Extracting Physics from Topologically Frozen Markov Chains

Urs Gerber[*], **Irais Bautista, Wolfgang Bietenholz, Héctor Mejía-Díaz**

Instituto de Ciencias Nucleares
Universidad Nacional Autónoma de México
A.P. 70-543, C.P. 04510 Distrito Federal, Mexico
E-mail: gerber@correo.nucleares.unam.mx,
irais.bautista@correo.nucleares.unam.mx,
wolbi@nucleares.unam.mx, he_mejia@yahoo.com.mx

Christoph P. Hofmann

Facultad de Ciencias, Universidad de Colima
Bernal Díaz del Castillo 340, Colima C.P. 28045, Mexico
E-mail: christoph.peter.hofmann@gmail.com

In Monte Carlo simulations with a local update algorithm, the auto-correlation with respect to the topological charge tends to become very long. In the extreme case one can only perform reliable measurements within fixed sectors. We investigate approaches to extract physical information from such topologically frozen simulations. Recent results in a set of σ-models and gauge theories are encouraging. In a suitable regime, the correct value of some observable can be evaluated to a good accuracy. In addition there are ways to estimate the value of the topological susceptibility.

The 32nd International Symposium on Lattice Field Theory,
23-28 June, 2014
Columbia University New York, NY

---

[*]Speaker.

1. Introduction

In many relevant models, the configurations are divided into topological sectors (for periodic boundary conditions). This includes the $O(N)$ models in $d = (N-1)$, all the 2d $CP(N-1)$ models, 2d and 4d Abelian gauge theory, and 4d Yang-Mills theories. The topology persists if we include fermions, hence this class of models also includes the Schwinger model, QED and QCD.

In the continuum formulation, a continuous deformation of a configuration (at finite Euclidean action) cannot change the topological charge $Q \in \mathbb{Z}$. On the lattice there are no topological sectors in this strict sense, but at fine lattice spacing the configurations of the above models occur in distinct sectors with local minima, separated by boundary zones of higher action. Thus it is possible — and often useful — to introduce topological sectors also in lattice field theory, although the definition of Q is somewhat ambiguous. For the $O(N)$ models that we are going to consider, the geometric definition [1] has the virtue that it naturally provides integer values of Q.

Most simulations in lattice field theory are performed with local update algorithms, such as the Metropolis algorithm for spin models, the heat-bath algorithm for pure gauge theories, and the Hybrid Monte Carlo algorithm for QCD with dynamical quarks. If there are well-separated topological sectors, such simulations may face a severe problem: a Markov chain hardly ever changes Q. Thus the simulation tends to get stuck in one topological sector for an extremely long computation time. Such a tremendous topological auto-correlation time was observed *e.g.* by the JLQCD Collaboration in their QCD simulations with dynamical overlap quarks [2]. For QCD simulations with non-chiral quarks (*e.g.* given by Wilson fermions) the problem has been less severe so far, *i.e.* for lattice spacings $a \gtrsim 0.05$ fm that have typically been used. However, in the future even finer lattices will be employed, and then this problem will become manifest.

So how can we measure the expectation value of some observable, $\langle \Omega \rangle$, or the topological susceptibility $\chi_t = \frac{1}{V} \left(\langle Q^2 \rangle - \langle Q \rangle^2 \right)$, if only topologically frozen simulations can be performed? [1]

Lüscher suggested open boundary conditions, so Q can change continuously [3]. This overcomes the problem, but giving up integer Q has disadvantages, like losing the link to aspects of field theory in the continuum, *e.g.* regarding the ε-regime of QCD.

