Multicanonical hybrid Monte Carlo algorithm: Boosting simulations of compact QED

G. Arnold and K. Schilling
Neumann Institute for Computing (NIC), c/o Research Center Jülich and DESY, Hamburg, D-52425 Jülich, Germany

Th. Lippert
Department of Physics, University of Wuppertal, D-42097 Wuppertal, Germany
(Received 24 September 1998; published 10 February 1999)

We demonstrate that substantial progress can be achieved in the study of the phase structure of four-dimensional compact QED by a joint use of hybrid Monte Carlo and multicanonical algorithms through an efficient parallel implementation. This is borne out by the observation of considerable speedup of tunnelling between the metastable states, close to the phase transition, on the Wilson line. We estimate that the creation of adequate samples (with order 100 flip-flops) becomes a matter of half a year's run time at 2 Gflops sustained performance for lattices of size up to 24^4. [S0556-2821(99)04403-3]

PACS number(s): 11.15.Ha, 12.20.Ds

I. INTRODUCTION

It appears exceedingly intriguing to define variants of QED by studying vacuum states other than the usual perturbative vacuum. Lattice techniques have the potential power to deal with this situation, whenever they provide us with phase transition points of second and higher orders.

It is embarrassing that lattice simulations of compact QED still have not succeeded to clarify the order of the phase transition near $\beta=1$, the existence of which was established in the classical paper of Guth [1]. This is mainly due to the failure of standard updating algorithms, such as metropolis, heat bath, or metropolis with reflections [2–5], to move the system at a sufficient rate between the observed metastable states near its phase transition. The tunneling rates decrease exponentially in L^3 and exclude the use of lattices large enough to make contact with the thermodynamic limit by finite size scaling (FSS) techniques [6].

In this paper, we propose to make use of the multicanonical algorithm (MUCA) [7] within the hybrid Monte Carlo (HMC) updating scheme [8] in order to boost the tunneling rates. Since both algorithms are inherently of a global nature, their combination will facilitate the parallelization of the MUCA which could not be achieved otherwise.

In the early days of simulations on the hypertorus, the U(1) phase transition was claimed to be second order [9–12], however, with increasing lattice sizes, metastabilities and double-peak action distributions became manifest, strongly hinting at its first-order character [13–15]. This picture is in accordance with various renormalization group studies [16,17].

However, the latent heat is found to decrease with the lattice size and the critical exponent ν is neither 0.25 (first order) nor 0.5 (trivially second order) [18]. These facts allow for two possible propositions: (1) the double-peak structure is a finite size effect and might vanish in the thermodynamic limit, leading to the signature of a second-order phase transition; (2) the phase transition is weakly first order; i.e., the correlation length ξ is finite, but large in terms of the available lattice extent L. This would fake, on small lattices, the signature of a second-order transition, and a stabilized value of the latent heat would only become visible in the thermodynamic limit $L>\xi$.

In the search for a lattice formulation of QED with a second-order transition point, the action was generalized to include a piece in the adjoint representation with coupling γ [19], in the expectation that the phase transition would be driven towards second order, at sufficiently small negative values of γ. However, simulations on the hypertorus, up to $\gamma=-0.4$, revealed the reappearance of a double peak on large enough lattices [20,21]. Thus, the hypothesis of a second-order phase transition at some finite negative value of γ [22–24] is again doubted; furthermore, renormalization group investigations indicated that the second-order phase transition is located at $\gamma\rightarrow-\infty$ [17].

Reference [25] has speculated that the mechanism behind the lattice heuristics of metastabilities is driven by monopole loops that wrap around the hypertorus. According to this scenario, the inefficiency of local updating algorithms to create and annihilate such monopole constellations causes their slowing down, in agreement with the earlier proposition (1). Results were presented in support of this view by switching to spherical lattices with trivial homotopy group where such wrapping loops are no more topologically stabilized [26–29]. But on spherical lattices equivalent to $L=26$ at $\gamma=-0.2$, double-peak structures have recently been reported to reappear [20,21], corroborating earlier observations with periodic boundary conditions at $\gamma=0$: the suppression of monopole loop penetration through the lattice surface turned out to be incapable of preventing the incriminated double-peak signal from showing up on large lattices, say, $L=32$ [30–32].

It appears that a clarification of the situation of the Wilson line is mandatory for further progress in the understanding of compact lattice QED. This challenge requires the design of more powerful updating algorithms. A promising method is based on simulated tempering [33–35], enlarging the Lagrangian by a monopole term whose coupling is treated as an additional dynamical variable. Multiscale update schemes in principle can alleviate the critical slowing down (CSD) which is associated with the increase of the correlation length ξ (as measured in a nonmixed phase) near the critical

coupling β_c [36]. However, the exponential supercritical slowing down (SCSD) which is a consequence of the surface tension at first-order phase transitions cannot be overcome by such type of scale-adapted methods. In that instance, one expects the autocorrelation times to grow exponentially with the system size due to the occurrence of two 3-dimensional interfaces, leading to

$$\tau_{\mathrm{SCSD}} \propto \exp(2\sigma L^3). \tag{1}$$

Valleau and co-workers [37–40] have shown how to generate arbitrary ''nonphysical'' sampling distributions. Their method, termed ''umbrella sampling,'' has been introduced to span large regions of phase diagrams. The method is capable of improving the efficiency of stochastic sampling for situations when dynamically nearly disconnected parts of phase space occur by biasing the system to frequent the dynamically depleted, connecting regions of configuration space. They interpreted their method as sampling ''a whole range of temperatures'' [37].

In recent years the idea of ''umbrella sampling'' has been popularized and extensively applied under the name ''multicanonical algorithm'' (MUCA) by Berg and Neuhaus [7,41–44] to the simulation of a variety of systems exhibiting first-order phase transitions [45–50]. In this procedure, the biasing weight $w(S)$ of a configuration with action S is dynamically adjusted (bootstrapped) such as to achieve a near-constant overall frequency distribution over a wide range of S within a *single* simulation.

