International Journal of Modern Physics C
Vol. 17, No. 6 (2006) 817–824

World Scientific
www.worldscientific.com

NUMERICAL COMPARISON OF WANG–LANDAU SAMPLING AND PARALLEL TEMPERING FOR MET-ENKEPHALIN

C. JUNGHANS* and U. H. E. HANSMANN†

John von Neumann Institute for Computing
Forschungszentrum Jülich, 52425 Jülich, Germany

**Institut für Theoretische Physik*
Universität Leipzig, 04109 Leipzig, Germany

†*Department of Physics*
Michigan technological University
Houghton, MI 49931, USA

**Christoph.Junghans@itp.uni-leipzig.de*
†*U.Hansmann@fz-juelich.de*

Received 6 January 2006
Revised 25 January 2006

We compare the efficiency of two prominent techniques for simulation of complex systems: parallel tempering and Wang–Landau sampling. We show that both methods are of comparable efficiency but are optimized for different platforms. Parallel tempering should be chosen on multi-processor system while Wang–Landau sampling is easier to implement on a single-processor computer.

Keywords: Extended ensemble; multicanonical Monte Carlo; Wang–Landau algorithm; replica Monte Carlo; parallel tempering; protein folding; multiple minima problem.

1. Introduction

One of the grand challenges in computational science is the study of proteins and their folding process by way of computer experiments. With regular methods such as Monte Carlo in a canonical ensemble[2] the molecule gets easily trapped in one of the many local minima in the rough energy landscape of a protein. Since the probability for escaping a local minimum in the canonical ensemble is given by $\exp(-\Delta E/k_B T)$ (with k_B the Boltzmann constant), the probability to escape a local minimum over an energy barrier of heights ΔE becomes vanishing small with decreasing temperature T. Various non-traditional Monte Carlo or molecular dynamics techniques have been designed to overcome this multiple minima problem (for a review, see, for instance, Ref. 1). Here we compare two prominent techniques. The first one is parallel tempering[3,4] where the molecule can escape local minima through a random walk in temperature space. On the other hand, in generalized ensemble[5] simulations artificial weights are introduced that do not depend on temperature. A common

variant of this approach is the Wang–Landau algorithm[6] which allows one to estimate the density of states and — in combination with a multicanonical simulation[7] — to determine thermal averages of physical quantities. As both techniques have become methods of choice in protein simulations it is worthwhile to compare their numerical efficiency. For this purpose, we have simulated the pentapeptide Met-enkephalin. Because of its small size this molecule has become an often used model in protein science for evaluating numerical techniques. We compare both methods by measuring (keeping the CPU time constant) in each case the number of *independent* visits of the known ground state of the molecule. Our results show that for this molecule both approaches are of comparable efficiency but at least two orders of magnitude faster than regular canonical simulations.

2. Methods

2.1. *Force field*

A prerequisite for every protein simulation is choice of a model that describes the protein and allows to calculate its energy as a function of its specific configuration. In the present paper we have chosen the ECEPP/3 force field[8] as implemented in the open source program package SMMP[9] (available from www.phy.mtu.edu/biophys/smmp.htm). Here, the energy of a protein:

$$E_{\text{tot}} = E_{LJ} + E_{el} + E_{hb} + E_{\text{tors}}$$

is the sum of the 12–6 Lennard–Jones term E_{LJ}:

$$E_{LJ} = \sum_{(i,j)} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right) a \,,$$

the electrostatic energy E_{el} as given by:

$$E_{el} = 332 \sum_{(i,j)} \frac{q_i q_j}{\epsilon r_{ij}} \,,$$

a term describing hydrogen-bonds:

$$E_{hb} = \sum_{(i,j)} \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) \,,$$

and finally, the torsion energy E_{tors} as given by:

$$E_{\text{tors}} = \sum_{l} U_l (1 \pm \cos(n_l \chi_l)) \,.$$

In this formula r_{ij} stands for the distance between the atoms i and j and q_i, q_j are the associated charges. The lth torsion is named by χ_l and ϵ is the dielectric constant of the environment. The constants A_{ij}, B_{ij}, C_{ij} and D_{ij} are parameters of the empirical potential. The energy E is measured in kcal/mol and r in Å.

