Z. Phys. C – Particles and Fields 27, 553–558 (1985)

Zeitschrift für Physik C Particles and Fields

Proposal of a New Upgrading Procedure for Monte Carlo Experiments

Hildegard Meyer-Ortmanns
II. Institut für Theoretische Physik, Universität, D-2000 Hamburg, Federal Republic of Germany

Received 14 September 1984

Abstract. We propose a new algorithm for Monte Carlo calculations in spin models or gauge theories. It is supposed to speed up the convergence to equilibrium in situations where long range correlations are important.

1. Introduction

A *n*ew Monte Carlo *u*pgrading *p*rocedure (NUP) is proposed for spin models as well as for pure gauge theories. It is worked out for the two-dimensional $O(3)$ sigma model, for the Ising spin theory in two and in three dimensions and for the $Z(2)$ gauge theory in three dimensions. Common to the last three models are their second order phase transitions, which occur at certain values of the inverse temperature (coupling) β. The critical temperature of the $O(3)$ sigma model is believed to be at $\beta=0$. Well established procedures used in Monte Carlo experiments are the heat and the Metropolis algorithms [1]. They satisfy certain conditions that guarantee the convergence of the distribution of configurations in the statistical ensemble towards the Boltzmann distribution, after a sufficiently large number of configurations has been generated. In both algorithms only a single spin- or gauge field variable is upgraded at a time. Doing this successively for all sites (links) of the lattice in a random or systematic way completes one Monte Carlo sweep. In the heat bath procedure, the choice of the new variable is independent of its old value, but weighted according to the temperature of the local heat reservoir the new variable is brought into contact with before it is selected. The Metropolis algorithm substitutes the new variable for the old one independently of the environment, but dependent on the local change in the action.

We believe that these properties of locality are responsible for the critical slowing down near second order phase transitions, where the correlation length ξ tends to infinity. In a model like the $O(3)$ sigma model they might prevent a transition from one topological sector to the next within a reasonable computing time.

NUP is proposed to speed up the convergence to the Boltzmann distribution using "long range upgrading" procedures in order to simulate long range correlations. 'Long range upgrading' means the simultaneous upgrading of spins or gauge fields of a whole sublattice Λ_s of the lattice Λ.

As an illustration of the idea consider the two-dimensional $Z(2)$ spin theory with action

$$S=\beta \sum_{\langle ij\rangle} (1-s(x^{(i)})\, s(x^{(j)})), \tag{1}$$

$s(x^{(i)})\in Z(2)=\{\pm 1\}$ are spins attached to sites $x^{(i)}$. The sum runs over all links $\langle ij\rangle$ of the lattice.

The situation near the critical point at $\beta_c=0.4407$ is well described by the droplet picture, which essentially says that in each island of up-spins there are smaller islands of down-spins and so on. The size of the islands ranges from one lattice spacing a to the correlation length $\xi \gg a$. In other words, Peierls contours [2] of each length scale are abundant. (A Peierls contour is a set of links $\langle ij\rangle$ with $s(x^{(i)})\cdot s(x^{(j)})=-1$. They are sufficient to characterize an arbitrary spin configuration uniquely up to global spin reflections.)

As a second example consider the three-dimensional $Z(2)$ gauge theory with action

$$S=\beta \sum_{p\in\Lambda} (1-\mathrm{Tr}\, U(\partial p)). \tag{2}$$

The sum runs over all plaquettes p of Λ. $U(\partial p)$ is the product of $Z(2)$ variables $U(b)$ associated to links b along the boundary ∂p of a plaquette p. This model undergoes a second order phase transition at β_c

$=0.7613$ [3]. The analogous quantities to Peierls contours are now $Z(2)$ vortices, i.e. sets of frustrated plaquettes p with $U(\partial p)=-1$. They characterize a $Z(2)$ gauge field configuration uniquely up to local gauge transformations. At the critical point vortices of each length scale are abundant. (The length of a vortex is defined as the number of plaquettes it consists of.)

