Random-cluster multihistogram sampling for the q-state Potts model

Weigel, M. , Janke, W. and Hu, C.-K.

Original citation & hyperlink:
Weigel, M. , Janke, W. and Hu, C.-K. (2002) Random-cluster multihistogram sampling for the q-state Potts model. Physical Review E, volume 65 : 036109
http://dx.doi.org/10.1103/PhysRevE.65.036109

DOI 10.1103/PhysRevE.65.036109
ISSN 1539-3755
ESSN 1550-2376

Publisher: American Physical Society

Random-cluster multi-histogram sampling for the q-state Potts model

Martin Weigel* and Wolfhard Janke†
Institut für Theoretische Physik, Universität Leipzig, Augustusplatz 10/11, 04109 Leipzig, Germany
Chin-Kun Hu‡
Institute of Physics, Academia Sinica, Nankang, Taipei, Taiwan 11529, Republic of China

(Dated: July 06, 2001)

Using the random-cluster representation of the q-state Potts models we consider the pooling of data from cluster-update Monte Carlo simulations for different thermal couplings K and number of states per spin q. Proper combination of histograms allows for the evaluation of thermal averages in a broad range of K and q values, including non-integer values of q. Due to restrictions in the sampling process proper normalization of the combined histogram data is non-trivial. We discuss the different possibilities and analyze their respective ranges of applicability.

PACS numbers: 05.10.Ln, 05.50.+q, 75.10.Hk

I. INTRODUCTION

During the last decade the question of how to make most efficient use of the data sampled during a Monte Carlo (MC) simulation has received an increasing amount of attention. The idea of *reweighting* [1] of time-series data from a single canonical simulation at a given fixed value of a coupling parameter (i.e., most commonly temperature or magnetic field) to nearby regions of the coupling-parameter space, allows for the analysis of thermal averages as continuous functions of external parameters and thus a much more precise determination of extremal, pseudo-critical points. As an extension of this, the combination of data from simulations at *different* points in the coupling-parameter space, commonly known as *multi-histogram technique* [2], in principle allows to get accurate estimates for thermal averages over a macroscopical region of couplings from a relatively small number of simulation (that, however, generally has to be increased with the size of the system).

The basic problem with collapsing data from different simulations is that of finding the correct relative normalization of the single histograms. Consider the sampled energy histogram $\hat{H}_{K_i}(E)$ of, e.g., an Ising model simulation at the coupling $K_i = J\beta_i$, consisting of N energy-measurements. The thermal average of an observable $A(E)$ at K_i is just given by the time-series average in the importance-sampling scheme and thus insensitive to the value of the partition function at that point. Combining two histograms, however, amounts to the combination of the temperature-independent expressions

$$Z_{K_i}\hat{H}_{K_i}(E)/N\, e^{K_i E} \tag{1}$$

for different simulations i, where the partition function Z_{K_i} appears as a normalization constant. Thus, for correct relative normalization of the histograms to be combined, one has to know the ratio of partition functions $Z_{K_{i_1}}/Z_{K_{i_2}}$ or, equivalently, differences in the free energy densities $f_{K_{i_1}} - f_{K_{i_2}}$ at the simulated couplings K_i. In Ref. [2] this problem has been solved by an iterative solution of self-consistency equations for the free energy differences at adjacent simulation couplings K_i.

Since the combination given in Eq. (1) is nothing but an estimator for the density of (energy) states $\Omega(E)$, multi-histograming data analysis amounts to estimating the density of states of the variable that is thermodynamically conjugate to the considered coupling parameter. Going to the random-cluster representation of the Potts model, i.e., its interpretation as correlated percolation model [3, 4, 5], the relevant density of states is given by the number $g(b, n)$ of bond configurations with b bonds and n clusters on the lattice. Apart from gaining control over *two* parameters, the thermal coupling K and the number of states q, this language suggests the use of cluster estimators for thermal averages like correlation functions, which are known to yield a variance-reduction in certain situations [6]. One of us [7, 8] has proposed a multi-histogram technique for the q-state Potts model and simulations at different temperatures making use of the sampling of cluster decompositions of the lattice as they occur in the Swendsen-Wang cluster-update algorithm [9]. There, the relative normalization of the individual histograms at couplings K_i is accomplished by making use of the known absolute number of configurations with b active bonds on the lattice, which is just given by the binomial $\binom{\mathcal{E}}{b}$, $\mathcal{E}$ being the total number of bonds of the lattice. While this method appears advantageous at first sight and gives nice results from the cases of percolation ($q \to 1$) [10] and the Ising model ($q = 2$) [8], we find that this procedure is not the best choice of normalization for simulations of Potts models with q

*Electronic address: weigel@itp.uni-leipzig.de
†Electronic address: janke@itp.uni-leipzig.de
‡Electronic address: huck@phys.sinica.edu.tw

larger than 2 or 3 and propose a different approach for normalization to circumvent this problem.

The outline of the paper is as follows. In Section II we restate the multi-histogram approach of Ref. [8] in the random-cluster representation ("RC histograming"), which was originally formulated for simulations at fixed q only, and generalize it to simulations of multiple q values. Applying it to the $q = 10$ Potts model in two dimensions we find large deviations from the expected results. As an alternative, in Section III we propose an adaptively normalized RC multi-histograming ansatz. We discuss details of its implementation and present a comparative reweighting analysis for the $q = 10$ case. For energy-related observables we also compare histograming in the random-cluster language to the sampling of the *energy* density of states in the well-known framework of histograming in the energy/magnetization language ("EM histograming") [2]. Comparing both methods, in Section IV we track down the observed deviations with the first ansatz to be a result of the application of the above mentioned normalization condition. This problem can thus be resolved by the second ansatz. Finally, Section V contains our conclusions.

II. RC HISTOGRAMS AND ABSOLUTE NORMALIZATION

Consider the Hamiltonian of the q-state Potts model in zero magnetic field,

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} \delta(\sigma_i, \sigma_j), \quad \sigma_i = 1, \ldots, q, \tag{2}$$

on a general graph $\mathcal{G}$ with $\mathcal{N}$ sites and $\mathcal{E}$ bonds. Transforming to the random-cluster representation [3], the partition function becomes

$$Z \equiv \sum_{\{\sigma_i\}} e^{K \sum_{\langle i,j \rangle} \delta(\sigma_i, \sigma_j)} = \sum_{\mathcal{G}' \subseteq \mathcal{G}} (e^K - 1)^{b(\mathcal{G}')} \, q^{n(\mathcal{G}')}, \tag{3}$$

where the sum runs over all bond configurations $\mathcal{G}'$ on the graph (subgraphs), and $K = \beta J$ denotes the thermal coupling. Notice that the formulation (3) in contrast to that of Eq. (2) allows for a natural continuation of the model to *non-integer* values of the parameter q. Using the subgraph expansion of the q-state Potts in external field, one of us [5] has shown that the q-state Potts model can be considered as a bond-correlated percolation model (BCPM) with bond occupation probability $p = 1 - e^{-K}$. Eq. (3) can be rewritten as:

$$\begin{aligned} Z_{p,q}(\mathcal{G}) &= e^{K\mathcal{E}} \sum_{\mathcal{G}' \subseteq \mathcal{G}} p^{b(\mathcal{G}')} (1-p)^{\mathcal{E} - b(\mathcal{G}')} \, q^{n(\mathcal{G}')} \\ &= e^{K\mathcal{E}} \sum_{b=0}^{\mathcal{E}} \sum_{n=1}^{\mathcal{N}} g(b,n) \, p^b \, (1-p)^{\mathcal{E}-b} \, q^n, \end{aligned} \tag{4}$$

where $g(b, n)$ denotes the number of subgraphs of $\mathcal{G}$ with b activated bonds and n clusters resulting therefrom. This purely combinatorial quantity corresponds to the density of states of the BCPM.

