ELSEVIER

Physica A 305 (2002) 157–171

PHYSICA A

www.elsevier.com/locate/physa

Generalized parallel sampling

T.W. Whitfield, L. Bu, J.E. Straub *

Department of Chemistry, Boston University, Boston, MA 02215, USA

Abstract

We develop a generalized version of the parallel tempering algorithm, based upon the non-extensive thermostatistics of Tsallis and coworkers. The effectiveness of the method is demonstrated on a simple one-dimensional problem and on a Lennard-Jones cluster.

PACS: 05.10.Ln; 02.70.Tt; 02.70.Uu

Keywords: Parallel tempering; Quasiergodicity; Non-extensive thermostatistics

1. Introduction

Recently, the "parallel tempering" [1–7] method has gathered attention [8–12] as a way to mitigate the problem of "broken ergodicity" [13,14] (also called "quasi-ergodicity" [15]) which can present challenges to Monte Carlo simulations in the canonical ensemble. Broken ergodicity occurs when the simulation temperature is low enough that the configuration space of a system becomes effectively partitioned into important low-energy regions, or basins, connected by low transition probabilities. In such cases, the simulation time required to sample the complete set of basins becomes prohibitively long, and calculations of equilibrium averages become intractable. Broken ergodicity is characteristic of complex systems with rugged potential energy landscapes, like spin glasses, atomic clusters and biomolecules.

At high enough temperatures, however, transition probabilities become sufficiently large that standard Monte Carlo methods can sample all important basins. Parallel tempering (also called "replica-exchange" [3,4]), along with the related methods of "simulated tempering" [16–18] and "J-walking" [19,8], take advantage of the ergodicity present at higher temperatures by effectively allowing the exchange of configurations

* Corresponding author. Tel.: +1-617-353-6816; fax: +1-617-353-6466.
E-mail address: straub@bu.edu (J.E. Straub).

0378-4371/02/$ - see front matter
PII: S0378-4371(01)00656-2

between low and high temperatures. With the parallel tempering method in particular, this is accomplished by running a number of independent and simultaneous simulations of the same system at different temperatures, and allowing configurational exchanges between neighbouring (or perhaps just different [8]) temperatures. The set of temperatures chosen for a parallel tempering simulation is determined by the criteria that a Monte Carlo simulation at the highest temperature be ergodic, and that configurational exchanges between pairs of temperatures are accepted with appreciable frequency.

The J-walking method may be viewed as a special case of parallel tempering [8], where only two temperatures are used, and the set of configurations sampled at the high temperature is computed and stored beforehand. High-temperature configurations are then occasionally fed to the low-temperature simulation as trial moves. Although such "jump" moves have the same formal acceptance probability as do exchanges in the parallel tempering method, the exchange is not symmetric in the sense that the low-temperature configurations are not added to the stored set sampled from the higher temperature. In completely analogous fashion, the "q-jumping" method [20] feeds configurations from an ergodic simulation (within the standard Monte Carlo scheme) into one plagued with broken ergodicity. In the case of q-jumping, however, the stored configurations are sampled from a simulation which is ergodic not because of *thermally* enhanced rates of barrier crossing but due to sampling from a *non-Boltzmann* distribution. In particular, q-jumping uses conformational distributions from the nonextensive thermostatistics proposed by Tsallis [21]. Previous studies have successfully applied the q-jumping technique to atomic clusters [20], peptide folding [22] and molecular docking problems [23].

In this paper we present a generalization of the parallel tempering method, which uses extensivity, rather than temperature, to parameterize the different parallel simulations. The method is shown to be effective at overcoming problems of broken ergodicity in Monte Carlo simulations. We develop the method first on a simple one-dimensional system, and later apply it to a 13-atom Lennard-Jones cluster.

2. Methods

In this section we first review the nonextensive thermostatistics of Tsallis and coworkers [21,24–26] before introducing the generalized parallel sampling (GPS) method.

2.1. Generalized thermostatistics

The generalized thermostatistics proposed by Tsallis begins with the following definition of the generalized configurational entropy for an N-body system:

$$S_q = \frac{k}{q-1} \int \mathrm{d}\mathbf{r}^N p_q(\mathbf{r})(1 - [p_q(\mathbf{r})]^{q-1}) , \tag{1}$$

where q is a real number and S_q goes to the Gibbs–Shannon entropy, $S = -k \int \mathrm{d}\mathbf{r}^N p(\mathbf{r}) \times \ln p(\mathbf{r})$, for $q \to 1$. The configurational probability distribution function, $p_q(\mathbf{r}^N)$, is

determined by extremizing Eq. (1) subject to the constraints,

$$\int \mathrm{d}\mathbf{r}^N p_q(\mathbf{r}^N) = 1, \quad \int \mathrm{d}\mathbf{r}^N [p_q(\mathbf{r}^N)]^q U(\mathbf{r}^N) = U_q \,, \tag{2}$$

where $U(\mathbf{r}^N)$ is the potential energy. One finds

$$p_q(\mathbf{r}^N) = \frac{1}{Z_q}(1-(1-q)\beta U(\mathbf{r}^N))^{1/(1-q)} \,, \tag{3}$$

where

$$Z_q = \int \mathrm{d}\mathbf{r}^N (1-(1-q)\beta U(\mathbf{r}^N))^{1/(1-q)} \tag{4}$$

is the generalized configurational partition function. The so-called "q-expectation value" for an operator O is defined as

$$\langle O \rangle_q = \int \mathrm{d}\mathbf{r}^N [p_q(\mathbf{r}^N)]^q O(\mathbf{r}^N) \,. \tag{5}$$

