Topological susceptibility of the 2D CP¹ or O(3) nonlinear σ model: Is it divergent or not?

Claudio Bonanno[*]

INFN Sezione di Firenze, Via G. Sansone 1, I-50019 Sesto Fiorentino, Firenze, Italy

Massimo D'Elia[†] and Francesca Margari[‡]

Università di Pisa and INFN Sezione di Pisa, Largo B. Pontecorvo 3, I-56127 Pisa, Italy

The topological susceptibility of 2D CP^{N-1} models is expected, based on perturbative computations, to develop a divergence in the limit $N \rightarrow 2$, where these models reduce to the well-known *nonlinear O(3) σ model*. The divergence is due to the dominance of instantons of arbitrarily small size and its detection by numerical lattice simulations is notoriously difficult, because it is logarithmic in the lattice spacing. We approach the problem from a different perspective, studying the behavior of the model when the volume is fixed in dimensionless lattice units, where perturbative predictions are turned into more easily checkable behaviors. After testing this strategy for $N = 3$ and 4, we apply it to $N = 2$, adopting at the same time a multicanonic algorithm to overcome the problem of rare topological fluctuations on asymptotically small lattices. Our final results fully confirm, by means of purely nonperturbative methods, the divergence of the topological susceptibility of the 2D CP¹ model.

I. INTRODUCTION

The 2D CP^{N-1} models are quantum field theories that play an important role in the study of the nonperturbative properties of gauge theories, as they share many intriguing features with 4D Yang-Mills theories, such as confinement, the existence of a nontrivial topological structure, and the related dependence on the topological parameter θ [1–3]. These theories are amenable to being treated exactly by analytic means in certain regimes but have also been extensively explored by means of Monte Carlo (MC) simulations on the lattice, since they constitute the perfect theoretical laboratory to test new numerical methods in view of an application to the more complicated physical gauge theories.

At large N, CP^{N-1} models admit a $1/N$ expansion which is similar to the 't Hooft large-N expansion of QCD. These models, however, admit an analytic solution in this regime, and the large-N limit of the vacuum energy $E(\theta)$ is well known both analytically and numerically [4–14].

An important difference between 2D CP^{N-1} models and 4D SU(N) Yang-Mills theories emerges in the opposite, small-N limit. Indeed, in the $N \rightarrow 2$ limit, where the theory becomes equivalent to the nonlinear O(3) σ model (which has been widely studied both analytically and numerically in the literature [15–38]), a pathological behavior emerges, which has no analog in the Yang-Mills case, where instead the approach from small to large N is much smoother [12,39–41].

The semiclassical picture predicts a divergence of the topological susceptibility χ for $N = 2$, which survives the renormalization procedure. Various studies have already tried to check this prediction by lattice numerical simulations, and, while there is a general consensus that the prediction is verified, the issue is not completely settled. For example, while various works about the nonlinear O(3) σ model found numerical evidence supporting that χ is divergent in the continuum limit (see, e.g., Refs. [33,37]), a recent investigation [42] from some of the authors of the present paper, considering both direct simulations at $N = 2$ and the $N \rightarrow 2$ limit of CP^{N-1} models, pointed out some difficulties in making a definite statement.

The main difficulty can be related to the fact that the divergence is of ultraviolet (UV) origin; i.e., it is related to the presence of semiclassical solutions with nonzero topological charge (instantons) at arbitrarily small scales. As a consequence, lattice studies need to check the emergence of a divergent behavior as the lattice spacing $a \rightarrow 0$: This task can be ambiguous, since the behavior

[*]claudio.bonanno@fi.infn.it
[†]massimo.delia@unipi.it
[‡]f.margari@studenti.unipi.it

could be barely distinguishable from a badly convergent behavior in a wide range of lattice spacings. For instance, even for $N = 3$ the finiteness of χ could be definitely established only recently (see, e.g., Ref. [42], where two different strategies led to consistent results).

The purpose of the present study is to develop a novel strategy, in order to make the problem better defined and reach more definite conclusions. In practice, we will approach the continuum limit, keeping the ratio between the UV and the infrared (IR) cutoffs fixed, i.e., working at fixed volume in lattice units and then considering the same procedure for different values of the dimensionless volume, a strategy which resembles some aspects of the determination of the step scaling beta function on the lattice (see, e.g., Ref. [43] for a recent review on the topic). As we will discuss in more detail in the following, within this framework the original divergent behavior of the topological susceptibility is turned into a convergent (as opposed to vanishing) continuum limit, which is much easier to check numerically, as indeed we will manage to do.

A drawback of this strategy is that one is forced to study volumes of arbitrarily small size in physical units, where topological fluctuations are extremely rare and a precise determination of the topological susceptibility could require an unfeasible statistics. This problem is easily solved by adopting a *multicanonical algorithm* [44], which has been recently employed to face the same issue in Refs. [45–48]. The general idea is to add a bias potential to the action, so that the probability of visiting suppressed topological sectors is enhanced; the MC averages with respect to the original distribution are then obtained by means of an *exact* standard reweighting procedure.

This paper is organized as follows. In Sec. II, we give a brief review about 2D CP^{N-1} models and their topological properties. In Sec. III, we describe our numerical setup and our strategy to compute the continuum limit of χ on asymptotically small lattices, including a description of the adopted multicanonical algorithm. In Sec. IV, we present and discuss our numerical results, including also an application of the same method to $N = 3$ and 4, in order to check consistency with previous results in the literature for these models [10,42,49,50]. Finally, in Sec. V, we draw our conclusions.

II. CONTINUUM THEORY

The Euclidean action of 2D CP^{N-1} models can be written in terms of a matter field $z(x)$, a complex N-component scalar field satisfying $\bar{z}(x)z(x) = 1$, and of an auxiliary nonpropagating U(1) gauge field A_μ. In the presence of the topological term, the action reads

$$S(\theta) = \int d^2x \left[\frac{N}{g} \bar{D}_\mu \bar{z}(x) D_\mu z(x) - i\theta q(x) \right], \quad (1)$$

where g is the 't Hooft coupling, $D_\mu = \partial_\mu + iA_\mu$ is the U(1) covariant derivative, and

$$Q = \int d^2x q(x) = \frac{1}{2\pi} \epsilon_{\mu\nu} \int d^2x \partial_\mu A_\nu(x) \in \mathbb{Z} \quad (2)$$

is the integer-valued topological charge.

The θ-dependent vacuum energy density, using the path-integral formulation of the theory, is given by

$$E(\theta) = -\frac{1}{V} \log \int [d\bar{z}][dz][dA] e^{-S(\theta)}, \quad (3)$$

where V is the 2D space-time volume. Assuming that $E(\theta)$ is an analytic function of θ around $\theta = 0$, one can Taylor expand it around this point; at leading order, one has [3,12]

$$E(\theta) - E(0) = \frac{1}{2}\chi\theta^2 + O(\theta^4), \quad (4)$$

where χ is the topological susceptibility:

$$\chi = \frac{1}{V} \langle Q^2 \rangle \Big|_{\theta=0}. \quad (5)$$

To better understand the origin of the divergence of χ in the $N \to 2$ limit, it is useful to recall that, in the semi-classical approximation, the path integral is evaluated by integrating fluctuations around instanton solutions, and it is reduced to an ordinary integral of the instanton density. At leading order, the instanton density of CP^{N-1} models is given, as a function of the instanton size, by [51]

$$d_I(\rho) \propto \rho^{N-3}. \quad (6)$$

For $N = 2$, $d_I(\rho) \sim 1/\rho$; i.e., it develops an UV divergence for $\rho \to 0$. This means that the divergence of χ in this case can be traced back to the proliferation of small-size instantons with vanishing size $\rho \to 0$, whose density grows proportionally to $1/\rho$.

III. NUMERICAL METHODS

In this section, we discuss various aspects related to the discretization of the models and of the observables, in particular, those related to topology and to the employed numerical strategies.

A. Discretization details

We discretized space-time through a square lattice with L^2 sites and periodic boundary conditions and the $\theta = 0$ continuum action (1) through the tree-level Symanzik-improved lattice action [9]

$$S_L = -2N\beta_L \sum_{x,\mu} \{c_1 \mathrm{Re}[\bar{U}_\mu(x)\bar{z}(x+\hat{\mu})z(x)]$$
$$+ c_2 \mathrm{Re}[\bar{U}_\mu(x+\hat{\mu})\bar{U}_\mu(x)\bar{z}(x+2\hat{\mu})z(x)]\}, \quad (7)$$

where $c_1 = 4/3$ and $c_2 = -1/12$ are improvement coefficients, $\beta_L \equiv 1/g_L$ is the inverse bare coupling, $z(x)$ are the matter fields, satisfying $\bar{z}(x)z(x) = 1$, and $U_\mu(x)$ are U(1) gauge link variables. Symanzik improvement cancels out logarithmic corrections to the leading continuum scaling [52], improving convergence toward the continuum limit.

