

# Large- $N$ $SU(N)$ Yang-Mills theories with milder topological freezing

Claudio Bonanno,<sup>a,b</sup> Claudio Bonati<sup>a</sup> and Massimo D'Elia<sup>a</sup>

<sup>a</sup> *Università di Pisa and INFN Sezione di Pisa,  
Largo Pontecorvo 3, I-56127 Pisa, Italy*

<sup>b</sup> *Centro Nazionale INFN di Studi Avanzati GGI,  
Largo Enrico Fermi 2, I-50125 Firenze, Italy*

*E-mail:* [claudio.bonanno@pi.infn.it](mailto:claudio.bonanno@pi.infn.it), [claudio.bonati@unipi.it](mailto:claudio.bonati@unipi.it),  
[massimo.delia@unipi.it](mailto:massimo.delia@unipi.it)

**ABSTRACT:** We simulate  $4d$   $SU(N)$  pure-gauge theories at large  $N$  using a parallel tempering scheme that combines simulations with open and periodic boundary conditions, implementing the algorithm originally proposed by Martin Hasenbusch for  $2d$   $CP^{N-1}$  models. That allows to dramatically suppress the topological freezing suffered from standard local algorithms, reducing the autocorrelation time of  $Q^2$  up to two orders of magnitude. Using this algorithm in combination with simulations at non-zero imaginary  $\theta$  we are able to refine state-of-the-art results for the large- $N$  behavior of the quartic coefficient of the  $\theta$ -dependence of the vacuum energy  $b_2$ , reaching an accuracy comparable with that of the large- $N$  limit of the topological susceptibility.

**KEYWORDS:**  $1/N$  Expansion, Lattice Quantum Field Theory, Nonperturbative Effects

**ARXIV EPRINT:** [2012.14000](https://arxiv.org/abs/2012.14000)

---

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Lattice setup</b>	<b>3</b>
2.1	Lattice action and lattice topological charge	3
2.2	Imaginary- $\theta$ method	4
2.3	Parallel tempering of volume defect	5
<b>3</b>	<b>Numerical results</b>	<b>7</b>
3.1	Parallel tempering: results and comparison	8
3.2	Analytic continuation and continuum limit	12
3.3	Large- $N$ limit	13
<b>4</b>	<b>Conclusions</b>	<b>17</b>

---

## 1 Introduction

The dependence of physical observables on the topological parameter  $\theta$  is one of the most interesting properties of four dimensional  $SU(N)$  pure-gauge theories. The parameter is coupled in the action to the topological charge

$$Q = \int d^4x q(x) = \frac{1}{64\pi^2} \varepsilon_{\mu\nu\rho\sigma} \int d^4x F_{\mu\nu}^a(x) F_{\rho\sigma}^a(x); \quad (1.1)$$

$\theta$ -dependence can be studied to achieve a better understanding of the non-perturbative features of Yang-Mills theories, but also has direct phenomenological implications for hadron physics in the limit of large number of colors [1–7].

A particularly interesting quantity, whose dependence on  $\theta$  has been thoroughly investigated, is the vacuum energy density  $E(\theta)$ , which is formally defined by the relation

$$E(\theta) \equiv -\frac{1}{V} \log \int [dA] e^{-S_{\text{YM}}[A] + i\theta Q[A]}, \quad (1.2)$$

where  $V$  is the four-dimensional euclidean space-time volume. The functional form of  $E(\theta)$  is known only for very specific theories, like for QCD close to the chiral limit [7]. It is customary to consider a Taylor expansion of  $E(\theta)$  around  $\theta = 0$ . Since  $Q$  is odd under a CP transformation, only even powers of  $\theta$  appear in the expansion, which can be parametrized in the form

$$E(\theta) - E(0) = \frac{1}{2} \chi \theta^2 \left( 1 + \sum_{n=1}^{\infty} b_{2n} \theta^{2n} \right). \quad (1.3)$$

The coefficients of this expansion are related to the cumulants of the topological charge distribution at  $\theta = 0$ , for instance  $\chi = \langle Q^2 \rangle_{\theta=0} / V$ , where  $\chi$  is the topological susceptibility.

While the exact numerical values of the coefficients  $\chi$  and  $b_{2n}$  are generically unknown, something is known about their dependence on the number of colors  $N$ , at least when  $N$  is large enough. Indeed, assuming that a non-trivial  $\theta$ -dependence is present in the large- $N$  limit, it is possible to fix the large  $N$  scaling form of the energy density  $E(\theta, N) = N^2 \bar{E}(\theta/N)$  [4, 6, 8], implying

$$\chi = \bar{\chi} + O\left(\frac{1}{N^2}\right), \tag{1.4}$$

$$b_{2n} = \frac{\bar{b}_{2n}}{N^{2n}} + O\left(\frac{1}{N^{2(n+1)}}\right). \tag{1.5}$$

The value of  $\bar{\chi}$  is related to the mass of the  $\eta'$  meson by the Witten-Veneziano formula [4, 5], which provides the estimate  $\bar{\chi} \simeq (180 \text{ MeV})^4$ . Analytic estimates of  $\bar{\chi}$  and of the  $\bar{b}_{2n}$  coefficients are available only for two dimensional models [9–13]. Given the non-perturbative nature of  $\theta$ -dependence, the numerical lattice approach is the natural tool to investigate such topics quantitatively, and in particular to test large- $N$  predictions [11, 13–27]. There are however some non-trivial computational challenges that have to be faced, especially in the large  $N$  regime.

The first problem is related to the measure of the coefficients  $b_{2n}$ . The task is challenging by itself since, contrary to what happens for  $\chi$ , these coefficients approach zero as  $N$  is increased. In addition, the simplest estimators available for  $\theta = 0$  simulations are based on the cumulants of the topological charge distribution, which are not self-averaging quantities, leading to a bad signal-to-noise ratio for large volumes. Such problem can be solved by introducing an explicit source term in the action, coupled to the topological charge density, which corresponds in practice to an imaginary  $\theta$  term [22, 28–36]. The method was exploited for the determination of the  $b_{2n}$  coefficients first in ref. [22] and later developed and applied in several works to both  $SU(N)$  Yang-Mills theory [13, 25–27] and two dimensional  $CP^{N-1}$  models [37, 38].

The second problem is the so-called “freezing problem”, it represents a well known problem in a wide range of theories sharing the presence of topological excitations [17, 39–49] and it will be the main topic of this work. When adopting local update algorithms on lattices with periodic boundary conditions, the topological modes experience a severe critical slowing down when approaching the continuum limit, with the autocorrelation time of the topological charge which grows approximately exponentially as a function of  $1/a$  [17, 43]. As a consequence gauge configurations stay fixed (frozen) in a given topological sector for an exceedingly large amount of Monte Carlo time, thus preventing a correct sampling of the path-integral distribution. This problem becomes even worse in the large- $N$  limit, since at fixed lattice spacing  $a$  the autocorrelation time of the topological charge seems to grow exponentially with the number of colors [17, 37, 43].

Despite the fact that a definitive solution to the topological freezing problem has been obtained only in toy models [48], several strategies have been proposed to reduce its severeness, in particular by reducing the exponential critical slowing down to a polynomial critical slowing down [44, 45, 48, 50], or to extract information even from completely frozen configurations [51]. A popular method, proposed in ref. [44], is to adopt open boundary

conditions for the gauge fields in the time direction instead of the usual periodic ones. The presence of an open boundary eliminates barriers among different topological sectors even in the continuum, and one expects the critical slowing down of the topological modes to be essentially diffusive, with an autocorrelation time increasing as  $1/a^2$  approaching the continuum limit. Using this method, however, one also breaks translation invariance and loses completely any notion of global topological charge. Therefore  $\chi$  and  $b_{2n}$  can only be estimated from the integral of the  $(2n + 2)$ -point connected correlators of the topological charge density on the bulk of the lattice.

A different strategy, that keeps the advantages of the open boundary conditions without breaking translation invariance, has been proposed by M. Hasenbusch in ref. [52], where it was tested for two dimensional  $CP^{N-1}$  models. The basic idea of this method is to combine periodic and open boundary conditions in a parallel tempering framework, using the copies with open or partially open boundary conditions as sources of topological fluctuations for the copy with periodic boundary conditions, which is the one on which measures are performed. To reduce the number of copies to be used in the parallel tempering, open boundary conditions are not enforced along all the temporal boundaries, but only in a limited space region, that will be referred to as the “defect” in the following.

In ref. [38] it was shown that in two dimensional  $CP^{N-1}$  models, for large  $N$  values, the adoption of the Hasenbusch algorithm in combination with the imaginary- $\theta$  method allows to achieve impressive improvements compared to previously available results for the  $\theta$ -dependence of these models. The aim of the present work is to test the same setup in the case of four dimensional  $SU(N)$  Yang-Mills theories at zero temperature, comparing its performance with that of the standard simulations. In doing this we will also refine the state-of-the-art results about the large- $N$  behavior of the  $b_2$  coefficient.

This paper is organized as follows: in section 2 we discuss our lattice setup, along with the parallel tempering algorithm and the imaginary- $\theta$  method, in section 3 we show the numerical results obtained with the parallel tempering and finally in section 4 we draw our conclusions.