To sample the configurational distribution $[p_q(\mathbf{r}^N)]^q$, the following Monte Carlo acceptance probability has been used [27,20]:

$$p = \min\left[1, \left(\frac{p_q(\mathbf{r}^N_{new})}{p_q(\mathbf{r}^N_{old})}\right)^q\right] = \min\left[1, \left(\frac{\mathrm{e}^{-\beta\bar{U}(\mathbf{r}^N_{new})}}{\mathrm{e}^{-\beta\bar{U}(\mathbf{r}^N_{old})}}\right)\right] \,, \tag{6}$$

where

$$\bar{U}(\mathbf{r}^N) = \frac{q}{\beta(q-1)} \ln(1-(1-q)\beta U(\mathbf{r}^N)) \tag{7}$$

defines an effective potential. In order to guarantee that $p_q(\mathbf{r}^N)$ be real it is customary to introduce a constant shift, ε ($\geqslant$ the ground state energy), to the potential energy,

$$\bar{\mathrm{U}}(\mathbf{r}^N) = \frac{q}{\beta(q-1)} \ln(1-(1-q)\beta[U(\mathbf{r}^N)+\varepsilon]) \,. \tag{8}$$

In all the simulations presented here, ε is the ground state energy of the system.

Note that $\lim_{q\to 1}\bar{U} = U$. The acceptance probability therefore goes to that of the standard Metropolis Monte Carlo method as $q \to 1$. Instead, for $q \neq 1$, while much of the structure of the Boltzmann–Gibbs statistical mechanics remains intact [26], thermodynamic state functions like the entropy and internal energy are no longer extensive functions of the system. In particular, q parameterizes the degree of extensivity of state functions [25]. For example, for $q > 1$ the entropy [Eq. (1)] is subadditive, while it is superadditive for $q < 1$.

For enhanced sampling methods in Monte Carlo simulations, it is useful to consider the effect on the configurational distribution (or on the effective potential) of having q different from unity. For $q < 1$, distributions become narrower and more focused around

minima [26,28], while for $q > 1$ they become broader and exhibit greater probability in barrier regions of $U(\mathbf{r}^N)$.

2.2. Generalized parallel sampling

The delocalized nature of the Tsallis distributions may be exploited to sample Boltzmann statistics in systems for which ergodicity is broken in standard Monte Carlo simulations. One approach is to sample from $[p_q(\mathbf{r}^N)]^q$ and reweight averages [29]. That is, the Tsallis distributions may be used in umbrella sampling [30]. An alternative approach is the q-jumping method [20], in which configurations from a $q > 1$ simulation are occasionally supplied as trial moves for a simulation at $q = 1$. These "q-jump" moves are accepted with probability

$$p = \min\left[1, \frac{\mathrm{e}^{-\beta U(\mathbf{r}'^N)}\mathrm{e}^{-\beta \bar{\mathrm{U}}(\mathbf{r}^N)}}{\mathrm{e}^{-\beta U(\mathbf{r}^N)}\mathrm{e}^{-\beta \bar{\mathrm{U}}(\mathbf{r}'^N)}}\right], \tag{9}$$

where the primed coordinates are taken from the $q > 1$ distribution. Inspection of Eq. (9) immediately makes apparent the possibility of sampling the $q = 1$ and $q > 1$ distributions in *parallel*. When sampling from Tsallis distributions in parallel, exchanges between configurations at different values of q (alternatively, the qs may be exchanged) are accepted with probability

$$p = \min\left[1, \frac{\mathrm{e}^{-\beta \bar{\mathrm{U}}(\mathbf{r}'^N;q)}\mathrm{e}^{-\beta \bar{\mathrm{U}}(\mathbf{r}^N;q')}}{\mathrm{e}^{-\beta \bar{\mathrm{U}}(\mathbf{r}^N;q)}\mathrm{e}^{-\beta \bar{\mathrm{U}}(\mathbf{r}'^N;q')}}\right]. \tag{10}$$

It is easy to show that Eq. (10) leads to detailed balance being satisfied for the composite simulation [8].

The scheme we are proposing here, generalized parallel sampling (GPS), consists of performing a series of simultaneous simulations at different values of q. The simulations are independent, apart from occasional exchanges accepted according to Eq. (10). When exchange moves are not attempted, the acceptance rule for particle moves is given by Eq. (6), after choosing an appropriate ε.

One can immediately see how GPS is analogous with parallel tempering. With parallel tempering, choosing the set of temperatures to include is a crucial component in efficiently calculating equilibrium properties for complex systems. In general, one of the temperatures must be high enough such that the important configuration space is rapidly sampled, and the simulation should include enough temperatures so that exchanges between them are frequently accepted. Maintaining a consistent acceptance ratio for exchanges implies a higher density of temperatures in regions where the heat capacity is large [4–6].

With GPS, one has these same technical concerns in determining a set of qs. To explore this point further, we rewrite Eq. (10) as

$$p = \min[1, \mathrm{e}^{-\Delta}], \tag{11}$$

where

$$\Delta \approx \beta\delta\left[\overline{\mathrm{U}}'(\mathbf{r}'^N) - \overline{\mathrm{U}}'(\mathbf{r}^N)\right], \tag{12}$$

and $\overline{\mathrm{U}}'(\mathbf{r}^N)$ is just the first derivative of $\overline{\mathrm{U}}(\mathbf{r}^N)$ with respect to q,

$$\overline{\mathrm{U}}'(\mathbf{r}^N) = \frac{1}{q(q-1)}\left[\frac{q^2}{\beta}F(\mathbf{r}^N) - \overline{\mathrm{U}}(\mathbf{r}^N)\right]. \tag{13}$$

We have set $q' = q + \delta$ and

$$F(\mathbf{r}^N) \equiv \frac{\beta[U(\mathbf{r}^N) + \varepsilon]}{\left(1 - (1-q)\beta\left[U(\mathbf{r}^N) + \varepsilon\right]\right)}. \tag{14}$$

