitted to our data. In Fig. S7 we showed that the sixth order (highest order with one fitted coefficient) perturbative result gives a good description of both the pressure and the trace anomaly. This pressure function is the basis of our result at high temperatures. We included the bottom threshold using the method described in Sec. S3.3. The full 2+1+1+1 flavor QCD contribution we show in Fig. S9.

e) W^\pm, Z^0 and the Higgs boson The bosonic version of Eq. (S14) can be used as a first estimate:

$$p/T^4 = \frac{1}{2\pi^2} \sum_i g_i \left(\frac{m_i}{T}\right)^2 \sum_{k=1}^{\infty} \frac{1}{k^2} K_2 \left(\frac{mk}{T}\right). \quad (\text{S15})$$

Ref. [S61] goes beyond this and adds the one-loop electroweak corrections. The one loop corrections become noticeable at the temperature of approximately $T \gtrsim 90$ GeV. For this correction we use the data of Ref. [S61].

f) *The electroweak transition* For the electroweak epoch we quote the results of Ref. [S59]. They use perturbation theory, dimensional reduction [S55] and the results of 3D simulations to estimate the equation of state of the Standard Model near the electroweak transition. This is a continuation of the earlier work [S61]. Although at the time continuum extrapolated electroweak lattice input was not yet available [S32], the final continuum extrapolation shows a very mild lattice spacing dependence [S31].

Adding all components from a) to f) we arrive at our final result for $g_\rho(T)$ and $g_s(T)$ that we plot in the main text. Here we give a cubic spline parametrization for $g_\rho(T)$ and the ratio of $g_\rho(T)$ and $g_s(T)$, see Table S2.

$\log_{10} T/\text{MeV}$	$g_\rho(T)$	$g_\rho(T)/g_s(T)$
0.00	10.71	1.00228
0.50	10.74	1.00029
1.00	10.76	1.00048
1.25	11.09	1.00505
1.60	13.68	1.02159
2.00	17.61	1.02324
2.15	24.07	1.05423
2.20	29.84	1.07578
2.40	47.83	1.06118
2.50	53.04	1.04690
3.00	73.48	1.01778
4.00	83.10	1.00123
4.30	85.56	1.00389
4.60	91.97	1.00887
5.00	102.17	1.00750
5.45	104.98	1.00023

Table S2: Data set on the logarithmic scale that can be used with simple cubic spline interpolation to find a parametrization for the entire Standard Model. The spline's typical deviation from g_ρ is about 1%, and 0.3% for the ratio.

Ref. [S59] has constructed an equation of state of the universe. In their work only the electroweak theory was based on lattice simulations. Here we replace the earlier perturbative deliberations on the QCD epoch by fully controlled lattice QCD result, which we conveniently parametrize. In Fig. S10 we show this non-perturbative effect by comparing our result to the published data set in Ref. [S59].

Our final result for the quantities in Eq. (S12), and their ratios is shown in Fig. 1 of the main paper.

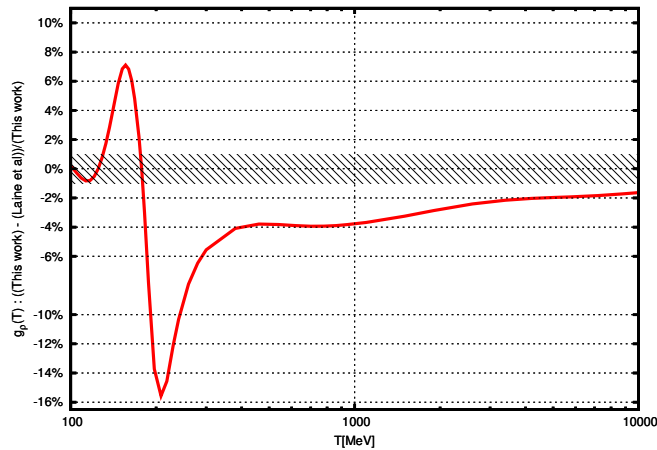


Figure S10: Our results and a previous estimate [S59]. We computed the relative difference of the two results. Before and after the QCD epoch there is agreement. The discrepancy in the range between 0.1 and 10 GeV is explained by the lack of lattice QCD input in [S59]. The dashed region around $\pm 1\%$ indicates the systematics of our parametrization.

S5 Overlap simulations

One of the results of this paper, namely the mass dependence of the topological susceptibility, is obtained using overlap fermions. In this Section we give a short summary of the numerical simulations with overlap quarks used in this work. We also describe our method to determine the Lines of Constant Physics with overlap fermions.

Our setup is based on the one used in Refs. [S20, S17]. For completeness we give a brief summary here:

- tree level Symanzik improved gauge action with gauge coupling parameter β .
- three flavors of overlap quarks. The sign function in the Dirac operator D_{ov} is computed using the Zolotarev approximation. The Dirac operator is constructed from a Wilson operator D_W with mass -1.3 . The quark fields are coupled to two step HEX smeared [S36] gauge fields.
- two flavors of Wilson fermions using the above D_W operator.
- two scalar fields with mass 0.54.

The extra Wilson-fermion fields are required to fix the topological charge [S47] and to avoid difficulties when topology changing is required [S44, S42]. These fields are irrelevant in the continuum limit. Note, that their action does not constrain the topology but suppresses the probability of the low lying modes of D_W . In a continuous update algorithm, no eigenmode of D_W can cross zero, which is equivalent to having a fixed topology. The role of the boson fields is to cancel the ultraviolet modes of the extra Wilson fermions. These boson fields are also irrelevant in the continuum limit.

S5.1 Odd flavor algorithm

The main difference to the works in [S20, S17] is, that here we use $n_f = 2 + 1$ flavors instead of $n_f = 2$. The simulations are done using the standard Hybrid Monte Carlo (HMC) algorithm. For the strange quark we use the chiral decomposition suggested first in Ref. [S12] and later in [S28]. The square of the Hermitian Dirac operator $H_{ov}^2 = (\gamma_5 D_{ov})^2$ can be decomposed as $H_{\pm}^2 = P_{\pm} H_{ov}^2 P_{\pm}$ with $P_{\pm} = (1 \pm \gamma_5)/2$. It is trivial to show that the determinant of the one flavor Dirac operator is $\det D_{ov} \sim \det H_{\pm}^2$ where the proportionality constants depend only on the topological charge. Since we do our runs at fixed topology, these constants can be factored out from the partition function and become irrelevant. Therefore a straightforward HMC with either of the H_{\pm}^2 operators corresponds to simulating a single flavor at fixed topology. The actual implementation is very simple: the pseudo-fermion generated at the beginning of each trajectory of the HMC has to be projected to one of the chirality sectors.

We tested the simulation code by comparing its results to a brute force update, in which the configurations are picked from a pure gauge heatbath and subsequently reweighted by the exact determinant. Since the determinant calculation is expensive, we ran the test on a small lattice, 4^4 . We found a perfect agreement between the two updating algorithms, as shown in Figure S11. We did tests by running the code with two copies of $n_f = 1$ fields and comparing the results obtained with our previous code for $n_f = 2$ and agreement was found in this case, too.

S5.2 Determination of the LCP

We now show how to determine the LCP for physical quark masses in the overlap formalism. It usually requires zero temperature simulations at the physical point, which is for overlap quarks prohibitively expensive with today's computer resources. Fortunately, these are not needed. Here we present an alternative, cost efficient strategy to determine the physical LCP with overlap quarks, based on the physical LCP, that is already known in a different, less expensive fermion formulation. In our case this will be an LCP with staggered fermions, which we have described in the previous Section, see Equation (S1).

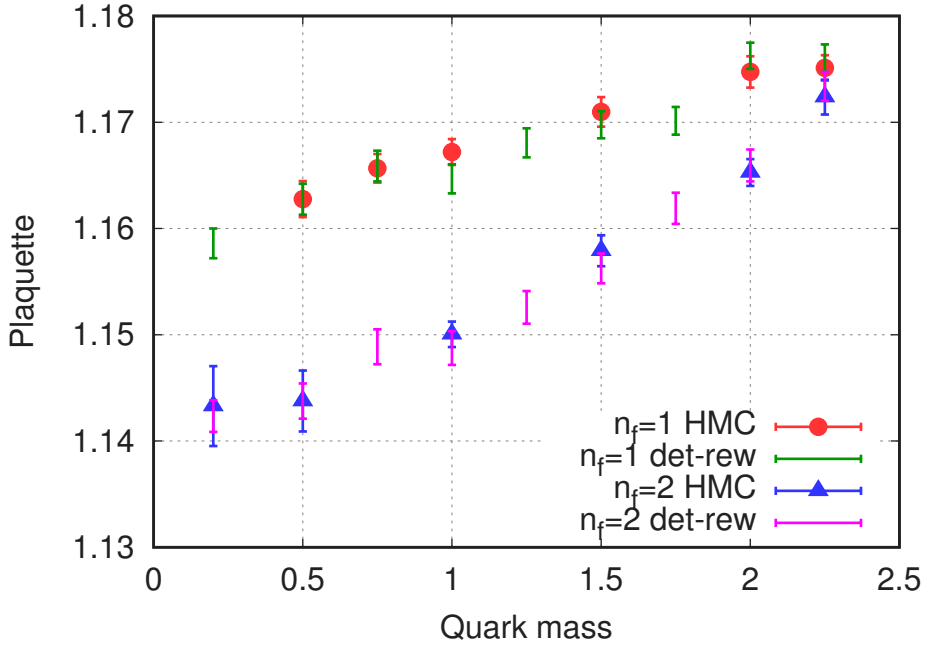


Figure S11: Comparison of the plaquette obtained by two different updating algorithms: the HMC algorithm and determinant reweighting. Shown are the average plaquette in the simulation as a function of the quark mass with both updating algorithms. The plot shows $n_f = 1$ and $n_f = 2$ flavor overlap simulations on a test lattice of size 4^4 .

The idea is to use the inexpensive three-flavor symmetric $n_f = 3$ theory as a bridge between the staggered and the overlap LCP's. In Subsection S1.3 we already determined the pion mass and the w_0 scale in the $n_f = 3$ theory in the continuum limit using staggered quark simulations¹. This can be used to construct an $n_f = 3$ overlap LCP, by tuning the quark mass for each gauge coupling so that $m_\pi w_0 \equiv m_\pi^{(3)} w_0^{(3)}$, and in this way we get the quark mass function $m_s^{ov}(\beta)$. This function is of course not the same as $m_s^{st}(\beta)$, as they are obtained with different fermion discretizations. The important point is, that both define the same physics, e.g. they both give the same $m_\pi w_0$ in the continuum limit. To close the definition of this three-flavor overlap LCP one has to measure $w_0^{ov}(\beta)$, the w_0 -scale as a function of the coupling. From this non-physical LCP one can get a physical $n_f = 2 + 1$ LCP with overlap fermions as follows

$$m_s = m_s^{ov}(\beta), \quad m_{ud} = R \cdot m_s^{ov}(\beta), \quad a = w_0^{(3)} / w_0^{ov}(\beta). \quad (\text{S16})$$

Here the value of the lattice spacing at the physical point was obtained by dividing the three-flavor continuum value of w_0 in physical units by the dimensionless three-flavor w_0^{ov} -scale measured in the overlap simulations.

For the $n_f = 3$ flavor overlap LCP we performed simulations at parameters listed in Table S3. From those we determined the quark mass by interpolating to the point, where $m_\pi w_0 = 0.552$. Then we used the formulas of Equation (S16) to obtain the lattice spacing and quark mass parameters for each β . The results are shown in Figure S12. Finally we fitted a three-parameter curve to these points to interpolate to β values, where no simulations were performed. These interpolations are also shown in Figure S12.

¹ Note, that the staggered theory also contained the charm quark, whereas the overlap simulations not. The continuum values of $m_\pi^{(3)}$ and $w_0^{(3)}$ are expected to be insensitive to the presence of the charm.

β	m	$N_t \times N_s$	ntraj
3.80	0.150,0.130	16×32	1000
3.90	0.120,0.100	16×32	1500
4.00	0.090,0.070	16×32	2000
4.05	0.070,0.055	16×32	1200
4.10	0.042,0.032	24×48	2200
4.20	0.042,0.032	24×48	2000
4.30	0.030,0.025	32×64	1600
4.40	0.020,0.030	32×64	1400

Table S3: Gauge coupling parameter, quark mass, lattice size and number of trajectories for the three flavor overlap simulations at zero temperature, that were used to determine the LCP.

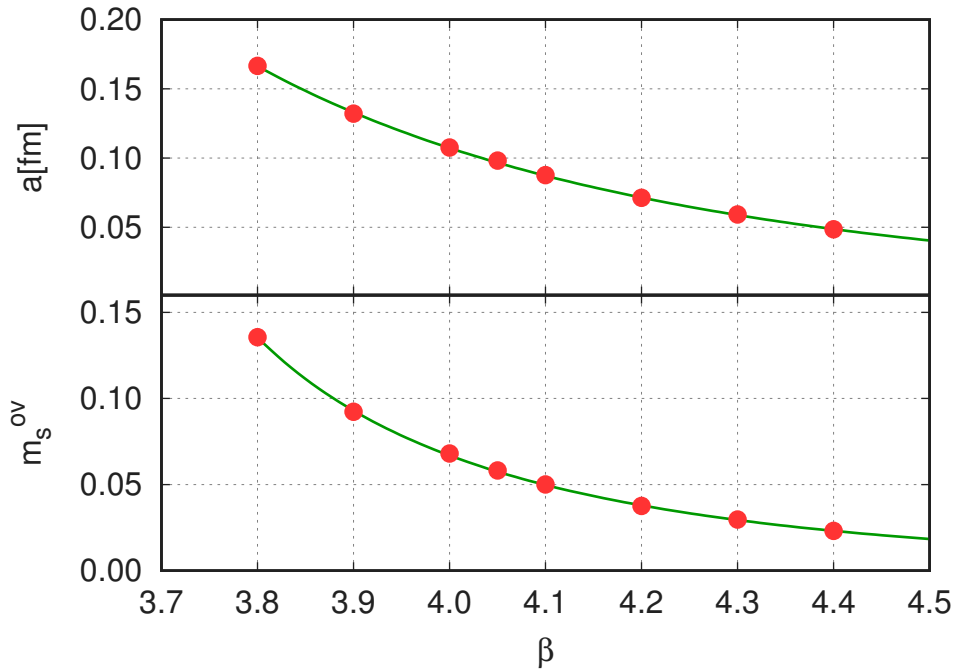


Figure S12: Lines of constant physics in the physical $n_f = 2 + 1$ flavor theory with overlap fermions. The upper plot shows the lattice spacing, the lower plot the strange mass parameter, both as the function of β . The light mass can be obtained as $m_{ud} = R \cdot m_s^{\text{ov}}$. The errors are smaller than the symbol size, the lines are smooth interpolations between the points.

S6 Eigenvalue reweighting technique

In the present Section we demonstrate how cut-off effects arise in the topological susceptibility with staggered quarks and present a method to efficiently suppress them.

The cut-off effects are strongly related to the zero-modes. To understand their importance, we first note that in the quark determinant for every zero-mode each dynamical flavor contributes a factor m_f , the corresponding quark mass. In this way gauge configurations with zero modes are strongly suppressed in the path integral, especially if the quark masses are small. Due to the index theorem, this also implies that light dynamical quarks strongly suppress higher topological sectors and thus the topological susceptibility.

On the lattice, however, there can be strong cut-off effects in this suppression. This is because the suppression factor is not $2m_f$ but $2m_f + i\lambda_0$, where λ_0 is the given would-be zero mode of the staggered Dirac operator, D_{st} ². The lack of exact zero modes can thus introduce strong cut-off effects and slow convergence to the continuum limit. Indeed, as long as the typical would-be zero eigenvalues are comparable to or larger than the lattice bare quark mass m_f , higher topological sectors are much less suppressed on the lattice than in the continuum.

To improve the situation, even at finite lattice spacing we can identify the would-be zero modes and restore their continuum weight in the path integral. In case of rooted staggered quarks this amounts to a reweighting of each configuration with a weight factor

$$w[U] = \prod_f \prod_{n=1}^{2|Q[U]|} \prod_{\sigma=\pm} \left(\frac{2m_f}{\sigma i\lambda_n[U] + 2m_f} \right)^{n_f/4} \quad (\text{S17})$$

where the second product runs over the would-be zero eigenvalues of the staggered Dirac operator with positive imaginary part. The third product takes into account the $i\lambda \rightarrow -i\lambda$ symmetry of the eigenvalue spectrum. The $n_f/4$ factor takes rooting into account, the factor 2 next to $|Q|$ together with the \pm symmetry make up for the fact that in the continuum limit the staggered zero modes become four-fold degenerate [S39].