Obviously, the MUCA in principle offers a powerful handle to deal with SCSD. It remains then a practical question whether one can indeed proceed to large lattices by boosting tunneling rates from the SCSD behavior [Eq. (1)] to the peak efficiency of local Monte Carlo methods (characterized by $O((L^4)^2)$ complexity]. This leads us to the key point of this paper: it is a severe shortcoming of the multicanonical algorithm that its implementation is seemingly restricted to sequential computers, as it requires knowledge of the *global* action, even during local updating. We will show that the HMC algorithm is from the very outset able to implement the MUCA in a parallel manner.

In Secs. II A and II B, we will give a short review of the MUCA and the HMC algorithms. In Sec. II C, we merge the MUCA with the HMC algorithm. From our ongoing simulation project of U(1) theory on the Wilson line [51], we determine the tunneling efficiency compared to the standard metropolis algorithm which in our case is complemented by three reflection steps. In Sec. III, we shall present our results for lattice sizes up to 16^4 and predict the tunnelling rates for lattice sizes up to 24^4, as would be required for a proper FSS.

II. MULTICANONICAL HYBRID MONTE CARLO ALGORITHM

The hybrid Monte Carlo algorithm [8,52,53] produces a global trial configuration by carrying out a molecular dynamics evolution of the field configuration very close to the surface of constant action. Subsequently, a Monte Carlo decision is imposed which is based on the global action difference ΔS, being small enough to be frequently accepted. Within the HMC algorithm, all degrees of freedom can be changed simultaneously and hence in parallel. This then provides a straightforward path to implement the MUCA as part of the HMC algorithm on parallel machines:[1] one just uses the values of the global action, as provided by the HMC algorithm, to compute the bias function $w(S)$ for the MUCA.

A. Multicanonical algorithm

The ''canonical'' Monte Carlo algorithm generates a sample of field configurations $\{\phi\}$, within a Markov process according to the Boltzmann weight

$$P_\beta(S) = \frac{1}{Z_{\mathrm{can}}} e^{-\beta S(\phi)}, \tag{2}$$

which follows from maximizing the entropy with respect to all possible probability distributions $P[\phi]$. The partition function Z normalizes the total probability to 1,

$$Z_\beta = \int [d\phi] e^{-\beta S[\phi]}. \tag{3}$$

S is the action (the energy in the case of statistical mechanics) and β the coupling (or inverse temperature $1/kT$).

The canonical action density, which in general exhibits a double-peak structure at a first-order phase transition, can be rewritten as

$$N_{\mathrm{can}}(S,\beta) = \rho(S) e^{-\beta S}, \tag{4}$$

with the spectral density $\rho(S)$ being independent of β. Usually, $\int N_{\mathrm{can}}(S,\beta) dS$ is set to 1.

The multicanonical approach aims at generating a flat action density

$$N_{\mathrm{MUCA}}(S,\beta) = \mathrm{const} \quad \text{for } S_{\min} \leqslant S \leqslant S_{\max}, \tag{5}$$

in a range of S that covers the double peaks at the first-order phase transition.

Therefore, instead of sampling canonically according to $e^{-\beta S}$, one modifies the sampling by a weight factor W_{MUCA}:

$$W_{\mathrm{MUCA}}(S,\beta) = \begin{cases} \dfrac{1}{\rho(S_{\min})} e^{\beta S_{\min}} & \text{for } S < S_{\min}, \\ \dfrac{1}{\rho(S)} e^{\beta S} & \text{for } S_{\min} \leqslant S \leqslant S_{\max}, \\ \dfrac{1}{\rho(S_{\max})} e^{\beta S_{\max}} & \text{for } S > S_{\max}, \end{cases} \tag{6}$$

[1] A first attempt in this direction has been made in Ref. [54] in the framework of the Higgs-Yukawa model.

which is constant outside the relevant action range. This is equivalent to a net sampling according to the *yet* unknown probability distribution

$$w(S)=\frac{1}{\rho(S)} \quad \text{for } S_{\min}\leqslant S\leqslant S_{\max}. \tag{7}$$

Since $W(S,\beta)$ is unknown at the begin of the simulation, it is instrumental for the MUCA to follow Münchhausen and start from good ''guesstimates'' [43]. We shall do so by starting from an observed histogram of the canonical action density, $\hat{N}_{\rm can}(S,\hat{\beta}_c)$, [see Eq. (4)] at the supposed location of the phase transition[2]: $\hat{\beta}_c$. From the action density, we compute $\hat{W}_{\rm MUCA}(S,\hat{\beta}_c)$ according to Eq. (6). The sampling then proceeds with the full MUCA weight,

$$\hat{P}_{\rm MUCA}(S)\propto e^{-\hat{\beta}_c S}\hat{W}_{\rm MUCA}(S,\hat{\beta}_c)\propto e^{-[\hat{\beta}_c+\hat{\beta}(S)]S-\hat{\alpha}(S)}. \tag{8}$$

This latter formulation can be interpreted as a simulation proceeding at varying couplings (temperatures), hence the name ''*multicanonical*.''

In order to compute expectation values of observables $\mathcal{O}$, one has to reweight the resulting action density eventually by the factor $\hat{W}_{\rm MUCA}(S,\hat{\beta}_c)$, which reconstitutes the proper canonical density:

$$\langle\hat{\mathcal{O}}_{\hat{\beta}_c}\rangle=\frac{\sum_i \mathcal{O}^i_{\hat{\beta}_c}\dfrac{1}{\hat{W}_{\rm MUCA}(S_i,\hat{\beta}_c)}}{\sum_i \dfrac{1}{\hat{W}_{\rm MUCA}(S_i,\hat{\beta}_c)}}. \tag{9}$$

Additionally, $\langle\hat{\mathcal{O}}_{\hat{\beta}_c}\rangle$ simulated at $\hat{\beta}_c$ can be reweighted to any desired β (following [55]), given that the corresponding region of phase space has been covered by the MUCA simulation. We emphasize that Eq. (9) is only useful complemented by a proper error analysis. The canonical error computed from the multicanonical ensemble has been elaborated in Ref. [56].

Note that there are many possible choices for the form of the multicanonical weight. Just for technical reasons we require it to be continuous in S. One can either guess an analytic function, or choose a polygonal approximation such as given in Eq. (8). In this case, the multicanonical weight is expressed in terms of the functions $\hat{\beta}(S)$ and $\hat{\alpha}(S)$ which are actually characteristic functions of the bins. $\hat{\beta}(S)$ can be considered as an effective temperature [37].