2.2. *Parallel tempering*

Detailed energy functions such as the one described in the last sub-section are a mixture of attractive and repulsive interactions. The resulting rough energy landscape leads to extremely slow thermalization of protein simulations at low temperatures. One way to overcome this difficulty is parallel tempering[3] (also known as replica exchange method or Multiple Markov chains), a technique that was first applied to protein studies in Ref. 4. In parallel tempering one considers an artificial system built up of N *non-interacting* replicas of the molecule, each at a different temperature T_i. In addition to standard Monte Carlo or molecular dynamics moves that act only on one replica (i.e., the molecule at a fixed temperature), an exchange of conformations between two copies i and $j = i + 1$ is allowed with probability

$$P(C_i \to C_j) = \min(1, \exp(-\beta_i E(C_j) - \beta_j E(C_i) + \beta_i E(C_i) + \beta_j E(C_j))) \,. \quad (1)$$

The exchange of conformations will at low temperatures lead to a faster convergence of the Markov chain than is observed in regular canonical simulations with only local moves. This is because the resulting random walk in temperatures allows the configurations to move out of local minima and cross energy barriers. The expectation value of a quantity Q at temperature T is calculated as:

$$\langle Q \rangle(T) = \frac{\sum_i \delta_{T_i,T} Q_i}{\sum_i \delta_{T_i,T}} \,.$$

Here the sum goes over all measurements of Q in the simulation and δ is the Kronecker symbol. Expectation values at interpolating temperatures are calculated by re-weighting.[10]

2.3. *Wang–Landau-Algorithm*

Originally, the Wang–Landau-Algorithm has been designed to estimate the density of states $\Omega(E(C))$ by performing a random walk in energy space. At start, each energy is assigned the same weight $W(E) = 1$ and one perform a extended Metropolis update with the transition probability

$$P(C \to C') = \min\left(1, \frac{W(E(C'))}{W(E(C))}\right) \,.$$

After visiting a configuration with energy E, its weight is divided by a shape factor $f > 1$ (we start with $f = e$):

$$W(E) \to \frac{W(E)}{f} \,.$$

This procedure is repeated for a certain number of sweeps, in our case 100,000 sweeps. Afterwards, the shape factor f is replaced by $f' = \sqrt{f}$ and the simulations. We increase the number of sweeps in each iteration by 10%. Successively, the shape factor is lowered till $f - 1 \leq \varepsilon$ with ε a pre-chosen tiny number. At this point, the density of states $\Omega(E)$ can be estimated by $1/W(E)$. Notice that this algorithm does

not fulfill detailed balance as the weights are changing during the simulation. For convenience, the density of states is often determined only within a given interval corresponding to the relevant temperature range of the system. For Met-enkephalin we have chosen -20 kcal/mol $\leq E \leq 20$ kcal/mol. Higher energies are suppressed with a Boltzmann factor of temperature $T = 1000$ K. Because in the Wang–Landau algorithm the estimate of the density of states has an uncontrollable error (as the technique does not realizes a Markov chain), we do a final run with high statistics and fixed weights. Under these conditions, the Wang–Landau sampling becomes a multicanonical simulation.[7] Expectation values and errors of physical quantities can be calculated over a range of temperatures through re-weighting as the distribution of the energies is now given by:

$$P(E) = W(E)\Omega(E) \approx \text{const}\,.$$

The expectation value at a temperature T is given by:

$$\langle Q \rangle(T) = \frac{\sum_i Q(C_i) W(E(C_i))^{-1} \mathrm{e}^{-E(C_i)/k_B T}}{\sum_i W(E(C_i))^{-1} \mathrm{e}^{-E(C_i)/k_B T}}$$

Here, the sum goes over all measurements of configurations C.

3. Results

We have compared the two algorithms by simulating the pentapeptide Met-enkephalin (Tye–Gly–Gly–Phe–Met) which in the last years has become an often used test case in protein folding simulations. For parallel tempering, we choose ten temperatures between 100 K and 1000 K that are equidistant in $1/T$. We find that each of the 10 energy histograms overlaps with its neighbors allowing therefore exchange between neighboring temperatures with sufficient high probability. As a consequence, the parallel tempering protocol leads to a random walk of the replicas through all ten temperatures. We have made 100 000 sweeps for thermalisation and 1 000 000 sweeps for measurement per replica. Figure 1 demonstrates that within the first 10 000 sweeps the lowest temperature ($T = 100$ K) has been visited by every replica at least once. The complementing plot is shown in Fig. 2. This plot follows "replica 4" over the first 10 000 sweeps and shows that each of the ten temperatures has been visited at least once. The total $10 \times 1\,100\,000$ sweeps require 4.25 h on 10 CPUs (Power PC+ 1400 MHz of the supercomputer JUMP at FZ Jülich).

