

Non-Gaussianities in the topological charge distribution of the SU(3) Yang-Mills theory

Marco Cè,¹ Cristian Consonni,² Georg P. Engel,^{3,*} and Leonardo Giusti³

¹*Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy and INFN, Sezione di Pisa, Largo B. Pontecorvo 3, 56127 Pisa, Italy*

²*Dipartimento di Ingegneria e Scienza dell'Informazione, Università di Trento, Via Sommarive 9, 38123 Povo, Trento, Italy*

³*Dipartimento di Fisica, Università di Milano-Bicocca, and INFN, Sezione di Milano-Bicocca, Piazza della Scienza 3, 20126 Milano, Italy*

(Received 28 June 2015; published 6 October 2015)

We study the topological charge distribution of the SU(3) Yang-Mills theory with high precision in order to be able to detect deviations from Gaussianity. The computation is carried out on the lattice with high statistics Monte Carlo simulations by implementing a naive discretization of the topological charge evolved with the Yang-Mills gradient flow. This definition is far less demanding than the one suggested from Neuberger's fermions and, as shown in this paper, in the continuum limit its cumulants coincide with those of the universal definition appearing in the chiral Ward identities. Thanks to the range of lattice volumes and spacings considered, we can extrapolate the results for the second and fourth cumulant of the topological charge distribution to the continuum limit with confidence by keeping finite volume effects negligible with respect to the statistical errors. Our best results for the topological susceptibility is $t_0^2\chi = 6.67(7) \times 10^{-4}$, where t_0 is a standard reference scale, while for the ratio of the fourth cumulant over the second, we obtain $R = 0.233(45)$. The latter is compatible with the expectations from the large N_c expansion, while it rules out the θ behavior of the vacuum energy predicted by the dilute instanton model. Its large distance from 1 implies that, in the ensemble of gauge configurations that dominate the path integral, the fluctuations of the topological charge are of quantum *nonperturbative* nature.

DOI: [10.1103/PhysRevD.92.074502](https://doi.org/10.1103/PhysRevD.92.074502)

PACS numbers: 12.38.Gc, 11.15.Ha

I. INTRODUCTION

The discovery of a fermion operator [1] that satisfies the Ginsparg-Wilson (GW) relation [2] triggered a breakthrough in our understanding of the topological effects in Quantum Chromodynamics (QCD) and in the Yang-Mills theory [3–7]. This progress made it possible to give a precise and unambiguous implementation of the Witten-Veneziano formula [5,8–10].

In lattice QCD a naive definition of the topological charge density needs to be combined with an unambiguous renormalization condition. The cumulants of the charge, for instance the susceptibility, also require additional subtractions of short-distance singularities to make them integrable distributions. If the topological charge density is defined as suggested by GW fermions [1,3,4], however, its bare lattice expression and those of the corresponding cumulants have finite and unambiguous continuum limits as they stand, which in turn satisfy the anomalous chiral Ward identities [5–7]. By combining a series of those identities, the cumulants can be written as integrated correlation functions of scalar and pseudoscalar density chains or combination of them [6,7,11].

In this form a particular regularization is not required anymore to prove that no renormalization factor or subtractions of short-distance singularities are required. These expressions provide a universal definition of the susceptibility and of the higher cumulants which satisfy the anomalous chiral Ward identities [7].

Recently a new definition of the topological charge was found [12], whose cumulants have a finite and unambiguous continuum limit [12,13]. It is a naive discretization of the charge evolved with the Yang-Mills gradient flow. It is particularly appealing because its numerical evaluation is significantly cheaper than the one for the definition suggested by Neuberger's fermions. Here we show that in the Yang-Mills theory the cumulants defined this way coincide, in the continuum limit, with those of the universal definition appearing in the anomalous chiral Ward identities of QCD. By implementing the gradient-flow definition, we compute the topological susceptibility in the continuum limit with a precision 5 times better than the reference computation with the Neuberger's definition [14]. We then determine the ratio of the fourth cumulant over the second one in the continuum limit by keeping for the first time all systematics, especially finite volume effects, negligible with respect to the statistical errors. As a byproduct we also perform an interesting universality test at the

*Present address: Institute for Sustainable Technologies–AEE INTEC, Feldgasse 19, 8020 Gleisdorf, Austria.

per mille level by comparing the values of the topological susceptibility at different flow times.

II. PRELIMINARIES IN THE CONTINUUM

Starting from the ordinary fundamental gauge field

$$B_\mu|_{t=0} = A_\mu, \quad (2.1)$$

where $A_\mu = A_\mu^a T^a$ (see Appendix A for the generator conventions), the Yang-Mills gradient flow evolves the gauge field as a function of the flow time $t \geq 0$ by solving the differential equation [12]

$$\partial_t B_\mu = D_\nu G_{\nu\mu} + \alpha_0 D_\mu \partial_\nu B_\nu, \quad (2.2)$$

$$G_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu - i[B_\mu, B_\nu], \quad D_\mu = \partial_\mu - i[B_\mu, \cdot], \quad (2.3)$$

with α_0 being the parameter which determines the gauge. Here we focus on the gradient-flow evolution of the topological charge density defined as¹

$$q^t = \frac{1}{32\pi^2} \epsilon_{\mu\nu\rho\sigma} \text{tr}[G_{\mu\nu} G_{\rho\sigma}], \quad (2.4)$$

and of the corresponding topological charge,

$$Q^t = \int d^4x q^t(x), \quad (2.5)$$

where $\epsilon_{\mu\nu\rho\sigma}$ is the four-index totally antisymmetric tensor and the trace is over the color index. Under a generic variation δB_μ of a given gauge field configuration, the topological charge density changes as

$$\delta q^t = \partial_\rho \tilde{w}_\rho^t, \quad \tilde{w}_\rho^t = \frac{1}{8\pi^2} \epsilon_{\rho\mu\nu\sigma} \text{tr}[G_{\mu\nu} \delta B_\sigma]; \quad (2.6)$$

see, for instance, Ref. [15]. If we now specify

$$\delta B_\mu = \partial_t B_\mu \delta t, \quad (2.7)$$

it is straightforward to show that

$$\partial_t q^t = \partial_\rho w_\rho^t, \quad w_\rho^t = \frac{1}{8\pi^2} \epsilon_{\rho\mu\nu\sigma} \text{tr}[G_{\mu\nu} D_\alpha G_{\alpha\sigma}], \quad (2.8)$$

where w_ρ^t is a local dimension-five gauge-invariant pseudovector field. This in turn implies that for a given gauge field configuration

$$\partial_t Q^t = 0, \quad (2.9)$$

an equation which reflects the topological nature of Q^t .

¹If not explicitly indicated, the superscript t on the quantities evolved with the gradient flow is always > 0 .

When $q^t(x)$ is inserted in a correlation function, Eq. (2.8) implies

$$\langle q^t(x) O(y) \rangle = \langle q^{t=0}(x) O(y) \rangle + \partial_\rho \int_0^t dt' \langle w_\rho^{t'}(x) O(y) \rangle \quad (x \neq y), \quad (2.10)$$

where $O(y)$ is any finite (multi)local operator inserted at a physical distance from x . The lhs of Eq. (2.10) is finite thanks to the fact that a gauge-invariant local composite field constructed with the gauge field evolved at positive flow time is finite [12,13]. Since there are no local composite fields of dimension $d < 5$ with the symmetry properties of $w_\rho^t(x)$, the integrand on the rhs of Eq. (2.10) diverges at most logarithmically when $t' \rightarrow 0$. This implies that the quantity

$$\langle q^{t=0}(x) O(y) \rangle \equiv \lim_{t \rightarrow 0} \langle q^t(x) O(y) \rangle \quad (x \neq y), \quad (2.11)$$

is finite; i.e. the limit on the rhs exists for any finite operator $O(y)$. The Eq. (2.11) can be taken as the definition of $q^{t=0}(x)$, i.e. the renormalized topological charge density operator at $t = 0$. The latter satisfies the proper singlet chiral Ward identities when fermions are included; see the next section for an explicit derivation. It is worth noting that Eq. (2.10) implies that the small- t expansion of $q^t(x)$ is of the form

$$\langle q^t(x) O(y) \rangle = \langle q^{t=0}(x) O(y) \rangle + \mathcal{O}(t) \quad (x \neq y). \quad (2.12)$$