## 2 Lattice setup

In this section we introduce the discretizations adopted for the action and the topological charge, we present a summary of the imaginary- $\theta$  method and discuss the parallel tempering algorithm employed in our simulations.

### 2.1 Lattice action and lattice topological charge

We discretize the Yang-Mills action on an hyper-cubic lattice of size  $L$  and with periodic boundary conditions in every direction (see section 2.3 for the defect) using the standard Wilson action:

$$S_W = -\frac{\beta}{N} \sum_{x, \mu > \nu} \Re \text{Tr} \{ \Pi_{\mu\nu}(x) \}, \tag{2.1}$$

where  $\Pi_{\mu\nu}(x) \equiv U_\mu(x)U_\nu(x + \hat{\mu})U_\mu^\dagger(x + \hat{\nu})U_\nu^\dagger(x)$  is the plaquette operator.

For the topological charge (1.1), we adopt the simplest discretization with definite parity, the so-called *clover* discretization:

$$Q_{\text{clov}} = \frac{1}{2^9 \pi^2} \sum_{x, \mu, \nu, \rho, \sigma} \varepsilon_{\mu\nu\rho\sigma} \text{Tr} \{ C_{\mu\nu}(x) C_{\rho\sigma}(x) \}, \quad (2.2)$$

where  $C_{\mu\nu}(x)$  is a discretization of the field strength given by the sum of all the 4 plaquettes centered in the site  $x$  and lying on the  $\mu$ - $\nu$  plane.  $Q_{\text{clov}}$  is generically non-integer and it is related, configuration by configuration, to the physical charge  $Q$  by [53]:

$$Q_{\text{clov}} = ZQ + \eta, \quad (2.3)$$

where  $Z$  is a finite renormalization constant that approaches 1 in the continuum limit, and  $\eta$  is a stochastic noise due to ultraviolet (UV) fluctuations at the scale of the lattice spacing. Using the variance of  $Q_{\text{clov}}$  to estimate the topological susceptibility would require to take into account both multiplicative and additive renormalizations, which can be avoided by using one of the several smoothing procedures that have been proposed in the literature, such as the gradient flow [54, 55] or cooling [56–62], which are all known to agree with each other when properly matched [63, 64]. In this work we use cooling due to its simplicity and numerical effectiveness.

We denote by  $Q_{\text{clov}}^{\text{cool}}$  the topological charge obtained by measuring the observable eq. (2.2) on a configuration to which a certain number of cooling steps have been applied. To assign an integer topological charge  $Q_L$  to each configuration we follow ref. [17], defining

$$Q_L = \text{round} \left\{ \alpha Q_{\text{clov}}^{\text{cool}} \right\}, \quad (2.4)$$

where “round” denotes the rounding to the closest integer and the value of  $\alpha$  is fixed by minimizing

$$\left\langle \left( \alpha Q_{\text{clov}}^{\text{cool}} - \text{round} \left\{ \alpha Q_{\text{clov}}^{\text{cool}} \right\} \right)^2 \right\rangle, \quad (2.5)$$

so that the maxima in the distribution of  $\alpha Q_{\text{clov}}^{\text{cool}}$  are located approximately at integer values; such fixing is performed at  $\theta = 0$  and then adopted also for  $\theta \neq 0$ . The topological susceptibility computed using  $Q_L$  becomes stable (i.e. independent of the number of cooling steps  $n_{\text{cool}}$ ) after  $n_{\text{cool}} \sim 10$ , moreover such threshold reveals to be weakly dependent on the lattice spacing, thus we chose  $n_{\text{cool}} = 20$  to define the topological charge in all simulations, verifying also the stability of all continuum extrapolations if a different value of  $n_{\text{cool}}$  is used.

## 2.2 Imaginary- $\theta$ method

As anticipated in the introduction, the imaginary- $\theta$  method is a technique that is useful to estimate the topological susceptibility and especially the coefficients  $b_{2n}$  introduced in eq. (1.3). In this section we provide a short summary of this computational method, referring to ref. [25] for more details. The idea of the method is to introduce an imaginary  $\theta$  term, in order to avoid a sign problem, and to extract  $\chi$  and  $b_{2n}$  from the dependence on  $\theta$  of the cumulants of the topological charge distribution: the method wins over the standard computation at  $\theta = 0$  since now the information on the  $b_{2n}$  parameters is contained in all

cumulants, including the lowest order ones. The procedure is most conveniently explained by working formally in the continuum: the continuum euclidean action can be written in the form

$$S(\theta_I) = S_{\text{YM}} - \theta_I Q, \tag{2.6}$$

where  $\theta_I = i\theta$ . The dependence on  $\theta_I$  of the cumulants of the topological charge distribution can be computed from the derivatives of  $E(\theta)$  in eq. (1.2), properly continued to the imaginary axis, and can be expressed in terms of  $\chi$  and  $b_{2n}$  using eq. (1.3). As an example, the explicit expressions of the first few cumulants as a function of  $\theta_I$  read:

$$\begin{aligned} \frac{k_1(\theta_I)}{V} &= \chi \left[ \theta_I - 2b_2\theta_I^3 + 3b_4\theta_I^5 + O(\theta_I^6) \right], \\ \frac{k_2(\theta_I)}{V} &= \chi \left[ 1 - 6b_2\theta_I^2 + 15b_4\theta_I^4 + O(\theta_I^5) \right], \\ \frac{k_3(\theta_I)}{V} &= \chi \left[ -12b_2\theta_I + 60b_4\theta_I^3 + O(\theta_I^4) \right], \\ \frac{k_4(\theta_I)}{V} &= \chi \left[ -12b_2 + 180b_4\theta_I^2 + O(\theta_I^3) \right], \end{aligned} \tag{2.7}$$

where  $V$  is the space-time volume and the first fourth cumulants of the topological charge are

$$\begin{aligned} k_1 &= \langle Q \rangle, \\ k_2 &= \langle Q^2 \rangle - \langle Q \rangle^2, \\ k_3 &= \langle Q^3 \rangle - 3 \langle Q^2 \rangle \langle Q \rangle + 2 \langle Q \rangle^3, \\ k_4 &= \langle Q^4 \rangle - 4 \langle Q^3 \rangle \langle Q \rangle - 3 \langle Q^2 \rangle^2 + 12 \langle Q^2 \rangle \langle Q \rangle^2 - 6 \langle Q \rangle^4. \end{aligned} \tag{2.8}$$

All these averages are computed by using the weight  $e^{-S(\theta_I)}$  in the path-integral.

Let us now describe what changes on the lattice: the lattice action is

$$S_L(\theta_L) = S_W - \theta_L Q_{\text{clov}}, \tag{2.9}$$

where the  $\theta$ -term is discretized by using the non-smoothed clover charge  $Q_{\text{clov}}$  defined in the previous subsection, and  $\theta_L$  is the bare imaginary- $\theta$  coupling. The reason for using  $Q_{\text{clov}}$  is that with this choice standard heatbath and overrelaxation algorithms can be used in the update. Relations analogous to eq. (2.7) can be obtained, where  $\theta_I = Z\theta_L$  and  $Z$  is the renormalization constant appearing in eq. (2.3). Measuring the cumulants for several values of  $\theta_L$  we can thus fit the values of  $\chi$ ,  $b_{2n}$  and  $Z$ . We explicitly note that the cumulants are not affected by the renormalization of  $Q_{\text{clov}}$  since they are evaluated by using the smoothed and rounded charge  $Q_L$  introduced in the previous section (see ref. [38] for a more detailed discussion on this point).

### 2.3 Parallel tempering of volume defect

The lattice action in eq. (2.9), as the standard Wilson action, is linear in each of the link variables, hence standard heatbath [65, 66] and overrelaxation [67] algorithms can be applied to update the gauge configurations when using the imaginary- $\theta$  method. However,

as anticipated in the introduction, the local nature of this updating procedure results in a slowing down of the topological modes, which is exponential both in the inverse lattice spacing and in the number  $N$  of colors. This makes very difficult to perform controlled continuum extrapolations for large values of  $N$ , since even simulations with moderately small values of the lattice spacing become prohibitively expensive as  $N$  is increased.

To mitigate this problem, we adopt, in this work, the parallel tempering algorithm proposed for the  $CP^{N-1}$  models in ref. [52], where this algorithm has been shown to perform as well as simulations with open boundaries while bypassing their complications related to finite-size effects. Moreover, as shown for  $CP^{N-1}$  models in ref. [38], parallel tempering can be easily applied in combination with the imaginary- $\theta$  method discussed in the previous subsection.

In this algorithm we consider  $N_r$  identical systems, differing only for the boundary conditions on a cuboid defect located on a given spatial slice. In particular, the boundary conditions imposed on the links that orthogonally cross the defect are chosen so that the different copies interpolate between periodic boundary conditions (pbc) and open boundary conditions (obc). Each system is evolved independently using standard local algorithms, and different copies are exchanged from time to time with a Metropolis step, so that the strong reduction of the autocorrelation time achieved in the obc copy is transferred to the pbc one, on which the measure of the cumulants of the topological charge  $Q_L$  is performed. Since the injection (or ejection) of topological charge in the system is mainly triggered by the update of the links close to the defect, it is convenient [52] to alternate updating sweeps over the whole lattice with hierarchic updates over sub-regions of the lattice centered around the defect. In particular, we updated more frequently small hyper-rectangular regions centered around the defect.

