# Make Life Simple: Unleash the Full Power of the Parallel Tempering Algorithm

Elmar Bittner, Andreas Nußbaumer, and Wolfhard Janke

*Institut für Theoretische Physik and Centre for Theoretical Sciences (NTZ), Universität Leipzig, Postfach 100 920, D-04009 Leipzig, Germany*

(Received 20 May 2008; published 26 September 2008)

We introduce a new update scheme to systematically improve the efficiency of parallel tempering simulations. We show that, by adapting the number of sweeps between replica exchanges to the canonical autocorrelation time, the average round-trip time of a replica in temperature space can be significantly decreased. The temperatures are not dynamically adjusted as in previous attempts but chosen to yield a 50% exchange rate of adjacent replicas. We illustrate the new algorithm with results for the Ising model in two and the Edwards-Anderson Ising spin glass in three dimensions.

DOI: 10.1103/PhysRevLett.101.130603 PACS numbers: 05.10.Ln, 02.70.−c, 05.50.+q, 75.50.Lk

The parallel tempering (PT), or replica exchange, simulation technique [1–4] provides an efficient method to investigate systems with rugged free-energy landscapes [5], particularly at low temperatures. Initially, applications of the method were limited to problems in statistical physics. By now, however, PT and its extensions are used in many disciplines, e.g., biomolecules [6–12], bioinformatics [13], zeolite structure solution [14], classical and quantum frustrated spin systems [15,16], spin glasses [3,4,17–20], and QCD [21–23]. The use of PT in interdisciplinary fields spanning physics, chemistry, biology, engineering, and material sciences rapidly increases.

In a PT simulation, one generates many replicas of Monte Carlo (MC) Markov chains or molecular dynamics (MD) trajectories at different temperatures in parallel. At regular intervals, an attempt is made to exchange the configurations of different, usually adjacent replicas, which is accepted with probability

$$P_{\mathrm{PT}}(E_1, \beta_1 \to E_2, \beta_2) = \min[1, \exp(\Delta\beta\Delta E)], \quad (1)$$

where $\Delta\beta = \beta_2 - \beta_1$ is the difference between the inverse temperatures of the two replicas and $\Delta E = E_2 - E_1$ their energy difference. The acceptance probability is smaller the larger the temperature difference or the system size gets. For PT simulations to be most efficient, each replica should spend the same amount of time at each temperature. To this end, several strategies have been proposed in the past years [24–30], but an efficient selection of optimal temperature intervals is still an open problem. In the physically appealing protocol proposed by Katzgraber *et al.* [28], the optimal temperatures are determined from the flow in temperature space; that is, the rate of round trips between low and high temperatures is maximized by systematically readjusting the temperatures.

Unfortunately, their initial recursion is rather complex and needs a significant amount of CPU time. Therefore, we do not use the idea of maximum flow and rather employ the concept of a constant acceptance rate between adjacent replicas, which can be calculated from

$$A(1 \to 2) = \sum_{E_1, E_2} P_{\beta_1}(E_1) P_{\beta_2}(E_2) P_{\mathrm{PT}}(E_1, \beta_1 \to E_2, \beta_2), \quad (2)$$

where $P_{\beta_i}(E_i)$ is the probability for replica $i$ at $\beta_i$ to have the energy $E_i$ (the subscript is the replica index). Using this formula we can calculate, starting from $\beta_1$, a set of inverse temperatures $\beta_i$ which satisfy $A(i \to i+1) = \text{const}$. For systems with a diverging specific heat, one obtains a high density of replicas around the critical temperature; i.e., the difference between the inverse temperature of two adjacent replicas is small. For high values of $\beta$, i.e., low temperatures, the difference between energy distributions at different temperatures becomes small, and therefore $\Delta\beta$ increases. Furthermore, for small $\beta$ values, $\Delta E$ decreases and the spacing between the replicas grows.