Here we investigate approaches where periodic boundaries, and therefore $Q \in \mathbb{Z}$, are preserved. In the framework of non-linear σ-models, we test methods to extract physical results from Markov chains, which are permanently trapped in a single topological sector, hence numerical measurements are available only at fixed Q. We start with a procedure to determine χ_t from the correlation of the topological charge density, which was introduced by Aoki, Fukaya, Hashimoto and Onogi [4]. Then we probe a way to assemble an expectation value $\langle \Omega \rangle$ from topologically restricted results $\langle \Omega \rangle_{|Q|}$. That approach is based on the Brower-Chandrasekharan-Negele-Wiese (BCNW) formula [5], which also yields an estimate for χ_t.

2. Correlation of the topological charge density

Ref. [4] derived an approximate formula for the correlation of the topological charge density q, at topological charge $\pm Q$ and large separation $|x|$ (we now use lattice units),

$$\langle q_0 \, q_x \rangle_{|Q|, \, |x| \gg 1} \approx -\frac{\chi_t}{V} + \frac{Q^2}{V^2} \, . \tag{2.1}$$

---

[1]V is the volume, and we will deal with parity symmetric models, where $\langle Q \rangle = 0$.

The derivation assumes $\langle Q^2 \rangle$ to be large, and $|Q|/\langle Q^2 \rangle$ to be small.[2] Therefore we will limit our considerations to the sectors with $|Q| \leq 2$. We are going to consider the 1d $O(2)$ model and the 2d $O(3)$ model, and the explicit condition for $\langle Q^2 \rangle = \chi_t V$ will be tested.

In our simulations we use the Wolff cluster algorithm [7], which performs non-local cluster updates. Hence we can also measure χ_t directly, which is useful for testing this method in view of other models (in particular gauge theories), where no efficient cluster algorithm is available. Preliminary results were anticipated in Ref. [8], and Ref. [9] presented before a related study (with different densities) in 2-flavour QCD.

The 1d $O(2)$ model, or quantum rotor, describes a free quantum mechanical scalar particle on the circle S^1. We use periodic boundary conditions in Euclidean time over the size L. The continuum formulation deals with an angle $\varphi(x)$, where $\varphi(0) = \varphi(L)$. The lattice variables are the angles φ_x, $x = 1, \ldots L$, with $\varphi_1 = \varphi_{L+1}$. We define the nearest site difference as

$$\Delta \varphi_x = (\varphi_{x+1} - \varphi_x) \bmod 2\pi \in (-\pi, \pi] \,, \tag{2.2}$$

i.e. the modulus function acts such that it minimises the absolute value. This yields the (geometrically defined) topological charge density q_x and the topological charge Q,

$$q_x = \frac{1}{2\pi} \Delta \varphi_x \,, \quad Q = \sum_{x=1}^{L} q_x \in \mathbb{Z} \,. \tag{2.3}$$

We now give the continuum action and the three lattice actions — standard action, Manton action [10] and constraint action [11] — that we studied,

$$S_{\text{continuum}}[\varphi] = \frac{\beta}{2} \int_0^L dx \, \dot{\varphi}(x)^2 \,, \quad S_{\text{standard}}[\varphi] = \beta \sum_{x=1}^{L} (1 - \cos \Delta \varphi_x) \,,$$

$$S_{\text{Manton}}[\varphi] = \frac{\beta}{2} \sum_{x=1}^{L} \Delta \varphi_x^2 \,, \quad S_{\text{constraint}}[\varphi] = \begin{cases} 0 & \Delta \varphi_x < \delta \quad \forall x \\ +\infty & \text{otherwise} \end{cases} . \tag{2.4}$$

The parameter β corresponds here to the moment of inertia, and δ is the constraint angle. The continuum limit is attained at $\beta \to \infty$ and $\delta \to 0$, respectively. In the limit $L \to \infty$, χ_t and the correlation length ξ are known analytically for all four actions in eqs. (2.4) [6, 11].

Figures 1 and 2 show results for the standard action, the Manton action and the constraint action at different sizes L and parameters β and δ. They are all in excellent agreement with the prediction based on eq. (2.1) (horizontal lines), even down to $\langle Q^2 \rangle < 1$.

