

Dynamical de Sitter phase and nontrivial holonomy in strongly coupled gauge theories in an expanding universe

Ariel R. Zhitnitsky

Department of Physics & Astronomy, University of British Columbia, Vancouver, B.C. V6T 1Z1, Canada
(Received 9 June 2015; published 18 August 2015)

We discuss a new scenario for early cosmology when the inflationary de Sitter phase emerges dynamically. This genuine quantum effect occurs as a result of dynamics of the topologically nontrivial sectors in a strongly coupled QCD-like gauge theory in an expanding universe. We test these ideas by explicit computations in hyperbolic space $\mathbb{H}_k^3 \times \mathbb{S}_{k-1}^1$. We argue that the key element for this idea to work is the presence of nontrivial holonomy computed along \mathbb{S}_{k-1}^1 . The effect is nonlocal in nature, nonanalytical in coupling constant and cannot be described in terms of any local propagating degree of freedom such as scalar inflaton field $\Phi(x)$. We discuss some profound phenomenological consequences of this scenario for inflationary cosmology. We also suggest to test these ideas in a tabletop experiment by measuring some specific corrections to the Casimir pressure in the Maxwell theory formulated on a topologically nontrivial manifold.

DOI: 10.1103/PhysRevD.92.043512

PACS numbers: 11.15.-q, 11.15.Kc, 11.15.Tk

I. INTRODUCTION: MOTIVATION

The main motivation for the present studies is the proposal that inflationary de Sitter phase [1–3] may be dynamically generated as a result of presence of the topologically nontrivial sectors in expanding universe. Inflaton in this framework [4,5] is an auxiliary topological nonpropagating field with no canonical kinetic term, similar to known topologically ordered phases in condensed matter systems. This auxiliary field effectively describes the dynamics of the topological sectors $|k\rangle$ in a gauge theory (coined as $\overline{\text{QCD}}$ in [4,5]) in expanding Universe.

This picture should be contrasted with conventional proposals reviewed in [2,3] when the de Sitter behavior is achieved in quantum field theory (QFT) by assuming the existence of a new scalar local field $\Phi(x)$ with a nonvanishing potential energy density $V(\Phi)$. The shape of this potential energy can be adjusted in a such a way that the contribution to energy density ϵ and pressure p is in agreement with observations. In different words, the scale parameter $a(t)$ and the equation of state during the inflation take the following approximate form,

$$a(t) \sim \exp(Ht), \quad \epsilon \approx -p. \quad (1)$$

The key ingredient of the proposal [4,5] is a conjecture that the vacuum energy in context of the Friedmann-Lemaître-Robertson-Walker (FLRW) universe has the following expansion at small $H \ll \Lambda_{\overline{\text{QCD}}}$

$$E_{\text{FLRW}}(H) \sim \left[\Lambda_{\overline{\text{QCD}}}^4 + H \Lambda_{\overline{\text{QCD}}}^3 + \mathcal{O}(H^2) \right], \quad (2)$$

when the first nonvanishing term is linear $\sim H$, rather than (commonly accepted) quadratic $\sim H^2$ in the Hubble

constant. If this conjecture turns out to be correct, than the Friedman equation assumes the form

$$H^2 \simeq \frac{8\pi G}{3} \Delta E, \Rightarrow H_0 \simeq \frac{8\pi G}{3} \Lambda_{\overline{\text{QCD}}}^3$$

$$\Delta E \equiv [E_{\text{FLRW}}(H) - E_{\text{Mink}}] \sim H, \quad (3)$$

which automatically leads to a nontrivial solution with constant H_0 , and as a consequence, to a desired de Sitter behavior (1).

There are two critical elements in writing Eq. (3). The first one, as we already mentioned, is related to the expansion (2), see few comments on this conjecture below. The second key element is a paradigm that the relevant definition of the energy in an expanding background which enters the Friedman equation is the difference $\Delta E(H) \equiv [E(H) - E_{\text{Mink}}]$, similar to the computation of the Casimir pressure when the observable energy is the difference similar to ΔE . This element in our analysis is not a new proposal identifying ΔE with gravitating energy from the Friedman equation. In fact, in the present context such a definition for the vacuum energy was advocated long ago in 1967 by Zeldovich [6] for the first time. Later on such definition for the relevant energy $\Delta E \equiv (E_{\text{FLRW}} - E_{\text{Mink}})$ which enters the Friedman equations has been advocated from different perspectives in a number of papers, see, e.g., relatively recent works [7–11], see also review article [12] with a large number of references on original papers. Essentially, this prescription implies that ΔE may only depend on properties of the external gravitational background, while the conventional contributions computed in Minkowski flat space-time (such as the QCD vacuum energy or the Higgs potential in electroweak theory) are automatically subtracted by this

prescription.¹ We shall not elaborate on a number of subtle points related to this prescription in the present work referring to the original papers and review article [12].

The main topic of the preset paper is the analysis of another critical element, briefly mentioned above, and leading to (3). There is a well-known, conventional, and generally accepted argument which suggests that the expansion (2) starts with a quadratic $\sim H^2$, rather than the linear $\sim H$ term. The argument is based on fundamental principles of *locality* and general covariance, see original papers [15,16], recent review [12], and some comments [4] with pros and cons of these arguments. Indeed, the curvature scalar R for FLRW universe is quadratic in H ,

$$|R| = 6 \left(\frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} \right) = 12H^2 + 6\dot{H}, \quad (4)$$

when $\dot{H} \sim \mathcal{O}(H^2)$, see [12]. Therefore, if the infrared (IR) behavior of the system is entirely determined by the local characteristics, such as curvature scalar R and/or higher order derivative terms $R^2, R_{\mu\nu}R^{\mu\nu}$, then the corrections to the energy (2) indeed must be proportional to even powers H^{2n} as correctly argued in [12,15,16].

However, the main essence of the proposal [4,5] is precisely the observation that the conventional assumption on locality might be badly violated in strongly coupled gauge theories. The basic reason for such violation is well known and well understood, at least in Minkowski space-time. The energy (2) is generated due to the tunneling events between $|k\rangle$ topological sectors, which formulated in terms of inherently *nonlocal* large gauge transformation operator \mathcal{T} . Furthermore, this energy has nondispersive nature, i.e., it cannot be formulated in terms of any *local* propagating degrees of freedom.² Transition from Minkowski space-time to time dependent background (2) obviously will not modify the nature and origin of this type of energy. Rather, a transition to FLRW universe introduces some background-dependent corrections to the same type

¹A somewhat similar, but not identically the same subtraction procedure has been suggested recently in Refs. [13,14], the so-called “vacuum energy sequestering” proposal. The prescription [13,14] is also inherently nonlocal, similar to the crucial role of nonlocality in our framework realized in terms of the holonomy (5). In fact, our computation of the vacuum energy (2) as discussed below, is based on evaluation of the holonomy (5) along the entire history of the universe, which resembles in spirit the computations of the so-called “historic averages” in Refs. [13,14]. Furthermore, we have to keep the volume of the system to be finite in the computations for the infrared regularization of the theory. It is akin to that from Refs. [13,14] where the finite volume is also required property for the consistency of the procedure.

²This energy can be expressed in terms of the contact term in the topological susceptibility, determined by the IR physics and boundary conditions. The corresponding physics has been well understood using the lattice numerical simulations in strong coupling regime, see [4] for references and details.

of energy (2), which was coined as “strange energy” in [4,5] due to its unconventional origin as mentioned above.

One should comment here that this feature of *nonlocality* when the system is not completely characterized by a local physics is very similar to the well-known property in topologically ordered phases in condensed matter physics wherein an expectation value of a local operator does not fully characterize the ground state of the system. Instead, one should use some *nonlocal* variables for proper characterization of the system.

The main subject of the present work is to elaborate and clarify a number of nontrivial questions related to the nonlocality in QFT and generation of the linear $\sim H$ term in (2) in some simplified models.³ The basic point of our discussions is that a gauge QFT (when one should sum over all topological sectors $|k\rangle$ in the definition of the partition function) is not fully described by the local characteristics, such as curvature (4). In particular, the linear dependence on the background may enter (2) through other characteristics of the system such as holonomy

$$U(\mathbf{x}) = \mathcal{P} \exp \left(i \int_0^\beta dx_4 A_4(x_4, \mathbf{x}) \right), \quad (5)$$

which is gauge invariant but nonlocal object as it depends on the boundary conditions. We shall argue below that precisely the nontrivial holonomy in gauge theories plays a key role in the mechanism which could generate the “strange energy” (2). It is very hard technical problem to compute the nonperturbative energy (2) in a time-dependent background characterized by parameter H , see footnote 3 for clarification. However, one can simplify the problem by considering the sensitivity of a gauge system to some external dimensional parameters characterizing the gravitational background, such as κ , see definition below. This parameter plays a role similar to the Hubble constant H in FLRW universe (2). Our goal is to study the dependence of the “strange energy” (2) as a function of κ in the limit of small $\kappa \rightarrow 0$ in some simple settings where such computations can be performed.

The basic idea is as follows. We would like to consider hyperbolic space \mathbb{H}_κ^3 with the constant negative curvature $-\kappa^2$. As we discuss below, there is a conformal equivalence between $(\mathbb{R}^4 - \mathbb{R}^2)$ and $\mathbb{H}_\kappa^3 \times \mathbb{S}_{\kappa^{-1}}^1$ where $\mathbb{S}_{\kappa^{-1}}^1$ denotes the circle of radius κ^{-1} . The holonomy (5) is computed precisely along a closed loop $\mathbb{S}_{\kappa^{-1}}^1$. Our goal is to study the first nontrivial correction $\sim \kappa$ to the nonperturbative

³Here and in what follows we use term “linear in H correction” as a generic feature to distinguish a nontrivial background from the trivial Euclidean space. This dimensional parameter should not be literally identified with the Hubble constant. Rather, it could be any other dimensional parameter which characterizes the system, such as the size of torus related to the nontrivial holonomy (5) with $\beta \sim H^{-1}$.

energy (2) in the limit of small $\kappa \rightarrow 0$ corresponding to smooth transition to Euclidean space \mathbb{R}^4 . We should recover the Euclidean results when κ identically vanishes.

The key observation is that the topological configurations with nontrivial holonomy (5) produce a finite contribution to the energy density (2) with corrections being linearly proportional to κ . Such effects cannot be expressed in terms of any local operators such as curvature (4). Rather, it is generated due to topological vacuum configurations with nontrivial holonomy (5), not expressible in terms of local observables. This is precisely the reason why the generic arguments [12,15,16] based on locality simply do not apply here.

Therefore, our computations of the linear correction $\sim \kappa$ in the vacuum energy density using simplified model with $\mathbb{H}_\kappa^3 \times \mathbb{S}_{\kappa^{-1}}^1$ background can be thought of as a strong argument supporting our conjecture on the linear correction $\sim H$ in a generic FLRW universe (2)—in both cases the linear correction is not associated with the local curvature operator (4).

One may wonder how a bulk property (such as vacuum energy density) in a gapped theory could be ever sensitive to such global characteristics as radius of the circle of $\mathbb{S}_{\kappa^{-1}}^1$? The answer lies not in the local, but global properties of the space. Imagine that we study the Aharonov-Casher effect. We insert an external charge into a superconductor in which the electric field is exponentially suppressed $\sim \exp(-r/\lambda)$ with λ being the penetration depth. Nevertheless, a neutral magnetic fluxon will be still sensitive to an inserted external charge at arbitrary large distances in spite of the screening of the physical field (which is equivalent to the presence of a gap in our system). This genuine quantum effect is purely topological and nonlocal in nature and can be explained in terms of the dynamics of the gauge sectors which are responsible for the long range dynamics. Imagine now that we study the same effect but in a different background. The corresponding topological sectors will be modified due to the variation of the external background. However, this modification cannot be described in terms of any local dynamical fields, as there are not any propagating long range fields in the system since the physical electric field is screened. For this simplified example, the dynamics of the “strange energy” as a function of κ is determined by the modifications of topological sectors when the background varies. The effect is obviously nonlocal in nature as the Aharonov-Casher effect itself is a nonlocal phenomenon.