In our simulations the location of the defect is the tridimensional region

$$D = \left\{ x_1 = L - a, 0 \leq x_2 < L_d^{(2)}, 0 \leq x_3 < L_d^{(3)}, 0 \leq x_4 < L_d^{(4)} \right\},$$

however, after every hierarchic update, we perform a random translation of the pbc copy by one lattice spacing, thus effectively moving the location of the defect. For the sake of the simplicity we use a cubic defect  $L_d^{(2)} = L_d^{(3)} = L_d^{(4)} \equiv L_d$  and it is sufficient to choose  $L_d$  equal to a few lattice spacings to obtain satisfactory performances. For a discussion on how the choice of  $L_d$  affects the efficiency of the algorithm, see section 3.1.

In order to specify how the different boundary conditions across the defect are implemented, it is convenient to rescale each link of every replica according to

$$U_\mu^{(r)}(x) \rightarrow K_\mu^{(r)}(x)U_\mu^{(r)}(x),$$

where  $U_\mu^{(r)}(x)$  indicates a link of the  $r^{\text{th}}$  replica and the explicit expression of  $K_\mu^{(r)}(x)$  is:

$$K_\mu^{(r)}(x) = \begin{cases} c(r), & \text{if } \mu \neq 1 \text{ and } x \in D, \\ 1, & \text{otherwise,} \end{cases}$$

so that only the links crossing the volume defect are affected by its presence. For the pbc replica (corresponding to  $r = 0$ ) we have  $c(0) = 1$ , for the obc replica (corresponding to

$r = N_r - 1$ ) we have  $c(N_r - 1) = 0$ , for  $0 < r < N_r - 1$  the value of  $c(r)$  interpolates between 0 and 1. With these notations the action of the  $r^{\text{th}}$  copy reads

$$S_L^{(r)}(\theta_L) = S_W^{(r)} + S_\theta^{(r)}(\theta_L) \\ = -\frac{\beta}{N} \sum_{x, \mu > \nu} K_{\mu\nu}^{(r)}(x) \Re \text{Tr} \left\{ \Pi_{\mu\nu}^{(r)}(x) \right\} - \theta_L Q_{\text{clov}} \left[ U_\mu^{(r)}(x) \right],$$

where  $K_{\mu\nu}^{(r)}(x)$  is a short-hand for

$$K_{\mu\nu}^{(r)}(x) \equiv K_\mu^{(r)}(x) K_\nu^{(r)}(x + \hat{\mu}) K_\mu^{(r)}(x + \hat{\nu}) K_\nu^{(r)}(x).$$

Note that we chose to keep the  $\theta$  term insensitive to the presence of the defect, which only affects the Wilson part of the action. Exploratory simulations performed by modifying also the  $\theta$  term provided evidence that this choice does not significantly affects the performance of the algorithm, analogously to what was found for  $CP^{N-1}$  models in ref. [38]. This is not surprising since the barriers between the topological sectors, responsible of the critical slowing down, stem essentially from the Wilson term.

The swap of replicas was proposed after every step of hierarchic update, for every couple of adjacent copies  $r$  and  $r + 1$  (differentiating the cases of even and odd  $r$  in order to avoid synchronization problems), and was accepted with the Metropolis probability

$$p = \min \{ 1, \exp \{ -\Delta S \} \} \\ = \min \{ 1, \exp \{ -S_W^{\text{swap}} + S_W^{\text{no swap}} \} \}, \tag{2.10}$$

where the  $\theta$ -term, which is not affected by the defect, does not enter the acceptance probability. Since boundary conditions of different replicas differ only on a sub-region of the lattice, to compute  $\Delta S$  it is sufficient to sum the contributions to the action of the plaquettes centered on sites lying in a hyper-cuboid region centered around the defect and extending one lattice spacing from it.

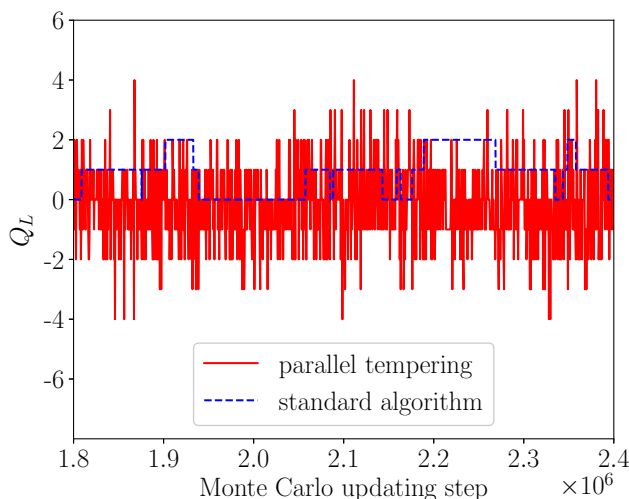
For the parallel tempering to be effective in decorrelating the topological modes of the pbc copy there must be no bottleneck for a configuration in the obc copy to be swapped toward the pbc one. In order to guarantee a random walk without barrier of the configurations between the different replicas it is thus convenient to choose the constants  $c(r)$  entering  $K_\mu^{(r)}(x)$  in such a way that the acceptance ratio is constant for all the proposed swaps:  $p(0, 1) = p(1, 2) = \dots = p(N_r - 2, N_r - 1)$ .

### 3 Numerical results

In order to compare the performances of parallel tempering with the ones of the standard algorithm, we performed simulations for  $N = 4$  and  $6$  with both algorithms and for several values of  $\beta$  at  $\theta_L = 0$ , measuring the autocorrelation time of  $Q_L^2$ . We then performed simulations at non-zero  $\theta_L$  with parallel tempering for each value of  $\beta$ , to estimate  $\chi$ ,  $b_2$  and  $b_4$  using the imaginary- $\theta$  method.

In table 1 we summarize the parameters of the performed simulations. Lattice sizes have been chosen to ensure that  $L\sqrt{\sigma} \gtrsim 3$ , where  $\sigma \simeq (440 \text{ MeV})^2$  is the string tension, so





**Figure 1.** Monte Carlo time evolution of the lattice topological charge  $Q_L$  for a run with  $N = 6$ ,  $\beta = 25.75$  and  $\theta_L = 0$ . For the comparison to be fair, data for the parallel tempering case are plotted as a function of the total number of global updates performed on all the replicas, i.e. 17 times the number of parallel tempering updates.

as to have finite-size effects under control [13, 68]. Statistics reported in table 1 refer to the parallel tempering simulations and are reported in number of parallel tempering steps. A single step of tempering update consists first of all of a complete update of each replica, using 5 lattice sweeps of over-relaxation [67] followed by 1 lattice sweep of heat-bath [65, 66], both implemented *à la* Cabibbo-Marinari [69], i.e., updating all the  $N(N - 1)/2$  diagonal  $SU(2)$  subgroups of  $SU(N)$ . After this “global” update step we perform an iteration on the sub-lattices entering the hierarchical update (see section 2.3), each iteration consisting of

- a local update sweep of the sub-lattices for every replica, using the same combination of local algorithms adopted for the global update;
- the parallel tempering swap proposal;
- a random translation of the pbc copy by one lattice spacing.

Since each system is updated using the same procedure and the time required for hierarchic updates, swap and translations is negligible with respect to the time of the global update, the total numerical effort of a single parallel tempering step is  $\sim N_r$  times larger than the one required for the local update in the standard setup.

### 3.1 Parallel tempering: results and comparison

Just by inspection of the time histories of the topological charge it is simple to realize that parallel tempering substantially reduces the topological freezing allowing us to perform simulations at values of the lattice spacings which would have been otherwise prohibitive with the standard algorithm. An example of Monte Carlo time evolution of  $Q_L$  obtained

$N = 4$						
$\beta$	$L/a$	$a\sqrt{\sigma}$	$L\sqrt{\sigma}$	$\theta_L^{\max}$	Stat. $\theta = 0$	Stat. $\theta \neq 0$
11.104	16	0.1981(5)*	3.17	15	255k	787k
11.347	20	0.1590(6)	3.17	15	1.39M	2.44M

$N = 6$						
$\beta$	$L/a$	$a\sqrt{\sigma}$	$L\sqrt{\sigma}$	$\theta_L^{\max}$	Stat. $\theta = 0$	Stat. $\theta \neq 0$
24.768	12	0.2912(11)	3.57	15	103k	257k
24.845	12	0.2801(13)	3.41	15	113k	166k
25.056	12	0.2499(10)	3.04	15	228k	280k
25.394	14	0.2143(8)	3.00	15	513k	553k
25.750	16	0.1878(18)	3.00	17.5	1.12M	1.84M

**Table 1.** Summary of simulation parameters. Simulations at non-zero values of  $\theta_L$  were performed in steps of  $\Delta\theta_L = 2$  up to  $\theta_L = 10$  and in steps of  $\Delta\theta_L = 2.5$  for  $\theta_L > 10$ . The last column refers to the total statistics accumulated for all imaginary- $\theta$  simulations. The defect length was, in all cases,  $L_d/a = 2$ . All simulations for  $N = 4$  were performed using  $N_r = 10$ , corresponding to a constant swap probability  $p$  around 20%, while simulations for  $N = 6$  used  $N_r = 17$ , corresponding to  $p \approx 30\%$ . Lattice spacings are taken from ref. [70] or interpolation/extrapolation of data thereof, except for the one marked with \*, which comes from ref. [68].

with parallel tempering for  $N = 6$ ,  $\beta = 25.75$  and  $\theta_L = 0$  is shown in figure 1, where we compare it with the evolution obtained with the standard algorithm.