For determining the weights of the final run we have made 25 iterations of Wang–Landau sampling. We extend the length of a run by a factor 1.1 in each iteration starting with 100 000 sweeps in the first iteration and ending with 1 080 000 sweeps in the last. The known ground state was first found in the 6th iteration. Weights are determined for energy bins of size 1 kcal/mol. After the 25th recursion run (at which $\epsilon = 2.98 \cdot 10^{-8}$) we continued with a multicanonical run of 10 0000 000 sweeps. This run lead indeed to a sufficiently flat histogram, indicating that the simulation now performs indeed a random run in energy space (see Fig. 3). For the recursion

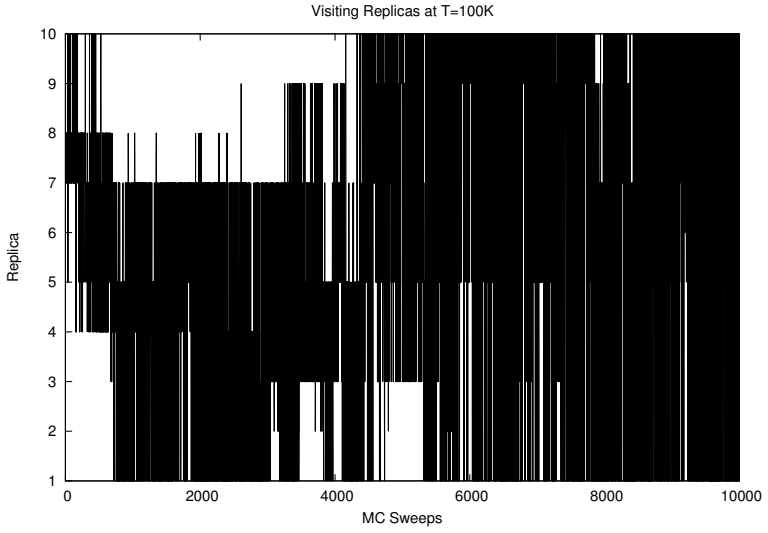

Fig. 1. Replicas visiting the temperature $T = 100$ K in first 10 000 sweeps.

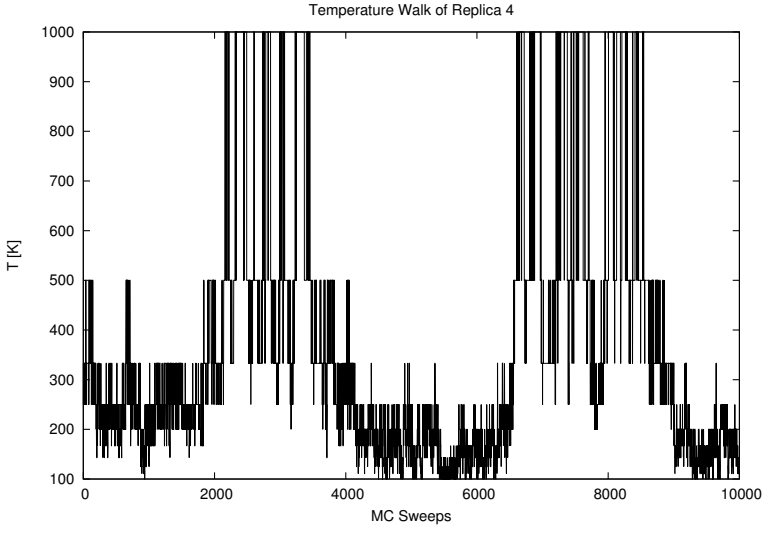

Fig. 2. The walk of replica "4" through temperature space for the first 10 000 sweeps.

15.69 h on a Intel P4 with 3 GHz was needed, and the final simulation required 12.82 h.