NUP as described in the next section shall simulate such situations in a few steps.

2. The Algorithm NUP

We use the following notations.

x is a coordinate vector on the d-dimensional lattice Λ, $d=2,3$, depending on the model. The lattice volume is $L^d=N$ with N lattice sites. A block $B(x_c^{(i)}, d_n)$ is the union of all sites of a d-dimensional sublattice $\Lambda_s\subset\Lambda$, $x_c^{(i)}$ is the block center. (i) labels the different choices of block centers. d_n is the block size, n $(1\leqq n\leqq n_{\max})$ labels the different choices of block sizes. A block $B_g(x_c^{(i)}, d_n)$ is the union of all links $\langle jk\rangle$ pointing in positive directions $\mu=1,2,3$ from sites $x^{(j)}\in B(x_c^{(i)}, d_n)$. A block plaquette $B_p(x_c^{(i)}, d_n, x_\rho)$ is the union of d_n^2 plaquettes $p(x_\rho)$ lying in the plane $x_\rho=\text{const}(\rho\in\{1,2,3\})$ such that at least two links of ∂p belong to $B_g(x_c^{(i)}, d_n)$. To each plaquette $p(x_\rho)\in B_p(.,.,.)$ we associate a link $\langle x\, x+\kappa\rangle$ pointing in positive direction κ orthogonal to μ and ν from a site which is corner of the plaquette. This set of links is called $B_l(x_c^{(i)}, d_n, x_\rho)$. For the definitions of B, B_g, B_p, and B_l compare Fig. 1a–d.

$s(x)\in Z(2)$ $[O(3)]$ are spin variables attached to sites x. We denote by $U(b)\in Z(2)$ gauge field variables living on links b, by $\sigma(B)$ $Z(2)$ "flips" $\in\{\pm 1\}$ applied to spins $s(x)$ of a block $B(x_c^{(i)}, d_n)$; $\sigma(B_l)$ stands for flips applied to gauge fields $U(b)\in Z(2)$ with $b\in B_l$. $R(B)\in O(3)$ represent rotations of spins $s(x)$ with $x\in B(x_c^{(i)}, d_n)$. The flips or rotations will be chosen according to a certain probability distribution which will be specified below.

Then NUP consists of the following steps.

1. Choose a block center $x_c^{(i)}$ from all sites, randomly, with a uniform distribution.

2. In the same way choose a block size d_n between $1\leqq d_n\leqq L$.

3. Now specify the rotation or flip, respectively. We will discuss two alternatives which reduce to the heat bath- and the Metropolis algorithms in the special case of $d_n=1$. Step 3 depends on the model. First we concentrate on spin models.

3.1.1. The $O(3)$ sigma model with action

$$S=\beta\sum_{\langle ij\rangle} s(x^{(i)})\, s(x^{(j)}), \tag{3}$$

where $s(x^{(i)})$ are three component unit vectors. Now choose $R(B)\in O(3)$, randomly, with a probability measure dP, given by

$$dP(R(B)) = \frac{1}{K}\exp\{-\beta\sum_{\langle lm\rangle\in\Xi} s(x^{(l)})\, R(B)\, s(x^{(m)})\}\, dR(B). \tag{4}$$

$dR(B)$ is the normalized Haar measure on $O(3)$. Ξ is the set of links $\langle lm\rangle$ where one and only one site of $\partial\langle lm\rangle$ belongs to B, cp. Fig. 2.

3.1.2. For the $Z(2)$ spin theory in two and three dimensions with action given by (1) replace $R\in O(3)$ by $\sigma\in Z(2)$. Instead of $dR(B)$ insert the discrete Haar measure on $Z(2)$.