The primed co-ordinates are sampled from the distribution characterized by q'. Averaging over configurations and approximating derivatives as forward differences gives

$$\langle\Delta\rangle_q \approx \beta\delta^2\frac{\mathrm{d}\langle\overline{\mathrm{U}}'\rangle_q}{\mathrm{d}q}. \tag{15}$$

For $q-1$ small, we can expand the logarithm in $\overline{\mathrm{U}}(\mathbf{r}^N)$ and substitute into equation (13). Including terms up to first order gives

$$\langle\Delta\rangle_q \approx \delta^2\frac{\mathrm{d}\langle F\rangle_q}{\mathrm{d}q}. \tag{16}$$

Two things are apparent from this result: first, to keep Δ of order unity, we should reduce the spacing between adjacent qs in regions where the derivative appearing in equations (16) is larger and second, divergences in the derivative will prohibit exchanges between configurations in different phases.

3. Results and discussion

The efficacy of the GPS method was first tested on a simple one-dimensional potential. Features of the technique were also explored with the 13-atom Lennard-Jones cluster. In both systems, comparisons are made with standard Metropolis Monte Carlo methods and with parallel tempering.

3.1. Model potential

In this section we explore the proposed scheme using the following one-dimensional "udder" potential [20]:

$$U(x) = x^2 - a\mathrm{e}^{-b(x+2)^2} - c\mathrm{e}^{-d(x-2)^2} - f\mathrm{e}^{-g(x-3)^2} + \varepsilon\,. \tag{17}$$

Looking at Fig. 1, we can see that $U(x)$ has four minima, including two relatively deep wells around $x=2$ and the global minimum at $x=-1.98673$. We have set $a=f=15$, $b=c=10$, $d=g=3$ and $\varepsilon=11.0265$ (setting the global minimum energy to zero). The barrier heights vary, with the largest one being ≈ 13.

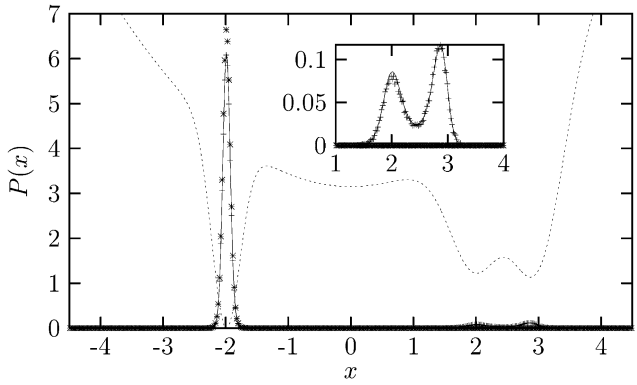

Fig. 1. Probability distributions for the one-dimensional "udder" potential. The exact Boltzmann distribution at $T=1$, indicated with a solid line, is compared with that calculated from standard Monte Carlo techniques ($*$) and the GPS method ($+$). The potential, $U(x)$, has been rescaled to fit in the figure, and is indicated with a dashed line.

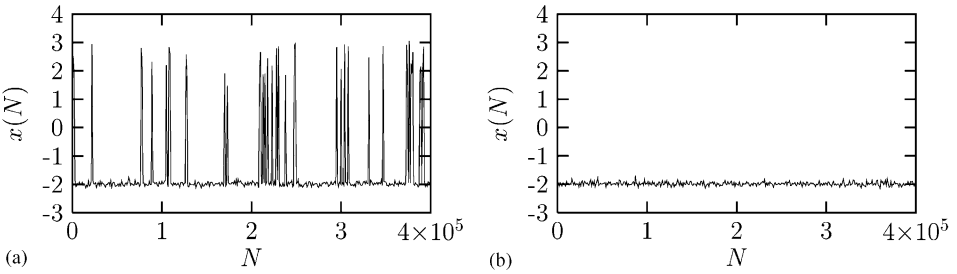

Fig. 2. Position of the walker versus number of Monte Carlo moves at $q=1$ for (a) a GPS simulation over the one-dimensional "udder" potential and (b) the corresponding standard Monte Carlo simulation. The total number of Monte Carlo moves, N, is the same for both simulations.

We have run a GPS simulation, using three parallel random walkers at $T=1$, and compared with standard Monte Carlo simulations. For both simulations, the box size for attempting displacement moves was adjusted so as to maintain a 50% acceptance ratio. Exchanges between walkers at different q ($q_1=1, q_2=1.9, q_3=2$) were attempted with probability 0.1. In this GPS simulation, and all exchange-type simulations throughout this work, exchanges were only attempted between simulations at neighbouring values of q. The neighbouring pairs which attempted exchanges were chosen at random from a uniform distribution. After fixing q_3 at 2, which has been shown to be a good choice for enhanced sampling in this system [20], the intermediate value of q was chosen to maximize the rate of self-averaging (see following section). In Figs. 1 and 2, we see that the standard Monte Carlo simulation becomes trapped in one of the deep wells, while the GPS simulation makes excursions across the barrier and accurately samples the equilibrium distribution.