The computation of the weights requires knowledge of the global and not just the local change in action for each single update step. For this reason, even for a local action, one cannot perform local updating moves in parallel, such as the well-known checkerboard pattern. As a consequence, the MUCA is not able to be made parallel for local update algorithms. For a remedy, we propose here to *go global* and utilize the HMC updating procedure.

[2]Quantities with carets refer to stochastic estimates.

B. Hybrid Monte Carlo algorithm

The HMC algorithm consists of two parts: the hybrid molecular dynamics (HMD) algorithm evolves the degrees of freedom by means of molecular dynamics (MD) which is followed by a global metropolis decision to render the algorithm exact.

In addition to the gauge fields $\phi_\mu(x)$ one introduces a set of statistically independent canonical momenta $\pi_\mu(x)$, chosen at random according to a Gaussian distribution $\exp(-\pi^2/2)$. The action $S[\phi]$ is extended to a guidance Hamiltonian

$$\mathcal{H}[\phi,\pi]=\frac{1}{2}\sum_{\mu,x}\pi^2_\mu(x)+\beta S[\phi]. \tag{10}$$

Starting with a configuration (ϕ,π) at MD time $t=0$, the system moves through phase space according to the equations of motion

$$\dot{\phi}_\mu=\frac{\partial\mathcal{H}}{\partial\pi_\mu}=\pi_\mu,$$
$$\dot{\pi}_\mu=-\frac{\partial\mathcal{H}}{\partial\phi_\mu}=-\frac{\partial}{\partial\phi_\mu}[\beta S], \tag{11}$$

leading to a proposal configuration (ϕ',π') at time $t=\tau$. Finally this proposal is accepted in a global metropolis step with probability

$$P_{acc}=\min(1,e^{-\Delta\mathcal{H}}),\ \text{with}\ \Delta\mathcal{H}=\mathcal{H}[\phi',\pi']-\mathcal{H}[\phi,\pi]. \tag{12}$$

The equations of motion are integrated numerically with finite step size Δt along the trajectory from $t=0$ up to $t=N_{\rm MD}$. Using the leapfrog scheme as symplectic integrator the discretized version of Eq. (11) reads

$$\phi^{n+1}=\phi^n+\Delta t\cdot\pi^n-\frac{\Delta t^2}{2}\frac{\partial}{\partial\phi}(\beta S[\phi^n])$$
$$\pi^{n+1}=\pi^n-\frac{\Delta t}{2}\frac{\partial}{\partial\phi}(\beta S[\phi^n])$$
$$-\frac{\Delta t}{2}\frac{\partial}{\partial\phi}(\beta S[\phi^{n+1}]). \tag{13}$$

Here we have presented the scheme with both the momenta and the gauge fields defined at full time steps,[3], $t=n\Delta t$.

[3]Note that the actual implementation computes the momenta at half time steps according to the sequence

$$\pi(t+\Delta t/2)=\pi(t-\Delta t/2)-\Delta t\frac{\partial}{\partial\phi}\{\beta S[\phi(t)]\},$$
$$\phi(t+\Delta t)=\phi(t)+\Delta t\ \pi(t+\Delta t/2),$$

initialized and finished by a half-step in π [8]. Each sequence approximates the correct $\mathcal{H}$ with an error of $O(\Delta t^3)$.

To ensure that the Markov chain of gauge field configurations reaches a *unique* fixed point distribution $\exp(-S[\phi])$ one must require the updating procedure to fulfill *detailed balance*, which is guaranteed by the iterative map of Eq. (13), $f:(\phi,\pi)\rightarrow(\phi',\pi')$ being, time reversible $f(\phi',-\pi')=(\phi,-\pi)$ and measure preserving $[d\phi']\times[d\pi']=[d\phi][d\pi]$.

It is easy to prove that these two conditions also hold for the multicanonical action. Note that the guidance Hamiltonian, Eq. (10), defining the MD may differ from the acceptance Hamiltonian in Eq. (12), which produces the equilibrium distribution proper. In the following, we shall exploit this freedom to develop two variant mergers of the MUCA and HMC algorithm.

C. Merging the MUCA and HMC algorithm for compact QED (MHMC)

We consider a multicanonical HMC for pure four-dimensional U(1) gauge theory with standard Wilson action defined as

$$S[\phi]=\sum_{x,\nu>\mu}\{1-\cos[\theta_{\mu\nu}(x)]\}, \tag{14}$$

where

$$\theta_{\mu\nu}(x)=\phi_\mu(x)+\phi_\nu(x+\hat{\mu})-\phi_\mu(x+\hat{\nu})-\phi_\nu(x)$$

is the sum of link angles that contribute to one of six plaquettes interacting with the link angle $\phi_\mu(x)$.

Equation (8) suggests to consider an effective action $\hat{S}$ including the ''multicanonical potential'' V_{MUCA}:

$$\hat{S}=\hat{\beta}_c S+\hat{V}_{\mathrm{MUCA}}(S,\hat{\beta}_c), \tag{15}$$

with

$$\begin{aligned}\hat{V}_{\mathrm{MUCA}}(S,\hat{\beta}_c)&=\log\left(\frac{1}{\hat{W}_{\mathrm{MUCA}}(S,\hat{\beta}_c)}\right)\\&=\hat{\beta}(S)S+\hat{\alpha}(S).\end{aligned} \tag{16}$$

There are two natural options to proceed from here.

Method 1 performs molecular dynamics using the canonical guidance Hamiltonian

$$\mathcal{H}_{\mathrm{MD}}=\frac{1}{2}\sum\ \pi^2+\hat{\beta}_c S,$$

with standard action S. The resulting gluonic force is given by

$$\begin{aligned}F(x)&=\dot{\pi}_\mu(x)\\&=\hat{\beta}_c\sum_{\nu\neq\mu}[\sin\theta_{\mu\nu}(x-\hat{\nu})-\sin\theta_{\mu\nu}(x)].\end{aligned} \tag{17}$$

Method 2 makes use of the multicanonical potential as a driving term within the Hamiltonian

$$\mathcal{H}_{\mathrm{MD}}=\frac{1}{2}\sum\ \pi^2+\hat{S},$$

inducing an additional drift term

$$\begin{aligned}\dot{\pi}_\mu(x)&=F(x)-\frac{\partial}{\partial\phi_\mu(x)}\hat{V}_{\mathrm{MUCA}}(S[\phi(x)],\hat{\beta}_c)\\&=[\hat{\beta}_c+\hat{\beta}(S)]\sum_{\nu\neq\mu}[\sin\theta_{\mu\nu}(x-\hat{\nu})-\sin\theta_{\mu\nu}(x)],\end{aligned} \tag{18}$$

with the effective $\hat{\beta}$ as defined in Eq. (8).