As an illustration, we shall first consider MC simulations of the two-dimensional (2D) Ising model where the density of states and hence (2) can be calculated exactly [31]. For all reasonably chosen rates $A(1 \to 2)$, the replica flow from high to low temperatures and vice versa turns out to be very slow, at least when a local update scheme, e.g., the Metropolis algorithm, is used for each of the replicas. The replica flow through the temperature space shows a significant drop around the critical temperature. In Fig. 1, we show as an example for an acceptance rate of 50% the fraction of replicas which have visited most recently the smallest $\beta$ value and wander "up" in the inverse temperature space. This sharp drop-off behavior led Katzgraber *et al.* [28] to their feedback-optimized update scheme (FBO-PT), in which they readjust the temperatures by analyzing the local diffusivity.

We, on the other hand, want to remove the unwanted behavior at $\beta_c$, while keeping the temperatures fixed at their initial values. Looking at the trajectory of an arbitrarily chosen replica in temperature space shown in the upper plot of Fig. 2, we see a clear block structure, where the border of the blocks coincides with the critical temperature. Such a block structure is related to a bottleneck in

 

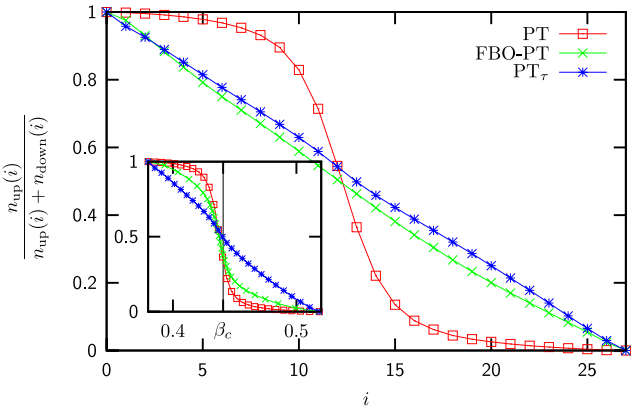


FIG. 1 (color online). Fraction of replicas which wander from the smallest $\beta_i$ to the largest as a function of the replica index $i$ for the 2D Ising model ($L = 80$). The simulations without optimization exhibit a sharp decline close to $\beta_c$, as one can see in the inset. Taking the canonical correlation times $\tau_{\rm can}$ into account (PT$_\tau$), the fraction decreases, for the same set of temperatures, almost linearly.

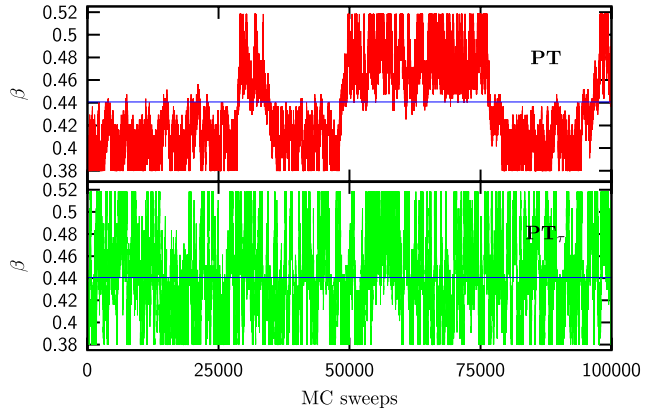


FIG. 2 (color online). Time series of an arbitrarily chosen replica on its way through the inverse temperature space of the 2D Ising model ($L = 80$). The upper plot shows the result of a PT simulation and the lower that of a PT$_\tau$ simulation with $N_{\rm local} = \tau_{\rm can}$. The horizontal lines indicate $\beta_c$, the infinite-volume critical point. The blocks in the time series are a signal of the increasing autocorrelation times due to critical slowing down.

the flow through the temperature space, or, in other words, for a replica starting from a high temperature, it is hard to overcome this bottleneck and move to the low temperature region. A plausible explanation of this observation is as follows: Toward the critical temperature, the autocorrelation time increases due to critical slowing down, and therefore two exchanged replicas stay in phase space close to each other. It is hence more likely that these two replicas exchange again.