The readers interested in the cosmological consequences, rather than in technical computational details may directly jump to Sec. III D where we list the main results of this framework. Section IV is our conclusion where we suggest to test some of the ideas presented in this work in a tabletop experiment by measuring some specific corrections to the Casimir vacuum energy in the Maxwell theory formulated on 4-torus with the nontrivial holonomy.

II. THE NATURE OF “STRANGE” ENERGY (2) IN EUCLIDEAN SPACE \mathbb{R}^4

The main goal of this section is to review a number of crucial elements relevant for our studies. We start in Sec. II A with explanation of a highly nontrivial nature of the “strange energy” (2) in the Euclidean space time. We continue in Sec. II B, by clarifying the crucial role of the holonomy (5) in generating such type of energy. We present few technical results in Secs. II C, II D. Finally, in Sec. II E we elaborate on nonlocal features of relevant vacuum configurations saturating the “strange energy” in cosmological context. The corresponding analysis will play an important role in our main Sec. III when we study the hyperbolic space $\mathbb{H}_\kappa^3 \times \mathbb{S}_{\kappa^{-1}}^1$ and analyze the properties of the “strange” energy as a function of κ at small $\kappa \rightarrow 0$.

A. The topological susceptibility and contact term

We start our short overview on the “strange” nature of the vacuum energy (2) by reviewing a naively unrelated topic—the formulation and resolution of the so-called $U(1)_A$ problem in strongly coupled QCD [17–19]. We introduce the topological susceptibility χ which is ultimately related to the vacuum energy $E_{\text{vac}}(\theta = 0)$ as follows⁴

$$\chi = \left. \frac{\partial^2 E_{\text{vac}}(\theta)}{\partial \theta^2} \right|_{\theta=0} = \lim_{k \rightarrow 0} \int d^4 x e^{ikx} \langle T \{ q(x), q(0) \} \rangle, \quad (6)$$

where θ parameter enters the Lagrangian along with topological density operator $q(x) = \frac{1}{16\pi^2} \text{tr} [F_{\mu\nu} \tilde{F}^{\mu\nu}]$ and $E_{\text{vac}}(\theta)$ is the “strange” vacuum energy density which represents the first term in expansion (2) corresponding to the flat space-time background. This θ -dependent portion of the vacuum energy (computed at $\theta = 0$) has a number of unusual properties as we review below. The corresponding properties are easier to explain in terms of the correlation function (6), rather than in terms of the vacuum energy $E_{\text{vac}}(\theta = 0)$ itself. The relation between the two is given by Eq. (6).

First of all, the topological susceptibility χ does not vanish in spite of the fact that $q = \partial_\mu K^\mu$ is total divergence. This feature is very different from any conventional correlation functions which normally must vanish at zero momentum if the corresponding operator can be represented as total divergence.

Second, any physical $|n\rangle$ state gives a negative contribution to this diagonal correlation function

$$\begin{aligned} \chi_{\text{dispersive}} &\sim \lim_{k \rightarrow 0} \int d^4 x e^{ikx} \langle T \{ q(x), q(0) \} \rangle \\ &\sim \lim_{k \rightarrow 0} \sum_n \frac{\langle 0|q|n\rangle \langle n|q|0\rangle}{-k^2 - m_n^2} \simeq - \sum_n \frac{|c_n|^2}{m_n^2} \leq 0, \end{aligned} \quad (7)$$

⁴We use the Euclidean notations where path integral computations are normally performed.

where m_n is the mass of a physical $|n\rangle$ state, $k \rightarrow 0$ is its momentum, and $\langle 0|q|n\rangle = c_n$ is its coupling to topological density operator $q(x)$. At the same time the resolution of the $U(1)_A$ problem requires a positive sign for the topological susceptibility (6), see the original reference [19] for a thorough discussion,

$$\chi_{\text{nondispersive}} = \lim_{k \rightarrow 0} \int d^4x e^{ikx} \langle T\{q(x), q(0)\} \rangle > 0. \quad (8)$$

Therefore, there must be a contact contribution to χ , which is not related to any propagating physical degrees of freedom, and it must have the “wrong” sign. The “wrong” sign in this paper implies a sign which is opposite to any contributions related to the physical propagating degrees of freedom (7). The “strange energy” in this paper implies the θ dependent portion of the energy (2), (6) which *cannot* be formulated in terms of conventional propagating degrees of freedom as it has pure nondispersive nature according to Eqs. (7), (8).

In the framework [17] the contact term with “wrong” sign has been simply postulated, while in Refs. [18,19] the Veneziano ghost (with a “wrong” kinetic term) had been introduced into the theory to saturate the required property (8).

Third, the contact term (8) has the structure $\chi \sim \int d^4x \delta^4(x)$. The significance of this structure is that the gauge variant correlation function in momentum space

$$\lim_{k \rightarrow 0} \int d^4x e^{ikx} \langle K_\mu(x), K_\nu(0) \rangle \sim \frac{k_\mu k_\nu}{k^4} \quad (9)$$

develops a topologically protected “unphysical” pole which does not correspond to any propagating massless degrees of freedom, but nevertheless must be present in the system. Furthermore, the residue of this pole has the “wrong sign.” This “wrong sign” is due to the Veneziano ghost contribution saturating the nondispersive term in gauge invariant correlation function (8),

$$\langle q(x)q(0) \rangle \sim \langle \partial_\mu K^\mu(x), \partial_\nu K^\nu(0) \rangle \sim \delta^4(x). \quad (10)$$

We conclude this review-type subsection with the following remark. The entire framework, including the singular behavior of $\langle q(x)q(0) \rangle$ with the “wrong sign,” has been well confirmed by numerous lattice simulations in a strong coupling regime, and it is accepted by the community as a standard resolution of the $U(1)_A$ problem. Furthermore, it has been argued long ago in Ref. [20] that the gauge theories may exhibit the “secret long range forces” expressed in terms of the correlation function (9) with topologically protected pole at $k = 0$.

Finally, in a weakly coupled gauge theory (the so-called “deformed QCD” model [21]) where all computations can be performed in theoretically controllable way one can

explicitly test every single element of this entire framework, including the topologically protected pole (9), the contact term with “wrong sign,” etc., see Refs. [22,23] for the details. In particular, one can explicitly see that the Veneziano ghost is in fact an auxiliary topological field which saturates the vacuum energy and the topological susceptibility χ . What is more important for the present studies is that one can explicitly see that the holonomy (5) plays a crucial role in generating the “strange” vacuum energy and contact term in topological susceptibility.

While all these unusual features of the vacuum energy are well-known and well-supported by numerous lattice simulations (see, e.g., [23] for a large number of references on original lattice results) the analytical understanding of these properties in strong coupling regime is still lacking. In the next subsection we review some known results on this matter specifically emphasizing the role of the holonomy (5) in the analytical computations. Precisely a nontrivial holonomy (5) may play a crucial role in generating the linear correction $\sim H$ in Eq. (2) as we argue in Sec. III D. This is the key technical element which pinpoints the source of linear corrections $\sim H$ not expressible in terms of any local operators such as curvature (4).

In conclusion we should comment that the vacuum energy in the electroweak (EW) sector of the standard model is not sensitive to the topological features of the EW gauge fields of the standard model (W_μ^\pm, Z_μ, A_μ) as these topological properties play no role in dynamics of these fields, in huge contrast with QCD. This is due to the fact that the mass gap of the non-Abelian EW gauge bosons is resulted from the Higgs mechanism, in contrast with QCD where the mass gap and the vacuum energy are dynamically generated by the topological fluctuations with nontrivial holonomy (5). Therefore, the subtraction procedure, formulated in Introduction, automatically removes all the vacuum energy related to the EW sector of the standard model without any corrections $\sim H$. The linear corrections are specific to the strongly coupled QCD with its topological features formulated in terms of the nonlocal operators, the holonomy (5), and large gauge transformations operator T as discussed above.

B. The holonomy (5) and generation of the “strange” energy in Euclidean space

The key role in our discussions will play the behavior of holonomy $U(\mathbf{x})$ at spatial infinity, the Polyakov line,

$$L = \mathcal{P} \exp \left(i \int_0^\beta dx_4 A_4(x_4, |\mathbf{x}| \rightarrow \infty) \right). \quad (11)$$

The operator $\text{Tr}L$ classifies the self-dual solutions which may contribute to the path integral at finite temperature $T \equiv \beta^{-1}$, including the low temperature limit $T \rightarrow 0$. There is a well-known generalization of the standard self-dual instantons to nonzero temperature, which corresponds to

the description on $\mathbb{R}^3 \times \mathbb{S}^1$ geometry. These are the so-called periodic instantons, or calorons [24] studied in detail in [25]. These calorons have trivial holonomy, which implies that the $\text{Tr}L$ assumes values belonging to the group center \mathbb{Z}_N for the $SU(N)$ gauge group.

A more general class of the self-dual solutions with nontrivial holonomy (11), the so-called KvBLL calorons were constructed more recently [26,27]. In this case the holonomy (11) in general, is not reduced to the group centre $\text{Tr}L \notin \mathbb{Z}_N$. The fascinating feature of the KvBLL calorons is that they can be viewed as a set of N monopoles of N different types. Normally, one expects that monopoles come in $N - 1$ different varieties carrying a unit magnetic charge from each of the $U(1)$ factors of the $U(1)^{N-1}$ gauge group left unbroken by vacuum expectation value due to nontrivial holonomy (11). There is an additional, so called Kaluza-Klein (KK) monopole which carries magnetic charges and instanton charge. All monopoles' charges are such that when a complete set of different types of monopoles is present, the magnetic charges exactly cancel, and the configuration of N different monopoles carries a unit instanton charge.

It has been known since [25] that the gauge configurations with nontrivial holonomy are strongly suppressed in the partition function. Therefore, naively KvBLL calorons cannot produce a finite contribution to the partition function. However, this naive argument is based on consideration of the individual KvBLL caloron, or finite number of them. If one considers a grand canonical assemble of these objects than their density is determined by the dynamics, and the old argument of Ref. [25] breaks down. The corresponding objects in this case may in fact produce a finite contribution to the partition function. A self-consistent computations in a weak coupling regime supporting this picture have been carried out in the so-called “deformed QCD” model [21]. One can explicitly see how N different types of monopoles with nontrivial holonomy (11) which carry fractional topological charge $\pm 1/N$ produce confinement, generate the “strange” vacuum energy (6) and associated with this energy the topological susceptibility (10) with known, but highly unusual properties reviewed above in Sec. II, see [22,23] for the technical details on these computations.

In the strong coupling regime we are interested in, the corresponding analytical computations have never been completed. There is a limited number of partial analytical and numerical results [28–32] on computations of moduli space and one loop determinant, controlling the dynamics and interaction properties of the constituents in a large ensemble of KvBLL calorons. We review these basic technical results next in Sec. II C, as they will play an important role in our analysis below.

While a complete analytical solution in strong coupling regime is still lacking, nevertheless there is a number of hints supporting the basic picture that the KvBLL configurations

with nontrivial holonomy (11) and representing N different types of monopoles with fractional topological charges $\pm 1/N$ saturate the “strange” vacuum energy (6) and associated with this energy the topological susceptibility (10) in a very much the same way as it happens in “deformed QCD” model where all computations are performed in a theoretically controllable regime [21–23].