For both options, the Hamiltonian governing the accept/reject decision, Eq. (12), reads equally,

$$\mathcal{H}_{\mathrm{acc}}=\frac{1}{2}\sum\ \pi^2+\hat{S}. \tag{19}$$

The latter method is governed by the dynamics underlying the very two peak structure: as one can see from Fig. 5, below, V_{MUCA} is repelling the system out of the hot (cold) phase towards the cold (hot) phase, thus increasing its mobility and enhancing flip-flop activity.

We comment that the implementation of method 2 requires the computation of the global action [to adjust the correct multicanonical weight, Eq. (8)] at each integration step along the trajectory of molecular dynamics to guarantee reversibility. In the polygon approximation, this amounts to a determination of the effective coupling $\hat{\beta}$ at each time step in MD. Note that $\hat{\alpha}$ does not influence MD but enters into the global metropolis decision, Eq. (12). Method 1 is much simpler, running at *fixed* trial coupling $\hat{\beta}_c$ and avoiding the effort of computing $N_{\mathrm{MD}}-2$ global sums while traveling along the trajectories. It turned out that of both versions of the MHMC algorithm, method 2 performs better than method 1, when autocorrelation and difference in computational effort are taken into account. Thus, we continue our analysis by investigation of method 2.

III. RESULTS

In order to evaluate the efficiency of the MHMC algorithm and to set the stage for a proper extrapolation, we generated time series of the U(1) action on lattices of size 6^4 up to 16^4. The runs are summarized in Table I. We applied method 2 as being the more promising one on real large lattices.

To arrive at an action density $N_{\mathrm{MUCA}}(S,\beta)$ which is approximately flat in the desired region between the two peaks it is crucial to find a good estimate $\hat{N}_{\mathrm{can}}(S,\hat{\beta}_c)$ of the canonical action density. However, it becomes more and more delicate for large volumes to find the proper multicanonical weight $\hat{P}^L_{\mathrm{MUCA}}(S)$ [Eq. (8)]. Figure 1 shows the evolution of the action density as the lattice size increases.

For lattices $<16^4$ it was sufficient to perform a short canonical run to generate an action density $\hat{N}_{\mathrm{MUCA}}$ suitable to

TABLE I. Total numbers of sweeps carried out both for the MRS and MHMC algorithm at different lattice sizes L and couplings β.

L	β	No. sweeps (MRS)	No. sweeps (MHMC)	Δt	N_{MD}
6	1.001600	1.460.000	1.650.000	0.120	10
8	1.007370	1.320.000	1.430.000	0.093	13
10	1.009300	1.030.000	560.000	0.071	17
12	1.010150	680.000		0.060	20
	1.010143	1.790.000	1.160.000		
14	1.010300	1.430.000		0.050	24
	1.010668	900.000	990.000		
16	1.010800	1.210.000		0.045	26
	1.010753	750.000	760.000		

compute a proper weight factor $\hat{W}_{\mathrm{MUCA}}$. On the 16^4 system, however, the resulting multicanonical distribution becomes quite sensitive to the choice of the weight. Therefore, in the case of large lattices ($L \geqslant 16$) we cannot rely on canonical simulations to start with. Even if we perform $O(10^6)$ sweeps using a standard metropolis update with three reflection steps,[4] SCSD prevents a sufficiently accurate determination of the phase weight. Therefore, we install a recursive procedure: from a previous guess $\hat{N}_{\mathrm{i}}(S,\hat{\beta}_c^i)$ we go through the MHMC algorithm and arrive at $\hat{N}_{\mathrm{i+1}}(S,\hat{\beta}_c^{i+1})$. This computational scheme is initialized by a standard canonical ''short run.'' We found that one such learning cycle is sufficient. Figure 2 illustrates the evolution of the multicanonical action density on the 16^4 lattice.

On larger volumes the determination of a good guess can be considerably boosted by a crank-up extrapolation that starts from smaller systems [43]. In Fig. 3, we display the quality of ''flatness'' of $\hat{N}_{\mathrm{MUCA}}(S,\beta)$ achieved in our investigations for the various lattice sizes.

A. Tunneling behavior

With our estimate for $\hat{N}^{16}_{\mathrm{MUCA}}(S)$ at $\beta = 1.010753$, depicted in Fig. 2, we have generated the time history of the action per site, $s = S/6V$, as shown in the upper part of Fig. 4. For reference, we have included the time history from the MRS algorithm on the same lattice. The figure demonstrates the success of the MHMC algorithm: the method provides us with a gain factor in tunneling rate of about one order of magnitude on the 16^4 lattice.

In order to quantify this achievement, we introduce the average flip time τ_{flip}, a quantity that is readily measurable. τ_{flip} is defined as follows: we histogram the time series of s using N bins, as illustrated in Fig. 5. A suitable number is $N = 8$. The total binning range is adjusted such that 99.9% of the events are covered symmetrically by the eight bins. A flip (flop) is given when the system travels from I_8 to I_1 (and vice versa). τ_{flip} is defined as the inverse number of the sum of flips and flops multiplied by the total number of trajectories. In Table II, τ_{flip} is given for the various lattices. The error in τ_{flip} has been computed by a jackknife error analysis.

[4]The metropolis algorithm with reflection steps (MRS) is considered as a very effective local update algorithm for U(1) [5].