In the following we will be interested in the cumulants of the topological charge

$$C_n^t = \int d^4x_1 \dots d^4x_{2n-1} \langle q^t(x_1) \dots q^t(x_{2n}) \rangle_c, \quad (2.13)$$

which, thanks to Eq. (2.9), are expected to be independent of the flow time for $t \geq 0$ with the limit $t \rightarrow 0$ which requires some care due to the possible appearance of short-distance singularities. It is the aim of the next section to address this question by using the lattice, a regularization where the theory can be nonperturbatively defined. We will supplement the theory with extra degenerate valence quarks of mass m , and we will consider the (integrated) correlator of a topological charge density with a chain made of scalar and pseudoscalar densities [7] defined as

$$\langle q^{t=0}(0) P_{51}(z_1) S_{12}(z_2) S_{23}(z_3) S_{34}(z_4) S_{45}(z_5) \rangle, \quad (2.14)$$

where S_{ij} and P_{ij} are the scalar and the pseudoscalar renormalized densities with flavor indices i and j . Power counting and the operator product expansion predict that there are no non-integrable short-distance singularities when the coordinates of two or more densities in (2.14) tend to coincide among themselves or with 0. When only

one of the densities is close to $q^{t=0}(0)$, the operator product expansion predicts the leading singularity to be

$$q^{t=0}(x)S_{ij}(0) \xrightarrow{x \rightarrow 0} c(x)P_{ij}(0) + \dots, \quad (2.15)$$

where $c(x)$ is a function which diverges as $|x|^{-4}$ when $|x| \rightarrow 0$, and the dots indicate subleading contributions. An analogous expression is valid for the pseudoscalar density. Being the leading short-distance singularity in the product of fields $q^{t=0}(x)S_{ij}(0)$, its Wilson coefficient $c(x)$ can be computed in perturbation theory. By using Eq. (2.10), to all orders in perturbation theory² we can write

$$\begin{aligned} \langle q^{t=0}(x)S_{ij}(0)O(y) \rangle &= \langle q^t(x)S_{ij}(0)O(y) \rangle \\ &\quad - \partial_\rho \int_0^t dt' \langle w_\rho^{t'}(x)S_{ij}(0)O(y) \rangle, \end{aligned} \quad (2.16)$$

where again $O(y)$ is any finite (multi)local operator inserted at a physical distance from 0 and x . When $t > 0$, the first member on the rhs of Eq. (2.16) has no singularities when $|x| \rightarrow 0$. If present, the singularity has to come from the second term, and therefore $c(x)$ must be of the form

$$c(x) = \partial_\rho u_\rho(x) \quad (2.17)$$

which does not contribute to the integral (over all coordinates) of the correlation function (2.14).

III. CUMULANTS OF THE TOPOLOGICAL CHARGE ON THE LATTICE

On the lattice the Yang-Mills gradient-flow equation can be written as a first-order differential equation [12]

$$\partial_t V_\mu(x) = -g_0^2 \{ \partial_{x,\mu} S(V) \} V_\mu(x), \quad V_\mu(x)|_{t=0} = U_\mu(x), \quad (3.1)$$

where the Wilson action S and the link differential operators $\partial_{x,\mu}$ are defined in Appendix A together with other conventions. The gauge field evolved at positive flow time $V_\mu(x)$ is smooth on the scale of the cut-off. When inserted at a physical distance, the gauge-invariant local composite fields constructed with the evolved gauge field are finite as they stand. Remarkably their universality class is determined only by their asymptotic behavior in the classical continuum limit [12,13]. At $t > 0$ any decent definition of the topological charge density is therefore finite. The same line of argumentation applies to the cumulants of the topological charge. At $t > 0$ short-distance

singularities cannot arise because of the exponential damping of the high-frequency components of the fields enforced by the flow evolution.

It remains to be shown, however, that the cumulants of the topological charge distribution defined at $t > 0$ satisfy the proper singlet chiral Ward identities when fermions are included in the theory. To show this it is sufficient to work with a particular discretization of the topological charge, and then appeal to the above mentioned universality argument for the other definitions. The GW discretizations have a privileged rôle since at $t = 0$ the lattice bare cumulants are finite, and they satisfy the singlet chiral Ward identities when fermions are included in the theory.

A. Ginsparg-Wilson definition of the charge density

The definition of the topological charge density suggested by GW fermions is [3,4,16]

$$a^4 q_N^t(x) = -\frac{\bar{a}}{2} \text{tr}[\gamma_5 D(x, x)], \quad (3.2)$$

where we indicate it with a subscript N since, for concreteness, we take $D(x, y)$ to be the Neuberger-Dirac operator given in Appendix A in which each link variable $U_\mu(x)$ is replaced by the corresponding evolved one $V_\mu(x)$ when $t > 0$. Since there are no other operators of dimension $d \leq 4$ which are pseudoscalar and gauge-invariant, it holds that

$$\lim_{a \rightarrow 0} Z_q \langle q_N^t(0) q_N^{t=0}(x) \rangle = \text{finite}, \quad (3.3)$$

where Z_q is a renormalization constant which is at most logarithmically divergent, while $q_N^t(0)$ is finite as it stands. This in turn implies that

$$\lim_{a \rightarrow 0} Z_q a^4 \sum_x \langle q_N^t(0) q_N^{t=0}(x) \rangle = \text{finite}, \quad (3.4)$$

since there are no short-distance singularities that contribute to the integrated correlation function because $q_N^t(0)$ is evolved at positive flow time. By supplementing the theory with extra degenerate valence quarks of mass m , and by replacing in Eq. (3.4) the topological charge at $t = 0$ with its density-chain expression [7] we obtain

$$\begin{aligned} a^4 \sum_x \langle q_N^t(0) q_N^{t=0}(x) \rangle &= -m^5 a^{20} \sum_{z_1, \dots, z_5} \langle q_N^t(0) P_{51}(z_1) \\ &\quad \times S_{12}(z_2) S_{23}(z_3) S_{34}(z_4) S_{45}(z_5) \rangle, \end{aligned} \quad (3.5)$$

where S_{ij} and P_{ij} are the scalar and the pseudoscalar densities with flavor indices i and j . Written as in Eq. (3.5), power counting and the operator product expansion predict that there are no non-integrable short-distance singularities

²Since the function $|x|^{-4} \ln(x^2)^{-p}$ is integrable for $p > 1$, the singularity needs to be determined only up to some finite order.

when the coordinates of two or more densities tend to coincide. The rhs of Eq. (3.5) is finite as it stands, and it converges to the continuum limit with a rate proportional to a^2 . This in turn implies that the limits on the lhs of Eqs. (3.3) and (3.4) are reached with the same rate if Z_q is set to any fixed (g_0 -independent) value. Since in the classical continuum limit Neuberger's definition in Eq. (3.2) has the same asymptotic behavior of the definition in Eq. (2.4) [17,18], we may set $Z_q = 1$ in which case

$$\lim_{a \rightarrow 0} \langle q_N^{t=0}(x) O_L(y) \rangle = \langle q^{t=0}(x) O(y) \rangle \quad (x \neq y), \quad (3.6)$$

where $O_L(y)$ is a discretization of the generic finite continuum operator $O(y)$. Once inserted in correlation functions at a physical distance from other (renormalized) fields, $q_N^{t=0}(x)$ does not require any renormalization in the Yang-Mills theory. It is finite as it stands, and it satisfies the singlet Ward identities when fermions are included in the theory. It is interesting to note that Eqs. (2.12) and (3.6) imply

$$\langle q^t(x) O(y) \rangle = \langle q_N^{t=0}(x) O_L(y) \rangle + \mathcal{O}(a^2) + \mathcal{O}(t), \quad (3.7)$$

where in general discretization effects depend on t . We could have arrived at Eq. (3.6) by following a procedure analogous to the one in the continuum; see Eqs. (2.8)–(2.12). To all orders in perturbation theory, or in general when Neuberger's operator is differentiable with respect to the gauge field [19], the change of the topological charge density with respect to the flow time can be written, analogously to Eq. (2.8), as [20,21] (see also [22])

$$\partial_t q_N^t(x) = \partial_\rho^* w_{N,\rho}^t(x), \quad (3.8)$$

where $w_{N,\rho}^t(x)$ is a discretization of the dimension-five gauge-invariant pseudovector operator $w_\rho^t(x)$, and ∂_ρ^* is the backward finite-difference operator.

B. Ginsparg-Wilson definition of the charge cumulants

The Neuberger's definition of the topological charge is given by

$$Q_N^t \equiv a^4 \sum_x q_N^t(x), \quad (3.9)$$

and its cumulants are defined as

$$C_{N,n}^t = a^{8n-4} \sum_{x_1, \dots, x_{2n-1}} \langle q_N^t(x_1) \dots q_N^t(x_{2n-1}) q_N^t(0) \rangle_c. \quad (3.10)$$

For $t = 0$ the cumulants have an unambiguous universal continuum limit as they stand and, when fermions are included, they satisfy the proper singlet chiral Ward

identities [5–7]. They are the proper quantities to be inserted in the Witten-Veneziano relations for the mass and scattering amplitudes of the η' meson in QCD [5,8–10]. It is far from being obvious that $C_{N,n}^{t=0}$ coincide with those defined at positive flow time, since the two definitions may differ by additional finite contributions from short-distance singularities.