In this limit, approached taking $\beta_L \to \infty$, a vanishing lattice spacing $a \to 0$ can be traded for a divergent lattice correlation length $\xi_L \equiv \xi/a \underset{a\to 0}{\sim} 1/a$. In order to fix a, in this work we chose the second moment correlation length ξ, defined in the continuum theory as

$$\xi^2 \equiv \frac{1}{\int G(x)d^2x} \int G(x)\frac{|x|^2}{4}d^2x, \quad (8)$$

where $G(x)$ denotes the two-point connected correlation function of the projector $P_{ij}(x) \equiv z_i(x)\bar{z}_j(x)$:

$$G(x) \equiv \langle P_{ij}(x)P_{ij}(0)\rangle - \frac{1}{N}. \quad (9)$$

A lattice discretization of Eq. (8) can be obtained from the Fourier transform $\tilde{G}_L(p)$ of $G_L(x)$, which is the lattice counterpart of Eq. (9) [53]:

$$\xi_L^2 = \frac{1}{4\sin^2(\pi/L)}\left[\frac{\tilde{G}_L(0,0)}{\tilde{G}_L(2\pi/L,0)} - 1\right]. \quad (10)$$

B. Topology on the lattice and smoothing

There are several possible discretizations Q_L of the topological charge (2), all yielding the same continuum limit for the topological susceptibility and other quantities relevant to θ dependence, when discretization effects are properly taken care of. Generally speaking, lattice definitions are related to the continuum one by [54,55]

$$Q_L = Z_Q(\beta_L)Q, \quad (11)$$

where $Z_Q(\beta_L)$ is a finite multiplicative renormalization factor. For this reason, lattice discretizations of Q are, in general, not integer valued. The most simple discretization can be defined in terms of the plaquette $\Pi_{\mu\nu}(x) \equiv U_\mu(x)U_\nu(x+\hat{\mu})\bar{U}_\mu(x+\hat{\nu})\bar{U}_\nu(x)$ as

$$Q_{\mathrm{plaq}} = \frac{1}{2\pi}\sum_x \mathrm{Im}[\Pi_{12}(x)]. \quad (12)$$

However, it is possible to work out *geometric* discretizations of the topological charge [9,16], which always

result in integer values for every configuration, i.e., definitions with $Z_Q = 1$. In particular, we adopted the geometric definition that can be built from the link variables $U_\mu(x)$ [9]:

$$Q_U = \frac{1}{2\pi}\sum_x \mathrm{Im}\{\log[\Pi_{12}(x)]\} \in \mathbb{Z}. \quad (13)$$

Although Q_U has $Z_Q = 1$, renormalization effects are still present when computing χ because of *dislocations* [10,56]. Dislocations are UV fluctuations of the background gauge field that make establishing the winding number of the configuration ambiguous. The net effect is that dislocations result in an additive renormalization when computing the lattice topological susceptibility [57,58]. Such renormalization diverges in the continuum limit and, thus, must be removed.

Dislocations being the result of UV fluctuations at the scale of the lattice spacing, computing the geometric charge on smoothed configurations is sufficient to remove their unphysical contribution while preserving the background topological structure of the gauge fields. Indeed, smoothing brings a configuration closer to a local minimum of the action, thus dumping UV fluctuations while, at the same time, preserving the physical topological signal.

Many different smoothing algorithms have been proposed in the literature, such as stout smearing, gradient flow, or cooling, all giving consistent results when properly matched with each other (see Refs. [59,60] for more details). For this reason, we chose cooling for its numerical cheapness. This method consists in a sequence of n_{cool} steps in which the configuration approaches a local minimum of the action by iteratively aligning both link variables $U_\mu(x)$ and site variables $z(x)$ to their relative local force. Since the choice of the action that is locally minimized during cooling is irrelevant [60], we adopted the unimproved one for this purpose, meaning that the local forces along which the $U_\mu(x)$ and $z(x)$ fields are aligned are computed from the action in Eq. (7) with $c_1 = 1$ and $c_2 = 0$. In the end, thus, we define

$$Q_L = Q_U^{(\mathrm{cool})},$$
$$a^2\chi = \frac{\langle Q_L^2\rangle}{L^2}. \quad (14)$$

It is worth mentioning that smoothing methods act as diffusive processes, thus modifying the UV behavior of the fields below a *smoothing radius* r_s which is proportional to the square root of the amount of smoothing performed (e.g., to $\sqrt{n_{\mathrm{cool}}}$ in our case). When χ is finite, the choice of n_{cool} is not critical, because the physical topological signal is well separated from the length scale r_s introduced by the smoothing procedure. As a consequence, in such cases χ exhibits a plateau upon increasing n_{cool} above a certain

threshold, and no residual dependence on n_{cool} is observed on continuum-extrapolated results.

The pathological case of the CP^1 model is, instead, different in this respect, since we exactly aim at probing the sensitivity of χ to the contribution of small instantons, which are, however, smoothed away below r_s [see Ref. [33], where the dependence of the topological susceptibility of the nonlinear O(3) σ model on the gradient flow time is discussed]. In particular, in our setup where $L = l/a$ is kept fixed as $a \to 0$, the quantity $\sqrt{n_{cool}}/L = r_s/l$ is a relevant parameter, and we expect the continuum limit of χ to depend on its value. Thus, we will extrapolate our results toward $r_s/l \to 0$ in order to ensure that no relevant contribution coming from small length scales is lost.

Concerning updating algorithms, CP^{N-1} models at small N do not require any particular strategy to decorrelate the topological charge [42], contrary to the large-N case, which is plagued by a severe *topological critical slowing down* [13,14,39,61–64]. As a matter of fact, CP^{N-1} models at small N are dominated by small instantons, which are more easily decorrelated by means of local field updates. Thus, we will adopt local updating algorithms such as over-relaxation (OR) and over-heat-bath (HB) [9]. An issue is, however, represented by the dominance of the $Q = 0$ sector on small physical volumes, which is discussed in more detail in Sec. III E.

C. Continuum limit at fixed volume in lattice units

The expectation value of a generic observable $\mathcal{O}$ scales toward the continuum limit according to

$$\langle \mathcal{O} \rangle_L(\xi_L) = \langle \mathcal{O} \rangle_{cont} + c\xi_L^{-2} + o(\xi_L^{-2}), \qquad (15)$$

where finite lattice spacing corrections to continuum scaling are expressed as inverse powers of $1/\xi_L$. However, the continuum scaling of topological observables is modified at small N, due to the presence of small-size topological fluctuations.

Such modifications have been worked out in Ref. [42], assuming the perturbative computation of the instanton size distribution $d_I(\rho) \propto \rho^{N-3}$ and that topological fluctuations are dominated by a noninteracting gas of small-size instantons and anti-instantons. Under these assumptions, the number of (anti-)instantons n_I (n_A) is distributed as a Poissonian with $\langle n_I \rangle = \langle n_A \rangle \propto l^2 \int_{\rho_{min}}^{\rho_{max}} \rho^{N-3} d\rho$, where the integral is taken from a UV scale ρ_{min}, proportional to the lattice spacing a, to an IR scale ρ_{max}, proportional to the correlation length ξ. Then,

$$\langle Q^2 \rangle \propto \langle (n_I - n_A)^2 \rangle = 2\langle n_I \rangle \propto l^2 \int_{\rho_{min}}^{\rho_{max}} \rho^{N-3} d\rho, \quad (16)$$

and, thus,

$$\xi^2 \chi = \xi^2 \frac{\langle Q^2 \rangle}{l^2} \propto \begin{cases} \rho_{max}^{N-2} - \rho_{min}^{N-2}, & (N > 2), \\ \log(\frac{\rho_{max}}{\rho_{min}}), & (N = 2). \end{cases} \quad (17)$$

From Eq. (17), taking into account that $\rho_{min} \propto a$ and $\rho_{max} \propto \xi$, one can predict the following behaviors for the topological susceptibility when the continuum limit is approached at fixed physical lattice volume $V = l^2$ (hence, on lattices satisfying $L/\xi_L \gg 1$):

$$\xi^2 \chi(x) = A_2 \log(B_2 x) + C_2 x^2 + O(x^4) \quad (N = 2),$$
$$\xi^2 \chi(x) = A_3 + B_3 x + C_3 x^2 + O(x^4) \qquad (N = 3),$$
$$\xi^2 \chi(x) = A_N + C_N x^2 + O(x^4) \qquad (N \geq 4),$$

where $x = 1/\xi_L \propto a$. Hence, for the CP^1 model, the divergence of the topological susceptibility should appear as a logarithm of the UV cutoff $1/a$, which may be difficult to distinguish from a regular power-law behavior in a.

To overcome this issue, we investigate the continuum limit of χ in the small-N limit performing lattice simulations at fixed volume in lattice units, i.e., fixing $L = l/a$. Using this approach, we have the following predictions:

$$\xi^2 \chi = \xi^2 \frac{\langle Q^2 \rangle}{l^2} \propto \begin{cases} a^{N-2}[(\frac{L}{R})^{N-2} - 1] & (N > 2), \\ \log(\frac{L}{R}) & (N = 2), \end{cases} \quad (18)$$

where R is an effective parameter accounting for the ratio between the maximum and the minimum instanton sizes which can live on the same lattice, which is expected in this case to be proportional to $l = aL$ (since $L \ll \xi_L$), i.e., $\rho_{max}/\rho_{min} \equiv L/R$, with R independent of L. We stress that Eq. (18) has been obtained by multiplying $\langle Q^2 \rangle/l^2$ for the squared correlation length ξ^2 obtained on large physical volumes (i.e., in the limit $L/\xi_L \gg 1$), so that the only dependence on L of the continuum limit of $\xi^2 \chi$ comes from χ alone.

These semiclassical considerations point out that, when the continuum topological susceptibility computed on lattices with $L/\xi_L \gg 1$ is finite, the continuum limit of χ at fixed L is expected to vanish as a for $N = 3$ or as a^2 for $N = 4$; cf. Eq. (18). This is due to the fact that, when $a \to 0$ at fixed L, the physical lattice size vanishes proportionally to a, and any topological fluctuation on physical scales disappears.

On the other hand, if the continuum limit of χ taken at fixed $L/\xi_L \gg 1$ is logarithmically divergent, as predicted by semiclassical computations, we expect to approach a constant and finite value for χ when, instead, the continuum limit is taken at fixed L; cf. again Eq. (18). In this case, topological fluctuations damped because of the decreasing IR cutoff are exactly balanced, as $a \to 0$, by new topological fluctuations appearing at arbitrarily small UV scales. This means that, with this strategy, the divergent continuum limit of χ predicted by semiclassical

computations is mapped into a nonvanishing finite continuum limit, which should be more amenable to be tested by numerical methods.

D. Determining ξ_L close to the continuum limit

As we stressed above, the values of ξ_L needed for our determination of $\xi^2 \chi = \xi_L^2 a^2 \chi$ are those that would be obtained in the infinite volume limit, i.e., on lattices of size L such that $L/\xi_L \gg 1$. This is barely feasible within the framework of our numerical strategy, aimed at reaching very large values of ξ_L but keeping L fixed; i.e., we would need to perform additional simulations on unfeasible large lattices for a reliable numerical determination of ξ_L.

