

# Efficient parallel tempering for first-order phase transitions

T. Neuhaus,<sup>1,2,\*</sup> M. P. Magiera,<sup>3,†</sup> and U. H. E. Hansmann<sup>1,‡</sup>

<sup>1</sup>John von Neumann Institute for Computing, Forschungszentrum Jülich, 52425 Jülich, Germany

<sup>2</sup>Institute für Physik, Johannes Gutenberg Universität Mainz, Mainz, Germany

<sup>3</sup>Department of Physics, Theoretical Physics, University Duisburg-Essen, 47048 Duisburg, Germany

(Received 25 April 2007; revised manuscript received 16 August 2007; published 9 October 2007)

We present a Monte Carlo algorithm that facilitates efficient parallel tempering simulations of the density of states  $g(E)$ . We show that the algorithm eliminates the supercritical slowing down in the case of the  $Q=20$  and  $Q=256$  Potts models in two dimensions, typical examples for systems with extreme first-order phase transitions. As recently predicted, and shown here, the microcanonical heat capacity along the calorimetric curve has negative values for finite systems.

DOI: [10.1103/PhysRevE.76.045701](https://doi.org/10.1103/PhysRevE.76.045701)

PACS number(s): 02.70.Rr, 64.60.Cn, 05.50.+q

Calorimetric experiments on nuclear matter [1] as well as on sodium (NaCl) clusters [2] have highlighted the need for understanding the thermodynamics of finite-sized systems. In a *microcanonical* description [3–6] the fundamental quantity is the *density of states*  $g(E)$  and the microcanonical entropy  $S(E)=k_B \ln g(E)$  defines the temperature curve  $T(E)=[\partial S(E)/\partial E]^{-1}$ . Note that in mixed phase systems exist regions where the microcanonical temperature  $T(E)$  decreases with increasing  $E$ . The resulting negative values for the caloric specific heat  $C_V(E)=1/V \partial E/\partial T(E)$  at constant volume  $V$  are observed in experiments, too [1]. Currently there is no mature theory that would explain these data.

Exploring such phenomena on a computer is extremely difficult. Large free energy barriers correspond to less probable and therefore seldom sampled configurations. The problem was addressed in statistical physics through multicannonical ensemble simulations [7] and later Wang-Landau sampling [8,9]. However, application of these methods is not always straightforward. Parallel tempering [10] is much simpler to implement but fails in systems with strong first-order phase transitions. Here, we present multiple Gaussian modified ensemble (MGME) sampling [11] to combine the ease of use of parallel tempering with the possibility of simulating first-order phase transitions. Our test cases are two-dimensional (2D) Potts models [12] on  $V=L^2$  boxes with  $Q=20$  and  $Q=256$ . Their partition function is given by a sum over all configurations

$$Z = \sum_{\text{conf.}} \exp\left(+\beta \sum_{nn} \delta_{q_i, q_j}\right) \quad (1)$$

with energies  $E$  ranging on the interval  $-2V, \dots, 0$  and with  $\beta=1/T$ . As these models exhibit very strong first-order transitions they should lead for finite systems to the predicted negative values of the caloric specific heat. MGME uses ideas from the Gaussian ensembles approach [13,14]. Here, the Boltzmann weight is modified at the transition point  $\beta_T$  by a Gaussian in the energy that relies on two parameters  $E_0$  and  $\Delta E_0$ :

$$p_B(\beta_T, E) = e^{-\beta_T E - [(E - E_0)/\Delta E_0]^2}. \quad (2)$$

The bimodal probability distribution function  $P(E)$  of the energy can be turned into a unimodal distribution, provided that the value of the Gaussian width parameter  $\Delta E_0$  is “small.” In MGME simulations the partition function is written as a product,

$$Z = \prod_{i=1}^{N_{\text{rep}}} Z_i. \quad (3)$$

The weights  $W_i(E)$  in the  $N_{\text{rep}}$  multiplicative terms  $Z_i = \sum_E g(E) \exp[-W_i(E)]$  are chosen such that for normalized functions  $P(E)$ ,

$$\sum_E \min[P_i(E), P_{i+1}(E)] \geq 0.63. \quad (4)$$

This overlap criterion is taken between all neighboring factors  $i$  and  $i+1$  ensuring a broad sampling of all energies.