The Swendsen-Wang cluster-update algorithm generates bond configurations drawn from the equilibrium canonical distribution of this model. Thus, the probability for the occurrence of a subgraph with b bonds and n clusters is given by

$$P_{p,q}(b,n) = W_{p,q}^{-1}(\mathcal{G}) \, g(b,n) \, p^b \, (1-p)^{\mathcal{E}-b} \, q^n, \tag{5}$$

which in turn is the expectation value of the normalized sampled histogram of bond configurations, i.e., $P_{p,q}(b,n) = \langle \hat{H}_{p,q}(b,n)/N \rangle$, where N denotes the length of the time series of measurements. Here, we separated the common factor $\exp(K\mathcal{E})$ from the partition function:

$$Z_{p,q}(\mathcal{G}) = e^{K\mathcal{E}} \, W_{p,q}(\mathcal{G}). \tag{6}$$

An estimator for the density of states $g(b, n)$ is therefore given by

$$\hat{g}(b,n) = W_{p,q}(\mathcal{G}) \frac{\hat{H}_{p,q}(b,n)}{p^b \, (1-p)^{\mathcal{E}-b} \, q^n \, N}. \tag{7}$$

Since the reduced partition function $W_{p,q}(\mathcal{G})$ is *a priori* unknown, the correct normalization of this estimator is not known at the beginning. Probably the most obvious way of fixing the normalization would be to estimate the reduced partition function $W_{p,q}(\mathcal{G})$ directly from Eq. (7). One can do better than that, however, by considering the accumulated density $g(b)$, which is obviously just a binomial [7],

$$g(b) = \sum_n g(b,n) = \binom{\mathcal{E}}{b}. \tag{8}$$

Imposing this restriction on the estimate $\hat{g}(b, n)$ also, one arrives at

$$\hat{C}_{p,q}(b) \equiv \frac{\hat{W}_{p,q}(\mathcal{G})}{p^b \, (1-p)^{\mathcal{E}-b} N} = \frac{\binom{\mathcal{E}}{b}}{\sum_n \hat{H}_{p,q}(b,n) \, q^{-n}}, \tag{9}$$

so that the absolute values of $\hat{g}(b, n)$ are now fixed by $\mathcal{E}$ independent normalization conditions, one for each number of active bonds b. Thus we have the following estimate for the density of states [8]:

$$\hat{g}(b,n) = \hat{C}_{p,q}(b) \, \hat{H}_{p,q}(b,n) \, q^{-n}. \tag{10}$$

Now, we want to combine the estimates $\hat{g}^{(i)}(b, n)$ from several simulations at different parameters (p_i, q_i), i.e., we want to do multi-histograming in both parameters, p and q. Then, we have

$$\hat{g}(b,n) = \sum_i \alpha_i(b,n) \, \hat{g}^{(i)}(b,n), \quad \sum_i \alpha_i(b,n) = 1. \tag{11}$$

Since we want to minimize the variance $\hat{\sigma}^2[\hat{g}]$ of the final estimate and the different simulations are statistically independent, the correct choice of the weights α_i obviously is given by

$$\alpha_i(b,n) = \frac{1/\hat{\sigma}^2[\hat{g}^{(i)}(b,n)]}{\sum_i 1/\hat{\sigma}^2[\hat{g}^{(i)}(b,n)]}. \tag{12}$$

From Eq. (10) we have

$$\begin{aligned}\hat{\sigma}^2[\hat{g}^{(i)}(b,n)] &= \hat{C}^2_{p_i,q_i}(b)\, q_i^{-2n}\, \hat{\sigma}^2[\hat{H}_{p_i,q_i}(b,n)] \\ &\approx \hat{C}^2_{p_i,q_i}(b)\, q_i^{-2n}\, \hat{H}_{p_i,q_i}(b,n), \end{aligned} \quad (13)$$

such that the variance-optimized estimate for $g(b,n)$ becomes

$$\hat{g}(b,n) = \frac{\binom{\mathcal{E}}{b} \sum_i \sum_\mu \hat{H}_{p_i,q_i}(b,\mu)\, q_i^{-\mu}\, q_i^n}{\sum_j \left[\sum_\nu \hat{H}_{p_j,q_j}(b,\nu)\, q_j^{-\nu}\right]^2 q_j^{2n}\, [\hat{H}_{p_j,q_j}(b,n)]^{-1}}. \quad (14)$$

In writing this expression we allow for several approximations: first, we treat $\hat{C}_{p_i,q_i}(b)$ as a parameter in Eq. (13) instead of taking its own variance into account; this is justified by the clear suppression of variance of this quantity as a sum as compared to the the variance of its summands $\hat{H}_{p_i,q_i}(b,n)$. Secondly, we take $\hat{\sigma}^2[\hat{H}(b,n)] = \hat{H}(b,n)$, i.e., we treat the individual bins (b,n) as independently distributed according to an uncorrelated $1/N$ statistics, which will in general not be exactly fulfilled. Since those assumptions only affect the variance of the final estimate, however, and do not introduce a bias, we consider them justified. Finally, we do not take autocorrelations between successive measurements (b,n) into account, i.e., we assume here and in the following that measurements in the sampling process are taken with a frequency around $1/\tau_{\rm int}$, where $\tau_{\rm int}$ denotes the integrated autocorrelation time, resulting in an effectively uncorrelated time series.

From the partition function Eq. (4) we infer the following cluster-language estimators for the free energy density $f = F/\mathcal{N}$, the per-site internal energy $u = U/\mathcal{N}$ and specific heat $c_v = C_v/\mathcal{N}$,

$$\begin{aligned}\hat{f} &= -\frac{1}{K\mathcal{N}} \ln \hat{Z}_{p,q}(\mathcal{G}), \\ \hat{u} &= -\frac{1}{p\mathcal{N}} \langle b \rangle_{\hat{g}}, \\ \hat{c}_v &= \frac{K^2}{p^2\mathcal{N}} \left[\langle (b - \langle b \rangle_{\hat{g}})^2 \rangle_{\hat{g}} - (1-p) \langle b \rangle_{\hat{g}} \right], \end{aligned} \quad (15)$$

cf. Appendix A. Here, the estimated expectation value of an observable $O(b,n)$ is defined as

$$\begin{aligned}\langle O(b,n) \rangle_{\hat{g}} &\equiv \hat{Z}^{-1}_{p,q}(\mathcal{G})\, e^{K\mathcal{E}} \sum_{b=0}^{\mathcal{E}} \sum_{n=1}^{\mathcal{N}} \hat{g}(b,n) \\ &\quad \times p^b\, (1-p)^{\mathcal{E}-b}\, q^n\, O(b,n). \end{aligned} \quad (16)$$

For the evaluation of magnetic observables one has to distinguish percolating clusters, denoted by indices π, from non-percolating, finite clusters, denoted by indices ϕ. Let $\{c(G')\}$ be the set of clusters of a subgraph G' of the lattice and $n_c(G')$ the number of sites in cluster c of G'. Then, consider the following microcanonical averages:

$$m_1^\pi(b,n) \equiv \frac{1}{\mathcal{N} g(b,n)} \sum_{\substack{G' \subseteq \mathcal{G}, b(G')=b, \\ n(G')=n}} \sum_{\{c^\pi(G')\}} n_c^\pi(G'), \quad (17)$$