To overcome this problem, we will first look for the onset of the asymptotic scaling region where ξ_L scales as predicted by the two-loop perturbative beta function [2]:

$$-a \frac{d\beta_L^{-1}}{da} = -\frac{\beta_L^{-2}}{2\pi} \left(1 + \frac{\beta_L^{-1}}{\pi N}\right) \qquad (19)$$

and then make use of such scaling to extend the determination of ξ_L within this region.

Integrating Eq. (19) to obtain the running of the quantity $2\pi\beta_L(a)$, it is possible to obtain the dynamically generated scale of the lattice theory (7) $\Lambda_L^{(\text{Sym})}$ in lattice units [9]:

$$\Lambda_L^{(\text{Sym})} = \frac{1}{a}[(2\pi\beta_L)^{2/N}\exp\{-2\pi\beta_L\}] \equiv \frac{1}{a}f(\beta_L). \quad (20)$$

The latter equation can be turned into a perturbative expression for the mass gap $M \equiv \xi^{-1}$ by multiplying both sides for ξ:

$$\Lambda_L^{(\text{Sym})}/M = \xi_L f(\beta_L). \qquad (21)$$

Equation (21) being the result of a perturbative computation, we expect the ratio $M/\Lambda_L^{(\text{Sym})}$ to approach a constant value plus $O(1/\beta_L)$ corrections in the asymptotic region $\beta_L \to \infty$. Assuming such corrections to be negligible, Eq. (21) allows one to compute ξ_L at arbitrarily large values of the bare coupling once its value $\xi_L^\star$ for a certain coupling $\beta_L^\star$ is fixed:

$$\xi_L(\beta_L) = f(\beta_L^\star) \frac{\xi_L^\star}{f(\beta_L)}. \qquad (22)$$

In the following, ξ_L will be first determined numerically on large lattices (satisfying $L/\xi_L \gg 1$) and for a feasible range of values of β_L; then, by matching results to asymptotic scaling prediction in Eq. (21), we will choose a $\beta_L^\star$ for which Eq. (22) is reliable and determine ξ_L accordingly for larger values of β_L. More details about our choice of $\beta_L^\star$ and on the check of the stability of ξ_L varying this choice can be found in Appendix A.

E. Dominance of the $Q = 0$ sector and multicanonical algorithm

Another drawback of working at fixed L is the dominance of the $Q = 0$ sector, which introduces the necessity of collecting unfeasible statistics to achieve a precise computation of the topological susceptibility on asymptotically small lattice volumes. In order to better clarify this statement we remark that, according to Eq. (18), even if χ diverges for $N = 2$ as expected from the semiclassical approximation, $\langle Q^2 \rangle$ is expected to vanish, for fixed L, as $1/\xi_L^2$ in the continuum limit. If $\langle Q^2 \rangle \ll 1$, then $P(Q = 0) \gg P(|Q| = 1) \gg P(|Q| = 2) \gg \cdots$; in this regime, the variance of the topological charge distribution $P(Q)$ can be approximated as

$$\langle Q^2 \rangle = V\chi \simeq \frac{P(|Q| = 1)}{P(Q = 0)}; \qquad (23)$$

i.e., to compute χ, we need to estimate with great precision a vanishing probability to visit $|Q| = 1$ sectors. This requires a growing and unfeasible numerical effort, since we need a sufficient number of fluctuations of Q to obtain $\langle Q^2 \rangle$ with a given target precision.

In order to overcome this problem, we will adopt the multicanonical algorithm. This approach was recently employed in the context of 4D gauge theories to enhance topological fluctuations at finite temperature; see, e.g., Refs. [45,47]. The main idea behind the multicanonic approach is to modify the probability distribution of the topological charge $P(Q) \to P_{\text{mc}}(Q) = P(Q)w(Q)$, where $w(Q)$ is a known Q-dependent weight function, in order to enhance the probability of visiting suppressed topological sectors. Since the relative error on Eq. (23) scales as the inverse of the square root of the number of $|Q| = 1$ events $\sim N_{\text{meas}}P(|Q| = 1)$, enhancing $P(|Q| = 1)$ with respect to $P(Q = 0)$ by a known factor of w_1/w_0 reduces the relative error on χ by a factor of $\sqrt{w_1/w_0}$.

In analogy with lattice QCD simulations [45–47], we introduce the weights $w(Q)$ by adding a topological potential $V_{\text{topo}}(Q_{\text{mc}})$ to the lattice action:

$$S_L \to S_L + V_{\text{topo}}(Q_{\text{mc}}) \Rightarrow w(Q) = e^{-V_{\text{topo}}(Q_{\text{mc}})}, \quad (24)$$

where Q_{mc} is a suitable discretization of the topological charge, which does not necessarily need to coincide with the one that is used to measure it.

Expectation values with respect to the original distribution are then *exactly* recovered by the following standard reweighting procedure:

$$\langle O \rangle = \frac{\langle O e^{V_{\text{topo}}(Q_{\text{mc}})} \rangle_{\text{mc}}}{\langle e^{V_{\text{topo}}(Q_{\text{mc}})} \rangle_{\text{mc}}}. \qquad (25)$$

We stress that the relation in Eq. (25) among expectation values computed with and without the bias potential is

exact; thus, any choice for $V_{\text{topo}}(Q_{\text{mc}})$ will in the end give the correct result for $\langle O \rangle$. Therefore, this strategy does not introduce any further source of uncertainty. For this reason, the discretization of the topological charge Q_{mc} and the bias potential V_{topo} can be chosen with some arbitrariness.

However, the choice of $V_{\text{topo}}(Q_{\text{mc}})$ can affect the efficiency of the algorithm, and, in particular, one would like to avoid possible issues due to a poor overlap between the starting and the biased path-integral distributions. This could happen, e.g., if V_{topo} is too strong. In that case, a sort of spontaneous breaking of the CP symmetry occurs [47], meaning that configurations with $Q \neq 0$ occur with overwhelming frequency with respect to $Q = 0$ ones and that $\langle Q \rangle \neq 0$, thus disrupting importance sampling and leading to uncontrolled effects on the correct estimation of statistical errors in the evaluation of the ratio of expectation values in Eq. (25).

However, such pathological cases can be easily avoided by tuning the topological potential through short test runs, ensuring that importance sampling is not disrupted. In particular, if the MC evolution of the topological charge is still dominated by the $Q = 0$ sector, the symmetry properties of the distribution are preserved (i.e., $\langle Q \rangle = 0$), and $Q \neq 0$ sectors (which are those giving contribution to the averages of interest) are explored more frequently, and then reweighted, which enhances (rather than disrupts) importance sampling for the observables of interest. The estimate of the statistical error on Eq. (25), which proceeds usually through a bootstrap analysis, will then be reliable if the number of tunnelings in and out of the $Q = 0$ topological sector is statistically significant, as is always the case in our simulations. The tuning of the potential was done following the procedure outlined in Ref. [47]. More details about our choices for Q_{mc} and V_{topo} and about our implementation of the multicanonical algorithm can be found in Appendix B.

IV. NUMERICAL RESULTS

In this section, we will first discuss results for the topological susceptibility for $N = 4$ and $N = 3$, showing that our strategy gives compatible results with previous findings in the literature [10,42,49,50]. Then, we show the behavior of $\xi^2 \chi$ in the continuum limit at fixed L for $N = 2$, established adopting the multicanonic algorithm. Finally, we conclude our study by comparing results achieved at fixed L with those obtained at fixed physical volume ($L/\xi_L \gg 1$).

A. Results for $N = 4$ and $N = 3$

In order to calibrate our strategy, we first consider the cases $N = 4$ and 3, for which we expect from semiclassical computations, and we actually know from previous lattice results [10,42,49,50], that the topological susceptibility is finite.

Therefore, according to Eq. (18), we expect to observe a vanishing continuum limit

$$\xi^2 \chi(x) \underset{x \to 0}{\sim} x^c, \qquad x = 1/\xi_L, \qquad (26)$$

where $c = 2$ for $N = 4$ and $c = 1$ for $N = 3$.

Following the strategy discussed in Sec. III C, we performed lattice simulations, keeping the volume fixed in lattice units on lattices with $L = 50$, exploring several values of β_L, and reaching values of ξ_L of the order of $\sim 10^3$. Our MC updating step in this case consisted of four lattice sweeps of OR and one lattice sweep of HB updating steps: In the following, we will simply call this combination "standard MC step." The computation of the topological susceptibility in lattice units via Eq. (14) was performed every ten MC steps and after $n_{\text{cool}} = 50$ cooling steps, while ξ_L was computed via Eq. (22).

In Table I, we report a complete summary of the parameters of the performed simulations for $N = 4$ and 3 along with the generated statistics and the obtained results for ξ_L, $a^2 \chi$, and $\xi^2 \chi$.