In order to quantitatively characterize the gain achieved with parallel tempering and optimize its efficiency, it is useful to study the autocorrelation time of the topological susceptibility. We use as definition of the autocorrelation time for the generic observable  $\mathcal{O}$  the expression [71]

$$\tau(\mathcal{O}) = \frac{1}{2} \left( \frac{\Delta_{\mathcal{O}}^{\text{binned}}}{\Delta_{\mathcal{O}}^{\text{naive}}} - 1 \right)^2, \tag{3.1}$$

where  $\Delta_{\mathcal{O}}^{\text{binned}}$  is the error associated to  $\langle \mathcal{O} \rangle$  by using a self-consistent binning analysis, while  $\Delta_{\mathcal{O}}^{\text{naive}}$  is the usual standard error of the mean for independent identically distributed samples. The autocorrelation time of the topological susceptibility, however, does not take into account the increased computational effort of the parallel tempering algorithm with respect to the standard local algorithms. As a figure of merit for the computational effectiveness of parallel tempering, it is thus convenient to introduce the effective autocorrelation time

given by

$$\tau_{\text{pt}}(\mathcal{O}) = \tau(\mathcal{O})N_r. \tag{3.2}$$

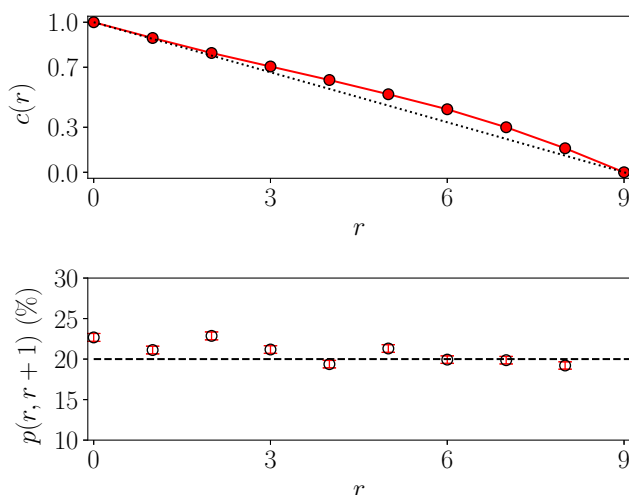
As discussed in the previous section, in every parallel tempering simulation we tuned the parameters  $c(r)$  in such a way that the acceptance  $p(r, r + 1)$  of the Metropolis swap move between the replicas  $r$  and  $r + 1$  is approximately independent of  $r$ . This tuning was performed using test simulations at  $\theta_L = 0$  (acceptances do not depend on  $\theta_L$ ), and in figure 2 we show an example of the behavior of  $c(r)$  for a run with  $N = 4$  and  $\beta = 11.347$ . Deviations from the linear behavior appear to be small in the optimal  $c(r)$  values, however using these values  $\tau_{\text{pt}}(Q_L^2)$  is about half the one obtained by using a simple linear interpolation. Once  $p(r, r + 1)$  is almost independent of  $r$ , configurations move freely among the different replicas following a random walk, as shown in figure 3.

The constant value  $p$  that is reached by  $p(r, r + 1)$  after the tuning of  $c(r)$  obviously depends on the number  $N_r$  of replicas used. We did not perform a systematic investigation of the dependence on  $p$  of the numerical effectiveness of parallel tempering, however this dependence seems to be quite mild. Indeed by increasing the number of replicas  $N_r$ , the constant acceptance probability  $p$  grows and the autocorrelation time  $\tau(Q_L^2)$  is reduced, however also the computational cost increases with  $N_r$ . The net effect is that  $\tau_{\text{pt}}(Q_L^2)$  is largely insensitive to the specific value of  $p$ , at least as far as it is not too close to 0 or 1, as we verified in some test simulations. For example, simulations performed with  $N = 4$  and  $\beta = 11.104$  using  $p \simeq 20\%$  (achieved with  $N_r = 10$ ) or  $p \simeq 30\%$  (corresponding to  $N_r = 12$ ) provided consistent values of  $\tau_{\text{pt}}(Q_L^2)$ : 72(10) and 78(18) respectively. The value of  $p$  was kept fixed while approaching the continuum limit, and in all cases we found that this could be achieved by using the same value of  $N_r$  for the different lattice spacings (at fixed physical volume).

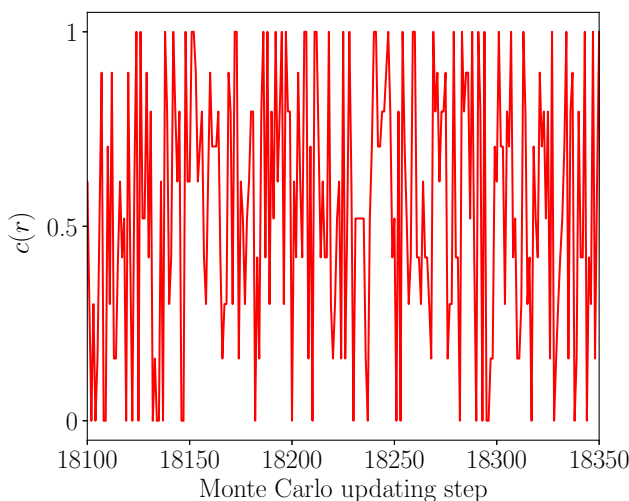
Most of the simulations reported in this work used  $L_d/a = 2$  for the size of the defect, a value that is sufficient to drastically reduce the freezing problem. To investigate the dependence of  $\tau_{\text{pt}}$  on  $L_d$  some more simulations have been performed with  $L_d/a = 3$  and  $L_d/a = 1$  for the case  $N = 6$ , always keeping the swap acceptance probability  $p$  fixed to about 30%, which requires to scale the number of replicas approximately as  $\sim \sqrt{L_d^3}$ .

A complete list of the obtained autocorrelation times  $\tau_{\text{pt}}(Q_L^2)$  is reported in table 2, where they are also compared with the results obtained in ref. [13] using just local algorithms. The scaling of  $\tau_{\text{pt}}(Q_L^2)$  with  $1/(a\sqrt{\sigma})$  for the case  $N = 6$  is instead shown in figure 4.

Simulations performed for  $N = 6$  at  $\beta = 24.845$  (corresponding to the points at  $1/(a\sqrt{\sigma}) \simeq 3.57$  in figure 4) using  $L_d/a = 1, 2, 3$  show that  $L_d/a = 2$  is the optimal choice for this value of the coupling. As can be seen from figure 4, autocorrelation times extracted from simulations performed at fixed  $L_d/a = 2$  are much smaller than the corresponding ones obtained from simulation using local update algorithms, also for smaller values of the lattice spacing. However  $\tau_{\text{pt}}(Q_L^2)$  still seems to scale exponentially with the inverse lattice spacing. This is due to the fact that, by approaching the continuum limit at fixed  $L_d/a$ , the size of the defect in physical units is reduced, and the mechanism of injection of topological charge through the defect becomes less and less efficient.



**Figure 2.** Behavior of  $c(r)$  as a function of the replica index  $r$  compared to a simple linear behavior (figure above), along with the corresponding acceptances ( $\sim 20\%$ ) for the swap between copies  $r$  and  $r + 1$  (figure below). Data refer to a run with  $N = 4$ ,  $\beta = 11.347$  and  $\theta_L = 0$ .



**Figure 3.** Random walk of a configuration among different replicas during a parallel tempering run for  $N = 4$ ,  $\beta = 11.347$  and  $\theta_L = 0$ . Replicas are parametrized by their value of  $c(r)$  and the Monte Carlo time is expressed in units of the parallel tempering update step defined in the text.

If instead  $L_d$  is kept fixed in physical units while approaching the continuum limit, one generically expects a polynomial critical slowing down in  $1/(a\sqrt{\sigma})$ . To investigate this point we performed additional simulations at  $\beta = 25.75$  using  $L_d/a = 3$ , in order to have at this lattice spacing a defect of the same physical size as the one corresponding to  $L_d/a = 2$  at  $\beta = 24.845$  (in both the cases  $L_d\sqrt{\sigma} \sim 0.56$ ). The outcome of this test is that, despite a  $\approx 33\%$  reduction of the lattice spacing, the effective autocorrelation time  $\tau_{pt}(Q_L^2)$  is compatible in the two cases, as reported in table 2 and shown in figure 4.