To verify this idea, we use a toy model based on the bivariate Gaussian process with $0 \leq \rho < 1$ [32]:

$$e_i = \rho e_{i-1} + \sqrt{1-\rho^2}\, e_i', \quad i \geq 1, \tag{3}$$

where $e_0 = e_0'$ and the $e_i'$ are *independent* Gaussian random variables satisfying $\langle e_i' \rangle = 0$ and $\langle e_i' e_j' \rangle - \langle e_i' \rangle^2 = \delta_{ij}$. By iterating this recursion, it follows that the autocorrelation function is $A(k) = \langle e_0 e_k \rangle = \rho^k \equiv e^{-k/\tau_{\rm exp}}$, where $\tau_{\rm exp} = -1/\ln\rho$ is the exponential autocorrelation time. It can be shown that with increasing $\tau_{\rm exp}$ the mean step size decreases, i.e., $\langle |e_{i+1} - e_i| \rangle = 2\sqrt{(1-\rho)/\pi}$, such that the system moves slower through the one-dimensional phase space, and this is what we are interested in.

Using the stochastic process (3), we are able to approximate for any realistic model the movement in energy space during a parallel tempering simulation. From the energy distribution of initial canonical simulations, we obtain for each of the replicas at $\beta_i$ the mean and variance which, after a trivial shift and rescaling, can be reproduced with (3). Next, we exploit the freedom in the model to adjust $\tau_{\rm exp}$ for each temperature which allows us to investigate the dependence of the flow through temperature space on the autocorrelation times. In general, simulations near a second-order phase transition are affected by critical slowing down, i.e., an increasing autocorrelation time $\tau_{\rm can} \sim \xi^z$, where $\xi$ denotes the (spatial) correlation length and $z$ is the dynamical critical exponent. To take this into account, we set $\tau_{\rm exp}$ to the canonical autocorrelation time $\tau_{\rm can}$ of the energy measured in the independent simulations. Together with the mean and variance, this specifies the parameters of the replicated process (3).

By fitting to 2D Ising model MC data, our first finding comes from a comparison of the autocorrelation times for iterations of (3) with and without the PT routine. As expected, the autocorrelation times for the PT simulation are much smaller. The flow through temperature space looks similar as for the 2D Ising model depicted in Fig. 1. We also find a pronounced decline around the pseudocritical point $\beta_c$. The reason for this behavior is, as already anticipated above, the slowed down dynamics near $\beta_c$. That means, after two adjacent replicas in the vicinity of $\beta_c$ have been exchanged, they will stay close to each other, and changing them back to the original state is more likely than an exchange with another replica. If the dynamics is even slower (by simply tuning $\tau_{\rm exp}$ larger), a complete trapping can be observed, and the replicas do not move from low to high temperatures at all.

By systematically varying the inputted autocorrelation times, our toy model suggests that an easy way to cure this problem is to increase the number of local updates between the PT exchanges proportional to the autocorrelation time of the initial (non-PT) simulation for a given temperature.

This general strategy will be now first tested for the 2D Ising model. For system sizes up to $L = 80$, we use the exact energy distributions [31] to calculate a set of inverse temperatures $\{\beta_i\}$ with an acceptance rate $A(i \to i+1) = 0.5$ starting from $\beta_1 = 0.38$. To cover almost the same

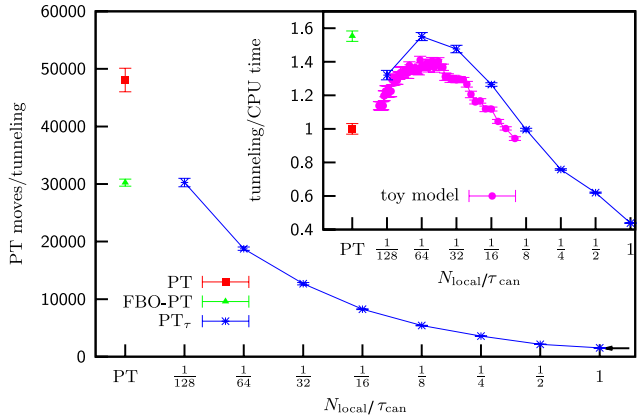


FIG. 3 (color online). PT moves per tunneling as a function of $N_{\rm local}$ for the 2D Ising model ($L = 80$), approaching for $N_{\rm local} \approx \tau_{\rm can}$ the unbiased random walk limit. The inset shows the actually needed computing time in units of the total run time of the standard PT simulation.