For the clarity of the presentation we start by focusing on the lowest cumulant, the topological susceptibility $C_{N,1}^t$. At $t = 0$, by replacing one of the two $q_N^{t=0}$ with its density-chain expression [7], we obtain

$$a^4 \sum_x \langle q_N^{t=0}(0) q_N^{t=0}(x) \rangle = -m^5 a^{20} \sum_{z_1, \dots, z_5} \langle q_N^{t=0}(0) P_{51}(z_1) \times S_{12}(z_2) S_{23}(z_3) S_{34}(z_4) S_{45}(z_5) \rangle. \quad (3.11)$$

When the susceptibility is written in this form, the discussion toward the end of Sec. II and, in particular, Eq. (2.17) guarantee that there are no contributions from short-distance singularities. This result, together with the fact that $Z_q = 1$, implies that

$$\lim_{t \rightarrow 0} \lim_{a \rightarrow 0} a^4 \sum_x \langle q_N^t(x) q_N^{t=0}(0) \rangle = \lim_{a \rightarrow 0} a^4 \sum_x \langle q_N^{t=0}(x) q_N^{t=0}(0) \rangle. \quad (3.12)$$

By replacing on the lhs $q_N^{t=0}(0)$ with the evolved one, no further short-distance singularities are introduced and we arrive at the final result,

$$\lim_{t \rightarrow 0} \lim_{a \rightarrow 0} a^4 \sum_x \langle q_N^t(x) q_N^t(0) \rangle = \lim_{a \rightarrow 0} a^4 \sum_x \langle q_N^{t=0}(x) q_N^{t=0}(0) \rangle. \quad (3.13)$$

By replacing $2n - 1$ of the charges in the n th cumulant with their density-chain definitions, the very same line of argumentation can be applied. The Eq. (3.13), together with the independence up to harmless discretization effects of $C_{N,n}^t$ from the flow time for $t > 0$ [12], implies that the continuum limit of $C_{N,n}^t$ coincides with the one of $C_{N,n}^{t=0}$. The cumulants of the topological charge distribution defined at $t > 0$ thus satisfy the proper singlet chiral Ward identities when fermions are included [5–7]. They are the proper quantities to be inserted in the Witten-Veneziano relations for the mass and scattering amplitudes of the η' meson in QCD [5,8–10].

C. Universality at positive flow time

For $t > 0$ different lattice definitions of the topological charge density belong to the same universality class if they share the same asymptotic behavior in the classical continuum limit [12,13]. In the rest of this paper we are

TABLE I. Overview of the ensembles and statistics used in this study. For each lattice we give the label, $\beta = 6/g_0^2$, the reference scale t_0/a^2 , the spatial extent of the lattice, the lattice spacing, the number N_{conf} of independent configurations generated, the number of sweeps `nit` required to space them, and the tolerances `eerr`, `q2err` and `q4err` on the primary observables considered (see main text).

Lattice	β	t_0/a^2	L/a	L [fm]	a [fm]	N_{conf}	<code>nit</code>	<code>eerr</code>	<code>q2err</code>	<code>q4err</code>
A_1	5.96	2.79	10	1.0	0.102	36000	30	0.19	0.0005	0.0024
B_1			12	1.2		144000		0.45		0.005
C_1			13	1.3		280000		0.42		0.0068
D_1			14	1.4		505000		0.74		0.01
E_1			15	1.5		880000		0.89		0.012
F_1			16	1.6		1440000		1.04		0.015
B_2	6.05	3.78	14	1.2	0.087	144000	60	0.31	0.0005	0.005
D_2			17	1.5		144000		0.045		0.01
B_3	6.13	4.87	16	1.2	0.077	144000	90	0.25	0.0005	0.005
D_3			19	1.5		144000		0.058		0.01
B_4	6.21	6.20	18	1.2	0.068	144000	250	0.20	0.0005	0.005
D_4			21	1.4		144000		0.042		0.01

interested in the naive definition of the topological charge density defined as³

$$q^t(x) = \frac{1}{64\pi^2} \epsilon_{\mu\nu\rho\sigma} G_{\mu\nu}^a(x) G_{\rho\sigma}^a(x), \quad (3.14)$$

where the field strength tensor $G_{\mu\nu}^a(x)$ is defined as [23]

$$G_{\mu\nu}^a(x) = -\frac{i}{4a^2} \text{tr}[(Q_{\mu\nu}(x) - Q_{\nu\mu}(x))T^a], \quad (3.15)$$

with

$$\begin{aligned} Q_{\mu\nu}(x) = & V_\mu(x)V_\nu(x + a\hat{\mu})V_\mu^\dagger(x + a\hat{\nu})V_\nu^\dagger(x) \\ & + V_\nu(x)V_\mu^\dagger(x - a\hat{\mu} + a\hat{\nu})V_\nu^\dagger(x - a\hat{\mu})V_\mu(x - a\hat{\mu}) \\ & + V_\mu^\dagger(x - a\hat{\mu})V_\nu^\dagger(x - a\hat{\mu} - a\hat{\nu}) \\ & \times V_\mu(x - a\hat{\mu} - a\hat{\nu})V_\nu(x - a\hat{\nu}) \\ & + V_\nu^\dagger(x - a\hat{\nu})V_\mu(x - a\hat{\nu})V_\nu(x + a\hat{\mu} - a\hat{\nu})V_\mu^\dagger(x). \end{aligned} \quad (3.16)$$

In the Yang-Mills theory, $q^{t=0}(x)$ requires a multiplicative renormalization constant⁴ when inserted in correlation functions at a physical distance from other operators [24]. The cumulants of the corresponding topological charge $Q^{t=0} \equiv a^4 \sum_x q^{t=0}(x)$, defined analogously to Eq. (3.10), have additional ultraviolet power-divergent

³We use the same notation for the naive definition of the field strength tensor, of the topological charge and of its density on the lattice and in the continuum, since any ambiguity is resolved from the context.

⁴This renormalization constant can be fixed by enforcing the analog of Eqs. (3.3) and (3.6).

singularities, and they do not have a well defined continuum limit.

The density $q^t(x)$ in Eq. (3.14) shares with $q_N^t(x)$ the same asymptotic behavior in the classical continuum limit [17,18]. Since for $t > 0$ short-distance singularities cannot arise, $C_{N,n}^t$ and C_n^t tend to the same continuum limit. The results in the previous section then imply that the continuum limit of the naive definition of C_n^t , at positive flow time, coincides with the universal definition which satisfies the chiral Ward identities when fermions are added [5–7]. It is interesting to note, however, that at fixed lattice spacing there can be quite some differences. For instance, the topological susceptibility defined at $t > 0$ with the naive definition is not guaranteed to go to zero in the chiral limit at finite lattice spacing in presence of fermions [25].

IV. NUMERICAL SETUP

For the numerical computation we discretize the SU(3) Yang-Mills theory with the standard Wilson plaquette action on a finite four-dimensional lattice with spacing a , with the same L/a size in all four spacetime directions, and with periodic boundary conditions imposed on the gauge fields, see Appendix A for details. The basic Monte Carlo update of each link variable implements the Cabibbo-Marinari scheme [26], by sweeping the full lattice with one heat bath update followed by $L/(2a)$ sweeps of over-relaxation updates.

A. Ensembles generated

We have simulated three series of lattices in order to estimate and remove the systematic effects due to the finiteness of the lattice spacing and volume, see Table I for details. In the first series $\{A_1, B_1, \dots, F_1\}$ the inverse coupling $\beta = 6/g_0^2$ is kept fixed so that the lattice spacing is approximatively 0.1 fm, while the physical volume

increases from $(1.0 \text{ fm})^4$ to $(1.6 \text{ fm})^4$. The number N_{conf} of independent gauge configurations generated scales with L^8 to ensure that the relative statistical error on R , the ratio of the fourth over the second cumulant of the topological charge distribution see Eq. (4.2), is always at the 10% level [27]. In the second series $\{B_1, \dots, B_4\}$ the physical volume is kept approximately fixed, while the spacing is decreased down to 0.068 fm . The volume is always $(1.2 \text{ fm})^4$ to guarantee that finite-size effects on R are within the statistical errors, while the computational cost remains affordable. In the third series $\{D_1, \dots, D_4\}$ is again the physical volume which is kept approximately fixed, always at least $(1.4 \text{ fm})^4$, to guarantee that finite-size effects in the reference scale t_0 and in the topological susceptibility χ , see Eq. (4.2), are within their (smaller) statistical errors. In both cases the measurements at the four lattice spacings are used to estimate discretization effects in the observables and to extrapolate them away in the continuum limit.

