# Hybrid algorithm for Metropolis simulations at low temperatures: Specific heat of the Coulomb glass

A. Möbius
*Institute for Solid State and Materials Research Dresden, D-01171 Dresden, Germany*

P. Thomas
*Department of Physics and Materials Sciences Center, Philipps University, Renthof 5, D-35032 Marburg, Germany*

(Received 28 June 1996)

We present an algorithm which suppresses the problem that correlation times increase rapidly at low temperatures $T$ to a large extent. It bridges the gap between Metropolis simulation and analytical statistical mechanics: The set of system states is decomposed into low- and high-energy subsets, and very fast transitions within the low-energy subset are introduced guaranteeing immediate thermalization therein. Thus, even if there is an appreciable number of local minima, equilibrium properties can be studied for $T$ decreasing continuously from high down to arbitrarily low values. The efficiency of this algorithm is demonstrated for the Coulomb glass. [S0163-1829(97)06809-4]

The Metropolis algorithm[1] has proved to be an efficient tool for the simulation of various systems.[2] But diverging correlation times restrict its applicability at low $T$ and close to phase transitions. As $T\rightarrow 0$, the portion of unsuccessful attempts rapidly tends to 100%; in addition, the number of transitions required to escape a local minimum diverges. Several methods have been proposed to treat the problems of large correlation times, e.g., $n$ fold way algorithm,[3,4] cluster algorithms,[5,6] a cluster version of the $n$ fold way algorithm,[7] multicanonical algorithms,[8] simulated tempering,[9] and, in the context of simulated annealing, genetic algorithms.[10] Here, we present a hybrid procedure for low $T$ based on dividing the configuration space into a low-energy subset, treated by directly performing the appropriate summations over all states included, and a high-energy subset studied by Monte Carlo techniques; for a short preliminary version of this work see Ref. 11. First, we describe the algorithm in a general formulation. Then, we illustrate it by studying temperature and disorder dependences of the specific heat of the three-dimensional Coulomb glass.

We consider an arbitrary system with a discrete spectrum of system states $i$ having energies $E[i]$. Let the transition probability per unit time from state $i$ to state $j$ be $w[i,j]$. The states which can be reached within one primitive step starting from $i$ are determined by the dynamics chosen for the simulation. We take $N$ such states, $j_1^i, j_2^i, \ldots, j_N^i$, into account, where $N$ does not depend on $i$. Thus $w[i,j]=0$ for all states $j$ not belonging to this set; some of the transitions $i\rightarrow j_\nu^i$ might be forbidden too. As in many Metropolis simulations, devoted solely to the study of equilibrium properties, the $w[i,j]$ are assumed to depend only on the energy change $\Delta[i,j]=E[j]-E[i]$,

$$w[i,j]=w^{(0)}[i,j]\begin{cases}1 & \text{if } \Delta[i,j]\leqslant 0,\\ \exp\{-\Delta[i,j]/T\} & \text{if } \Delta[i,j]>0,\end{cases} \tag{1}$$

where $w^{(0)}[i,j]=w^{(0)}[j,i]$ equals either $1/N$ or zero. The Boltzmann constant is taken to be 1 throughout.

The $w[i,j]$ define the dynamics of the simulation. Note that, if solely mean values of equilibrium observables are to be obtained, the dynamics can be modified but detailed balance must be maintained. Thus we introduce additional transitions within a subset of $M$ local minima, or, in general, of low-energy states, $\mathcal{S}^{(l)}=\{i_1^{(l)}, i_2^{(l)}, \ldots, i_M^{(l)}\}$. Previous attempts in this direction are, e.g., $J$-walker algorithm,[12] and simulated tunneling.[13] We ascribe to the transitions within $\mathcal{S}^{(l)}$ rates which are very large compared to the $w[i,j]$ introduced in the above paragraph, see Ref. 7. This simplifies the numerical problem enormously: Thermalization within $\mathcal{S}^{(l)}$ happens immediately after anyone of its states is reached; as long as $\mathcal{S}^{(l)}$ is not left, all equilibrium observables can be assumed to have the mean value related to $\mathcal{S}^{(l)}$.