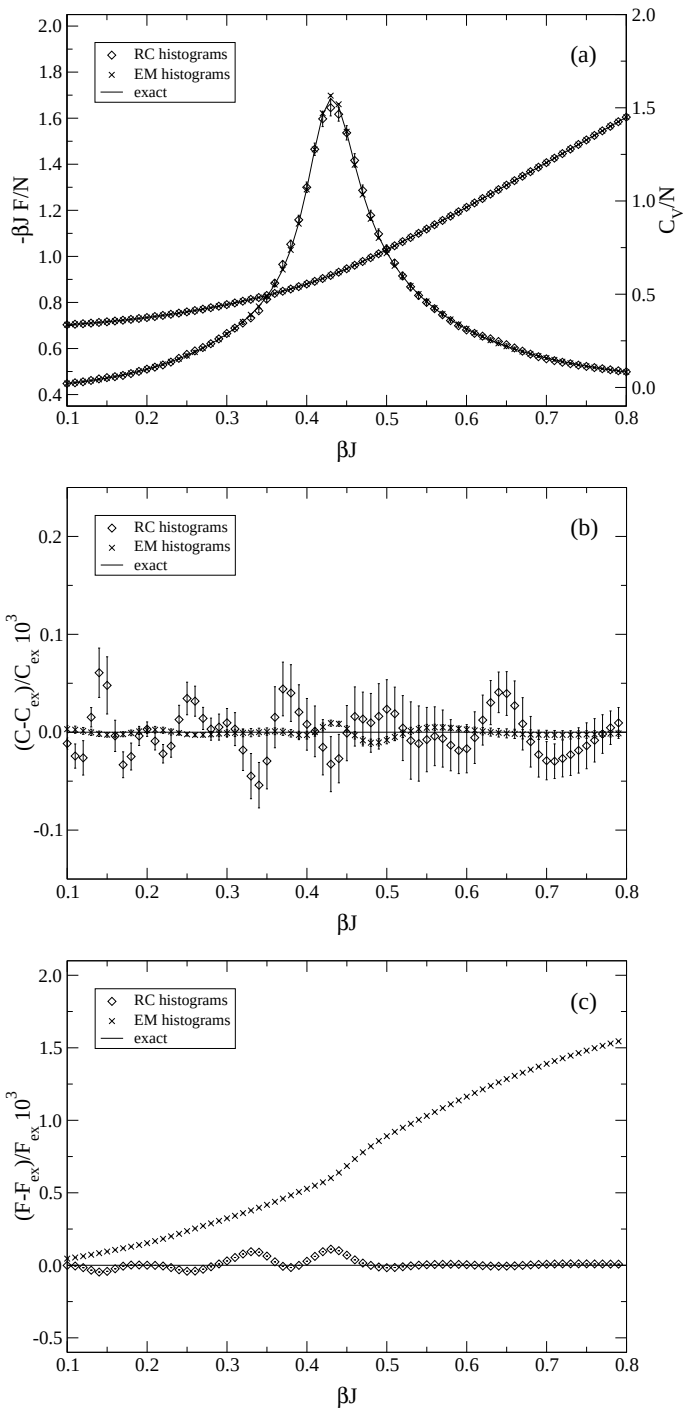

FIG. 1: Results from the EM and RC multi-histogram analyses of time series from 9 cluster update simulations of the two-dimensional $q = 2$ Potts model on a $\mathcal{N} = 16^2$ lattice. (a) Free energy density (left scale) and specific heat (right scale) as a function of the coupling $K = \beta J$ as compared to the exact solution of Ref. [11]. (b) Relative deviation of the results for the specific heat from the exact solution for both methods. (c) Relative deviation for the free energy. All results shown are re-scaled from the $q = 2$ Potts model to the Ising model formulation to fit the results from Ref. [11].

i.e., the average number of sites in percolating clusters

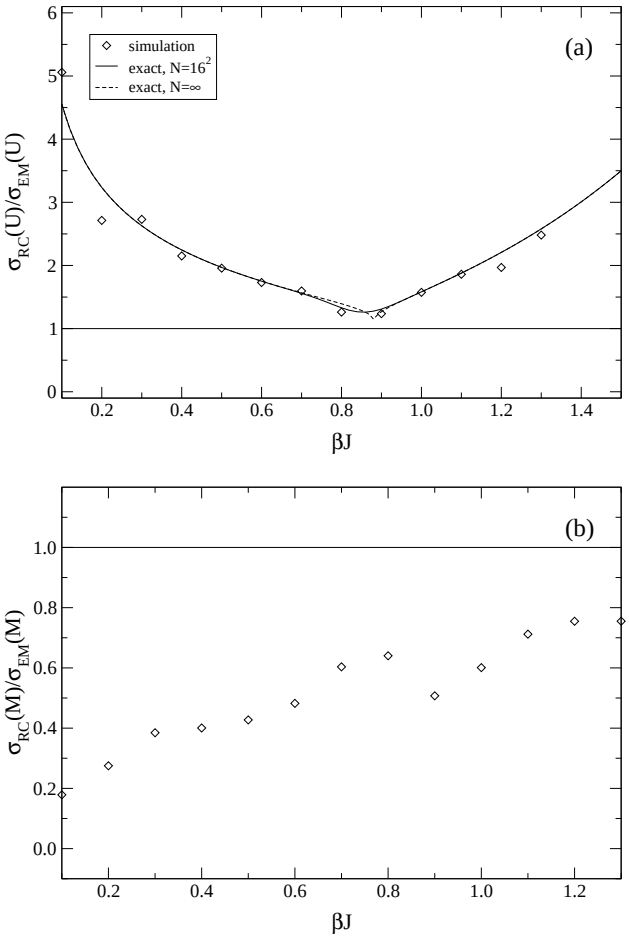

FIG. 2: Ratio of standard deviations for estimates of (a) the internal energy U and (b) the spontaneous magnetization $\tilde{M}$ of the $q = 2$ Potts model on a $\mathcal{N} = 16^2$ lattice and RC and EM histogram analyses as a function of the coupling $K = \beta J$. The variances are estimated by a "jackknife" time series analysis [12]. The solid line of (a) shows the exact result of Eq. (24) and Ref. [11], the dashed line that for $\mathcal{N} \to \infty$ from Ref. [13]. In (b) we use the definition (A9) for $K < 0.8$ and (A8) for $K \geq 0.8$.

for subgraphs with b active bonds and n clusters,

$$m_2^\pi(b,n) \equiv \frac{1}{\mathcal{N}^2 g(b,n)} \sum_{\substack{\mathcal{G}' \subseteq \mathcal{G}, b(G')=b, \\ n(G')=n}} \left[\sum_{\{c^\pi(G')\}} n_c^\pi(G') \right]^2, \tag{18}$$

i.e., the mean square number of sites in non-percolating clusters for those subgraphs, and

$$m_3^\phi(b,n) \equiv \frac{1}{\mathcal{N} g(b,n)} \sum_{\substack{\mathcal{G}' \subseteq \mathcal{G}, b(G')=b, \\ n(G')=n}} \sum_{\{c^\phi(G')\}} [n_c^\phi(G')]^2, \tag{19}$$

i.e., the mean squared sum of the size of non-percolating clusters. These microcanonical averages obviously can be estimated by adding $\sum_{\{c^\pi(G')\}} n_c^\pi(G')$

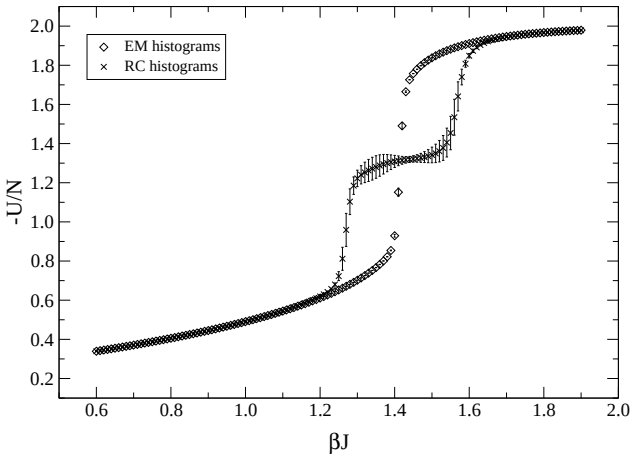

FIG. 3: Internal energy of the two-dimensional $q = 10$ Potts model on a $\mathcal{N} = 16^2$ square lattice with periodic boundary conditions as given by the random-cluster (RC) and energy representation (EM) multi-histogram analyses from simulations for different thermal couplings K. The transition point of the infinite system is given by $K_t = \ln(1 + \sqrt{10}) \approx 1.426$ [14].

to $\hat{M}_{1,p,q}^\pi(b,n)$ for each bond configuration G' with $\{b(G') = b, n(G') = n\}$, where $n_c^\pi(G')$ should be taken 0 for non-percolating bond configurations, adding $[\sum_{\{c^\pi(G')\}} n_c^\pi(G')]^2$ to $\hat{M}_{2,p,q}^\pi(b,n)$, and adding $\sum_{\{c^\phi(G')\}} [n_c^\phi(G')]^2$ to $\hat{M}_{3,p,q}^\phi(b,n)$ for each such observed configuration. Then, if we define

$$\begin{aligned} \hat{H}(b,n) &= \textstyle\sum_{p_i,q_i} \hat{H}_{p_i,q_i}(b,n), \\ \hat{M}_1^\pi(b,n) &= \textstyle\sum_{p_i,q_i} \hat{M}_{1,p_i,q_i}^\pi(b,n), \\ \hat{M}_2^\pi(b,n) &= \textstyle\sum_{p_i,q_i} \hat{M}_{2,p_i,q_i}^\pi(b,n), \\ \hat{M}_3^\phi(b,n) &= \textstyle\sum_{p_i,q_i} \hat{M}_{3,p_i,q_i}^\phi(b,n), \end{aligned} \tag{20}$$

we have the following estimates for $m_1^\pi(b,n)$, $m_2^\pi(b,n)$, and $m_3^\phi(b,n)$,

$$\hat{m}_{1/2/3}^{\pi/\phi}(b,n) = \frac{\hat{M}_{1/2/3}^{\pi/\phi}(b,n)}{\mathcal{N}\hat{H}(b,n)}, \tag{21}$$

which, finally, result in the following expressions for the (zero-field) magnetization $\tilde{m}$ and the magnetic susceptibility $\tilde{\chi}$,

$$\begin{aligned} \hat{\tilde{m}} &= \frac{q-1}{q} \left\langle \hat{m}_1^\pi(b,n) \right\rangle_{\hat{g}} + \frac{1}{q}, \\ \hat{\tilde{\chi}} &= \mathcal{N} \left(\frac{q-1}{q} \right)^2 \left[\left\langle m_2^\pi \right\rangle_{\hat{g}} - \left\langle m_1^\pi \right\rangle_{\hat{g}}^2 \right] \\ &\quad + \frac{q-1}{q^2} \left\langle m_3^\phi \right\rangle_{\hat{g}}, \end{aligned} \tag{22}$$

cf. Appendix A. Note, that we simply add up histograms from different simulations in Eqs. (20) and (21) without using any reweighting factors in p and q. This is correct since the conditional probability of the occurrence

of, say, a given number of sites in percolating clusters in a subgraph with b active bonds and n clusters does no longer depend on p and q. The order parameter m of the Potts model is given by [14]

$$m = \frac{q\tilde{m} - 1}{q - 1}, \tag{23}$$

and the corresponding, rescaled susceptibility is $\chi = [(q-1)/q]^2\tilde{\chi}$.