In what follows we assume that the “strange” vacuum energy (6) and associated with this energy the topological susceptibility (10) is indeed saturated by fractionally charged monopoles with $Q = \pm 1/N$ which are constituents of KvBLL caloron with nontrivial holonomy (11). With this assumption in hand the question which is addressed in this work is as follows. How does the “strange energy” vary in a hyperbolic space $\mathbb{H}_\kappa^3 \times \mathbb{S}_{\kappa^{-1}}^1$ as a function of dimensional parameter κ ? The difference between the original $\mathbb{R}^3 \times \mathbb{S}^1$ and $\mathbb{H}_\kappa^3 \times \mathbb{S}_{\kappa^{-1}}^1$ spaces is the curvature of the hyperbolic space $R[\mathbb{H}_\kappa^3] \sim \kappa^2$ at $\kappa \rightarrow 0$. If we find a linear dependence on κ at small κ it would be a strong argument supporting our conjecture (2) on linear dependence of “strange” vacuum energy as a function of external parameter. Such linear scaling obviously implies that this background-dependent correction is not generated by any local operators such as curvature (4), but rather is generated by nonlocal operators (5), (11) which is sensitive to the global characteristics of the background.

C. Nontrivial holonomy (11) in Euclidean space. Few technical details

In this section we present few formulas derived in Euclidean space in order to compare them with parallel expressions obtained in the hyperbolic space $\mathbb{H}_\kappa^3 \times \mathbb{S}_{\kappa^{-1}}^1$ from Sec. III. The corresponding comparison will allow us to study the dependence of the “strange” energy as a function of κ at small $\kappa \rightarrow 0$.

We start from analysis of the KvBLL configurations with nontrivial holonomy (11). We use $SU(2)$ gauge group in our discussions to simplify notations, though the generalization for $SU(N)$ is also known. The KvBLL caloron can be represented as a combination of two monopoles. The first monopole is a conventional Bogomolny-Prasad-Sommerfeld (BPS) monopole, and at large separations between the constituents in Hedgehog gauge can be represented as follows

$$\begin{aligned} A_4^M(r_i) &= \left(v \coth(vr) - \frac{1}{r} \right) \frac{r_a \tau^a}{2r} \\ A_i^M(r_i) &= \left(1 - \frac{vr}{\sinh(vr)} \right) \epsilon_{ijk} \frac{r_j \tau^k}{2r^2}, \end{aligned} \quad (12)$$

where we adopted the notations from Refs. [28–34] to coin this constituent as M -monopole. Parameter v in this formula is an arbitrary number which is determined by the holonomy (11). The classical moduli space is a circle,

$$v \in \mathbb{S}^1 : 0 \leq v \leq \frac{2\pi}{\beta}, \quad (13)$$

such that v is an angular variable. For any $v \neq 0$ the gauge group is broken to $U(1)$. In other words, parameter v plays the role of the vacuum expectation value of the Higgs field $\Phi(x)$, which is represented in this system by the A_4^M -component of the gauge field as expressed by Eq. (12). One should emphasize that we study gluodynamics, without a scalar Higgs field in the system. Nevertheless, the A_4 -component of the gauge field with nonvanishing expectation value v plays exactly the same role as the Higgs field in adjoint representation in the standard BPS equations. Therefore, it is not a mystery that the r dependence of the $A_4(r_i)$ -component in Eq. (12) is identically the same as in conventional BPS construction for the Higgs $\Phi(x)$ field.⁵

The most important property of solution (12) is the behavior of A_4^M at large distances, which is convenient to represent in the unitary gauge:

$$A_4^M(r \rightarrow \infty) \rightarrow \left(v - \frac{1}{r}\right) \frac{\tau^3}{2}, \quad (14)$$

where we also keep the Coulomb-like correction $1/r$ along with the leading term $\sim v$ because this Coulomb long range interaction will play an important role in our future discussions. One can explicitly see that the holonomy (11) is

$$\frac{1}{2} \text{Tr} L = \cos\left(\frac{\beta v}{2}\right) = \cos(\pi \nu), \quad (15)$$

where we introduced the dimensionless parameter $\nu \equiv \frac{\beta v}{2\pi}$. The holonomy belongs to the group center $\frac{1}{2} \text{Tr} L = \pm 1$ when v assumes its boundary points ($v = 0, \frac{2\pi}{\beta}$), in which case it is called the trivial holonomy.

The second type of monopole is the so-called L monopole which can be constructed from (12) as follows [26,27]. First, one should replace $v \rightarrow \frac{2\pi}{\beta} - v$, which is equivalent to replacement $\nu \rightarrow (1 - \nu)$. Then, one should make a ‘‘large’’ (improper) gauge transformation

$$U_{\text{large}}(x_4) = \exp\left(i \frac{\tau_3}{2} \frac{2\pi x_4}{\beta}\right). \quad (16)$$

As it is known the ‘‘large’’ gauge transformation should be treated differently from ‘‘small’’ (proper) gauge transformations because any two field configurations related by ‘‘large’’ gauge transformation do not belong to the same gauge orbit. Nevertheless, the transformation (16) preserves

⁵Our convention for normalization is $|\Phi|^2 \equiv 2\text{Tr}(\Phi)^2 = 2\text{Tr}[A_4(r \rightarrow \infty)]^2 = v^2$.

the periodic boundary conditions because $U_{\text{large}}(x_4=0) = -U_{\text{large}}(x_4=\beta)$. The final step is to perform the reflection $v \rightarrow -v$ in order to restore the original vacuum expectation value (13). It is implemented by the discrete transformation $U_{\text{reflection}} = \exp(i\tau^2\pi/2)$. The resulting configuration is the L monopole (the Kaluza-Klein monopole in the original terminology). Its asymptotic behavior is

$$A_4^L(r \rightarrow \infty) \rightarrow \left(v + \frac{1}{r}\right) \frac{\tau^3}{2}, \quad (17)$$

which should be contrasted with (14) with an opposite sign for a Coulomb term. It corresponds to the opposite magnetic charges of the M and L monopoles. Therefore, the action S , topological charge Q , and magnetic charge q for M and L monopoles are

$$\begin{aligned} S^M &= \frac{8\pi^2}{g^2} \nu, & Q^M &= \nu, & q^M &= +1, & \nu &\equiv \frac{\beta v}{2\pi}, \\ S^L &= \frac{8\pi^2}{g^2} (1-\nu), & Q^L &= (1-\nu), & q^L &= -1, \end{aligned} \quad (18)$$

while the monopole’s mass m is determined as $S = m\beta$ such that $m = \frac{4\pi}{g} v$ when m is expressed in terms of v . One can explicitly see from (18) that the classical action $S = (S^M + S^L)$ for the KvBLL configuration consisting L and M monopoles does not depend on v and coincides with action of the conventional periodic instanton [24,25] with $Q^M + Q^L = 1$ and action $S = 8\pi^2/g^2$. On the quantum level the partition function, of course, will depend on v . We review the relevant results from Refs. [28–34] on this matter below in Sec. II D. Specifically, we want to pinpoint few crucial elements which differ between Euclidean expressions and corresponding formulas written in hyperbolic space. Precisely this difference as we shall argue in Sec. III is responsible for the linear in κ correction in expression (2) describing the ‘‘strange’’ energy.

D. The grand canonical ensemble of monopoles with nontrivial holonomy (11)

In the semiclassical approximation the partition function of the gluodynamics is represented by the statistical ensemble of an arbitrary number of interacting monopoles and antimonopoles of all kinds. The corresponding picture is well tested in the weakly coupled gauge theory, the so-called ‘‘deformed QCD’’ where all important elements such as the generation of the strange energy, the topological susceptibility, the contact term, etc., have been explicitly computed [21–23]. The key lesson from those studies is the crucial role of the holonomy (11) and nonlocality in generation of all these effects. In fact one can argue [23]

that the system belongs to a topologically ordered phase as a result of these nonlocal effects.⁶

Similar computations in strongly coupled regime have not been performed yet. Nevertheless, one should expect a very similar behavior of the strongly coupled ensemble of KvBLL calorons represented by a set of their monopole constituents as discussed above in Sec. II C. In other words, we assume that KvBLL calorons with nontrivial holonomy are responsible for generation of the strange energy, the topological susceptibility, the contact term, and many other highly nontrivial features of the system. One should emphasize that the corresponding contributions are finite in the infrared (IR) in the large volume limit and small temperature, $(V, \beta) \rightarrow \infty$, in contrast with conventional instanton computations. We do not claim that the semiclassical approximation adopted here is justified in the strongly coupled regime. In fact, it is expected that the fluctuations with typical scales $\sim \Lambda_{\overline{\text{QCD}}}$ change some numerical estimates. However, as we shall argue below the corresponding fluctuations with scales $\sim \Lambda_{\overline{\text{QCD}}}$ can not modify the contributions from the far IR regions with typical scales $\sim \kappa$, which is precisely the subject of the present work.

Therefore, we follow [28–34] and represent the grand partition function for the ensemble as follows

$$\begin{aligned} \mathcal{Z} = & \sum_{K_L K_M K_{\bar{L}} K_{\bar{M}}} \prod_{i_L=1}^{K_L} \prod_{i_M=1}^{K_M} \prod_{i_{\bar{L}}=1}^{K_{\bar{L}}} \prod_{i_{\bar{M}}=1}^{K_{\bar{M}}} \\ & \times \int \frac{f d^3 x_{i_L}}{K_L!} \frac{f d^3 x_{i_M}}{K_M!} \frac{f d^3 y_{i_{\bar{L}}}}{K_{\bar{L}}!} \frac{f d^3 y_{i_{\bar{M}}}}{K_{\bar{M}}!} e^{-V(x-y)} \\ & \cdot \det G[x] \cdot \det G[y], \end{aligned} \quad (19)$$

where f is the “fugacity” of the monopoles. The fugacity has been computed in [28] in terms of the fundamental parameters of the system such as $\Lambda_{\overline{\text{QCD}}}$, β , and holonomy v defined by Eq. (13). One should emphasize that fugacity f is dimensional parameter which is sensitive to all scales of the problem as it is expressed in terms of zero as well as nonzero modes, see below. The 3 dimensional coordinates x_{i_M}, x_{i_L} and $y_{i_{\bar{L}}}, y_{i_{\bar{M}}}$ describe the positions of M, L monopoles and \bar{L}, \bar{M} antimonopoles correspondingly. The $G[x]$ is a $(K_L + K_M) \times (K_L + K_M)$ and $G[y]$ is a $(K_{\bar{L}} + K_{\bar{M}}) \times (K_{\bar{L}} + K_{\bar{M}})$ matrices describing the moduli space. Their explicit forms are given in Refs. [28,29]. These matrices represent the standard zero mode contributions, and highly

sensitive to the IR physics as they depend on the holonomy and the long range Coulomb interactions between the monopoles. The corresponding interactions can be traced from the asymptotical behavior of the monopole’s solutions (14) and (17). Finally, potential $V(x - y)$ describes the interaction between monopoles and antimonopoles of the entire ensemble. The corresponding interactions, along with $G[x]$ and $G[y]$ are also long ranged. These elements of the partition function are highly sensitive to the IR behavior and to the boundary conditions.

There are many subtle points in writing (19) which shall not be discussed here. We refer to the original works for the discussions and references. For the moment we ignore the interaction between the monopoles and antimonopoles, $V(x - y)$. As we argue below in Sec. III C, the corresponding interaction may change the numerical results, but cannot modify our main claim on the structure of the correction $\sim \kappa$.

The only relevant element for our future studies is the presence of the long range forces entering $G[x], G[y]$. The corresponding interactions are effectively canceled in computation of the free energy due to the total neutrality condition,⁷ as argued in [28,29]. As a result of neutrality the number of different types of monopoles is the same in each given configuration. In other words, the partition function under these assumptions decouples for monopoles and antimonopoles, $\mathcal{Z} = \mathcal{Z}_+ \cdot \mathcal{Z}_-$, and takes the following simple form [28,29]:

$$\mathcal{Z}_{\pm} = \sum_{K_L K_M} \frac{(4\pi f V)^{K_L + K_M}}{K_L! K_M!} \nu^{K_M} (1 - \nu)^{K_L}, \quad (20)$$

where V is the 3-volume of the system. The combination $\nu^{K_M} (1 - \nu)^{K_L}$ which enters the partition function (20) comes from a zero mode determinant which itself is expressed in terms of the classical actions of the constituents (18). In the large volume limit the sum is saturated by very large K such that the partition function can be evaluated using saddle point approximation.

