

**Original citation:**

Allen, Michael P. and Quigley, David. (2013) Some comments on Monte Carlo and molecular dynamics methods. *Molecular Physics: An International Journal at the Interface Between Chemistry and Physics* . pp. 1-6. ISSN 1362-3028

**Permanent WRAP url:**

<http://wrap.warwick.ac.uk/56225/>

**Copyright and reuse:**

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions. Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

**Publisher's statement:**

This is an Author's Accepted Manuscript of an article published in *Molecular Physics: An International Journal at the Interface Between Chemistry and Physics*, 2013, copyright Taylor & Francis, available online at:  
<http://www.tandfonline.com/10.1080/00268976.2013.817623>.

**A note on versions:**

The version presented here may differ from the published version or, version of record, if you wish to cite this item you are advised to consult the publisher's version. Please see the 'permanent WRAP url' above for details on accessing the published version and note that access may require a subscription.

For more information, please contact the WRAP Team at: [wrap@warwick.ac.uk](mailto:wrap@warwick.ac.uk)

warwick**publications**wrap  
  
highlight your research

<http://go.warwick.ac.uk/lib-publications>

## RESEARCH ARTICLE

Some Comments on Monte Carlo and Molecular Dynamics  
Methods

Michael P Allen and David Quigley  
 Department of Physics, University of Warwick  
 Coventry CV4 7AL, United Kingdom  
 (Received 00 Month 200x; final version received 00 Month 200x)

We highlight some links between molecular dynamics and Monte Carlo algorithms used to simulate condensed matter systems. Special attention is paid to the question of sampling the desired statistical ensemble.

## 1. Introduction

The intention of this paper is to point out some connections between commonly used algorithms in molecular simulation: broadly speaking, methods that are intermediate between Monte Carlo and molecular dynamics. Although the methods are prominent in the molecular simulation literature, and links between them are occasionally pointed out, these inter-relationships do not seem to be as widely known as they should be. More specifically, the properties of various classes of algorithms may often be found in the mathematical and computational literature (see for example [1–3]) while not becoming widely known amongst practitioners of the techniques. We shall consider two common classes of methods: those designed to sample states from the canonical ensemble, and those aimed at flat-histogram sampling.

## 2. Canonical Ensemble Sampling

We shall confine our interest to modelling the *configurational* properties of a system of  $N$  atoms, interacting by a potential energy  $U(\mathbf{r})$ . Here, and henceforth,  $\mathbf{r}$  stands for the complete set of coordinates  $\{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N\}$ . Our interest in the momenta  $\mathbf{p} = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N\}$ , and in dynamical properties of the system, will be restricted to the possibility of using them to explore configuration space in an efficient way. Suppose that we wish to sample configurations from the canonical ensemble, i.e. the Boltzmann distribution

$$\varrho(\mathbf{r}) \propto \exp(-\beta U(\mathbf{r})) \quad (1)$$

where  $\beta = 1/k_B T$ ,  $k_B$  is Boltzmann's constant and  $T$  is the chosen temperature. Standard Monte Carlo methods are based on a Markov chain connecting successive states with a transition matrix  $\pi$  satisfying detailed balance [4–6]

$$\varrho(\mathbf{r}^{\text{old}})\pi^{\text{old} \rightarrow \text{new}} = \varrho(\mathbf{r}^{\text{new}})\pi^{\text{new} \rightarrow \text{old}}. \quad (2)$$

where

$$\pi^{\text{old} \rightarrow \text{new}} = \alpha^{\text{old} \rightarrow \text{new}} P_{\text{acc}}^{\text{old} \rightarrow \text{new}}, \quad (3)$$