As a first comparative test for the method we performed a Swendsen-Wang cluster MC simulation for the Ising model case ($q = 2$) on a small, simple-cubic $\mathcal{N} = 16^2$ lattice with periodic boundary conditions. We gathered histograms from 9 different simulations at the couplings $K_i = 0.1, 0.2, \ldots, 0.8$ and $K = K_c = \frac{1}{2}\ln(1+\sqrt{2})$, where the couplings are given in the language of the Ising model in this case, i.e., are half of the couplings of the corresponding $q = 2$ Potts model. Each run sampled $2^{17} = 131\,072$ bond configurations resulting in corresponding time series of (b, n) and of the energy/magnetization pairs (E, M) for comparison with the EM histograming method. Thus, any differences in the results must be solely due to the method of data analysis, the underlying simulation data being exactly identical. For the EM histograms throughout this paper we use a multi-histogram analysis according to Ref. [2] very similar to that presented for the RC histograms in Sec. III. Figure 1 shows the results in comparison to the exact expression for F and C_v on square lattices as given by Kaufman [15] and analysed by Ferdinand and Fisher [11]. Statistical errors for both analysis schemes were evaluated using the "jackknife" error estimation technique [12]. The relative deviations $(\hat{C}_v - C_v)/C_v$ from the exact result are noticeably larger for the RC histogram analysis, however in agreement with statistical errors in both cases, cp. Fig. 1(b). The same holds true for the internal energy U given by the different estimates, which is not shown in Fig. 1. Note that for the energy related observables from Eq. (15) only the limiting distribution $P_{p,q}(b)$ is needed, which, in general, has a different width than the distribution of energies $P_{p,q}(E)$, thus leading to different variances.

In fact, by inspection of the random-cluster expression for the specific heat Eq. (15) and comparison with its definition in the energy language as $C_v = K^2(\langle E^2\rangle - \langle E\rangle^2)$ we can infer the following relation between the variances of energy estimates in the RC and EM schemes:

$$\frac{\sigma^2_{RC}(U)}{\sigma^2_{EM}(U)} = 1 - K^2\,\frac{1-p}{p}\,\frac{U}{C_v} \;\geq\; 1. \tag{24}$$

Thus, energy estimates from RC histograms are always less precise than those from EM histograms, regardless of the temperature. Figure 2(a) shows the ratio of jackknife-estimated variances of the two different estimates of internal energy, compared to the result from Eq. (24) with the exact expressions for U and C_v for the $q = 2$ case inserted [11]. Note, that from Fig. 2(a) this quantity seems to have extremely small finite-size corrections. As a reminder, this shows clearly that cluster estimators are not always improved estimators [16], but sometimes "deteriorated estimators". Note, however, that this effect will decrease with increasing number of states q, at least in the transition region, since the singularity in C_v sharpens in this limit, whereas the energies U always stay in the range $0 \leq U/\mathcal{N} \leq 2$. For the $q = 10$ model is has been observed that at the transition point $P_{p,q}(b)$ is almost indistinguishable from $P_{p,q}(E)$, when suitably rescaled [17]. The minimum of the exact curve of Fig. 2 at the critical point is somewhat in contrast to the usual notion that cluster estimators work best off criticality [6]; this result, however, applies to the spin-spin correlation function at medium and long distances and to magnetic observables like the susceptibility, which is the integral of the correlation function, whereas the internal energy U constitutes the extreme short distance limit of this quantity. For the magnetic observables m and χ the situation is reversed, the variance of the RC estimators being strongly reduced as compared to the EM estimators, cp. Fig. 2(b). Note that in contrast to the EM case, the RC estimators provide a single consistent definition of m and χ for both, the broken and unbroken phases, cf. Appendix A.

For the free energy, on the other hand, deviations for the RC method are by far smaller than those of the EM method, cp. Fig. 1(c). Moreover, deviations are not covered by statistical errors in the latter case, a fact we will comment on later in the next Section. In the EM case F is being fixed by making contact with the non-interacting limit $K = 0$ resp. $p = 0$, where

$$Z_{p=0,q}(\mathcal{G}) = \sum_E \Omega(E) = q^{\mathcal{N}}, \tag{25}$$

so that $-KF(p = 0)/\mathcal{N} = \ln q$. This equation corresponds to the normalization condition Eq. (8). It is obvious that having a normalization condition for each number b of active bonds and therefore, implicitly, for each (microcanonical) temperature in the RC case allows for accurate estimation of the free energy even far away from $K = 0$, whereas for EM histograms the results deteriorate with the distance from the only normalization point $K = 0$. Thus, for sampling free energies the RC multi-histogram technique normalized by Eq. (8) seems to be a good choice.

As a slightly less trivial example we performed simulations for the $q = 10$ Potts model on the same lattice, which exhibits a strongly first-order phase transition. It is well known that cluster algorithms are not efficient to reduce the "super-critical" (exponentially strong) slowing down of the local MC dynamics at first-order transitions. For the small lattice under consideration, however, autocorrelation times are still quite moderate, so that one gets reliable results without having to resort to more sophisticated methods like multi-canonical simulations [17, 18]. We gathered data from 11 single-histogram simulations at couplings $K_i = 0.8, 0.9, \ldots, 1.8$ with $2^{20} = 1\,048\,576$ measurements each. Figure 3 shows the quite astonish-

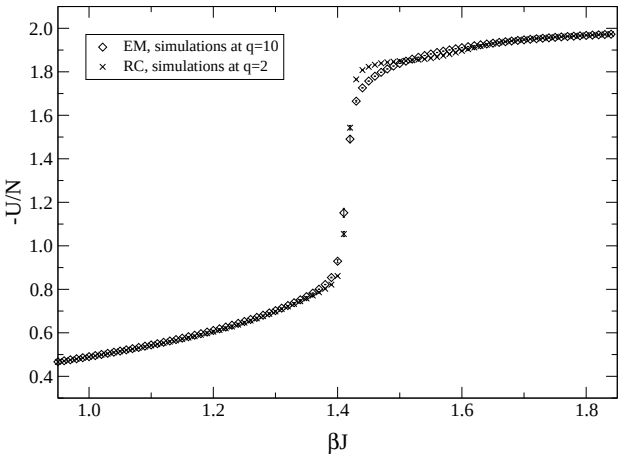

FIG. 4: Internal energy of the $q = 10$ Potts model on a $\mathcal{N} = 16^2$ square lattice, reweighted from the $q = 2$ model simulations shown in Fig. 1 using multiple RC histograms according to Eq. (14). The $q = 10$ results from the EM multi-histogram analysis are shown for comparison.

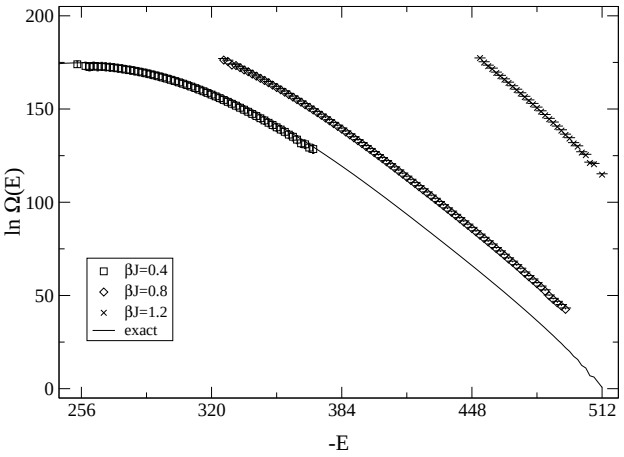

FIG. 5: Density of states for the $q = 2$ Potts model in two dimensions on a $\mathcal{N} = 16^2$ lattice from single-histogram cluster-update simulations at coupling $K = 0.4$, 0.8 and 1.2 using the estimator Eq. (26). The solid line shows the exact result of Ref. [20].

ing results for the internal energy from this simulation data using the analyses in the RC and EM languages, respectively. Naturally, we do not have exact results to compare with in this case; nevertheless, the results from the EM analysis are completely in agreement with our expectations and also well compatible with results from previous simulations [19]. So, obviously, the results from the RC histogram analysis are strikingly wrong — and in a way that is clearly not covered by the present statistical errors. Obviously, the results for the specific heat, which are not shown, look even worse, with a pronounced, unphysical double-peak resulting from the deviations in internal energy shown in Fig. 3.