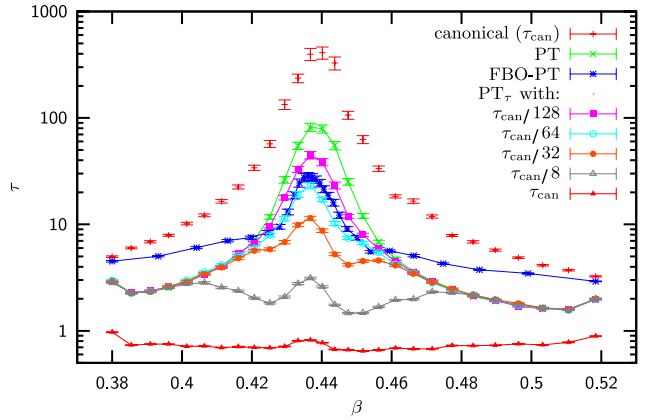


FIG. 4 (color online). Autocorrelation times as a function of inverse temperature for the canonical simulations, the standard PT update scheme, the FBO-PT method, and five different runs of our improved PT update scheme (PT$_\tau$) for the 2D Ising model ($L = 80$).

temperature interval for different system sizes $L$, the number of replicas $N$ has to increase with $L$ [33]. For this set of inverse temperatures, we perform short independent Metropolis MC simulations to estimate the canonical autocorrelation times $\tau_{\rm can}(\beta)$ of the energy, together with the mean and width of the energy distribution. In the actual simulations, we then use the usual PT update scheme with only one important modification; namely, we choose the number of sweeps $N_{\rm local}(\beta)$ between the attempts to exchange the configurations proportional to $\tau_{\rm can}(\beta)$ [$N_{\rm local}(\beta) = 1$ for standard PT and FBO-PT simulations]. The larger the number of sweeps between the exchange attempts, the smaller the correlation between adjacent replicas. Therefore, one has to find a compromise between accuracy and computer time, which can be easily achieved by using our toy model (which runs orders of magnitude faster than the actual simulations). To illustrate this we include in Fig. 3 a comparison for different choices of $N_{\rm local}$.

The main plot of Fig. 3 shows the number of PT moves necessary for an arbitrarily chosen replica to move from the highest to the lowest temperature and back again. In the following, such a round trip will be called tunneling. We clearly see that with an increasing number of sweeps per replica the tunneling time converges to the value of an unbiased random walk (indicated by the arrow in the lower right corner) consisting of two legs of length $(N - 1)$. The limit for one round trip is hence given by $2(N - 1)^2$. If we choose $N_{\rm local}(\beta) = \tau_{\rm can}(\beta)$, the correlation between adjacent replicas is negligible, and each replica performs a random walk through temperature space (see the lower plot in Fig. 2). Furthermore, the sweeps needed for a tunneling event are close to the theoretical value, as is also reflected in the inset in Fig. 1, where we show that the fraction of replicas moving up in the inverse temperature is an almost linear function of $\beta$. This is a major difference from FBO-PT [28], where the linear relation holds only for the fraction of replicas moving up as a function of the *replica index*. In the inset in Fig. 3, we compare the computational cost of our improved PT (denoted by PT$_\tau$) with that for standard PT and FBO-PT, showing that for moderate values of $N_{\rm local}$ the computational effort is the same for both improved methods. To keep the comparison fair, we have excluded the additional computer time needed for FBO-PT to determine the set of inverse temperatures and for PT$_\tau$ to obtain the local autocorrelation times. If one increases $N_{\rm local}$, the ratio of tunnelings per CPU time decreases; i.e., above a certain threshold value of $N_{\rm local}$, the computational effort of PT$_\tau$ increases faster than the improvement of the tunneling speed.