3.1.3. In the case of a $Z(2)$ gauge theory one could think of flipping all variables $U(b)\in Z(2)$ with $b\in B_g(x_c^{(i)}, d_n)$. The probability density would be given by $(1/K')\exp\{-\beta\sum_{p\in\mathscr{P}_g}\sigma(B_g)\, U(\partial p)\}$, where $\mathscr{P}_g$ is the set of plaquettes p with an odd number of links

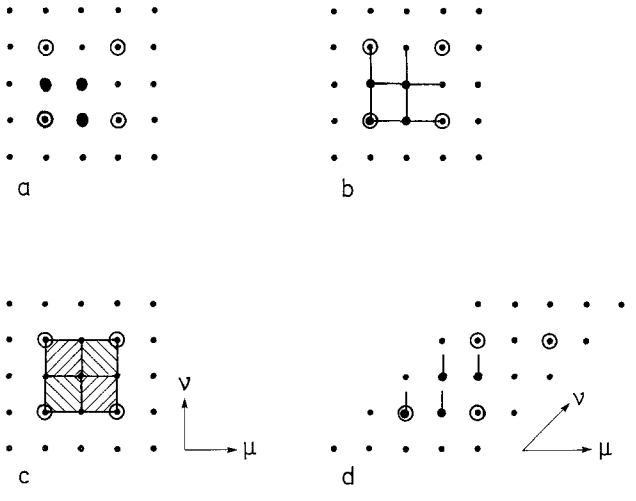


Fig. 1a–d. Sets B (**a** with fat dots), B_g (**b**), B_p (**c**), and B_l (**d**) as defined in the text. The circled dots are block centers

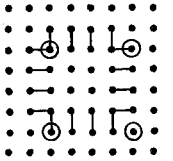


Fig. 2. Set Ξ as defined in the text for $d_n=4$

$b \in B_g$ in ∂p. However, this proposal does not include the usual heat bath algorithm for $d_n = 1$. Therefore it is suggested to use the following alternative [4].

3.1.3.1. For a fixed block $B(x_c^{(i)}, d_n)$ choose a block plaquette $B_p(x_c^{(i)}, d_n, x_\rho)$, i.e. choose x_ρ, randomly, uniformly distributed from the set $\{x_{\rho_1}, \ldots, x_{\rho_{d_n}}\}$, $\rho \in \{1,2,3\}$. For $d_n = 2$ and $d = 3$, $\{\ldots\}$ contains six elements.

3.1.3.2. Flip all variables $U(b)$ with $b \in B_l(x_c^{(i)}, d_n, x_\rho)$ according to the measure

$$dP(\sigma(B_l)) = \frac{1}{k''} \exp\Big[-\beta \sum_{p \in \mathscr{P}_V} \sigma(B_l)\, U(\partial p)\Big]\, d\sigma(B_l) \qquad (5)$$

$d\sigma(B_l)$ is the discrete Haar measure on $Z(2)$. $\mathscr{P}_V$ is the set of plaquettes which contain an odd number of links of B_l. It is shown in Fig. 3. If all $U(\partial p) = +1$ and $\sigma(B_l) = -1$, $\mathscr{P}_V$ coincides with a planar $Z(2)$ vortex V, which was generated by the flip along the "boundary" of the block plaquette B_p. Arbitrary vortices which wind through three dimensions can be composed of planar vortices.

The probability densities of (4) and (5) can be generated numerically in the usual way. If the notations of (4) and (5) are simplified for a moment to

$$dx = g(y)\, dy, \qquad (6)$$

first determine the maximum $g_{\max}(y)$, $y \in [y_{\min}, y_{\max}]$. Next select randomly $y_r \in [y_{\min}, y_{\max}]$ from a uniform distribution between $y_{\min}$ and $y_{\max}$. Compute $g(y_r)$. Now select $g_r \in [0, g_{\max}(y)]$, again randomly, from a uniform distribution between 0 and $g_{\max}(y)$. If $g(y_r) > g_r$ accept y_r, otherwise reject it. Finally the set of numbers y_r will be distributed with a probability density $g(y)$. In this way we can generate sets $\{R^r\}$, $(\{\sigma^r\})$ of rotations R^r, (flips σ^r), $1 \leqq r \leqq r_{\max}$, in several cycles of NUP, which are distributed according to (4) and (5), respectively. One cycle of NUP consists of steps 1–7, cp. below.