TABLE I. Summary of simulation parameters and results obtained for $N = 4$, 3 and $L = 50$. The correlation length ξ_L is computed according to Eq. (22) with $\beta_L^\star = 1.35$ and 1.455 for $N = 4$, 3, respectively (see Appendix A for more details). Reported values of χ are computed after $n_{\text{cool}} = 50$ cooling steps. Statistics is expressed in millions (M), and measures are taken every ten standard MC steps ($= 4$ OR $+1$ HB lattice updating sweeps).

| N	β_L	$\xi_L \times 10^{-3}$	$a^2\chi \times 10^9$	$\xi^2\chi \times 10^3$	Statistics
4	1.35	0.08262(70)	283.0(2.7)	1.932(38)	52M
	1.40	0.11108(94)	89.3(1.5)	1.102(27)	
	1.45	0.1494(13)	28.55(87)	0.638(22)	
	1.50	0.2012(17)	7.80(43)	0.316(18)	
	1.55	0.2709(23)	2.88(27)	0.211(20)	
	1.60	0.3651(31)	0.98(17)	0.130(22)	
	1.65	0.4922(42)	0.42(11)	0.102(27)	
	1.70	0.6639(56)	0.107(54)	0.047(24)	
	1.75	0.8959(76)	0.031(19)	0.025(15)	
3	1.50	0.12765(68)	529.1(4.1)	8.62(11)	25M
	1.55	0.17099(92)	223.8(2.7)	6.54(10)	
	1.60	0.2292(12)	90.8(1.7)	4.77(10)	
	1.65	0.3074(16)	38.4(1.1)	3.62(11)	
	1.70	0.4126(22)	15.66(73)	2.67(13)	
	1.75	0.5541(30)	6.87(47)	2.11(15)	
	1.80	0.7445(40)	2.78(15)	1.54(8)	102M
	1.85	1.0009(54)	1.18(10)	1.18(10)	
	1.90	1.3461(72)	0.501(76)	0.91(14)	76M
	1.95	1.8114(97)	0.146(37)	0.48(12)	
	2.00	2.438(13)	0.083(40)	0.50(24)	
	2.05	3.284(18)	0.025(11)	0.27(12)	128M
	2.10	4.424(24)	0.0156(83)	0.31(16)	
	2.15	5.963(32)	0.0094(70)	0.33(25)	
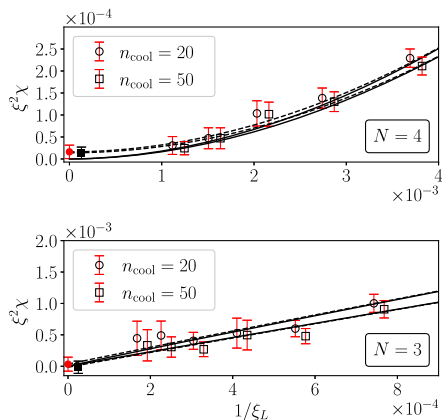
FIG. 1. Continuum extrapolation of $\xi^2\chi$ for $N = 4$ (top) and $N = 3$ (bottom). Solid and the dashed lines represent, respectively, best fits obtained using fit function (27) setting $a_0 = 0$ and treating it as a free parameter. Determinations obtained for different values of n_{cool} have been slightly shifted to improve readability. Full points in $1/\xi_L = 0$ represent continuum-extrapolated determinations.

We start our discussion from the CP³ model. We extrapolated the quantity $\xi^2\chi$ toward the continuum limit, fitting the ξ_L dependence of $\xi^2\chi$ according to the fit function

$$f(x) = a_0 + a_1 x^c, \qquad x = 1/\xi_L, \qquad (27)$$

where c is a free exponent.

In order to check that the continuum limit is indeed vanishing, we considered two cases: the case when a_0 is treated as a free parameter and the one when a_0 is fixed to zero. In the former case, our data turn out to be well compatible with a vanishing continuum limit, as the best fit yields $\tilde{\chi}^2/\mathrm{d.o.f.} = 6.6/6$ and $a_0 = 1.4(1.3) \times 10^{-5}$, which is compatible with zero within its statistical error. Moreover, the exponent $c = 1.96(6)$ turns out to be compatible with 2, which is in agreement with the semiclassical prediction in Eq. (26) and with results of Ref. [42]. Also, fixing $a_0 = 0$ gives a very good description of our data, as the best fit gives $\tilde{\chi}^2/\mathrm{d.o.f.} = 7.8/7$ and a compatible exponent $c = 1.91(4)$. In Fig. 1, we show such continuum extrapolations for $N = 4$ considering all available determinations in Table I.

Varying the fit range, the value of n_{cool} used to compute Q_L, or the coupling $\beta_L^\star$ used to fix ξ_L did not result in any appreciable variation of our final results. We can, thus, conclude that our findings are perfectly compatible with previous results in the literature pointing out a finite continuum limit for $\chi(N = 4)$ (see, e.g., Refs. [10,42,50]).

We now repeat the same analysis for $N = 3$. The continuum extrapolation of $\xi^2\chi$ data for $N = 3$ according to fit function (27) is depicted in Fig. 1. The best fit in the whole available range yields $a_0 = (-2 \pm 10) \times 10^{-5}$, $c = 0.98(4)$, and $\tilde{\chi}^2/\mathrm{d.o.f.} = 4.4/11$. Also, performing

the best fit fixing $a_0 = 0$ perfectly describes our data, giving $c = 0.99(2)$ and $\tilde{\chi}^2/\mathrm{d.o.f.} = 4.4/12$.

Again, we observe no dependence of our continuum-extrapolated results on the choice of n_{cool}, of $\beta_L^\star$, or of the fit range. Therefore, also in this case, our strategy gives compatible results both with semiclassical expectations and with previous numerical results in the literature for $\chi(N = 3)$ (see, e.g., Refs. [42,49]).

B. Results for the topological susceptibility of the CP¹ model from the multicanonical algorithm

In order to precisely assess the continuum behavior of $\xi^2\chi(N = 2)$, we pushed our investigation on the $L = 50$ lattice up to ξ_L as large as $\sim 10^6$. Reliably computing the susceptibility for such fine lattice spacings is an unfeasible task with standard methods due to the dominance of the $Q = 0$ sector previously explained, while it was made possible by the adoption of the multicanonical algorithm, which allowed us to largely improve the number of fluctuations of Q_L observed during MC simulations. Illustrative examples for $\beta_L = 2.50$ and 3.00 are shown in Fig. 2.

This has, in turn, allowed us to largely reduce the computational power needed to determine χ with a given
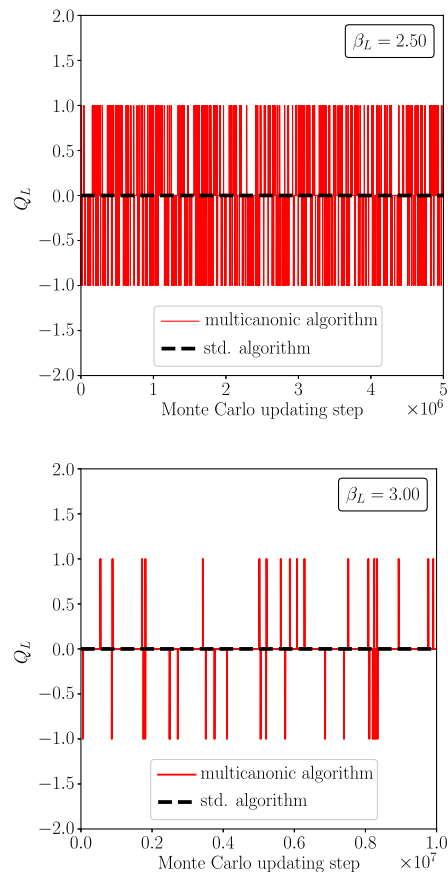

FIG. 2. Evolution of the geometric charge Q_L computed after $n_{\mathrm{cool}} = 50$ cooling steps for $N = 2$ obtained with the standard and the multicanonical algorithm.
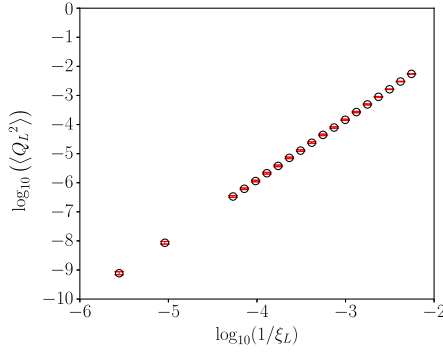
FIG. 3. Behavior of $\langle Q_L^2 \rangle$, measured after $n_{\mathrm{cool}} = 50$ cooling steps, as a function of $1/\xi_L$ for $N = 2$ and $L = 50$.

precision. As an example, let us consider our largest β_L, for which $\langle Q^2 \rangle \sim 10^{-9}$ (cf. Fig. 3). Using Eq. (23) and assuming that the error on χ scales as the inverse of the square root of $N_{\mathrm{meas}} P(|Q| = 1)$, we can estimate that, to reach the same $\sim 2\%$ relative error on the susceptibility achieved with the multicanonical algorithm, with the standard algorithm we would have needed a statistics larger by about a factor of ~ 100; i.e., we gained 2 orders of magnitude in terms of computational power.

For this reason, we adopted the multicanonic algorithm for $\beta_L \geq 2.20$, i.e., for $\xi_L \gtrsim 3 \times 10^3$, where $\langle Q^2 \rangle \lesssim 10^{-5}$. A summary of all the performed simulation and the obtained results for $N = 2$ is reported in Table II.

To extrapolate our finite-ξ_L determinations toward the continuum limit, we consider again the fit function ansatz in Eq. (27), and we perform a best fit of all available data for $\xi^2 \chi$ as a function of $1/\xi_L$, considering both fixed $a_0 = 0$ and a_0 as a free parameter.

While in the latter case such a best fit provides a very good description of our numerical results, giving

$$a_0 = 0.031(2),$$

$$c = 0.20(2),$$

$$\tilde{\chi}^2/\mathrm{d.o.f.} = 8.0/16,$$

the best fit performed fixing $a_0 = 0$ yields a $\tilde{\chi}^2/\mathrm{d.o.f.} = 42.7/17$, thus clearly providing a bad description of our data. Narrowing the fit range by, e.g., excluding the point at the smallest value of ξ_L ($\beta_L = 1.70$) does not improve the result, as we still obtain $\tilde{\chi}^2/\mathrm{d.o.f.} = 32.6/16$. On the other hand, the quality of the fit with a_0 free remains very good, as excluding the point for our smallest ξ_L yields $a_0 = 0.031(3)$ with $\tilde{\chi}^2/\mathrm{d.o.f.} = 7.5/15$. A comparison between the continuum limits taken at fixed L in the whole available range is displayed in Fig. 4.