To proceed with estimations we first represent $k!$ using the Stirling formula

$$\frac{1}{k!} = e^{-\ln k!} \simeq \frac{1}{\sqrt{2\pi k}} e^{-K \ln K + K}. \quad (21)$$

The next step is to replace the sum by the integrals

⁶Unfortunately we cannot use the “deformed QCD” model to address the question formulated in the present work. This is because we cannot implement κ parameter into this model because the size of \mathbb{S}^1 in “deformed QCD” model [21] must be small to keep the system in the weakly coupled regime, while in hyperbolic space the radius of $\mathbb{S}^1_{1/\kappa}$ must be large as it is correlated with our large 3d volume \mathbb{H}^3_{κ} .

⁷The consequences on neutrality condition has been slightly corrected recently in [32–34]. The correct statement is not that the neutrality condition implies that the total charge is zero for each given configuration. Rather, the correct statement is that the corresponding expectation value of the charge vanishes while charge number density itself may still fluctuate.

$$\sum_{K_L K_M} \rightarrow \int dK \int dQ, \quad K \equiv K_L + K_M, \quad Q \equiv K_M - K_L, \quad (22)$$

where K describes the total number of monopoles in a given configuration, while Q describes the magnetic charge of a given configuration as M and L monopoles have opposite charges according to (18). Using saddle point approximation one arrives to the following expression for the partition function in terms of the saddle value K_0 which saturates (20)

$$\mathcal{Z}_{\pm} \simeq e^{K_0} \cdot \int \frac{e^{-\frac{Q^2}{2K_0}} dQ}{\sqrt{2\pi K_0}}, \quad K_0 = 8\pi f V \sqrt{\nu(1-\nu)}. \quad (23)$$

Therefore the final expression for the partition function and the free energy for the vacuum ground state in this approximation assumes the form

$$\mathcal{Z}_{\pm} \simeq \exp[4\pi f V], \quad f = \frac{4\pi \Lambda_{\text{QCD}}^4}{g^4 T} \\ F_{\text{vac}} = -T \ln \mathcal{Z} = -\frac{32\pi^2}{g^4} \Lambda_{\text{QCD}}^4 V, \quad (24)$$

where we substitute the ‘‘confining’’ value for the holonomy $\nu = 1/2$ which minimizes the free energy.

Few comments are in order. The expectation value $\langle Q \rangle = 0$ obviously vanishes, such that system is neutral. However, the fluctuations of the Q^2 do not vanish, in agreement with [32,33], but strongly suppressed for a large volume system, as expected

$$\sqrt{\frac{\langle Q^2 \rangle}{K_0}} \simeq \sqrt{\frac{1}{K_0}} \sim \frac{1}{\sqrt{V}} \rightarrow 0. \quad (25)$$

The free energy (24) is finite at zero temperature limit. Furthermore, $\ln \mathcal{Z}$ is proportional to the 4-volume $\sim V/T$ of the system demonstrating the expected extensive scaling at low temperature. One should emphasize that these computations (based on configurations with nontrivial holonomy) generate the IR finite and well-defined contributions to different observables expressed in terms of fundamental parameters of the theory, in contrast, for example, with instanton computations. The dynamically generated ‘‘confining’’ value for the holonomy $\nu = 1/2$ is also a highly nontrivial phenomenon—it leads to a proper behavior for the Polyakov’s loop, the Wilson loop and the string tension [28,29], see also Appendix A with few historical and terminological comments on fractionally charged constituents, previously emerged in the literature in different contexts and different systems.

Furthermore, this ‘‘confining’’ value for the holonomy $\nu = 1/2$ leads to a consistent resolution of the so-called

$U(1)_A$ problem formulated in terms of the topological susceptibility (8) and the θ dependence of the ‘‘strange’’ vacuum energy (6). Indeed, the introduction of the θ term into the Lagrangian changes the fugacity for the monopole $f \rightarrow f e^{i\theta/2}$ and antimonopoles $f \rightarrow f e^{-i\theta/2}$. This modification follows from the fact that the topological charges for monopoles and antimonopoles assume the magnitude $Q = \pm 1/2$ for ‘‘confining’’ value of the holonomy $\nu = 1/2$ as it follows from quantum numbers for monopoles (18). The antimonopoles assume the opposite sign for the topological charge as they are antiself-dual solutions. This modification leads to replacement of expression (24) by the following formula which is valid for $|\theta| \leq \pi$:

$$F_{\text{vac}}(\theta) = -\frac{32\pi^2}{g^4} \Lambda_{\text{QCD}}^4 V \cdot \cos\left(\frac{\theta}{2}\right). \quad (26)$$

The topological susceptibility now can be easily computed by differentiating (26) twice with respect to θ with the result

$$\chi = \frac{1}{V} \frac{\partial^2 F_{\text{vac}}(\theta)}{\partial^2 \theta} \Big|_{\theta=0} = \frac{8\pi^2}{g^4} \Lambda_{\text{QCD}}^4. \quad (27)$$

Finally the vacuum energy (26) per unit volume $F_{\text{vac}}/V \sim \Lambda_{\text{QCD}}^4$ is precisely the first term entering the expression (2). It has all the features of the ‘‘strange energy’’ briefly described in Sec. II A in a model-independent generic way. The mechanism based on the KvBLL configurations reviewed above precisely generates all these required properties. Similar formulas can be easily generalized for arbitrary number of colors N when $F_{\text{vac}}(\theta) \sim N^2 \cos(\frac{\theta}{N})$ and $\chi \sim 1$, which is consistent with conventional resolution of the $U(1)_A$ problem in large N limit.

We do not claim to have derived any new results in the present subsection. Rather, we just reproduced and explained the known results [28,29,32,33] in a slightly different and simplified manner in order to analyze the role of similar vacuum configurations in the cosmological context next in Sec. III. Furthermore, we do not claim that the corresponding computations in strongly coupled regime are exact. In fact, we expect the corrections to be order of one to the fugacity f and all other numerical coefficients such as (24), (27) discussed above. However, we do not anticipate any drastic qualitative changes of this framework as a result of these possible corrections. In particular, we expect that the free energy generated by these configurations remains finite in the IR and demonstrates the extensive behavior at low temperature $T \rightarrow 0$ as presented above. Precisely these features will play a crucial role in our arguments on small modification of this ‘‘strange’’ vacuum energy with tiny variation of the background to be considered in Sec. III D.

E. Interpretation. Cosmological context

There are many important elements related to the computations reviewed in previous sections. In what follows we would like to make only a very few comments which will be relevant for our studies in cosmological context on IR sensitivity of the system.

- (1) First of all, the positive sign in (27) unambiguously implies that the corresponding configurations saturating the topological susceptibility (and related the θ -dependent portion of the vacuum energy) cannot be identified with any propagating degrees of freedom in accordance with (7). Indeed, the computations reviewed above explicitly show that the relevant configurations are the KvBLL calorons with nontrivial holonomy describing the tunneling events between topologically distinct but physically identical winding states, rather than propagating gluons. All effects are obviously nonanalytical in coupling constant $\sim \exp(-1/g^2)$ and cannot be seen in perturbation theory.
- (2) One can view the relevant topological configurations as the 3d magnetic monopoles wrapping around time direction. This leads to the nonvanishing holonomy (11) and nonvanishing topological charge (18) of the constituents defined in 4d space-time. In the limit of $T \rightarrow 0$ the “confining” value $\nu = 1/2$ for the holonomy implies that the parameter ν which determines the monopole’s mass in conventional 3d theory tends to zero in this limit, $\nu = \pi T \rightarrow 0$. At the same time the monopole generates a finite contribution (24) to the path integral due to its finite 4d action (18) resulting from very long path $\sim T^{-1}$ in time direction. It would be misleading to interpret the confinement and other features (discussed above in Sec. II D) generated by these configurations as the condensation of the monopoles. It would be more appropriate to use term “percolation” as the configurations described above do correlate at arbitrary large distances, but they obviously do not form a “condensate” in conventional condensed matter terminology.
- (3) In the cosmological context such configurations are highly unusual objects: they obviously describe the nonlocal physics as the holonomy (11) is a nonlocal object. Indeed, the holonomy defines the dynamics along the entire history of evolution of the system in the given confined phase: from the very beginning to the very end. There are no contradictions with causality in the system as there are not any physical degrees of freedom to propagate along this path at $\beta \rightarrow \infty$, see item (i). above. Indeed, this entire gauge configuration is a mere saddle point in Euclidean (imaginary time) path integral computation which describes the instantaneous tunneling event, rather than propagation of a physical degree of freedom capable of carrying an information/signal.
- (4) Further to this point, the extensive property of free energy $\beta F_{\text{vac}} \sim V/T$ at $\beta \rightarrow \infty$ is a highly nontrivial phenomenon in this framework as the 4-volume V/T appears in this description due to few important steps. First, one should regularize the moduli space by cutting off each given configuration in the IR. Second, one should sum over all configurations (20) by using saddle point approximation in the large volume limit, which eventually leads to (24). This “emergent” extensive property is drastically different from conventional approaches to cosmology when the free energy is determined by the Lagrangian density $L[\Phi]$ integrated over the 4-volume $\int d^4x$. In this last case the extensive property is a trivial manifestation of the system formulated from the very beginning in terms of the local field $\Phi(x)$. It shows one more time that generation of the “strange” energy (2) is a highly nonlocal nonperturbative effect when the volume of the system could be very large, but still finite to proceed with computations (24) in this approach. Essentially, the finite local energy density of the system (24) in this framework is determined by the entire time evolution $\beta \rightarrow \infty$ in confined phase, which is obviously a nonlocal procedure. Still, it does not contradict the causality, see item (iii) above.
- (5) Last but not least, all these highly nontrivial, nonlocal features listed above emerge only at $T < T_c$ when the configurations with nontrivial holonomy (11) start to play the dominant role in the dynamics. At high temperatures the contribution of the configurations with nontrivial holonomy can be completely ignored as they do not contribute to the partition function in thermodynamical limit. This property of drastic variation of “strange” energy (2) with temperature around T_c may play an important role in cosmological context as we discuss in the next section.

III. NONTRIVIAL HOLONOMY AND HYPERBOLIC SPACE $\mathbb{H}_\kappa^3 \times \mathbb{S}_{\kappa^{-1}}^1$

The main goal of this section is to generalize the results presented above in Secs. II C, II D to hyperbolic space $\mathbb{H}_\kappa^3 \times \mathbb{S}_{\kappa^{-1}}^1$ to argue that the correction to the free energy (24) are linearly proportional to κ at small $\kappa \rightarrow 0$. In this limit it is quite obvious that all features listed in Sec. II E on nature of the “strange energy,” including its nonlocal nature, remain the same in this limit $\kappa \rightarrow 0$ as the system is almost 4d Euclidean space, with very tiny deviations $\sim \kappa$ which we wish to recover. The corresponding linear dependence on κ would strongly support our conjecture that the corrections in Eq. (2) are linearly proportional to the Hubble constant,⁸

⁸See footnote 3 with some clarification on terminology.

which dynamically drags our Universe to the de Sitter state as Friedman equation (3) suggests.

We discuss the relevant gauge configurations in hyperbolic space in Sec. III A, while the grand canonical ensemble of such hyperbolic monopoles will be studied in Sec. III B where we discuss the crucial distinctions between hyperbolic and Euclidean monopoles. Precisely this difference eventually leads to a tiny $\sim\kappa$ variation from the Euclidean results. The corresponding deviation is expressed in terms of the fugacity in Sec. III C, which ultimately leads to slight modification of the vacuum energy. We list some profound cosmological consequences of this modification of the vacuum energy with background in Sec. III D.