The decisive point is the construction of $\mathcal{S}^{(l)}$. We begin with an initial $\mathcal{S}^{(l)}$ obtained by sampling local minima distributed in the configuration space. Therefore, the system is repeatedly quenched starting from randomly chosen states. Then, we iterate $\mathcal{S}^{(l)}$ modifying the idea of Ref. 14, viz., to utilize the Metropolis algorithm to obtain low-energy states: We start at a sufficiently high $T$, and diminish $T$ in the course of the simulation stepwise. Before starting the integration over a given time interval, we adapt the dynamics, i.e., we introduce very fast transitions connecting all states $\in\mathcal{S}^{(l)}$ with each other. During the integration, the lowest-energy states ''visited'' are stored in some ''temporary'' memory, $\mathcal{S}_{\text{temp}}^{(l)}$. After the end of the interval has been reached, $\mathcal{S}^{(l)}$ is updated by choosing $M$ lowest-energy states[16] out of $\mathcal{S}^{(l)}\cup\mathcal{S}_{\text{temp}}^{(l)}$.

The system stays within $\mathcal{S}^{(l)}$ only during a fraction of the simulation time. This fraction vanishes exponentially at high $T$, it increases with decreasing $T$, and finally tends to 100% as $T\rightarrow 0$. At sufficiently low $T$, the consideration of only $\mathcal{S}^{(l)}$ yields good approximations of the mean values of equilibrium observables. Thus our approach provides a smooth transition between ''standard'' Metropolis algorithm at high $T$, and simple summations at low $T$. It differs substantially from Ref. 7, where $\mathcal{S}^{(l)}$, updated in each step, in-

 

cludes a complete history. Moreover, two further efficient improvements of Ref. 7 should be noted: (a) The energy threshold criterion in updating $\mathcal{S}^{(l)}$ provides a natural mechanism of gradual forgetting. (b) The overhead for handling $\mathcal{S}^{(l)}$ is far lower here, we have to update this subset only a few times per $T$ point.

The usual Metropolis simulation is based on two ideas: (i) substituting temporal for ensemble averaging, and (ii) coarse graining the time in order to simplify the decisions to be performed in simulating the stochastic process. However, as $T\to 0$, (ii) proves to be ineffective: The number of possible choices remains finite whereas the number of unsuccessful attempts per successful one diverges. Therefore, abandoning (ii), we ''directly'' determine the dwelling time at the actual state, considering the time as a continuous variable. Such an approach was published first in the form of the $n$ fold way algorithm;[3] for surveys and the generalization to clusters of states see Refs. 4, 15, and 7, respectively.

Each elementary step starts with classifying the actual state $i$ according to whether or not it belongs to $\mathcal{S}^{(l)}$. Consider first $i\notin\mathcal{S}^{(l)}$: The dwelling time is an exponentially distributed random variable with the mean value

$$\tau[i]=1\Big/\sum_{\nu=1}^{N} w[i,j_\nu^i]. \tag{2}$$

In order to perform the random decision for the next state $j_\kappa^i$ the cumulated probabilities

$$s_\kappa[i]=\sum_{\nu=1}^{\kappa} w[i,j_\nu^i] \tag{3}$$

are calculated. A random number $\xi$, equally distributed in $\{0,1\}$, is chosen, and that $\kappa$ is determined for which

$$s_{\kappa-1}[i]\leqslant\xi s_N[i]<s_\kappa[i] \tag{4}$$

holds (where $s_0[i]=0$). Thus, the mean value of an observable $O$ is calculated integrating over $O[i(t)]$.