$\alpha^{\text{old} \rightarrow \text{new}}$  being the probability of randomly selecting a trial move and  $P_{\text{acc}}^{\text{old} \rightarrow \text{new}}$  the probability of accepting it. Applying eqn (2) gives the acceptance criterion

$$P_{\text{acc}}^{\text{old} \rightarrow \text{new}} = \min \left( 1, \frac{\varrho(\mathbf{r}^{\text{new}}) \alpha^{\text{new} \rightarrow \text{old}}}{\varrho(\mathbf{r}^{\text{old}}) \alpha^{\text{old} \rightarrow \text{new}}} \right) = \min \left( 1, e^{-\beta \Delta U} \frac{\alpha^{\text{new} \rightarrow \text{old}}}{\alpha^{\text{old} \rightarrow \text{new}}} \right) \quad (4)$$

for  $\text{new} \neq \text{old}$ , where  $\Delta U = U(\mathbf{r}^{\text{new}}) - U(\mathbf{r}^{\text{old}})$ , and a residual probability for rejected moves  $\pi^{\text{old} \rightarrow \text{old}} = 1 - \sum_{\text{new}} \pi^{\text{old} \rightarrow \text{new}}$  [7, 8].

The Hybrid Monte Carlo (HMC) method [9] uses a technique based on molecular dynamics to sample this ensemble. An auxiliary set of momenta  $\mathbf{p}$  is introduced, and selected at random from the Maxwell-Boltzmann distribution:

$$\mathbf{p}^{\text{old}} = \sqrt{mk_{\text{B}}T} \mathcal{G}, \quad (5)$$

where  $\mathcal{G}$  is a set of  $3N$  independently sampled Gaussian random numbers with zero mean and unit variance, and  $m$  is the mass (assumed the same for all atoms). The standard molecular dynamics velocity Verlet algorithm [10] advances the configuration by one timestep as follows:

$$\mathbf{r}^{\text{new}} = \mathbf{r}^{\text{old}} + (\Delta t/m) \mathbf{p}^{\text{old}} + \frac{1}{2} (\Delta t^2/m) \mathbf{f}^{\text{old}}, \quad (6a)$$

$$\mathbf{p}^{\text{new}} = \mathbf{p}^{\text{old}} + \frac{1}{2} \Delta t (\mathbf{f}^{\text{old}} + \mathbf{f}^{\text{new}}). \quad (6b)$$

Here  $\mathbf{f} = -\nabla U$ ,  $\Delta t$  is the timestep, and we set  $\mathbf{r}^{\text{old}} = \mathbf{r}(t)$ ,  $\mathbf{r}^{\text{new}} = \mathbf{r}(t + \Delta t)$ , and similarly for  $\mathbf{p}$ ,  $\mathbf{f}$ . It is convenient to introduce the kinetic energy  $K = |\mathbf{p}|^2/2m = \sum_i |\mathbf{p}_i|^2/2m$ , and it is easily shown that the ratio of probability densities for selecting forward and reverse moves may be written as

$$\frac{\alpha^{\text{new} \rightarrow \text{old}}}{\alpha^{\text{old} \rightarrow \text{new}}} = \exp(-\beta \Delta K), \quad \Delta K = K(\mathbf{p}^{\text{new}}) - K(\mathbf{p}^{\text{old}}). \quad (7)$$

The symplectic nature of the velocity Verlet algorithm means that phase-space volume elements  $d\mathbf{r}d\mathbf{p}$  (which in principle should appear in this expression) cancel. Consequently, eqn (4) becomes

$$P_{\text{acc}}^{\text{old} \rightarrow \text{new}} = \min(1, e^{-\beta \Delta U} e^{-\beta \Delta K}) = \min(1, e^{-\beta \Delta H}), \quad (8)$$

where  $H = U + K$ , and this generates the desired distribution of eqn (1).