A. Holonomy and monopoles in hyperbolic space

The construction of the monopoles on hyperbolic space \mathbb{H}_κ^3 has been known since [35–37]. Furthermore, many other topological objects, including calorons, instantons, vortices, skyrmions have been constructed on hyperbolic space [38–43]. An important technical element which was used in these constructions is the conformal equivalence of \mathbb{R}^4 and $\mathbb{H}_\kappa^3 \times \mathbb{S}_{\kappa^{-1}}^1$. Indeed, this equivalence can be explicitly checked by the introduction of toroidal coordinates $(\rho, \theta, \phi, \chi)$ on \mathbb{R}^4 as follows:

$$\begin{aligned} x_\mu &= (x_1, x_2, x_3, x_4) \\ x_1 &= \frac{1}{\cosh(\kappa\rho) + \cos\chi} \sinh(\kappa\rho) \sin\theta \cos\phi, \\ x_2 &= \frac{1}{\cosh(\kappa\rho) + \cos\chi} \sinh(\kappa\rho) \sin\theta \sin\phi, \\ x_3 &= \frac{1}{\cosh(\kappa\rho) + \cos\chi} \sinh(\kappa\rho) \cos\theta, \\ x_4 &= \frac{1}{\cosh(\kappa\rho) + \cos\chi} \sin\chi. \end{aligned} \quad (28)$$

It is then easy to check that the metric on \mathbb{R}^4 becomes

$$ds^2(\mathbb{R}^4) \equiv dx_\mu dx^\mu = \frac{ds^2(\mathbb{H}_\kappa^3) + \kappa^{-2} d\chi^2}{(\cosh(\kappa\rho) + \cos\chi)^2}, \quad (29)$$

where $ds^2(\mathbb{H}_\kappa^3)$ is the metric on hyperbolic 3-space with spherical coordinates (ρ, θ, ϕ) :

$$ds^2(\mathbb{H}_\kappa^3) = d\rho^2 + \frac{\sinh^2(\kappa\rho)}{\kappa^2} (d\theta^2 + \sin^2\theta d\phi^2). \quad (30)$$

The holonomy in terms of these variables is computed along the circle $\mathbb{S}_{\kappa^{-1}}^1$ parametrized by χ , that is

$$U = \mathcal{P} \exp\left(\frac{i}{\kappa} \int_0^{2\pi} d\chi A_\chi\right), \quad (31)$$

where A_χ is the component of the gauge potential associated with coordinate χ , and it plays the same role as the A_4 , similar to the construction in Euclidean space (12). In both cases, the A_4 and A_χ assume nonvanishing expectation values, and play the same role as the Higgs field Φ in adjoint representation, as explained after Eq. (13), see also footnote 5 on our normalization. Formula (31) plays the same role as Eqs. (5), (11), while $d\chi/\kappa$ in Eq. (31) plays the role of dx_4 in evaluation of the holonomy computed along the circle \mathbb{S}^1 according to (13). At large $\rho \rightarrow \infty$ the A_χ approaches a nonvanishing constant value, similar to parameter v in Euclidean space (12).

With these remarks in mind, the explicit form for the BPS monopole in hyperbolic space \mathbb{H}_κ^3 in the unitary gauge can be written as follows [36,37]:

$$A_\chi^M(\rho) = [C\kappa \coth(C\kappa\rho) - \kappa \coth(\kappa\rho)] \frac{\tau^3}{2}, \quad (32)$$

where parameter C takes any value greater than 1. In formula (32) we limited ourselves by writing down only the $A_\chi^M(\rho)$ -component, which defines the boundary conditions at large ρ . We coin this solution as M monopoles in order to be consistent with the terminology introduced for the Euclidean counterparts (12).

While many topological objects, including calorons with trivial holonomy [41] have been constructed in hyperbolic space, as we already mentioned, an explicit construction of the calorons with nontrivial holonomy, which is analogous to the KvBLL solutions [26,27], has not been constructed yet. In what follows we shall assume that such solutions do exist, though we do not need their explicit form in our future discussions. An important point is that if such configurations exist than they must exhibit the same features which the KvBLL solutions demonstrate. Namely they could be viewed as a set of 2 different types of monopoles for $SU(2)$ gauge group. The first type is precisely the M -monopole (32) discussed above, while the second one, the L monopole can be constructed as described in Sec. II C for the Euclidean counterpart. We shall return to this construction later in the text, but first, we want to understand the physical meaning of the parameter C entering Eq. (32) by analyzing the limit $\kappa \rightarrow 0$ when the Euclidean monopole (12) is recovered.

To recover the Euclidean monopole solution one should take the limit $C \rightarrow \infty$ along with $\kappa \rightarrow 0$ with combination $C\kappa$ being fixed to be equal v . In this limit $A_\chi^M(\rho)$ becomes

$$A_\chi^M(\rho) = \left[v \coth(v\rho) - \frac{1}{\rho} \right] \frac{\tau^3}{2}, \quad C\kappa \equiv v, \quad (33)$$

where ρ should be identified with r in the Euclidean space. Expression (33) exactly coincides with (12) and its asymptotic behavior (14) in the unitary gauge.

We are now in a position to discuss the asymptotical behavior of (32) and the holonomy (31) for finite κ . Taking

$\rho \rightarrow \infty$ one arrives to the following expressions for $A_\chi^M(\rho)$ and holonomy

$$A_\chi^M(\rho \rightarrow \infty) \rightarrow [\kappa(C-1) + \mathcal{O}(e^{-\rho})] \frac{\tau^3}{2},$$

$$\frac{1}{2} \text{Tr} U(\rho \rightarrow \infty) = \cos \pi \nu, \quad \nu \equiv (C-1), \quad (34)$$

where we introduced parameter ν expressed in terms of original parameter C . It plays the same role as parameter ν discussed in the Euclidean construction (5), (18). The crucial observation here is that the asymptotic formula for $A_\chi^M(\rho)$ does not contain a long range Coulomb interaction ρ^{-1} in contrast with its Euclidean counterpart (14). Instead, there is an exponentially suppressed correction $\sim \exp(-\kappa\rho)$ in formula (34). One can interpret such drastic changes in the behavior of the solution as a result of screening of the magnetic field by the curvature in the hyperbolic space. In terms of the new parameter ν the solution (32) assumes the form

$$A_\chi^M(\rho) = ((\nu+1) \coth [(\nu+1)\kappa\rho] - \coth \kappa\rho) \frac{\kappa\tau^3}{2}. \quad (35)$$

Our next step is to recover the L monopole assuming that KvBLL caloron with nontrivial holonomy in the hyperbolic space exists, similar to construction [26,27] in the Euclidean space. We follow the same steps to reconstruct their properties presented in Sec. II C for the Euclidean monopoles. The first step is to replace $\nu \rightarrow (1-\nu)$. The second step is to make a “large” gauge transformation which assumes the following form in hyperbolic variables

$$U_{\text{large}}(\chi) = \exp\left(i \frac{\tau_3}{2} \chi\right). \quad (36)$$

As we already mentioned the “large” gauge transformations should be treated differently from “small” (proper) gauge transformations because any two field configurations related by “large” gauge transformation do not belong to the same gauge orbit. Nevertheless, the transformation (36) preserves the periodic boundary conditions for the fields in the adjoint representation because $U_{\text{large}}(\chi=0) = -U_{\text{large}}(\chi=2\pi)$. The final step is to perform the discrete transformation $U_{\text{reflection}} = \exp(i\tau^2\pi/2)$ to restore the original boundary conditions (34). The resulting configuration is the hyperbolic L - monopole:

$$A_\chi^L(\rho) = (1 - (\bar{\nu}+1) \coth [(\bar{\nu}+1)\kappa\rho] + \coth \kappa\rho) \frac{\kappa\tau^3}{2}, \quad (37)$$

where we introduced $\bar{\nu} \equiv (1-\nu)$ for convenience. One should emphasize that electric and magnetic fields of the L -monopoles do depend on a χ variable as a result of χ

dependent “large” gauge transformation (36). This is analogous to the L -monopole solution in Euclidean space (17) which is a time dependent configuration, and ceases to exist in static 3d space.

The classical action, topological and magnetic charges of the M and L constituents are determined by the boundary conditions (35) and (37) at large $\rho \rightarrow \infty$, similar to the Euclidean counterparts (18). The corresponding parameters S, Q, q obviously assume the same values (18) when expressed in terms of ν . These dimensionless parameters obviously cannot depend on dimensional parameter κ , including $\kappa \rightarrow 0$ limit. In fact, the corresponding formula relating Q and q for BPS M -type monopole (35), which identically coincides with the Euclidean expression, was derived in hyperbolic space in the original work [35], while the relation between S and Q is a direct consequence of self-duality of Yang-Mills equations. The monopole’s mass (the total energy of the configuration), being a dimensional parameter, does depend on κ . Mass m satisfies an obvious relation $S = m \frac{2\pi}{\kappa}$, similar to its Euclidean counterpart, see text after Eq. (18). Explicitly, in terms of the original parameter C it is given by $m = \frac{4\pi}{g} \kappa(C-1)$, which reduces to its Euclidean form $m = \frac{4\pi}{g} v$ when one takes the corresponding limit (33).

B. Grand partition function for monopoles in the hyperbolic space

With our main assumption that the calorons with nontrivial holonomy exist in hyperbolic space, similar to Euclidean KvBLL construction [26,27], one should expect that the corresponding grand partition function \tilde{Z} has the following form, which is analogous to expression (19) discussed in a previous section:

$$\tilde{Z} = \sum_{K_L K_M K_{\bar{L}} K_{\bar{M}}} \prod_{i_L=1}^{K_L} \prod_{i_M=1}^{K_M} \prod_{i_{\bar{L}}=1}^{K_{\bar{L}}} \prod_{i_{\bar{M}}=1}^{K_{\bar{M}}} \int_{\mathbb{H}_\kappa^3} \frac{\bar{f} \sqrt{g} d^3 \bar{x}_{i_L} \bar{f}}{K_L!} \frac{\bar{f} \sqrt{g} d^3 \bar{x}_{i_M} \bar{f}}{K_M!} \frac{\bar{f} \sqrt{g} d^3 \bar{y}_{i_{\bar{L}}} \bar{f}}{K_{\bar{L}}!} \frac{\bar{f} \sqrt{g} d^3 \bar{y}_{i_{\bar{M}}} \bar{f}}{K_{\bar{M}}!} \times e^{-V(\bar{x}-\bar{y})} \cdot \det G[\bar{x}] \cdot \det G[\bar{y}], \quad (38)$$

where \bar{f} is the “fugacity” of the hyperbolic monopoles. The corresponding dimensional parameter is highly sensitive to many details of monopole’s structure and their interactions with other monopoles. It is obviously different from its cousin fugacity f computed in the Euclidean space as discussed in Sec. II D. We shall estimate the corrections to \bar{f} later in Sec. III C where we argue that the difference $(\bar{f} - f) \sim \kappa$ is linear in κ at small $\kappa \rightarrow 0$ for configurations with nontrivial holonomy.

The next item to discuss from formula (38) is the 3-dimensional coordinates $\bar{x}_{i_M}, \bar{x}_{i_L}$ and $\bar{y}_{i_{\bar{L}}}, \bar{y}_{i_{\bar{M}}}$. They describe the positions of M, L monopoles and \bar{L}, \bar{M} antimonopoles

correspondingly. These coordinates play the same role as in formula (19) where the only difference is that the distance between constituents is computed using the metric (30). The $G[\bar{x}]$ is a $(K_L + K_M) \times (K_L + K_M)$ and $G[\bar{y}]$ is a $(K_{\bar{L}} + K_{\bar{M}}) \times (K_{\bar{L}} + K_{\bar{M}})$ matrices describing the moduli space. The difference with corresponding Euclidean expressions is that the behavior at large distances is not the Coulomb like, but rather the exponentially suppressed as we already mentioned (34). This is because the matrices $G[\bar{x}]$ and $G[\bar{y}]$ are computed from the corresponding zero modes in the background of the monopoles. At the same time, the zero modes, as usual, are fixed by the corresponding classical solutions. Therefore, the asymptotical behavior of classical solutions (35) and (37) dictates the behavior of $G[\bar{x}]$ and $G[\bar{y}]$. In both cases the computations of the free energy is reduced to expression (20) as a result of the neutrality condition discussed in Sec. II D.