3.2. Measuring ergodicity

As useful indicators of ergodicity, we employ the "energy metric" and the "fluctuation metric" [31–33]. Both of these measures estimate the rate of self-averaging in equilibrium simulations. Self-averaging is a necessary but not sufficient condition that the ergodic hypothesis be satisfied. In Monte Carlo simulations, the rate of self-averaging for a given observable is expected to be proportional to the rate of configuration space sampling. Looking at the potential energy, we define these two measures as follows. Consider a pair of independent Monte Carlo "trajectories", a and b, starting from different initial conditions. Let $u_{ja}(N)$ be the average potential energy, after N moves, of particle j over trajectory a,

$$u_{ja}(N) = \frac{1}{N}\sum_{k=1}^{N} U_{ja}(k)\,, \tag{18}$$

where $U_{ja}(k)$ is the potential energy of particle j at the kth Monte Carlo move of trajectory a. The limiting value of $u_{ja}(N)$, averaged over all M particles in the system is

$$\overline{u_a(N)} = \frac{1}{M}\sum_{j=1}^{M} u_{ja}(N)\,. \tag{19}$$

The fluctuation metric is now defined as

$$\Omega_a(N) = \frac{1}{M}\sum_{j=1}^{M} [u_{ja}(N) - \overline{u_a(N)}]^2\,. \tag{20}$$

For ergodic systems, the average potential energy does not depend on the initial conditions. In that case, the asymptotic behaviour of the fluctuation metric is

$$\frac{\Omega(N)}{\Omega(0)} \approx \frac{1}{D_\Omega N}\,, \tag{21}$$

where D_Ω is a rate of self-averaging, or generalized diffusion constant, for the potential energy. That is, for ergodic sampling in the asymptotic regions, $\Omega(0)/\Omega(N)$ is linear in N, with the slope indicating the rate of self-averaging.

For the one-dimensional "udder" potential, where $M = 1$, we have looked at a simplified metric,

$$\Omega_a(N) = [u_a(N) - \overline{u_{ex}}]^2\,, \tag{22}$$

where

$$u_a(N) = \frac{1}{N}\sum_{k=1}^{N} U_a(k)\,, \tag{23}$$

and $\overline{u_{ex}}$ is the *exact* Boltzmann average for the potential energy. In Fig. 3, we compare a parallel tempering simulation with the GPS simulation. Using the fluctuation metric, both simulations were optimized for three walkers by first adjusting the simulation parameters. The GPS parameters are those given above, while the parallel tempering

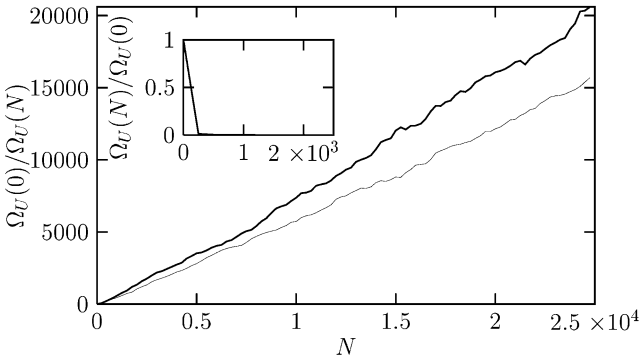

Fig. 3. Reciprocal potential energy fluctuation metric, $\Omega_U(0)/\Omega_U(0)$, for simulations of the one-dimensional "udder" potential. The heavy solid line corresponds to the GPS simulation and the light solid line to a parallel tempering simulation. All data are averages over 300 independent simulations. Inset is the potential energy fluctuation metric, which clearly decays to zero for both simulations in the limit that $N \to \infty$.

simulation was run at the temperatures ($T_1 = 1, T_2 = 2, T_3 = 25$). For both simulations, a 50% acceptance ratio was maintained for the particle displacement moves and exchanges were attempted with probability 0.1. In Fig. 3 we compare the reciprocal fluctuation metric for the low temperature, $q = 1$ ensemble, in both simulations. At this temperature, the barrier height is many times greater than the thermal energy. Nevertheless, we note that both simulations are ergodic, with GPS having a somewhat greater rate of self-averaging.

Another useful gauge of self-averaging is the energy metric, defined as

$$d(N) = \frac{1}{M} \sum_{j=1}^{M} [u_{ja}(N) - u_{jb}(N)]^2 \,. \tag{24}$$

The energy metric uses a pair of independent Monte Carlo "trajectories" to measure the rate of self-averaging. The fluctuation and energy metrics are related by

$$d(N) = \Omega_a(N) + \Omega_b(N) + X_{ab}(N) \,, \tag{25}$$

where the cross term, X_{ab}, measures the correlations between potential energy fluctuations on trajectories a and b. For ergodic systems, the cross term is [33],

$$X_{ab}(N) = -\frac{2}{M} \sum_{j=1}^{M} [u_{ja}(N) - \overline{u_a(N)}][u_{jb}(N) - \overline{u_b(N)}] \,, \tag{26}$$

which clearly goes to zero as the fluctuation metric relaxes. When the simulated system is not ergodic, however, Eq. (26) is no longer valid and X_{ab} may remain finite even when Ω_a and Ω_b both relax to zero. This may happen for example, when simulation a is self-averaging within a different basin from simulation b, and inter-basin motion is shut down. It has been noted [33] that whereas the fluctuation metric is a measure of relaxation processes occurring on a single trajectory, the energy metric can be used to probe rare crossing events between basins. This is a feature that proves useful in

evaluating the effectiveness of sampling methods in complex systems, like the atomic clusters investigated here.

3.3. Lennard-Jones cluster

To further explore the effectiveness of the GPS method, we sample the low-temperature structural properties of the 13-atom Lennard-Jones cluster. The system is defined by the familiar pairwise additive potential for M particles,

$$U(\mathbf{r}^M) = 4\epsilon \sum_{j>i}^{M} \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right] + U_c \,, \tag{27}$$

where σ and ϵ are the Lennard-Jones length and energy parameters, respectively. The distance between particles i and j is given by r_{ij} and U_c is a constraining potential given by

$$U_c = \sum_{i=1}^{M} u(\mathbf{r}_i) \,, \tag{28}$$

where

$$u(\mathbf{r}) = \begin{cases} \infty & |\mathbf{r} - \mathbf{r}_{cm}| > R_c \,, \\ 0 & |\mathbf{r} - \mathbf{r}_{cm}| < R_c \,, \end{cases} \tag{29}$$

where $\mathbf{r}_{cm}$ is the centre of mass position and we have used a confining radius of $R_c = 2.25$.