Eqn (6) does not exactly conserve  $H$ , but because of the existence of a conserved ‘shadow hamiltonian’  $H^{\ddagger}$  [11], differing from  $H$  by  $\mathcal{O}(\Delta t^2)$ , the change  $\Delta H = H^{\text{new}} - H^{\text{old}}$  is expected to be of the same order. Therefore, as  $\Delta t \rightarrow 0$ ,  $P_{\text{acc}}^{\text{old} \rightarrow \text{new}} \rightarrow 1$ , and the Metropolis step becomes unnecessary. Omitting this step gives a form of molecular dynamics incorporating the thermostat of Andersen [12]. If the acceptance/rejection step is included, it can be described as ‘Metropolised’ thermostatted molecular dynamics. There is a fair degree of flexibility in devising schemes that interpolate between the extremes of randomizing all the momenta, and preserving some of them from one step to the next [13–15].

Quite commonly, the acceptance/rejection step is omitted, but doing this will result in incorrect sampling, to an extent depending on  $\Delta t$ . To illustrate this, we

simulate a system of particles in cubic periodic boundary conditions, interacting through the simple interaction potential:

$$U = \sum_{i < j} u(|\mathbf{r}_i - \mathbf{r}_j|), \quad (9a)$$

$$u(r) = \begin{cases} \frac{1}{2}\alpha(1-r)^2 & r \leq 1 \\ 0 & r > 1 \end{cases}. \quad (9b)$$

This represents a soft repulsive interaction between fluid elements or particles, as commonly used in dissipative particle dynamics (DPD) simulations [16, 17]. The unit of length is chosen equal to the range of the potential, and the particle mass is also taken to be unity. We choose a repulsion strength parameter  $\alpha = 25$  and temperature  $k_B T = 1$ , with a number density  $N/V = 3$ , all characteristic of DPD simulations of water under ambient conditions [18]. For this test, a small number of particles  $N = 30$  is sufficient, and simulations were performed over a total run time of 500 reduced units, with timesteps between  $\Delta t = 0.005$  and  $\Delta t = 0.06$ , encompassing the values in common use for this kind of system.

As a measure of deviation from the desired canonical distribution, we consider the configurational temperature  $T_c$  defined by [19, 20]

$$\beta_c = \frac{1}{T_c} = \frac{\langle \nabla^2 U \rangle}{\langle \nabla U \cdot \nabla U \rangle}, \quad (10)$$

where the angle brackets represent a simulation average. In the canonical ensemble,  $T_c = T$ . Fig. 1 shows the results of implementing thermostatted molecular dynamics, with and without the Metropolisation step, as a function of timestep. The results are in agreement with earlier DPD studies [21]: without Metropolisation,  $T_c/T \approx [1 - (\frac{1}{2}\Omega\Delta t)^2]^{-1}$  where  $\Omega$  represents a mean-field (effective) harmonic vibration frequency. This function, with  $\Omega$  treated as a fitting parameter, is shown in Fig. 1. It is clear that errors in configurational properties can increase rapidly with  $\Delta t$ , and the results confirm earlier concerns regarding the size of timesteps commonly used in DPD simulations [22, 23].

Nearly a decade before the introduction of HMC, Rosky et al. [24] proposed the Smart Monte Carlo (SMC) method. They primarily considered the implementation of single-particle moves, but also discussed  $N$ -particle moves, which we focus on here. In this case, the trial move is written

$$\mathbf{r}^{\text{new}} = \mathbf{r}^{\text{old}} + \beta\Delta t D \mathbf{f}^{\text{old}} + \mathcal{R}, \quad \mathcal{R} = \sqrt{2\Delta t D} \mathcal{G} \quad (11)$$

where again all vectors have  $3N$  components, and  $\mathcal{G}$  was defined earlier. The single parameter of the method, written as the combination  $(\Delta t D)$ , deliberately highlights that the proposal step is identical with the standard Brownian Dynamics (BD) advancement algorithm [25].  $D$  represents the diffusion coefficient in the absence of any interactions, in which case the distribution of random displacements satisfies  $\langle \mathcal{R} \mathcal{R} \rangle = 2\Delta t D \mathbf{1}$ , where  $\mathbf{1}$  is the  $3N \times 3N$  unit matrix. Applying standard methods, Rosky et al. derived the move attempt probability ratio to take account of the bias introduced by the forces  $\mathbf{f}$ :