The various partition functions  $Z_i$  are distributed on the nodes of a parallel computer. Utilizing local spin updates, as well as parallel tempering swaps [10] between neighboring partition functions  $Z_i$  and  $Z_{i+1}$ , one records the multihistogram  $H_{\text{MH}}(E)$  [15,16] of energy  $E$  occurring in  $Z$ . This allows one to obtain a stochastic estimate of  $g(E)$  by way of reweighting:

$$\ln g(E) = \text{const} + \ln\{H_{\text{MH}}(E)\} - \ln\left\{\sum_{i=1}^{N_{\text{rep}}} e^{-W_i(E) - F_i}\right\}. \quad (5)$$

The peculiar choice of the constant 0.63 in Eq. (4) yields swap acceptance rates  $P_{\text{acc}} \approx 0.5$  for unimodal  $P(E)$  functions. The  $N_{\text{rep}}+1$  free constants  $F_i$ , and the value “const” are determined self-consistently.

We factorize the partition function  $Z$  of Eq. (3) into three factors  $Z=Z_{\text{low-T}}, Z_{\text{MGME}}$ , and  $Z_{\text{high-T}}$ . The factors  $Z=Z_{\text{low-T}}$  and  $Z_{\text{high-T}}$  ensure broad sampling in low and high temperature phases.  $Z_{\text{MGME}}$  covers the mixed phase at  $\beta_T$  for energies  $e_o V \leq E \leq e_d V$ . For  $Q=20$ , we have the energy density values  $e_d = -0.626\,529\,17$  and  $e_o = -1.820\,684\,43$ , while  $\beta_T = \ln[1 + \sqrt{Q}]$ . We write

$$Z_{\text{high-T}} = \prod_{i=1}^{N_1} \sum_{\text{conf.}} e^{-\beta_i E} \Theta\left(e - \frac{e_o + e_d}{2}\right), \quad (6)$$

\*t.neuhaus@fz-juelich.de

†m.magiera@uni-duisburg.de

‡u.hansmann@fz-juelich.de

$$Z_{\text{MGME}} = \prod_{i=N_1+1}^{N_2} \sum_{\text{conf.}} e^{-\beta_T E - [(E - E_0^i)/\Delta E_0]^2}, \quad (7)$$

$$Z_{\text{low-T}} = \prod_{i=N_2+1}^{N_{\text{rep}}} \sum_{\text{conf.}} e^{-\beta_i E} \Theta\left(\frac{e_o + e_d}{2} - e\right), \quad (8)$$

where  $\Theta(z)$  is the Heaviside step function. The parameters  $\beta_i$  for  $i=1, \dots, N_1$  and  $i=N_2+1, \dots, N_{\text{rep}}$ , as well as the  $E_0^i$  for  $i=N_1+1, \dots, N_2$ , are subject to the “boundary conditions”

$$\beta_1 = 0, \quad \beta_{N_1} = \beta_{N_2+1} = \beta_T, \quad \beta_{N_{\text{rep}}} = \beta_{\text{max}}, \quad (9)$$

$$E_0^{N_1+1} = e_d V, \quad E_0^{N_2} = e_o V. \quad (10)$$

The  $\beta_{\text{max}}$  is chosen such that the overlap with ground state probability distributions

$$\sum_E \min[P(\beta_{\text{max}}, E), P(\beta \rightarrow \infty, E)] \geq 0.5 \quad (11)$$

is not small. On 2D square boxes with linear extent  $L$  we select Gaussian width parameter  $\Delta E_0$  with  $\Delta E_0 = C_{\Delta E_0} L$ .

In the following, we implement MGME simulations for the 2D Potts models at  $Q=20$  and  $Q=256$  on square boxes with periodic boundary conditions. We compare exact findings [17] to high precision Monte Carlo simulations of Eq. (3). In particular, we use the planar interface tension  $\sigma_0$  between disordered and ordered phases at  $\beta_T$  [18]. For these models one has the values  $2\sigma_0=0.37072$  and  $2\sigma_0=1.90338$ . This interface tension leads to the asymptotic residual supercritical slowing

$$\lim_{L \rightarrow \infty} \frac{\ln(\tau_{\text{erg}})}{L} = 0.1346 \times 2\sigma_0, \quad (12)$$

from the free energy difference of lens shaped versus circular droplets at the droplet-strip shape transition [19]. Finally, we note that the  $Q$ -state Potts models exhibits a trivial ground state degeneracy of  $n_{\text{GS}}=g(E=-2V)=Q$  on boxes with periodic boundaries.