3.2.1,2,3. Alternatively to steps 3.1.1,2,3, which are a kind of block heat bath algorithm, one could define a "block Metropolis algorithm" in the following sense. Select randomly, e.g. according to a uniform distribution on the group manifolds, rotations R, flips σ. Next calculate the change in the action ΔS when all spin- (gauge field) variables of a block B (B_l) are flipped, respectively. Extract a (pseudo) random number r, generated with uniform probability distribution between 0 and 1. If $e^{-\Delta S} > r$ accept the selected rotations (flips), otherwise reject them.

4. Multiply all $s(x)$, $x \in B(x_c^{(i)}, d_n)$, with $R(B)$, $[\sigma(B)]$, while $s(x)$ for $x \notin B$ remain unchanged. (In the $Z(2)$ gauge theory multiply all $U(b)$, $b \in B_l(x_c^{(i)}, d_n, x_\rho)$, with $\sigma(B_l)$.)

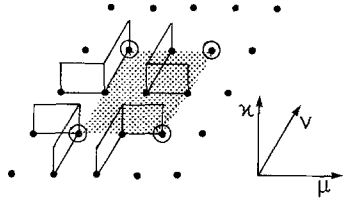
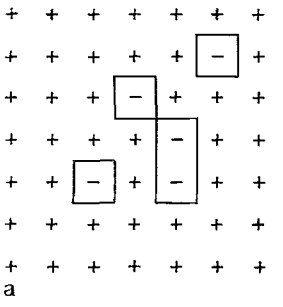
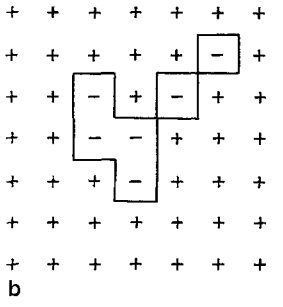


Fig. 3. Set $\mathscr{P}_V$. B_p is the shaded region





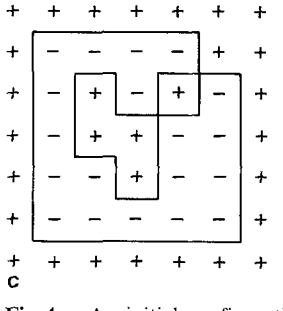


Fig. 4. a An initial configuration in a two-dimensional Ising spin system with spins up (+) and spins down (−). The full lines represent Peierls contours. **b** Effect of a flip $\sigma = -1$ of all spins of a 3×3 block, $d_n = 3$. The situation described in terms of Peierls contours is still similar to that of **a. c** A long Peierls contour is generated in addition, when $d_n = 5$, i.e. all spins of a 5×5 block are flipped

5. Write the just generated configuration on tape.

5a (only for the $Z(2)$ gauge theory). Repeat steps 3.1.3.1–5 for different values of x_ρ.

6. Repeat steps 2–5(a) for different block sizes d_n, until all (or most of) the values between 1 and L appeared in the sequence. It might be useful to weight the local fluctuations (corresponding to $d_n=1$) stronger than the "super long range" fluctuations with $d_n \sim L$. Then steps 2–5(a) should be repeated for $d_n=1$ several times in addition.

7. Repeat steps 1–6 for different choices of block centers, until most of the sites of Λ have been the center at least for one time. Steps 1–7 complete one cycle of NUP.

8. Repeat the cycle several times in order to generate distributions according to (4) and (5).

Figure 4 for the two-dimensional $Z(2)$ spin model shows the creation of Peierls contours in a spin configuration (Fig. 4a), when all spins in a block of size $d_1=3$ (Fig. 4b) and of size $d_2=5$ (Fig. 4c) are flipped.