$$\frac{\alpha^{\text{new} \rightarrow \text{old}}}{\alpha^{\text{old} \rightarrow \text{new}}} = \frac{\exp\left\{-\left|(-\Delta \mathbf{r}) - \beta\Delta t D \mathbf{f}^{\text{new}}\right|^2 / 4\Delta t D\right\}}{\exp\left\{-\left|\Delta \mathbf{r} - \beta\Delta t D \mathbf{f}^{\text{old}}\right|^2 / 4\Delta t D\right\}} \quad (12)$$

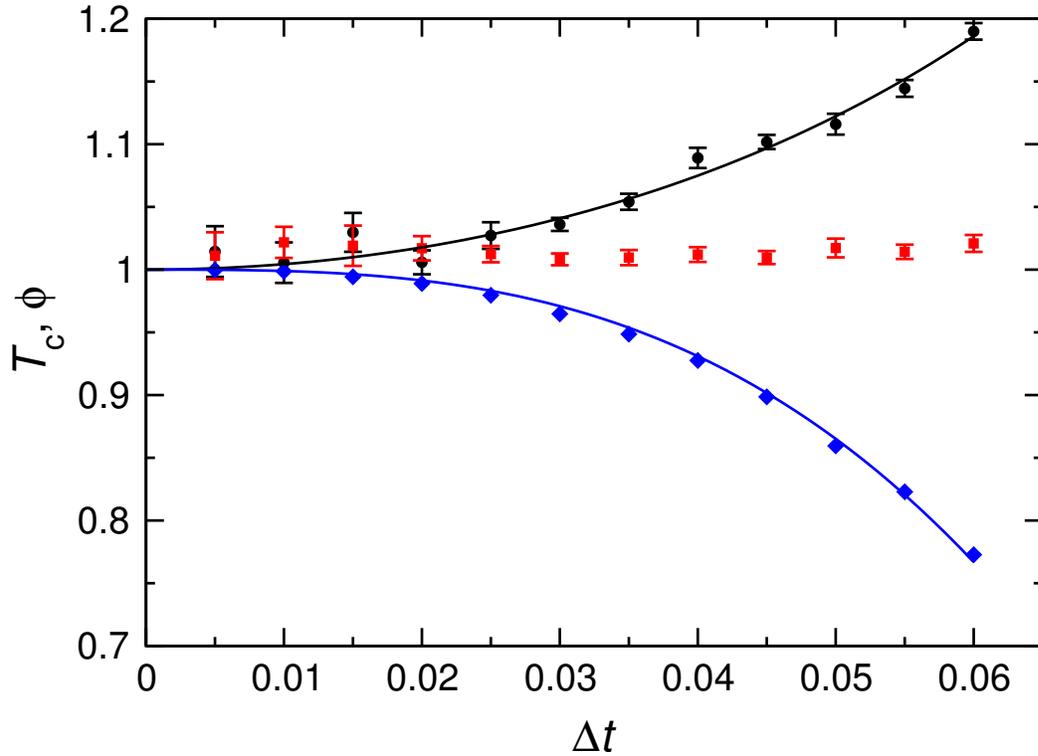


Figure 1. Configurational temperature  $T_c$  for non-Metropolised (circles, black) and Metropolised (squares, red) Andersen-thermostatted dynamics using the DPD potential, with  $T = 1$ . The fitted line is discussed in the text. Also shown is the acceptance ratio  $\phi$  (diamonds, blue) of Metropolis steps in the latter case; error bars for these values are smaller than the symbol size, and the line is a guide to the eye.