B. Computation of the observables

The primary observables that we have computed on each configuration at $t \geq 0$ are the energy density,

$$E^t = \frac{a^4}{4V} \sum_x F_{\mu\nu}^{a,t}(x) F_{\mu\nu}^{a,t}(x), \quad (4.1)$$

and the topological charge Q defined as in Sec. III C. The quantum averages we are interested in are

$$\langle E^t \rangle, \quad \chi^t = \frac{\langle [Q^t]^2 \rangle}{V}, \quad R^t = \frac{\langle [Q^t]^4 \rangle_c}{\langle [Q^t]^2 \rangle}. \quad (4.2)$$

To numerically integrate the Yang-Mills gradient flow, we have implemented a fourth order *Runge-Kutta-Munthe-Kaas (RKMK) method* [28–30]. It is a *structure-preserving* Runge-Kutta (RK) integrator, designed to exactly preserve the Lie group structure of the gradient flow equation (see Appendix B for details). On each lattice the field has been evolved approximately up to $t = 1.2t_0$, where t_0 is the reference flow-time value defined below. The observables in Eq. (4.2) have been computed with a flow-time resolution of $0.08a^2$ or smaller. The numerical integration of the flow equation introduces a systematic error in the gauge field values at positive flow time t , and thus in each observable. In our case, at asymptotically small values of the RK step size, it is proportional to ϵ^4 . There are, however, large fluctuations in the prefactor among the various gauge configurations (see Fig. 1). A reliable estimate of this systematics is achieved by monitoring the error configuration by configuration, and occasionally adapt the step ϵ . To do so we integrate the flow equation two times with steps ϵ and 2ϵ , where in our case $\epsilon = 0.08a^2$. Denoting with $E_j^{(\epsilon)}$ and $Q_j^{(\epsilon)}$ the basic

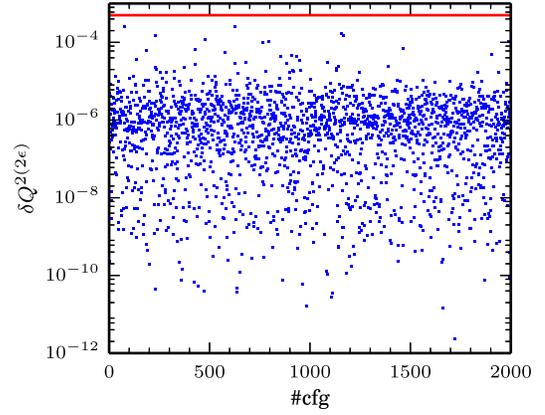


FIG. 1 (color online). History plot of the systematic error $\delta Q^{2,(2\epsilon)}$ at $t \approx t_0$ for the first 2000 configurations of D_4 . The red line indicates the bound $q2err = 0.0005$ of the systematic error enforced on all configurations.

observables E and Q respectively computed on the j th field configuration evolved with step size ϵ , at small enough ϵ the error is given by

$$\delta E_j^{(2\epsilon)} = |E_j^{(\epsilon)} - E_j^{(2\epsilon)}|, \quad \delta Q_j^{(2\epsilon)} = |Q_j^{(\epsilon)} - Q_j^{(2\epsilon)}|, \quad (4.3)$$

with both observables obviously measured at the same flow time. By applying linear propagation, the error on the average over all configurations is bounded by

$$\begin{aligned} \delta \bar{E}_j^{(2\epsilon)} &\leq \max_j(\delta E_j^{(2\epsilon)}), & \delta \bar{Q}^{2,(2\epsilon)} &\leq \max_j(|2Q_j| \delta Q_j^{(2\epsilon)}), \\ \delta \bar{Q}^{4,(2\epsilon)} &\leq \max_j(|4Q_j^3| \delta Q_j^{(2\epsilon)}), \\ \delta \bar{R}^{(2\epsilon)} &\leq \frac{1}{\bar{Q}^2} \max \left(\max_j(|4Q_j^3| \delta Q_j^{(2\epsilon)}), \right. \\ &\quad \left. \frac{\bar{Q}^4 + 3(\bar{Q}^2)^2}{\bar{Q}^2} \max_j(|2Q_j| \delta Q_j^{(2\epsilon)}) \right). \end{aligned} \quad (4.4)$$

At run time, for each configuration and each flow time, the systematic errors of the observables E , Q^2 and Q^4 are compared with the given tolerances $eerr$, $q2err$ and $q4err$ respectively. If one of the tests fails, the flow evolution is recomputed for that configuration with a new step size $\epsilon' = (1/2)\epsilon$ and new observables data, along with old $\epsilon = 2\epsilon'$ data, are used to estimate the systematic errors and compare them with the tolerances. If the test fails again, the field is evolved with $\epsilon'' = (1/2)\epsilon'$, and so on. This ensures that

$$\begin{aligned} \delta \bar{E}_j^{(2\epsilon)} &\leq eerr, & \delta \bar{Q}^{2,(2\epsilon)} &\leq q2err, \\ \delta \bar{Q}^{4,(2\epsilon)} &\leq q4err. \end{aligned} \quad (4.5)$$

The parameters $eerr$, $q2err$ and $q4err$ are chosen as a function of the target statistical error on the corresponding

TABLE II. Integrated autocorrelation times of the various observables in units of a single sweep of the update algorithm. They have been measured on dedicated runs made of 36 series of 1000 sweeps each.

Lattice	N_{conf}	t/a^2	τ_{int}^E	τ_{int}^Q	$\tau_{\text{int}}^{Q^2}$	$\tau_{\text{int}}^{Q^4}$
B_{1a}	36×1000	3.36	7.0(5)	9.1(7)	5.2(3)	4.39(25)
B_{2a}	36×1000	4.64	7.9(6)	17.4(18)	7.9(6)	6.3(4)
B_{3a}	36×1000	6.08	12.4(11)	30(4)	12.4(11)	9.2(7)
B_{4a}	36×1000	7.68	17.2(18)	73(13)	31(4)	22.1(25)

observables. If we set the upper limit for the systematic error to be roughly 10 times smaller than the statistical one, this condition is readily translated into a limit for eerr , q2err and q4err , see Table I for the values chosen for each lattice. The Eqs. (4.5) put bounds on the systematic errors for the coarser evolution, but the data evolved with the finer step size ϵ are actually those used in the final analysis. This choice is rather conservative in our case, being the actual error more than 1 order of magnitude smaller. For the quantity E^t the actual error turns out to be more than two orders of magnitude smaller with respect to the bound in Eq. (4.5), see Fig. 8 in Appendix B. We have, therefore, chosen larger values for eerr with respect to one given by the bound in Eq. (4.5).

C. Autocorrelation times

To measure the autocorrelation time of the various observables, we perform a dedicated run for each lattice $\{B_1, \dots, B_4\}$ where the gauge field configurations are separated by a single iteration of the update algorithm. Each series is replicated 36 times to increase statistical accuracy. The integrated autocorrelation times τ_{int} of the observables E , Q , Q^2 , and Q^4 , estimated as in Ref. [31], are reported in Table II and shown in Fig. 2. In the range of β values considered, Q has the largest autocorrelation time which increases rapidly toward the continuum limit [32]. To ensure that the measurements in the main runs are statistically independent, we have spaced them by `nit` sweeps of the lattice, see Table I.

V. PHYSICS RESULTS

A first analysis of the data reveals the effectiveness of the gradient flow in splitting the field space of the lattice theory into different topological sectors. In Fig. 3 we plot the histograms of the topological charge Q measured at different flow times on the lattice D_4 . In the plot on the top-left corner, the topological charge distribution at $t=0$ is a smooth function over noninteger values. By increasing the flow time, the configurations with charge close to integers become more and more probable. The spikes in the bottom-right plot turn out to be slightly shifted towards zero with respect to the integer values due to discretization effects. On the other lattices similar histograms are obtained.

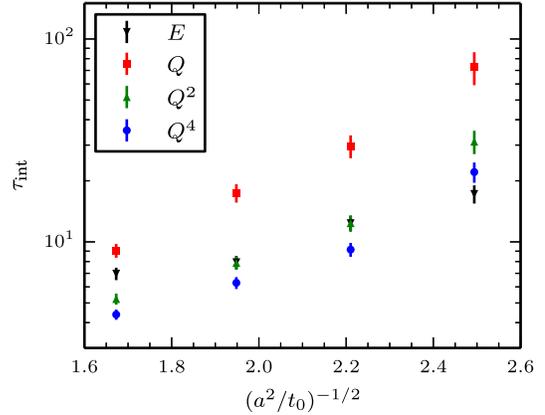


FIG. 2 (color online). The integrated autocorrelation times τ_{int} of the primary observables as a function of $(a^2/t_0)^{-1/2}$.

A. Scale setting

The reference flow time t_0 is defined through the implicit equation [12]

$$t^2 \langle E^t \rangle|_{t=t_0} = 0.3. \quad (5.1)$$

In the region of interest $t^2 \langle E^t \rangle$ grows approximatively as a linear function of t . Since we have computed $\langle E^t \rangle$ at flow times spaced by finite steps, we have solved equation (5.1) by interpolating linearly the two data points closest to t_0 . The results are reported in Table III, with the systematic error due to the interpolation being negligible. By comparing the values of t_0/a^2 obtained on the lattices $\{A_1, \dots, F_1\}$, finite-size effects are not visible at the level of 0.1 per mille in the statistical precision for $L \geq 1.4$ fm. We thus fix the lattice spacing at all values of β from t_0/a^2 determined on the lattices $\{D_1, \dots, D_4\}$.

In Table III the values of t_0/r_0^2 , where r_0 is the Sommer scale computed in [33], are also reported. As shown in Fig. 4, discretization effects in this ratio are indeed negligible with respect to the statistical errors dominated by the 0.3%–0.6% error on r_0/a . By extrapolating the results linearly in a^2/t_0 , we obtain in the continuum limit

$$\sqrt{8t_0}/r_0 = 0.941(7), \quad (5.2)$$

which corresponds to $t_0/r_0^2 = 0.1108(17)$. To express t_0 in physical units, we supplement the theory with quenched quarks. The value of $F_K r_0 = 0.293(7)$ from Ref. [34] together with $F_K = 109.6$ MeV leads to⁵

⁵Note that we use an updated determination for the physical value of F_K with respect to Ref. [34]. The change on $F_K r_0$ induced by the new tuning of the strange quark mass is negligible with respect to the statistical error quoted.