The final item to discuss is the factor \sqrt{g} which accounts for the curvature of the hyperbolic space \mathbb{H}_κ^3 . It enters along with the spatial volume of the system where computations are being performed. We put the system into a large volume of radius R such that the volume is

$$V = \frac{4\pi}{2\kappa^3} \left(\frac{\sinh 2\kappa R}{2} - \kappa R \right), \quad \sqrt{g} = \frac{\sin \theta \sinh^2 \kappa \rho}{\kappa^2}. \quad (39)$$

It reduces to the Euclidean expression $V = \frac{4\pi R^3}{3}$ in the limit of small κ with the corrections of order $\mathcal{O}(\kappa^2)$ which are consistently neglected in the present work. Therefore, by repeating all the steps leading to formula (24) we arrive to the following expression for the partition function and the free energy in hyperbolic space

$$\bar{Z}_\pm \simeq \exp[4\pi \bar{f} V], \quad \bar{F}_{\text{vac}} = -\frac{\kappa}{2\pi} \ln \bar{Z}, \quad (40)$$

where we identify \mathbb{S}^1 from $\mathbb{R}^3 \times \mathbb{S}^1$ geometry (reviewed in Sec. II) with circle $\mathbb{S}_{\kappa^{-1}}^1$ from $\mathbb{H}_\kappa^3 \times \mathbb{S}_{\kappa^{-1}}^1$ geometry presented in Sec. III A. In other words, we identify

$$T \equiv \frac{1}{\beta} = \frac{\kappa}{2\pi}, \quad (41)$$

such that the corresponding expressions for holonomy (5) and (31) coincide. With this identification the corresponding formulas (15) and (34) in terms of dimensionless parameter ν also coincide. In formula (40) we substitute the confining value for holonomy $\nu = 1/2$ as it has been done in the Euclidean space. This is because the free energy is minimized at $\nu = 1/2$ irrespectively to the value of the fugacity, which is indeed different for two different geometries. To conclude: the only difference between \bar{Z} and Z describing the system on $\mathbb{H}_\kappa^3 \times \mathbb{S}_{\kappa^{-1}}^1$ and $\mathbb{R}^3 \times \mathbb{S}^1$ geometries correspondingly is that the fugacities in these

two systems assume slightly different values at small $\kappa \rightarrow 0$, which is the subject of the next subsection.

C. Monopole's fugacity

We start our analysis with explanations on how the basic dimensional parameter, the fugacity, emerges in the system. This x independent dimensional parameter f effectively determines the dynamics of the system. This parameter essentially represents the density of the monopoles in the system. The classical action, the zero and nonzero mode contributions lead to the following expression for the monopole's fugacity f in terms of the fundamental parameters of the theory [28,29]:

$$f^2 = \left[\frac{4\pi\beta\Lambda_{\text{QCD}}^4}{g^4} \right]^2 \cdot c$$

$$c = \left\langle \frac{[1 + 2\pi\nu\bar{\nu}\frac{r_{12}}{\beta}]}{(\Lambda_{\text{QCD}}r_{12})^{2/3}} \left[1 + 2\pi\nu\frac{r_{12}}{\beta} \right]^{\frac{8}{3}\nu-1} \left[1 + 2\pi\bar{\nu}\frac{r_{12}}{\beta} \right]^{\frac{8}{3}\bar{\nu}-1} \right\rangle, \quad (42)$$

where brackets $\langle \dots \rangle$ imply averaging over separation r_{12} between M and L monopoles in ensemble (19) such as f is x -independent as it should.

A few comments are in order. Each KvBLL caloron is represented by the L and M monopoles and accompanied by 8 zero modes. It explains the major dimensional factor in Eq. (42), including Λ_{QCD}^8 . The remaining numerical dimensionless factor “ c ” entering (42) is order of 1. It includes factor $(\Lambda_{\text{QCD}})^{-2/3}$ which can be easily restored from the renormalization group analysis which requires that Λ_{QCD} enters (42) with power $(\Lambda_{\text{QCD}})^{22/3}$. Subsequently, this factor $(\Lambda_{\text{QCD}})^{-2/3}$ must be accompanied by a dimensional parameter, which at small temperatures could be nothing else but the separation distance $r_{12}^{-2/3}$ between the monopoles. In estimates of Refs. [29] the numerical coefficient c is assumed to be one, which precisely corresponds to the expression for f given in Eq. (24).

The next factor which can be easily explained is the first term in the numerator, $[1 + 2\pi\nu\bar{\nu}\frac{r_{12}}{\beta}]$. This term has been originally computed in [26] and reproduced in [28,29]. It is originated from the zero mode determinant. The crucial point here is that the algebraic dependence on r_{12} emerges as a result of long range Coulomb terms in the classical solutions (14) and (17). This is because the zero mode structure is unambiguously fixed by the classical solutions with the corresponding Coulomb terms. Another important element is that this term is proportional to the holonomy $\nu\bar{\nu}$. It implies that this term will not be generated for the configurations with trivial holonomy.

The nature of the next two terms in the numerator in Eq. (42) is much harder to explain because they are originated from the contributions of the nonzero modes.

The only comment we would like to make here is that these terms are also proportional to the holonomy, and cannot be generated by the configurations with trivial holonomy.

Now we are in a position to estimate the difference between the fugacity generated by monopoles in the Euclidean space versus hyperbolic monopoles at small $\kappa \rightarrow 0$. The estimation is convenient to represent in the following form

$$\frac{f}{\bar{f}} \simeq (1 + \Delta_{zm}) \cdot (1 + \Delta_{nzm}), \quad \Delta_{zm} \simeq \frac{\nu\bar{\nu}}{2} \frac{\kappa}{\Lambda_{\text{QCD}}}, \quad (43)$$

where factors Δ_{zm} and Δ_{nzm} describe the corrections due to the zero and nonzero modes correspondingly. The correction factor Δ_{zm} comes from the first term in the numerator (42), where we estimate $\langle r_{12} \rangle \sim \Lambda_{\text{QCD}}^{-1}$ as the only dimensional parameter in the system at small temperature. We also expressed the result in terms of κ rather than β to be consistent with identification (41). The crucial point in our estimate Δ_{zm} is that a similar correction in hyperbolic space is absent as the corresponding classical solutions (35) and (37) have exponentially suppressed asymptotic at large distances, in contrast with the Euclidean counterpart. Therefore, a term $\sim r_{12}$ cannot be generated in hyperbolic space, in contrast with the Euclidean case (42).

Unfortunately, a similar unambiguous conclusion cannot be reached regarding the nonzero mode contribution Δ_{nzm} . This is because both the Euclidean as well as hyperbolic monopoles may generate such contributions proportional to $\kappa \cdot r_{12}$, which eventually produce a desired correction $\sim \kappa / \Lambda_{\text{QCD}}$. While the Euclidean expression is known and is represented by two factors in the numerator in Eq. (42), a similar expression in hyperbolic space is not known simply because an explicit construction of the hyperbolic calorons with nontrivial holonomy, similar to KvBLL solution, is yet unknown. Therefore, for the numerical estimates in what follows we set $\Delta_{nzm} = 0$.

A few comment are in order. First, we want to argue that an unknown Δ_{nzm} correction cannot exactly cancel the computed Δ_{zm} term (43). Indeed, the Δ_{zm} structure is precisely fixed by the structure of the $SU(N)$ gauge group with $4N$ zero modes, while Δ_{nzm} varies and depends, in particular, on the presence of the matter fields, and other details of the system. In other words, a possible cancellation, if ever occurs, cannot be a universal phenomenon. Therefore, we use Δ_{zm} and disregard Δ_{nzm} as our order of magnitude estimate for ratio (43). It explicitly exhibits the linear in κ correction to the fugacity, which is the main result of the present work. This correction can be only generated by the configurations with nontrivial holonomy. In particular, conventional instantons and calorons with trivial holonomy may only generate the higher order corrections $\sim \mathcal{O}(\kappa^2)$ and do not contribute to the linear term (43).

We should emphasize that estimate (43) was derived under assumption that the interaction term $V(\bar{x} - \bar{y})$ in formula (38) vanishes. It corresponds to the ensemble containing exclusively the monopoles (or antimonopoles). Only in this case the partition function is exactly reduced to simple form (20) as a result of neutrality condition as argued in [28,29]. In reality the interaction plays a crucial numerical role at finite temperature as shown in [32–34]. However, our main claim is that the linear correction (43) may receive large numerical corrections, but it cannot be exactly canceled as a result of unaccounted interaction term $V(\bar{x} - \bar{y})$ in formula (38). The basic argument behind this claim is the same one as presented above and based on the observation that the interaction $V(\bar{x} - \bar{y})$ is highly sensitive to the matter content of the theory (the number of flavors and its masses in the system), while a Δ_{zm} in formula (43) is not sensitive to these modifications. The interaction $V(\bar{x} - \bar{y})$ may change a numerical coefficient in the estimate $\langle r_{12} \rangle \sim \Lambda_{\text{QCD}}^{-1}$ which enters (43), but cannot completely destroy this term. Therefore, our main claim [that the linear correction (43) will be generated] holds irrespectively to any type of monopole-antimonopole interactions $V(\bar{x} - \bar{y})$.

D. Linear corrections $\sim \kappa$ to the vacuum energy

The result (43) for monopole's fugacity can be translated into the statement on variation of the vacuum energy density in the bulk of space-time with a tiny variation of the background. Indeed, according to (24), (40), and (43) the relevant ratio for the vacuum energies at $\kappa \rightarrow 0$ for two different geometries can be represented as follows

$$\frac{E_{\text{vac}}[\mathbb{H}_\kappa^3 \times \mathbb{S}_{\kappa^{-1}}^1]}{E_{\text{vac}}[\mathbb{R}^3 \times \mathbb{S}^1]} \simeq \frac{\bar{f}}{f} \simeq \left(1 - \frac{\nu\bar{\nu}}{2} \cdot \frac{\kappa}{\Lambda_{\text{QCD}}} \right). \quad (44)$$

The same result can be represented in a more conventional form

$$\begin{aligned} E_{\text{vac}}[\mathbb{H}_\kappa^3 \times \mathbb{S}_{\kappa^{-1}}^1] &\simeq -\Lambda_{\text{QCD}}^4 \left(1 - \frac{\nu\bar{\nu}}{2} \cdot \frac{\kappa}{\Lambda_{\text{QCD}}} \right) \\ &\simeq -\Lambda_{\text{QCD}}^4 + \kappa \cdot \Lambda_{\text{QCD}}^3 \frac{\nu\bar{\nu}}{2}, \end{aligned} \quad (45)$$

where we omitted all irrelevant numerical factors in the expression for the vacuum energy in Euclidean space, but kept the relevant sign minus ($-$) in front, which is a well-known feature of QCD. Our final formula (45) is a precise analog (in a simplified model) for the vacuum energy (2) conjectured for the de Sitter space. As we emphasized in the Introduction, the significance of the linear correction in Eq. (2) is that the Friedman equation (3) unambiguously predicts a nontrivial solution with constant H_0 if the subtraction procedure is adopted as discussed in the Introduction. The constant solution H_0 automatically

corresponds to a desired de Sitter behavior (1), which might be relevant for the early Universe during the inflationary epoch, and in the present epoch for a description of the dark energy. A few comments are in order:

- (1) The difference between two geometries, $\mathbb{H}_\kappa^3 \times \mathbb{S}_{\kappa^{-1}}^1$ and $\mathbb{R}^3 \times \mathbb{S}^1$ when the sizes of $\mathbb{S}_{\kappa^{-1}}^1$ and \mathbb{S}^1 identically coincide according to (41) is the small curvature $\sim \kappa^2$ of the hyperbolic space. According to conventional arguments on locality as discussed in the Introduction it unambiguously suggests that all corrections must be proportional to the even powers κ^{2n} . However, we obviously observe a linear correction (45) in explicit computations.
- (2) This linear correction $\sim \kappa$ is generated by the configurations with nontrivial holonomy, which itself is nonlocal, but gauge invariant operator. Therefore, the standard arguments on locality, reviewed in the Introduction are badly violated by such configurations. One can see from Eq. (45) that the linear correction $\sim \kappa$ is explicitly proportional to the holonomy ν , which is the gauge invariant observable, not reducible to the curvature. In other words, this correction is generated by nonlocal configurations, and cannot be expressed in terms of local curvature $\sim \kappa^2$.
- (3) All effects discussed in the present work are non-analytical in coupling constant $\sim \exp(-1/g^2)$ and cannot be seen in perturbation theory.
- (4) The result (45) is consistent with the previous analysis in a weakly coupled “deformed QCD” model where one can study the sensitivity of the vacuum energy to the very large distances by putting the system into the box of size \mathbb{L} . It turns out [44] that the corrections to the vacuum energy are linear in inverse size $\sim \mathbb{L}^{-1}$. This model is very similar in all respects to the system studied in the present work because the vacuum energy in the “deformed QCD” model is also saturated by the monopoles with nontrivial holonomy. At the same time the conventional instantons with trivial holonomy produce only quadratic corrections $\sim \mathbb{L}^{-2}$ as noticed in [44].
- (5) The generation of the linear correction $\sim \kappa$ is also consistent with computations of Ref. [4] in the “deformed QCD” model where analysis was performed in terms of auxiliary topological nonpropagating field. In those computations the root of the phenomenon is the presence of the nontrivial holonomy and long range monopole’s field which eventually is responsible for generation of the linear correction.
- (6) The linear correction observed in our work is also consistent with the lattice simulations [45] when one studies the dependence of the vacuum energy on the size of the system.

- (7) Our results are also consistent with the lattice simulations [46] when the author studies the rate of particle production in the de Sitter background. The rate turns out to be linearly proportional to the Hubble constant $\sim H$, rather than H^2 . Our comment here is that the rate of the particle production in quantum field theory in general is determined by the imaginary part of the stress tensor, $\text{Im}[T_\mu^\nu]$, while the vacuum energy is related to the real part of the stress tensor, $\text{Re}[T_\mu^\nu]$. Analyticity suggests that both components must have the same corrections on H at small H . Therefore, the lattice measurements [46] of the linear dependence on H strongly suggest that the vacuum energy (which is determined by the real part of the same stress tensor) must also exhibit the same linear $\sim H$ correction. The corresponding lattice computations of the θ dependent portion of the vacuum energy and topological susceptibility in a time dependent background are possible in principle, but are technically much more involved than the analysis performed in Ref. [46].
- (8) Last but not least, the sign for the difference $\Delta E_{\text{vac}} \equiv E[\mathbb{H}_\kappa^3 \times \mathbb{S}_{\kappa^{-1}}^1] - E[g_{\mu\nu} = \delta_{\mu\nu}]$ is positive as one can see from Eq. (45). It corresponds to the positive sign for the cosmological constant (dark energy) in cosmological context.

There is a fundamental difference in signs with the conventional Casimir effect when the corresponding subtraction procedure typically leads to the negative, rather than positive, sign for the vacuum energy. This difference is due to the fact that the conventional Casimir vacuum energy is generated by the fluctuations of the physical propagating photons. It is drastically different from the vacuum energy computed in the present work when it is generated by the tunneling transitions between different topological sectors. As we explained in Sec. II A the corresponding vacuum energy cannot be expressed in terms of any propagating degrees of freedom as it has pure nondispersive nature. This is precisely the origin for the positive sign of the vacuum energy: $\Delta E_{\text{vac}} > 0$.

IV. CONCLUSION

The formal result of the present work can be expressed by Eqs. (44), (45), and we shall not repeat the comments listed in the last Sec III D explaining some important consequences of this result. If the same effect persists in a FLRW universe (2), which we expect to be the case, it may have a number of profound consequences for understanding of the past, present, and future evolution of our Universe.

First of all, the nontrivial holonomy (5) implies the presence of \mathbb{S}^1 as a part of our space-time of our Universe. It is an additional invariant characteristic of the manifold which cannot be reduced to the local curvature. In

construction discussed in the present work the corresponding \mathbb{S}^1 is identified with Euclidean time direction. This is because the original intention in early works on the subject (where the corresponding topological vacuum configurations were invented) was to analyze the temperature dependence of the QCD phase transition. In principle, similar \mathbb{S}^1 could also be a part of spatial coordinates. Such an assumption is definitely consistent with all known observations if the size of the \mathbb{S}^1 is sufficiently large $\gtrsim H^{-1}$ at present epoch, see [47] for the estimates in the given context. Furthermore, the linear correction enters formula (45) in the form of its absolute value $|\kappa|$ as it essentially describes the linear (positively defined) size of the corresponding manifold. In the context of the FLRW universe, a similar statement implies that the linear correction in H enters formula (2) in the form of its absolute value $|H|$, see footnote 3 for clarification. Therefore, it cannot lead to any T-violating effects, which one could suspect as $H = \dot{a}/a$ indeed is a T-odd parameter. Still, it generates the de Sitter behavior (1) as a result of this linear scaling.

Another profound consequence of this framework is as follows. The conventional scenarios of the eternal self-producing inflationary universes are always formulated in terms of a physical scalar dynamical inflaton field $\Phi(x)$. This problem with self-reproduction of the universe does not even emerge in our framework as there are not any fundamental scalar dynamical fields in the system responsible for inflation. Instead, the de Sitter behavior (1) in our framework is a pure quantum phenomenon, which is a consequence of the dynamics of the long ranged topological configurations with nontrivial holonomy, rather than a result of a physical fluctuating dynamical field. A “strange nature” of this type of energy manifests itself in terms of the “wrong” sign in the correlation function which cannot be formulated in terms of any local propagating degrees of freedom as explained in Sec. II A. The corresponding topological configurations which are responsible for this behavior may generate, as argued in this work, the linear in H correction in the Friedman equation (3) which eventually leads to the de Sitter behavior. Other problems formulated in terms of scalar inflaton field $\Phi(x)$ (such as large initial value $\Phi_{\text{in}} \gg M_{\text{PL}}$ for the inflaton) do not emerge in this framework, see [4,5] for the details.

Finally, we should mention that the energy described by a formula similar to Eq. (2) [which eventually leads to the de Sitter behavior (1)] has been previously postulated [48,49] as the driving force for the dark energy. The model has been (successfully) confronted with observations, see recent review papers [50,51] and many original references therein, where it has been claimed that this proposal is consistent with all presently available data, see also Ref. [52] for completeness. Our comment here is that history of evolution of the universe may repeat itself by realizing the de Sitter behavior twice in its history. The QCD-dynamics was responsible for the inflation in the

early universe, while the QCD dynamics is responsible for the dark energy in the present epoch.

We conclude this work (mainly devoted to the analysis of the topological configurations with typical energy scale $\Lambda_{\overline{\text{QCD}}}$) with the following comment related to a fundamentally different problem with a drastically different energy scale. Namely, as we discussed at length in this paper, the heart of the proposal is a fundamentally new type of energy (2), (24), (45) which cannot be expressed in terms of any propagating degrees of freedom. Rather, this novel contribution to the energy has nondispersive nature. The effect is formulated in terms of the tunneling processes between topologically different but physically identical states. This novel type of energy, in fact, has been well studied in the QCD lattice simulations in the flat background, see [4] for references on the original lattice results. Our comment relevant for the present study is that this fundamentally new type of energy can be, in principle, studied in a tabletop experiment by measuring some specific corrections to the Casimir vacuum energy in the Maxwell theory as suggested in [53–56]. This fundamentally new contribution to the Casimir pressure emerges as a result of tunneling processes, rather than due to the conventional fluctuations of the propagating photons with two physical transverse polarizations. This effect does not occur for the scalar field theory, in contrast with conventional Casimir effect which is operational for both: scalar as well as for Maxwell fields. The extra energy computed in [53–56] is the direct analog of the nondispersive contribution to the energy (2) which is the key player of the present work. In fact, an extra contribution to the Casimir pressure emerges in this system as a result of nontrivial holonomy similar to (5) for the Maxwell field. The nontrivial holonomy is enforced by the nontrivial boundary conditions imposed in Refs. [53–56].

ACKNOWLEDGMENTS

I am thankful to the participants of the workshop “Future prospects for fundamental particle physics and cosmology” (Simons Center for Geometry and Physics, Stony Brook, May 2015) where this work was presented, for their comments, questions, and suggestions. This research was supported in part by the Natural Sciences and Engineering Research Council of Canada.

APPENDIX: FEW COMMENTS ON FRACTIONALLY CHARGED CONSTITUENTS AND THE TERMINOLOGY

The constituents of the KvBLL configurations were originally [26,27] called the BPS (Bogomolnyi-Prasad-Sommerfeld) monopoles (14) and KK (Kaluza-Klein)-monopoles (17) correspondingly. These configurations were later on coined as M -dyons and L -dyons to emphasize that they carry the electric charges along with the magnetic charges

[28,29]. This (incorrect) interpretation is based on the observation that these configurations carry the topological charges (and naively the electric charges) along with magnetic charges. One should remember, however, that the monopoles in this construction are pseudoparticles living in 4d Euclidean space-time, rather than static 3d objects. The finite action and finite topological charge for these objects results from wrapping of the monopole's path along the Euclidean time direction S^1 with nontrivial holonomy (11). These objects do not carry a conventional static electric charge; nevertheless, they do carry the topological charges defined in 4d Euclidean space-time. Furthermore, the second types of the monopoles, the L monopoles are time dependent configurations, and do not exist as static objects in 3d Euclidean space. Therefore, we keep notations for letters M and L suggested in [28,29], but we use term “monopoles” rather than “dyons” in the present work. When the holonomy assumes its “confining” value $\nu = 1/2$ the topological charges of the constituents assume $Q = \pm 1/2$.

In the more generic case of the $SU(N)$ gauge group the topological charge $Q = \pm 1/N$ for “confining” holonomy such that a single KvBLL configuration can be thought of as a superposition of N different types of monopoles which carry N different types of magnetic charges and fractional $1/N$ topological charge such that the superposition carries an integer topological charge.

We believe a short historical detour on fractionalization of the topological charge in QFT is warranted here. In the given context fractional topological objects appear in 2 dimensional CP^{N-1} model [57] which were coined as *instanton quarks* (other names: instanton partons, fractional instantons). These quantum objects carry fractional topological charge $Q = \pm 1/N$, and they are very similar to the L and M monopoles discussed in this work. These objects do not appear individually in the path integral; instead, they appear as configurations consisting of N different objects with fractional charge $1/N$ such that the total topological charge of each configuration is always an integer. In this case $4Nk$ zero modes for k instanton solution is interpreted as 4 translation zero modes accompanied by every single instanton quark. The same counting holds, in fact, for any gauge group G , not limited to the $SU(N)$ case. While the instanton quarks emerge in the path integral

coherently, these objects are highly delocalized: they may emerge on opposite sides of the space-time or be close to each other with alike probabilities. Similar objects have been discussed in a number of papers in different contexts, including the topic of the present work, [21,22,26–29,58–63].

In particular, it has been argued that the well-established θ/N dependence in strongly coupled QCD [expressed by formula (26) for specific case $N = 2$] unambiguously implies that the relevant configurations in QCD must carry fractional topological charges in the confinement phase, see review preprint [59] and the references on earlier original results therein. The weakly coupled deformed QCD model [21–23] where computations are under complete theoretical control is a precise dynamical realization of this idea when the fractionally charged monopoles are responsible for confinement, saturate the topological susceptibility with a “wrong sign,” generate the “secrete long range forces,” suspected long ago [20], and provide other crucial elements which are known to exist in a strongly coupled regime as reviewed in Sec. II A.