Assume now $i\in\mathcal{S}^{(l)}$. The immediate thermalization within $\mathcal{S}^{(l)}$ permits us to integrate directly until leaving this subset: All equilibrium observables $O$ can be assumed to have the related mean value,

$$\langle O\rangle^{(l)}=\sum_{\mu=1}^{M} p_\mu^{(l)} O[i_\mu^{(l)}], \tag{5}$$

where

$$p_\lambda^{(l)}=\exp\{-E[i_\lambda^{(l)}]/T\}\Big/\sum_{\mu=1}^{M}\exp\{-E[i_\mu^{(l)}]/T\}. \tag{6}$$

The time, during which the system remains trapped within $\mathcal{S}^{(l)}$, is an exponentially distributed random variable too. Its mean value is given by

$$\tau^{(l)}=1\Big/\sum_{\mu=1}^{M} p_\mu^{(l)}\sum_{\nu=1}^{N} w[i_\mu^{(l)},j_\nu^{i_\mu^{(l)}}]\Theta[j_\nu^{i_\mu^{(l)}}], \tag{7}$$

where $\Theta[j]=0$ if $j\in\mathcal{S}^{(l)}$ and $\Theta[j]=1$ otherwise. The decision for the next state after leaving $\mathcal{S}^{(l)}$ is performed along similar lines as above. First, cumulated probabilities are calculated,

$$s_{\lambda\kappa}^{(l)}=\sum_{\mu=1}^{\lambda-1} p_\mu^{(l)}\sum_{\nu=1}^{N} w[i_\mu^{(l)},j_\nu^{i_\mu^{(l)}}]\Theta[j_\nu^{i_\mu^{(l)}}] + p_\lambda^{(l)}\sum_{\nu=1}^{\kappa} w[i_\lambda^{(l)}j_\nu^{i_\lambda^{(l)}}]\Theta[j_\nu^{i_\lambda^{(l)}}]. \tag{8}$$

Then the final state is determined. It is advantageous to perform this choice within two steps, first deciding which initial state $i_\mu^{(l)}$ is now to be left, and then choosing the final state $\notin\mathcal{S}^{(l)}$. Note, that the full set of $s_{\lambda\kappa}^{(l)}$ has to be calculated only once after updating $\mathcal{S}^{(l)}$, or changing $T$.

Completing the general part of this paper, we discuss the limitations of our algorithm. Due to modifying the transition rates, this approach is not appropriate for a direct study of dynamical properties. Moreover, as generally in such approaches, there is no guarantee that all relevant low-energy states are found, only for sufficiently small systems the related probability is close to 1. Therefore, concerning the system size, there is always a bound limiting the applicability region — although this bound can be surprisingly large as demonstrated below. For larger systems, the simulation usually fails for one of the following reasons: (i) all initial archive states can be separated ''from the most important valleys by too high barriers,'' or (ii) the simulation can get stuck around some local minimum $\notin\mathcal{S}^{(l)}$. This applicability bound depends on the model studied, as well as on the strength of the corresponding disorder, and on the integration time. However, determining the latter by means of an automatic accuracy control, see below, increases considerably the efficiency, and enlarges therefore the applicability region.

The advantages of the algorithm presented in the first part of this paper are now demonstrated studying the specific heat of the Coulomb glass. This semiclassical model describes a diluted impurity band, i.e., a disordered system of interacting localized states with negligible quantum tunneling between them. For related reviews see Refs. 17 and 18. Here we focus on a half-filled impurity band. It is represented by the Hamiltonian

$$H=\sum_\alpha \varepsilon_\alpha n_\alpha+\frac{1}{2}\sum_{\alpha\neq\beta}\frac{(n_\alpha-1/2)(n_\beta-1/2)}{|\vec{x}_\alpha-\vec{x}_\beta|}, \tag{9}$$

where $n_\alpha\in\{0,1\}$ denote the occupation numbers of states localized at sites $\vec{x}_\alpha$. We assume these sites to form a regular three-dimensional lattice of $L^3$ sites, and impose periodic boundary conditions.[1] Elementary charge, lattice spacing, and dielectric constant are all taken to be 1. The values of the random potential $\varepsilon_\alpha$ are uniformly distributed between $-B/2$ and $B/2$.