We first compare the mean values of energy and the specific heat defined by:

$$\langle c_V \rangle(T) = \beta^2 \frac{\langle E^2 \rangle(T) - \langle E \rangle^2(T)}{N},$$

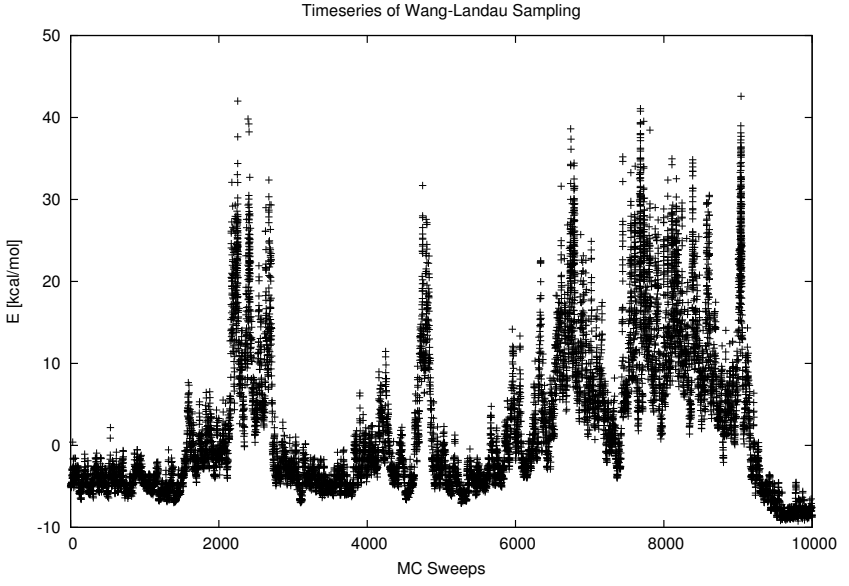

Fig. 3. Time series of energies for the final Wang-Landau simulation. Shown are only the first 10 000 sweeps. Note the random walk through energy covering the low energy region but visiting also from time to time the high energy phase.

Table 1. The mean values of energy and the heat capacity shown good agreement with their error bars have overlapping.

	Parallel tempering		Wang–Landau sampling	
T [K]	$\langle E\rangle_{\mathrm{PT}}$ [kcal/mol]	$\langle c_V\rangle_{\mathrm{PT}}$ [1]	$\langle E\rangle_{\mathrm{WL}}$ [kcal/mol]	$\langle c_V\rangle_{\mathrm{WL}}$ [1]
100	−9.390 (61)	4.76 (17)	−9.414 (86)	4.66 (18)
111	−8.861 (70)	4.87 (11)	−8.890 (96)	4.88 (16)
125	−8.196 (66)	4.80 (12)	−8.209 (95)	4.88 (16)
143	−7.359 (47)	4.74 (13)	−7.344 (85)	4.78 (15)
167	−6.242 (36)	4.86 (12)	−6.202 (73)	4.84 (10)
200	−4.616 (38)	5.17 (8)	−4.567 (59)	5.16 (10)
250	−1.900 (49)	5.79 (6)	−1.854 (48)	5.76 (8)
333	3.043 (48)	5.91 (5)	3.082 (41)	5.86 (5)
500	10.699 (18)	3.47 (2)	10.728 (26)	3.49 (2)
1000	22.433 (126)	1.85 (6)	21.395 (83)	1.81 (2)

where N is the number of amino acid, in our case $N = 5$. The calculated values are similar and have the expected monotonic behavior (see Table 1). The error bars are estimated with the help of the jackknife technique.[11,12] Hence, both methods lead to comparable results and therefore are likely not prone to systematic errors. In fact, the values in Table 1 agree within the errorbars with that of previous investigations using different techniques.

Table 2. The number of tunnel events for every replica in the parallel tempering simulation and the total number of tunnel events for the parallel tempering simulation and Wang–Landau sampling. A tunnel event is a evolution of the system from energy $E = -10$ kcal/mol to energy $E = 18$ kcal/mol and back.