None of the  $P_i(E)$  for single  $Z_i$  in Eq. (3) should depart from a *unimodal* shape as the existence of several “peaks” leads to metastability. Therefore, the value of the Gaussian width parameter  $\Delta E_0$  within  $Z_{\text{MGME}}$  of Eq. (7) cannot be too large. For the same reason, the Heaviside step functions of Eq. (6) and Eq. (8) turn all probabilities within  $Z_{\text{high-T}}$  and  $Z_{\text{low-T}}$  into unimodal distributions in energy. The Gaussian’s width parameter  $C_{\Delta E_0}$  was chosen as  $C_{\Delta E_0}=2$  for the  $Q=20$  case, and  $C_{\Delta E_0}=1$  for  $Q=256$ . Thus any simulation within a  $Z_i$  is far away from one of a microcanonical ensemble. At the same time,  $\Delta E_0$  values are small enough to inhibit multimodal  $P(E)$  functions. The total number of replica scales as  $N_{\text{rep}}(L)=\text{const } L$ , where we find  $\text{const} \approx 2.2$  at  $Q=20$  and  $\text{const} \approx 3.4$  at  $Q=256$ . Apparently the quantity  $N_{\text{rep}}(L)$  for 2D Potts models with first-order phase transitions behaves as in critical systems. Using Beale’s exact density of states [20] we find  $N_{\text{rep}}(L) \approx 1.3L$  for  $P_{\text{acc}} \approx 0.5$  parallel tempering simulations in the 2D Ising model. We also expect that  $\Delta E_0 \propto L$  and  $N_{\text{rep}}(L) \propto L^{d-1}$  is the proper choice for MGME

simulations of first-order phase transitions in  $d > 2$ .

Once the Gaussian width is fixed,  $Z = Z_{\text{low-T}}$ ,  $Z_{\text{MGME}}$ , and  $Z_{\text{high-T}}$  turn into functions of one unknown parameter, either  $\beta_i$  or  $E_0^i$ . We can simulate any of the  $Z_i$  with standard Monte Carlo and utilizing single histogram reweighting for the overlap equation (4) we determine the parameter of  $Z_{i+1}$ . This iterative procedure determines the parameter partition  $\beta_i$  and  $E_0^i$  in a robust way, and for the considered models it only requires a marginal CPU amount in units of the total cost.

Our main observable is the density of state  $g(E)$  which can be estimated by multihistogram reweighting, see Eq. (5). Once  $g(E)$  is known, all other quantities as the specific heat  $C_V(E) = 1/V \partial E / \partial T(E)$  and the ground state degeneracy  $n_{\text{GS}} = g(E=-2V)$  follow. The reweighting in Eq. (5) depends on the ratios  $F_i/F_{i+1}$  and an overall constant *const* that can be eliminated by the normalization  $\sum_E g(E) = Q^V$ . We obtain the ratios  $F_i/F_{i+1}$  through a self-consistent iteration as described in [15,16], and secondly also via the method of Bennett [21],

$$\frac{F_i}{F_{i+1}} = \ln \left( \frac{\langle \mathcal{M}[W_i(E_{i+1}) - W_{i+1}(E_{i+1})] \rangle_{i+1}}{\langle \mathcal{M}[W_{i+1}(E_i) - W_i(E_i)] \rangle_i} \right). \quad (13)$$

Here,  $\mathcal{M}[x]$  can be the Metropolis function  $\mathcal{M}[x] = \min[\exp(-x), 1]$ , or the Fermi function  $\mathcal{M}[x] = [\exp(x) + 1]^{-1}$ . We have checked for the 2D Ising model and the Potts models that the two versions of  $F_i/F_{i+1}$  yield statistically consistent results.

We simulate the partitions functions  $Z_i$  of Eq. (3) employing local heat bath steps [22]. The total Monte Carlo statistics is  $0.5 \times 10^9 V$  local spin updates for given box sizes  $L \leq 64$  at volume  $V=L^2$ . With one sweep consisting of  $V$  local Monte Carlo steps, this implies a statistics of  $0.5 \times 10^9 / N_{\text{rep}}$  Monte Carlo sweeps on each  $Z_i$  corresponding to a total statistics of order one to  $10^7$  sweeps. Parallel tempering swaps between spin configurations of  $Z_i$  and  $Z_{i+1}$  are performed one swap per sweep. Error bars are calculated with jackknife binning using 10 bins, and we use the generalized feedback shift register random number generator R1279.