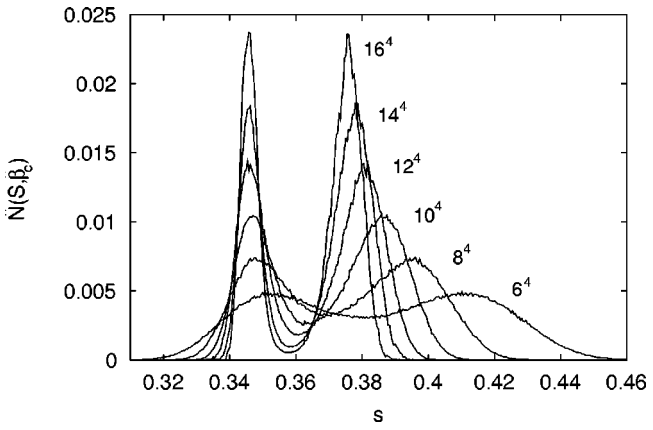

FIG. 1. The canonical action densities, Eq. (4), on 6^4 up to 16^4 lattices, reweighted to their respective $\hat{\beta}_c$ values, here defined via the equal height of the histograms.

B. Scaling behavior

With the results for τ_{flip} on lattices up to 16^4 we are in the position to estimate the scaling behavior of the MHMC algorithm in comparison to standard MRS updates. Figure 6 shows τ_{flip} both for the MHMC algorithm and MRS as a function of the lattice size L at their respective β values, as listed in Table II.

According to the expected exponential behavior of $\tau^{\mathrm{MRS}}_{\mathrm{flip}}$ which, in the asymptotic regime $L \to \infty$, should be given by Eq. 1, we perform a χ^2 fit with the ansatz

$$\tau^{\mathrm{MRS}}_{\mathrm{flip}} = a L^b e^{cL^3}. \tag{20}$$

It yields the following parameter values:

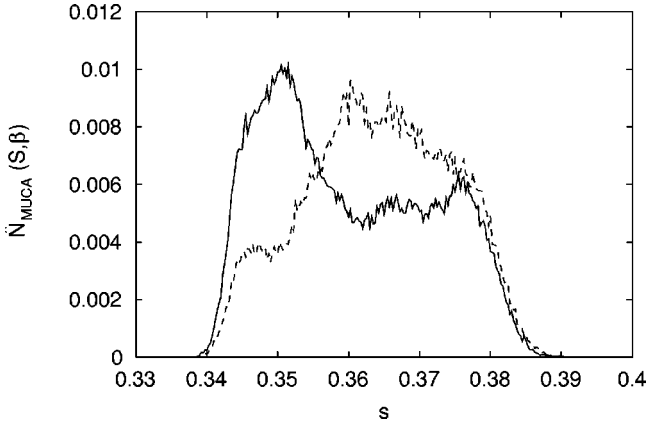

FIG. 2. Evolution of multicanonical action density. Starting from a canonical run a first MHMC simulation is performed (dashed line), and, based on this result, the final run (solid line) is carried out. Both curves are computed on the 16^4 lattice at $\beta = 1.010753$.

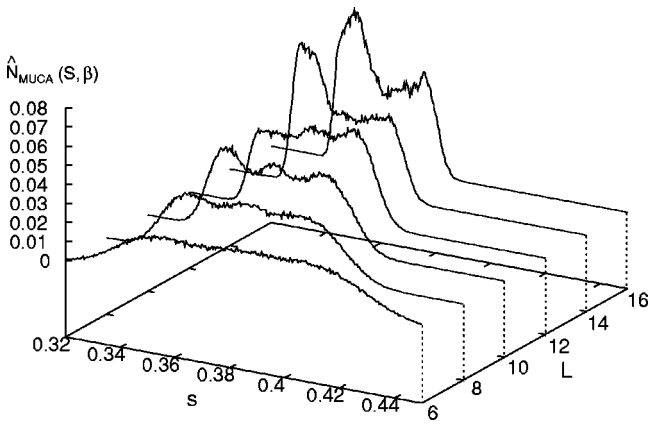

FIG. 3. $\hat{N}_{\mathrm{MUCA}}(S,\beta)$ at the run-parameter β as given in Table II.

$$a = 11.9(3.7),$$

$$b = 2.01(18),$$

$$c = 6.7(8)\times 10^{-4}, \tag{21}$$

with $\chi^2_{\mathrm{per\ DOF}} = 0.897$. As a result, we find a clear exponential SCSD behavior for the MRS algorithm.[5]

On the other hand, for the tunneling times of the MHMC algorithm, we expect a monomial dependence in L:

$$\tau_{\mathrm{flip}}^{\mathrm{MHMC}} = p\ \ L^q. \tag{22}$$

We obtain, for the fit parameters,

$$p = 11.6(1.6), \tag{23}$$

$$q = 2.238(68). \tag{24}$$

The power law ansatz is well confirmed by the fit quality with $\chi^2_{\mathrm{per\ DOF}} = 0.795$.

We also took the pessimistic ansatz and tried to detect a potentially exponential increase of $\tau_{\mathrm{flip}}^{\mathrm{MHMC}}$. The exponential fit gives $\chi^2_{\mathrm{per\ DOF}} = 0.975$. As can be seen in Fig. 6, the exponential contribution remains suppressed in the extrapolation. A potentially dominating exponential behavior for the MHMC algorithm can only be detected in future MHMC simulations on larger lattices. In other words, the parallel MHMC algorithm *is* capable of overcoming SCSD in compact QED in practical simulations, at least up to lattices sizes $\approx 24^4$.

IV. COST ESTIMATES FOR A FSS STUDY

Finally, we try to assess the computer effort required to perform a FSS study on a series of lattices ranging up to 24^4.