To compare our improved PT$_\tau$ with other methods, one should not only look at the computational cost but also at the accuracy that is achieved for the same number of measurements. An easy way to check this is to measure the autocorrelation time $\tau$. In Fig. 4, we show the autocorrelation times of the 2D Ising model with $L = 80$ for standard PT, FBO-PT, and our PT$_\tau$ algorithm with different choices of $N_{\rm local}(\beta)$. The improvement gained by using PT instead of simulating each temperature independently is almost 1 order of magnitude in the region around the critical point. If one rearranges the inverse temperatures according to the FBO-PT algorithm, one finds even smaller autocorrelation times around $\beta_c$, but the improvement away from criticality is less pronounced than for the standard PT method. Taking in PT$_\tau$ the local autocorrelation times $\tau_{\rm can}(\beta)$ into account, we can decrease $\tau$ systematically. For $N_{\rm local}(\beta) = \tau_{\rm can}(\beta)/64$, where for all temperatures the autocorrelation times of the PT$_\tau$ simulation are slightly smaller than for FBO-PT, the computational effort is almost equal for the two methods. If we use $N_{\rm local}(\beta) = \tau_{\rm can}(\beta)$, then the autocorrelation times are smaller than

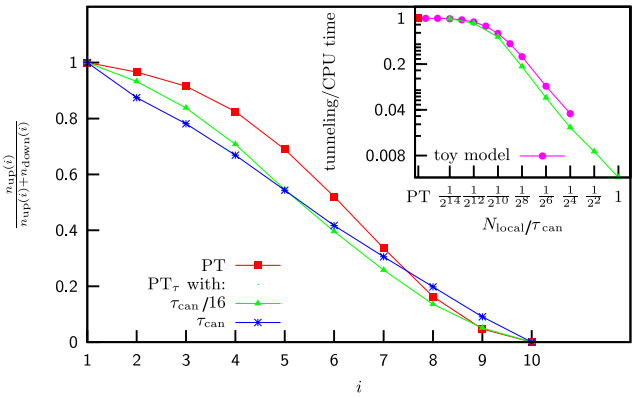


FIG. 5 (color online). Fraction of replicas which wander from the smallest $\beta_i$ to the largest as a function of the replica index $i$ for the 3D Edwards-Anderson Ising spin-glass model ($L = 6$, averaged over 20 disorder realizations).

unity for all temperatures, and the resulting time series are nearly uncorrelated, but the computational costs are clearly too high to make this choice useful.

We close with a brief remark on applications of our PT$_\tau$ algorithm to a MC study of the 3D Edwards-Anderson Ising spin-glass model on a $L = 6^3$ lattice simulated in a temperature range from 0.75 to 1.7 around $T_c \sim 1.15$. Using the same procedure as described above, we find also here an improvement of the replica flow from high to low temperatures, i.e., from the disordered to the spin-glass phase (see Fig. 5). However, the additional computational effort to gain this improvement is significant due to the exponential increase of the autocorrelation time with decreasing temperature. Therefore, one has to carefully tune the balance between used computer time and quality of results.

In summary, we discovered a remarkable block building structure in PT simulations, revealed the mechanism behind it, and showed how to cure this problem by taking into account the temperature dependence of autocorrelation times. This demonstrates how easily the quality of PT simulation data can be improved in both MC and MD studies.

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) under Grants No. JA483/22-1 and No. JA483/23-1, the EU RTN-Network "ENRAGE" under Grant No. MRTN-CT-2004-005616, and the computer-time Grant No. hlz10 of NIC Jülich.