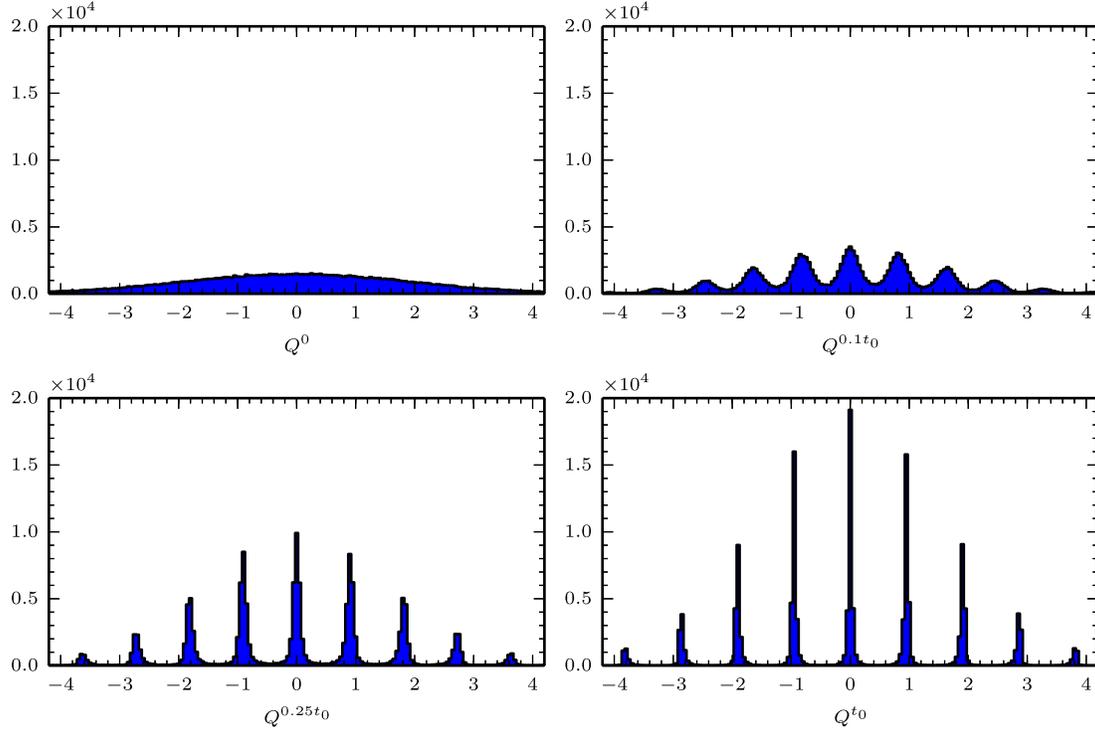


FIG. 3 (color online). Histograms of the topological charge distribution measured on the lattice D_4 at different flow times.

$$t_0 = (0.176(4) \text{ fm})^2, \quad (5.3)$$

the error being dominated by the one on $F_K r_0$.

B. Topological susceptibility

The full set of results for the topological charge moments and cumulants are given in Table IV. They

TABLE III. Results for the reference flow time t_0/a^2 and the ratio t_0/r_0^2 . The error on the latter is dominated by the 0.3%–0.6% relative error on r_0/a quoted in [33].

Lattice	t_0/a^2	t_0/r_0^2
A_1	2.995(4)	0.1195(9)
B_1	2.7984(9)	0.1117(9)
C_1	2.7908(5)	0.1114(9)
D_1	2.7889(3)	0.1113(9)
E_1	2.78892(23)	0.1113(9)
F_1	2.78867(16)	0.1113(9)

Lattice	t_0/a^2	t_0/r_0^2
B_2	3.7960(12)	0.1114(9)
B_3	4.8855(15)	0.1113(10)
B_4	6.2191(20)	0.1115(11)
D_2	3.7825(8)	0.1110(9)
D_3	4.8722(11)	0.1110(10)
D_4	6.1957(14)	0.1111(11)

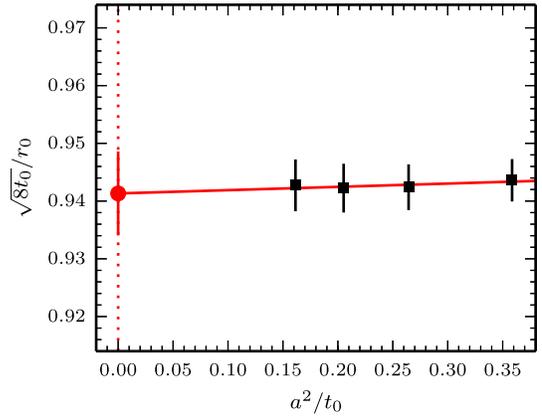


FIG. 4 (color online). Continuum limit extrapolation of $\sqrt{8t_0}/r_0$ computed on the lattices $\{D_1, \dots, D_4\}$. The errors are dominated by the 0.3%–0.6% relative error on r_0/a quoted in [33].

are computed⁶ at the reference flow time t_0 by linearly interpolating the numerical data as described in the previous section.

In Fig. 5 we show the values of the topological susceptibility $\chi = \langle Q^2 \rangle / V$ from the lattices $\{A_1, \dots, F_1\}$

⁶Unless explicitly indicated, the gradient flow time at which the topological quantities are computed throughout this and the next sections is $t = t_0$.

TABLE IV. Results for the various topological observables measured at flow time t_0 on all lattices simulated.

Lattice	$\langle Q^2 \rangle$	$\langle Q^4 \rangle$	$\langle Q^4 \rangle_c$	R
A_1	0.701(6)	1.75(4)	0.273(20)	0.39(3)
B_1	1.617(6)	8.15(7)	0.30(4)	0.187(24)
C_1	2.244(6)	15.50(10)	0.40(5)	0.177(23)
D_1	3.028(6)	28.14(14)	0.63(7)	0.209(23)
E_1	3.982(6)	48.38(18)	0.81(9)	0.202(23)
F_1	5.167(6)	80.90(22)	0.81(11)	0.157(22)
B_2	1.699(7)	9.07(9)	0.41(5)	0.24(3)
D_2	3.686(14)	41.6(4)	0.83(19)	0.22(5)
B_3	1.750(7)	9.58(9)	0.39(5)	0.22(3)
D_3	3.523(13)	37.8(3)	0.56(17)	0.16(5)
B_4	1.741(7)	9.44(9)	0.35(5)	0.20(3)
D_4	3.266(12)	32.7(3)	0.68(15)	0.21(5)

as a function of the linear extension of the lattice. For $L \geq 1.4$ fm, finite-size effects turn out to be below our target statistical error of approximately 0.5%. The continuum value of $t_0^2\chi$ can thus be obtained by extrapolating

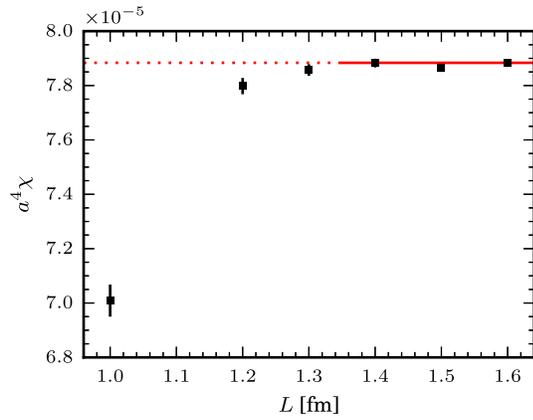
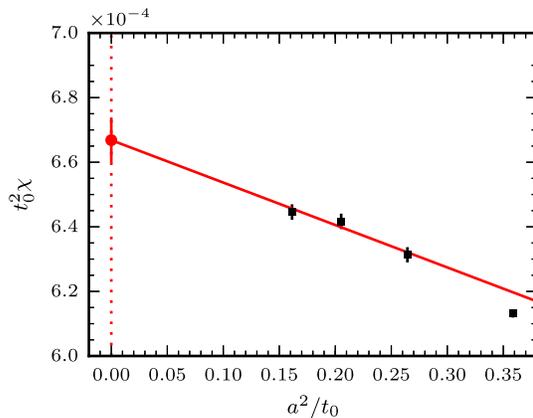


FIG. 5 (color online). Values of $a^4\chi$ as a function of L for the series $\{A_1, \dots, F_1\}$.



the results from the lattices $\{D_1, \dots, D_4\}$, see left plot of Fig. 6. The Symanzik effective theory analysis predicts discretization errors to start at $O(a^2)$, and indeed the four data points are compatible with a linear behavior in a^2 . A linear fit of all of them gives as intercept $t_0^2\chi = 6.75(4) \times 10^{-4}$ with a significance of $\chi^2/\text{dof} = 1.26$. A quadratic fit gives $t_0^2\chi = 6.49(18) \times 10^{-4}$ with $\chi^2/\text{dof} = 0.38$, and with a coefficient of the quadratic term compatible with zero within the statistical errors. By restricting the linear fit to the three points at the finer lattice spacings, we obtain

$$t_0^2\chi = 6.67(7) \times 10^{-4}, \quad (5.4)$$

with $\chi^2/\text{dof} = 0.88$, which is our best result for this quantity. It is 5 times more precise than the determination which uses the Neuberger's definition of the topological charge [14].