Since as one of its major strengths in the RC approach we have the possibility of reweighting in the parameter q also, as a first clue to the reason for this conspicuous failure we show the outcome of using the $q = 2$ simulation data from above for determining the internal energy of the $q = 10$ case, cp. Fig. 4. The agreement with the direct EM analysis of the $q = 10$ simulations is remarkably good considering the large distance in q between the simulation and analysis points. Comparing Figs. 3 and 4 it is quite natural to suspect that the application of the normalization condition (8) is not a proper choice for simulation data from larger q models.

III. RC HISTOGRAMS AND ADAPTIVE NORMALIZATION

To understand this normalization problem let us shortly go back to the sampling of the energy density of states $\Omega(E)$ for the case of the two-dimensional $q = 2$ (Ising) model. Here, exact results are not only available for thermal averages, but for $\Omega(E)$ itself [20]. Using the $K = 0$ normalization condition (25), a single-histogram estimator for the density of states in the energy language would be given by

$$\hat{\Omega}(E) = 2^{\mathcal{N}} \, \frac{\hat{H}_K(E)\, e^{KE}}{\sum_E \hat{H}_K(E)\, e^{KE}}. \tag{26}$$

This works quite well in the high-temperature phase and for small lattices. For lower temperatures, however the histogram loses contact with the normalization point $K = 0$, resulting in large deviations from the correct normalization, cp. Fig. 5. Clearly, each simulation samples only a rather small window of energy space; from the exponential in the denominator of Eq. (26), however, configurations near the maximal energy $E = -\mathcal{N}$ receive the largest weight in the sum, so that missing those configurations, which is the case for large K, results in an exponentially wrong normalization factor (linear in $\ln \Omega(E)$). In other words, the absolute normalization condition (25) reweights the histogram data to the point $K = 0$, which will have no reliable outcome if the overlap between the histograms at the simulation coupling and at $K = 0$ is too small or even vanishes. Note also, that the statistical error bars given in Fig. 5 do not reflect this fundamental failure, although it is statistical in nature. This is due to the fact that the usual implementation of error estimation schemes for histograms take the error of histogram bins without entries to be zero, whereas according to $1/N$ statistics it should in some sense be considered infinitely large.

Expecting a similar sampling-related normalization failure for the RC histograms normalized by the condition (8) let us relax this absolute normalization and apply an adaptive normalization scheme as in the original EM multi-histograming formulation of Ref. [2]. Consider single-histogram estimates of the partition function from

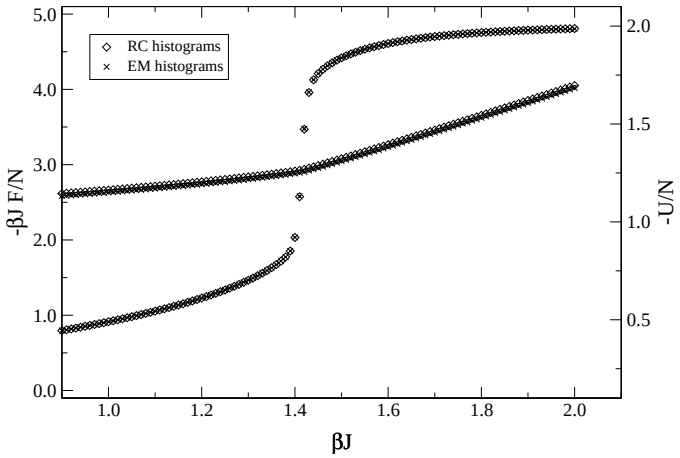

FIG. 6: Internal energy and free energy of the $q = 10$ Potts model on a $\mathcal{N} = 16^2$ square lattice as computed by the adaptively normalized RC multi-histograming scheme according to Eqs. (29) and (30). The results from the EM multi-histogram analysis of the same data are shown for comparison.

simulations at (p_i, q_i) and define reduced free energies $\mathcal{F}_i$ as

$$F_i = -\frac{1}{K_i} \ln Z_{p_i,q_i}(\mathcal{G}) = -\mathcal{E} - \frac{\mathcal{F}_i}{K_i}. \tag{27}$$

From Eq. (7) estimates of the density $g(b, n)$ from the single histograms $\hat{H}_{p_i,q_i}(b, n) \equiv \hat{H}^{(i)}(b, n)$ are given by:

$$\hat{g}^{(i)}(b, n) = e^{\mathcal{F}_i} \frac{\hat{H}^{(i)}(b, n)}{p_i^b \, (1 - p_i)^{\mathcal{E}-b} \, q_i^n \, N_i}. \tag{28}$$

Once again combining these estimates in a variance-optimized way as above in Section II, treating the $\mathcal{F}_i$ as parameters with zero variance, we arrive at the following expression

$$\hat{g}(b, n) = \frac{\sum_i N_i \, e^{-\mathcal{F}_i} \, p_i^b \, (1 - p_i)^{\mathcal{E}-b} \, q_i^n}{\sum_j N_j^2 \, e^{-2\mathcal{F}_j} \, p_j^{2b} \, (1 - p_j)^{2(\mathcal{E}-b)} \, q_j^{2n} [\hat{H}^{(i)}(b, n)]^{-1}}. \tag{29}$$

Now, from this estimate one has the following *a posteriori* relation for computation of the parameters $\mathcal{F}_i$:

$$e^{\mathcal{F}_i} = \sum_{b=0}^{\mathcal{E}} \sum_{n=1}^{\mathcal{N}} \hat{g}(b, n) \, p_i^b \, (1 - p_i)^{\mathcal{E}-b} \, q_i^n. \tag{30}$$

Eqs. (29) and (30) form a pair of equations to be solved self-consistently for the determination of the parameters $\mathcal{F}_i$, which can be straightforwardly iterated by plugging in the results for $\mathcal{F}_i$ from Eq. (29) into Eq. (30) and vice versa. One can improve on that by applying more sophisticated iteration schemes like, e.g., the Newton-Raphson iteration [21]. We find, however, that the radius of convergence of this method is quite small; therefore, we adaptively revert to the simple iteration if the procedure leaves the Newton-Raphson convergence region. It is obvious that for the iteration to converge, one needs some overlap between the $\hat{H}(b, n)$ histograms between "adjacent" simulations, i.e., at least pairwise overlap. Apart from this restriction, however, we find this iterative scheme to be very well-behaved, converging rapidly in every case that fulfils the overlap-condition.

To get started, we use first-guess values of the $\mathcal{F}_i$ from thermodynamic integration. Assume that the simulation points $(i) = (P_i, q_i)$ are ordered such that the histograms of (i) and $(i + 1)$ have reasonable overlap; then

$$e^{\mathcal{F}_i} = \sum_{b=0}^{\mathcal{E}} \sum_{n=1}^{\mathcal{N}} \frac{\hat{H}^{(i-1)}(b, n)}{N_{i-1}} \, \frac{p_i^b \, (1 - p_i)^{\mathcal{E}-b} \, q_i^n}{p_{i-1}^b \, (1 - p_{i-1})^{\mathcal{E}-b} \, q_{i-1}^n} \, e^{\mathcal{F}_{i-1}} \tag{31}$$

is a good starting point for the described iteration scheme. $\mathcal{F}_1$ can be chosen arbitrarily, since the given pair of equations is obviously invariant under a global shift $\mathcal{F}_i \to \mathcal{F}_i - \mathcal{F}_1$. Thus, we have determined the final estimate $\hat{g}(b, n)$ only up to a global factor. To fix this last normalization we propose two different possibilities; on the one hand, we can use the free model limit, i.e., evaluate $\mathcal{F}_{p=0,q}$ from Eq. (30) and use Eq. (25) for any q:

$$\mathcal{F}_{p=0,q} = \ln Z_{p=0,q}(\mathcal{G}) - K\mathcal{E} = \mathcal{N} \ln q. \tag{32}$$

On the other hand, also the $q = 1$ partition function is trivial,

$$\mathcal{F}_{p,q=1} = \ln \left[\sum_{b=0}^{\mathcal{E}} \binom{\mathcal{E}}{b} p^b \, (1 - p)^{\mathcal{E}-b} \right] = 0, \tag{33}$$

and can serve as normalization point for arbitrary p. In practice the best choice depends on the set of simulated couplings (p_i, q_i): for large-q simulations one might want to resort to Eq. (32), while otherwise Eq. (33) should be the better choice.