Let us now turn to the most important part of the reweighting: the definition of the would-be zero modes. Since we are interested in the topological susceptibility, we identify the number of these modes with the magnitude of the topological charge $2|Q|$ as obtained from the gauge field after using the Wilson flow, see Section S1. We investigated two specific choices for the would-be zero modes. In the first approach we took the $2|Q|$ eigenmodes that have the largest magnitude of chirality among the eigenmodes with the appropriate sign of chirality, positive if $Q < 0$ and negative if $Q > 0$. In the second approach we took the $2|Q|$ eigenmodes with smallest magnitude. These two approaches are equivalent in the continuum limit, where zero-modes are exactly at zero and their chirality is unity. In practical simulations they give very similar results, we use the second approach in our analysis.

Since in the continuum limit the would-be zero eigenvalues get closer to zero, the reweighting factors tend to unity and in the continuum limit we recover the original Dirac operator. In this way, even at finite lattice spacings the proper suppression of higher topological charge sectors is restored and cut-off effects are strongly reduced resulting in much faster convergence in the continuum limit. For completeness let us note, that the above modification corresponds to a non-local modification of the path integral³. In the following we provide several pieces of numerical evidence for the correctness of the approach.

In Figure S13 we plot the distribution of the eigenvalues corresponding to the would-be zero modes at a temperature of $T = 240$ MeV for different lattice spacings. The distributions get narrower and their center moves towards zero as the lattice spacing is decreased. In Figure S14 we show the expectation value of the reweighting factors in the first few topological sectors. In the continuum limit $\langle w \rangle_Q = 1$ should be fulfilled in each sector. The results nicely converge to 1.

² Note, that our normalization of D_{st} is such, that in the free field continuum limit it approaches $2\cancel{\partial}$.

³ In this respect it stands on a footing similar to another method, which also modifies the quark determinant and which we also use in our staggered simulations: determinant rooting. As of today there is ample theoretical and numerical evidence for the correctness of the staggered rooting. See [S38] and its follow ups.

In most of our runs, especially at large temperatures and small quark masses, the weights were much smaller than 1. As a result there are orders of magnitude differences between the topological susceptibility with and without reweighting. It is therefore important to illustrate how the brute force approach breaks down if the lattice spacing is large and how the correct result is recovered for very small lattice spacings. In the following show two examples, Figures S15 and S16, where the standard method produces cut-off effects so large, that a reliable continuum extrapolation is not possible. In contrast the lattice spacing dependence of the reweighted results is much milder. To make sure that the reweighted results are in the a^2 -scaling regime, for both cases we present a non-standard approach to determine the topological susceptibility and compare their results to those of the reweighting approach.

In the first case, Figure S15 the temperature is just at the transition point, $T = 150$ MeV, where we expect to get a value close to the zero temperature susceptibility. This suggests that in this case the cut-off effects of the standard method can be largely eliminated by performing the continuum limit of the ratio $\chi(T, a)/\chi(T = 0, a)$, where the finite temperature result is divided by the zero temperature one at the same lattice spacing. We call this approach “ratio method”, see e.g. [S13]. As it can be seen in the Figure, this is indeed the case. The so obtained continuum extrapolation is nicely consistent with the value from the reweighting approach.

In the second case, Figure S16, we have a temperature well above the transition, $T = 300$ MeV. We see again, that the standard method produces results with large cut-off effects. The ratio method seems to perform better, however the apparent scaling is misleading. Although a nice continuum extrapolation can be done from lattice spacings $N_t = 8, 10$ and 12 , the $N_t = 16$ result is much below the extrapolation curve. The reweighting produces a result that is an order of magnitude smaller. In Sections S7, S8 and S9 we introduce a new method, called “integral method”, which is tailored for large temperatures. The so obtained result, where no reweighting is applied, agrees nicely with the reweighted one in the continuum limit.

These results provide numerical evidence for our expectations: the reweighting does not only produce a correct continuum limit, it also eliminates the large cut-off effects of staggered fermions.

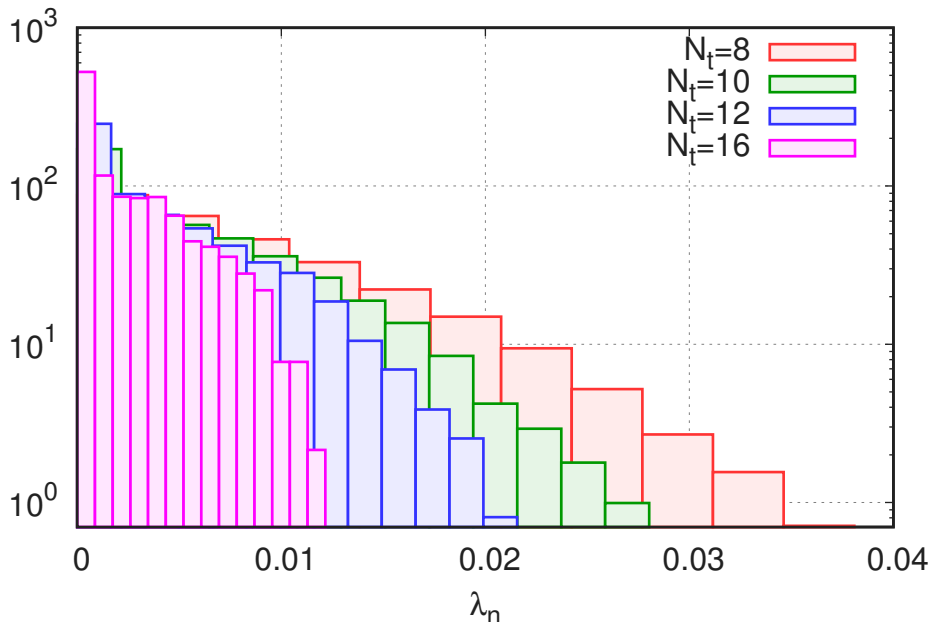


Figure S13: The probability distribution of the eigenvalues corresponding to the would-be zero modes obtained using the chirality method described in the text. The different colors refer to different lattice spacings. The plot shows $n_f = 2 + 1 + 1$ flavor staggered simulations at a temperature of $T = 240$ MeV.

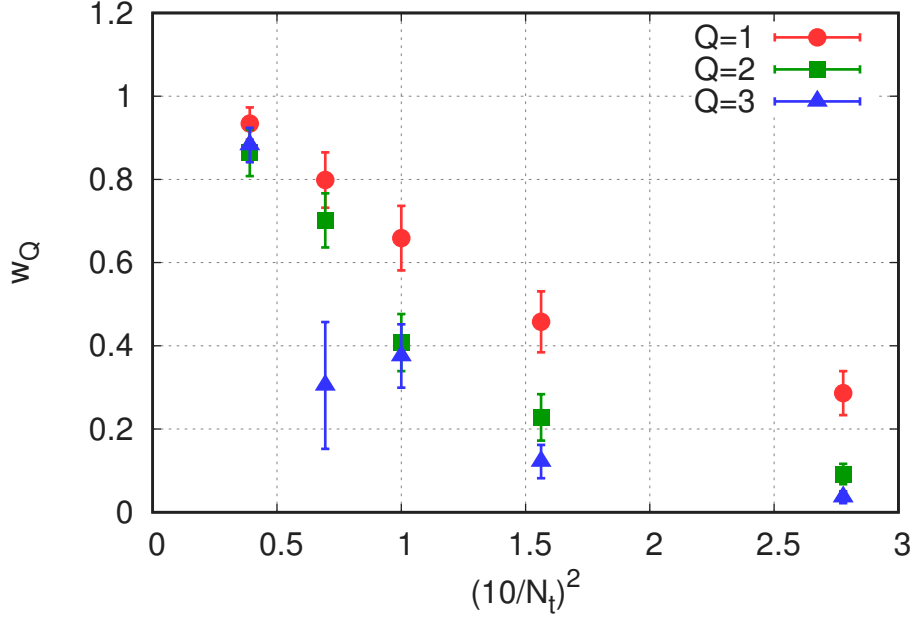


Figure S14: Expectation value of the weight factors in different topological sectors, $\langle w \rangle_Q$, as the function of the lattice spacing squared. The plot shows $n_f = 3 + 1$ flavor staggered simulations at a temperature of $T = 300$ MeV.

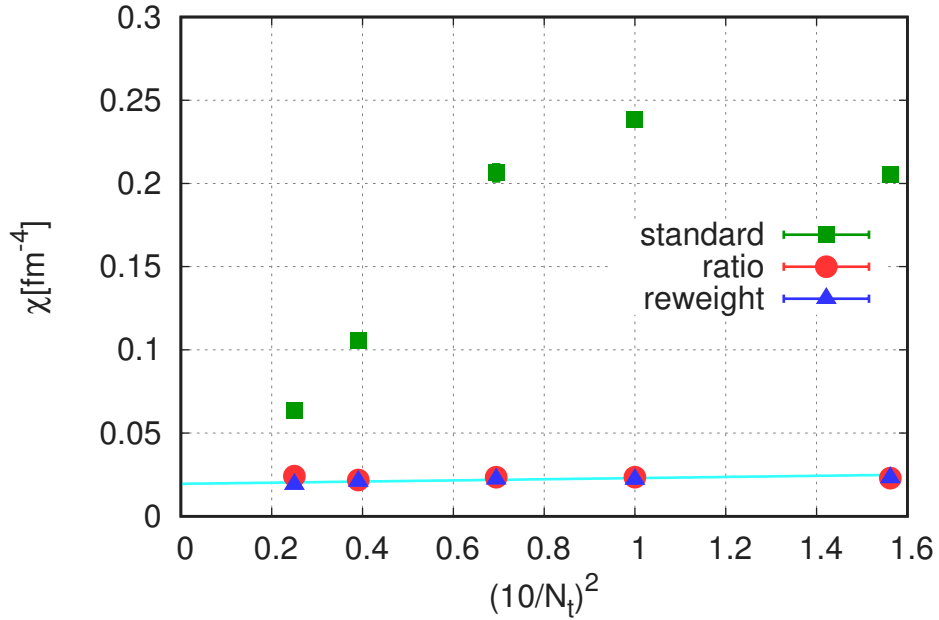


Figure S15: Lattice spacing dependence of the topological susceptibility obtained from three different methods described in the text: standard, ratio and reweighting. For the last method a continuum extrapolation is also shown. At this relatively small temperature the standard (“brute force”) method still cannot provide three lattice spacings, which extrapolate to the proper continuum limit, though they correspond to very fine lattices with $N_t = 12, 16$ and 20 . The plot shows $n_f = 2 + 1 + 1$ flavor staggered simulations at a temperature of $T = 150$ MeV.

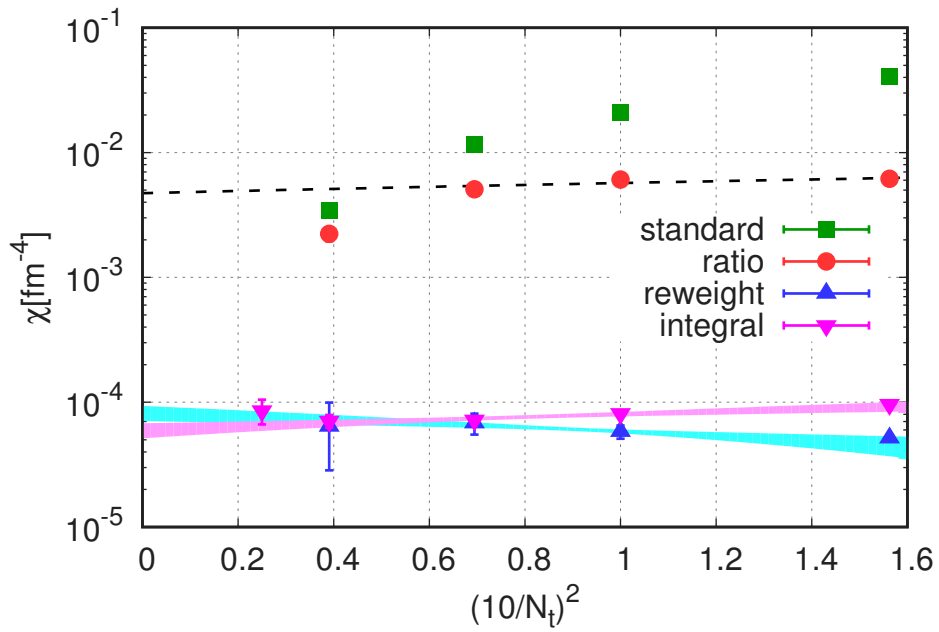


Figure S16: Lattice spacing dependence of the topological susceptibility obtained from four different methods described in the text: standard, ratio, reweighting and integral. For the ratio method a misleading continuum extrapolation using $N_t = 8, 10$ and 12 is shown with dashed line. For the reweighting and integral methods continuum extrapolations are shown with bands. The plot shows $n_f = 2 + 1 + 1$ flavor staggered simulations at a temperature of $T = 300$ MeV.

S7 Fixed sector integral in the pure gauge theory

To measure the topological susceptibility in the conventional way one needs configurations in sectors with non-zero topology. This gets increasingly difficult for two reasons: first with increasing the temperature the weight of the non-zero topology sectors decreases rapidly with the temperature. The necessary computer time increases with the inverse of the susceptibility. Secondly, as the lattice spacing is decreased, the change between topological sectors gets very improbable, see e.g. [S70], which increases the computer time demand even more.

In this section we present a novel way to measure the topological susceptibility, which is especially useful for large temperatures. To illustrate the new approach in this section we will work in the pure gauge theory. The gauge configurations are generated with a probability proportional to $\exp(-\beta S_g)$, where β is the gauge coupling parameter and S_g is the gauge action. In the next two sections we apply the method for staggered and overlap setups.

S7.1 A novel approach

There are many proposals to increase the tunneling between the topological sectors, see e.g. [S65, S62, S66], however, we do not pursue these here. Instead we forbid topological sector changes and determine the relative weight of the topological sectors by measuring the Q dependence of certain observables.⁴ More precisely we consider the following differentials:

$$b_Q \equiv \frac{d \log Z_Q/Z_0}{d \log T} = \frac{d\beta}{d \log a} \langle S_g \rangle_{Q-0}, \quad (\text{S18})$$

where Z_Q is the partition function of the system restricted to topological sector Q . In the continuum limit the sectors are unambiguously defined, however, on the lattice several different definitions are possible, for our particular choice on Z_Q see the next subsection. In Equation (S18) we also introduced the notation $\langle O \rangle_{Q-0} = \langle O \rangle_Q - \langle O \rangle_0$ for the difference of the expectation of an observable between the sectors Q and 0. Equation (S18) gives a renormalized quantity, the ultraviolet divergences cancel in the difference of the gauge actions. The important observation is, that the necessary statistics to reach a certain level of precision on b_Q 's does not depend on the temperature.

To obtain the relative weights Z_Q/Z_0 , we just have to integrate Equation (S18) in the temperature. For that we start from a temperature T_0 , where the standard approach (or the eigenvalue reweighting for the case of full QCD) is still feasible and determine the relative weights of the sectors Z_Q/Z_0 from a direct measurement. Then by measuring the b_Q 's for higher temperatures, where the direct measurement would become prohibitively expensive, we can use the following integral to obtain the Z_Q/Z_0 's:

$$Z_Q/Z_0|_T = \exp \left(\int_{T_0}^T d \log T' b_Q(T') \right) Z_Q/Z_0|_{T_0}. \quad (\text{S19})$$

If the temperature is high enough, the contribution of $Q > 1$ sectors can be neglected and the susceptibility is given by $\chi = 2Z_1/(Z_0 N_s^3 N_t a^4)$. Then the rate of change of the susceptibility b is given by:

$$b \equiv \frac{d\chi}{d \log T} = b_1 - 4, \quad (\text{S20})$$

where the term -4 takes into account, that the physical volume also changes with the temperature. To derive the Stefan-Boltzmann limit of Equation (S20), we can use that for large temperatures $\beta = 33 \log a/(4\pi^2)$. The gauge action difference is given by the classical action of one instanton $\langle S_g \rangle_{1-0} = 4\pi^2/3$. Up to lattice artefacts we get $b = 7$ in the Stefan-Boltzmann limit.

⁴A few hours after the submission of the present paper to the arXiv a paper appeared by J. Frison et al. discussing essentially the same method [S46], though only in the quenched approximation using coarse lattices.