The cumulants of the topological charge are expected to be t -independent in the continuum limit. In the right plot of Fig. 6 we show the topological susceptibility computed at various flow times normalized to its value at t_0 . The data points have statistical errors which range from 0.1 to 1 per mille due to the correlation between the numerator and the denominator. At finite lattice spacing discretization effects are clearly visible, and they depend on t . When each set of data is extrapolated to the continuum limit with a quadratic function in a^2/t_0 , the intercepts are all compatible with 1 within the statistical errors which, depending on t , range from 0.5 to 5 per mille. We can also compare our result in Eq. (5.4) with the one obtained almost ten years ago with the Neuberger's definition of the topological charge [14]. If we use Eqs. (5.2) and (5.4), we obtain

$$r_0^4\chi = 0.0544(18), \quad (5.5)$$

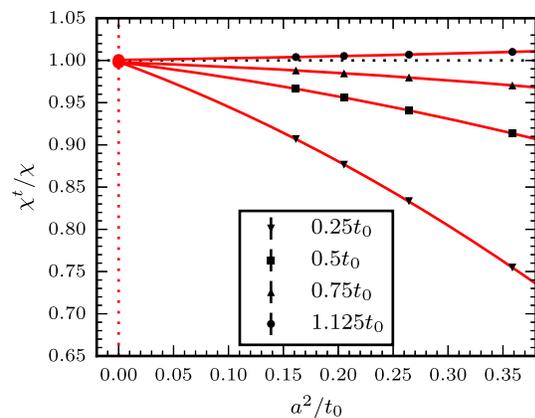


FIG. 6 (color online). Right: the dimensionless quantity $t_0^2\chi$ as a function of a^2/t_0 , and its extrapolation to the continuum limit. Left: the ratio χ^t/χ (errors are smaller than symbols) as a function of a^2/t_0 for several values of t , and its extrapolation to the continuum limit.

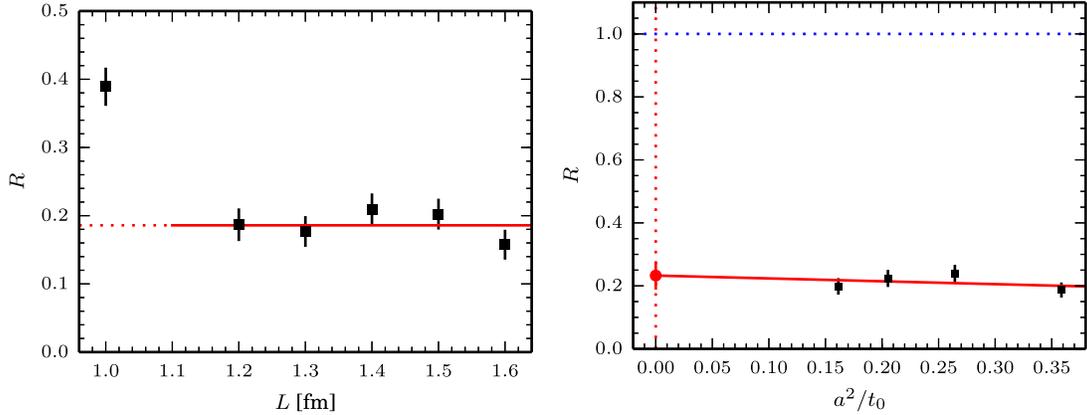


FIG. 7 (color online). Left: values of R at flow time t_0 versus L for the series $\{A_1, \dots, F_1\}$. Right: the quantity R as a function of a^2/t_0 and its extrapolation to the continuum limit; the dotted blue line is the dilute instanton gas model prediction $R = 1$.

which differs by less than 1.5 standard deviations⁷ from the result in Eq. (11) of Ref. [14]. It is interesting to note that after ten years from the first computation of χ in the continuum limit [14], we moved from an unsolved problem to a universality test at the per mille level.⁸

By using the result in Eq. (5.3), the value of χ in physical units is given by⁹

$$\chi = (180.5(5)(43) \text{ MeV})^4, \quad (5.6)$$

where the first error is statistical from Eq. (5.4), while the second is the one from the uncertainty in the scale in Eq. (5.3). If we use the physical value of t_0 determined in QCD with $N_f = 2$ and $N_f = 2 + 1$ flavours [37,38], we obtain a value of χ in physical units which differs (downwards) by 10%–20% per linear dimension. This is the size of the ambiguity which is expected when results of the Yang-Mills theory are expressed in physical units.

C. The ratio R

The values of $R = \langle Q^4 \rangle_c / \langle Q^2 \rangle^2$ from the lattices $\{A_1, \dots, F_1\}$ are shown in the left plot of Fig. 7 as a function of L . Since our target statistical error is approximately 10%, a linear extension of $L \geq 1.2$ fm is enough for finite-size effects to be within errors. Given the increase with L^8 of the computational cost of R , we have chosen to determine its continuum limit by extrapolating the data from the lattices $\{B_1, \dots, B_4\}$, see left plot of Fig. 7. Also in this case the Symanzik effective theory analysis predicts discretization errors to start at $O(a^2)$, and indeed the four

data points are compatible with a linear behavior in a^2 . A fit to a constant of all of them gives $R = 0.210(13)$ with a significance of $\chi^2/\text{dof} = 0.83$. A linear fit in a^2/t_0 gives

$$R = 0.233(45), \quad (5.7)$$

which is our best result for this quantity. The significance of the fit is $\chi^2/\text{dof} = 1.1$, and the slope is compatible with zero.

The value in Eq. (5.7) is compatible with the one obtained with the Neuberger's definition in Ref. [27], albeit with an error 2.5 times smaller. It is also relevant to note that a systematic study of finite-size effects was not carried out in Ref. [27], and finite-size effects were estimated and added to the final error.

VI. CONCLUSIONS

The θ dependence of the vacuum energy, or equivalently the functional form of the topological charge distribution, is a distinctive feature of the ensemble of gauge configurations that dominate the path integral of a Yang-Mills theory. The value of $R = 0.233(45)$ in Eq. (5.7) rules out the θ behavior predicted by the dilute instanton gas model. Its large distance from 1 implies that, in the ensemble of gauge configurations that dominate the path integral, the fluctuations of the topological charge are of quantum *nonperturbative* nature. The large N_c expansion does not provide a sharp prediction for R . Its small value, however, is compatible with being a quantity suppressed as $1/N_c^2$ in the limit of large number of colors N_c . The value of R found here is related via the Witten-Veneziano mechanism to the leading anomalous contribution to the $\eta' - \eta'$ elastic scattering amplitude in QCD. It is one of the low-energy constants which enter the effective theory of QCD when its Green functions are expanded simultaneously in powers of momenta, quark masses and $1/N_c$.

⁷This value takes into account the fact that the same determination of r_0 is used in the two computations.

⁸A first test of universality for χ was already presented in Ref. [35] with statistical errors more than 1 order of magnitude larger than those obtained here. Results with similar large statistical errors were recently obtained in Ref. [36].

⁹Note that in Ref. [14], $F_K = 113.1$ MeV was used to set the scale in physical units.

The Yang-Mills gradient flow is an extremely powerful tool for studying the topological properties of the theory. It provides a reference scale and a sensible definition of the topological charge which are cheap to be computed numerically. With a modest numerical effort by today standards, it allowed us to compute the dimensionless ratio $t_0^2\chi = 6.67(7) \times 10^{-4}$ with a relative error of roughly 1% in the continuum limit, i.e. 5 times smaller than the one of the previous reference computation with the Neuberger's definition. The Yang-Mills gradient flow is clearly an interesting tool to study the topological properties of the Yang-Mills vacuum as a function of N_c .

As proven in this paper, in the continuum limit the cumulants of topological charge defined by the Yang-Mills gradient flow coincide with those of the universal definition appearing in the chiral Ward identities. This in turn implies that this definition of the topological charge is the correct one for studying the θ dependence of the vacuum of QCD at zero and nonzero temperature. If computed in thermal (full) QCD, its cumulants can be directly related, for instance, to the axion dynamics without further renormalization.

ACKNOWLEDGMENTS

We thank M. Lüscher for several illuminating discussions which were instrumental to deriving the results in Secs. II and III and for giving us many suggestions to improve the first version of the paper. Simulations have been performed on the PC-cluster Turing and Wilson at Milano-Bicocca and on Tramontana at Pisa. We thank these institutions for the computer resources and the technical support. G. P. E. and L. G. acknowledge partial support by the MIUR-PRIN Contract No. 20093BMNPR and by the INFN SUMA project.

APPENDIX A: DEFINITION AND CONVENTIONS

The Lie algebra of the SU(3) group may be identified with the linear space of all hermitian traceless 3×3 matrices. In the basis T^a , $a = 1 \dots 8$, with

$$\text{tr}[T^a] = 0, \quad T^{a\dagger} = T^a, \quad (\text{A1})$$

the elements of the algebra are linear combinations of them with real coefficients. The structure constants f^{abc} in the commutator relation

$$[T^a, T^b] = if^{abc}T^c \quad (\text{A2})$$

are real and totally antisymmetric in the indices if the normalization condition

$$\text{tr}[T^a T^b] = \frac{1}{2}\delta^{ab} \quad (\text{A3})$$

is imposed.