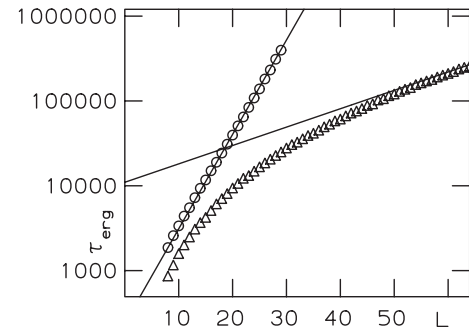


FIG. 1. Ergodicity autocorrelation times  $\tau_{\text{erg}}$  in units of sweeps for  $Q=256$  (circles) and  $Q=20$  (triangles) as a function of  $L$ . The straight lines have slopes as given in Eq. (12)

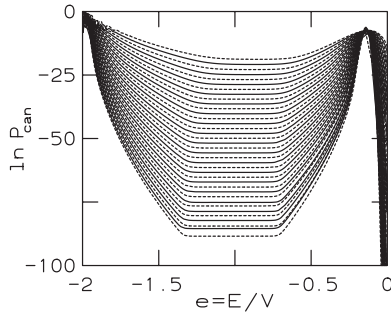


FIG. 2. Logarithmized energy distribution functions at  $\beta_T$  for  $L=8,9,\dots,43,44$  at  $Q=256$ . Dashed curves belong to even values of  $L$ , solid ones to odd values.

One important measure of computational efficiency in broad histogram sampling methods is the ergodicity time scale  $\tau_{erg}$  for the completion—in the mean—of one single “tunneling event” from the low to high—or high to low—temperature regions, see also Ref. [24,25]. We display in Fig. 1 our final  $\tau_{erg}$  results in units of sweeps in a logarithmic scale for the  $Q=20$  and  $Q=256$  Potts models. We only select runs that are longer than ten times  $\tau_{erg}$ . The figure contains straight lines that correspond to residual supercritical slowing down with slopes as given through Eq. (12). The data collapse of  $\tau_{erg}$  onto the prediction indicates the absence of supercritical slowing down, i.e., there are likely no time scales diverging as  $e^{2 \times \sigma_0 \times L}$ .

We have determined the density of states, and in Fig. 2 we display the probability distribution function  $P_{can}(E) = \ln g(E) - \beta_T E$  for the  $Q=256$  case. The number of ground states  $n_{GS} = g(E=-2V)$  is determined from a fit to all data. We find  $n_{GS} = 19.998(19)$  ( $Q=20$ ) and  $n_{GS} = 255.94(33)$  ( $Q=256$ ). In the  $Q=20$  Potts model, see Table I, we compare finite-size measurements of pseudocritical couplings  $\beta_{CV}$ , related to the specific heat maximum, with earlier findings from multicanonical simulations [23]. The measurements agree within statistical error bars. These results and previous findings provide a high-precision test of our simulations and analysis methods. Table I also contains a datum on the fraction of the Monte Carlo statistics (third row) that was spent in mixed phase simulations. The statistical error bars within both studies are of comparable size, if they are scaled to equal numbers of sweeps. Thus MGME, multicanonical, and

TABLE I. Finite-size pseudocritical couplings  $\beta_{CV}$  as mentioned in the text for  $Q=20$  from current studies (MGME07) and from a work of Berg *et al.* of the year 1993 [23]: (BB11972).

$L$	Origin	Sweeps/ $10^6$	$\beta_{CV}$
16	BBN1993	104.4	1.036470(150)
16	MGME07	202.6	1.036366(103)
24	BBN1993	129.1	1.034820(100)
24	MGME07	203.6	1.034681(085)
32	BBN1993	40.5	1.033170(220)
32	MGME07	201.3	1.033031(055)

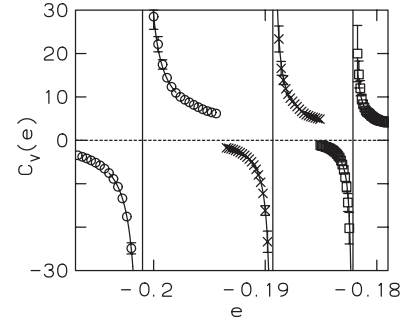


FIG. 3. Caloric specific heat  $C_V(e)$  for the  $Q=256$  Potts model on  $50^2$  (circles),  $70^2$  (tilted crosses), and  $90^2$  lattices (squares) as a function of  $e=E/V$ .

Wang-Landau simulation methods have comparable efficiency in computing  $\beta_{CV}$  as well as  $g(E)$ .