It is interesting to observe that the best fit with a_0 free yields $\tilde{\chi}^2/\mathrm{d.o.f.} \simeq 0.5$, i.e., smaller than 1. A possible explanation is that, all values of ξ_L being obtained from the same two-loop scaling equation, results for $\xi^2 \chi$ at

TABLE II. Summary of simulation parameters and results obtained for $N = 2$ and $L = 50$. The correlation length ξ_L is computed according to Eq. (22) with $\beta_L^\star = 1.70$ (see Appendix A for more details). Reported values of χ are computed after $n_{\mathrm{cool}} = 50$ cooling steps. Statistics is expressed in millions or billions (M or G, respectively), and measures are taken every ten standard MC steps ($= 4$ OR $+1$ HB lattice updating sweeps) or every ten multicanonic steps (see Appendix B for more details).

| N	Algorithm	β_L	$\xi_L \times 10^{-3}$	$a^2 \chi \times 10^9$	$\xi^2 \chi \times 10^3$	Statistics
2	Standard	1.70	0.17991(78)	2207.4(4.8)	71.45(64)	51M
		1.75	0.2393(10)	1205.7(3.6)	69.03(63)	
		1.80	0.3185(14)	652.6(2.7)	66.21(64)	
		1.85	0.4243(18)	359.0(2.0)	64.62(66)	
		1.90	0.5656(25)	196.7(1.5)	62.92(72)	
		1.95	0.7545(33)	107.6(1.1)	61.24(82)	
		2.00	1.0072(44)	58.15(57)	58.99(77)	102M
		2.05	1.3453(58)	31.59(34)	57.17(79)	153M
		2.10	1.7980(78)	17.63(20)	57.00(82)	256M
		2.15	2.404(10)	9.50(12)	54.90(83)	410M
	Multicanonic	2.20	3.217(14)	5.086(59)	52.64(76)	81M
		2.25	4.307(19)	2.864(35)	53.12(79)	133M
		2.30	5.768(25)	1.516(18)	50.43(75)	266M
		2.35	7.729(34)	0.849(11)	50.73(79)	595M
		2.40	10.362(45)	0.4576(56)	49.13(74)	566M
		2.45	13.897(60)	0.2475(30)	47.80(71)	640M
		2.50	18.645(81)	0.1346(16)	46.78(70)	1G
		2.80	109.643(48)	0.003448(68)	41.45(89)	7G
		3.00	359.6(1.6)	0.0003093(69)	39.98(96)	16G
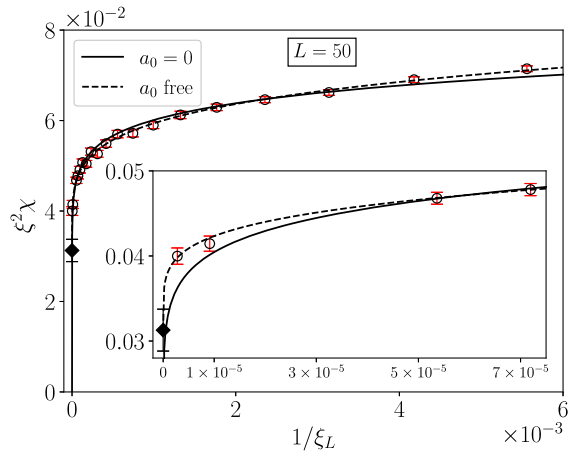
FIG. 4. Continuum extrapolation of $\xi^2\chi$ for $N = 2$, $L = 50$, and $n_{\rm cool} = 50$. Solid and the dashed lines represent, respectively, best fits obtained using fit function (27) setting $a_0 = 0$ and treating it as a free parameter. The full point in $1/\xi_L = 0$ represents the nonvanishing continuum-extrapolated determination.
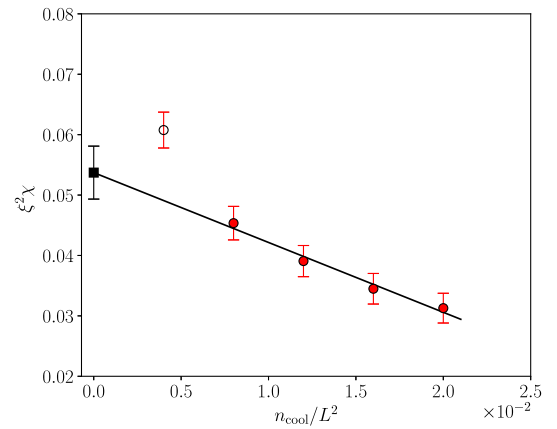

FIG. 5. Zero-cooling extrapolation of continuum-extrapolated results for $\xi^2\chi(n_{\rm cool})$ for $N = 2$ and $L = 50$ obtained for $n_{\rm cool} = 20, 30, 40, 50$ (full points). Determination for $n_{\rm cool} = 10$ has been excluded from the fit (empty point). The square full point at $n_{\rm cool} = 0$ represents our zero-cooling extrapolation according to fit function (28) without keeping into account correlations among determinations of $\xi^2\chi$ for different values of $n_{\rm cool}$. For the final result, see the text and Table IV.

different values of β_L are slightly correlated; thus, our result for the $\tilde{\chi}^2/{\rm d.o.f.}$ is actually underestimated.

To check if this explanation is reasonable, we repeated the best fits previously discussed computing the error on $\xi^2\chi$ without considering the error on ξ_L, so that the mentioned correlation becomes irrelevant in the evaluation of the $\tilde{\chi}^2/{\rm d.o.f.}$ Treating a_0 as a free parameter, we obtain $a_0 = 0.033(2)$ with $\tilde{\chi}^2/{\rm d.o.f.} = 17/16$, i.e., a perfectly agreeing result but with an $\mathcal{O}(1)$ reduced chi squared. The best fit with fixed $a_0 = 0$ is instead further disproved, as it yields an even larger reduced chi squared: $\tilde{\chi}^2/{\rm d.o.f.} = 104/17$.

In summary, these results point out that $\xi^2\chi$ behaves in the continuum limit in perfect agreement with the semi-classical prediction; cf. Eq. (18).

In order to check that all systematics are under control, we repeated this analysis, varying the number of cooling steps $n_{\rm cool}$, changing the value of $\beta_L^\star$, and narrowing the fit range. While again we observe that the latter two choices do not produce any appreciable change in the obtained result for $a_0 \neq 0$, as any observed variation of this parameter is much smaller compared to its statistical error, we observe a systematic drift of our continuum extrapolations for $\xi^2\chi$ as $n_{\rm cool}$ is increased (see Fig. 5).

Naively, one could think that taking the continuum limit at a fixed value of $n_{\rm cool} = (r_s/a)^2$ would result in a vanishing smoothing radius. However, since we took $a \to 0$, fixing $L = l/a$, the quantity $n_{\rm cool}/L^2 = (r_s/l)^2$ is kept constant in our continuum extrapolation and does not disappear from the game. The fact that $\xi^2\chi$ decreases increasing $n_{\rm cool}$ can be easily understood in these terms: $n_{\rm cool}$ fixes r_s in lattice spacing units, and, hence, results should eventually become independent of $n_{\rm cool}$ for a theory

with no UV divergences. However, because of the divergent small-instanton density and of the fixed ratio between the UV and the IR cutoffs $\sqrt{n_{\rm cool}}/L = r_s/l$, the fraction of topological signal which is smoothed away becomes eventually finite and independent of the lattice spacing but increases as r_s/l increases.

In order to provide the correct final result for $\xi^2\chi(N = 2)$ including the full UV contribution, the correct thing to do is to extrapolate continuum results toward the $n_{\rm cool} \to 0$ limit. To do so, we extrapolated our continuum determinations for $\xi^2\chi$ assuming the following scaling function:

$$\xi^2\chi\left(\frac{n_{\rm cool}}{L^2}\right) = \xi^2\chi\left(\frac{n_{\rm cool}}{L^2} = 0\right) + A\frac{n_{\rm cool}}{L^2}. \qquad (28)$$

This fit function is justified on the basis of the argument explained in Ref. [65]. Such argument is, strictly speaking, proven within the gradient flow formalism. However, since it has been shown that performing $n_{\rm cool}$ cooling steps is numerically equivalent to flow for a time $\tau_{\rm flow} = k n_{\rm cool}$ with k constant [e.g., $\tau_{\rm flow} = n_{\rm cool}/3$ in the 4D SU(3) pure-gauge theory with the Wilson action] [59], we expect this argument to also apply in our case.

In the continuum theory, any operator $\mathcal{O}_{\rm smooth}$ computed on smoothed fields can be expressed in terms of operators computed on the nonsmoothed ones by the operator product expansion formalism. The leading-order contribution is simply given by $\mathcal{O}$ computed on nonsmoothed fields (apart from a multiplicative renormalization constant), and higher-order contributions coming from contaminating higher-dimensional operators are suppressed as suitable compensating powers of the amount of smoothing performed. In the case of the topological susceptibility, the

relevant operator to be considered is just the topological charge density $q(x)$, since $\chi = \int d^2x \langle q(x)q(0)\rangle$. In this case, the renormalization constant appearing in front of the leading-order term is just 1 because of the nonrenormalizability of the topological charge in the continuum theory, while the next-to-leading-order term is suppressed as $n_{\rm cool} \propto r_s^2$ [65]. This justifies the ansatz given in Eq. (28).

The result of the best fit of our data with ansatz (28) is shown in Fig. 5. A linear term in $n_{\rm cool}$ nicely describes our data for $n_{\rm cool} > 10$ ($\chi^2/{\rm d.o.f.} = 0.32/2$). Including $n_{\rm cool} = 10$ instead yields a much larger $\chi^2/{\rm d.o.f.} = 7.4/3$, thus providing a worse description of our data. Excluding further points (e.g., $n_{\rm cool} = 20, 30$) gives compatible results within the errors with the one obtained excluding $n_{\rm cool} = 10$, thus justifying our choice for the fit range. Our final zero-cooling extrapolation turns out to be $\xi^2\chi(n_{\rm cool} = 0) = 0.054(4)$, i.e., clearly different from zero. The latter result has been obtained by performing the continuum extrapolation at fixed $n_{\rm cool}$ followed by the $n_{\rm cool}/L^2 \to 0$ limit on $\mathcal{O}(1000)$ bootstrap resamplings extracted for each value of ξ_L, each one of the same size of the corresponding original dataset.

C. Checking the L dependence and the thermodynamic limit

In Sec. IV B, we have shown that our results for the topological susceptibility are compatible with the log-divergent continuum limit predicted by semiclassical arguments. Our numerical evidence has been obtained on lattices with fixed $L = 50$, i.e., with vanishing volume in the continuum limit, and is based on the ansatz, stemming from perturbative computations, reported in Eqs. (17) and (18). Therefore, as a last step along our investigation, it is useful to check the dependence on L appearing in this ansatz and, moreover, that results are consistent with those obtained in standard simulations approaching the thermodynamic limit, i.e., for fixed $l = La$ and $L \gg \xi_L$, such as those reported in our previous study in Ref. [42].