τ_{flip} being readily accessible, we relate this quantity to the

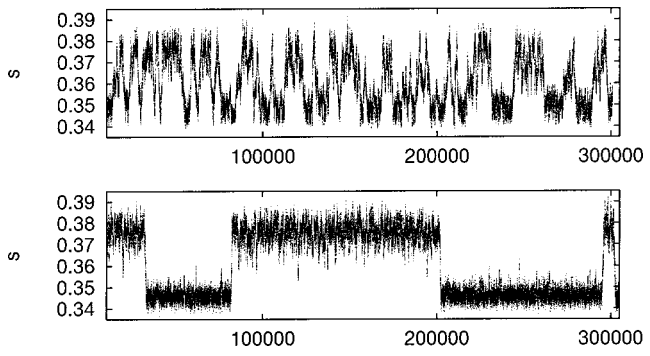

FIG. 4. Time history of the 16^4 system at $\beta = 1.010753$ for the MHMC algorithm (top) and MRS (bottom).

effective integrated autocorrelation time, τ_{eff}, defined in Ref. [56] by

$$\sigma^2_{\mathrm{MUCA}} = \sigma^2_{\mathrm{can}} \frac{2\tau_{\mathrm{eff}}}{N_{\mathrm{ts}}}. \tag{25}$$

σ^2_{MUCA} is the squared error of the observable $\mathcal{O}$ computed from the multicanonical ensemble; see Eq. (12). σ^2_{can} denotes the canonical variance of $\mathcal{O}$ (computed from the reweighted canonical ensemble) and N_{ts} is the length of the multicanonical time series. We can determine σ^2_{MUCA} in a numerically quite stable way from jackknife blocking. From the time series of S on the 14^4 and a 16^4 MUCA system, with about 550000 and 300000 entries, respectively, we have determined τ_{eff} to be

$$\tau^{14}_{\mathrm{eff}} = 35(10)\times 10^{-3}\times \tau^{14}_{\mathrm{flip}},$$

$$\tau^{16}_{\mathrm{eff}} = 41(12)\times 10^{-3}\times \tau^{16}_{\mathrm{flip}}. \tag{26}$$

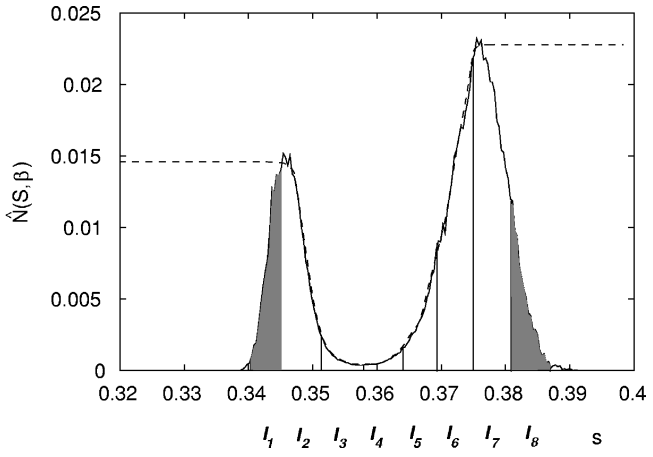

FIG. 5. Action density of the 16^4 system at $\beta = 1.010753$ as a function of $s = S/6V$ from which $\hat{V}_{\mathrm{MUCA}}$ is derived [the dashed line shows $\exp(\hat{V}_{\mathrm{MUCA}})$]. The subdivision of the support of the action density into eight intervals is introduced in order to define τ_{flip}.

[5]One is tempted to extract the interfacial surface tension σ from the fit to the MRS data. We find $\sigma = 3.35(39)\ \ 10^{-4}$.

TABLE II. τ_{flip} for the MRS and MHMC algorithm measured at the simulated β's.

L	β	$\tau_{\text{flip}}^{\text{MRS}}$	$\tau_{\text{flip}}^{\text{MHMC}}$
6	1.001600	508(12)	650(20)
8	1.007370	1023(60)	1173(50)
10	1.009300	2474(117)	2006(84)
12	1.010143	5470(770)	3260(440)
14	1.010668	16400(3300)	5090(630)
16	1.010753	44800(9700)	6350(860)

As a result, we found $\tau_{\text{flip}} \approx 0.038(8) \times \tau_{\text{eff}}$.[6] From here on we can estimate the number of decorrelated subsamples (independent measurements) out of a time series of length N_{ts} to be roughly

$$N_{\text{indep}} = \frac{N_{\text{ts}}}{2\,\tau_{\text{eff}}} = \frac{N_{\text{ts}}}{79(16)\ \ 10^{-3}\tau_{\text{flip}}}. \tag{27}$$

Assuming the inverse square root of N_{indep} to be an upper bound to the relative error r of an observable $\mathcal{O}$, we arrive at

$$K = \frac{79(16)\times 10^{-3}}{r^2}, \tag{28}$$

with K being the required number of flips to achieve a relative error $< r$.

We thus conclude that $O(100)$ flip-flops might allow one to determine quantities like the specific heat and the Binder-Landau cumulant with a relative error of 3%.

Obviously, the costs of MHMC and MRS simulations increase with the volume of the lattice, $V = L^4$. Additionally, for the MHMC algorithm, we want to keep the average acceptance probability of the leapfrog scheme constant. To this end, we have to lower the step size according to $\Delta t \sim V^{-1/4}$. In a detailed tuning investigation we have confirmed that the scaling rule of constant acceptance probability [57] leads to optimal performance. From a spectral analysis of the molecular dynamics we can find an optimized trajectory length N_{MD} (in the sense that the average acceptance probability is maximized) fulfilling $\Delta t\, N_{\text{MD}} = \text{const}(\beta)$, with only a slight dependence on β near the phase transition. We choose step-size Δt according to

$$\langle P_{\text{acc}} \rangle = \text{erfc}[\,c(\beta) V^{1/2} \Delta t^2] = \text{const}, \tag{29}$$

adjusted such that the product $\tau_{\text{int}} N_{\text{MD}}$ is minimized finally. In our case, the optimal acceptance probability is 65% [58].

Figure 7 confirms the scaling of the MHMC algorithm: the ratio of measured CPU times for a sweep, $t_{\text{MUCA}}/t_{\text{MRS}}$,

[6]Note that τ_{flip} strongly depends on the difference between I_8 and I_1 in Fig. 5. It remains to be confirmed that the ratio between τ_{eff} and τ_{flip} does not vary too much going to larger lattices.