$N = 4$						
$\beta$	$a\sqrt{\sigma}$	$L_d/a$	$p(\%)$	$N_r$	$\tau_{\text{pt}}(Q_L^2)$	$\tau_{\text{std}}(Q_L^2)$
11.104	0.1981(5)	2	20	10	72(10)	140(10)*
			30	12	78(18)	
11.347	0.1590(6)	2	20	10	380(80)	1000(200)

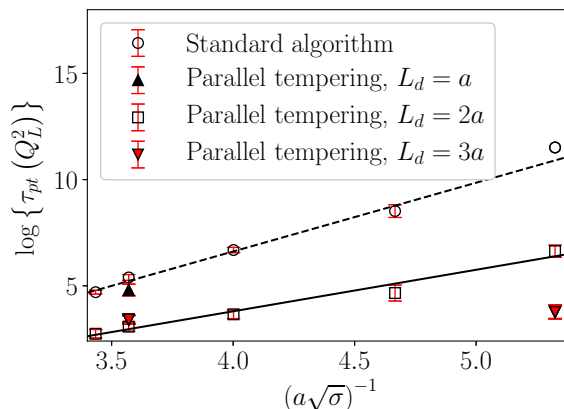
$N_c = 6$						
$\beta$	$a\sqrt{\sigma}$	$L_d/a$	$p(\%)$	$N_r$	$\tau_{\text{pt}}(Q_L^2)$	$\tau_{\text{std}}(Q_L^2)$
24.768	0.2912(11)	2	30	17	16(3)	110(10)*
24.845	0.2801(13)	1	30	7	120(30)	220(30)*
		2		17	22(5)	
		3		29	30(6)	
25.056	0.2499(10)	2	30	17	39(8)	800(100)*
25.394	0.2143(8)	2	30	17	110(40)	5000(1500)
25.750	0.1874(8)	2	30	17	760(200)	$\sim 10^{5**}$
		3		30	43(11)	

**Table 2.** Results for the autocorrelation time of  $Q_L^2$  obtained by using the standard and the Hasenbusch algorithm. Quantities denoted with \* are taken from ref. [13], where the procedure used for the local update was the same of the present work. The result denoted by \*\* is just a rough estimate, since it was impossible to obtain a reliable value even after  $\sim 2.5$ M trajectories.

These results still do not permit to make a clear assessment about the scaling and the optimal tuning of the parallel tempering algorithm towards the continuum limit, however, altogether they give a strong indication that it works exceedingly well, compared to standard algorithms, in reducing topological freezing, and that the best scaling is obtained keeping  $L_d$  in the range  $0.2 - 0.3$  fm. All this is consistent with what is observed in two-dimensional  $CP^{N-1}$  models [38, 52], where the continuum limit is performed at fixed  $L_d/\xi$ , i.e. at fixed physical size of the defect.

### 3.2 Analytic continuation and continuum limit

In table 3 we summarize the results obtained for the topological observables  $\chi$  and  $b_2$  at different values of  $N$  and of the lattice spacing, obtained by fitting the  $\theta$  dependence of the



**Figure 4.** Scaling of  $\tau_{pt}(Q_L^2)$  with the inverse lattice spacing obtained by using the local algorithms or parallel tempering for  $N = 6$ . The scaling of the autocorrelation time obtained with the standard algorithm and with parallel tempering at fixed  $L_d/a = 2$  are both compatible with an exponential scaling in  $1/a$  (dashed and solid line). Best fits performed with the fit function  $\log\{\tau_{pt}\} = k_0 + k_1/(a\sqrt{\sigma})$  yield  $k_1 = 3.24(20)$  and  $1.95(17)$  for the standard and the parallel tempering updates respectively.

cumulants as described in section 2.2. An example of imaginary- $\theta$  fit of the cumulants is shown in figure 5 for the case  $N = 6$  and  $\beta = 25.75$ .

In all the cases we found sufficient to fit the first three cumulants, as the addition of the fourth one did not change the obtained results. Moreover, in all cases we found the  $O(\theta_L^6)$  term in the expansion of the vacuum energy to be well compatible with zero since no signal above zero is observed for  $b_4$ . In particular, we find  $|b_4(N = 4)| \cdot 10^5 \lesssim 15$  and  $|b_4(N = 6)| \cdot 10^5 \lesssim 30$ . For this reason, results for  $a^4\chi$  and  $b_2$  reported in table 3 have been obtained by neglecting  $b_4$  in eqs. (2.7). Finally we note that correlations between the different cumulants are small and do not significantly affect the result of the fit, as we explicitly checked by performing both correlated and uncorrelated fits.

We used our data for  $N = 4$  and  $N = 6$ , as well as data obtained for larger lattice spacings taken from ref. [13], to extrapolate continuum results for  $\chi/\sigma^2$  and  $b_2$ . For the topological susceptibility the improvement with respect to previously available results is only marginal, since the dominant source of error comes from the string tension  $\sigma$  used to set the scale. For  $b_2$ , instead, we achieved a substantial improvement of the state of the art, both for  $N = 4$  and  $N = 6$ . In particular for  $N = 6$  parallel tempering allowed us to reach much finer lattice spacings than the ones used in previous studies. In this way we could perform for the first time a controlled continuum extrapolation of  $b_2$  in this case, while in ref. [13] only a reasonable confidence interval was reported. In table 4 we summarize our continuum limits, while in figure 6 we report our continuum extrapolations.

### 3.3 Large- $N$ limit

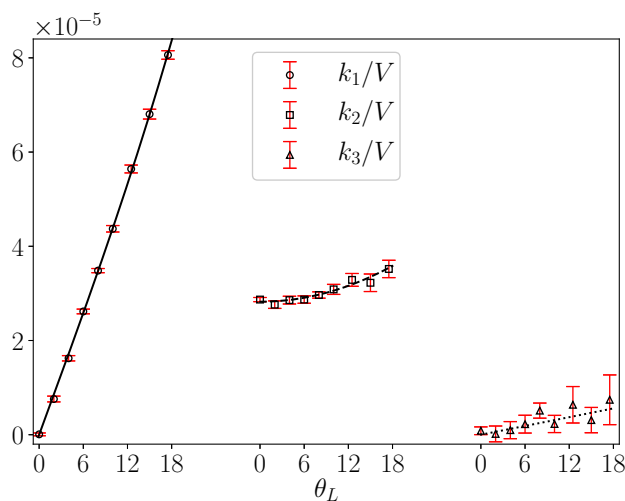
In this section we revisit the large- $N$  extrapolation on the basis of our improved results in particular we report our estimates of  $\bar{\chi}$  and  $\bar{b}_2$  introduced in eq. (1.4). Let us start from the topological susceptibility: following large- $N$  expectations, we fitted our data for  $N \geq 3$

$N = 4$				
$\beta$	$a\sqrt{\sigma}$	$Z$	$a^4\chi \cdot 10^5$	$b_2 \cdot 10^4$
11.104	0.1981(5)	0.14742(96)	4.183(28)	-137.0(7.5)
11.347	0.1590(6)	0.1747(23)	1.691(22)	-128(11)

$N = 6$				
$\beta$	$a\sqrt{\sigma}$	$Z$	$a^4\chi \cdot 10^5$	$b_2 \cdot 10^4$
24.768	0.2912(11)	0.10300(52)	18.371(93)	-77.2(8.3)
24.845	0.2801(13)	0.10945(68)	15.205(94)	-73.6(9.1)
25.056	0.2499(10)	0.12053(88)	9.719(69)	-72.9(8.6)
25.394	0.2143(8)	0.1382(12)	5.120(44)	-65.1(7.6)
25.750	0.1878(13)	0.1518(15)	2.816(27)	-58.7(7.4)

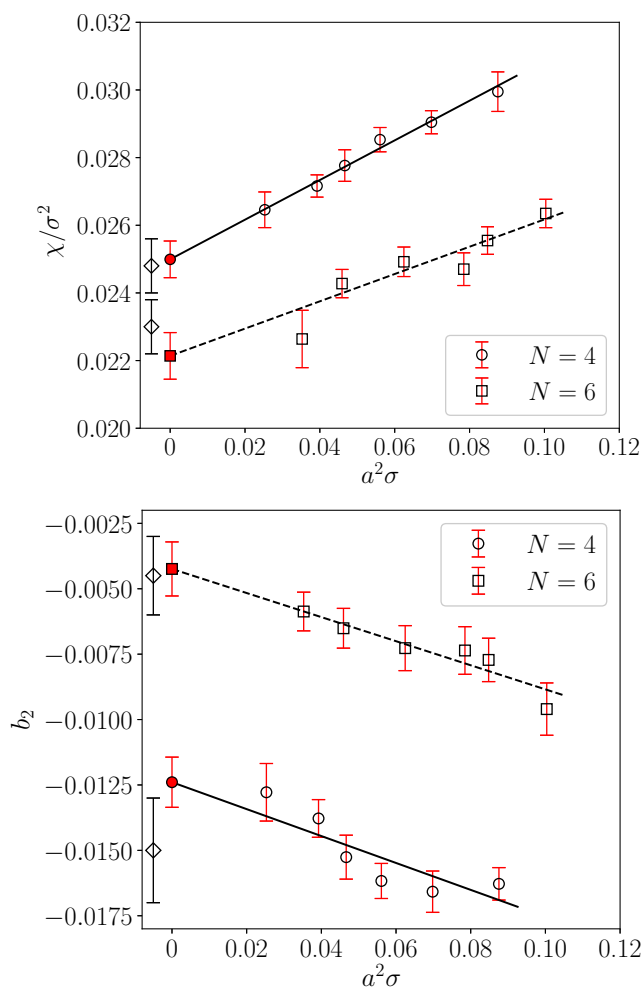
**Table 3.** Summary of the results obtained using the imaginary- $\theta$  fit for  $N = 4$  and 6 by considering up to  $O(\theta_L^4)$  terms in the Taylor expansion of the vacuum energy, i.e., neglecting  $b_4$  in eqs. (2.7).