As already discussed in Sec. III C, we have the following prediction [see Eq. (17)]:

$$\xi^2\chi = C\log\left(\frac{\rho_{\max}}{\rho_{\min}}\right), \qquad (29)$$

where $\rho_{\max}$ ($\rho_{\min}$) are the maximum (minimum) instanton size that can be observed on the given lattice. On a small lattice with fixed $L \ll \xi_L$, we expect $\rho_{\max} \propto L$, while on a large lattice with $L \gg \xi_L$ we expect $\rho_{\max}$ to be fixed by some physical IR cutoff, and, hence, $\rho_{\max} \propto \xi_L$ in lattice spacing units. Regarding $\rho_{\min}$, instead, we expect it to be proportional to the lattice spacing, with a proportionality constant independent of L as long as $L \gg 1$. Putting these considerations together, we expect

$$\xi^2\chi(\xi_L) \underset{\xi_L \to \infty}{\sim} C\log\left(\frac{\xi_L}{R}\right), \qquad L \gg \xi_L, \qquad (30)$$

$$\xi^2\chi(L) \underset{\xi_L \to \infty}{\sim} C\log\left(\frac{L}{R}\right), \qquad L \ll \xi_L, \qquad (31)$$

TABLE III. Summary of simulation parameters and results obtained for $N = 2$ and $L = 100, 200$. The correlation length ξ_L is computed according to Eq. (22) with $\beta_L^\star = 1.70$. Reported values of χ are computed after $n_{\rm cool} = 50$ cooling steps. Statistics is expressed in millions (M), and measures are taken every ten standard MC steps ($= 4$ OR $+1$ HB lattice updating sweeps).

|			$L = 100$			$L = 200$		
N	β_L	ξ_L	$a^2\chi \times 10^9$	$\xi^2\chi \times 10^3$	Statistics	$a^2\chi \times 10^9$	$\xi^2\chi \times 10^3$	Statistics
2	1.70	0.17991(78)	3912(10)	126.6(1.2)	9M	5620(22)	181.9(1.7)	2M
	1.75	0.2393(10)	2149.1(7.9)	123.0(1.2)		3097(16)	177.3(1.8)	
	1.80	0.3185(14)	1193.5(5.9)	121.1(1.2)		1717(12)	174.2(2.0)	
	1.85	0.4243(18)	652.9(4.4)	117.5(1.3)		948.5(9.4)	170.7(2.3)	
	1.90	0.5656(25)	365.9(3.4)	117.1(1.5)		525.0(6.9)	168.0(2.7)	
	1.95	0.7545(33)	198.9(2.5)	113.3(1.7)		282.9(5.4)	161.0(3.4)	
	2.00	1.0072(44)	108.3(1.9)	109.9(2.2)		160.4(4.1)	162.7(4.4)	
	2.05	1.3453(58)	61.0(1.5)	110.4(2.8)		88.5(2.3)	160.1(4.4)	
	2.10	1.7980(78)	33.81(56)	109.3(2.0)	37M	48.2(1.2)	155.8(4.2)	4M
	2.15	2.404(10)	18.23(39)	105.4(2.4)		27.57(87)	159.4(5.2)	8M
	2.20	3.217(14)	10.01(21)	103.6(2.4)	75M	14.40(43)	149.0(4.7)	
	2.25	4.307(19)	5.61(16)	104.0(3.1)		9.28(43)	172.0(8.2)	15M
	2.30	5.768(25)	3.08(12)	102.5(4.2)		4.43(26)	147.5(8.6)	
	2.35	7.729(34)	1.745(93)	104.3(5.6)		2.53(22)	151(13)	
	2.40	10.362(45)	0.955(72)	102.6(7.8)		1.30(12)	139(13)	
	2.45	13.897(60)	0.542(48)	104.6(9.3)		0.767(86)	148(17)	
	2.50	18.645(81)	0.355(47)	123(16)		0.458(74)	159(26)	
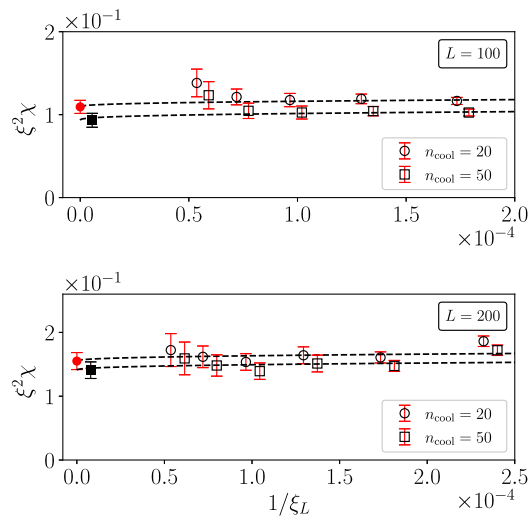
FIG. 6. Extrapolation toward the continuum limit of $\xi^2\chi$ for $N = 2$ and $L = 100$, 200 for $n_{\text{cool}} = 20$, 50. Dashed lines represent best fits obtained using fit function (27). The full point in $1/\xi_L = 0$ represents the nonvanishing continuum-extrapolated determination.

where R and $\bar{R}$ are two effective parameters which are different in the two cases, while the prefactor C is expected (and we will actually check) to be the same, since it just comes from the (unknown) prefactor of the instanton density $d_I(\rho) \propto 1/\rho$.

To extract C from finite L results, thus, we need to study the L dependence of the finite continuum limit of $\xi^2\chi(N = 2)$. For this reason, we also performed simulations for $L = 100$ and $L = 200$. The only difference compared to the $L = 50$ investigation discussed above is that, for these lattices, we do not employ the multicanonical algorithm, since the logarithmic UV divergence is assumed *a priori*; hence, we do not need extremely precise data to disprove a convergent behavior. In Table III, we summarize the parameters of the simulations performed for $L = 100$ and 200.

The computation of $\xi^2\chi$ for $L = 100$, 200 has been done following the same lines as Sec. IV B. First, we extrapolate our results toward the continuum limit at fixed value of n_{cool}. Continuum extrapolations at fixed $n_{\text{cool}} = 20$, 50 for $L = 100$, 200 are shown in Fig. 6. As a further consistency check, we also verified that the free exponent c appearing in

TABLE IV. Double-extrapolated results for $\xi^2\chi(L)$ ($1/\xi_L \to 0$ followed by $n_{\text{cool}}/L^2 \to 0$) and determinations of the exponent c appearing in fit function (27).

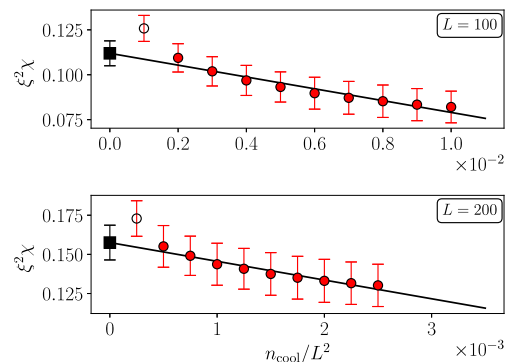| L	$\xi^2\chi(n_{\text{cool}}/L^2 = 0)$	Exponent c
50	0.054(4)	0.20(2)
100	0.109(9)	0.35(14)
200	0.15(2)	0.39(20)

FIG. 7. Zero-cooling extrapolation of continuum-extrapolated results for $\xi^2\chi(n_{\text{cool}})$ for $N = 2$ and $L = 100$, 200 obtained for $n_{\text{cool}} = 20, 30, \ldots, 100$ (full points). Determinations for $n_{\text{cool}} = 10$ have been excluded from the fit (empty point). The square full point at $n_{\text{cool}} = 0$ represents our zero-cooling extrapolation according to fit function (28) without keeping into account correlations among determinations of $\xi^2\chi$ for different values of n_{cool}. For the final result, see the text and Table IV.

the fit function in Eq. (27) was compatible within the errors in all cases; cf. Table IV.

Then, we extrapolate such continuum determinations toward the zero-cooling limit. Again, our results for $L = 100$, 200 are nicely described by a linear function in n_{cool}/L^2; cf. Fig. 7. Our final results for $\xi^2\chi(n_{\text{cool}}/L^2 = 0)$, for $L = 50$, 100, and 200, are collected in Table IV.

Finally, we performed a best fit of our results for $\xi^2\chi(n_{\text{cool}}/L^2 = 0)$ as a function of the fixed lattice size L according to Eq. (31) to determine the prefactor C.