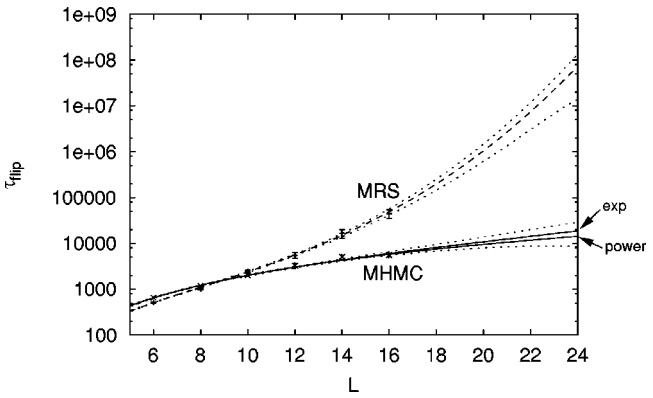

FIG. 6. Tunneling times for the MRS (exponential fit) and MHMC algorithm (lower convex curve is power law; upper convex curve is exponential fit). The errors of the two exponential fits are depicted as dotted lines. The error of the power law fit, Eq. (22), is not visible on this scale.

increases quite linear with $L = V^{1/4}$. Minor deviations from this behavior stem from the use of suboptimal run parameters Δt and N_{MD}.

The product $\tau_{\text{flip}} \times t_{\text{MUCA}}$ reflects the efficiency of the MHMC algorithm. Table III lists the effective gain factors achieved taking the MHMC algorithm instead of the MRS. The two columns refer to the power law and exponential extrapolations for the MHMC algorithm.

Let us finally translate these factors into real costs: in Fig. 8, we extrapolate the sustained CPU time in Tflop hours required to generate one flip. We conclude that the integrated CPU time to generate the required 100 flips with the MHMC algorithm on a 24^4 lattice amounts to about 3 Tflop hours of sustained CPU time.

V. SUMMARY AND OUTLOOK

We have demonstrated that the fully parallel MHMC algorithm is a very effective tool which is able to overcome SCSD as present in the pronounced metastabilities of four-dimensional U(1) gauge theory. A FSS study up to a lattice size of $L = 24$ with about 100 flip events for each lattice is

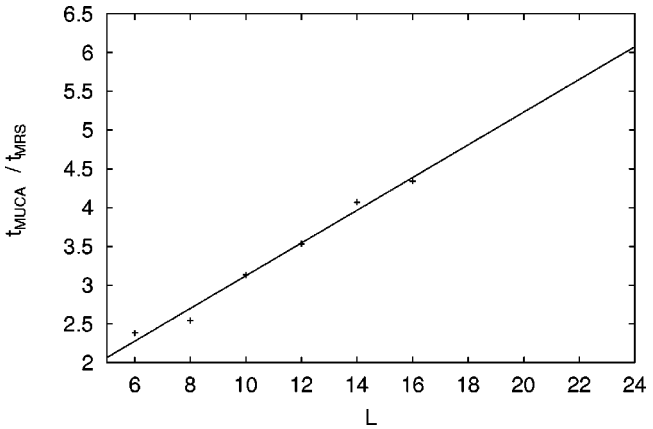

FIG. 7. Ratio of CPU times per sweep, $t_{\text{MUCA}}/t_{\text{MRS}}$ and linear fit. Errors are not visible on this scale.

TABLE III. Gain factor for the MHMC algorithm over the MRS as function of the linear lattice extension. The prediction for $\tau_{\mathrm{flip}}^{\mathrm{MHMC}}$ is based on a power law ansatz (row 1) and an exponential ansatz (row 2).

L	power	exponential
16	1.9(3)	1.9(3)
18	5.5(1.5)	5.1(1.6)
20	21.0(8.7)	18.4(9.0)
22	112(67)	92(65)
24	856(700)	652(638)

feasible within half a year run time, given a sustained performance of about 2 Gflops, due to the improvements achieved by the MHMC algorithm. These performance figures should be obtainable on a 32-node partition of a Cray T3E-600.

Less well known is the influence of the delicate part of the MHMC algorithm, i.e., the determination of a suitable estimate for V_{MUCA}, which is carried out in an iterative manner. So far, we have encouraging experiences on the 16^4 lattice. The success of the crank-up procedures described in Ref. [43] gives us hope that the P_{MUCA} determination will carry through with only marginal deterioration of the improvement factors estimated here.

The investigations presented form part of an ongoing study that aims at a conclusive FSS analysis of compact QED on the Wilson line [51].

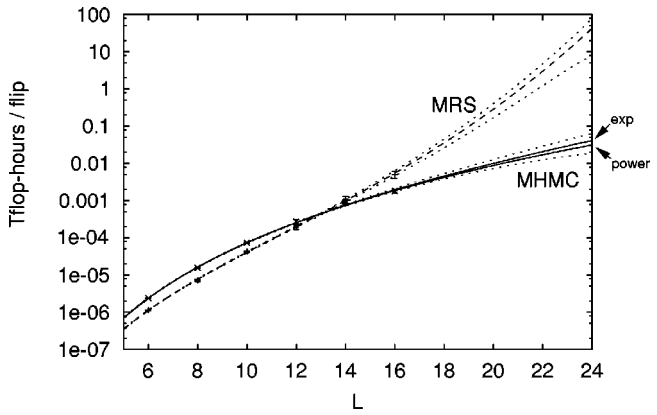

FIG. 8. Sustained CPU time in Tflop hours to generate one flip.

ACKNOWLEDGEMENTS

We thank P. Ueberholz and P. Fiebach for their friendly support. We are indebted to Thomas Riechmann and Dr. Claus-Uwe Linster of the Institut für Mathematische Maschinen at the computer center of Erlangen University, Germany, for providing us a substantial amount of computer time on their 32-node connection machine CM5. Without this support, the present investigation would not have been possible.