Next we address the 2d $O(3)$ model, or Heisenberg model, on square lattices of size $L \times L$, with classical spins $\vec{e}_x \in S^2$. Here we simulated the standard action and the constraint action, which are analogous to the formulations (2.4),

$$S_{\text{standard}}[\vec{e}] = \beta \sum_{x,\mu} (1 - \vec{e}_x \cdot \vec{e}_{x+\hat{\mu}}) \,, \quad S_{\text{constraint}}[\vec{e}] = \begin{cases} 0 & \vec{e}_x \cdot \vec{e}_{x+\hat{\mu}} > \cos \delta \quad \forall x, \mu = 1, 2 \\ +\infty & \text{otherwise} \end{cases} . \tag{2.5}$$

Also here we use the geometric definition of the topological charge, which is written down explicitly in Ref. [11]. Figure 3 shows results for the topological charge density correlation. Again the comparison to the prediction (2.1) works in all cases (in this context we don't have to worry about

---

[2]Actually this formula also involves a *kurtosis* term (which vanishes for Gaussian Q-distributions). However, its contribution is negligible in all examples that we considered, so here we skip that term.
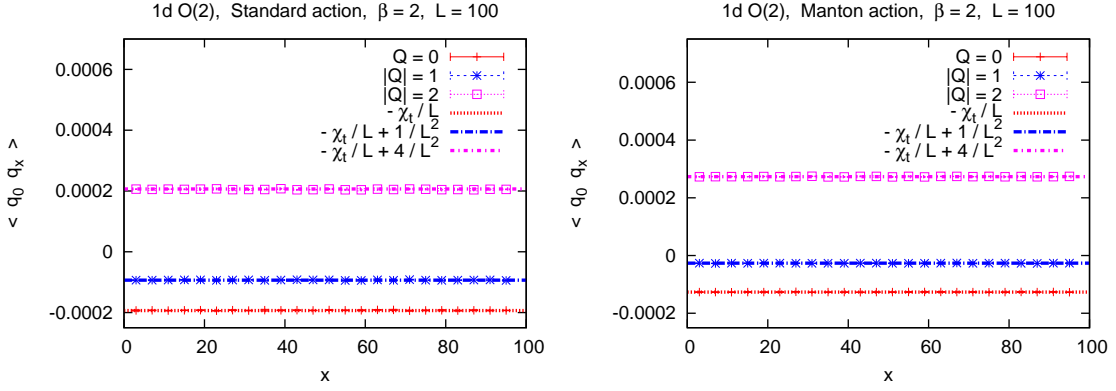