For the SU(3) Yang-Mills theory the standard Wilson plaquette action is given by

$$S[U] = \frac{\beta}{2} a^4 \sum_x \sum_{\mu, \nu} \left[1 - \frac{1}{3} \text{Re tr} \{ U_{\mu\nu}(x) \} \right], \quad (\text{A4})$$

where the trace is over the color index, $\beta = 6/g_0^2$ with g_0 the bare coupling constant, a is the lattice spacing, and the plaquette is defined as a function of the gauge links $U_\mu(x)$ as

$$U_{\mu\nu}(x) = U_\mu(x) U_\nu(x + a\hat{\mu}) U_\mu^\dagger(x + a\hat{\nu}) U_\nu^\dagger(x), \quad (\text{A5})$$

with $\mu, \nu = 0, \dots, 3$, $\hat{\mu}$ is the unit vector along the direction μ and x is the spacetime coordinate.

The Neuberger-Dirac operator is defined as [1]

$$D = \frac{1}{\bar{a}} \{ 1 + \gamma_5 \text{sign}(H) \},$$

$$H = \gamma_5 (aD_w - 1 - s), \quad \bar{a} = \frac{a}{1+s}, \quad (\text{A6})$$

where s is a real parameter in the range $|s| < 1$, and D_w is the Wilson-Dirac operator. It is defined as

$$D_w = \frac{1}{2} \{ \gamma_\mu (\nabla_\mu^* + \nabla_\mu) - a \nabla_\mu^* \nabla_\mu \}, \quad (\text{A7})$$

where

$$\nabla_\mu f(x) = \frac{1}{a} \{ U_\mu(x) f(x + a\hat{\mu}) - f(x) \}, \quad (\text{A8})$$

$$\nabla_\mu^* f(x) = \frac{1}{a} \{ f(x) - U_\mu^\dagger(x - a\hat{\mu}) f(x - a\hat{\mu}) \} \quad (\text{A9})$$

are the gauge-covariant forward and backward difference operators. The Neuberger-Dirac operator satisfies the GW relation

$$\gamma_5 D + D \gamma_5 = \bar{a} D \gamma_5 D. \quad (\text{A10})$$

The link differential operators acting on functions $f(U)$ of the gauge field are

$$\partial_{x,\mu}^a f(U) = \frac{d}{ds} f(e^{-isX} U) |_{s=0},$$

$$X_\nu(y) = \begin{cases} T^a & \text{if } (y, \nu) = (x, \mu) \\ 0 & \text{otherwise} \end{cases}. \quad (\text{A11})$$

While these depend on the choice of the generators T^a , the combination

$$\partial_{x,\mu} f(U) = T^a \partial_{x,\mu}^a f(U) \quad (\text{A12})$$

can be shown to be basis-independent.

APPENDIX B: RUNGE-KUTTA-MUNTHER-KAAS INTEGRATORS

Consider an ordinary differential equation,

$$\dot{y} = f(y)y, \quad y(0) = y_0, \quad (\text{B1})$$

where $y \in G$ for some Lie group G and $f(y): G \rightarrow \mathfrak{g}$, with \mathfrak{g} being the Lie algebra of G . Runge-Kutta-Munthe-Kaas methods [28–30] are *structure-preserving* Runge-Kutta methods designed to integrate numerically these equations on the group manifold, for a general introduction see Ref. [39]. The starting point is to write the solution of (B1) as

$$y(t) = \exp\{v(t)\}y(0), \quad (\text{B2})$$

and then solve the ordinary differential equation

$$\dot{v} = \text{dexp}_v^{-1}\{f(y)\}, \quad v(0) = 0, \quad (\text{B3})$$

where dexp_v^{-1} has the series expansion

$$\text{dexp}_v^{-1} = \sum_{k=0}^{\infty} \frac{B_k}{k!} \text{ad}_v^k = 1 + \frac{1}{2}[v, \cdot] + \frac{1}{12}[v, [v, \cdot]] + \dots, \quad (\text{B4})$$

with B_k being the Bernoulli numbers and $\text{ad}_v = [v, \cdot]$ the adjoint action. Since $v(t)$ takes values in the Lie algebra, the differential equation (B3) can be numerically integrated using an ordinary RK method. No extra conditions are needed, and any RK method of a given order can be used as a base for a RKMK method of the same order. The only complication is given by the operator dexp_v^{-1} , which can be substituted with its series expansion in Eq. (B4) suitably truncated according to the order of the method. The RKMK method of q th order with s stages is given by

for $i = 1, 2, \dots, s$:

$$\begin{aligned} u_i &= \sum_{j=1}^s a_{i,j} \tilde{k}_j \\ k_i &= hf\{\exp(u_i)y_0\} \\ \tilde{k}_i &= \text{dexpinv}(u_i, k_i, q) \\ v &= \sum_{j=1}^s b_j \tilde{k}_j \\ y_1 &= \exp\{v\}y_0, \end{aligned} \quad (\text{B5})$$

where $\text{dexpinv}(u, v, q)$ is the truncated series,

$$\text{dexpinv}(u, v, q) = \sum_{k=0}^{q-1} \frac{B_k}{k!} \text{ad}_u^k, \quad (\text{B6})$$

and $a_{i,j}, b_i$ are the coefficients of the q th-order s -stages RK method. The fourth-order RKMK method that we implemented is obtained starting from the very common fourth-order 4-stages RK method with coefficients, arranged in a Butcher tableau,

$$\begin{array}{c|ccc} 0 & & & \\ \frac{1}{2} & \frac{1}{2} & & \\ \frac{1}{2} & 0 & \frac{1}{2} & \\ 1 & 0 & 0 & 1 \\ \hline & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6} \end{array}, \quad (\text{B7})$$

introduced by Kutta himself. At first sight this method entails the computation of six different commutators of k_i structures. However, it is possible to reduce the number of independent commutators needed to only two. As explained in Ref. [40], this is due to the fact that, whereas the k_i are in general $\mathcal{O}(h)$, some combinations of them are higher order in h and so the corresponding commutators can be neglected. The resulting integration algorithm is

$$\begin{aligned} u_1 &= 0, & k_i &= hf\{\exp(u_i)y_0\} \\ u_2 &= \frac{1}{2}k_1, \\ u_3 &= \frac{1}{2}k_2 + \frac{1}{8}[k_1, k_2], \\ u_4 &= k_3, \\ v &= \frac{1}{6}k_1 + \frac{1}{3}k_2 + \frac{1}{3}k_3 + \frac{1}{6}k_4 - \frac{1}{12}[k_1, k_4], \\ y_1 &= \exp(v)y_0. \end{aligned} \quad (\text{B8})$$

Alternative RK methods for integrating (B1) are given by the Crouch-Grossman integrators [41,42]. They are a special case of so-called *commutator-free* Lie group methods [43]. The third-order algorithm described in Ref. [12] belongs to this class. The conditions which the coefficients need to satisfy, order by order, are computable up to arbitrary order [44]. They are given by the order conditions for a classical RK method, plus specific extra conditions. At fourth order, however, we did not find a coefficient scheme with the useful properties of the Lüscher's integrator in terms of exponential reusing.

1. Application to the Yang-Mills gradient flow

The Yang-Mills gradient flow equation (3.1) can be written as an ordinary first-order autonomous differential equation,

$$\dot{V}(t) = Z[V(t)]V(t), \quad V(0) = V_0. \quad (\text{B9})$$

where

$$Z[V(t)] = -g_0^2 \{ \partial_{x,\mu} S[V(t)] \}, \quad (\text{B10})$$

and the link differential operators are defined in Eq. (A11). The fourth-order RKMK method in (B8) reads

$$\begin{aligned} W_1 &= V(t), & Z_i &= eZ[W_i], \\ W_2 &= \exp \left\{ \frac{1}{2} Z_1 \right\} V(t), \\ W_3 &= \exp \left\{ \frac{1}{2} Z_2 + \frac{1}{8} [Z_1, Z_2] \right\} V(t), \\ W_4 &= \exp \{ Z_3 \} V(t), \\ V(t + a^2 e) &= \exp \left\{ \frac{1}{6} Z_1 + \frac{1}{3} Z_2 + \frac{1}{3} Z_3 + \frac{1}{6} Z_4 \right. \\ &\quad \left. - \frac{1}{12} [Z_1, Z_4] \right\} V(t). \end{aligned} \quad (\text{B11})$$

This method computes 4 times the force field $Z[W_i]$ and 4 times the Lie group exponential. The commutators are economically implemented exploiting structure constants of \mathfrak{g} . Each iteration needs space in memory for one auxiliary gauge field and three Z_i fields. Gauge fields are stored in memory with a full 3×3 complex matrix, which has 18 real components, for each link. A Z_i field is an element of $\mathfrak{su}(3)$, which is an eight-dimensional linear space, for each link. Thus, the method (B11) requires space for $(18 + 3 \times 8) \times 4V$ floating point numbers. Each exponential of a \mathfrak{g} -valued combination of $Z[W_i]$ reduce to $4V$

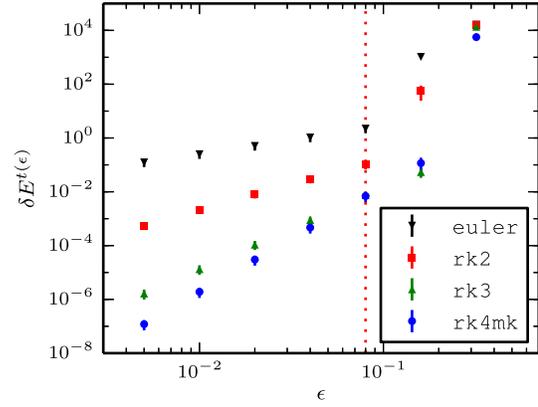


FIG. 8 (color online). Comparison of the numerical integration methods. The systematic error $\delta E^t(\epsilon)$, where $t = 3.2a^2$, is estimated from 100 configurations of the lattice B_1 by taking the difference between $E^t(\epsilon)$, evolved with step size ϵ , and $E^t(\epsilon/2)$, evolved with $\epsilon/2$.