This ratio is used in eqn (4); therefore, Rosky et al. [24] established that SMC is a Metropolised BD method. However, they also pointed out that the algorithm can be considered in the context of molecular dynamics, setting

$$\beta\Delta t D = \frac{1}{2}(\Delta t^2/m), \quad \sqrt{2\Delta t D} = \sqrt{mk_B T}(\Delta t/m) \quad (13)$$

in which case eqn (11) becomes *identical* with eqn (6a), combined with the Andersen prescription for selecting momenta, eqn (5). Eqn (12) may be written more simply, by recognising the denominator as  $\exp(-\beta K^{\text{old}})$ , and using eqn (6b) to define  $\mathbf{p}^{\text{new}}$ , in terms of which the numerator becomes  $\exp(-\beta K^{\text{new}})$ . Hence eqn (12) is the same as eqn (7). Although Ref [24] predated the velocity Verlet algorithm [10], and did not base its derivation on the evolution of a hamiltonian system, the SMC algorithm is exactly equivalent to HMC [9]. In the same way, the non-Metropolised advancement algorithms (BD, and MD with Andersen thermostating at each step) are identical: essentially eqn (13) replaces the diffusion coefficient by an inverse mass. Although these links have been noted before [14, 26], they seem not to be widely known.

We note a further extension of the two equivalent descriptions that may be of interest. BD simulations frequently use a configuration-dependent diffusion tensor  $\mathbf{D}(\mathbf{r})$  to approximate the effects of hydrodynamics [27]:

$$\mathbf{r}^{\text{new}} = \mathbf{r}^{\text{old}} + \Delta t \nabla \cdot \mathbf{D} + \beta \Delta t \mathbf{D} \cdot \mathbf{f}^{\text{old}} + \mathcal{R} \quad (14)$$

where now the random displacements  $\mathcal{R}$  are selected from the  $3N$ -dimensional *correlated* Gaussian distribution, with zero means, and  $3N \times 3N$  covariance matrix  $\langle \mathcal{R} \mathcal{R} \rangle = 2\Delta t \mathbf{D}$ . Usually  $\mathbf{D}$  is composed of an  $N \times N$  set of  $3 \times 3$  block matrices for each pair of atoms, for example the Oseen or Rotne-Prager tensor [28],

which depend on the separation vector between them. The generation of the correlated random displacements  $\mathcal{R}$  is a nontrivial numerical problem:  $\mathbf{D}$ , while being symmetric, is not usually sparse, band-structured, or otherwise simple to handle. Various approaches to this problem have been proposed [29, 30]. Reformulating this as HMC is straightforward and instructive: a configuration-dependent mass tensor is introduced by

$$\mathbf{M}^{-1}(\mathbf{r}) = (2\beta/\Delta t)\mathbf{D}(\mathbf{r})$$

and is used to define a kinetic energy  $K = \frac{1}{2}\mathbf{p} \cdot \mathbf{M}(\mathbf{r})^{-1} \cdot \mathbf{p}$ . A reversible, symplectic, algorithm for solving the equations of motion derived from the corresponding hamiltonian has been given [13]. The correct sampling of correlated momenta in the Andersen thermostating step is non-trivial, if  $\mathbf{M}$  is non-diagonal, but can be tackled in the same way as the random displacement in Brownian Dynamics with hydrodynamic interactions. The idea of adjusting the masses, to improve the efficiency of molecular dynamics, goes back to Bennett [31], and there has been a recent revival of interest in this approach in the biomolecular simulation community [32–35]. It would be instructive to test deviations from the canonical ensemble (and possibly correct them by adding a Metropolisation step), when using these approaches.