[1] A. H. Guth, Phys. Rev. D **21**, 2291 (1980).
[2] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[3] M. Creutz, L. Jacobs, and C. Rebbi, Phys. Rev. D **20**, 1915 (1979).
[4] S. Adler, Phys. Rev. D **23**, 2901 (1981).
[5] B. Bunk, ''Proposal for U(1) update'' (unpublished); and (private communication).
[6] M. E. Fisher and M. N. Barber, Phys. Rev. Lett. **28**, 1516 (1971).
[7] B. A. Berg and T. Neuhaus, Phys. Lett. B **267**, 249 (1991).
[8] S. Duane *et al.*, Phys. Lett. B **195**, 216 (1987).
[9] B. Lautrup and M. Nauenberg, Phys. Lett. **95B**, 63 (1980).
[10] G. Bhanot, Phys. Rev. D **24**, 461 (1981).
[11] K.-H. Mútter and K. Schilling, Nucl. Phys. **B200**, 362 (1982).
[12] R. Gupta, A. Novotny, and R. Cordery, Phys. Lett. B **172**, 86 (1986).
[13] J. Jersák, T. Neuhaus, and P. M. Zerwas, Phys. Lett. **133B**, 103 (1983).
[14] V. Azcoiti, G. di Carlo, and A. Grillo, Phys. Lett. B **268**, 101 (1991).
[15] G. Bhanot, Th. Lippert, K. Schilling, and P. Ueberholz, Nucl. Phys. **B378**, 633 (1991).
[16] C. B. Lang, Nucl. Phys. **B28** [FS18], 255 (1987).
[17] K. Decker, A. Hasenfratz, and P. Hasenfratz, Nucl. Phys. **B295** [FS21], 21 (1988).
[18] B. Klaus and C. Roiesnel, Phys. Rev. D **58**, 114509 (1998).
[19] G. Bhanot, Nucl. Phys. **B205**, 168 (1992).
[20] I. Campos, A. Cruz, and A. Tarancón, Phys. Lett. B **424**, 328 (1998).
[21] I. Campos, A. Cruz, and A. Tarancón, Nucl. Phys. **B528**, 325 (1998).
[22] H. G. Evertz, J. Jersák, T. Neuhaus, and P. M. Zerwas, Nucl. Phys. **B251**, 279 (1985).
[23] J. Cox, W. Franzki, J. Jersák, C. B. Lang, T. Neuhaus, and P. W. Stephenson, Nucl. Phys. **B499**, 371 (1997).
[24] J. Cox, W. Franzki, J. Jersák, C. B. Lang, T. Neuhaus, A. Seyfried, and P. W. Stephenson, Nucl. Phys. B (Proc. Suppl.) **63**, 691 (1998).
[25] V. Grósch *et al.*, Phys. Lett. **162B**, 171 (1985).
[26] C. B. Lang and T. Neuhaus, Nucl. Phys. B (Proc. Suppl.) **34**, 543 (1994).
[27] C. B. Lang and T. Neuhaus, Nucl. Phys. **B431**, 119 (1994).
[28] J. Jersák, C. B. Lang, and T. Neuhaus, Nucl. Phys. B (Proc. Suppl.) **42**, 672 (1995).
[29] M. Baig and H. Fort, Phys. Lett. B **332**, 428 (1994).
[30] A. Bode, Th. Lippert, and K. Schilling, Nucl. Phys. B (Proc. Suppl.) **34**, 1205 (1994).
[31] Th. Lippert, A. Bode, V. Bornyakov, and K. Schilling, Nucl. Phys. B (Proc. Suppl.) **42**, 684 (1995).
[32] Th. Lippert, V. Bornyakov, A. Bode, and K. Schilling, *Monopoles in Compact U(1) - Anatomy of the Phase Transition*,

edited by H. Toki, Y. Mizuno, H. Suganuma, T. Suzuki, and O. Miyamura, in *Proceedings of the International RCNP Workshop on Color Confinement and Hadrons (CONFINEMENT 95)*, Osaka, Japan, 1995 (World Scientific, Singapore, 1995), pp. 247–254.
[33] E. Marinari and G. Parisi, Europhys. Lett. **19**, 451 (1992).
[34] W. Kerler, C. Rebbi, and A. Weber, Phys. Rev. D **50**, 6984 (1994).
[35] W. Kerler, C. Rebbi, and A. Weber, Nucl. Phys. **B450**, 452 (1995).
[36] S. L. Adler, G. Bhanot, Th. Lippert, K. Schilling, and P. Ueberholz, Nucl. Phys. **B368**, 745 (1992).
[37] G. M. Torrie and J. P. Valleau, J. Comput. Phys. **23**, 187 (1977).
[38] G. M. Torrie and J. P. Valleau, J. Chem. Phys. **66**, 1402 (1977).
[39] J. P. Valleau, J. Comput. Phys. **91**, 193 (1996).
[40] K. Ding and J. P. Valleau, J. Chem. Phys. **98**, 3306 (1993).
[41] B. A. Berg and T. Neuhaus, Phys. Rev. Lett. **68**, 9 (1992).
[42] B. A. Berg, J. Stat. Phys. **82**, 323 (1996).
[43] B. A. Berg, in *Proceedings of the International Conference on Multiscale Phenomena and Their Simulations*, Bielefeld, 1996, edited by F. Karsch, B. Monien, and H. Satz (World Scientific, Singapore, 1997).
[44] B. A. Berg, Nucl. Phys. B (Proc. Suppl.) **63**, 982 (1998).
[45] B. Grossmann *et al.*, Phys. Lett. B **293**, 175 (1992).
[46] K. Rummukainen, Nucl. Phys. **B390**, 621 (1993).
[47] W. Janke and T. Sauer, Nucl. Phys. B (Proc. Suppl.) **34**, 771 (1994).
[48] W. Janke and T. Sauer, J. Stat. Phys. **78**, 759 (1995).
[49] U. H. E. Hansmann and Y. Okamoto, hep-lat/9411005.
[50] T. Neuhaus, Report No. Wuppertal WUB 96-30, hep-lat/9608043.
[51] G. Arnold, Th. Lippert, and K. Schilling (unpublished).
[52] S. A. Gottlieb, W. Liu, D. Toussaint, R. L. Renken, and R. L. Sugar, Phys. Rev. D **35**, 2531 (1987).
[53] Th. Lippert, in *Proceedings of Workshop of the Graduiertenkolleg at the University of Wuppertal on ''Field Theoretical Tools for Polymer and Particle Physics,''* edited by H. Meyer-Ortmanns, and A. Klümper, (Springer, Berlin, 1998), p. 122.
[54] A. Seyfried, Ph.D. thesis, Wuppertal University, 1998; and (private communication).
[55] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1988); **63**, 1658E (1989).
[56] W. Janke and T. Sauer, J. Stat. Phys. **78**, 759 (1995).
[57] M. J. Creutz, *Quantum Fields on the Computer*, Vol. 11 of Advanced Series on Directions in High Energy Physics (World Scientific, Singapore, 1992).
[58] G. Arnold, Th. Lippert, and K. Schilling (unpublished).