Our data are very well described by a log-divergent function of the lattice size L, as shown in Fig. 8, and we obtain
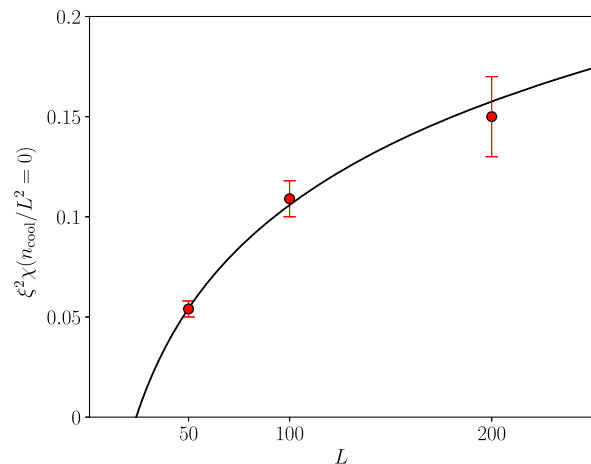
$$C = 0.074(11), \tag{32}$$

FIG. 8. Best fit of $\xi^2\chi(n_{\text{cool}}/L^2 = 0)$ as a function of L according to Eq. (31). Best fit gives $\tilde{\chi}^2/\text{d.o.f.} = 0.23/1$.
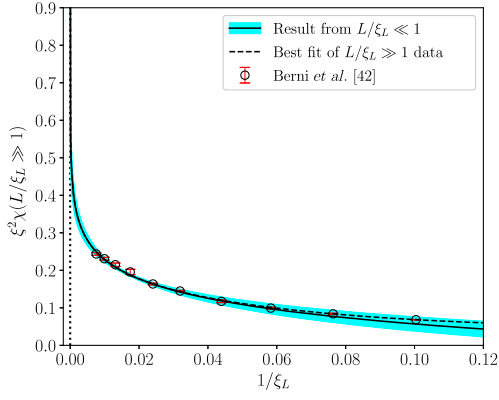
FIG. 9. Results for $\xi^2\chi$ reported in Ref. [42] for $L/\xi_L \gg 1$ and $n_{\text{cool}} = 20$. The dashed line represents best fit of these data according to fit function $f(x) = -C\log(\bar{R}x) + C_2 x^2$, where $x = 1/\xi_L$ and the result for C is reported in Eq. (34). The solid line represents the curve $g(x) = -C\log(Rx) + \log(R/\bar{R})$, where $x = 1/\xi_L$ and C is reported in Eq. (32). The shadowed area represents the error band on $g(x)$.

$$R = 24(3). \tag{33}$$

It is now interesting to compare these results with those of Ref. [42], obtained in the thermodynamic limit $L/\xi_L \gg 1$.

Extrapolating the results for $L/\xi_L \gtrsim 12$ reported in that work toward the continuum limit, and according to the divergent fit function in Eq. (30) plus $O(\xi_L^{-2})$ corrections, we obtain

$$C = 0.074(2), \tag{34}$$

$$\bar{R} = 4.7(3). \tag{35}$$

As expected, while the constants R and $\bar{R}$ are different, the prefactors C of the logarithms turn out to be in perfect agreement among each other.

In Fig. 9, we show the $L/\xi_L \gg 1$ determinations for $\xi^2\chi$ of Ref. [42] along with their best fit according to Eq. (30) plus $O(\xi_L^{-2})$ corrections. On top of these, we plot the curve $C\log(\xi_L/R) + \log(R/\bar{R})$, using the value of C in Eq. (32), i.e., coming from the logarithmic best fit of the fixed-L results obtained in this work, reported in Table IV. The two curves collapse on top of each other.

In conclusion, the comparison carried out in this subsection provides solid numerical evidence that results obtained by fixed L simulations contain information which is consistent, as for the UV behavior of the topological susceptibility, with what would be obtained in the thermodynamic infinite volume limit.

V. CONCLUSIONS

The purpose of the present study was that of providing numerical evidence for the predicted divergent behavior in the continuum limit of the topological susceptibility of the CP^1 model. The same problem has been considered by several past studies; the novelty of the present investigation is to approach the continuum limit at fixed volume in dimensionless lattice units: This maps a logarithmically divergent behavior, which can be barely distinguishable from a badly convergent behavior over a wide range of lattice spacings, into a convergent behavior with a non-vanishing continuum limit, which is more amenable to be checked numerically with a well-defined conclusion.

After checking that this method reproduces the results obtained with standard strategies for the CP^2 and the CP^3 theories, we applied it to our target model, implementing at the same time a multicanonical algorithm in order solve the problem of rare fluctuations of the topological charge on asymptotically small lattices. The use of the multicanonical algorithm revealed essential, since it reduced the computational effort by up to 2 orders of magnitude for the smallest explored lattice spacings.

Our results show that the continuum limit of the topological susceptibility of the CP^1 model obtained at fixed L, and after extrapolation to zero-cooling steps, is indeed nonvanishing, as predicted by semiclassical computations. Moreover, repeating the same computation for different values of L, we observe that the obtained nonvanishing determinations of $\xi^2\chi$ grow proportionally to $\log L$, with a prefactor consistent with previous lattice results: That provides evidence that our investigation at fixed L is perfectly consistent with what would be obtained in the thermodynamic infinite volume limit; however, it permits one, at the same time, to definitely disprove the possibility of a convergent behavior for χ.

APPENDIX A: ASYMPTOTIC SCALING CHECK

To check if our assumption of being in the asymptotic scaling region is correct, we consider the quantity $M/\Lambda_L^{(\text{Sym})} \equiv [\xi_L f(\beta_L)]^{-1}$, which is expected to be constant plus $O(1/\beta_L)$ corrections for $\beta_L \to \infty$. For $N = 2$, the exact value of M in the continuum is known and can be expressed in terms of the dynamically generated scale of the Symanzik theory $\Lambda_L^{(\text{Sym})}$ by combining results of Refs. [9,23]:

TABLE V. Simulation summary of the additional runs performed to check asymptotic scaling for $N = 2$, 3, and 4.

N	β_L	L	ξ_L	L/ξ_L
2	1.70	360	141.93(26)	2.5
		500	162.13(63)	3
		600	170.17(66)	3.5
		700	174(1)	4
		800	177(1)	4.5
		1024	179.81(87)	5.7
		1450	181(2)	8
		∞	179.91(78)	∞
3	1.32	562	44.75(29)	12.5
	1.455	1250	98.19(53)	12.4
4	1.20	436	34.03(18)	12.7
	1.30	766	61.76(34)	12.5
	1.35	1030	82.62(70)	12.5

$$\left.\frac{M}{\Lambda_L^{(\text{Sym})}}(N = 2)\right|_{\text{exact}} \simeq 21.7. \qquad (\text{A1})$$

To test asymptotic scaling for CP¹, CP², and CP³ models, we consider results for ξ_L as a function of β_L of Ref. [42]. Furthermore, we also added higher-ξ_L data to this analysis, which are reported in Table V.

For $N = 3$ and 4, we chose the lattice size requiring that $L/\xi_L \gtrsim 12$, which is enough to ensure that finite size effects are well under control. For $N = 2$, instead, we computed ξ_L for several lattice sizes and extrapolated it toward the thermodynamic limit by fitting its L dependence according to

$$\xi_L(L) = \xi_L^{(\infty)}(1 - a e^{-bL/\xi_L}), \qquad (\text{A2})$$

where $\xi_L^{(\infty)}$ is the desired quantity and a and b are additional fit parameters. In Fig. 10, we display the quantity $M/\Lambda_L^{(\text{Sym})} = [\xi_L f(\beta_L)]^{-1}$ as a function of $1/\beta_L$ for, respectively, the CP¹, CP², and CP³ models.

For $N = 4$ and 3, the quantity $M/\Lambda_L^{(\text{Sym})}$ reaches a plateau asymptotically. Thus, we choose $\beta_L^\star(N = 4) = 1.35$ and $\beta_L^\star(N = 3) = 1.455$ to fix ξ_L via Eq. (22). For $N = 2$, despite the wider range of $1/\beta_L$ explored, we observe a slower approach to the asymptotic scaling regime probably due to larger $O(1/\beta_L)$ corrections in this case. Nonetheless, we observe that the obtained results for ξ_L using Eq. (22) do not show an appreciable dependence on the choice of $\beta_L^\star$, showing that our procedure to fix the scale is solid even in this case. As an example, for $N = 2$ and $\beta_L = 3.00$ we have $\xi_L = 354.7(1.5)$ if $\beta_L^\star = 1.65$ and $\xi_L = 359.6(1.6)$ if $\beta_L^\star = 1.70$; i.e., the two determinations agree within $\sim$2.2 standard deviations. Therefore, we choose $\beta_L^\star(N = 2) = 1.70$ to fix the scale in this case.
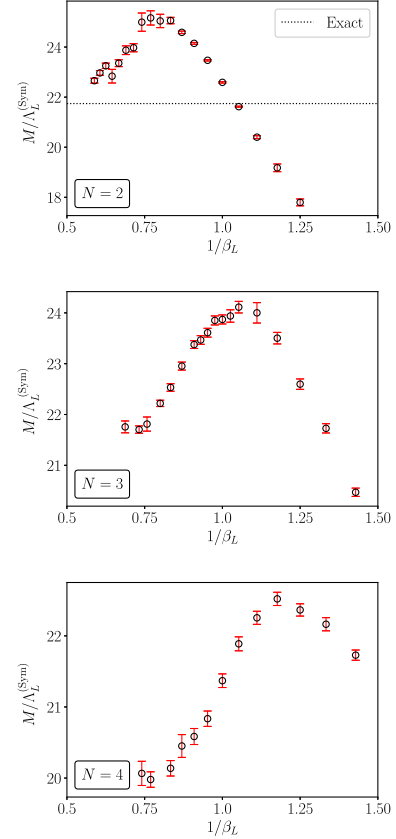

FIG. 10. Check of the asymptotic scaling of ξ_L for the CP¹, CP², and CP³ models. The figures show the behavior of $M/\Lambda_L^{(\text{Sym})} = [\xi_L f(\beta_L)]^{-1}$ as a function of the inverse coupling $1/\beta_L$. For $N = 2$, the dotted line displays the exact analytic result for the continuum limit of $M/\Lambda_L^{(\text{Sym})}(N = 2)$ in Eq. (A1).

APPENDIX B: MULTICANONICAL ALGORITHM DETAILS

The topological bias potential was chosen according to the same functional form adopted in Ref. [47]:

$$V_{\text{topo}}(x) = \begin{cases} -\sqrt{(Bx)^2 + C}, & \text{if } |x| \le Q_{\max}, \\ -\sqrt{(BQ_{\max})^2 + C}, & \text{if } |x| > Q_{\max}. \end{cases} \qquad (\text{B1})$$

Here, B, C, and $Q_{\max}$ are free parameters that can be calibrated through short preliminary runs to improve the performances of the multicanonic algorithm. The employed values of B varied between $\sim$5 and $\sim$10 for $\beta_L \in [2.2, 3]$, while the choice of C and $Q_{\max}$ turned out to be not critical; thus, we used $C = 0.05$ and $Q_{\max} = 12$ for all β_L. An illustrative example of the functional form in Eq. (B1) is shown in Fig. 11.