### 3. Flat Histogram Sampling

We now turn to methods for sampling flat histograms, based on molecular dynamics. Once more, we confine our interest to configurational properties. A key quantity is the density of states  $W(U)$ , defined such that  $W(U)dU$  is proportional to the number of accessible states of the system having potential energy between  $U$  and  $U + dU$ . This is related to the configurational entropy, and hence to the inverse temperature  $\beta(U)$  through

$$\beta(U) = \left( \frac{\partial S/k_B}{\partial U} \right)_V = \left( \frac{\partial \ln W(U)}{\partial U} \right)_V \quad (15a)$$

$$S = k_B \ln W(U) = k_B \int^U dU' \beta(U'). \quad (15b)$$

An ensemble whose energies occur with a distribution function  $p(U)$  will correspond to a configurational distribution of states  $\varrho(U(\mathbf{r}))$  satisfying  $p(U) = W(U)\varrho(U)$ , and so for a flat energy distribution

$$p_{\text{flat}}(U) = \text{constant} \quad \Rightarrow \quad \varrho_{\text{flat}}(U(\mathbf{r})) \propto W(U)^{-1} \propto e^{-S(U)/k_B}.$$

A Monte Carlo algorithm for sampling  $\varrho_{\text{flat}}(U(\mathbf{r}))$  is to choose trial moves  $\mathbf{r}^{\text{old}} \rightarrow \mathbf{r}^{\text{new}}$ , in an unbiased way, and accept/reject them with probability

$$P_{\text{acc}}^{\text{old} \rightarrow \text{new}} = \min \left( 1, \frac{W(U^{\text{old}})}{W(U^{\text{new}})} \right) = \min(1, e^{-\Delta S/k_B}),$$

where  $\Delta S = S(U^{\text{new}}) - S(U^{\text{old}})$ . If the function  $S(U)$  is not known ahead of time, it must be built up progressively during the simulation. Multicanonical sampling [36] and entropy sampling [37] provide possible methods of doing this, but here we specialize to the approach of Wang and Landau [38]. A running histogram of

values  $S(U)$  is updated  $S(U) \rightarrow S(U) + \delta S$  for each value of  $U$  visited during the simulation, thereby reducing the acceptance probability of future moves to that energy, and the increment  $\delta S$  is progressively reduced in a way that is intended to lead to a converged, flat, histogram of energies visited.

Various methods based on molecular dynamics may be devised to generate flat histograms with respect to prescribed order parameters [39–42]. The dynamical sampling of flat energy distributions was originally proposed by Hansmann et al. [43] (see also [44]). The original papers describe (separately) BD (termed Langevin dynamics in Ref. [43]), MD, and HMC algorithms. Here we focus on HMC. As in Sec. 2, a kinetic energy  $K = |\mathbf{p}|^2/2m$  is defined in terms of auxiliary momenta  $\mathbf{p}$ , and used in a formal hamiltonian

$$H^{\text{flat}} = T_0 S(U(\mathbf{r})) + K(\mathbf{p}) \quad (16)$$

where  $T_0$  is an arbitrary temperature. The system may be sampled by thermostatted molecular dynamics based on  $H^{\text{flat}}$ , which now take a form involving scaled forces (see also [45]):

$$\dot{\mathbf{r}} = \mathbf{p}/m \quad (17a)$$

$$\dot{\mathbf{p}} = -\nabla(T_0 S(U)) = -T_0 \left( \frac{\partial S}{\partial U} \right) \nabla U(\mathbf{r}) = (\beta(U)/\beta_0) \mathbf{f}. \quad (17b)$$

Here  $\beta_0 = 1/k_B T_0$ . These are solved step by step. The Andersen thermostat re-selects all initial momenta  $\mathbf{p}^{\text{old}}$  from a distribution at temperature  $T_0$ . Then the advancement step proceeds

$$\mathbf{r}^{\text{new}} = \mathbf{r}^{\text{old}} + (\Delta t/m) \mathbf{p}^{\text{old}} + \frac{1}{2}(\Delta t^2/m) k_B T_0 \beta^{\text{old}} \mathbf{f}^{\text{old}}, \quad (18a)$$

$$\mathbf{p}^{\text{new}} = \mathbf{p}^{\text{old}} + \frac{1}{2} \Delta t k_B T_0 (\beta^{\text{old}} \mathbf{f}^{\text{old}} + \beta^{\text{new}} \mathbf{f}^{\text{new}}). \quad (18b)$$