As we have already mentioned, the statistics can be kept constant with increasing the temperature to reach the same level of precision on b_1 . However, with increasing the spatial size N_s , the statistics has to be increased as N_s^3 , the computer time as N_s^6 . This can be understood as follows: the gauge action difference between sectors 1 and 0 will be approximately given by the action of one instanton, which remains constant with increasing volume. The gauge action S_g itself however increases with the volume and the cancellation in Equation (S18) gets more severe. This volume squared scaling problem can be mildened by putting more and more topological charge into the box with increasing box size. If the topological objects are localized, then for large volumes the action difference between sectors 1 and 0 can be obtained from the difference between sectors Q and 0:

$$\langle S_g \rangle_{1-0} = \langle S_g \rangle_{Q-0} / |Q|. \quad (\text{S21})$$

This results in Q -fold increase in the signal-to-noise ratio, which translates into a Q^2 -fold decrease in the necessary computing time. The total gain in computer time over the “brute force” approach will be discussed in the next subsection.

S7.2 Numerical tests

We have carried out several numerical simulations to test the new approach. We used the Wilson-plaquette action. For scale setting we took the r_0 -scale parameterized as in [S41]. To convert the temperatures into units of T_c we took $r_0 T_c = 0.75$. Configurations were generated by overrelaxation/heatbath steps. To implement the fixed topology setup we added a Metropolis step to the end of each update, that rejected configurations if the topological charge escaped from a predefined region (see later). Since the Metropolis step is a global update, one has to make sure, that the overrelaxation/heatbath steps separately satisfy detailed balance, see [S49].

The topological charge was defined using the standard clover definition after applying a Wilson-flow for a flow time of $(8T^2)^{-1}$. There is a certain degree of ambiguity in defining the topological sectors, this ambiguity disappears in the continuum limit. For each N_t we first explored the density of states of Q , by running simulations, that were not allowed to go below some fixed value of Q . For small Q values we found sharp peaks, for an example see Figure S17, which was made on 8×16^3 lattices. To define the $Q = 0, 1, 2, \dots$ sectors, we constrain the Q such, that the zeroth, first, second, \dots peaks are in the middle of the allowed regions. We performed fixed topology simulations between these boundaries using the above Metropolis step. Usually it was enough to fix only the boundary, which is closer to 0, since the system did not attempt to cross the other boundary. For small Q 's we achieved an acceptance ratio, which was around 70% or better.

With increasing Q the peaks get broader and for large Q 's the distributions in the above defined sectors do not show a peak any more. It can also happen, that a simulation gets trapped on the sector boundary with a small acceptance ratio. We discarded such Q 's and simulations in our analysis. With approaching the continuum limit the peaks get sharper. Simulations that are trapped on a sector boundary, no longer occur. This can also be achieved by using gauge actions that suppress the topological tunneling, like the tree-level Symanzik, Iwasaki or DBW2 actions [S29].

An important issue in fixed topology simulations is the ergodicity. Usually we ran the simulations in 16 streams. The starting configurations were picked from a simulation at a low temperature, where topology decorrelated on a timescale of few updates. Therefore the streams can be regarded as independent. After sufficiently many updates the gauge action was consistent between the different streams. As an example, in Figure S18 we show the result of fixed topology runs on an 8×16^3 lattice. Plotted is b_Q from Equation (S18). The odd- Q sectors are not shown. The result was obtained from 20k updates per stream. The results obtained from different streams are all consistent with each other. The Q -dependence is consistent with a linear increase of the gauge action difference with Q , see Equation (S21). The lines represent the fit to all streams and charges assuming Equation (S21).

Let us now consider the finite spacing and size effects. The upper plot in Figure S19 shows $b_1 - 4$ as a function of the lattice spacing squared in a fixed physical volume, whereas the middle panel shows it as a

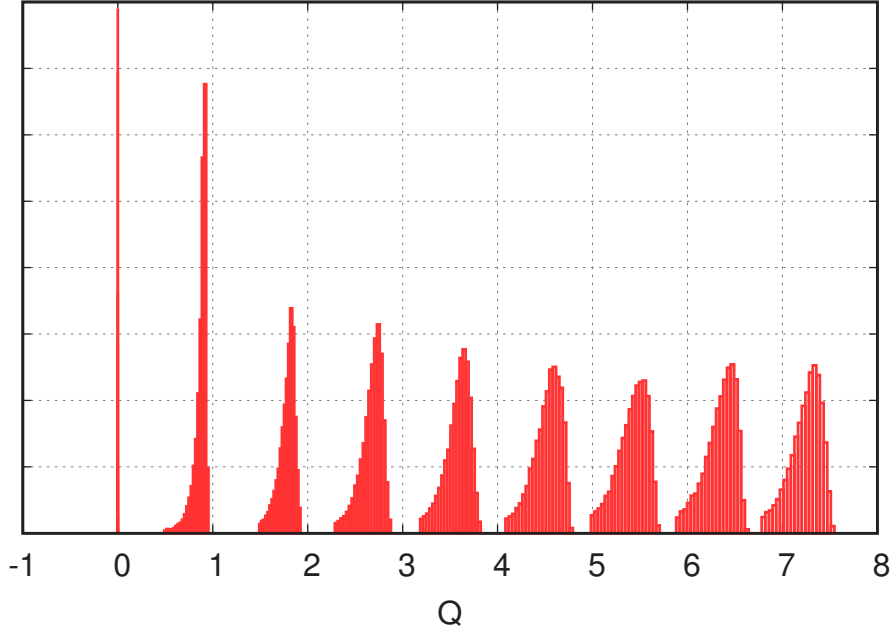


Figure S17: Histograms of the topological charge from fixed sector simulations. The relative weight of the sectors is determined by measuring the gauge action difference, see text. The plot shows pure gauge theory simulations on 8×16^3 lattices at $T = 5T_c$ temperature.

function of the aspect ratio N_s/N_t . Starting from aspect ratio ≈ 3 , we see no significant finite size effects. Note, that starting from aspect ratio 6, the boxes are large enough to accommodate non-perturbative length scales. We see no difference between boxes with perturbative and non-perturbative size. The runs of the finite spacing and volume scans are fitted jointly with a formula, that takes into account both effects linearly. The fits are shown with solid lines on the plots. For the exponent we obtain $b = 7.1(3)$ in the continuum and infinite volume limit at $T = 6T_c$. This is in with our previous estimate from the direct method [S14].

Figure S20 shows the temperature dependence of the decay exponent $b = b_1 - 4$ obtained from 8×32^3 simulations. Again we find agreement between the new data and the direct approach. At one temperature we did a simulation on an 8×64^3 lattice, where the exponent was obtained from measuring the difference between the $Q = 8$ and 0 sectors, $b = b_8/8 - 4$. We see no significant finite size effect. To get the Z_1/Z_0 ratio we performed a direct simulation at a temperature of $T_0 = 1.2T_c$. From this temperature we integrated up the b curve to obtain the Z_1/Z_0 as the function of temperature, up to $7T_c$, see Figure S20. The result can be compared to the lattice result obtained from the direct method [S14] and we find a good agreement both for the exponent and the susceptibility itself.

We also calculated the prediction of the dilute instanton gas approximation (DIGA). The necessary formulas can be found in eg. [S69]. To convert the result into units of T_c we used $T_c/\Lambda_{\overline{MS}} = 1.26$ from [S16]. Three different renormalization scales were used to test the scheme dependence: 1, 2 and $1/2$ times πT . For the exponent b we see a good agreement for temperatures above $\sim 4T_c$, for smaller temperatures the lattice tends to give smaller values than the DIGA. In case of the susceptibility the DIGA underestimates the lattice result by about an order of magnitude, this was already observed in [S14]. The ratio at $T = 2.41T_c$ is $K = 11.1(2.6)$, where the error is dominated by the error of the lattice calculation.

Figure S20 was made using 30 million 8×32^3 and 1 million 8×64^3 update sweeps. The cost of a simulation at $T = 7T_c$ using the standard method can be estimated from [S14]: it would require about 250 million updates on 8×64^3 lattices or about 2 billion on 8×32^4 , two orders of magnitude more, than with the novel method.

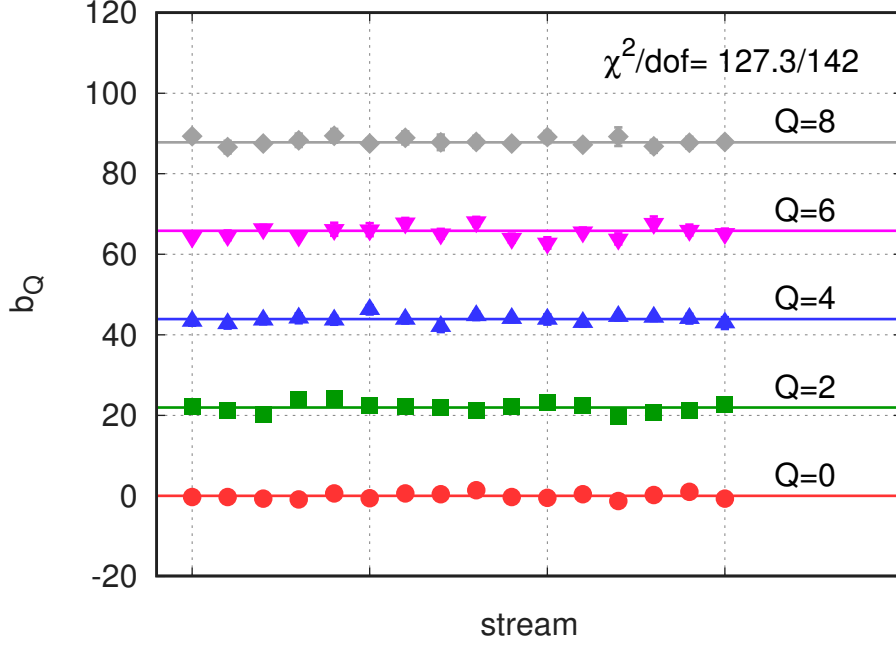


Figure S18: Gauge action difference as defined in Equation (S18). The different points correspond to independent simulations and different topological sectors. A good fit can be obtained assuming ergodicity and that the action difference scales linearly with the topological charge, see Equation (S21). The plot shows pure gauge theory simulations on 8×16 lattices at $T = 5T_c$ temperature.

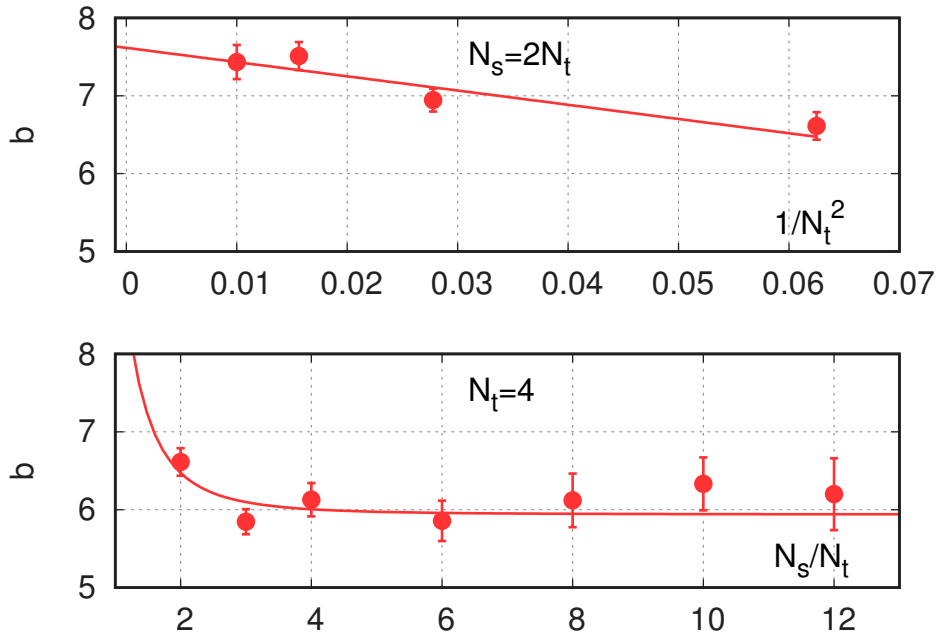


Figure S19: Lattice spacing (top) and finite volume (bottom) dependence of the decay exponent of the topological susceptibility b . The lines are fits taking into account leading order artefacts. The plot shows pure gauge theory simulations at $T = 6T_c$ temperature.

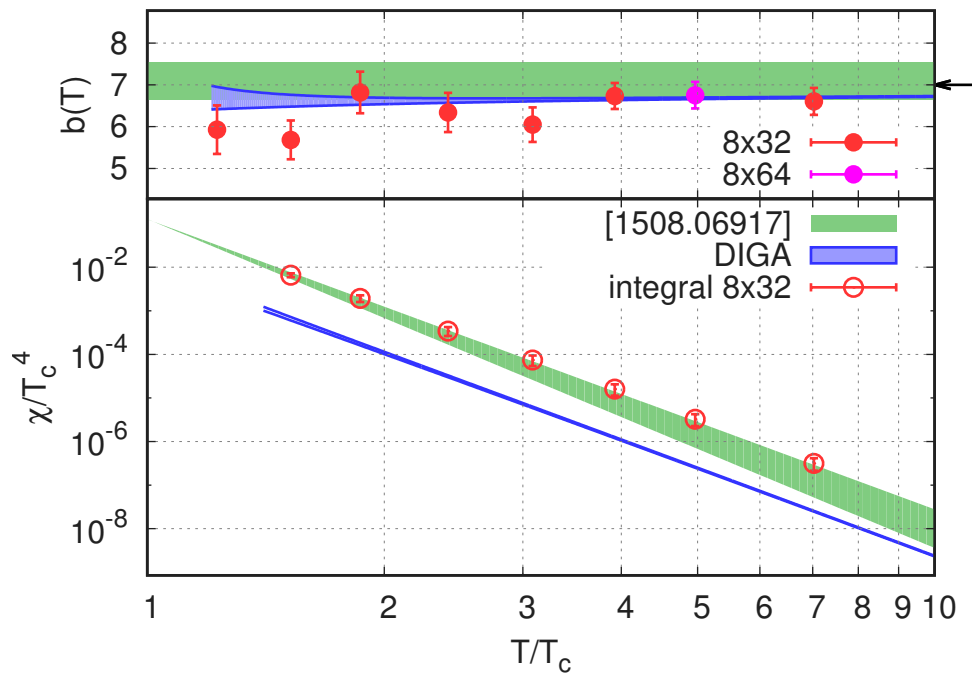


Figure S20: Topological susceptibility in the pure gauge theory. Results shown from an earlier direct simulation [S14], from the DIGA and from the novel fixed Q integral method. Upper plot is the decay exponent b , the lower the susceptibility itself. The arrow indicates the Stefan-Boltzmann limit.

S8 Fixed sector integral with staggered fermions

The method presented in Section S7 can be trivially generalized in the presence of fermions. The definition of b_Q is still given by Equation (S18). In the fixed- N_t approach changing the temperature is achieved via changing the lattice spacing, which requires a simultaneous change of β and the mass parameters m_f , to keep the system on the LCP, see Equation (S1). Then for b_Q we obtain:

$$b_Q \equiv \frac{d \log Z_Q / Z_0}{d \log T} = \frac{d\beta}{d \log a} \langle S_g \rangle_{Q=0} + \sum_f \frac{d \log m_f}{d \log a} m_f \langle \bar{\psi} \psi_f \rangle_{Q=0}. \quad (\text{S22})$$

Besides the gauge action S_g , we also have to measure the chiral condensate $\bar{\psi} \psi_f$ of each flavor. The full expression is a renormalized quantity, and so is the chiral condensate difference multiplied by the quark mass. To obtain the susceptibility we have to apply the same integral as in the pure gauge case, see Equation (S19).

Now let us look at the Stefan-Boltzmann limit of the decay exponent of the susceptibility. We can neglect the contribution of the $Q > 1$ sectors, so the decay exponent is $b = b_1 - 4$. The gauge action difference is the same as in the pure gauge case $4\pi^2/3$. The gauge parameter depends on the lattice spacing as $\beta = (33 - 2n_f) \log a / (4\pi^2)$ and the mass parameter as $\log m_f = \log a$ up to logarithmic corrections in a . The difference in the chiral condensate between sectors $Q = 1$ and 0 comes entirely from the presence of the zero mode, which gives a $\langle \bar{\psi} \psi_f \rangle_{1=0} = 1/m_f$. Altogether we have $b = (33 - 2n_f)/3 - 4 + n_f$ in the high temperature limit.

The statements of Section S7 about the computer time scaling with the volume and the possibility of using $Q > 1$ sectors also apply in the case of dynamical fermions. We used $Q = 0$ and 1 in this work, this is sufficient, since the topological susceptibility with dynamical fermions is tiny.