**Figure 5.** Best fit of the first 3 cumulants (solid, dashed and dotted line respectively) for  $N = 6$ ,  $\beta = 25.750$  and  $\theta_L \in [0, 17.5]$ , obtained considering up to  $O(\theta_L^4)$  terms in the Taylor expansion of the vacuum energy, i.e., neglecting  $b_4$  in eqs. (2.7). The best fit yields  $\chi^2/\text{dof} = 14.4/24$ .

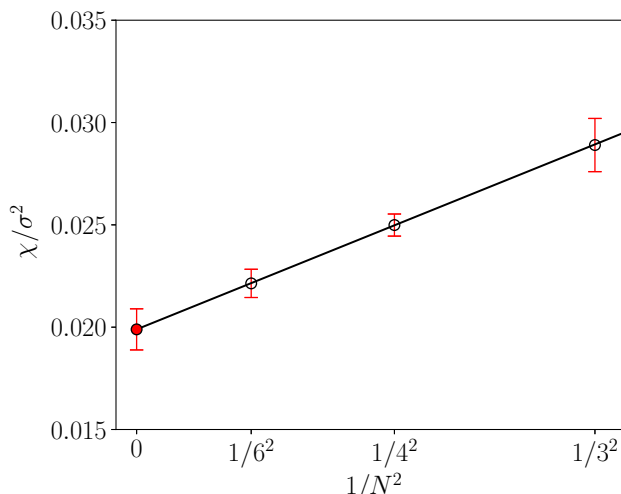
$N$	$\chi/\sigma^2$	$b_2$
3	0.0289(13)	-0.0216(15)
4	0.02499(54)	-0.01240(96)
6	0.02214(69)	-0.0042(10)

**Table 4.** Summary of continuum extrapolations for  $N = 3, 4$  and  $6$ . Values for  $N = 3$  are taken from ref. [25].



**Figure 6.** Continuum extrapolations of  $\chi/\sigma^2$  (above) and  $b_2$  (below) for  $N = 4$  and  $6$  (solid and dashed line respectively) obtained fitting linear corrections in  $a^2\sigma$  to the continuum limit. The reported best fits yield, respectively for  $N = 4$  and  $6$ ,  $\chi^2/\text{dof} = 0.8/4$  and  $3.7/4$  for  $\chi/\sigma^2$  and  $\chi^2/\text{dof} = 4.9/4$  and  $1.2/4$  for  $b_2$ . The diamond points represent the determinations reported in ref. [13] for  $N = 4$  and  $6$ .





**Figure 7.** Extrapolation of  $\chi/\sigma^2$  towards the large- $N$  limit using fit function  $\chi/\sigma^2 = \bar{\chi}/\sigma^2 + k/N^2$ . Best fit yields  $\bar{\chi}/\sigma^2 = 0.0199(10)$  and  $k = 0.082(17)$ .

using the functional form:

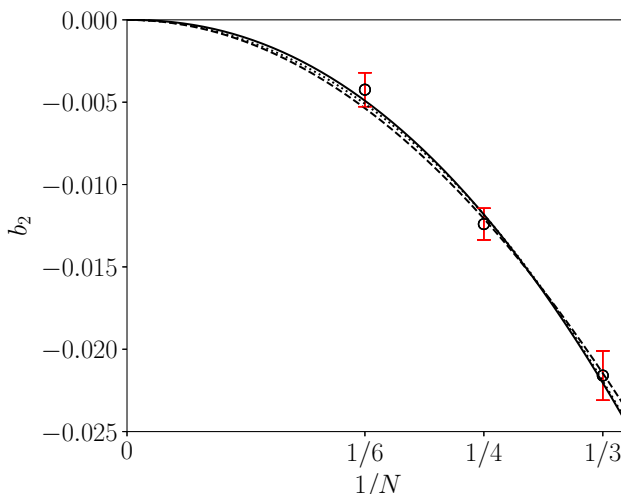
$$\frac{\chi}{\sigma^2} = \frac{\bar{\chi}}{\sigma^2} + \frac{k}{N^2} + O\left(\frac{1}{N^4}\right). \quad (3.3)$$

Our data are in agreement with the expected large- $N$  scaling and we find the result  $\bar{\chi}/\sigma^2 = 0.0199(10)$ ; the best fit is shown in figure 7 together with numerical results. As already observed, our result does not improve the previous determination  $\bar{\chi}/\sigma^2 = 0.0209(11)$  of ref. [13] as the main source of errors comes from the string tension used to set the scale. Using  $\Lambda_{\text{large-}N}/\sqrt{\sigma} = 0.525(2)$  [72] and  $\Lambda_{\text{large-}N} = 242(10)$  MeV [73] to convert to physical units we get  $\bar{\chi}^{1/4} = 173(8)$  MeV, in agreement with the prediction  $\bar{\chi}^{1/4} \simeq 180$  MeV obtained from the Witten-Veneziano formula.

We now pass to the discussion of the large- $N$  behavior of  $b_2$ . According to the standard large- $N$  arguments, we expect a behavior of the type:

$$b_2 = \frac{\bar{b}_2}{N^2} + \frac{\bar{b}_2^{(1)}}{N^4} + O\left(\frac{1}{N^6}\right). \quad (3.4)$$

To test this prediction we perform a best fit of our data with  $N \geq 3$  using the power law  $b_2(N) = \bar{b}_2/N^c$ , obtaining a perfect agreement with expectations, since the exponent results  $c = 2.17(26)$ , which improves the previous result  $c = 2.0(4)$  reported in ref. [13]. The obtained best fit is shown in figure 8. By fixing the exponent  $c = 2$  and fitting our data with just the leading behavior  $b_2 = \bar{b}_2/N^2$  in the ranges  $N \geq 3$  and  $N \geq 4$ , we obtain the results  $\bar{b}_2 = -0.1931(98)$  and  $\bar{b}_2 = -0.192(14)$  respectively. Since the curve profiles obtained in these two cases are practically indistinguishable, we only show the former in figure 8. As our final result, we quote the value  $\bar{b}_2 = -0.193(10)$ , which improves on the previous determination  $\bar{b}_2 = -0.23(3)$  of ref. [13].



**Figure 8.** Extrapolation of  $b_2$  towards the large- $N$  limit. The solid line represents the best fit obtained using fit function  $b_2 = \bar{b}_2/N^c$  in the whole range; the dashed line represents the best fit obtained using the same fit function but with fixed  $c = 2$  in the whole range; the dotted line represents the best fit obtained in the whole range including a further  $\bar{b}_2^{(1)}/N^4$  term in the fit function. The best fits yield, respectively,  $\bar{b}_2 = -0.238(79)$ ,  $-0.1931(98)$  and  $-0.179(31)$ . The free exponent results  $c = 2.17(26)$ , while the next-to-leading correction is  $\bar{b}_2^{(1)} = -0.17(35)$ .

#### 4 Conclusions

In this work we investigated the  $\theta$ -dependence of  $SU(N)$  Yang-Mills theories at zero temperature using the parallel tempering algorithm proposed by Hasenbusch in ref. [52]. This algorithm was originally tested in two dimensional  $CP^{N-1}$  models, and a first extension of the original proposal has already been performed in ref. [38] (still for  $CP^{N-1}$  models), by extending the parallel tempering approach to simulations at imaginary  $\theta$  values. In the present work we implemented the same setup in  $SU(N)$  Yang-Mills theory, thus proving the feasibility of the approach also for computationally more demanding models, and improving the state of the art results for the  $\theta$  dependence of these models in the large  $N$  limit.

The idea of the method is to simulate many independent identical systems differing only for the boundary conditions imposed on a cubic defect  $L_d \times L_d \times L_d$ , which are chosen to interpolate between open (obc) and periodic boundary conditions (pbc). Each replica evolves independently, and swaps among them are proposed from time to time in order to transfer configurations from the obc to the pbc replica. In this way a drastic reduction of the autocorrelation time of the topological charge is achieved, while avoiding the complication related to the breaking of translation invariance connected to the adoption of open boundary conditions, since measures are performed on the pbc replica.

By using the parallel tempering algorithm we got an impressive reduction of the autocorrelation times of topological observables, which for the smallest lattice spacing used was of at least two order of magnitude when taking into account also the larger computational complexity. A nice feature of the algorithm is that this gain was obtained without optimally tuning all the possible parameters entering the update, which proves the robustness

of the approach. The most relevant parameter to be fixed is clearly the size of the defect, and we verified that for the cases studied in this paper a size in the range 0.2–0.3 fm is sufficiently close to optimal to obtain a huge reduction of the critical slowing down.

The possibility of performing simulations at smaller lattice spacing than in previous studies allowed us to achieve a substantial improvement in the determination of  $\theta$ -dependence beyond the leading  $O(\theta^2)$  order. In particular, we improved the accuracy of the determination of the coefficients  $b_2$  for both  $N = 4$  and  $N = 6$ , in the last case also performing a controlled continuum limit extrapolation, which was not possible with previously available results. These data confirm that  $b_2$  scales with the number of colors in a way that is consistent with the leading behavior predicted by large- $N$  arguments: data for  $N \geq 3$  are perfectly compatible with a scaling of the form  $b_2 = \bar{b}_2/N^2$ , with  $\bar{b}_2 = -0.193(10)$ , while a best fit according to  $b_2 = \bar{b}_2/N^c$  returns the value  $c = 2.17(26)$  for the free exponent. This shows that the scaling of our data is consistent with the leading expected behavior, and it is thus reasonable to neglect sub-leading corrections. We however explicitly note that the accuracy is still not high enough to exclude possible unconventional scenarios like the one put-forward in ref. [74], or to better investigate if critical corrections emerge at small  $N$  [75].