Figure 1: The topological charge density correlation over a separation of x lattice spacings for the standard action (on the left) and for the Manton action (on the right), both at $L = 100$ and $\beta = 2$. This implies $\xi = 2.779$, $\langle Q^2 \rangle = 1.936$ for S_{standard}, and $\xi = 4.000$, $\langle Q^2 \rangle = 1.266$ for S_{Manton}. For comparison, we show the prediction based on eq. (2.1), where we insert the measured values of χ_t.
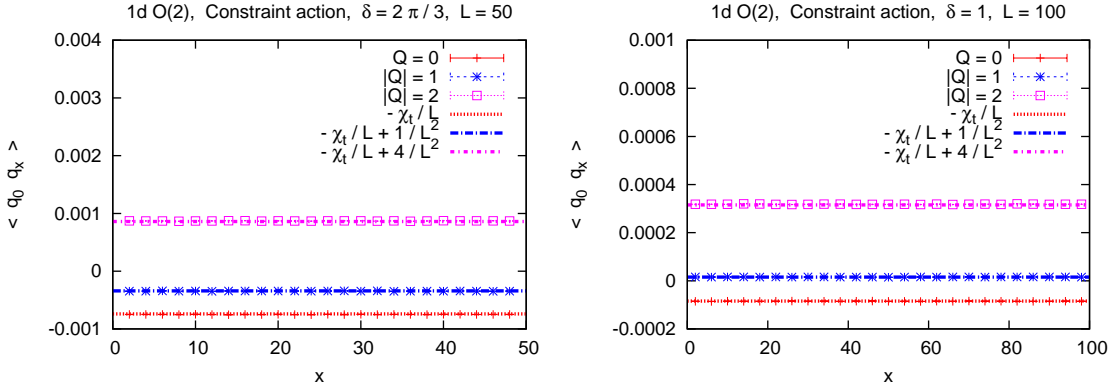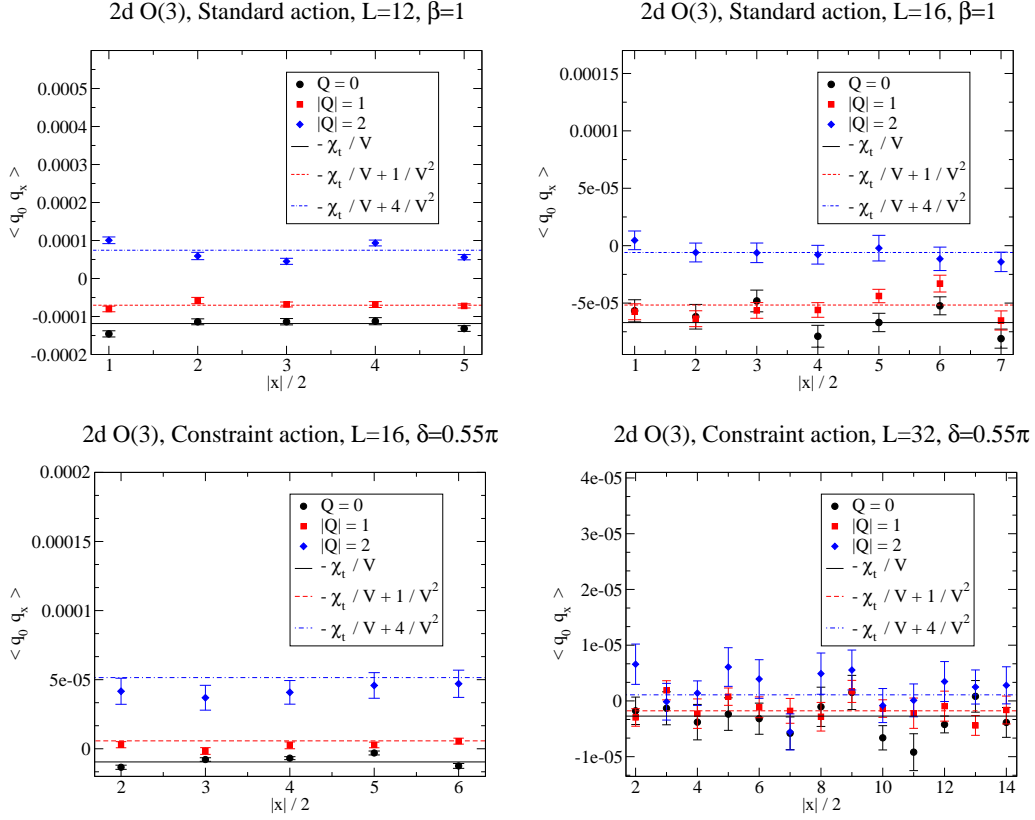

Figure 2: The topological charge density correlation over a separation of x lattice spacings for the constraint action at $\delta = 2\pi/3$, $L = 50$ (on the left, with $\xi = 1.132$, $\langle Q^2 \rangle = 1.852$), and $\delta = 1$, $L = 100$ (on the right, with $\xi = 5.793$, $\langle Q^2 \rangle = 0.844$). Again we compare with the prediction (2.1), using the measured χ_t.

the fact that $\chi_t \cdot \xi^2$ diverges logarithmically in the continuum limit). However, we also observe that for increasing volume it becomes soon difficult to resolve a clear signal from the statistical noise (as required for the determination of χ_t), even with the huge statistics provided by the cluster algorithm. The quantitative results will be given in Ref. [13].