Now, we can re-consider the internal energy of the $q = 10$ case from above with the new, adaptively normalized RC multi-histograming scheme. Figure 6 shows internal energy and free energy from this analysis as compared to the EM multi-histogram approach. As far as the error estimates are concerned, we apply the jackknife process to the whole iteration run, i.e., the iteration scheme for fixing the weights $\mathcal{F}_i$ is done for each jackknife block of data separately, taking full account of statistical errors. Clearly, now the results from both approaches perfectly agree, the deviations of Fig. 4 have vanished. As anticipated in Sec. II, also the cluster estimator for the internal energy performs noticeably better than in the $q = 2$ case, such that — at least in the critical region — it is quite comparable in precision to the EM estimator, cp. Fig. 7.

For the free energy it is obvious that with the adaptive normalization scheme of RC histograms we lose the especially high precision throughout the whole K region obtained by the application of the sum rule (8) in Fig. 1 for the Ising model. To amend this, having fixed the relative normalization of the single histograms adaptively,

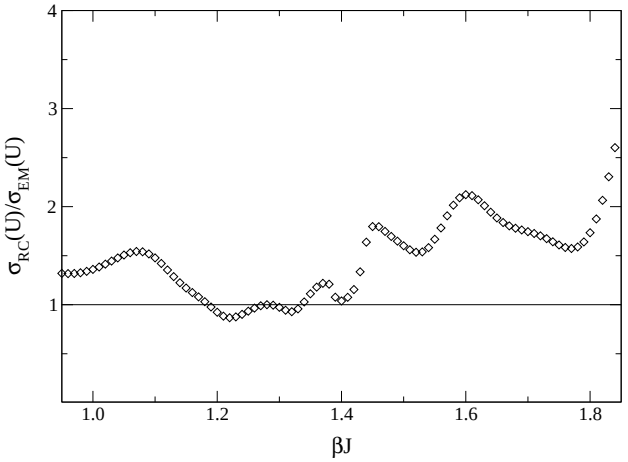

FIG. 7: Ratio of standard deviations for estimates of the internal energy U of the two-dimensional $q = 10$ Potts model on a $\mathcal{N} = 16^2$ lattice and RC and EM multi-histogram analyses as a function of the coupling $K = \beta J$. The variances are estimated by a "jackknife" time-series analysis [12]. The RC histograms are normalized according to Eqs. (29) and (30).

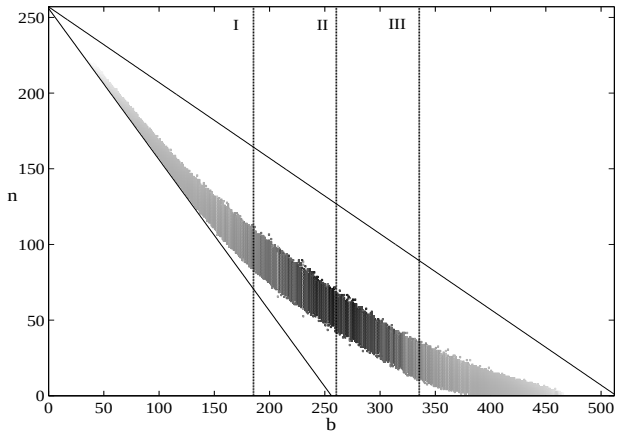

FIG. 8: Density plot of the relative differences $[\hat{g}^{(\mathrm{abs})} - \hat{g}^{(\mathrm{rel})}]/\hat{g}^{(\mathrm{rel})}$ of the density of states as sampled from the $q = 10$ Potts model on a $\mathcal{N} = 16^2$ lattice by the absolutely normalized RC histograming scheme of Eq. (14) ($\hat{g}^{(\mathrm{abs})}$) and the adaptively normalized scheme of Eq. (29) ($\hat{g}^{(\mathrm{rel})}$), respectively. Dark shading indicates that $\hat{g}^{(\mathrm{abs})}(b,n) > \hat{g}^{(\mathrm{rel})}(b,n)$ and vice versa.

one might consider applying the sum rule (8) to the final result $\hat{g}(b,n)$ instead of using the normalizations Eq. (33) or Eq. (32). This, however, gives results looking almost identical to those shown above in Fig. 3, i.e., the large deviations reappear, which clearly reveals the source they are resulting from.

IV. COMPARISON OF THE METHODS

The effect of this normalization problem should also be clearly seen in the final estimates for the density of states $g(b,n)$ from the two RC histograming methods. In Fig. 8 we show a density plot of the relative differences of the estimated density of states $\hat{g}(b,n)$ from the absolutely normalized histograming scheme of Eq. (14), $\hat{g}^{(\mathrm{abs})}(b,n)$, and of the adaptively normalized scheme of Eqs. (29) and (30), $\hat{g}^{(\mathrm{rel})}(b,n)$, i.e., the quantity

$$\hat{\Delta}(b,n) \equiv \frac{\hat{g}^{(\mathrm{abs})}(b,n) - \hat{g}^{(\mathrm{rel})}(b,n)}{\hat{g}^{(\mathrm{rel})}(b,n)}. \tag{34}$$

Note that the range of possible value pairs (b,n) is restricted by two simple bounds in the (b,n) plane. First, starting from the point $(b = 0, n = \mathcal{N})$ each added bond can at most reduce the number of clusters by one, namely by connecting two previously unconnected clusters, i.e., one has

$$n \geq \mathcal{N} - b. \tag{35}$$

On the other hand, starting from the "opposite" point $(b = \mathcal{E}, n = 1)$ one has $b\mathcal{N}/\mathcal{E}$ bonds per site, so that for producing a new cluster one must at least remove $\mathcal{N}/\mathcal{E}$ bonds,

$$n - 1 \leq \frac{\mathcal{N}}{\mathcal{E}}(\mathcal{E} - b), \tag{36}$$

or, for the square lattice,

$$n \leq \mathcal{N} - \frac{b}{2} + 1. \tag{37}$$

Apart from single points near those bounds, all configurations within this triangle can actually appear in a Potts model simulation with non-vanishing probability.

Now, from Fig. 8 it is obvious, given that the estimate $\hat{g}^{(\mathrm{rel})}(b,n)$ is correct up to an overall factor, that the absolutely normalized histograming estimate $\hat{g}^{(\mathrm{abs})}(b,n)$ gives too large estimates for b values near the centre $b = \mathcal{N}$ as compared to the other regions of b (dark shading in Fig. 8). Then, considering again the deviation in internal energy shown in Fig. 3, its origin becomes clear: the histogram $\hat{H}_{p,q}(b,n)$ for a simulation somewhat below the transition point will be centred around the line $b = b_{\mathrm{I}}$ in Fig. 8; then, using the density of states estimate $\hat{g}^{(\mathrm{abs})}(b,n)$ for evaluating U, the parts of the histogram lying to the right of $b = b_{\mathrm{I}}$ will have too large weight as compared to the values $b < b_{\mathrm{I}}$, thus by Eq. (15) resulting in a too large estimate for the internal energy U. On the other hand, for couplings above the transition point the histogram will be centred around $b = b_{\mathrm{III}}$ with too large weights for $b < b_{\mathrm{III}}$, leading to estimates for U that are too low. Directly in the vicinity of the transition point, deviations in normalization are symmetric with respect to the histogram, which will be centred around $b = b_{\mathrm{II}}$, thus leading to an unbiased estimate for U. This is exactly the behavior found in Fig. 3. Finally, contemplating on the reason for the deviations in normalization

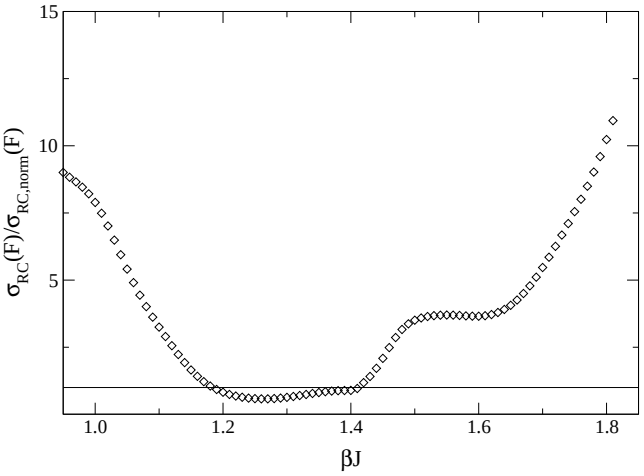

FIG. 9: Ratio of standard deviations of estimates of the free energy from the adaptively normalized RC histograming scheme of Eq. (29) with ($\sigma_{\mathrm{RC,norm}}$) and without ($\sigma_{\mathrm{RC}}$) a final application of the sum rule Eq. (8) to the density of states after determining the weights. The used time series data includes both, the $q = 2$ and the $q = 10$ simulations reported above.

shown in Fig. 8 in the first place, it becomes obvious that they have the same origin as those shown in Fig. 5. The exponential factor q^{-n} from the sum rule Eq. (9) attaches large weight to the configurations with small number of clusters n; if, however, histograms miss entries for small n, as is the case for histograms in the transition region $b \approx \mathcal{N}$ of Fig. 8, the sum $\sum_n \hat{H}_{p,q}(b,n)\, q^{-n}$ will become too small, resulting in too large normalization factors $\hat{C}_{p,q}(b)$.