A further refinement of present results, including also new values of  $N$ , would thus be welcome in the future: the algorithm proposed in ref. [52], and extended to  $SU(N)$  gauge theories in this study, will permit such systematic refinement.

## Acknowledgments

Numerical simulations have been performed on the MARCONI machine at CINECA, based on the agreement between INFN and CINECA (under projects INF19\_npqcd and INF20\_npqcd).

**Open Access.** This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

## References

- [1] G. 't Hooft, *A planar diagram theory for strong interactions*, *Nucl. Phys. B* **72** (1974) 461 [[INSPIRE](#)].
- [2] G. 't Hooft, *Symmetry breaking through Bell-Jackiw anomalies*, *Phys. Rev. Lett.* **37** (1976) 8 [[INSPIRE](#)].
- [3] E. Witten, *Instantons, the quark model, and the  $1/n$  expansion*, *Nucl. Phys. B* **149** (1979) 285 [[INSPIRE](#)].
- [4] E. Witten, *Current algebra theorems for the U(1) Goldstone boson*, *Nucl. Phys. B* **156** (1979) 269 [[INSPIRE](#)].
- [5] G. Veneziano, *U(1) without instantons*, *Nucl. Phys. B* **159** (1979) 213 [[INSPIRE](#)].
- [6] E. Witten, *Large  $N$  chiral dynamics*, *Annals Phys.* **128** (1980) 363 [[INSPIRE](#)].

- [7] P. Di Vecchia and G. Veneziano, *Chiral dynamics in the large  $N$  limit*, *Nucl. Phys. B* **171** (1980) 253 [INSPIRE].
- [8] E. Witten, *Theta dependence in the large  $N$  limit of four-dimensional gauge theories*, *Phys. Rev. Lett.* **81** (1998) 2862 [hep-th/9807109] [INSPIRE].
- [9] A. D’Adda, M. Lüscher and P. Di Vecchia, *A  $1/N$  expandable series of nonlinear sigma models with instantons*, *Nucl. Phys. B* **146** (1978) 63 [INSPIRE].
- [10] M. Campostrini and P. Rossi,  *$1/N$  expansion of the topological susceptibility in the  $CP^{N-1}$  models*, *Phys. Lett. B* **272** (1991) 305 [INSPIRE].
- [11] L. Del Debbio, G.M. Manca, H. Panagopoulos, A. Skouroupathis and E. Vicari, *Theta-dependence of the spectrum of  $SU(N)$  gauge theories*, *JHEP* **06** (2006) 005 [hep-th/0603041] [INSPIRE].
- [12] P. Rossi, *Effective Lagrangian of  $CP^{N-1}$  models in the large  $N$  limit*, *Phys. Rev. D* **94** (2016) 045013 [arXiv:1606.07252] [INSPIRE].
- [13] C. Bonati, M. D’Elia, P. Rossi and E. Vicari,  *$\theta$  dependence of 4D  $SU(N)$  gauge theories in the large- $N$  limit*, *Phys. Rev. D* **94** (2016) 085017 [arXiv:1607.06360] [INSPIRE].
- [14] B. Alles, M. D’Elia and A. Di Giacomo, *Topological susceptibility at zero and finite  $T$  in  $SU(3)$  Yang-Mills theory*, *Nucl. Phys. B* **494** (1997) 281 [Erratum *ibid.* **679** (2004) 397] [hep-lat/9605013] [INSPIRE].
- [15] B. Alles, M. D’Elia and A. Di Giacomo, *Topology at zero and finite  $T$  in  $SU(2)$  Yang-Mills theory*, *Phys. Lett. B* **412** (1997) 119 [hep-lat/9706016] [INSPIRE].
- [16] L. Del Debbio, L. Giusti and C. Pica, *Topological susceptibility in the  $SU(3)$  gauge theory*, *Phys. Rev. Lett.* **94** (2005) 032003 [hep-th/0407052] [INSPIRE].
- [17] L. Del Debbio, H. Panagopoulos and E. Vicari, *Theta dependence of  $SU(N)$  gauge theories*, *JHEP* **08** (2002) 044 [hep-th/0204125] [INSPIRE].
- [18] M. D’Elia, *Field theoretical approach to the study of theta dependence in Yang-Mills theories on the lattice*, *Nucl. Phys. B* **661** (2003) 139 [hep-lat/0302007] [INSPIRE].
- [19] B. Lucini, M. Teper and U. Wenger, *Topology of  $SU(N)$  gauge theories at  $T \simeq 0$  and  $T \simeq T_c$* , *Nucl. Phys. B* **715** (2005) 461 [hep-lat/0401028] [INSPIRE].
- [20] 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** (2007) 094510 [arXiv:0705.2352] [INSPIRE].
- [21] E. Vicari and H. Panagopoulos, *Theta dependence of  $SU(N)$  gauge theories in the presence of a topological term*, *Phys. Rept.* **470** (2009) 93 [arXiv:0803.1593] [INSPIRE].
- [22] H. Panagopoulos and E. Vicari, *The 4D  $SU(3)$  gauge theory with an imaginary  $\theta$  term*, *JHEP* **11** (2011) 119 [arXiv:1109.6815] [INSPIRE].
- [23] M. Cè, C. Consonni, G.P. Engel and L. Giusti, *Non-Gaussianities in the topological charge distribution of the  $SU(3)$  Yang-Mills theory*, *Phys. Rev. D* **92** (2015) 074502 [arXiv:1506.06052] [INSPIRE].
- [24] M. Cè, M. García Vera, L. Giusti and S. Schaefer, *The topological susceptibility in the large- $N$  limit of  $SU(N)$  Yang-Mills theory*, *Phys. Lett. B* **762** (2016) 232 [arXiv:1607.05939] [INSPIRE].
- [25] C. Bonati, M. D’Elia and A. Scapellato,  *$\theta$  dependence in  $SU(3)$  Yang-Mills theory from analytic continuation*, *Phys. Rev. D* **93** (2016) 025028 [arXiv:1512.01544] [INSPIRE].

- [26] C. Bonati, M. Cardinali and M. D’Elia,  $\theta$  dependence in trace deformed SU(3) Yang-Mills theory: a lattice study, *Phys. Rev. D* **98** (2018) 054508 [[arXiv:1807.06558](#)] [[INSPIRE](#)].
- [27] C. Bonati, M. Cardinali, M. D’Elia and F. Mazziotti,  $\theta$ -dependence and center symmetry in Yang-Mills theories, *Phys. Rev. D* **101** (2020) 034508 [[arXiv:1912.02662](#)] [[INSPIRE](#)].
- [28] G. Bhanot and F. David, The phases of the O(3)  $\sigma$  model for imaginary  $\theta$ , *Nucl. Phys. B* **251** (1985) 127 [[INSPIRE](#)].
- [29] V. Azcoiti, G. Di Carlo, A. Galante and V. Laliena, New proposal for numerical simulations of theta vacuum-like systems, *Phys. Rev. Lett.* **89** (2002) 141601 [[hep-lat/0203017](#)] [[INSPIRE](#)].
- [30] B. Alles and A. Papa, Mass gap in the 2D O(3) non-linear sigma model with a  $\theta = \pi$  term, *Phys. Rev. D* **77** (2008) 056008 [[arXiv:0711.1496](#)] [[INSPIRE](#)].
- [31] M. Imachi, M. Kambayashi, Y. Shinno and H. Yoneyama, The  $\theta$ -term,  $CP^{N-1}$  model and the inversion approach in the imaginary  $\theta$  method, *Prog. Theor. Phys.* **116** (2006) 181 [[INSPIRE](#)].
- [32] S. Aoki et al., The electric dipole moment of the nucleon from simulations at imaginary vacuum angle  $\theta$ , [arXiv:0808.1428](#) [[INSPIRE](#)].
- [33] B. Alles, M. Giordano and A. Papa, Behavior near  $\theta = \pi$  of the mass gap in the two-dimensional O(3) non-linear sigma model, *Phys. Rev. B* **90** (2014) 184421 [[arXiv:1409.1704](#)] [[INSPIRE](#)].
- [34] M. D’Elia and F. Negro,  $\theta$  dependence of the deconfinement temperature in Yang-Mills theories, *Phys. Rev. Lett.* **109** (2012) 072001 [[arXiv:1205.0538](#)] [[INSPIRE](#)].
- [35] M. D’Elia and F. Negro, Phase diagram of Yang-Mills theories in the presence of a  $\theta$  term, *Phys. Rev. D* **88** (2013) 034503 [[arXiv:1306.2919](#)] [[INSPIRE](#)].
- [36] M. D’Elia, M. Mariti and F. Negro, Susceptibility of the QCD vacuum to CP-odd electromagnetic background fields, *Phys. Rev. Lett.* **110** (2013) 082002 [[arXiv:1209.0722](#)] [[INSPIRE](#)].
- [37] C. Bonanno, C. Bonati and M. D’Elia, Topological properties of  $CP^{N-1}$  models in the large- $N$  limit, *JHEP* **01** (2019) 003 [[arXiv:1807.11357](#)] [[INSPIRE](#)].
- [38] M. Berni, C. Bonanno and M. D’Elia, Large- $N$  expansion and  $\theta$ -dependence of 2d  $CP^{N-1}$  models beyond the leading order, *Phys. Rev. D* **100** (2019) 114509 [[arXiv:1911.03384](#)] [[INSPIRE](#)].
- [39] B. Alles, G. Boyd, M. D’Elia, A. Di Giacomo and E. Vicari, Hybrid Monte Carlo and topological modes of full QCD, *Phys. Lett. B* **389** (1996) 107 [[hep-lat/9607049](#)] [[INSPIRE](#)].
- [40] P. de Forcrand, M. Garcia Perez, J.E. Hetrick and I.-O. Stamatescu, Topology of full QCD, *Nucl. Phys. B Proc. Suppl.* **63** (1998) 549 [[hep-lat/9710001](#)] [[INSPIRE](#)].
- [41] B. Lucini and M. Teper, SU( $N$ ) gauge theories in four-dimensions: exploring the approach to  $N = \infty$ , *JHEP* **06** (2001) 050 [[hep-lat/0103027](#)] [[INSPIRE](#)].
- [42] D.B. Leinweber, A.G. Williams, J.-B. Zhang and F.X. Lee, Topological charge barrier in the Markov chain of QCD, *Phys. Lett. B* **585** (2004) 187 [[hep-lat/0312035](#)] [[INSPIRE](#)].
- [43] L. Del Debbio, G.M. Manca and E. Vicari, Critical slowing down of topological modes, *Phys. Lett. B* **594** (2004) 315 [[hep-lat/0403001](#)] [[INSPIRE](#)].
- [44] M. Lüscher and S. Schaefer, Lattice QCD without topology barriers, *JHEP* **07** (2011) 036 [[arXiv:1105.4749](#)] [[INSPIRE](#)].