The specific heat of the Coulomb glass $c(T)$ has been studied numerically, including arbitrary correlations, in the literature along two different lines. On the one hand, the Metropolis algorithm was used in Refs. 14, 19, and 20. These papers are restricted to $T\gtrsim 2\times 10^{-2}$ due to the dramatically increasing effort as $T\to 0$. On the other hand, the

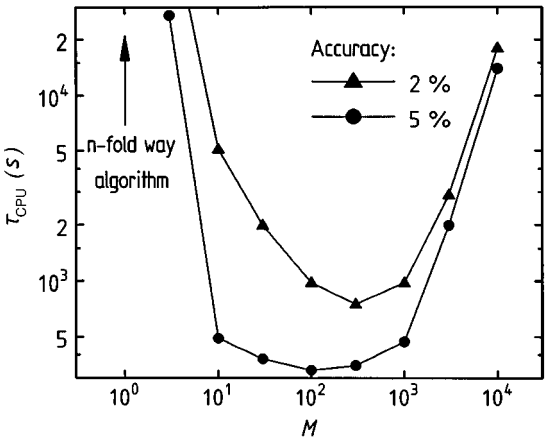


FIG. 1. Dependence of the CPU time $\tau_{\mathrm{CPU}}$, needed for investigating three samples of $4^3$ sites, in dependence on the number of states $M$ in the reservoir $\mathcal{S}^{(l)}$. $\tau_{\mathrm{CPU}}$ is measured in processor seconds at an HP K200. The aspired error bounds of $c(T)$ were chosen to be 5% and 2%, respectively.

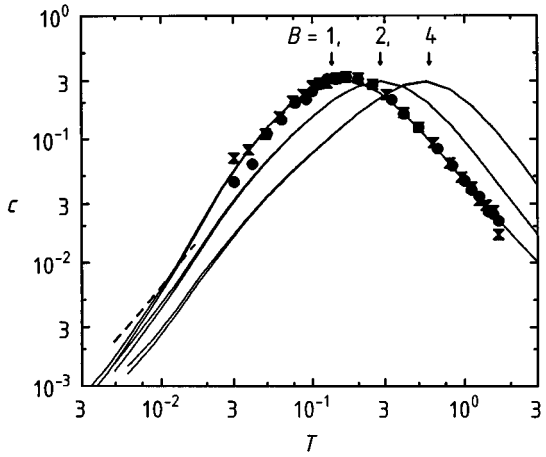


FIG. 2. Temperature dependence of the specific heat of the three-dimensional Coulomb glass for three different degrees of disorder. 50 samples of $10^3$ sites were studied in each case. The pairs of curves cover the 1-$\sigma$ region of ensemble averaging. They result from sets of $T$ points where neighboring ones differ by a factor $\sqrt{2}$. For comparison, data from a previous Metropolis simulations (Ref. 14) are included (circles and double triangles), as well as the results for $L=8$ from a study based on constructing large sets of system states (Ref. 21) (dashed line).

specific heat was investigated by first obtaining large sets of low-energy system states and then calculating $c$ by means of analytical statistical mechanics. The principal feasility of this approach was demonstrated in Ref. 14. Based on a more efficient numerical method to obtain low-energy states, a corresponding detailed study of the influence of correlations was accomplished in Ref. 21. The $T$ region investigated could be extended down to $5\times10^{-3}$ in Refs. 14 and 21, where this bound results from severe finite size effects. However, due to finite size effects and high-energy cutoff, the $c(T)$ results for fixed $L$ covered only about a factor of 3 of the $T$ scale so that the total curve had to be assembled from the pieces obtained for different $L$.[21]

The advantages of both these approaches are combined by the hybrid algorithm presented here. To test it, we performed simulations including into the original dynamics only nearest-neighbor hops and transitions of one electron between the sample and a reservoir at infinity. The specific heat was determined according to $c(T)=(\langle H^2\rangle-\langle H\rangle^2)/(T^2L^3)$, where the accuracy was controlled by means of the following procedure. We decompose the time interval into 30 equal parts, and check whether the dispersion of the related values $\langle c\rangle_n$ is sufficiently small and, moreover, the value measured over the full interval $\langle c\rangle$ does not deviate too much from the mean value of the $\langle c\rangle_n$. An initial $\mathcal{S}^{(l)}$ was obtained by repeated relaxation down to states which are stable with respect to particle exchange with a reservoir at infinity, arbitrary one-particle, and compact two-particle hops inside the sample.