Thus, for the application of the sum rule (8) to the final result from the adaptively normalized RC histograming scheme to work reliably, one always has to include histograms from small-q simulations, like percolation ($q \to 1$) or the Ising model ($q = 2$), that produce configurations with relatively small numbers of clusters n. To illustrate this, we combined the data from the $q = 2$ and $q = 10$ simulations reported above, used the histograming scheme Eqs. (29) and (30) to get results for $q = 10$ and applied the sum rule Eq. (8) afterwards, i.e., the normalization of the single histograms was found adaptively, whereas the total histogram was normalized by the sum rule (8); this yields results for the internal and free energies indistinguishable from those of the pure $q = 10$ results of Fig. 6. For the free energy, however, the size of statistical errors is largely affected by the final normalization, cf. Fig. 9. For most of the couplings shown, the estimate from the finally sum-rule-normalized density of states is up to about 10 times more accurate in terms of the statistical errors. The presence and size of such a gain for a given coupling is not mainly physically motivated, but rather depends on the relation of the simulation points (q_i, p_i) to the points of data analysis.

V. CONCLUSIONS

We have considered multi-histogram data analyses of time series from cluster-update Monte Carlo simulations of the q-state Potts model in the random-cluster language. Generalizing the original formulation of Ref. [8] to the case of simulations of different number of states q, we found the original ansatz of absolutely normalizing the individual histograms with a geometrical sum rule for finite-length time series to produce large deviations from the expected behavior when applied to cases q larger then about 3 or 4 in two dimensions. We track this error down to a mismatch between exponential suppression of a part of the state space (b, n) and a simultaneous exponential enhancement of this region in the sum rule Eq. (9). To circumvent this problem, we propose a different ansatz, normalizing the histograms adaptively via a set of self-consistency equations aiming at the minimization of the variance of the final estimate of the density of states $g(b, n)$. Absolute normalization over the whole temperature region can still be maintained by making contact with the trivial partition function of the percolation limit $q \to 1$ or by combining large- and small-q data and applying the sum rule (8) after the adaptive normalization. This new approach does not exhibit the limitations of the absolutely normalized ansatz to small-q simulations.

Comparing the newly introduced, adaptively normalized random-cluster "RC" multi-histogram technique with multi-histograming in the energy/magnetization "EM" language, we can make the following statements: (a) The cluster variables (b, n) form the natural state space for the analysis of the Potts model. Using the Swendsen-Wang cluster-update algorithm, these numbers are automatically known as a by-product of the update-steps; no additional measurement steps are needed. (b) The RC representation allows for reweighting in both parameters, the thermal coupling p resp. K, and the number of states q, without systematical errors as in the partial transformation of Ref. [22]. Especially, the model can be considered for the case of non-integer q. It is easy to combine data from different-q simulations to enhance the accuracy for large q. (c) Cluster estimators occur naturally in the RC language. Although we found that short-distance observables like the internal energy and specific heat are sampled systematically less acurate by cluster estimators, this situation is reversed for observables sensitive to long-range order like the magnetization, susceptibility and correlation functions. Also, even short-range cluster estimators perform comparable to EM language estimators for larger q values, at least in the transition region. In the RC language, the magnetic observables can be defined consistently throughout the broken and unbroken phases, cf. Appendix A.

Apart from that, the combination of data from small and large q models can serve as a new method to cope with the super-critical slowing down at the first-order transitions for large q: for sufficiently large lattices sim-

ulation runs will entirely stay in one of the pure phases depending on the initial configuration and boundary conditions due to ergodicity breaking at the transition point. However, combining such data with smaller-q simulations from the second-order or weak first-order regime allows the adaptive normalization scheme to still find the correct normalization of the pure-phase histograms without real tunneling events. This approach is similar in spirit to the simulated tempering technique [23, 24].

One might also think of applying the multi-canonical [18] resp. multi-bondic [17] simulation approach or one of the related techniques to the sampling of the density of states $g(b,n)$. Especially, application of the absolute normalization Eq. (8) to this case might be of interest. This approach is currently under investigation. However, sampling the complete range of possible values in the (b,n) plane with sufficient accuracy is found to be a computationally very demanding problem. In contrast, in the current approach, we still stick to the physically sensible approach of importance-sampling, i.e., sampling the phase-space according to the local canonical weights. Furthermore, we are able to take full advantage of the computational gain of cluster algorithms, whereas the generalized-ensemble algorithms put forward so far employ local updates (apart from the multi-bondic algorithm of Ref. [17]).

As an interesting application of our ansatz we suggest the analysis of the tri-critical point q_c, where the order of the thermal transitions changes from second to first order, in three dimensions. There has been quite some debate about the location of this point, estimates ranging from $q_c = 2.15$ [25] to $q_c = 2.6$ [26]. Furthermore, the universality class, critical exponents etc. of this transition have not yet been properly analyzed. A test for the $q_c = 4$ case in two dimensions shows that our method is well suited for such an analysis. This problem will be considered in a forthcoming publication.

ACKNOWLEDGMENTS

The collaboration of the authors has been supported by DAAD-NSC grant No. D/9827248. MW acknowledges support by the DFG through the "Graduiertenkolleg QFT".

APPENDIX A: CLUSTER ESTIMATORS

Consider the Potts model coupled to an external magnetic field H with Hamiltonian [27]

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} \delta(\sigma_i, \sigma_j) - H \sum_i \delta(\sigma_i, 1). \qquad \text{(A1)}$$

Then, the random-cluster representation of the partition function on a graph $\mathcal{G}$ consisting of $\mathcal{N}$ sites and $\mathcal{E}$ edges (bonds) is given by

$$\begin{aligned} Z_{p,q}(\mathcal{G}, B) &= \sum_{\mathcal{G}' \subseteq \mathcal{G}} (e^K - 1)^{b(\mathcal{G}')} \prod_c [(q-1) + e^{Bn_c}] \\ &= e^{K\mathcal{E}} \sum_{\mathcal{G}' \subseteq \mathcal{G}} p^{b(\mathcal{G}')} (1-p)^{\mathcal{E} - b(\mathcal{G}')} \\ &\quad \times \prod_c [(q-1) + e^{Bn_c}], \end{aligned} \qquad \text{(A2)}$$

where the product runs over the set of clusters $\{c\}$ of the subgraph $\mathcal{G}'$, n_c is the number of sites in cluster c, $K = \beta J$ is the thermal coupling parameter and $B = \beta H$ denotes the reduced magnetic field. $p = 1 - e^{-K}$ is the probability for the activation of bonds.

The zero-field, per-site internal energy u is then given by

$$\begin{aligned} u &= -\frac{\partial}{\partial K} \left[\frac{\ln Z_{p,q}(\mathcal{G}, B=0)}{\mathcal{N}} \right] \\ &= -\frac{1}{\mathcal{N}} \left\langle \frac{\frac{\partial}{\partial K}[(e^K-1)^b]}{(e^K-1)^b} \right\rangle = -\frac{1}{p} \left\langle \frac{b}{\mathcal{N}} \right\rangle, \end{aligned} \qquad \text{(A3)}$$

which shows the close connection between the b and E distributions.