Thus we confirm that formula (2.1) is a valid approximation over a broad set of parameters. Nevertheless, in view of 4d quantum field theory its application is not promising, since for large volume it becomes very statistics demanding. That limitation is in agreement with the conclusion of an earlier study in the 2-flavour Schwinger model with dynamical overlap hypercube fermions (and with the plaquette gauge action) [12]: at $\beta = 5$, $V = 16 \times 16$ and fermion masses $m = 0.01 \dots 0.06$ a statistics of $O(1000)$ configurations in one topological sector was insufficient to determine χ_t in this way; to achieve this to about 2 digits would take at least $O(10^5)$ configurations.

3. Applications of the BCNW formula

We now turn to the more ambitious goal of evaluating an observable $\langle \Omega \rangle$, when only some values $\langle \Omega \rangle_{|Q|}$ — at various $|Q|$ and volumes — are available. To this end, we use an approximate

Figure 3: The topological charge density correlation in the 2d $O(3)$ model on $L \times L$ lattices for the standard action at $\beta = 1$ (above, $\xi \simeq 1.3$; $\langle Q^2 \rangle = 2.46$ at $L = 12$ and $\langle Q^2 \rangle = 4.39$ at $L = 16$) and for the constraint action at $\delta = 0.55\pi$ (below, $\xi \simeq 3.5$; $\langle Q^2 \rangle = 0.63$ at $L = 16$ and $\langle Q^2 \rangle = 2.86$ at $L = 32$). Due to the definition of the topological charge density, we proceed in separation steps of 2 lattice spacings. In all cases, we show for comparison the prediction based on eq. (2.1), with the measured values of χ_t. It works well, even at $\langle Q^2 \rangle = 0.63$ the result is reasonable, although slight deviations from the prediction show up. But the last plot illustrates that for increasing volume the signal get lost in the statistical noise.

formula, which was derived in Ref. [5],

$$\langle \Omega \rangle_{|Q|} \approx \langle \Omega \rangle + \frac{c}{V \chi_t} \left(1 - \frac{Q^2}{V \chi_t} \right) . \tag{3.1}$$

Our input are measured values for the left-hand-side in various $|Q|$ and V, and a fit determines $\langle \Omega \rangle$, χ_t and c, where the former two are of interest. We refer to a regime of moderate V, where these three quantities practically take their infinite-volume values, but the $\langle \Omega \rangle_{|Q|}$ are still well distinct.

This is the beginning of an expansion in $1/\langle Q^2 \rangle$, hence $\langle Q^2 \rangle$ should be large, but what that means has to explored numerically. Moreover the assumption of a small value of $|Q|/\langle Q^2 \rangle$ is involved again (see also the re-derivation in Ref. [12]), hence we only use sectors with $|Q| \leq 2$. An extended expansion has been mentioned in Ref. [5] (first work), and explored in great detail in Refs. [14–16]. It involves further free parameters, and the crucial question if this improves the results for $\langle \Omega \rangle$ and χ_t was addressed in Refs. [14–16], and will be discussed further in Ref. [13].