In numerical simulations the statistical noise on the gauge action difference is much larger than on the chiral condensate difference. This is very similar to, what was already observed in the context of the equation of state [S22]. This inspired us to use the following strategy: evaluate the b_Q and the susceptibility at a quark mass, where the simulation is less expensive than at the physical point. We choose a point, the so-called three-flavor symmetric point, where the two light-quark masses were set to the physical strange mass: $m_{ud} \equiv m_{s,\text{phys}}$. At this point we determined χ using the eigenvalue reweighting method (see later). Then we carried out an integration in the light-quark mass from $m_{s,\text{phys}}$ down to the physical light-quark mass $m_{ud,\text{phys}} = m_{s,\text{phys}}/R$. In this way we could avoid calculating the expensive gauge action difference at the physical point.

We observed that there are huge lattice artefacts on the chiral condensate contribution, if a non-chiral fermion discretization is used. In the absence of exact zero modes the chiral condensate difference needs very fine lattices to reach the continuum limit. The lattice spacing dependence of the three flavor chiral condensate difference is shown on the data labeled by "std" in the upper panel Figure S21. We used $3 + 1$ flavor staggered quarks in the simulation at a temperature of $T = 750$ MeV. There is an order of magnitude increase in the condensate by going from the coarsest to the finest lattice spacing. In the middle panel of Figure S21 the temperature dependence of the three flavor condensate is shown for different lattice spacings. As the temperature increases the condensate, which is calculated in the standard way, approaches zero contrary to the expectation in the high temperature limit. This also happens for the charm quark condensate, although at a somewhat smaller pace, see lower panel. The vanishing of the chiral condensate difference decreases the decay exponent of the susceptibility by n_f in the standard approach, which largely explains the unexpectedly small exponent obtained in the recent lattice calculation [S13].

We present two independent approaches to solve this problem. One is to use a chiral fermion discretization to evaluate the chiral condensate difference, this is explained in a separate section, Section S9. The other is the modification of the path integral by the reweighting technique, that we already introduced in Section S6. Let us present the details of the reweighting here. The introduction of the reweighting

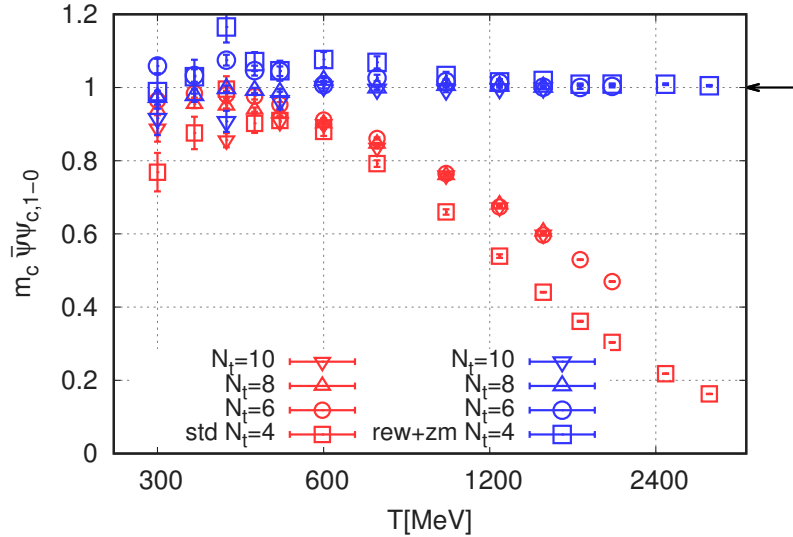
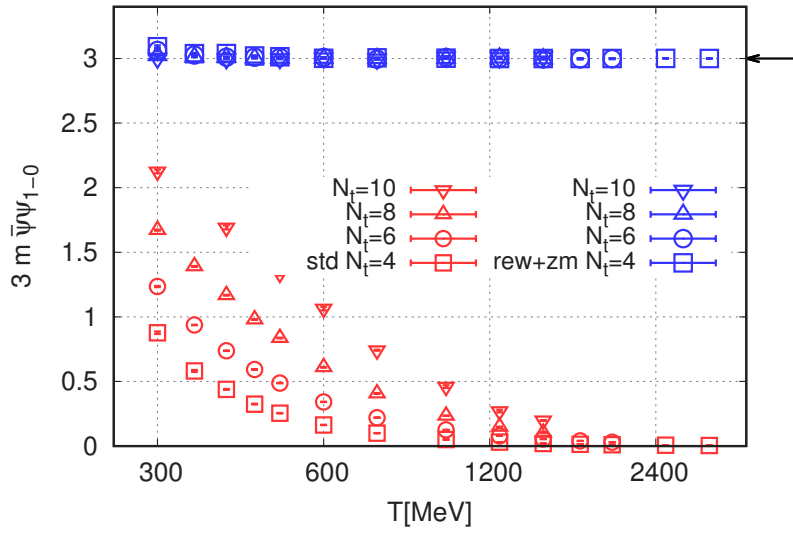
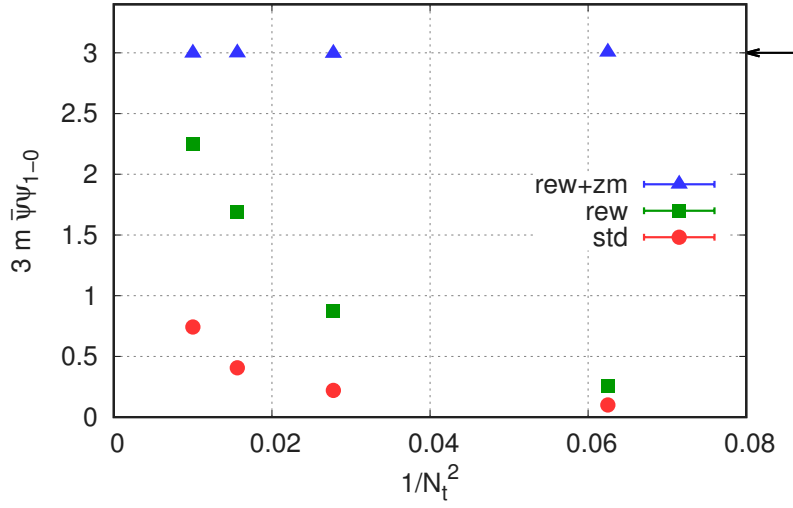


Figure S21: Chiral condensate difference between sectors $Q = 1$ and $Q = 0$ multiplied by the quark mass. The data labeled by “std” denotes the value calculated from the standard chiral condensate. The “rew” data is obtained by reweighting with the weights in Equation (S17). The “rew+zm” data includes the contribution of the zero modes, i.e. the mass dependence of the weight factors, see Equation (S25). The arrows indicate the Stefan-Boltzmann limit. The upper plot shows the difference as a function of the lattice spacing squared at $T = 750$ MeV temperature. The middle plot shows the difference as a function of the temperature, whereas the lower plot is the same for the charm quark. The plots show $n_f = 3 + 1$ flavor staggered simulations on $N_t = 4, 6, 8$ and 10 lattices.

factors $w[U]$ in Equation (S17) means, that our simulation corresponds to a modified partition function:

$$Z^{\text{rw}} = \int [dU] \exp(-\beta S_g) \cdot \prod_f \det(D_{st} + 2m_f)^{1/4} \cdot w[U]. \quad (\text{S23})$$

Note, that reweighting affects only the sectors with non-trivial topologies. This results in a modification of the expression for b_Q in Equation (S22) as:

$$b_Q^{\text{rw}} = \frac{d \log Z_Q^{\text{rw}} / Z_0^{\text{rw}}}{d \log T} = \frac{d\beta}{d \log a} \langle S_g \rangle_{Q=0}^{\text{rw}} + \sum_f \frac{d \log m_f}{d \log a} m_f \langle \bar{\psi} \psi_f \rangle_{Q=0}^{\text{rw}+\text{zm}}, \quad (\text{S24})$$

where $\langle \cdot \rangle_Q^{\text{rw}}$ denotes the fixed Q expectation value including the weights $w[U]$. The fermionic contribution has two parts, one coming from the reweighted chiral condensate and another from the mass dependence of the weight factors, which we call zero mode contribution:

$$\langle \bar{\psi} \psi_f \rangle_{Q=0}^{\text{rw}+\text{zm}} = \langle \bar{\psi} \psi_f \rangle_{Q=0}^{\text{rw}} + \frac{|Q|}{m_f} - \left\langle \frac{1}{2m_f} \sum_{n=1}^{2|Q|} \frac{4m_f^2}{\lambda_n^2[U] + 4m_f^2} \right\rangle_Q^{\text{rw}}. \quad (\text{S25})$$

In the end we have to measure three observables, the gauge action, the chiral condensate and the zero mode contribution, on the reweighted configurations.

The effects of reweighting and including zero mode contribution can also be seen in the panels of Figure S21. In the upper panel we see, that reweighting already improves the estimate of the fermionic contribution significantly and including the additional zero mode contribution reduces the lattice artefacts even further. In the middle panel of the figure we see that the reweighted condensate together with the zero mode contribution approaches the Stefan-Boltzmann limit for high temperatures, as expected. Indeed for the strange quark at our smallest temperature $T = 300$ MeV the Stefan-Boltzmann value is already reached, whereas the charm contribution is about 20% lower. The temperature dependence of the latter is plotted in the lower panel of Figure S21.

Now let's come to the lattice artefacts and finite size effects on the total decay exponent $b = b_1 - 4$, including not only the fermionic but also the gauge contribution, see Equation (S24). We did dedicated simulations at $T = 750$ MeV temperature: first we varied the lattice spacing in a fixed volume, secondly we changed the volume at a fixed lattice spacing. The upper panel of Figure S22 shows the lattice spacing dependence of b with and without reweighting. The lattice artefacts are larger without reweighting. The continuum extrapolations of the two data sets differ significantly. This difference is due to the problematic behaviour of the fermionic contribution without reweighting, as we explained before in detail. The lower panel of Figure S22 shows the finite size dependence of b . We did simulations with an aspect ratio up to and including 6. The box size corresponding to the largest aspect ratio is 1.6 fm, which is large enough to accommodate all non-perturbative length scales. Note, that the pion mass is $m_\pi^{(3)} \approx 710$ MeV in these $n_f = 3 + 1$ flavor simulations. We see no significant finite size effects even for an aspect ratio as small as 2. Although finite size and finite lattice spacing seem to be not significant at a first glance, these effects will be properly taken into account in our final analysis, where we perform a global fit over our data set with different N_t 's and aspect ratios, see Section S10.

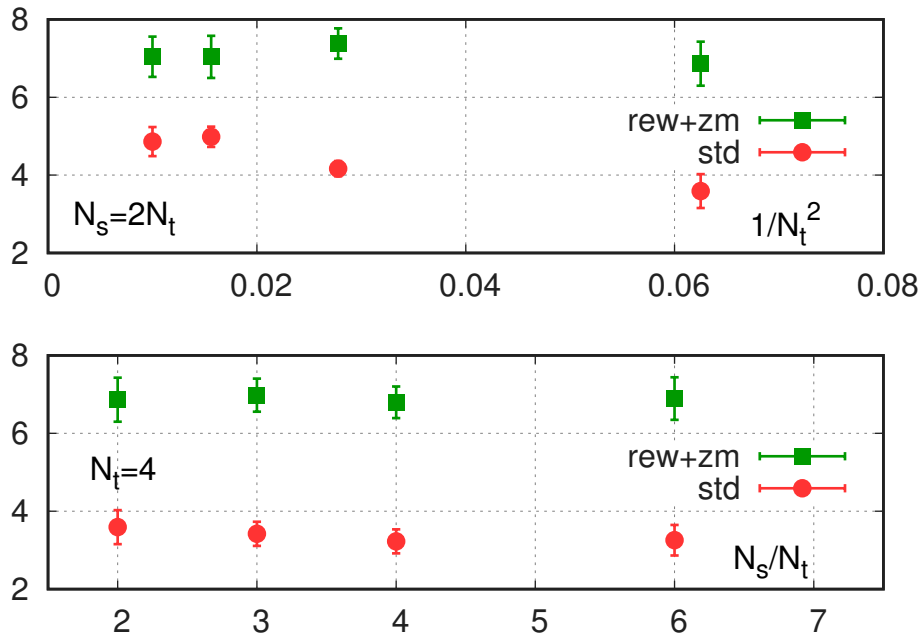


Figure S22: Lattice spacing (upper panel) and finite size (lower panel) dependence of the rate of change of the topological susceptibility b . The lattice spacing dependence is shown for $N_s = 2N_t$, while the finite volume dependence for $N_t = 4$. The labels “std” and “rew” indicate, whether the data set was obtained with the standard or with the reweighting method including the zero mode contribution. The plot shows $n_f = 3 + 1$ flavor staggered simulations at $T = 750$ MeV temperature.

S9 Fixed sector integral with overlap fermions

As demonstrated in Sections S6 and S8 staggered fermions produced huge lattice artefacts in the topological susceptibility and in the high temperature chiral condensate. An obvious remedy is to carry out simulations in a chirally symmetric discretization.

Determining the topological susceptibility with overlap fermions has a long history. The direct measurement is numerically difficult, since one has to deal with the non-analyticity of the overlap operator on the topological sector boundary. Though solutions exist [S44, S27, S42], they are somewhat cumbersome. Alternatively one can perform simulations in fixed topology [S47] and determine the topological susceptibility from the long distance behaviour of the topological charge correlator [S2]. This is a viable approach at zero temperature, but for high temperatures, where the susceptibility is small, one needs to measure the correlator with a very high precision.

Our new approach, presented in Section S8, also requires simulations with fixed topology. However we need to determine only the chiral condensate difference, the rest can be taken from direct simulations at parameters, where the direct approach is feasible. As we have seen the use of staggered fermions is complicated and difficult for this purpose. The difference, as we will show in this section, can be nicely measured in the overlap formulation. For algorithmic and other technical details we refer the reader to Section S5. Before showing results for the chiral condensate, we start with a previously unknown subtlety in fixed topology simulations with overlap fermions, which is related to configurations with a pair of an instanton and an anti-instanton.

S9.1 Instanton–anti-instanton (IA) configurations

The overlap topological charge is defined as the difference of the number of left and right handed zero modes of the overlap Dirac operator. A smooth instanton/anti-instanton produces a left/right handed zero mode in the spectrum. In practical simulations, one never encounters a configuration, where simultaneously left and right handed zero modes are present. This of course does not mean, that configurations with an IA pair are not allowed. If we look at smooth configurations, which contain a well-separated instanton and an anti-instanton, the overlap operator has a complex conjugate pair of overlap modes with very small but non-zero eigenvalues⁵.

For general configurations the definition of an IA pair or the number of IA pairs is of course not unambiguous. However for sufficiently high temperatures we observe very small modes, that are well separated from the rest of the non-zero modes, the latter being on the scale of the temperature. We looked at the topological charge distribution of such configurations and indeed observed the concentration of the charge into a positive and a negative lump, see Figure S23. Since such objects produce small complex conjugate pairs in the overlap operator spectrum, the value of the chiral condensate depends strongly on the presence of IA pairs. Therefore it is important to know, what is the fraction of configurations with IA pairs.

For an overlap fermion with a topology fixing term, configurations with a well-separated IA pair pose the following problem. Annihilating such a pair cannot proceed by simply removing the instanton and the anti-instanton one-by-one, since this would change the topological sector. Either they have to be removed simultaneously or they have to be brought to the same position, where they can annihilate. If the volume is large this latter can be difficult to achieve. In unfortunate cases we are stuck with an IA pair, and do not sample the probability distribution correctly. In our concrete numerical simulations we encountered this problem at only one parameter set: at the mass of the strange quark, $T = 300$ MeV and an aspect ratio of $N_s/N_t = 4$. For the history of the lowest eigenvalue in three different Monte-Carlo streams see Figure S24. “stream-1” contains no, “stream-3” one IA pair, in “stream-2” there was an IA annihilation after about 900 trajectories. For smaller masses and larger temperatures the runs always ended up without

⁵ One can prepare an artificial configuration with simultaneous left and right handed overlap zero modes, where an instanton is placed in the first half of the volume and the second half is obtained by CP transforming the first. Such configurations are expected to be a zero measure subset in the configuration space.