Finally, we display in Fig. 3 the caloric specific heat  $C_V(e)$  for the  $Q=256$  Potts model. As predicted earlier, this quantity can become negative. At least qualitatively, our plots look similar to the data in [1], which however correspond to a 3D experimental setting. As indicated by the vertical lines at  $e^*$  in the figure, we find a position of the singularity  $\Delta_e = e_d - e^*$ , which with increasing  $L$  decreases as  $L^{-d/(d+1)}$  in  $d=2$ . The quantity  $\Delta_e L^{d/(d+1)}$  has the values 0.8260(4), 0.8354(2), and 0.8436(3) on lattices of size  $50^2$ ,  $70^2$ , and  $90^2$ . It exhibits a mild finite-size dependence and is predicted to become constant in the  $L \rightarrow \infty$  limit. The  $C_V(e)$  discontinuities in Fig. 3 are consistent with the singular form

$$C_V(e) = \frac{\text{const}}{L^\alpha (e - e^*)} \quad (14)$$

with  $\alpha \approx 2.48$ . This implies that the width of the  $C_V(e)$  curves in Fig. 3 with increasing  $L$  decrease much faster than their positions  $\Delta_e = e_d - e^*$ . Such a sharpening at phase coexistence of droplets and vapor for atomic systems is predicted by Binder [26]. For other recent simulational data on the condensation and/or evaporation phase transition in the Ising model we refer to [27].

In conclusion, we have developed MGME sampling as a method for parallel tempering studies of first-order phase transitions and of mixed phases in finite-sized thermodynamic systems. It is particularly suited for the study of theories with a continuous energy function, as updates only require calculations of Gaussian's in the energy  $E$ , see Eq. (7). In addition, MGME sampling can be implemented easily on parallel computers. As the communication load is small, simulations of hundreds to thousands of processors are possible with little overhead. We therefore recommend MGME simulations as a tool for future studies in free energy landscapes of complex systems.

The calculations were done on the computers of the Neumann Institute for Computing, Forschungszentrum Jülich, Jülich, Germany. This work was supported in part by a research grant (Grant No. CHE-0313618) of the National Science Foundation (USA). T.N. thanks Professor K. Binder for useful discussions.

- [1] M. D'Agostino *et al.*, Phys. Lett. B **473**, 219 (2000).
- [2] M. Schmidt, R. Kusche, T. Hippler, J. Donges, W. Kronmüller, B. von Issendorff, and H. Haberland, Phys. Rev. Lett. **86**, 1191 (2001).
- [3] A. Hüller, Z. Phys. B: Condens. Matter **95**, 63 (1994).
- [4] A. Hüller, Z. Phys. B: Condens. Matter **93**, 401 (1994).
- [5] D. H. E. Gross, A. Ecker, and X. Z. Zhang, Ann. Phys. **5**, 446 (1996).
- [6] D. H. E. Gross, e-print arXiv:cond-mat/0206341.
- [7] B. A. Berg and T. Neuhaus, Phys. Rev. Lett. **68**, 9 (1992).
- [8] F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001).
- [9] F. Wang and D. P. Landau, Phys. Rev. E **64**, 056101 (2001).
- [10] K. Hukushima and K. Nemoto, J. Phys. Soc. Jpn. **65**, 1604 (1996).
- [11] T. Neuhaus and J. S. Hager, Phys. Rev. E **74**, 036702 (2006).
- [12] R. B. Potts, Proc. Cambridge Philos. Soc. **48**, 106 (1952).
- [13] M. S. S. Challa and J. H. Hetherington, Phys. Rev. Lett. **60**, 77 (1988).
- [14] M. Costeniuc, R. S. Ellis, H. Touchette, and B. Turkington, Phys. Rev. E **73**, 026105 (2006).
- [15] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1988).
- [16] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **63**, 1658 (1989).
- [17] R. J. Baxter, Solid State Phys. **6**, L445 (1973).
- [18] C. Borgs and W. Janke, J. Phys. I **2**, 2011 (1992).
- [19] T. Neuhaus and J. S. Hager, J. Stat. Phys. **113**, 47 (2003).
- [20] P. D. Beale, Phys. Rev. Lett. **76**, 78 (1996).
- [21] C. H. Bennett, J. Comput. Phys. **22**, 245 (1976).
- [22] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
- [23] A. Billoire, T. Neuhaus, and B. Berg, Nucl. Phys. B **396**, 779 (1993).
- [24] W. Nadler and U. H. E. Hansmann, Phys. Rev. E **75**, 026109 (2007).
- [25] S. Trebst, M. Troyer, and U. Hansmann, J. Chem. Phys. **124**, 174903 (2006).
- [26] K. Binder, Physica A **319**, 99 (2003).
- [27] A. Nussbaumer, E. Bittner, T. Neuhaus, and W. Janke, Europhys. Lett. **75**, 716 (2006).