The zero-field specific heat c_v follows from

$$\begin{aligned} c_v &= K^2 \frac{\partial^2}{\partial K^2} \left[\frac{\ln Z_{p,q}(\mathcal{G}, B=0)}{\mathcal{N}} \right] = \frac{K^2}{\mathcal{N}} \left[-\frac{1}{p^2} \langle b \rangle^2 + \left\langle \frac{\frac{\partial^2}{\partial K^2}[(e^K-1)^b]}{(e^K-1)^b} \right\rangle \right] \\ &= \frac{K^2}{p^2 \mathcal{N}} \left[\langle b^2 \rangle - \langle b \rangle^2 - (1-p)\langle b \rangle \right]. \end{aligned} \qquad \text{(A4)}$$

In the thermodynamic limit, the per-site, zero-field "magnetization" $\tilde{m} = \langle \sum_i \delta(\sigma_i, 1)/\mathcal{N} \rangle$ is given by

$$\tilde{m} = \lim_{B \to 0+} \lim_{\mathcal{N} \to \infty} \frac{\partial}{\partial B} \left[\frac{\ln Z_{p,q}(\mathcal{G}, B)}{\mathcal{N}} \right]$$

$$= \lim_{B\to 0+}\lim_{\mathcal{N}\to\infty} \frac{e^{K\mathcal{E}}}{Z_{p,q}(\mathcal{G},B)} \sum_{\mathcal{G}'\subseteq\mathcal{G}} p^{b(\mathcal{G}')}(1-p)^{\mathcal{E}-b(\mathcal{G}')} \prod_c [(q-1)+e^{Bn_c}] \sum_{c'} \frac{n'_c}{\mathcal{N}} \frac{e^{Bn'_c}}{(q-1)+e^{Bn'_c}}$$

$$= \lim_{B\to 0+}\lim_{\mathcal{N}\to\infty} \left[\left\langle \sum_{c^\pi} \frac{n_c^\pi}{\mathcal{N}} \frac{1}{(q-1)e^{-Bn_c^\pi}+1} \right\rangle + \left\langle \sum_{c^\phi} \frac{n_c^\phi}{\mathcal{N}} \frac{1}{(q-1)e^{-Bn_c^\phi}+1} \right\rangle \right]. \quad \text{(A5)}$$

Here, we split the cluster contributions of the subgraph into percolating clusters c^π and non-percolating, finite clusters c^ϕ. In the indicated order of taking the limits, first $\mathcal{N}\to\infty$ and then $B\to 0$, the factors $\exp(-Bn_c)$ take the values 0 and 1 for percolating and non-percolating clusters c, respectively. Thus, we arrive at

$$\tilde{m} = \left\langle \sum_{c^\pi} \frac{n_c^\pi}{\mathcal{N}} \right\rangle + \frac{1}{q} \left\langle \sum_{c^\phi} \frac{n_c^\phi}{\mathcal{N}} \right\rangle = \frac{q-1}{q} \left\langle \sum_{c^\pi} \frac{n_c^\pi}{\mathcal{N}} \right\rangle + \frac{1}{q}, \quad \text{(A6)}$$

which explicitly reflects the symmetry-breaking nature of the percolating configurations. For the order parameter m, which varies between 0 for the completely disordered state and 1 for the ground states, we find

$$m \equiv \frac{q\tilde{m}-1}{q-1} = \left\langle \sum_{c^\pi} \frac{n_c^\pi}{\mathcal{N}} \right\rangle. \quad \text{(A7)}$$

For finite lattices one can retain this definition since the notion of percolating and non-percolating clusters is still well-defined. Note, that this gives a consistent definition of the order parameter throughout the disordered and broken phases. In contrast, in the EM language one has to explicitly break symmetry in the low-temperature phase, which is usually done by defining

$$\tilde{m}_{K>K_t} = \left\langle \max_{1\le j\le q} \sum_i \delta(\sigma_i, j) \right\rangle, \quad \text{(A8)}$$

whereas for the unbroken phase one uses

$$\tilde{m}_{K\le K_t} = \left\langle \sum_i \delta(\sigma_i, 1) \right\rangle. \quad \text{(A9)}$$

Obviously, for finite lattices, the expectation values of the RC and EM definitions will not coincide exactly; critical exponents, however, will of course agree.

The zero-field susceptibility $\tilde{\chi}$ is given by

$$\tilde{\chi} = \lim_{B\to 0+}\lim_{\mathcal{N}\to\infty} \frac{\partial^2}{\partial B^2}\left(\frac{\ln Z(B)}{\mathcal{N}}\right) = \lim_{B\to 0+}\lim_{\mathcal{N}\to\infty} \left[-\frac{1}{\mathcal{N}Z(B)^2}\left(\frac{\partial Z(B)}{\partial B}\right)^2 + \frac{1}{\mathcal{N}Z(,B)}\frac{\partial^2 Z(B)}{\partial B^2} \right]$$

$$= -\mathcal{N}m^2 + \lim_{B\to 0+}\lim_{\mathcal{N}\to\infty} \left[\left\langle \sum_c \frac{n_c^2}{\mathcal{N}} \frac{1}{(q-1)e^{-Bn_c}+1} \right\rangle + \mathcal{N} \left\langle \left(\sum_c \frac{n_c}{\mathcal{N}} \frac{1}{(q-1)e^{-Bn_c}+1} \right)^2 \right\rangle + \left\langle \sum_c \frac{n_c^2}{\mathcal{N}} \frac{1}{[(q-1)e^{-Bn_c}+1]^2} \right\rangle \right]$$

$$= \mathcal{N}\left(\frac{q-1}{q}\right)^2 \left[\left\langle \left(\sum_{c^\pi} \frac{n_c^\pi}{\mathcal{N}} \right)^2 \right\rangle - \left\langle \sum_{c^\pi} \frac{n_c^\pi}{\mathcal{N}} \right\rangle^2 \right] + \frac{q-1}{q^2} \left\langle \sum_{c^\phi} \frac{{n_c^\phi}^2}{\mathcal{N}} \right\rangle. \quad \text{(A10)}$$

From Eq. (A10) one recognizes the widely used improved cluster estimator for the high-temperature phase, namely the last term. Note, however, that the original improved estimator includes *all* clusters here instead of only the non-percolating ones, which makes a difference for finite lattices. For finite lattices, once again, from Eq. (A10) we have a single definition for both, the unbroken and broken phases.

Alternatively defining the susceptibility corresponding to the order parameter m we get

$$\chi = \left(\frac{q}{q-1}\right)^2 \tilde{\chi} = \frac{1}{q-1} \left\langle \sum_{c^\phi} \frac{{n_c^\phi}^2}{\mathcal{N}} \right\rangle + \mathcal{N} \left[\left\langle \left(\sum_{c^\pi} \frac{n_c^\pi}{\mathcal{N}} \right)^2 \right\rangle - \left\langle \sum_{c^\pi} \frac{n_c^\pi}{\mathcal{N}} \right\rangle^2 \right]. \quad \text{(A11)}$$

REFERENCES

[1] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1988); Phys. Rev. Lett. **63**, 1658(E) (1989).

[2] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **63**, 1195 (1989).

[3] C. M. Fortuin and P. W. Kasteleyn, Physica **57**, 536 (1972).

[4] A. Coniglio and W. Klein, J. Phys. A **13**, 2775 (1980).

[5] C.-K. Hu, Phys. Rev. B **29**, 5103 and 5109 (1984); J. Phys. A **19**, 3067 (1986).

[6] U. Wolff, Nucl. Phys. B **300**, 501 (1988).

[7] C.-K. Hu, Phys. Rev. Lett. **69**, 2739 (1992).

[8] J.-A. Chen and C.-K. Hu, Phys. Rev. B **50**, 6260 (1994).

[9] R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **58**, 86 (1987).

[10] C.-K. Hu, Phys. Rev. B **46**, 6592 (1992).

[11] A. E. Ferdinand and M. E. Fisher, Phys. Rev. **185**, 832 (1969).

[12] B. Efron, *The Jackknife, the Bootstrap and Other Resampling Plans* (Society for Industrial and Applied Mathematics [SIAM], Philadelphia, 1982).

[13] L. Onsager, Phys. Rev. **65**, 117 (1944).

[14] F. Y. Wu, Rev. Mod. Phys. **54**, 235 (1982).

[15] B. Kaufman, Phys. Rev. **76**, 1232 (1949).

[16] U. Wolff, Nucl. Phys. B **334**, 581 (1990).

[17] W. Janke and S. Kappler, Phys. Rev. Lett. **74**, 212 (1995).

[18] B. A. Berg and T. Neuhaus, Phys. Rev. Lett. **68**, 9 (1992).

[19] W. Janke and S. Kappler, J. Phys. I (France) **7**, 663 (1997).

[20] P. D. Beale, Phys. Rev. Lett. **76**, 78 (1996).

[21] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C — The Art of Scientific Computing* (Cambridge University Press, Cambridge, 1992), 2nd ed.

[22] J. Lee and J. M. Kosterlitz, Phys. Rev. B **43**, 1268 (1991).

[23] E. Marinari and G. Parisi, Europhys. Lett. **19**, 451 (1992).

[24] W. Kerler and A. Weber, Phys. Rev. B **47**, 11 563 (1993).

[25] S. Grollau, M. L. Rosinberg, and G. Tarjus, cond-mat/0011483 preprint (2000).

[26] B. Nienhuis, E. K. Riedel, and M. Schick, Phys. Rev. B **23**, 6055 (1981).

[27] Note that in Ref. [5] a different coupling between magnetic field and spins is employed.