Our expectation is the following.

For a certain range of the temperature β^{-1}, where long range correlations become important, because long range topological excitations are present or a second order phase transition ($\xi \to \infty$) is close, NUP speeds up the convergence to equilibrium. It should be less CPU-time consuming than the usual algorithms to produce an importance sampling.

3. Proof of Detailed Balance

If the sequence of distributions of configurations generated in the Markov process converges to a Boltzmann distribution, the Boltzmann factor which appears in the path integral is absorbed in the selection of configurations. Expectation values of observables can be calculated as arithmetic averages. To ensure the convergence, the transition probabilities from one configuration to the next have to satisfy certain conditions which we recall below. We label the configurations of a discrete set of configurations by numbers μ, $\nu \in \mathbb{N}$, the members in the sequence of the Markov chain by k. $W(\nu \to \mu)$ denotes the probability density for the transition from $\{\nu\}_k \to \{\mu\}_{k+1}$. The density is understood as a density on the product space of group manifolds. $W(\nu \to \mu)$ depends on the specific algorithm. $\{\nu\}$ and $\{\mu\}$ are configurations that finally enter the calculation of expectation values. (In NUP they were collected on tape.)

To an ensemble of k configurations $\{\nu_1\}$, $\{\nu_2\}, \ldots, \{\nu_k\}$ in the sequence of configurations, generated so far in the Markov process, we associate a vector $\{\mathscr{V}_\mu(k)\}_{\mu \in \mathbb{N}}$, constructed in the following way.

$\mathscr{V}_\mu(k)$ = probability to find $\{\mu\}$ in the ensemble of k configurations

$$= \sum_{i=1}^{k} \delta_{\mu \nu_i}/k.$$

The dimension of $\mathscr{V}$ can be finite or infinite depending on the total number of configuration $\{\mu\}$, which are only assumed to be a countable set. The sequence of vectors $\{\mathscr{V}_\mu(1)\}$, $\{\mathscr{V}_\mu(2)\}, \ldots, \{\mathscr{V}_\mu(k)\}, \ldots$ converges to a unique limit given by the Boltzmann distribution, i.e.

$$\lim_{k \to \infty} \{\mathscr{V}_\mu(k)\} = \begin{pmatrix} e^{-\beta S(\{1\})} \\ e^{-\beta S(\{2\})} \\ \vdots \\ e^{-\beta S(\{\mu\})} \\ \vdots \end{pmatrix}, \tag{8}$$

if the conditions of normalization (i), ergodicity (ii), and balance (iii) are satisfied.

$$\text{(i)} \quad \sum_{\nu} W(\nu \to \mu) = 1. \tag{9}$$

(ii) If $e^{-\beta S(\{\nu\})}$, $e^{-\beta S(\{\mu\})} > 0$, there exists a number $k_0 \in \mathbb{N}$ such that $W^{k_0}(\nu \to \mu) > 0$ (corresponding to k_0 steps). (10)

$$\text{(iii)} \quad \sum_{\nu} W(\nu \to \mu) e^{-\beta S(\{\nu\})} = e^{-\beta S(\{\mu\})}. \tag{11}$$

In the next section we will show instead of (iii) the stronger condition of detailed balance (iiia)

$$\text{(iiia)} \quad W(\nu \to \mu) e^{-\beta S(\{\nu\})} = W(\mu \to \nu) e^{-\beta S(\{\mu\})}. \tag{12}$$

The proof of (12) will be outlined only in the notation of the $O(3)$ sigma model. Given a configuration $\{\nu\}$, we have to distinguish two cases for a transition (according to NUP) from $\{\nu\}_k$ to an arbitrary $\{\mu\}_{k+1}$. (μ may coincide with ν.)

1. The transition from $\{\nu\}$ to $\{\mu\}$ in one step is impossible, then (12) is fulfilled with vanishing left and right sides.