As our observables we consider the action density $s = \langle S \rangle / V$ and the magnetic susceptibility
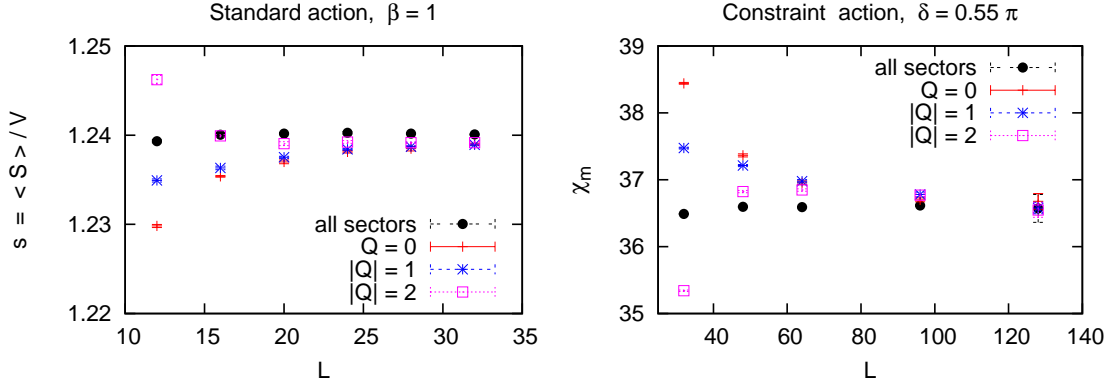
Figure 4: The action density $s = \langle S \rangle / V$ for the 2d $O(3)$ model on $L \times L$ lattice (standard action, $\beta = 1$, $\xi \simeq 1.3$), and the magnetic susceptibility with the constraint action at $\delta = 0.55\pi$ ($\xi \simeq 3.6$). We display the values measured in all sectors, and restricted to $|Q| = 0$, 1 or 2.

$\chi_{\mathrm{m}} = \langle \vec{M}^2 \rangle / V$ (where $\vec{M} = \sum_x \vec{e}_x$ is the magnetisation, and $\langle \vec{M} \rangle = \vec{0}$). Results for the 2d $O(3)$ model in $V = L \times L$ are shown in the plots of Figure 4, which reveal the aforementioned regimes of "moderate V". The fitting results involving the sectors $|Q| = 0$, 1, 2, and various ranges in L in those regimes, are given in Table 1. In particular we see an impressive precision of the values for χ_{m}, and also the fitting results for s and χ_{t} are quite good.

Standard action				directly measured
fitting range for L	$16 - 24$	$16 - 28$	$16 - 32$	in all sectors at $L = 32$
s	1.24038(12)	1.24027(8)	1.24015(5)	1.24008(5)
χ_{t}	0.0173(6)	0.0169(5)	0.0164(5)	0.01721(4)
Constraint action				directly measured
fitting range for L	$48 - 64$	$48 - 96$	$48 - 128$	in all sectors at $L = 128$
χ_{m}	36.56(4)	36.58(3)	36.57(2)	36.57(2)
χ_{t}	0.00262(17)	0.00256(16)	0.00259(14)	0.002790(5)

Table 1: Above: the action density $s = \langle S \rangle / V$ extracted from fits to the BCNW formula (3.1), at $|Q| \leq 2$ and various ranges of the L. For $L \geq 16$ the directly measured s stabilises. It is close to the fitting results. Below: the susceptibilities χ_{m} and χ_{t}, extracted from fits in various ranges of the L. For $L \geq 48$ the directly measured χ_{m} stabilises, and the results in distinct sectors converge quite well around $L = 96 \ldots 128$. The fits in moderate volumes lead to extremely precise values for χ_{m}. For both observables, also the fitting results for χ_{t} are correct within less than 2σ.

4. Conclusions

In simulations with local update algorithms and fine lattices, the Monte Carlo history tends to be confined to a single topological sector for an extremely long (simulation) time. This rises questions about the ergodicity (even within one sector). Here we do not address this conceptual issue; we trust the topologically restricted measurements of $\langle \Omega \rangle_{|Q|}$, and try to interpret them physically.