[1] R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **57**, 2607 (1986).
[2] C. J. Geyer, in *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface*, edited by E. M. Keramidas (Interface Foundation, Fairfax Station, VA, 1991), p. 156.
[3] K. Hukushima and K. Nemoto, J. Phys. Soc. Jpn. **65**, 1604 (1996).
[4] E. Marinari, G. Parisi, and J. J. Ruiz-Lorenzo, in *Spin Glasses and Random Fields*, edited by A. P. Young (World Scientific, Singapore, 1998), p. 59.
[5] *Rugged Free Energy Landscapes: Common Computational Approaches to Spin Glasses, Structural Glasses and Biological Macromolecules*, Lect. Notes Phys. Vol. 736, edited by W. Janke (Springer, Berlin, 2008).
[6] U. H. E. Hansmann, Chem. Phys. Lett. **281**, 140 (1997).
[7] Y. Sugita and Y. Okamoto, Chem. Phys. Lett. **314**, 141 (1999).
[8] C. Y. Lin, C. K. Hu, and U. H. Hansmann, Proteins: Struct. Funct. Genet. **52**, 436 (2003).
[9] Y. Yamada, Y. Ueda, and Y. Kataoka, J. Comput. Chem. Jpn. **4**, 127 (2005).
[10] A. Schug *et al.*, J. Phys. Condens. Matter **17**, S1641 (2005).
[11] T. Bataille *et al.*, Z. Kristallogr. Suppl. **23**, 9 (2006).
[12] H.-H. (Gavin) Tsai *et al.*, Proc. Natl. Acad. Sci. U.S.A. **102**, 8174 (2005).
[13] M. Habeck, M. Nilges, and W. Rieping, Phys. Rev. Lett. **94**, 018105 (2005).
[14] M. Falcioni and M. W. Deem, J. Chem. Phys. **110**, 1754 (1999).
[15] A. de Candia and A. Coniglio, Phys. Rev. E **65**, 016132 (2001).
[16] R. G. Melko, J. Phys. Condens. Matter **19**, 145203 (2007).
[17] E. Marinari, in *Advances in Computer Simulation*, edited by J. Kertész and I. Kondor (Springer-Verlag, Berlin, 1998), p. 50.
[18] K. Hukushima, H. Takayama, and H. Yoshino, J. Phys. Soc. Jpn. **67**, 12 (1998).
[19] H. G. Katzgraber, M. Palassini, and A. P. Young, Phys. Rev. B **63**, 184422 (2001).
[20] E. Bittner and W. Janke, Europhys. Lett. **74**, 195 (2006).
[21] B. Joó *et al.*, Phys. Rev. D **59**, 114501 (1999).
[22] E.-M. Ilgenfritz *et al.*, Phys. Rev. D **65**, 094506 (2002).
[23] G. Burgio *et al.*, Phys. Rev. D **75**, 014504 (2007).
[24] D. A. Kofke, J. Chem. Phys. **117**, 6911 (2002); **120**, 10 852 (E) (2004).
[25] C. Predescu, M. Predescu, and C. Ciabanu, J. Phys. Chem. B **109**, 4189 (2005).
[26] N. Rathore, M. Chopra, and J. J. de Pablo, J. Chem. Phys. **122**, 024111 (2005).
[27] A. Kone and D. A. Kofke, J. Chem. Phys. **122**, 206101 (2005).
[28] H. G. Katzgraber *et al.*, J. Stat. Mech. (2006) P03018.
[29] S. Trebst, M. Troyer, and U. H. E. Hansmann, J. Chem. Phys. **124**, 174903 (2006).
[30] D. Gront and A. Kolinski, J. Phys. Condens. Matter **19**, 036225 (2007).
[31] P. D. Beale, Phys. Rev. Lett. **76**, 78 (1996).
[32] W. Janke, in *Ageing and the Glass Transition*, Lect. Notes Phys. Vol. 716, edited by M. Henkel, M. Pleimling, and R. Sanctuary (Springer, Berlin, 2007), pp. 207–260.
[33] E. Bittner and W. Janke (to be published).