2. $\{\mu\}$ is a configuration that can be reached from $\{\nu\}$ within one step. In the case $\mu = \nu$, Eq. (12) is obviously true. If $\mu \neq \nu$ then at least one block B with center $x_c^{(i)}$ and size d_n exists such that all spins of B are rotated with R compared to those of $\{\nu\}$. (In the case of the $Z(2)$ spin model, the parameters $x_c^{(i)}$, d_n and σ are uniquely determined by $\{\mu\}$ for a

given $\{\nu\}$ for $\mu \neq \nu$; in the $O(3)$-case they are "almost always" unique (i.e. except for a set of measure zero) if $\mu \neq \nu$.) Then the probability $W(\nu \to \mu)$ to find just these parameters $x_c^{(i)}$, d_n and R during NUP is given by

$$W(\nu \to \mu) = \text{prob}(x_c^{(i)}) \cdot \text{prob}(d_n) \cdot \text{prob}(R) \tag{13}$$

with

$$\text{prob}(x_c^{(i)}) = 1/N, \tag{14}$$

$$\text{prob}(d_n) = 1/n_{\max} \tag{15}$$

if $n_{\max}$ is the maximal number of different block sizes, and

$$\text{prob}(R) = \frac{1}{k} \exp[-\beta \sum_{\langle lm \rangle \in \Xi} s(x^{(l)}) R s(x^{(m)})] \tag{16}$$

for NUP in the block heat bath version,

$$\text{prob}(R) = \begin{cases} \text{const} \cdot 1 & \text{if } \Delta S < 0 \\ \text{const} \cdot e^{-(S(\{\mu\}) - S(\{\nu\}))} & \text{if } \Delta S > 0 \end{cases} \tag{17}$$

with $\Delta S := S(\{\mu\}) - S(\{\nu\})$ as defined in steps 3.2.1, 2, 3 for NUP in the block Metropolis version.

Equations (14) and (15) hold, because $x_c^{(i)}$ and d_n were chosen randomly according to uniform distributions. They are symmetric in k and $k+1$ and drop out of both sides of (12). A systematic choice of $x_c^{(i)}$ and d_n could prevent detailed balance marking a certain direction in the process.

For $W(\mu \to \nu)$ we have instead of (16), (17)

$$\text{prob}(R^{-1}) = \frac{1}{k} \exp[-\beta \sum_{\langle lm \rangle \in \Xi} s(x^{(l)}) R^{-1} R s(x^{(m)})], \tag{18}$$

$$\text{prob}(R^{-1}) = \begin{cases} \text{const} \cdot 1 & \text{if } \Delta S < 0 \\ \text{const} \cdot e^{-(S(\{\nu\}) - S(\{\mu\}))} & \text{if } \Delta S > 0 \end{cases} \tag{19}$$

with $\Delta S := S(\{\nu\}) - S(\{\mu\})$, respectively.

Equation (12) follows now from (13)–(19) and the definition of S. This completes the proof of detailed balance.

The normalization condition (i) is fulfilled for a suitable choice of normalizations constants $k^{(\prime\prime)}$ in (4) and (5).

Condition (ii), i.e. ergodicity, is implied for NUP by the heat bath- and the Metropolis algorithms, since they are included in NUP for $d_n = 1$, respectively, for spin models as well as for the $Z(2)$ gauge theory.

A disadvantage of NUP may be the time consuming computation of factors $e^{-\beta\Sigma}$ or of ΔS, since the number of links (plaquettes) summed over in $\sum$ or in ΔS is proportional to the "boundary" of the block.

4. Remark on a Further Alternative of NUP

We would like to comment briefly on an alternative upgrading procedure, which is proposed in the same spirit as the previous NUP but differs in the probability distributions of R, σ, respectively. It consists in a heat bath upgrading of block spins in the following steps.