Furthermore, it has been argued in [60,63] that the confinement deconfinement phase transition within this framework can be interpreted as a Berezinskii-Kosterlitz-Thouless (BKT)-like phase transition: at $T > T_c$ the constituents prefer to organize a single caloron of a finite size. We coin this phase as a “molecular phase” which corresponds to a deconfined phase in conventional terminology. When one crosses the phase transition line at $T < T_c$ the constituents (which are called L, M monopoles in the present work) prefer to stay far away from each other. It corresponds to the dissociation of each caloron into N constituents, and we call this state the “ N component plasma phase” in 4d Euclidean space. This regime corresponds to the confined phase in conventional terminology when all constituents are delocalized in 4d Euclidean space. The gap in this confined phase is determined by the Debye correlation length of this 4d plasma. The arguments [60,63] are based on large N counting, but we believe that this picture holds for any finite N . Recent numerical studies [32–34] are capable, in principle, to bring these large N qualitative arguments into a solid theoretical framework.

-
- [1] A. Guth, Inflationary universe: A possible solution to the horizon and flatness problems, *Phys. Rev. D* **23**, 347 (1981); A. D. Linde, A new inflationary universe scenario: A possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems, *Phys. Lett.* **108B**, 389 (1982).
 [2] A. D. Linde, Inflationary cosmology, *Lect. Notes Phys.* **738**, 1 (2008).

- [3] V. Mukhanov, *Physical Foundation of Cosmology* (Cambridge University Press, Cambridge, England, 2005).
 [4] A. R. Zhitnitsky, Inflaton as an auxiliary topological field in a QCD-like system, *Phys. Rev. D* **89**, 063529 (2014).
 [5] A. R. Zhitnitsky, Cosmological perturbations in QCD-inflation. Estimates confronting the observations, including BICEP2, *Phys. Rev. D* **90**, 043504 (2014).

- [6] Y. B. Zeldovich, Cosmological constant and elementary particles, *Pis'ma Zh. Eksp. Teor. Fiz.* **6**, 883 (1967) [Cosmological constant and elementary particles, *JETP Lett.* **6**, 316 (1967)].
- [7] J. Bjorken, Emergent gauge bosons, [arXiv:hep-th/0111196](https://arxiv.org/abs/hep-th/0111196).
- [8] R. Schutzhold, Small Cosmological Constant from the QCD Trace Anomaly?, *Phys. Rev. Lett.* **89**, 081302 (2002).
- [9] F. R. Klinkhamer and G. E. Volovik, Self-tuning vacuum variable and cosmological constant, *Phys. Rev. D* **77**, 085015 (2008).
- [10] E. C. Thomas, F. R. Urban, and A. R. Zhitnitsky, The cosmological constant as a manifestation of the conformal anomaly?, *J. High Energy Phys.* **08** (2009) 043.
- [11] M. Maggiore, Zero-point quantum fluctuations and dark energy, *Phys. Rev. D* **83**, 063514 (2011).
- [12] J. Sola, Cosmological constant and vacuum energy: Old and new ideas, *J. Phys. Conf. Ser.* **453**, 012015 (2013).
- [13] N. Kaloper and A. Padilla, Sequestering the Standard Model Vacuum Energy, *Phys. Rev. Lett.* **112**, 091304 (2014).
- [14] N. Kaloper and A. Padilla, Vacuum energy sequestering: The framework and its cosmological consequences, *Phys. Rev. D* **90**, 084023 (2014); **90**, 109901 (2014).
- [15] I. L. Shapiro and J. Sola, On the scaling behavior of the cosmological constant and the possible existence of new forces and new light degrees of freedom, *Phys. Lett. B* **475**, 236 (2000).
- [16] I. L. Shapiro and J. Sola, Scaling behavior of the cosmological constant: Interface between quantum field theory and cosmology, *J. High Energy Phys.* **02** (2002) 006.
- [17] E. Witten, Current algebra theorems for the U(1) Goldstone boson, *Nucl. Phys.* **B156**, 269 (1979).
- [18] G. Veneziano, U(1) without instantons, *Nucl. Phys.* **B159**, 213 (1979).
- [19] P. Di Vecchia and G. Veneziano, Chiral dynamics in the large N limit, *Nucl. Phys.* **B171**, 253 (1980).
- [20] M. Luscher, The secret long range force in quantum field theories with instantons, *Phys. Lett. B* **78**, 465 (1978).
- [21] M. Ünsal and L. G. Yaffe, Center-stabilized Yang-Mills theory: Confinement and large N volume independence, *Phys. Rev. D* **78**, 065035 (2008).
- [22] E. Thomas and A. R. Zhitnitsky, Topological susceptibility and contact term in QCD. A toy model, *Phys. Rev. D* **85**, 044039 (2012).
- [23] A. R. Zhitnitsky, QCD as a topologically ordered system, *Ann. Phys. (Amsterdam)* **336**, 462 (2013).
- [24] B. J. Harrington and H. K. Shepard, Periodic Euclidean solutions and the finite temperature Yang-Mills gas, *Phys. Rev. D* **17**, 2122 (1978).
- [25] D. J. Gross, R. D. Pisarski, and L. G. Yaffe, QCD and instantons at finite temperature, *Rev. Mod. Phys.* **53**, 43 (1981).
- [26] T. C. Kraan and P. van Baal, Periodic instantons with nontrivial holonomy, *Nucl. Phys.* **B533**, 627 (1998).
- [27] K. M. Lee and C. H. Lu, SU(2) calorons and magnetic monopoles, *Phys. Rev. D* **58**, 025011 (1998).
- [28] D. Diakonov, N. Gromov, V. Petrov, and S. Slizovskiy, Quantum weights of dyons and of instantons with nontrivial holonomy, *Phys. Rev. D* **70**, 036003 (2004).
- [29] D. Diakonov and V. Petrov, Confining ensemble of dyons, *Phys. Rev. D* **76**, 056001 (2007).
- [30] P. Faccioli and E. Shuryak, QCD topology at finite temperature: Statistical mechanics of self-dual dyons, *Phys. Rev. D* **87**, 074009 (2013).
- [31] E. Shuryak and T. Sulejmanpasic, Holonomy potential and confinement from a simple model of the gauge topology, *Phys. Lett. B* **726**, 257 (2013).
- [32] Y. Liu, E. Shuryak, and I. Zahed, Confining Dyon-Anti-Dyon coulomb liquid model I, [arXiv:1503.03058](https://arxiv.org/abs/1503.03058).
- [33] Y. Liu, E. Shuryak, and I. Zahed, Light quarks in the screened Dyon-Anti-Dyon coulomb liquid model II, [arXiv:1503.09148](https://arxiv.org/abs/1503.09148).
- [34] R. Larsen and E. Shuryak, Interacting ensemble of the instanton-dyons and deconfinement phase transition in the SU(2) gauge theory, [arXiv:1504.03341](https://arxiv.org/abs/1504.03341).
- [35] M. F. Atiyah, in *Collected Works* (Clarendon Press, Oxford, 1988), Vol. 5.
- [36] A. Chakrabarti, Construction of hyperbolic monopoles, *J. Math. Phys. (N.Y.)* **27**, 340 (1986).
- [37] C. Nash, Geometry of hyperbolic monopoles, *J. Math. Phys.* **27**, 2160 (1986).
- [38] H. Garland and M. K. Murray, Why instantons are monopoles, *Commun. Math. Phys.* **121**, 85 (1989).
- [39] N. S. Manton and T. M. Samols, Skyrmions on S^{*3} and H^3 from instantons, *J. Phys. A* **23**, 3749 (1990).
- [40] M. Atiyah and P. Sutcliffe, Skyrmions, instantons, mass and curvature, *Phys. Lett. B* **605**, 106 (2005).
- [41] D. Harland, Hyperbolic calorons, monopoles, and instantons, *Commun. Math. Phys.* **280**, 727 (2008).
- [42] P. Sutcliffe, Hyperbolic vortices with large magnetic flux, *Phys. Rev. D* **85**, 125015 (2012).
- [43] N. S. Manton and P. M. Sutcliffe, Platonic hyperbolic monopoles, *Commun. Math. Phys.* **325**, 821 (2014).
- [44] E. Thomas and A. R. Zhitnitsky, Casimir scaling in gauge theories with a gap. Deformed QCD as a toy model, *Phys. Rev. D* **86**, 065029 (2012).
- [45] B. Holdom, From confinement to dark energy, *Phys. Lett. B* **697**, 351 (2011).
- [46] A. Yamamoto, Lattice QCD in curved spacetimes, *Phys. Rev. D* **90**, 054510 (2014).
- [47] F. R. Urban and A. R. Zhitnitsky, Cosmological constant, violation of cosmological isotropy and CMB, *J. Cosmol. Astropart. Phys.* **09** (2009) 018.
- [48] F. R. Urban and A. R. Zhitnitsky, The QCD nature of dark energy, *Nucl. Phys.* **B835**, 135 (2010).
- [49] F. R. Urban and A. R. Zhitnitsky, The cosmological constant from the QCD Veneziano topological, *Phys. Lett. B* **688**, 9 (2010).
- [50] R.-G. Cai, Y. Gong, and B. Wang, Some studies on dark energies, *Universe* **2**, 21 (2014).
- [51] R.-G. Cai, Z.-L. Tuo, Y.-B. Wu, and Y.-Y. Zhao, More on QCD topological dark energy, *Phys. Rev. D* **86**, 023511 (2012).
- [52] A. Gomez-Valent and J. Sola, Vacuum models with a linear and a quadratic term in H: Structure formation and number counts analysis, *Mon. Not. R. Astron. Soc.* **448**, 2810 (2015).
- [53] C. Cao, M. van Caspel, and A. R. Zhitnitsky, Topological Casimir effect in Maxwell electrodynamics on a compact manifold, *Phys. Rev. D* **87**, 105012 (2013).

- [54] A. R. Zhitnitsky, Maxwell theory on a compact manifold as a topologically ordered system, *Phys. Rev. D* **88**, 105029 (2013).
- [55] A. Zhitnitsky, Topological order and Berry connection for the Maxwell vacuum on a four-torus, *Phys. Rev. D* **90**, 105007 (2014).
- [56] A. R. Zhitnitsky, Dynamical Casimir effect in a small compact manifold for the Maxwell vacuum, *Phys. Rev. D* **91**, 105027 (2015).
- [57] V. A. Fateev, I. V. Frolov, and A. S. Schwarz, Quantum fluctuations of instantons in the nonlinear σ model, *Nucl. Phys.* **B154**, 1 (1979); B. Berg and M. Luscher, Computation of quantum fluctuations around multi-instanton fields from exact Green's functions: The CP^{n-1} case, *Commun. Math. Phys.* **69**, 57 (1979).
- [58] N. M. Davies, T. J. Hollowood, V. V. Khoze, and M. P. Mattis, Gluino condensate and magnetic monopoles in supersymmetric gluodynamics, *Nucl. Phys.* **B559**, 123 (1999).
- [59] A. R. Zhitnitsky, Confinement-deconfinement phase transition and fractional instanton quarks in dense matter, [arXiv: hep-ph/0601057](https://arxiv.org/abs/hep-ph/0601057).
- [60] A. Parnachev and A. R. Zhitnitsky, Phase transitions, theta behavior and instantons in QCD and its holographic model, *Phys. Rev. D* **78**, 125002 (2008).
- [61] B. Collie and D. Tong, The partonic nature of instantons, *J. High Energy Phys.* **08** (2009) 006.
- [62] S. Bolognesi and K. Lee, Instanton partons in 5-dim SU(N) gauge theory, *Phys. Rev. D* **84**, 106001 (2011).
- [63] A. R. Zhitnitsky, Conformal window in QCD for large numbers of colors and flavors, *Nucl. Phys.* **A921**, 1 (2014).