With the "udder" potential, we have seen how infrequent barrier crossing can become problematic in low-temperature simulations. For the atomic cluster, the radial distribution function can indicate if the correct lowest-energy structure is being sampled, or if the simulation is confined to a higher-energy local minimum.

A GPS simulation for LJ_{13} has been set up with five walkers arranged linearly between $q_1 = 1$ and $q_5 = 1.026$ and at a reduced temperature of $T = 0.02$. The highest value of q was chosen with the work of Hansmann and Okamoto [34,35] in mind, which shows that $q = 1 + q/n_f$, where n_f is the number of degrees of freedom of the system, represents an optimal choice of q for enhancing the sampling rate using Tsallis statistics.

In Fig. 4, we see that one of the requirements of a successful GPS simulation is met with this selection of qs: exchanges between simulations at neighbouring qs are frequently accepted. We track the cluster which started the GPS simulation at $q_1 = 1$. As the simulation progresses, the cluster is free to diffuse through all different values of q with no obvious bias.

Another key feature required for the success of the GPS simulation is apparent from Fig. 5. In that figure, we note that a standard Monte Carlo simulation (i.e., a single walker), performed using an effective potential with $q = 1.026$ leads to self-averaging of the potential energy. Both the fluctuation metric and the energy metric relax to zero, indicating not only effective local self-averaging, but basin-hopping as well.

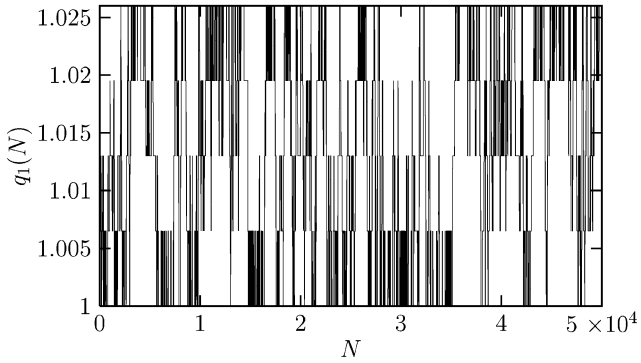

Fig. 4. Random walk in q throughout a GPS simulation of LJ_{13}. The value of q for the walker initially starting at $q = 1$ is plotted as a function of the total number of Monte Carlo steps, N.

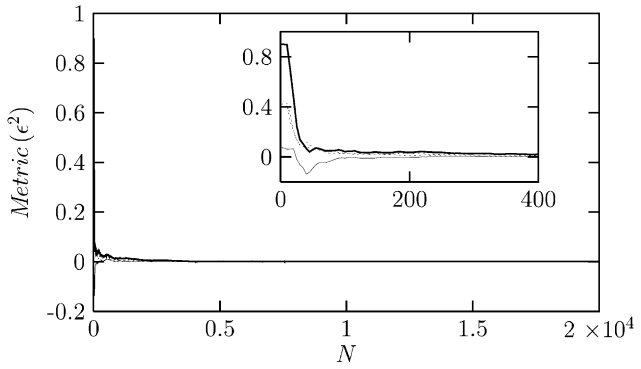

Fig. 5. Ergodic behaviour for standard Monte Carlo simulation of LJ_{13} at $T = 0.02$ and $q = 1.026$. Shown here are $d(N)$ (heavy solid line), $X_{ab}(N)$ (light solid line), $\Omega_a(N)$ and $\Omega_b(N)$ (dashed lines).

In order to ensure that the low-temperature self-averaging was not stymied by particle-exchange effects within the cluster, particle identities were shuffled during the simulations in which $d(N)$ and $\Omega(N)$ were calculated. Without such exchanges, simulations in which the system samples configurations corresponding to physically realistic harmonic motion about the global minimum structure do not appear to be self-averaging. Exchanging particle identities has no effect on the Monte Carlo decisions, nor on the structures that are being sampled.

In contrast to the simulation at $q = 1.026$, Fig. 6 shows the self-averaging properties of a standard Monte Carlo simulation at $q = 1$. While there is self-averaging around local minima, as indicated by the decay of $\Omega_a(N)$ and $\Omega_b(N)$ to zero, there is frustrated motion between local minima, as indicated by $d(N)$ remaining finite for large N.

Fig. 7 shows the self-averaging properties of the potential energy at $q = 1$ from a GPS simulation. As expected, based on the ergodicity of the $q = 1.026$ simulation, and the frequent exchanges between walkers at neighbouring qs, the GPS simulation is also

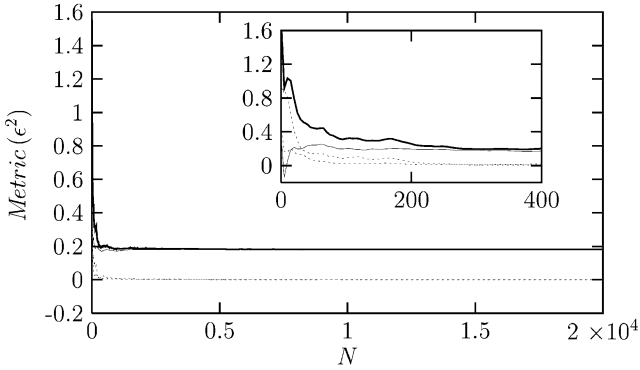

Fig. 6. Ergodic behaviour for standard Monte Carlo simulation of LJ_{13} at $T = 0.02$ and $q = 1$. Shown here are $d(N)$ (heavy solid line), $X_{ab}(N)$ (light solid line), $\Omega_a(N)$ and $\Omega_b(N)$ (dashed lines).