1. Choose a block size d_n, $1 \leqq d_n \leqq L$.

2. Define a block spin $s(x_B)$, e.g. according to

$$s(x_B) := \frac{\sum_{x \in B} s(x)}{\| \sum_{x \in B} s(x) \|}, \tag{20}$$

as is commonly used in spin models. x_B is a site on the block lattice, $s(x)$ are spins on the original lattice. The sum runs over all sites of a block B of size d_n.

3. Now apply the usual heat bath- or Metropolis algorithm to the block spins $s(x_B)$ with respect to the effective action S_{eff} which determines the interaction of block spins. Then $s(x_B) \to s'(x_B) =: R(B) s(x_B)$ which defines $R(B)$ for all blocks B of Λ (uniquely only under additional conditions).

4. After one sweep on the block lattice multiply all $s(x)$ on the original lattice with $R(B)$ for $x \in B$; write $\{s'(x)\}$ with $s'(x) := R(B) s(x)$ for $x \in B$ on tape.

5. Repeat steps 2–4 for several sweeps on the block lattice starting in each sweep from the last configuration $\{s'(x)\}$ which was calculated in 4.

6. Repeat steps 1–5 for different block sizes d_n including $d_n = 1$, which corresponds to the usual heat bath - or Metropolis algorithm.

Step 3 may involve a principal difficulty. Generally, the precise form of S_{eff} is unknown. Therefore conditions (i), (ii), (iii) of Sect. 3 cannot be proven as long as the systematic error, arising from a truncation of S_{eff}, is not under control.

5. Remark on the Use of Boundary Conditions

It is well known that certain topological configurations are excluded by the use of periodic boundary conditions. As an example consider the pure $SU(2)$ gauge theory with Wilson's action in a four-dimensional box. Define twisted boundary conditions by

$$\begin{aligned} U(b) &\to \sigma(b)\, U(b) \quad \text{for } b \in T \\ U(b) &\to \quad U(b) \quad \text{otherwise,} \end{aligned} \tag{21}$$

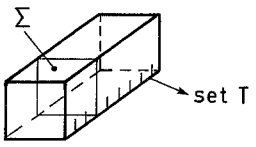


Fig. 5. Three-dimensional projection of a box Λ illustrating twisted boundary conditions

$U(b) \in SU(2)$, $\sigma(b) \in Z(2)$. T is the set of links in $\partial\Lambda$ which is coclosed in $\partial\Lambda$ and winds around Λ as shown in Fig. 5. $SU(2)$ vortices can be characterized by so called vortex souls. For the precise definition of vortex souls see [5]. Only an even number of vortex souls passing a cross section Σ (cp. Fig. 5) is compatible with periodic boundary conditions, while an odd number is forced into the box by twisted conditions. Measurements of long range quantities like correlation functions might be influenced by neglecting a whole class of $SU(2)$ configurations.

Therefore we propose (for whatever upgrading is used) to start from $n \geqq 1$ initial configurations with n topologically inequivalent boundary conditions. In the case of gauge groups $SU(2)$ and $Z(2)$ we would start from two configurations differing just by a twist.

Acknowledgement. I would like to thank Prof. G. Mack for useful discussions.

References

1. C.P. Yang: Proc. of Symposis in Applied Mathematics, Vol. XV, 351, Amer. Math. Soc., Providence RI (1963)
   M. Creutz: Phys. Rev. D 21, 2308 (1980); N. Metropolis et al.: J. Chem. Phys. **21**, 1087 (1953)
2. R. Peierls: Proc. Cam. Phil. Soc. **32**, 477 (1936)
3. R. Balian, J.M. Drouffe, C. Itzykson: Phys. Rev. **D 11**, 2098 (1975)
4. G. Mack: private communication
5. G. Mack: Properties of lattice gauge theory models at low temperature, DESY 80/03 (Jan. 1980). *In:* Recent developments in gauge theories, (G. 't Hooft et al., eds.) New York: Plenum Press 1980