To ensure reliability and efficiency of our program we performed several tests, all of them for $B=1$:

(i) The influence of the number of states $M$ kept in $\mathcal{S}^{(l)}$ was studied considering the ensemble average of three samples of $4^3$ sites. We calculated $c(T)$ for $14T$ values, starting at 0.2, and then diminishing $T$ by the factor $\sqrt{1/2}$ step by step down to $2.2\times10^{-3}$ (not accessible to standard Metropolis simulations). Two such series of runs were performed where the permitted fluctuations of $c(T)$ were restricted to 5 and 2 %, respectively. The results are given in Fig. 1 which relates the CPU time $\tau_{\mathrm{CPU}}$ to $M$. For small $M$, $\tau_{\mathrm{CPU}}$ diverges with decreasing $M$ due to capturing of the system close to local minima: The task could not be performed within reasonable $\tau_{\mathrm{CPU}}$ for $M=1$, which, due to the missing additional transitions within $\mathcal{S}^{(l)}$ and negligible overhead for handling $\mathcal{S}^{(l)}$, corresponds to the standard $n$ fold way algorithm. Even for $M=3$, $\tau_{\mathrm{CPU}}$ is by roughly two orders of magnitude larger than for the optimal $M$. For large $M$, $\tau_{\mathrm{CPU}}$ increases with $M$ because of the effort for treating $\mathcal{S}^{(l)}$. The optimal $M$ value as well as that region, where the sharp low $M$ increase of $\tau_{\mathrm{CPU}}$ starts, shift to higher $M$ as the permitted fluctuations get smaller.

(ii) To test how reliably the deepest states are found, and to check the automatic accuracy control, we performed 50 runs considering the same sample of $10^3$ sites. For this size, we estimated the number of local minima of the dynamics of the standard Metropolis simulation to exceed $10^{22}$. The $T$ values were those of Fig. 2; $M=10\,000$ was chosen; a random error of 5% was aspired. First reentrance into $\mathcal{S}^{(l)}$ occured at $T\approx3\times10^{-2}$. New states, later on included into $\mathcal{S}^{(l)}$, were detected between $7\times10^{-2}$ and $8\times10^{-3}$. Below roughly $8\times10^{-3}$, more than 95% of $\langle c\rangle$ originate from $\mathcal{S}^{(l)}$. We found the lowest 15 states of the final $\mathcal{S}^{(l)}$ to be the same for all runs, where 10 of them are local minima differing from each other by the occupation of up to 62 sites. That these 15 lowest states are very likely indeed complete was confirmed by an independent study according to Ref. 21. The remarkably small percentage of, on the average, only 0.16% missing states out of the lowest 100 ones clearly demonstrates the power of our approach. Moreover, the maximum relative mean square deviation of $c(T)$ was found to amount to 6%.

(iii) Figure 2 compares with previous studies. It shows the ensemble average for 50 samples of $10^3$ sites. Very good agreement is found with previous Metropolis data,[14] as well

as with data obtained directly by considering system states.[21] The advantages of our hybrid procedure are obvious: A broad low-$T$ region not accessible to usual Metropolis simulations could be investigated. In comparison with Ref. 21, the accuracy problems in assembling the total curve from pieces for different $L$ is avoided, and, moreover, the accessible $T$ region is enlarged by increasing the maximum sample size from 512 to 1000 sites.

Finally, the properties of the Coulomb glass themselves are discussed. The influence of the degree of disorder is studied in Fig. 2 by comparing results for $B=1$, 2, and 4. Finite size effects are accounted for as in Ref. 21. Figure 2 suggests three conclusions:

(i) There is a tendency towards $c(T)$ being independent of $B$ as $T\rightarrow 0$.

(ii) The description of $c(T\rightarrow 0)$ in terms of a power law with a characteristic exponent is not possible within the $T$ region studied.

(iii) For $T>0.03$, the increase of $B$ seems to correspond to a rescaling of the temperature.

Moreover, the present results, which have a considerably higher accuracy than previous data, confirm the physical analysis of the influence of correlations in Ref. 21.