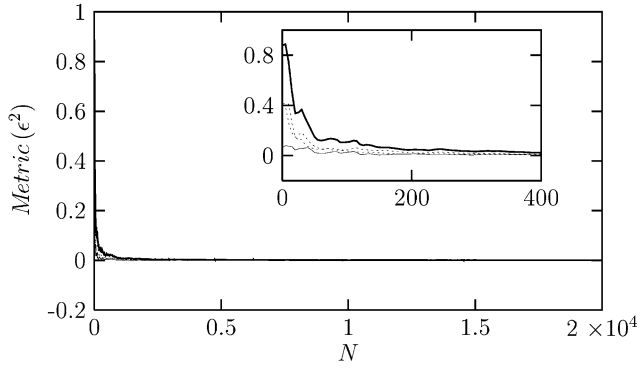

Fig. 7. Ergodic behaviour for GPS simulation of LJ_{13} at $T = 0.02$ with two walkers (at $q = 1$ and 1.026). Shown here are $d(N)$ (heavy solid line), $X_{ab}(N)$ (light solid line), $\Omega_a(N)$ and $\Omega_b(N)$ (dashed lines).

self-averaging. In Fig. 8, we see that the GPS technique accurately captures harmonic motion about the global minimum for this system.

The most remarkable thing about the data in Figs. 7 and 8, however, is that only two walkers were needed to generate it. Although the full five parallel walkers certainly solve the broken ergodicity problem in this simulation, using only the walkers at $q = 1$ and 1.026 is also effective. Looking at Fig. 9, we can appreciate why these two walkers suffice. Since the potential energy histograms for $q = 1$ and 1.026 substantially overlap one another, we expect exchanges to be frequently accepted. Note that a GPS simulation with only two walkers is essentially a q-jumping simulation.

It is interesting to compare the situation with that in a parallel tempering simulation of the same system. In Fig. 10, we plot potential energy distributions taken from a parallel tempering simulation of LJ_{13}. The parallel tempering simulation was performed with 40 walkers spaced between $T = 0.02$ and 0.5. Without including higher temperatures, the five temperatures (between $T = 0.02$ and $T = 0.0808$) used to generate the

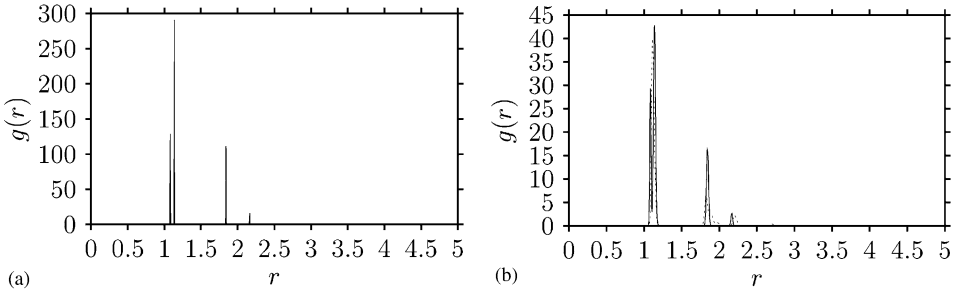

Fig. 8. Radial distribution function, $g(r)$, for LJ_{13} at (a) $T = 0$ (icosahedral global minimum structure) and (b) at $T = 0.02$. In (b), the correct radial distribution function from an ergodic GPS simulation at $T = 0.02$ is drawn with a solid line. The dashed line in (b) is a radial distribution function obtained during a quasi-ergodic simulation.

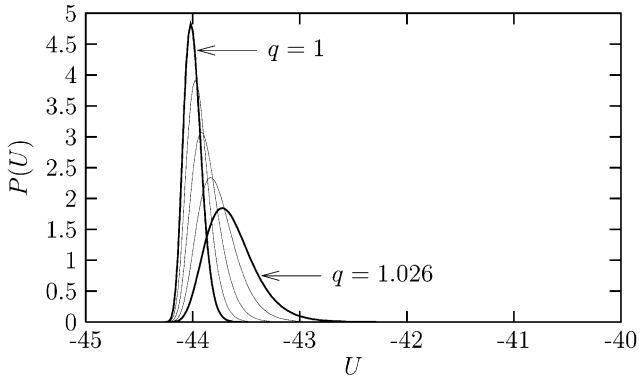

Fig. 9. Potential energy distributions for 5 walkers in a GPS simulation of LJ_{13}. The parameter q is uniformly spaced over the interval $1 \leqslant q \leqslant 1.026$.

histograms in Fig. 10 would not be equilibrium distributions. To understand this, we ran a standard Monte Carlo simulation on the cluster at $T = 0.0808$ (the highest temperature appearing in Fig. 10). At that temperature, the energy metric does not decay to zero, and the simulation is not ergodic. Even if the simulation at $T = 0.0808$ were ergodic, however, there is very little overlap between the distributions at $T = 0.02$ and 0.0808 and intermediate walkers would be required to ensure sufficient acceptance of exchanges.

Finally, to further explore the issue of selecting a good set of qs, and the guideline implied by Eq. (16), we ran a larger GPS simulation which included higher values of q. In Fig. 11, we plot the caloric curve for LJ_{13}, along with an analogous curve in q. The lowest point on each curve, corresponds to the $T = 0.02$, $q = 1$ structure that is characterized by the solid-like radial distribution function plotted in Fig. 8. For high enough values of both T and q, the cluster has a liquid-like structure. Indeed, the bend in the caloric curve (Fig. 11(a)) around $T = 0.3$ corresponds to the well known phase change for this cluster [19,36,37]. The transition as q is increased beyond $q \approx 1.032$,

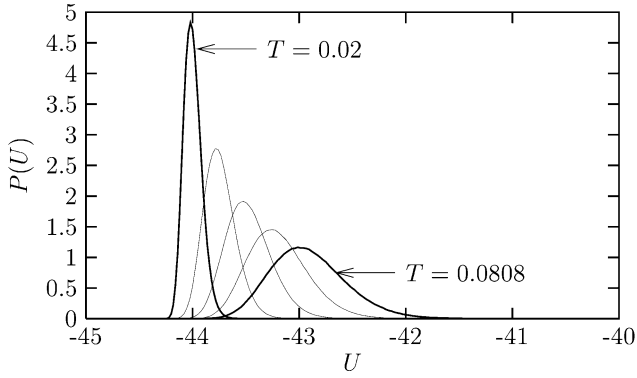

Fig. 10. Potential energy distributions for 5 walkers in a parallel tempering simulation of LJ_{13}. The temperature of the walkers plotted here is uniformly spaced over the interval $0.02 \leqslant T \leqslant 0.0808$.