In very large volumes V, the restricted expectation values all coincide with the physical result, $\langle \Omega \rangle_{|Q|} \equiv \langle \Omega \rangle$, cf. eq. (3.1), but in practical simulations such large volumes are often inaccessible. For smaller V, where $\langle \Omega \rangle$ is well converged to its large-V limit, but the $\langle \Omega \rangle_{|Q|}$ are still significantly distinct, the BCNW formula (3.1) often allows us to determine $\langle \Omega \rangle$ to a good accuracy, and it also

provides useful results for χ_t. It is favourable to employ only the sectors with $|Q| \leq 2$, and (roughly speaking) the method is successful if $\langle Q^2 \rangle \gtrsim 1.5$.

If we relax that requirement to $\langle Q^2 \rangle \gtrsim 2/3$, we can still measure χ_t from the topological charge density correlation $\langle q_0 \, q_x \rangle_{|Q|}$. The (theoretical) condition of a large separation $|x|$ turns out to be harmless in practice, but for increasing V the wanted signal decreases very rapidly. Therefore that method is hardly promising for 4d models, where — in reasonable volumes — the signal would most likely be overshadowed by statistical noise.

On the other hand, the BCNW formula *is* promising for applications in QCD, where typical simulations take place at $\langle Q^2 \rangle = O(10)$. This observation is supported by studies in the Schwinger model [12, 15, 17], the quantum rotor with a potential [14] and in 4d $SU(2)$ gauge theory [16], which will be reported in detail in Ref. [13].

References

[1] B. Berg and M. Lüscher, *Nucl. Phys.* **B 190** (1981) 412.

[2] H. Fukaya *et al.*, *Phys. Rev. Lett.* **98** (2007) 172001; *Phys. Rev.* **D 76** (2007) 054503.

[3] M. Lüscher, *JHEP* **1008** (2010) 071. M. Lüscher and S. Schaefer, *JHEP* **1107** (2011) 036.

[4] S. Aoki, H. Fukaya, S. Hashimoto and T. Onogi, *Phys. Rev.* **D 76** (2007) 054508.

[5] R. Brower, S. Chandrasekharan, J.W. Negele and U.-J. Wiese, *Nucl. Phys. B Proc. Suppl.* **106** (2002) 581; *Phys. Lett.* **B 560** (2003) 64.

[6] W. Bietenholz, R. Brower, S. Chandrasekharan and U.-J. Wiese, *Phys. Lett.* **B 407** (1997) 283.

[7] U. Wolff, *Phys. Rev. Lett.* **62** (1989) 361.

[8] I. Bautista, W. Bietenholz, U. Gerber, C.P. Hofmann, H. Mejía-Díaz and L. Prado, arXiv:1402.2668 [hep-lat].

[9] S. Aoki *et al.*, *Phys. Lett.* **B 665** (2008) 294.

[10] N. Manton, *Phys. Lett.* **B 96** (1980) 328.

[11] W. Bietenholz, U. Gerber, M. Pepe and U.-J. Wiese, *JHEP* **1012** (2010) 020. W. Bietenholz, M. Bögli, U. Gerber, F. Niedermayer, M. Pepe, F.G. Rejón-Barrera and U.-J. Wiese, arXiv:1309.6278 [hep-lat].

[12] W. Bietenholz, I. Hip, S. Shcheredin and J. Volkholz, *Eur. Phys. J.* **C 72** (2012) 1938.

[13] I. Bautista, W. Bietenholz, C. Czaban, A. Dromard, U. Gerber, C.P. Hofmann, H. Mejía-Díaz and M. Wagner, in preparation.

[14] A. Dromard and M. Wagner, arXiv:1309.2483 [hep-lat].

[15] C. Czaban and M. Wagner, arXiv:1310.5258 [hep-lat].

[16] A. Dromard and M. Wagner, arXiv:1404.0247 [hep-lat].
C. Czaban, A. Dromard and M. Wagner, *Acta Phys. Polon. Supp.* **7** (2014) 3, 551.

[17] W. Bietenholz and I. Hip, *PoS LATTICE2008 (2008) 079*; *J. Phys. Conf. Ser.* **378** (2012) 012041.