- [45] A. Laio, G. Martinelli and F. Sanfilippo, *Metadynamics surfing on topology barriers: the  $CP^{N-1}$  case*, *JHEP* **07** (2016) 089 [[arXiv:1508.07270](#)] [[INSPIRE](#)].
- [46] J. Flynn, A. Juttner, A. Lawson and F. Sanfilippo, *Precision study of critical slowing down in lattice simulations of the  $CP^{N-1}$  model*, [arXiv:1504.06292](#) [[INSPIRE](#)].
- [47] C. Bonati et al., *Axion phenomenology and  $\theta$ -dependence from  $N_f = 2 + 1$  lattice QCD*, *JHEP* **03** (2016) 155 [[arXiv:1512.06746](#)] [[INSPIRE](#)].
- [48] C. Bonati and M. D’Elia, *Topological critical slowing down: variations on a toy model*, *Phys. Rev. E* **98** (2018) 013308 [[arXiv:1709.10034](#)] [[INSPIRE](#)].
- [49] A. Athenodorou and M. Teper, *The glueball spectrum of SU(3) gauge theory in 3 + 1 dimensions*, *JHEP* **11** (2020) 172 [[arXiv:2007.06422](#)] [[INSPIRE](#)].
- [50] E. Vicari, *Monte Carlo simulation of lattice  $CP^{N-1}$  models at large  $N$* , *Phys. Lett. B* **309** (1993) 139 [[hep-lat/9209025](#)] [[INSPIRE](#)].
- [51] W. Bietenholz, P. de Forcrand and U. Gerber, *Topological susceptibility from slabs*, *JHEP* **12** (2015) 070 [[arXiv:1509.06433](#)] [[INSPIRE](#)].
- [52] M. Hasenbusch, *Fighting topological freezing in the two-dimensional  $CP^{N-1}$  model*, *Phys. Rev. D* **96** (2017) 054504 [[arXiv:1706.04443](#)] [[INSPIRE](#)].
- [53] M. Campostrini, A. Di Giacomo and H. Panagopoulos, *The topological susceptibility on the lattice*, *Phys. Lett. B* **212** (1988) 206 [[INSPIRE](#)].
- [54] M. Lüscher, *Trivializing maps, the Wilson flow and the HMC algorithm*, *Commun. Math. Phys.* **293** (2010) 899 [[arXiv:0907.5491](#)] [[INSPIRE](#)].
- [55] M. Lüscher, *Properties and uses of the Wilson flow in lattice QCD*, *JHEP* **08** (2010) 071 [*Erratum ibid.* **03** (2014) 092] [[arXiv:1006.4518](#)] [[INSPIRE](#)].
- [56] B. Berg, *Dislocations and topological background in the lattice  $O(3)$   $\sigma$  model*, *Phys. Lett. B* **104** (1981) 475 [[INSPIRE](#)].
- [57] Y. Iwasaki and T. Yoshie, *Instantons and topological charge in lattice gauge theory*, *Phys. Lett. B* **131** (1983) 159 [[INSPIRE](#)].
- [58] S. Itoh, Y. Iwasaki and T. Yoshie, *Stability of instantons on the lattice and the renormalized trajectory*, *Phys. Lett. B* **147** (1984) 141 [[INSPIRE](#)].
- [59] M. Teper, *Instantons in the quantized SU(2) vacuum: a lattice Monte Carlo investigation*, *Phys. Lett. B* **162** (1985) 357 [[INSPIRE](#)].
- [60] E.-M. Ilgenfritz, M.L. Laursen, G. Schierholz, M. Muller-Preussker and H. Schiller, *First evidence for the existence of instantons in the quantized SU(2) lattice vacuum*, *Nucl. Phys. B* **268** (1986) 693 [[INSPIRE](#)].
- [61] M. Campostrini, A. Di Giacomo, H. Panagopoulos and E. Vicari, *Topological charge, renormalization and cooling on the lattice*, *Nucl. Phys. B* **329** (1990) 683 [[INSPIRE](#)].
- [62] B. Alles, L. Cosmai, M. D’Elia and A. Papa, *Topology in 2D  $CP^{N-1}$  models on the lattice: a critical comparison of different cooling techniques*, *Phys. Rev. D* **62** (2000) 094507 [[hep-lat/0001027](#)] [[INSPIRE](#)].
- [63] C. Bonati and M. D’Elia, *Comparison of the gradient flow with cooling in SU(3) pure gauge theory*, *Phys. Rev. D* **89** (2014) 105005 [[arXiv:1401.2441](#)] [[INSPIRE](#)].

- [64] C. Alexandrou, A. Athenodorou and K. Jansen, *Topological charge using cooling and the gradient flow*, *Phys. Rev. D* **92** (2015) 125014 [[arXiv:1509.04259](#)] [[INSPIRE](#)].
- [65] M. Creutz, *Monte Carlo study of quantized SU(2) gauge theory*, *Phys. Rev. D* **21** (1980) 2308 [[INSPIRE](#)].
- [66] A.D. Kennedy and B.J. Pendleton, *Improved heat bath method for Monte Carlo calculations in lattice gauge theories*, *Phys. Lett. B* **156** (1985) 393 [[INSPIRE](#)].
- [67] M. Creutz, *Overrelaxation and Monte Carlo simulation*, *Phys. Rev. D* **36** (1987) 515 [[INSPIRE](#)].
- [68] L. Del Debbio, H. Panagopoulos, P. Rossi and E. Vicari, *Spectrum of confining strings in SU(N) gauge theories*, *JHEP* **01** (2002) 009 [[hep-th/0111090](#)] [[INSPIRE](#)].
- [69] N. Cabibbo and E. Marinari, *A new method for updating SU(N) matrices in computer simulations of gauge theories*, *Phys. Lett. B* **119** (1982) 387 [[INSPIRE](#)].
- [70] B. Lucini, M. Teper and U. Wenger, *Properties of the deconfining phase transition in SU(N) gauge theories*, *JHEP* **02** (2005) 033 [[hep-lat/0502003](#)] [[INSPIRE](#)].
- [71] B.A. Berg, *Markov chain Monte Carlo simulations and their statistical analysis*, [World Scientific](#), Singapore (2004), pg. 196.
- [72] A. Gonzalez-Arroyo and M. Okawa, *The string tension for large N gauge theory from smeared Wilson loops*, *PoS(LATTICE2012)221* (2012) [[arXiv:1212.3835](#)] [[INSPIRE](#)].
- [73] QCDSF and UKQCD collaborations, *Determination of Lambda in quenched and full QCD: an update*, *Nucl. Phys. B Proc. Suppl.* **140** (2005) 228 [[hep-lat/0409166](#)] [[INSPIRE](#)].
- [74] T. Vonk, F.-K. Guo and U.-G. Meißner, *Aspects of the QCD  $\theta$ -vacuum*, *JHEP* **06** (2019) 106 [*Erratum ibid.* **10** (2019) 028] [[arXiv:1905.06141](#)] [[INSPIRE](#)].
- [75] R. Kitano, N. Yamada and M. Yamazaki, *Is  $N = 2$  large?*, *JHEP* **02** (2021) 073 [[arXiv:2010.08810](#)] [[INSPIRE](#)].