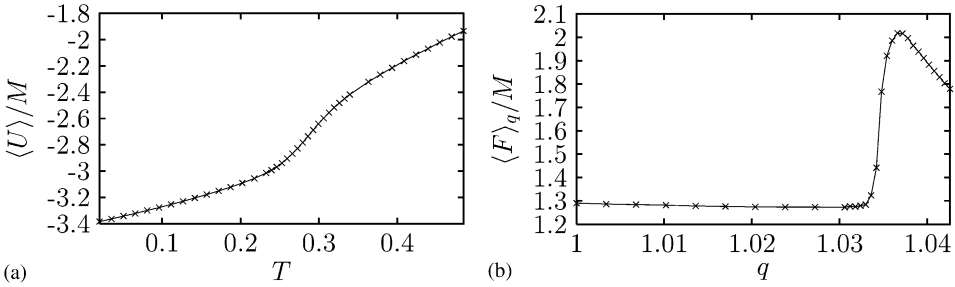

Fig. 11. (a) Caloric curve for LJ_{13} from a parallel tempering simulation at $q = 1$ (i.e., standard Gibbs–Boltzmann statistics). (b) The "q-expectation" value for the potential energy of LJ_{13}, plotted as a function of q, at $T = 0.02$.

however, is very sharp by comparison. Looking at Eq. (16), this would imply that one should use an extremely close spacing between neighbouring qs in that region.

4. Conclusions

The idea of generalizing the parallel tempering method to sampling from non-Boltzmann distributions [38,11] has been explored using Tsallis statistics. The method has effectively been applied to a one-dimensional problem and to an atomic cluster where, in both cases, broken ergodicity is present. We have shown that GPS can effectively sample low-temperature equilibrium distributions in the canonical ensemble. As such, GPS may be an effective technique for global optimization.

With the Lennard-Jones cluster, our method was more efficient than the parallel tempering method. Evidently in this system, increasing q has the effect, over a broad range, of facilitating barrier crossing from local minima to the global minimum. Certainly the

topology of the potential energy surface must help in determining how effective GPS can be.

As q is increased in the cluster, we have noted an abrupt phase change. This posed no difficulty in our simulations, since the sampling was ergodic for q below this transition. One can certainly imagine, however, cases in which it might be important to sample configurations across the transition region. If the transition is not too sharp, increasing the number of parallel simulations in that region should be effective. In parallel tempering and simulated tempering simulations, iterative schemes, which could certainly be applied to GPS, have been used to adjust the set of temperatures [39,4].

Acknowledgements

J.E.S. acknowledges the generous support of the National Science Foundation (CHE-9975494).

References

[1] C.J. Geyer, Markov chain Monte Carlo maximum likelihood, in: E.M. Keramidas (Ed.), Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface, Interface Foundation of North America, Fairfax Station, Virginia, 1991.
[2] M.C. Tesi, E.J. Janse van Rensburg, E. Orlandini, S.G. Whittington, Monte Carlo study of the interacting self-avoiding walk model in three dimensions, J. Stat. Phys. 82 (1996) 155.
[3] K. Hukushima, K. Nemoto, Exchange Monte Carlo method and application to spin glass simulations, J. Phys. Soc. Jpn 65 (1996) 1604.
[4] K. Hukushima, H. Takayama, K. Nemoto, Application of an extended ensemble method to spin glasses, Int. J. Mod. Phys. C 7 (1996) 337.
[5] E. Marinari, Optimized Monte Carlo methods, in: J. Kertész, I. Kondor (Eds.), Advances in Computer Simulation, Lecture Notes in Physics, Vol. 501, Springer, Berlin, 1998 (Chapter 3).
[6] E. Marinari, G. Parisi, J.J. Ruiz-Lorenzo, Numerical simulations of spin glass systems, in: A.P. Young (Ed.), Spin Glasses and Random Fields, Directions in Condensed Matter Physics, Vol. 12, World Scientific, Singapore, 1998 (Chapter 3).
[7] S.G. Whittington, MCMC methods in statistical mechanics: avoiding quasi-ergodic problems, in: N. Madras (Ed.), Monte Carlo Methods, Fields Institute Communications, No. 26, American Mathematical Society, Providence, RI, 2000.
[8] J.P. Neirotti, F. Calvo, D.L. Freeman, J.D. Doll, Phase changes in 38-atom Lennard-Jones clusters. I. A parallel tempering study in the canonical ensemble, J. Chem. Phys. 112 (2000) 10340.
[9] F. Calvo, J.P. Neirotti, D.L. Freeman, J.D. Doll, Phase changes in 38-atom Lennard-Jones clusters. II. A parallel tempering study of equilibrium and dynamic properties in the molecular dynamics and microcanonical ensembles, J. Chem. Phys. 112 (2000) 10350.
[10] S.B. Opps, J. Schofield, Extended state-space Monte Carlo methods, Phys. Rev. E 63 (2001) 056701.
[11] A. Mitsutake, Y. Sugita, Y. Okamoto, Generalized-ensemble algorithms for molecular simulations of biopolymers, Biopolymers 60 (2001) 96.
[12] E. Marinari, G. Parisi, F. Ricci-Tersenghi, F. Zultani, The use of optimized Monte Carlo methods for studying spin glasses, J. Phys. A 34 (2001) 383.
[13] R.G. Palmer, Broken ergodicity, Adv. Phys. 31 (1982) 669.
[14] R. Palmer, Broken ergodicity, in: D.L. Stein (Ed.), Lectures in the Sciences of Complexity, Santa Fe Institute Studies in the Sciences of Complexity, Vol. 1, Addison-Wesley Publishing Company, New York, 1989.
[15] J.P. Valleau, S.G. Whittington, A guide to Monte Carlo for statistical mechanics: 1. Highways, in: B.J. Berne (Ed.), Statistical Mechanics. Part A: Equilibrium Techniques, Modern Theoretical Chemistry, No. 5, Plenum, New York, 1977 (Chapter 4).