Tunnel events of every replica of parallel tempering										
Replica	1	2	3	4	5	6	7	8	9	10
Tunnel events	9	14	9	15	16	11	13	13	14	11
Total tunnel events of parallel tempering										125
Total tunnel events of Wang–Landau sampling										100

The efficiency of algorithms in protein simulations is given by their ability to sample *independent* low-energy configurations. A necessary condition is the generation of new configurations. Hence, the probability to accept a proposed new configuration should not be vanishing small. In our case, the acceptance rate was for both algorithms sufficiently large and of same order (see Table 3). However, a more distinguishing quantity to compare the efficiency of the two algorithm is the tunneling time. This quantity measures the time needed for the system for a complete walk through energy space from a low-energy configuration (defined by us as one with an energy $E \leq -10.0$ kcal/mol) into the high-energy region (energies with $E \geq 18$ kcal/mol) and back. Obviously, two configurations that are separated by such a tunnel event are independent from each other as the visit in the high-energy region ensures that they are no longer correlated. Hence, the number of tunneling events is a lower bound for the number of truely independent low-energy configurations that have been sampled in a simulation. Values for the two methods are listed in Table 2.

We find little difference in the tunneling times between Parallel tempering (125 Tunneling events) and the Wang–Landau sampling (100 Tunneling events). Hence, our data indicate that both techniques have a comparable efficiency. For comparison we have performed an additional simulation of Met-enkephalin in a canonical ensemble corresponding to a temperature $T = 100$ K. The lowest energy found in a simulation of 11 000 000 sweeps was -8.97 kcal/mol, so the mean value was too high and the system never explored the lower energy regions of the configuration space. Hence, only one *independent* ground state structure was found in this canonical simulation. In this sense, regular canonical simulations are at least two orders of magnitude slower than either parallel tempering or Wang–Landau Sampling. Obviously, the latter two are preferable in protein simulations. However, which of the two algorithms one should choose for a certain application will depend on the specific application. If a parallel computer is available, it may be preferable to use parallel tempering as this technique will lead to faster results. On the other hand, Wang–Landau sampling is easier to implement on a single-processor machine and here preferable.

Table 3. The acceptance rate at every temperature in parallel tempering simulation and the average acceptance rate of the parallel tempering simulation and Wang–Landau sampling. A update counted as accepted if the energy has changed.

	Acceptance rate at every temperature of parallel tempering									
T [K]	100	111	125	143	167	200	250	333	500	1000
acceptance rate	95.2%	98.1%	98.3%	98.6%	98.9%	99.1%	99.4%	99.7%	99.8%	99.9%
Average acceptance rate of parallel tempering										98.7%
Average acceptance rate of Wang–Landau sampling										96.8%

4. Conclusions

We have compared two prominent algorithms that are often used in protein studies. Both Parallel tempering and Wang–Landau sampling allow to obtain information over a large temperature range from a single simulation. Both sample *independent* low-energy configurations with comparable efficiency but are at least two orders of magnitude faster than a canonical simulation at a low temperature of $T = 100$ K. Hence, choice between both algorithms will depend on the equipment available and personal preferences of the researcher. As both techniques are numerically different they allow one also to cross-check simulational results in protein studies.

Acknowledgments

This work was supported in part by the National Science Foundation (CHE-0313618). C. J. thanks the Forschungszentrum Jülich for a summer fellowship.

References

1. Y. Iba, *Int. J. Mod. Phys. C* **12**, 623 (2001).
2. N. Metropolis *et al.*, *J. Chem. Phys.* **21**, 1087 (1953).
3. K. Hukushima and K. Nemoto, *J. Phys. Soc.* (*Jap.*) **65**, 1604 (1996).
4. U. H. E. Hansmann, *Chem. Phys. Lett.* **281**, 140 (1997).
5. G. M. Torrie and J. P. Valleau, *J. Phys. Comp. Phys.* **23**, 187 (1977).
6. F. Wang and D. P. Landau, *Phys. Rev. E* **64**, 056101 (2001).
7. B. A. Berg and T. Neuhaus, *Phys. Rev. Lett.* **68**, 9 (1992).
8. G. Némethy *et al.*, *J. Phys. Chem.* **96**, 6472 (1992).
9. F. Eisenmenger *et al.*, *Comput. Phys. Commun.* **138**, 192 (2001), available from www.phy.mtu.edu/biophys/smmp.htm.
10. A. M. Ferrenberg and R. H. Swendsen, *Phys. Rev. Lett.* **61**, 2635 (1988).
11. B. Efron, *The Jackknife, the Bootstrap and Other Resampling Plans* (SIAM, Philadelphia, 1982).
12. R. G. Miller, *Biometrika* **61**, 1 (1974).