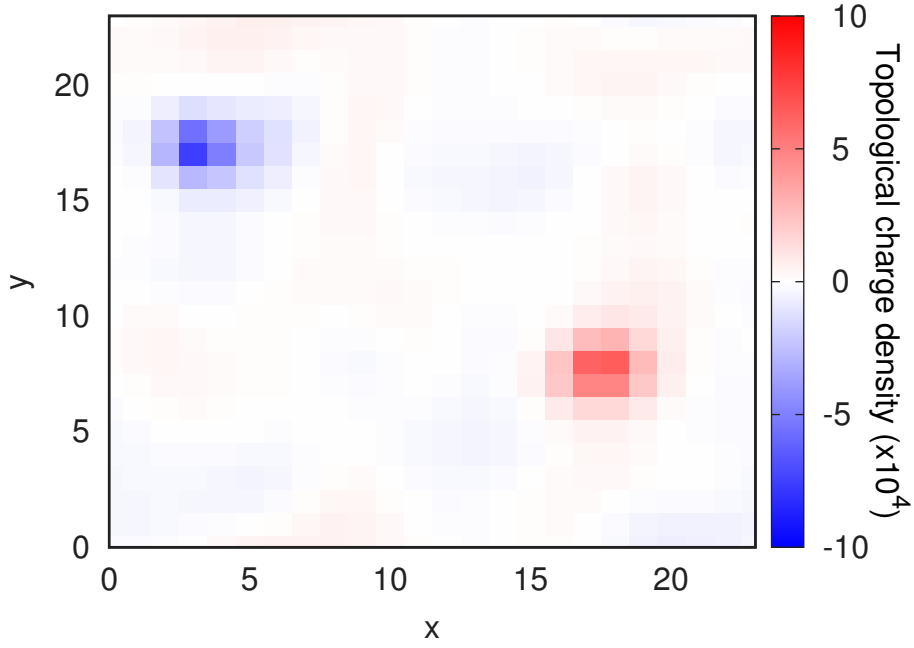


Figure S23: Topological charge density $q(x)$ distribution on an IA configuration. The $q(x)$ is averaged over the z and t coordinates and scaled by 10^4 . The plot shows a configuration from an $n_f = 3$ flavor overlap simulation on a 6×24^3 lattice at $T = 300$ MeV temperature.

having IA pairs after a short thermalization time.

It is interesting to look at the eigenvalues of the overlap kernel operator on IA configurations. In our case the kernel is a Wilson-Dirac operator with a negative mass: $D_W - m_W$. In the presence of an IA pair D_W will have two real modes lying between $0 \dots m_W$ and with opposite chiralities. To annihilate the IA pair, the two real modes have to be placed to a different region in the spectrum. For an illustration see the plot in Figure S25. The simplest way, i.e. to move them towards larger real values, is not possible. This is because, the effect of the topology fixing term is to forbid real modes to go through the point m_W . So the only way they can disappear is, to move into the complex plane. However due to the γ_5 -hermiticity of the D_W operator, complex eigenvalues have to come in complex conjugate pairs. So in order to leave the real axis, the two real eigenvalues first have to become degenerate, it is only then possible for them to go into the complex plane.

For the single problematic overlap run, mentioned above, we calculate the contribution of the IA configurations as follows. We measure the weight of such configurations with staggered fermions in the continuum limit and with the same physical parameters ($m = m_s$, $T = 300$ MeV and $LT = 4$). We use the direct approach without fixing the topology to generate configurations, which is still efficient at these parameters. To measure the number of topological objects we use a smeared overlap operator with kernel mass parameter $m_W = -1.3$. To define the number of IA pairs, N , we counted the number of complex conjugate pairs, for which the eigenvalue satisfied $|\lambda|^2 < 10^{-4}$. We checked that the results are not sensitive to small variations of the upper bound exponent. The results as a function of the lattice spacing can be seen in Figure S26. Although the probability of $N > 0$ configurations is non-negligible, somewhat less than 10% at this particular parameter set, their contribution drops out of the ratio Z_1/Z_0 and

$$\frac{Z_1}{Z_0} = \frac{Z_1}{Z_0} \Big|_{N=0} \quad (\text{S26})$$

holds to a very good accuracy in the continuum limit. We also find, that the weight of configurations depends only on the total number of topological objects, i.e. only on $Q + 2N$. Based on this, as we increase the temperature or decrease the mass, we expect that in a given topological sector the IA pairs will

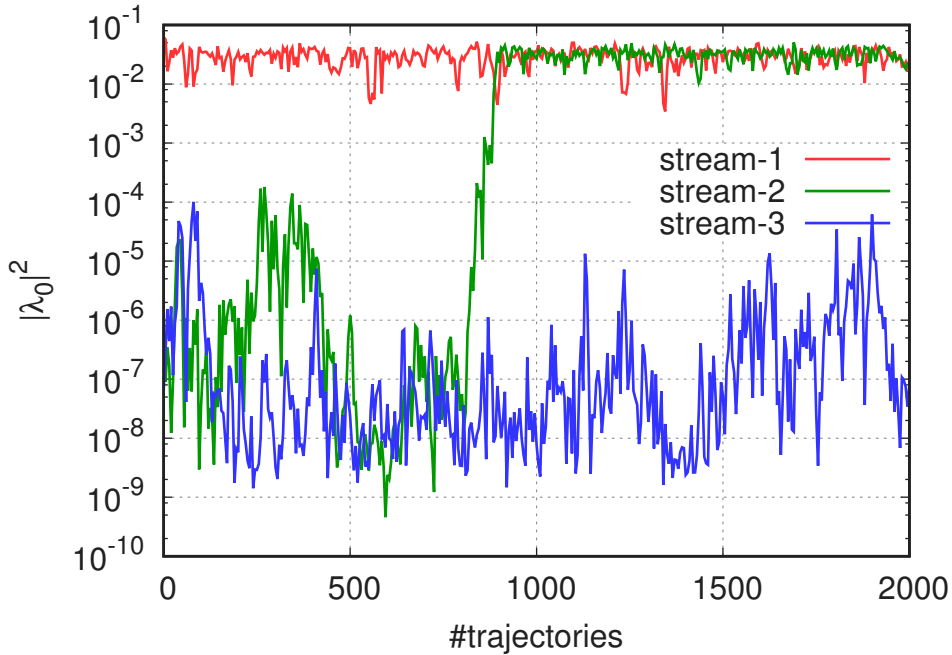


Figure S24: Lowest eigenmode squared $|\lambda|^2$ of the overlap Dirac operator in three different Monte-Carlo streams in the trivial topological sector. The plot shows $n_f = 3$ flavor overlap simulations on a 6×24^3 lattice at $T = 300$ MeV temperature.

have a decreasing contribution, like the contribution of non-trivial topological sectors decreases compared to the trivial sector. Since all our overlap runs use either a larger temperature or a smaller mass, than this particular simulation, we can safely neglect the contribution of IA configurations in all of our overlap simulations and use Equation (S26) to calculate sector weights.

S9.2 The chiral condensate difference

The chiral condensate can be decomposed into the eigenmodes of the Dirac-operator. The contribution of each topological mode is $1/m_f$. In the infinite temperature limit the rest of the spectrum does not contribute and therefore

$$m_f \langle \bar{\psi} \psi_f \rangle_{1-0} = 1. \quad (\text{S27})$$

For finite temperatures we expect corrections to this result. We also expect, that with decreasing quark mass the topological contribution will dominate and the corrections to Equation (S27) will get smaller. To investigate the size of these corrections we have carried out overlap simulations for a wide range of parameters. These are given in Table S4.

First we calculated the corrections to Equation (S27) for light-quark masses in the range $m_{ud}/m_s = 0.1 \dots 1$, we refer to it as “ m_{ud} -scan” in the table. We fixed the strange mass and used lattices of fixed size 6×12 and used three different temperatures. Then we looked at the lattice spacing dependence of the results, at $T = 300$ MeV at the three flavor point, we call it “ N_t -scan”. Finally at the same temperature and quark mass we investigated the finite size effects, these runs are called “ N_s -scan”. In all cases we found, that Equation (S27) holds with an accuracy of about one percent. The results are given in the last column of Table S4.

β	$N_s \times N_t$	m_{ud}	m_s	# ktraj	$\frac{1}{2}m_{ud}\langle\bar{\psi}\psi_{ud}\rangle_{1-0}$
m_{ud} -scan at $T = 300$ MeV					
3.99	12×6	0.0690	0.0690	10	1.00(1)
3.99	12×6	0.0460	0.0690	5	0.99(1)
3.99	12×6	0.0172	0.0690	8	1.00(1)
3.99	12×6	0.0069	0.0690	10	1.00(1)
m_{ud} -scan at $T = 450$ MeV					
4.19	12×6	0.0389	0.0389	10	1.00(1)
4.19	12×6	0.0291	0.0389	6	1.00(1)
4.19	12×6	0.0259	0.0389	3	1.00(1)
4.19	12×6	0.0195	0.0389	3	1.00(1)
4.19	12×6	0.0097	0.0389	3	1.00(1)
4.19	12×6	0.0049	0.0389	3	1.00(1)
m_{ud} -scan at $T = 650$ MeV					
4.38	12×6	0.0242	0.0242	5	1.00(1)
4.38	12×6	0.0181	0.0242	5	1.00(1)
4.38	12×6	0.0161	0.0242	3	1.00(1)
4.38	12×6	0.0121	0.0242	2	1.00(1)
4.38	12×6	0.0060	0.0242	2	1.00(1)
N_t -scan					
3.99	12×6	0.0690	0.0690	12	1.00(1)
4.13	16×8	0.0458	0.0458	29	1.02(2)
4.24	20×10	0.0342	0.0342	80	1.00(1)
N_s -scan					
3.99	12×6	0.0690	0.0690	12	1.00(1)
3.99	16×6	0.0690	0.0690	20	1.00(1)
3.99	20×6	0.0690	0.0690	32	1.02(1)
3.99	24×6	0.0690	0.0690	48	1.00(1)

Table S4: Gauge coupling parameter, lattice size, quark masses and number of thousand trajectories for the 2+1 flavor overlap simulations at finite temperature. Last column contains the chiral condensate difference.

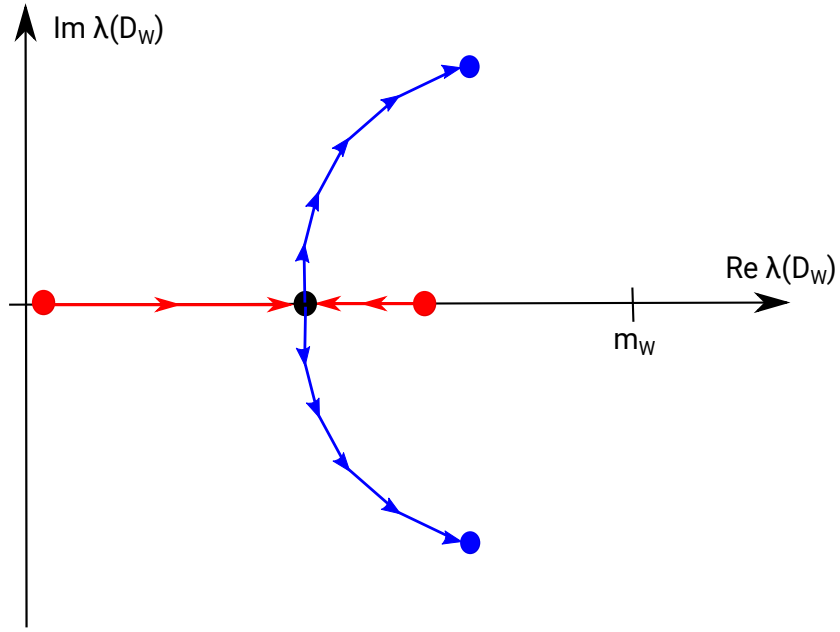


Figure S25: Illustration of an IA annihilation in the spectrum of the Wilson-Dirac operator, which is used in the kernel of the overlap Dirac operator.

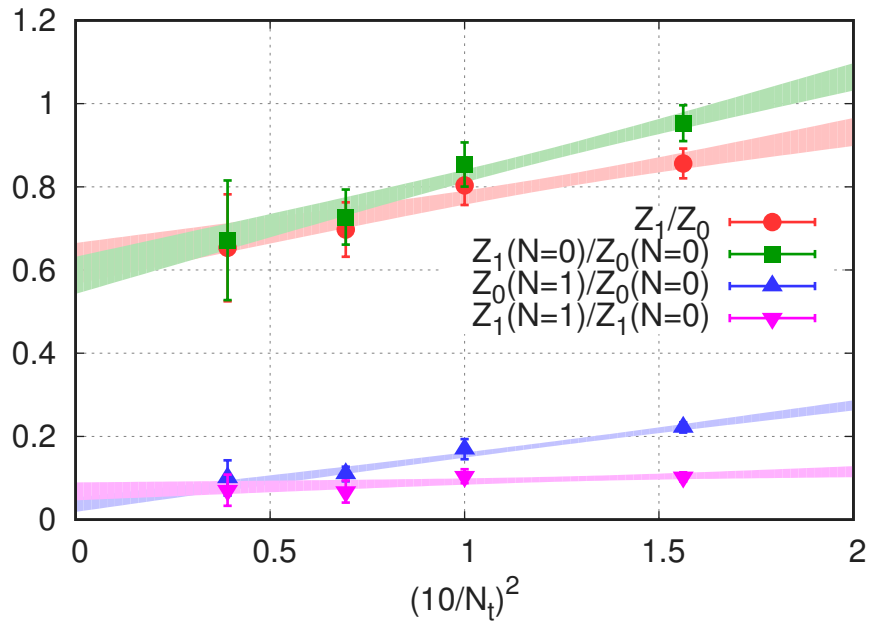


Figure S26: Continuum extrapolation of partition function ratios. The lower index in Z_Q stands for the topological charge, N for the number of IA pairs. The plot shows $n_f = 3 + 1$ flavor staggered simulations on lattices with $N_s/N_t = 4$ at $T = 300$ MeV temperature.

S10 Analysis for the topological susceptibility

We combine all approaches developed in the other sections to obtain our final result for the continuum extrapolated topological susceptibility at the physical point. Figure 2 in the main text shows this result including statistical and systematic error estimates.

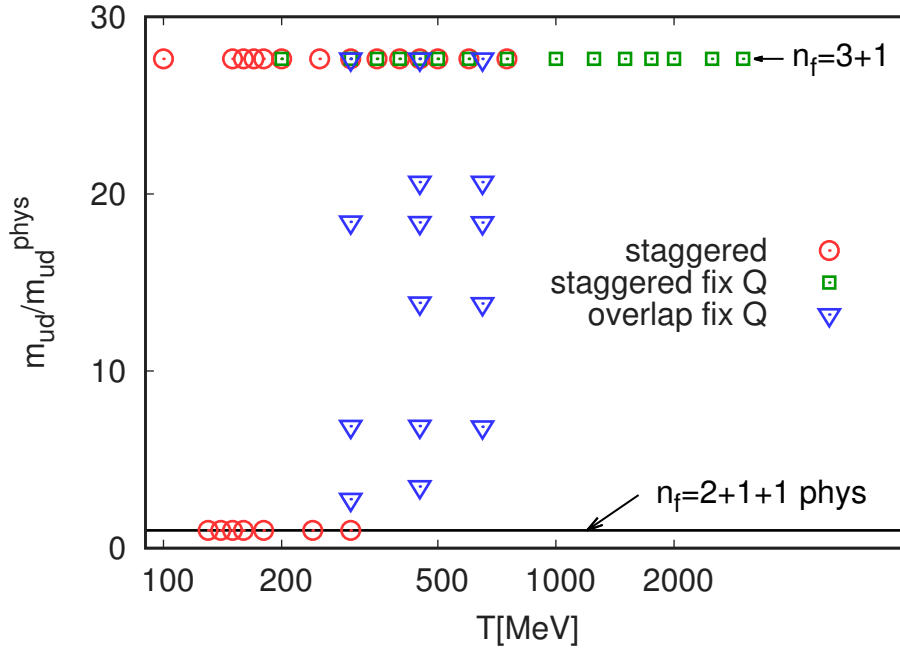


Figure S27: Summary plot of simulation points to determine χ . There are staggered simulations in the $n_f = 3 + 1$ and $n_f = 2 + 1 + 1$ flavor theories, which are then connected by overlap simulations.