$\mathfrak{su}(3)$ matrices exponentials, which can be computed economically exploiting the Cayley-Hamilton theorem as described in [45].

In the left plot of Fig. 8 the RKMK method is compared to lower-order Runge-Kutta methods, such as the third-order method `rk3` found in [12]. The comparison is done averaging over 100 configurations at $\beta = 5.96$ on a 12^4 lattice evolved at $t = 3.2a^2$. The `rk4mk` algorithm scales correctly as a fourth-order method. However, the prefactor appears to be larger, thus the new method is more precise with respect to `rk3` in Ref. [12] for $\epsilon \lesssim 0.1$.

-
- [1] H. Neuberger, Vector-like gauge theories with almost massless fermions on the lattice, *Phys. Rev D* **57**, 5417 (1998).
- [2] P. H. Ginsparg and K. G. Wilson, A remnant of chiral symmetry on the lattice, *Phys. Rev D* **25**, 2649 (1982).
- [3] P. Hasenfratz, V. Laliena, and F. Niedermayer, The index theorem in QCD with a finite cut-off, *Phys. Lett. B* **427**, 125 (1998).
- [4] M. Lüscher, Exact chiral symmetry on the lattice and the Ginsparg-Wilson relation, *Phys. Lett. B* **428**, 342 (1998).
- [5] L. Giusti, G. C. Rossi, M. Testa, and G. Veneziano, The $U_A(1)$ problem on the lattice with Ginsparg-Wilson fermions, *Nucl. Phys.* **B628**, 234 (2002).
- [6] L. Giusti, G. C. Rossi, and M. Testa, Topological susceptibility in full QCD with Ginsparg-Wilson fermions, *Phys. Lett. B* **587**, 157 (2004).
- [7] M. Luscher, Topological effects in QCD and the problem of short distance singularities, *Phys. Lett. B* **593**, 296 (2004).
- [8] E. Witten, Current algebra theorems for the U(1) Goldstone boson, *Nucl. Phys.* **B156**, 269 (1979).
- [9] G. Veneziano, U(1) without instantons, *Nucl. Phys.* **B159**, 213 (1979).
- [10] E. Seiler, Some more remarks on the Witten-Veneziano formula for the η' mass, *Phys. Lett. B* **525**, 355 (2002).
- [11] L. Giusti and M. Lüscher, Chiral symmetry breaking and the Banks-Casher relation in lattice QCD with Wilson quarks, *J. High Energy Phys.* **03** (2009) 013.
- [12] M. Lüscher, Properties and uses of the Wilson flow in lattice QCD, *J. High Energy Phys.* **08** (2010) 071; **03** (2014) 092(E).
- [13] M. Lüscher and P. Weisz, Perturbative analysis of the gradient flow in non-Abelian gauge theories, *J. High Energy Phys.* **02** (2011) 051.
- [14] L. Del Debbio, L. Giusti, and C. Pica, Topological Susceptibility in the SU(3) Gauge Theory, *Phys. Rev. Lett.* **94**, 032003 (2005).
- [15] A. M. Polyakov, *Gauge Fields and Strings: 3 (Contemporary Concepts in Physics)* (1987).

- [16] H. Neuberger, Exactly massless quarks on the lattice, *Phys. Lett. B* **417**, 141 (1998).
- [17] Y. Kikukawa and A. Yamada, Weak coupling expansion of massless QCD with a Ginsparg-Wilson fermion and axial U(1) anomaly, *Phys. Lett. B* **448**, 265 (1999).
- [18] K. Fujikawa, A continuum limit of the chiral Jacobian in lattice gauge theory, *Nucl. Phys.* **B546**, 480 (1999).
- [19] P. Hernandez, K. Jansen, and M. Lüscher, Locality properties of Neuberger's lattice Dirac operator, *Nucl. Phys.* **B552**, 363 (1999).
- [20] M. Lüscher, Topology and the axial anomaly in Abelian lattice gauge theories, *Nucl. Phys.* **B538**, 515 (1999).
- [21] M. Lüscher, Lattice regularization of chiral gauge theories to all orders of perturbation theory, *J. High Energy Phys.* **06** (2000) 028.
- [22] M. Lüscher, Chiral gauge theories revisited, *Subnucl. Ser.* **38**, 41 (2002).
- [23] S. Caracciolo, P. Menotti, and A. Pelissetto, One loop analytic computation of the energy momentum tensor for lattice gauge theories, *Nucl. Phys.* **B375**, 195 (1992).
- [24] B. Alles, M. D'Elia, and A. Di Giacomo, Topological susceptibility at zero and finite T in SU(3) Yang-Mills theory, *Nucl. Phys.* **B494**, 281 (1997).
- [25] M. Bruno, S. Schaefer, and R. Sommer (ALPHA Collaboration), Topological susceptibility and the sampling of field space in $N_f = 2$ lattice QCD simulations, *J. High Energy Phys.* **08** (2014) 150.
- [26] N. Cabibbo and E. Marinari, A new method for updating SU(N) matrices in computer simulations of gauge theories, *Phys. Lett. B* **119**, 387 (1982).
- [27] L. Giusti, S. Petrarca, and B. Taglienti, Theta dependence of the vacuum energy in the SU(3) gauge theory from the lattice, *Phys. Rev. D* **76**, 094510 (2007).
- [28] H. Munthe-Kaas, Lie-Butcher theory for Runge-Kutta methods, *BIT* **35**, 572 (1995).
- [29] H. Munthe-Kaas, Runge-Kutta methods on Lie groups, *BIT* **38**, 92 (1998).
- [30] H. Munthe-Kaas, High order Runge-Kutta methods on manifolds, *Applied Numerical Mathematics* **29**, 115 (1999).
- [31] U. Wolff (ALPHA Collaboration), Monte Carlo errors with less errors, *Comput. Phys. Commun.* **156**, 143 (2004).
- [32] L. Del Debbio, H. Panagopoulos, and E. Vicari, θ dependence of SU(N) gauge theories, *J. High Energy Phys.* **08** (2002) 044.
- [33] M. Guagnelli, R. Sommer, and H. Wittig (ALPHA Collaboration), Precision computation of a low-energy reference scale in quenched lattice QCD, *Nucl. Phys.* **B535**, 389 (1998).
- [34] J. Garden, J. Heitger, R. Sommer, and H. Wittig (ALPHA and UKQCD Collaboration), Precision computation of the strange quark's mass in quenched QCD, *Nucl. Phys.* **B571**, 237 (2000).
- [35] M. Lüscher and F. Palombi, Universality of the topological susceptibility in the SU(3) gauge theory, *J. High Energy Phys.* **09** (2010) 110.
- [36] K. Cichy, E. Garcia-Ramos, K. Jansen, K. Ottnad, and C. Urbach (ETM Collaboration), Non-perturbative test of the Witten-Veneziano formula from lattice QCD, [arXiv: 1504.07954](https://arxiv.org/abs/1504.07954).
- [37] S. Borsanyi *et al.*, High-precision scale setting in lattice QCD, *J. High Energy Phys.* **09** (2012) 010.
- [38] M. Bruno and R. Sommer (ALPHA Collaboration), On the N_f -dependence of gluonic observables, *Proc. Sci., LATTICE2013* (2014), 321.
- [39] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, Springer Series in Computational Mathematics (Springer, New York, 2006).
- [40] H. Munthe-Kaas and B. Owren, Computations in a free Lie algebra, *Phil. Trans. R. Soc. A* **357**, 957 (1999).
- [41] P. E. Crouch and R. Grossman, Numerical integration of ordinary differential equations on manifolds, *J. Nonlinear Sci.* **3**, 1 (1993).
- [42] B. Owren and A. Marthinsen, Runge-kutta methods adapted to manifolds and based on rigid frames, *BIT* **39**, 116 (1999).
- [43] E. Celledoni, A. Marthinsen, and B. Owren, Commutator-free Lie group methods, *Future Gener. Comput. Syst.* **19**, 341 (2003).
- [44] B. Owren, Order conditions for commutator-free Lie group methods, *J. Phys. A* **39**, 5585 (2006).
- [45] M. Lüscher, SU(3) matrix functions, <https://luscher.web.cern.ch/luscher/notes/su3fcts.pdf>, 2009.