Our implementation of the multicanonical algorithm follows the lines of Ref. [45]. First, we generate a candidate new lattice configuration by performing a standard updating step and ignoring the Q-dependent bias potential. Then,
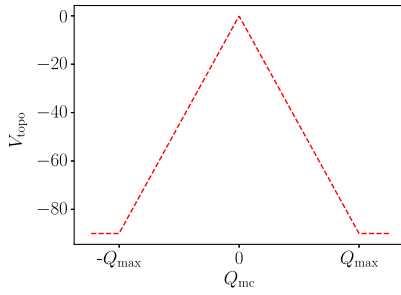
FIG. 11. Illustrative example of the bias potential $V_{\text{topo}}(Q_{mc})$ in Eq. (B1) with $B = 7.5$ and $C = 0.05$ and with a cut at $Q_{\max} = 12$.

we accept the updated configuration by performing a standard Metropolis test:

$$p = \min\{1, \exp(-\Delta V_{\text{topo}})\},$$

where

$$\Delta V_{\text{topo}} \equiv V_{\text{topo}}(Q_{\text{mc}}^{(\text{new})}) - V_{\text{topo}}(Q_{\text{mc}}^{(\text{old})})$$

is the variation of the bias potential before and after the update. After running some preliminary simulations, we found that the optimal implementation to have higher Metropolis acceptances was to perform the Metropolis test after each single-link update $U_\mu(x)$ instead of performing it after a whole standard MC step (i.e., after five sweeps of the

whole lattice). Moreover, we also found that proposing single-site and single-link updates stochastically was more effective to obtain higher Metropolis acceptances than performing lattice sweeps.

It is easy to verify that, for any starting updating step, our choice respects *detailed balance* for the original distribution. Moreover, when considering the path-integral probability distribution obtained with the modified action in Eq. (24), our multicanonical updating step with the addition of the Metropolis test respects detailed balance, too.

Finally, regarding the topological charge discretization Q_{mc}, our choice is $Q_{\text{mc}} = Q_U$, i.e., the geometric definition in Eq. (13) computed without performing any cooling step. This choice allows one to avoid the full computation of Q_{mc} (necessary to compute the Metropolis probability) every time an update of a link variable $U_\mu(x)$ is proposed, as with this choice one can directly compute $\Delta Q_{\text{mc}} = Q_{\text{mc}}^{(\text{new})} - Q_{\text{mc}}^{(\text{old})}$ in terms of the new link and its relative staples.

With this setup, we obtained mean Metropolis acceptances larger than 90%, and we found that a multicanonical MC step required an $\approx 85\%$ larger numerical effort compared to a standard MC step. After taking into account such overhead, we found that the multicanonic algorithm allowed to gain up to 2 orders of magnitude in terms of computational power compared to the standard algorithm when $\langle Q^2 \rangle \ll 1$.

[1] A. D'Adda, M. Lüscher, and P. Di Vecchia, Nucl. Phys. **B146**, 63 (1978).

[2] M. Shifman, *Advanced Topics in Quantum Field Theory* (Cambridge University Press, Cambridge, England, 2012), pp. 171–268, 361–367.

[3] E. Vicari and H. Panagopoulos, Phys. Rep. **470**, 93 (2009).

[4] E. Witten, Ann. Phys. (N.Y.) **128**, 363 (1980).

[5] E. Witten, Phys. Rev. Lett. **81**, 2862 (1998).

[6] P. Rossi and E. Vicari, Phys. Rev. D **48**, 3869 (1993).

[7] P. Rossi, Phys. Rev. D **94**, 045013 (2016).

[8] M. Campostrini and P. Rossi, Phys. Lett. B **272**, 305 (1991).

[9] M. Campostrini, P. Rossi, and E. Vicari, Phys. Rev. D **46**, 2647 (1992).

[10] M. Campostrini, P. Rossi, and E. Vicari, Phys. Rev. D **46**, 4643 (1992).

[11] M. Campostrini and P. Rossi, Riv. Nuovo Cimento **16**, 1 (1993).

[12] C. Bonati, M. D'Elia, P. Rossi, and E. Vicari, Phys. Rev. D **94**, 085017 (2016).

[13] C. Bonanno, C. Bonati, and M. D'Elia, J. High Energy Phys. 01 (2019) 003.

[14] M. Berni, C. Bonanno, and M. D'Elia, Phys. Rev. D **100**, 114509 (2019).

[15] D. Forster, Nucl. Phys. **B130**, 38 (1977).

[16] B. Berg and M. Lüscher, Nucl. Phys. **B190**, 412 (1981).

[17] F. D. M. Haldane, Phys. Lett. **93A**, 464 (1983).

[18] F. D. M. Haldane, Phys. Rev. Lett. **50**, 1153 (1983).

[19] G. Bhanot and F. David, Nucl. Phys. **B251**, 127 (1985).

[20] F. D. M. Haldane, J. Appl. Phys. **57**, 3359 (1985).

[21] P. B. Wiegmann, Phys. Lett. **152B**, 209 (1985).

[22] I. Affleck and F. D. M. Haldane, Phys. Rev. B **36**, 5291 (1987).

[23] P. Hasenfratz, M. Maggiore, and F. Niedermayer, Phys. Lett. B **245**, 522 (1990).

[24] M. D'Elia, F. Farchioni, and A. Papa, Nucl. Phys. **B456**, 313 (1995).

[25] M. D'Elia, F. Farchioni, and A. Papa, Phys. Rev. D **55**, 2274 (1997).

[26] M. Blatter, R. Burkhalter, P. Hasenfratz, and F. Niedermayer, Phys. Rev. D **53**, 923 (1996).

[27] R. Burkhalter, M. Imachi, Y. Shinno, and H. Yoneyama, Prog. Theor. Phys. **106**, 613 (2001).

[28] D. Controzzi and G. Mussardo, Phys. Rev. Lett. **92**, 021601 (2004).

[29] B. Alles and A. Papa, Phys. Rev. D **77**, 056008 (2008).

[30] D. Nogradi, J. High Energy Phys. 05 (2012) 089.

[31] V. Azcoiti, G. Di Carlo, E. Follana, and M. Giordano, Phys. Rev. D **86**, 096009 (2012).

[32] B. Alles, M. Giordano, and A. Papa, Phys. Rev. B **90**, 184421 (2014).

[33] W. Bietenholz, P. de Forcrand, U. Gerber, H. Mejía-Díaz, and I. O. Sandoval, Phys. Rev. D **98**, 114501 (2018).

[34] F. Bruckmann, K. Jansen, and S. Kühn, Phys. Rev. D **99**, 074501 (2019).

[35] T. Sulejmanpasic and Y. Tanizaki, Phys. Rev. B **97**, 144201 (2018).

[36] T. Sulejmanpasic, D. Göschl, and C. Gattringer, Phys. Rev. Lett. **125**, 201602 (2020).

[37] S. Thomas and C. Monahan, Proc. Sci. LATTICE2021 (**2022**) 076.

[38] M. Marino, R. Miravitllas, and T. Reis, arXiv:2205.04495.

[39] C. Bonanno, C. Bonati, and M. D'Elia, J. High Energy Phys. 03 (2021) 111.

[40] R. Kitano, N. Yamada, and M. Yamazaki, J. High Energy Phys. 02 (2021) 073.

[41] R. Kitano, R. Matsudo, N. Yamada, and M. Yamazaki, Phys. Lett. B **822**, 136657 (2021),

[42] M. Berni, C. Bonanno, and M. D'Elia, Phys. Rev. D **102**, 114519 (2020).

[43] L. Del Debbio and A. Ramos, Phys. Rep. **920**, 1 (2021).

[44] B. A. Berg and T. Neuhaus, Phys. Lett. B **267**, 249 (1991).

[45] P. T. Jahn, G. D. Moore, and D. Robaina, Phys. Rev. D **98**, 054512 (2018).

[46] C. Bonati and M. D'Elia, Phys. Rev. E **98**, 013308 (2018).

[47] C. Bonati, M. D'Elia, G. Martinelli, F. Negro, F. Sanfilippo, and A. Todaro, J. High Energy Phys. 11 (2018) 170.

[48] A. Athenodorou, C. Bonanno, C. Bonati, G. Clemente, F. D'Angelo, M. D'Elia, L. Maio, G. Martinelli, F. Sanfilippo, and A. Todaro, J. High Energy Phys. 10 (2022) 197.

[49] D. Petcher and M. Lüscher, Nucl. Phys. **B225**, 53 (1983).

[50] Y. Lian and H. B. Thacker, Phys. Rev. D **75**, 065031 (2007).

[51] M. Lüscher, Nucl. Phys. **B200**, 61 (1982).

[52] K. Symanzik, Nucl. Phys. **B226**, 187 (1983).

[53] S. Caracciolo and A. Pelissetto, Phys. Rev. D **58**, 105007 (1998).

[54] M. Campostrini, A. Di Giacomo, and H. Panagopoulos, Phys. Lett. B **212**, 206 (1988).

[55] F. Farchioni and A. Papa, Phys. Lett. B **306**, 108 (1993).

[56] B. Berg, Phys. Lett. **104B**, 475 (1981).

[57] B. Alles, M. D'Elia, A. Di Giacomo, and R. Kirchner, Phys. Rev. D **58**, 114506 (1998).

[58] M. D'Elia, Nucl. Phys. **B661**, 139 (2003).

[59] C. Bonati and M. D'Elia, Phys. Rev. D **89**, 105005 (2014).

[60] C. Alexandrou, A. Athenodorou, and K. Jansen, Phys. Rev. D **92**, 125014 (2015).

[61] L. Del Debbio, G. M. Manca, and E. Vicari, Phys. Lett. B **594**, 315 (2004).

[62] M. Lüscher and S. Schaefer, J. High Energy Phys. 07 (2011) 036.

[63] M. Hasenbusch, Phys. Rev. D **96**, 054504 (2017).

[64] C. Bonanno, M. D'Elia, B. Lucini, and D. Vadacchino, Phys. Lett. B **833**, 137281 (2022).

[65] L. Altenkort, A. M. Eller, O. Kaczmarek, L. Mazur, G. D. Moore, and H.-T. Shu, Phys. Rev. D **103**, 114513 (2021).