In Figure S27 we show the simulation points that were used in the analysis. The plot shows the temperature – light-quark mass (m_{ud}) plane. Four simulation sets can be distinguished:

1. $n_f = 3 + 1$ flavor staggered simulations in the region $T = 100 \dots 750$ MeV
2. $n_f = 3 + 1$ flavor staggered simulations at fixed topology in the region $T = 200 \dots 3000$ MeV
3. $n_f = 2 + 1$ flavor overlap simulations at fixed topology for temperatures $T = 300, 450$ and 650 MeV building a bridge between the three flavor and the physical theories
4. $n_f = 2 + 1 + 1$ flavor staggered simulations at the physical point for temperatures $T = 130 \dots 300$ MeV

The main feature of our strategy is that the majority of the simulations are done in the three flavor symmetric theory (1. and 2.) instead of gathering statistics at the physical point. Starting from a temperature of $T \sim 300$ MeV the difference between the two can be taken into account by rescaling the topological sector weights, which is justified by the results of the overlap simulations (3.) connecting the two theories. The main observation is, as expected, that for large temperatures χ scales with the mass. Since it works on the one percent level already at 300 MeV, there is no need to go beyond 650 MeV for these bridging simulations. In the transition region the scaling behavior of the susceptibility with the quark mass is expected to change, so for temperatures $T \lesssim 300$ MeV we still resort to direct simulations at the physical point (4.).

S10.1 Topological susceptibility for $n_f = 3 + 1$ flavors

In the region between $T = 100$ and 750 MeV we have performed new simulations at six different lattice spacings, $N_t = 6, 8, 10, 12, 16$ and 20 for direct measurements of χ . The aspect ratio was set to $N_s/N_t =$

4. The simulation points together with the statistics are given in Table S5. On these configurations we calculated the low-lying eigenvalues of the Dirac operator.

$N_t \times N_s$	T[MeV]	ktraj
6×24	200	170
	250	260
	300	360
	350	380
	400	480
	450	490
	500	490
	600	490
	750	490
8×32	200	100
	300	100
	350	70
	400	70
	450	140
	500	100
	600	480
	750	460

$N_s \times N_t$	T[MeV]	ktraj
10×40	200	140
	250	80
	300	150
	350	400
	400	400
	450	860
	500	1300
	12×48	200
250		200
300		270
350		410
400		470
500		340
16×64	200	10
	250	30
	300	180
20×80	200	10

Table S5: $n_f = 3+1$ flavor staggered simulations points for direct measurements of χ . Lattice geometry, temperature and number of trajectories are given (in thousands).

We employ the reweighting method, as described in Section S6, to decrease the lattice artefacts. The topological charge was measured using Wilson-flow. The systematic errors were determined from varying the definition of the charge (plain or rounded to the nearest integer), from the choice of points considered in the continuum extrapolation (including a change of the fit range), and from the parametrization of the lattice artefacts. This gave us in total 20-64 fits per temperature. We used them in our histogram method to determine the systematic uncertainties [S33, S15]. In Figure S28 we show the continuum extrapolated results both with and without reweighting. The large errors without reweighting are coming from the continuum extrapolation.

To reach the temperature region that is needed for axion phenomenology we employ the fixed sector integral method developed in Sections S7 and S8. In this way no extrapolation in the temperature is needed. For this we generated configurations in topological sectors $Q = 0$ and 1 , at four different lattice spacings $N_t = 4, 6, 8$ and 10 and aspect ratios $N_s/N_t = 2 \dots 6$. The simulation points are given in Table S6. The number of trajectories are given for all temperatures together.

$N_t \times N_s$	T[MeV]	Mtraj
4×8	300, 350, 400, 450, 500, 600, 750, 1000, 1250, 1500, 1750, 2000, 2500, 3000	1
4×12	750, 1500	2
4×16	750, 1500	3
4×24	750, 1500	9
6×12	300, 350, 400, 450, 500, 600, 750, 1000, 1250, 1500, 1750, 2000	30
8×16	300, 350, 400, 450, 500, 600, 750, 1000, 1250, 1500	56
10×20	300, 400, 500, 600, 750, 1000, 1250, 1500	20

Table S6: $n_f = 3 + 1$ flavor staggered simulation points with fixed topology, $Q = 0, 1$ for measuring χ with the fixed sector integral method. Lattice geometry, temperature values and total number of trajectories are given (in millions).

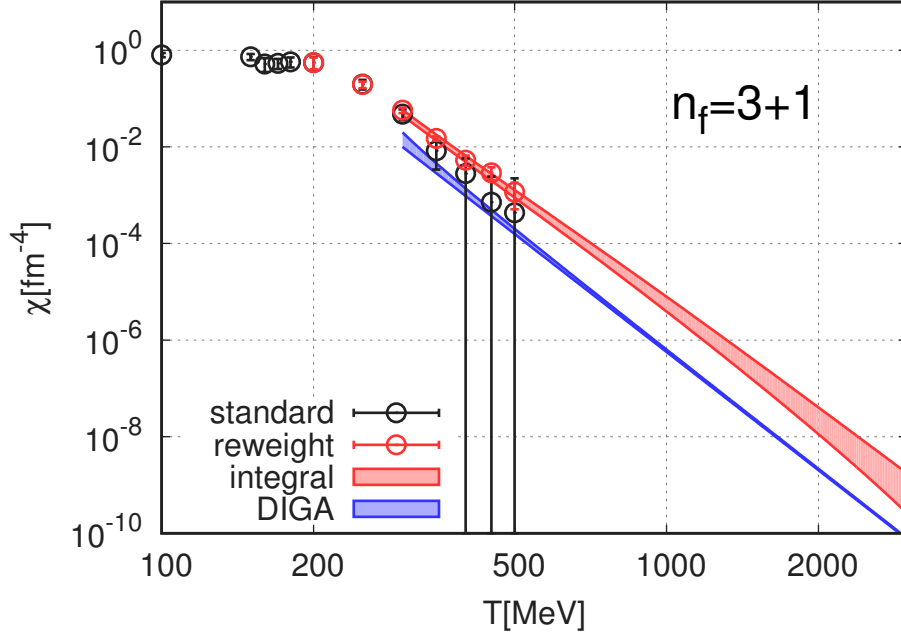


Figure S28: Continuum extrapolated topological susceptibility as a function of the temperature in the three flavor symmetric theory. The result was obtained from $n_f = 3 + 1$ flavor staggered simulations after performing the continuum limit. Lattice results of two direct methods (standard and reweighted) and of the fixed sector integral are shown. Also shown is the prediction of the DIGA.

We first determined the continuum extrapolation of the exponent $b = d \log \chi / d \log T$. For this we use a combined fit to all of the data which includes also infinite volume extrapolation. We use 16 ansätze which are combinations of four choices of the temperature dependence (polynomials in $1/T$ up to third order), two choices for the parametrization of the lattice artefacts (none or linear in $1/N_t^2$) and two for the parametrization of finite volume effects (none or linear in $(N_t/N_s)^3$). The different fits are combined using the Akaike Information Criterion (AIC) and also yield a systematic error estimate. For a more detailed exposition of this procedure, see [S15]. The result is shown in Figure S29. We also calculated the prediction of the DIGA for $n_f = 3 + 1$ flavors. For this we took the strange mass from [S34] and $m_c/m_s = C$ from Equation (S1), since this ratio was used in the simulations. We took $\Lambda_{\overline{\text{MS}}}^{(4)}$ from [S68] to convert the perturbative results to physical units. The renormalization scale dependence was estimated by using three different scales: 1 , $1/\sqrt{2}$ and $\sqrt{2}$ times πT . The continuum extrapolated lattice results agrees with the DIGA for temperatures above $T \sim 1000$ MeV, whereas for smaller T 's the lattice gives somewhat smaller exponents. This behaviour is qualitatively similar to the quenched case.

To obtain χ an integration in the temperature has to be performed, see Equation (S19). We start with a continuum extrapolated value from the low temperature region and then use the continuum extrapolation of the data for the exponent b to perform the integration. The systematic error is derived from the systematic error at the starting point of the integration and from three choices of the starting point (300, 350 and 400 MeV). The result is then plotted together with the direct simulations at lower temperature in Figure S28 and also with the prediction of the DIGA. Similarly to the quenched case, Figure S20, the lattice result is considerably larger than the DIGA prediction. At $T = 500$ MeV the ratio of the two is $K = 6.1(1.1)(0.7)$, where the first error comes from the lattice, the second from the scheme dependence of the DIGA.

S10.2 Topological susceptibility for $n_f = 2 + 1 + 1$ flavors

For the zero temperature susceptibility we applied a method based on leading order chiral perturbation theory to remove the lattice artefacts. This is described in Section S1.

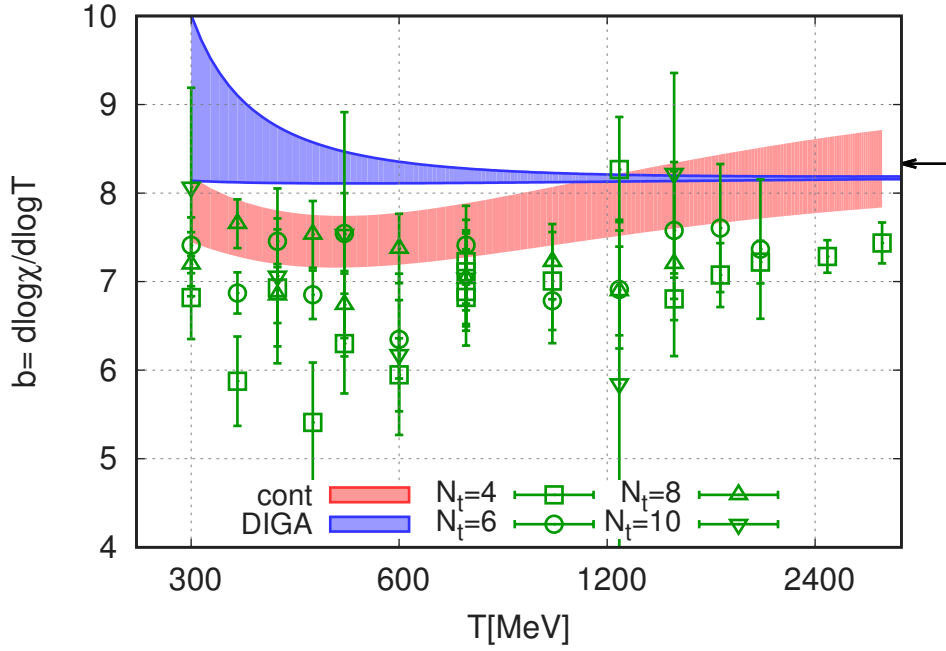


Figure S29: Temperature dependence of the exponent $b = d \log \chi / d \log T$. The plot shows $n_f = 3 + 1$ flavor staggered simulations on $N_t = 4, 6, 8$ and 10 lattices. The red band is the continuum extrapolation. The prediction of the $n_f = 3 + 1$ flavor DIGA is given by the blue band. The arrow shows the Stefan-Boltzmann limit.

At finite temperature we can start from the topological sector weights in the $n_f = 3 + 1$ theory. The result for $n_f = 2 + 1 + 1$ flavors is obtained by performing an integration in the light-quark mass. For example, the relative weight of sectors $Q = 0$ and 1 can be calculated as:

$$\frac{Z_1}{Z_0} \Big|_{2+1+1} = \exp \left(\int_{m_{ud}^{phys}}^{m_s^{phys}} d \log m_{ud} m_{ud} \langle \bar{\psi} \psi_{ud} \rangle \right) \cdot \frac{Z_1}{Z_0} \Big|_{3+1} \quad (\text{S28})$$

The overlap simulations in Section S9 provided ample evidence, that above $T = 300$ MeV, to a very good precisions the integrand is given by the number of light flavors $m_{ud} \langle \bar{\psi} \psi_{ud} \rangle = 2$. Thus the sector weights at the physical point are given by the following scaling:

$$\frac{Z_1}{Z_0} \Big|_{2+1+1} = R^2 \cdot \frac{Z_1}{Z_0} \Big|_{3+1}, \quad (\text{S29})$$

with R given in Equation (S1).

Equation (S28) can of course be used at any temperature, but the simple scaling with R^2 will not work e.g. at low temperatures or in the transition region. To be on the safe side for temperatures below $T = 300$ MeV we decided to fall back on the direct measurement at the $n_f = 2 + 1 + 1$ physical point, which turned out to be feasible even on the already existing statistics from [S11]. As was described in Section S6 reliable continuum extrapolations could only be performed after applying the reweighting procedure.

S10.3 Topological susceptibility - full result

There are two effects, that are missing in the $n_f = 2 + 1 + 1$ flavor theory, and have to be taken into account to obtain the full result for χ : the presence of the bottom quark and the mass difference between the up and the down quarks.

As we have seen in Section S8, the charm contribution to the decay exponent b has almost reached

$T[\text{MeV}]$	$-\log_{10}(\chi[\text{fm}^{-4}])$
100	-1.66(5)
120	-1.65(7)
140	-1.75(9)
170	-2.18(5)
200	-2.72(6)
240	-3.39(6)
290	-4.11(6)
350	-4.74(6)
420	-5.34(7)
500	-5.90(8)
600	-6.49(9)
720	-7.08(11)
860	-7.67(13)
1000	-8.17(15)
1200	-8.79(17)
1500	-9.56(20)
1800	-10.20(23)
2100	-10.75(26)
2500	-11.38(28)
3000	-12.05(33)

Table S7: Topological susceptibility of QCD taking into account the effect of the up, down, strange, charm and bottom quarks.

the high temperature limit at $T = 300$ MeV. We also found that the charm starts to contribute to the equation of state at $T \sim 250$ MeV (see Section S4). We therefore expect that the bottom contribution starts to be appreciable at temperatures above $\sim m_b/m_c \times 250$ MeV. To take into account the bottom contribution, we added $1/3$ to the $n_f = 2 + 1 + 1$ flavor exponent for temperatures higher than some threshold temperature. The value $1/3$ is the contribution of an extra flavor to the high temperature limit. Then we integrated the so obtained b and finally rescaled the results according to Equation (S29). We have chosen three different threshold temperatures: $T = 1.0, 1.5$ and 2.0 GeV. The resulting variation in χ was added to the systematic error.

Although isospin violating effects are typically on the level of 1%, the topological susceptibility is a notable exception. This is because, the susceptibility is proportional to the product of the quark masses. Therefore the topological susceptibility is a factor of

$$\frac{4m_u m_d}{(m_u + m_d)^2} \approx 0.88 \quad (\text{S30})$$

smaller than in the isospin symmetric, $n_f = 2 + 1 + 1$ flavor case. The quark mass values were taken from [S45]. To take isospin violation into account we scaled the isospin symmetric results by this factor for all temperatures.

Our final result for $\chi(T)$ is shown in Figure 2 of the main text. We also tabulate the base-10 logarithm $-\log_{10} \chi(T)$ for a couple of temperature values in Table S7.

S11 Axion dark matter from misalignment

The details of axion production via the misalignment mechanism are well described in the literature (see e.g. [S76]) but for completeness we briefly discuss our calculations.

In order to calculate the amount of axions produced we have to solve the equation of motion for the $A(x)$ axion field or equivalently for the $\theta(x) = A(x)/f_A$ axionic angle in an expanding universe:

$$\frac{d^2\theta}{dt^2} + 3H(T)\theta\frac{d\theta}{dt} + \frac{dV(\theta)}{d\theta} = 0, \quad (\text{S31})$$

where $V(\theta) = m_A^2(T)(1 - \cos\theta) = \chi(T)/f_A^2(1 - \cos\theta)$ is the temperature dependent axion potential. Since we focus on the misalignment mechanism, we assume that θ changes slowly in space on the relevant scales. Spatial fluctuations and defects lead to a string contribution which we do not discuss here.

The expansion is governed by the Friedmann equations:

$$H^2 = \frac{8\pi}{3M_{Pl}^2}\rho \quad (\text{S32})$$

$$\frac{d\rho}{dt} = -3H(\rho + p) = -3HsT \quad (\text{S33})$$

where ρ, p and s are the energy density, pressure and entropy density of the early universe and M_{Pl} is the Planck mass. At the temperatures where axion production happens, the contribution of axions to these densities can be neglected. ρ and s can be expressed as:

$$\rho = \frac{\pi^2}{30}g_\rho T^4 \quad s = \frac{2\pi^2}{45}g_s T^3 \quad (\text{S34})$$

using the effective number of degrees of freedom of Figure 1. Since we determined the energy density and entropy density for a wide temperature range, the solution of these equations yields the following relation between the age of the universe (t) and its temperature:

$$\frac{dt}{dT} = -M_{Pl}\sqrt{\frac{45}{64\pi^3}}\frac{1}{T^3g_s(T)\sqrt{g_\rho(T)}}\left(T\frac{dg_\rho(T)}{dT} + 4g_\rho(T)\right) \quad (\text{S35})$$

With the help of this expression, Equation (S31) can be rewritten in terms of temperature derivatives:

$$\frac{d^2\theta}{dT^2} + \left[3H(T)\frac{dt}{dT} - \frac{d^2t}{dT^2}/\frac{dt}{dT}\right]\frac{d\theta}{dT} + \frac{\chi(T)}{f_A^2}\left(\frac{dt}{dT}\right)^2\sin\theta = 0 \quad (\text{S36})$$

We solve this equation by numerical integration with some initial angle θ_0 and vanishing first derivative. When the temperature is large, the θ angle is frozen to its initial value. It starts to roll down the potential around T_{osc} which is defined as $3H(T_{\text{osc}}) = m_A(T_{\text{osc}})$. At the same time the axion number density (n_A) starts to increase. After a few oscillations its ratio to the entropy density (n_A/s) converges to a finite value which is then conserved for the rest of the evolution. Figure S30 shows T_{osc} for a large range of axion masses/couplings. Note that a coupling close to the Planck scale results in a T_{osc} below the QCD phase transition which emphasizes the need for the equation of state and $\chi(T)$ even for these low temperatures. Below the transition $T_{\text{osc}} \propto m_A^{0.47}$ while above the transition $T_{\text{osc}} \propto m_A^{0.17}$.