[16] E. Marinari, G. Parisi, Simulated tempering: a new Monte Carlo scheme, Europhys. Lett. 19 (1992) 451.
[17] A.P. Lyubartsev, A.A. Martsinovski, S.V. Shevkunov, P.N. Vorontsov-Velyaminov, New approach to Monte Carlo calculation of the free energy: method of expanded ensembles, J. Chem. Phys. 96 (1992) 1776.
[18] C.J. Geyer, E.A. Thompson, Annealing Markov chain Monte Carlo with applications to ancestral inference, J. Am. Stat. Assoc. 90 (1995) 909.
[19] D.D. Frantz, D.L. Freeman, J.D. Doll, Reducing quasi-ergodic behaviour in Monte Carlo simulations by j-walking: applications to atomic clusters, J. Chem. Phys. 93 (1990) 2769.
[20] I. Andricioaei, J.E. Straub, On Monte Carlo and molecular dynamics methods inspired by Tsallis statistics: methodology, optimization, and application to atomic clusters, J. Chem. Phys. 107 (1997) 9117.
[21] C. Tsallis, Possible generalization of Boltzmann–Gibbs statistics, J. Stat. Phys. 52 (1988) 479.
[22] Y. Pak, S. Wang, Folding of a 16-residue helical peptide using molecular dynamics simulation with Tsallis effective potential, J. Chem. Phys. 111 (1999) 4359.
[23] Y. Pak, S. Wang, Application of a molecular dynamics simulation method with a generalized effective potential to the flexible molecular docking problems, J. Phys. Chem. B 104 (2000) 354.
[24] E.M.F. Curado, C. Tsallis, Generalized statistical mechanics: connection with thermodynamics, J. Phys. A 24 (1991) L69.
[25] B.M. Boghosian, Thermodynamic description of the relaxation of two-dimensional turbulence using Tsallis statistics, Phys. Rev. E 53 (1996) 4754.
[26] C. Tsallis, R.S. Mendes, A.R. Plastino, The role of constraints within generalized nonextensive statistics, Physica A 261 (1998) 534.
[27] I. Andricioaei, J.E. Straub, Generalized simulated annealing algorithms using Tsallis statistics: application to conformational optimization of a tetrapeptide, Phys. Rev. E 33 (1996) 3055.
[28] B.M. Boghosian, Navier–Stokes equations for generalized thermostatistics, Braz. J. Phys. 29 (1999) 91.
[29] J.E. Straub, I. Andricioaei, Exploiting Tsallis statistics in: P. Deuflhard, J. Hermans, B. Leimkuhler, A.E. Mark, S. Reich, R.D. Skeel (Eds.), Computational Molecular Dynamics: Challenges, Methods, Ideas. Proceedings of the Second International Symposium on Algorithms for Macromolecular Modelling, Berlin, May 21–24, 1997, Lecture Notes in Computational Science and Engineering, Vol. 4, Springer, Berlin, 1998.
[30] G.M. Torrie, J.P. Valleau, Nonphysical sampling distributions in Monte Carlo free-energy estimation: umbrella sampling, J. Comput. Phys. 23 (1977) 187.
[31] D. Thirumalai, R.D. Mountain, T.R. Kirkpatrick, Ergodic behaviour in supercooled liquids and glasses, Phys. Rev. A 39 (1989) 3563.
[32] D. Thirumalai, R.D. Mountain, Probes of equipartition in nonlinear Hamiltonian systems, J. Stat. Phys. 57 (1989) 789.
[33] D. Thirumalai, R.D. Mountain, Ergodic convergence properties of supercooled liquids and glasses, Phys. Rev. A 42 (1990) 4574.
[34] U.H.E. Hansmann, Y. Okamoto, Generalized-ensemble Monte Carlo method for systems with rough energy landscape, Phys. Rev. E 56 (1997) 2228.
[35] U.H.E. Hansmann, F. Eisenmenger, Y. Okamoto, Stochastic dynamics simulations in a new generalized ensemble, Chem. Phys. Lett. 297 (1998) 374.
[36] C.J. Tsai, K.D. Jordan, Use of an eigenmode method to locate the stationary points on the potential energy surfaces of selected argon and water clusters, J. Phys. Chem. 97 (1993) 11 227.
[37] E. Yurtsever, F. Calvo, Many-body effects on the melting and dynamics of small clusters, Phys. Rev. B 62 (2000) 9977.
[38] I. Andricioaei, J.E. Straub, Computational methods for the simulation of classical and quantum many body systems sprung from nonextensive thermostatistics, in: S. Abe, Y. Okamoto (Eds.), Nonextensive Statistical Mechanics and its Applications, Lecture Notes in Physics, Vol. 560, Springer, Berlin, 2001 (Chapter IV).
[39] W. Kerler, P. Rehberg, Simulated-tempering procedure for spin-glass simulations, Phys. Rev. E 50 (1994) 4220.