Concluding, we have presented a hybrid algorithm bridging the gap between Metropolis algorithm and analytical methods of statistical mechanics. It proved to be very successful in investigating the specific heat of the Coulomb glass, where the low-temperature problems of Metropolis simulations were dramatically diminished by taking advantage of the large memory of modern computers. The algorithm proposed is expected to be useful also for the study of other equilibrium properties of this model. Since the Coulomb glass is not special compared to other disordered systems — in particular, it exhibits a complicated hierarchy of metastable states — this algorithm should be advantageous for the study of low-temperature equilibrium properties of a broad class of disordered systems. Moreover, parts of our approach to the long-standing problem of handling situations with many local minima, should be of considerable advantage for completing physical methods of combinatorial optimization, as simulated annealing. Finally, we stress that the algorithm presented is well suited for the implementation on massive parallel computers.

This work was supported by the German Israeli Foundation for Scientific Research and Development under Grant No. I-0331-256.07/93. A considerable part of the results was obtained during a stay of A.M. at the Weizmann Institute Rehovot. We are grateful to the Albert Einstein Centre for Theoretical Physics at the Weizmann Institute of Science for enabling this visit. A.M. thanks the colleagues at the Physics Department of the Weizmann Institute, in particular A. Finkel'stein, for their warm-hearted hospitality. Discussions with S.L. Drechsler, S. Haas, S. Kobe, S. Mertens, M. Ortuño, M. Pollak, M. Richter, and M. Schreiber have been very useful.

[1] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).

[2] *Applications of the Monte Carlo Method in Statistical Physics*, edited by K. Binder, Topics in Current Physics Vol. 36 (Springer-Verlag, Berlin, 1984).

[3] A.B. Bortz, M.H. Kalos, and J.L. Lebowitz, J. Comput. Phys. **17**, 10 (1975).

[4] M.A. Novotny, Comput. Phys. **9**, 46 (1995), and references therein.

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

[6] H.G. Evertz, G. Lana, and M. Marcu, Phys. Rev. Lett. **70**, 875 (1993).

[7] W. Krauth and O. Pluchery, J. Phys. A **27**, L715 (1994).

[8] B.A. Berg and T. Neuhaus, Phys. Rev. Lett. **68**, 9 (1992); B.A. Berg and T. Celik, *ibid.* **69**, 2292 (1992).

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

[10] S.W. Mahfoud and D.E. Goldberg, Parallel Comput. **21**, 1 (1995), and references therein.

[11] A. Möbius and P. Thomas, Czech. J. Phys. **46**, Suppl. S4, 2237 (1996).

[12] D.D. Frantz, D.L. Freeman, and J.D. Doll, J. Chem. Phys. **93**, 2769 (1990).

[13] P. Rujan, Z. Phys. B **73**, 391 (1988).

[14] M. Schreiber and K. Tenelsen, Europhys. Lett. **21**, 697 (1993); K. Tenelsen and M. Schreiber, Phys. Rev. B **49**, 12 662 (1994).

[15] *Monte Carlo Method in Statistical Physics*, 2nd ed., edited by K. Binder, Topics in Current Physics Vol. 7 (Springer-Verlag, Berlin, 1986), pp. 30–34.

[16] It proved to be advantageous to include into $\mathcal{S}^{(l)}$ a minimum number of local minima, using two different energy bounds, the one for local minima being clearly higher than that one for arbitrary states.

[17] M. Pollak and M. Ortuño, in *Electron-Electron Interactions in Disordered Systems*, edited by A.L. Efros and M. Pollak (North-Holland, Amsterdam, 1985), p. 287.

[18] B.I. Shklovskii and A.L. Efros, *Electronic Properties of Doped Semiconductors*, edited by M. Cardona, Springer Series in Solid-State Sciences Vol. 45 (Springer, Berlin, 1984).

[19] J.H. Davies, P.A. Lee, and T.M. Rice, Phys. Rev. B **29**, 4260 (1984).

[20] S.D. Baranovskii, A.A. Uzakov, and A.L. Efros, Zh. Eksp. Teor. Fiz. **83**, 756 (1982) [Sov. Phys. JETP **56**, 422 (1982)].

[21] A. Möbius and M. Pollak, Phys. Rev. B **53**, 16 197 (1996).