We start the numerical solution at $T = 5T_{\text{osc}}$. The oscillation starts around T_{osc} . We detect this by looking for the first sign change of θ which happens at T_s . We then extract n_A/s by averaging

$$\frac{n_A}{s}(T) = \frac{45}{2\pi^2}\frac{f_A^2}{m_A g_s T^3}\left[\frac{1}{2}\left(\frac{d\theta}{dT}/\frac{dt}{dT}\right)^2 + \frac{\chi(T)}{f_A^2}(1 - \cos\theta)\right] \quad (\text{S37})$$

for the temperature range $0.8 - 0.2T_s$. Throughout the solution of Equation (S36) we fix f_A (or equivalently m_A) and use our results for $\chi(T)$ and $\rho(T)$. The present axion energy density is obtained by using the conservation of n_A/s :

$$n_{A;\text{today}} = \frac{n_A(T)}{s(T)} s_{\text{today}} \quad \rho_{A;\text{today}} = m_A n_{A;\text{today}} \quad (\text{S38})$$

The current entropy of the universe is dominated by photons and neutrinos:

$$s_{\text{today}} = \frac{2\pi^2}{45} (2T_\gamma^3 + 6\frac{7}{8}T_\nu^3) = \frac{2\pi^2}{45} \frac{43}{11} T_\gamma^3 \quad (\text{S39})$$

where $T_\gamma = 2.725K$ is the cosmic microwave background temperature. This axion energy density has to be compared to the critical density or its dark matter component:

$$\Omega_{\text{axion}} = \frac{\rho_{a;\text{today}}}{\rho_{\text{crit}}} \quad R_A = \frac{\Omega_{\text{axion}}}{\Omega_{DM}} \quad (\text{S40})$$

In the pre-inflation scenario a single θ_0 and m_A (or f_A) determines R_A uniquely. Assuming $R_A = 1$ results in the curve in Figure 3 of the main text. In the post-inflation scenario all θ_0 angles are present with equal probabilities after the Peccei-Quinn transition and we have to average over them:

$$\overline{R_A}(m_A) = \frac{1}{2\pi} \int_{-\pi}^{\pi} R_A(\theta_0, m_A) d\theta_0 \quad (\text{S41})$$

For the whole m_A range which is relevant for the post-inflation scenario, we found that up to a few per mil the average angle of $\theta_0 = 2.155$ can be used:

$$\overline{R_A}(m_A) = R_A(\theta_0 = 2.155, m_A) \quad (\text{S42})$$

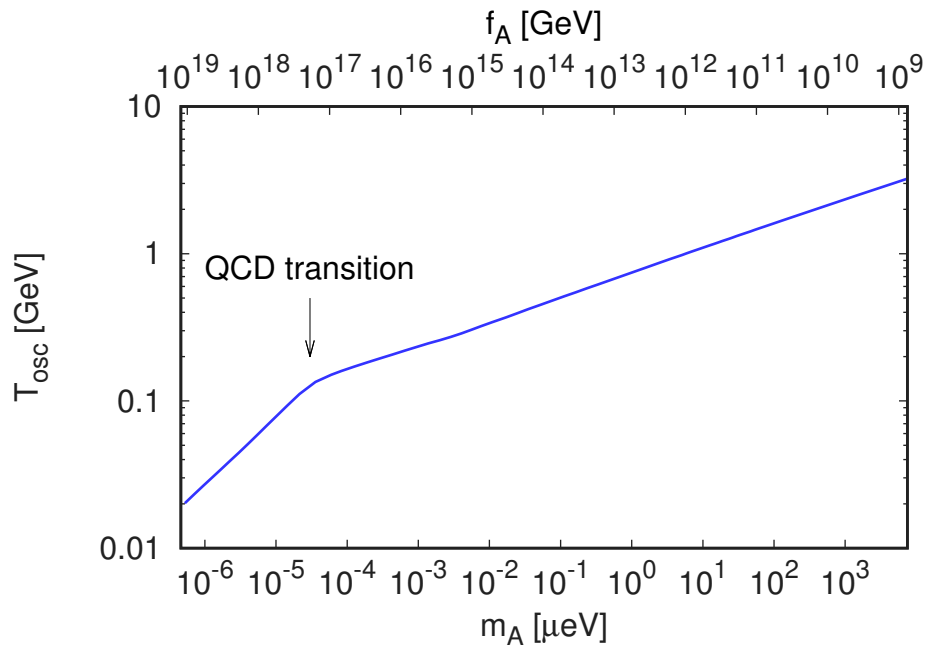


Figure S30: The oscillation temperature as a function of the axion mass. For this figure we assume that all the observed dark matter comes from the axions. Within this pre-inflation scenario the roll-down comes from a single θ_0 angle. The bend on the figure represents the QCD transition temperature. It reflects the very different behaviour of $\chi(T)$. Above the QCD transition χ rapidly drops, whereas below the QCD transition it has a much milder –almost constant– behaviour.

S12 Experimental searches for the axion in the predicted mass region

Using lattice QCD to determine the cosmological equation of state (EoS) and the temperature dependence of the topological susceptibility $\chi(T)$ the present paper showed that post-inflation dark matter axions have a mass between 50 and 1500 μeV (see Figure 3). In this section we show what should be done experimentally to approach and explore this mass region and detect dark matter axions. We conclude that, though presently operating and planned next generation experiments are not able to cover the predicted mass region, it is possible to design experiments, which offer discovery potential already in the near future.

Axions in the predicted 50-1500 μeV mass range are extremely challenging to detect. Recently, their theoretical appeal has been increasingly recognised and new techniques and experiments have been proposed. These include axion dark matter searches exploiting the excitation of atomic transitions in tuneable Rydberg atoms [S72] and electron spin precession [S7], but also purely laboratory searches for virtual axion long-range mediated forces [S4, S26]. Unfortunately, these are model dependent. The success of the former depends on a large axion-electron coupling and the latter upon the existence of new sources of CP violation beyond the Standard Model.

The most promising venue is to exploit the axion coupling to photons, $\mathcal{L}_{a\gamma} = -\theta\mathbf{E}\cdot\mathbf{B}\alpha C_{A\gamma}/(2\pi)$, with $C_{A\gamma}$ being an $\mathcal{O}(1)$ model-dependent constant. The simplest model compatible with the post-inflation scenario is the Kim-Shifman-Vainshtein-Zakharov (KSVZ) axion [S58, S71], which has $C_{A\gamma} = -1.92$ and we take it as our benchmark. The local axion dark matter field oscillations $\theta(t) \sim \theta_0 \cos(m_a t)$ in a homogeneous magnetic field \mathbf{B}_e generate an electric field $\mathbf{E}_\theta = \theta(t)\mathbf{B}_e C_{A\gamma}\alpha/(2\pi)$.

The haloscope experiment of Sikivie [S73] uses this field to drive the resonant mode of a microwave cavity when the oscillation frequency ($\nu_A = m_A/2\pi$) coincides with its resonant frequency. Several collaborations have already employed this technique. The local dark matter density $\rho_{\text{dm}} \simeq 0.3 \text{ GeV}/\text{cm}^3$ fixes the amplitude of the oscillations $\theta_0 = \sqrt{2\rho_{\text{dm}}/\chi(0)} \simeq 3.6 \times 10^{-19}$ and the electric field $|\mathbf{E}_\theta| = 1.2 \times 10^{-12} (|\mathbf{B}_e|/10 \text{ T}) \text{ V/m}$. Since the precise axion mass still remains unknown, the cavity has to be tuned to scan over the desired mass range. The bandwidth of the signal follows from the velocity dispersion of dark matter particles in the galactic halo $\Delta\nu \sim \nu_A/Q_A$ with $Q_A^{-1} = \langle (v/c)^2 \rangle / 2 \sim 10^{-6}$. In the mass range of interest, there are still $10^6 \log(1500/50) \sim 3.4 \times 10^6$ channels to be explored in the frequency range $\nu_a = 12\text{-}363 \text{ GHz}$. The power extracted from the cavity on resonance is given by

$$P_A = \kappa V m_A \max\{Q, Q_A\} \mathcal{G} |\mathbf{E}_\theta|^2 / 2, \quad (\text{S43})$$

where $\mathcal{G} = (\int dV \mathbf{E}_m \cdot \mathbf{B}_e)^2 / (|\mathbf{B}_e|^2 V \int dV |\mathbf{E}_m|^2)$ is the geometric overlap between the electric field of the cavity mode \mathbf{E}_m with the background B-field, Q the quality factor, V the volume of the cavity, and κ the coupling coefficient (ratio of the power extracted to the full cavity losses), optimally set to ~ 0.5 . The integration time required to find this signal with a given signal-to-noise-ratio S/N within the thermal and amplifier noise fluctuations is given by Dicke's radiometer equation

$$\Delta t = \Delta\nu \left(\frac{T_{\text{sys}} S}{P_A N} \right)^2 \quad (\text{S44})$$

where T_{sys} is the system noise.

As an example, ADMX is a state of the art and only fully commissioned experiment [S5]. It utilises a cylindrical cavity (1 m long, 0.5 m diameter, $Q \sim 10^5$) in an 8 T magnetic solenoid in a dilution refrigerator reaching 100 mK and SQUID amplifiers with noise close to the quantum limit. A measurement campaign of three years is being started and has the sensitivity to find dark matter axions in the pre-inflation scenario in the region labelled ADMX in Fig. S31. Generation 2 (G2) experiments to reach higher frequencies are currently under preparation by the ADMX HF-group and the Center of Axion and Precision Physics (CAPP) in South Korea. Our estimated sensitivities with $Q \sim 10^6$ and with cavities operated in fields of

up to 20 T may discover axions in the pre-inflation scenario up to $m_A = 30 \mu\text{eV}$, see G2 region in Fig. S31. The post-inflation scenario predicted in this paper may only be partially explored by presently envisaged Generation 3 (G3) experiments. Still, this would require magnetic fields as strong as 40 T and combining signals of several tuneable cavities. At this moment it is not clear at all if the required technologies will ever be available for such a search. Nevertheless we include this G3 region in Fig. S31.

An alternative method was proposed in a recent paper [S52]. A spherical mirror in a strong magnetic field was shown to emit electromagnetic waves of frequency ν_A that focus at the center of curvature in response to the oscillating axion dark matter field. The power per unit dish area is

$$\frac{P_A}{A} = \frac{|\mathbf{E}_\theta|^2}{2} = 2.2 \times 10^{-27} \frac{\text{W}}{\text{m}^2} \left(\frac{C_{A\gamma} |\mathbf{B}_e|}{10 \text{ T}} \right)^2, \quad (\text{S45})$$

too small for a wide-band search. However, it has been pointed out that the power can be enhanced by exploiting a dielectric planar mirror made of a sequence of N dielectrics [S54]. An equivalent power per area is emitted by each dielectric interface and can be added up coherently. This increases the power by a factor $4N^2$ (using a mirror at one end) which can be focused into a microwave receiver by a parabolic dish like the one used in [S75]. Detuning the dielectric thickness from $\lambda/2$, the dielectrics become partially reflecting, the power stored builds up like in a resonant cavity and the boost factor can be increased significantly so that a realistic axion search becomes feasible [S67].

Here, we envision a variable set of 20-40 sapphire dielectric slabs of $\sim 1 \text{ mm}$ thickness and 1 m^2 transverse area placed in a 10 T magnetic field with a planar mirror at one side. The distance between the dielectrics can be adjusted to have boost factors of order $\beta \sim 10^5$ in a relatively broad band $\mathcal{O}(50)$ MHz [S67]. Typically, one needs $d \sim \lambda/2$, which ranges between 3.1 cm and 2.4 mm in the axion mass range $40 - 250 \mu\text{eV}$. For $m_A = 250 \mu\text{eV}$, the coherence length of the axion field reaches $(m_A v)^{-1} \sim 1 \text{ m}$. Thus, coherent detection with such large axion-photon transducers can be severely hampered for larger masses. A 3 year measurement campaign with such an apparatus may scan the 50-100 μeV range with sensitivity to KSVZ axions with commercial low-noise high-electron-mobility transistor (HEMT) amplifier technology and up to 250 μeV with quantum limited detection. The reach of such a tuneable dielectric mirror is shown in Fig. S31 as a green-yellow band. The feasibility of an experiment of this type is currently being assessed at the Max-Planck-Institute for Physics in Munich.

This problem of coherence suggests us to reconsider the spherical dish antenna idea. Reaching sensitivity to KSVZ axions at 1 meV with a plain dish in a magnetic field of 10 T requires a sensitivity of 2.2 photons/($\text{m}^2 \text{ day}$). Again, the yield can be increased by a factor $4N^2$ when a few dielectric slabs are mounted (at adjustable relative distances) in front of each (planar) mirror element. With a boost of only $\mathcal{O}(100)$ and a total mirror area of 5 m^2 the photon rate would increase to 10^{-2} Hz with appears technically feasible for devices cooled down to temperatures of 10 mK and operating near the quantum limit [S7]. Still, it will be very challenging to shield the entire setup sufficiently against thermal noise.

In summary, we have shown that the region of high axion masses predicted by the QCD lattice calculations in this paper remains largely unexplored by presently operating and planned next generation experiments. This is mostly because of technical and practical limitations, particularly when searches over large ranges of axion masses are attempted. However, new experimental directions, some of which have been discussed in this paper and still being very challenging, may offer discovery potential already in the near future.

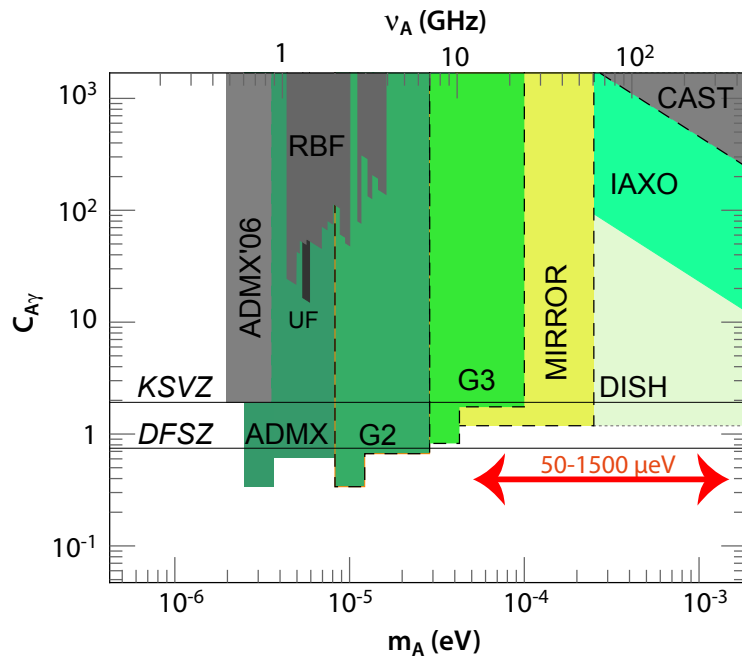


Figure S31: Sensitivity reach of the experiments discussed in the text as well as the International Axion Observatory (IAXO) [S3] (colored regions) with the current exclusion limits from previous cavity experiments: ADMX, RBF, UF and CAST (grey regions). We also show $C_{A\gamma}$ for the KSVZ and Dine-Fischler-Srednicki-Zhitnitsky (DSFZ) [S30, S79] models most interesting in the post-inflation and pre-inflation scenarios, respectively, and the range of m_A that could fit the dark matter abundance in the post-inflation scenario as follows from this work.

References

- [S1] Jens O. Andersen et al. “NNLO hard-thermal-loop thermodynamics for QCD”. In: *Phys. Lett.* B696 (2011), pp. 468–472. arXiv:1009.4644 [hep-ph].
- [S2] Sinya Aoki et al. “Finite volume QCD at fixed topological charge”. In: *Phys. Rev.* D76 (2007), p. 054508. arXiv:0707.0396 [hep-lat].
- [S3] E. Armengaud et al. “Conceptual Design of the International Axion Observatory (IAXO)”. In: *JINST* 9 (2014), T05002. arXiv:1401.3233 [physics.ins-det].
- [S4] Asimina Arvanitaki and Andrew A. Geraci. “Resonantly Detecting Axion-Mediated Forces with Nuclear Magnetic Resonance”. In: *Phys. Rev. Lett.* 113.16 (2014), p. 161801. arXiv:1403.1290 [hep-ph].
- [S5] S. J. Asztalos et al. “A SQUID-based microwave cavity search for dark-matter axions”. In: *Phys. Rev. Lett.* 104 (2010), p. 041301. arXiv:0910.5914 [astro-ph.CO].
- [S6] M. F. Atiyah and I. M. Singer. “The Index of elliptic operators. 1”. In: *Annals Math.* 87 (1968), pp. 484–530.
- [S7] R. Barbieri et al. “Searching for galactic axions through magnetized media: the QUAX proposal”. In: (2016). arXiv:1606.02201 [hep-ph].
- [S8] A. Bazavov et al. “Equation of state in (2+1)-flavor QCD”. In: *Phys. Rev.* D90 (2014), p. 094503. arXiv:1407.6387 [hep-lat].
- [S9] A. Bazavov et al. “Topological susceptibility with the asqtad action”. In: *Phys. Rev.* D81 (2010), p. 114501. arXiv:1003.5695 [hep-lat].
- [S10] A. Bazavov et al. “Update on the 2+1+1 flavor QCD equation of state with HISQ”. In: *PoS LATTICE2013* (2014), p. 154. arXiv:1312.5011 [hep-lat].
- [S11] R. Bellwied et al. “Fluctuations and correlations in high temperature QCD”. In: *Phys. Rev.* D92.11 (2015), p. 114505. arXiv:1507.04627 [hep-lat].
- [S12] Achim Bode et al. “First experiences with HMC for dynamical overlap fermions”. In: *Lattice fermions and structure of the vacuum. Proceedings, NATO Advanced Research Workshop, Dubna, Russia, October 5-9, 1999*. 1999, pp. 65–68. arXiv:hep-lat/9912043 [hep-lat].
- [S13] Claudio Bonati et al. “Axion phenomenology and θ -dependence from $N_f = 2 + 1$ lattice QCD”. In: *JHEP* 03 (2016), p. 155. arXiv:1512.06746 [hep-lat].
- [S14] S. Borsanyi et al. “Axion cosmology, lattice QCD and the dilute instanton gas”. In: *Phys. Lett.* B752 (2016), pp. 175–181. arXiv:1508.06917 [hep-lat].
- [S15] Sz. Borsanyi et al. “Ab initio calculation of the neutron-proton mass difference”. In: *Science* 347 (2015), pp. 1452–1455. arXiv:1406.4088 [hep-lat].
- [S16] Sz. Borsanyi et al. “Precision SU(3) lattice thermodynamics for a large temperature range”. In: *JHEP* 07 (2012), p. 056. arXiv:1204.6184 [hep-lat].
- [S17] Sz. Borsanyi et al. “QCD thermodynamics with continuum extrapolated dynamical overlap fermions”. In: (2015). arXiv:1510.03376 [hep-lat].
- [S18] Szabolcs Borsanyi et al. “Full result for the QCD equation of state with 2+1 flavors”. In: *Phys. Lett.* B730 (2014), pp. 99–104. arXiv:1309.5258 [hep-lat].
- [S19] Szabolcs Borsanyi et al. “High-precision scale setting in lattice QCD”. In: *JHEP* 09 (2012), p. 010. arXiv:1203.4469 [hep-lat].
- [S20] Szabolcs Borsanyi et al. “QCD thermodynamics with dynamical overlap fermions”. In: *Phys. Lett.* B713 (2012), pp. 342–346. arXiv:1204.4089 [hep-lat].

- [S21] Szabolcs Borsanyi et al. “The QCD equation of state and the effects of the charm”. In: *PoS LATTICE2011* (2011), p. 201. arXiv:1204.0995 [hep-lat].
- [S22] Szabolcs Borsanyi et al. “The QCD equation of state with dynamical quarks”. In: *JHEP* 11 (2010), p. 077. arXiv:1007.2580 [hep-lat].
- [S23] Florian Burger et al. “Towards thermodynamics with $N_f = 2 + 1 + 1$ twisted mass quarks”. In: *PoS Lattice2013* (2013), p. 153. arXiv:1311.1631 [hep-lat].
- [S24] Michael Cheng. “Charm Quarks and the QCD Equation of State”. In: *PoS LAT2007* (2007), p. 173. arXiv:0710.4357 [hep-lat].
- [S25] M. A. Clark et al. “Improving dynamical lattice QCD simulations through integrator tuning using Poisson brackets and a force-gradient integrator”. In: *Phys. Rev. D* 84 (2011), p. 071502. arXiv:1108.1828 [hep-lat].
- [S26] Nicol Crescini et al. “The QUAX- $g_p g_s$ experiment to search for monopole-dipole Axion interaction”. In: (2016). arXiv:1606.04751 [physics.ins-det].
- [S27] N. Cundy et al. “Numerical methods for the QCD overlap operator IV: Hybrid Monte Carlo”. In: *Comput. Phys. Commun.* 180 (2009), pp. 26–54. arXiv:hep-lat/0502007 [hep-lat].
- [S28] Thomas DeGrand and Stefan Schaefer. “Simulating an arbitrary number of flavors of dynamical overlap fermions”. In: *JHEP* 07 (2006), p. 020. arXiv:hep-lat/0604015 [hep-lat].
- [S29] Thomas A. DeGrand, Anna Hasenfratz, and Tamas G. Kovacs. “Improving the chiral properties of lattice fermions”. In: *Phys. Rev. D* 67 (2003), p. 054501. arXiv:hep-lat/0211006 [hep-lat].
- [S30] Michael Dine, Willy Fischler, and Mark Srednicki. “A Simple Solution to the Strong CP Problem with a Harmless Axion”. In: *Phys. Lett. B* 104 (1981), pp. 199–202.
- [S31] Michela D’Onofrio and Kari Rummukainen. “Standard model cross-over on the lattice”. In: *Phys. Rev. D* 93.2 (2016), p. 025003. arXiv:1508.07161 [hep-ph].
- [S32] Michela D’Onofrio, Kari Rummukainen, and Anders Tranberg. “Sphaleron Rate in the Minimal Standard Model”. In: *Phys. Rev. Lett.* 113.14 (2014), p. 141602. arXiv:1404.3565 [hep-ph].
- [S33] S. Durr et al. “Ab-Initio Determination of Light Hadron Masses”. In: *Science* 322 (2008), pp. 1224–1227. arXiv:0906.3599 [hep-lat].
- [S34] S. Durr et al. “Lattice QCD at the physical point: light quark masses”. In: *Phys. Lett. B* 701 (2011), pp. 265–268. arXiv:1011.2403 [hep-lat].
- [S35] S. Durr et al. “Lattice QCD at the physical point: Simulation and analysis details”. In: *JHEP* 08 (2011), p. 148. arXiv:1011.2711 [hep-lat].
- [S36] Stephan Durr. “Logarithmic link smearing for full QCD”. In: *Comput. Phys. Commun.* 180 (2009), pp. 1338–1357. arXiv:0709.4110 [hep-lat].
- [S37] Stephan Durr. “Taste-split staggered actions: eigenvalues, chiralities and Symanzik improvement”. In: *Phys. Rev. D* 87.11 (2013), p. 114501. arXiv:1302.0773 [hep-lat].
- [S38] Stephan Durr. “Theoretical issues with staggered fermion simulations”. In: *PoS LAT2005* (2006), p. 021. arXiv:hep-lat/0509026 [hep-lat].
- [S39] Stephan Durr, Christian Hoelbling, and Urs Wenger. “Staggered eigenvalue mimicry”. In: *Phys. Rev. D* 70 (2004), p. 094502. arXiv:hep-lat/0406027 [hep-lat].
- [S40] Stephan Durr et al. “Lattice QCD at the physical point meets SU(2) chiral perturbation theory”. In: *Phys. Rev. D* 90.11 (2014), p. 114504. arXiv:1310.3626 [hep-lat].
- [S41] Stephan Durr et al. “Precision study of the SU(3) topological susceptibility in the continuum”. In: *JHEP* 04 (2007), p. 055. arXiv:hep-lat/0612021 [hep-lat].
- [S42] G. I. Egri et al. “Topology with dynamical overlap fermions”. In: *JHEP* 01 (2006), p. 049. arXiv:hep-lat/0510117 [hep-lat].

- [S43] G. Endrodi et al. “The Equation of state at high temperatures from lattice QCD”. In: *PoS LAT2007* (2007), p. 228. arXiv:0710.4197 [hep-lat].
- [S44] Z. Fodor, S. D. Katz, and K. K. Szabo. “Dynamical overlap fermions, results with hybrid Monte Carlo algorithm”. In: *JHEP* 08 (2004), p. 003. arXiv:hep-lat/0311010 [hep-lat].
- [S45] Z. Fodor et al. “Up and down quark masses and corrections to Dashen’s theorem from lattice QCD and quenched QED”. In: (2016). arXiv:1604.07112 [hep-lat].
- [S46] J. Frison et al. “Topological susceptibility at high temperature on the lattice”. In: (2016). arXiv:1606.07175 [hep-lat].
- [S47] Hidenori Fukaya et al. “Lattice gauge action suppressing near-zero modes of H(W)”. In: *Phys. Rev. D* 74 (2006), p. 094505. arXiv:hep-lat/0607020 [hep-lat].
- [S48] Maarten F. L. Golterman. “STAGGERED MESONS”. In: *Nucl. Phys. B* 273 (1986), pp. 663–676.
- [S49] M. Hasenbusch. “Speeding up finite step size updating of full QCD on the lattice”. In: *Phys. Rev. D* 59 (1999), p. 054505. arXiv:hep-lat/9807031 [hep-lat].
- [S50] Vicente Hernandez, Jose E. Roman, and Vicente Vidal. “SLEPc: A Scalable and Flexible Toolkit for the Solution of Eigenvalue Problems”. In: *ACM Trans. Math. Software* 31.3 (2005), pp. 351–362.
- [S51] Mark Hindmarsh and Owe Philipsen. “WIMP dark matter and the QCD equation of state”. In: *Phys. Rev. D* 71 (2005), p. 087302. arXiv:hep-ph/0501232 [hep-ph].
- [S52] Dieter Horns et al. “Searching for WISPy Cold Dark Matter with a Dish Antenna”. In: *JCAP* 1304 (2013), p. 016. arXiv:1212.2970 [hep-ph].
- [S53] N. Ishizuka et al. “Operator dependence of hadron masses for Kogut-Susskind quarks on the lattice”. In: *Nucl. Phys. B* 411 (1994), pp. 875–902.
- [S54] Joerg Jaeckel and Javier Redondo. “Resonant to broadband searches for cold dark matter consisting of weakly interacting slim particles”. In: *Phys. Rev. D* 88.11 (2013), p. 115002. arXiv:1308.1103 [hep-ph].
- [S55] K. Kajantie et al. “Generic rules for high temperature dimensional reduction and their application to the standard model”. In: *Nucl. Phys. B* 458 (1996), pp. 90–136. arXiv:hep-ph/9508379 [hep-ph].
- [S56] K. Kajantie et al. “The Pressure of hot QCD up to $g_6 \ln(1/g)$ ”. In: *Phys. Rev. D* 67 (2003), p. 105008. arXiv:hep-ph/0211321 [hep-ph].
- [S57] F. Karsch, E. Laermann, and A. Peikert. “The Pressure in two flavor, (2+1)-flavor and three flavor QCD”. In: *Phys. Lett. B* 478 (2000), pp. 447–455. arXiv:hep-lat/0002003 [hep-lat].
- [S58] Jihn E. Kim. “Weak Interaction Singlet and Strong CP Invariance”. In: *Phys. Rev. Lett.* 43 (1979), p. 103.
- [S59] M. Laine and M. Meyer. “Standard Model thermodynamics across the electroweak crossover”. In: *JCAP* 1507.07 (2015), p. 035. arXiv:1503.04935 [hep-ph].
- [S60] M. Laine, G. Nardini, and K. Rummukainen. “First order thermal phase transition with 126 GeV Higgs mass”. In: *PoS LATTICE2013* (2014), p. 104. arXiv:1311.4424 [hep-lat].
- [S61] Mikko Laine and York Schroder. “Quark mass thresholds in QCD thermodynamics”. In: *Phys. Rev. D* 73 (2006), p. 085009. arXiv:hep-ph/0603048 [hep-ph].
- [S62] Alessandro Laio, Guido Martinelli, and Francesco Sanfilippo. “Metadynamics Surfing on Topology Barriers: the CP(N-1) Case”. In: (2015). arXiv:1508.07270 [hep-lat].
- [S63] Ludmila Levkova. “Effects of the charm quark on the QCD equation of state”. In: *PoS LAT2009* (2009), p. 170. arXiv:0910.3006 [hep-lat].
- [S64] Martin Luscher. “Properties and uses of the Wilson flow in lattice QCD”. In: *JHEP* 08 (2010). [Erratum: *JHEP*03,092(2014)], p. 071. arXiv:1006.4518 [hep-lat].

- [S65] Martin Luscher and Stefan Schaefer. “Lattice QCD without topology barriers”. In: *JHEP* 07 (2011), p. 036. arXiv:1105.4749 [hep-lat].
- [S66] Simon Mages et al. “Lattice QCD on Non-Orientable Manifolds”. In: (2015). arXiv:1512.06804 [hep-lat].
- [S67] A Millar et al. *Layered dielectric haloscopes: a new way to detect axion dark matter*, MPP-2016-141, in preparation.
- [S68] K. A. Olive et al. “Review of Particle Physics”. In: *Chin. Phys.* C38 (2014), p. 090001.
- [S69] A. Ringwald and F. Schrempp. “Confronting instanton perturbation theory with QCD lattice results”. In: *Phys. Lett.* B459 (1999), pp. 249–258. arXiv:hep-lat/9903039 [hep-lat].
- [S70] Stefan Schaefer, Rainer Sommer, and Francesco Virotta. “Critical slowing down and error analysis in lattice QCD simulations”. In: *Nucl. Phys.* B845 (2011), pp. 93–119. arXiv:1009.5228 [hep-lat].
- [S71] Mikhail A. Shifman, A. I. Vainshtein, and Valentin I. Zakharov. “Can Confinement Ensure Natural CP Invariance of Strong Interactions?” In: *Nucl. Phys.* B166 (1980), pp. 493–506.
- [S72] P. Sikivie. “Axion Dark Matter Detection using Atomic Transitions”. In: *Phys. Rev. Lett.* 113.20 (2014), p. 201301. arXiv:1409.2806 [hep-ph].
- [S73] P. Sikivie. “Experimental Tests of the Invisible Axion”. In: *Phys. Rev. Lett.* 51 (1983). [Erratum: *Phys. Rev. Lett.* 52,695(1984)], pp. 1415–1417.
- [S74] Mark Srednicki, Richard Watkins, and Keith A. Olive. “Calculations of Relic Densities in the Early Universe”. In: *Nucl. Phys.* B310 (1988), p. 693.
- [S75] Jun'ya Suzuki et al. “Hidden photon CDM search at Tokyo”. In: *Proceedings, 11th Patras Workshop on Axions, WIMPs and WISPs (Axion-WIMP 2015)*. 2015. arXiv:1509.00785 [hep-ex].
- [S76] Olivier Wantz and E. P. S. Shellard. “Axion Cosmology Revisited”. In: *Phys. Rev.* D82 (2010), p. 123508. arXiv:0910.1066 [astro-ph.CO].
- [S77] Hantao Yin and Robert D. Mawhinney. “Improving DWF Simulations: the Force Gradient Integrator and the M6bius Accelerated DWF Solver”. In: *PoS LATTICE2011* (2011), p. 051. arXiv:1111.5059 [hep-lat].
- [S78] Cheng-xing Zhai and Boris M. Kastening. “The Free energy of hot gauge theories with fermions through g^{*5} ”. In: *Phys. Rev.* D52 (1995), pp. 7232–7246. arXiv:hep-ph/9507380 [hep-ph].
- [S79] A. R. Zhitnitsky. “On Possible Suppression of the Axion Hadron Interactions. (In Russian)”. In: *Sov. J. Nucl. Phys.* 31 (1980). [*Yad. Fiz.* 31,497(1980)], p. 260.