where  $\beta^{\text{old}} = \beta(U^{\text{old}})$  etc. Finally, the Metropolis criterion may be applied:

$$P_{\text{acc}}^{\text{old} \rightarrow \text{new}} = \min(1, e^{-\Delta K/k_B T_0} e^{-\Delta S/k_B}), \quad (19)$$

although this step is often omitted. It is important to realize that the value chosen for  $T_0$  has no physical significance at all: the scaling  $T_0 \rightarrow \chi^2 T_0$  (and hence  $\mathbf{p}^{\text{old}} \rightarrow \chi \mathbf{p}^{\text{old}}$ ) together with  $\Delta t \rightarrow \chi^{-1} \Delta t$ , leaves the above equations completely unchanged. The combination of  $T_0$ ,  $m$ , and  $\Delta t$  simply serves to generate trial displacements with a physically reasonable variance. It follows that the generated distribution is not a canonical one corresponding to the reference temperature  $T_0$ .

These equations must be combined with a prescription for refining  $\beta(U)$  towards the thermodynamically correct inverse temperature. The simplest approach is to follow the Wang-Landau Monte Carlo prescription. Tabulate the entropy on a regular grid in  $U$ , of spacing  $\delta U$ . At each step, identify the bin  $k$  corresponding to the current value of  $U$ , and update the entropy histogram  $S_k \rightarrow S_k + \delta S$ . When values of  $\beta(U)$  and  $S(U)$  are required for eqns (18), (19), use a suitable interpolation

formula, for example:

$$\begin{aligned}
 S(U) &\approx S_k + \left( \frac{S_{k+1} - S_{k-1}}{2\delta U} \right) (U - U_k) + \frac{1}{2} \left( \frac{S_{k+1} - 2S_k + S_{k-1}}{\delta U^2} \right) (U - U_k)^2 \\
 \Rightarrow \beta(U) &= \left( \frac{\partial S}{\partial U} \right) \approx \left( \frac{S_{k+1} - S_{k-1}}{2\delta U} \right) + \left( \frac{S_{k+1} - 2S_k + S_{k-1}}{\delta U^2} \right) (U - U_k)
 \end{aligned}$$

where  $U_k$  is the nearest tabulated value. At intervals, depending on the flatness of the sampled potential energy distribution, the increment  $\delta S$  is reduced, just as in Wang-Landau. A slightly different approach [46, 47], termed Statistical Temperature Molecular Dynamics (STMD), stores  $\beta(U)$  or  $T(U)$  in a table, rather than  $S(U)$ . This may be more convenient, and is commonly used (see e.g. [48]), but close examination reveals that the numerical method is the same as that obtained by storing  $S(U)$ , so we refer to this as STMD as well. Of course, there is plenty of scope for applying different interpolation and smoothing algorithms, whichever method is used to tabulate the histograms.

The function  $\beta(U) = (\partial S / \partial U)$  is equivalent to the *microcanonical* ensemble configurational temperature [19, 49–52]

$$\beta_c = \left\langle \nabla \cdot \left( \frac{\nabla U}{\nabla U \cdot \nabla U} \right) \right\rangle = \left\langle \frac{\nabla^2 U}{\nabla U \cdot \nabla U} - 2 \frac{\nabla U \nabla U : (\nabla \nabla U)}{(\nabla U \cdot \nabla U)^2} \right\rangle,$$

The expression  $\nabla \nabla U$  is the  $3N \times 3N$  Hessian matrix, which is doubly contracted with  $\nabla U$ , to give the second term. This extra term (compared with eqn (10)) is  $\mathcal{O}(1/N)$  for an  $N$ -atom system. As a simple illustration, following [53], we consider a set of  $2N$  atoms of unit mass in three dimensions, linked into  $N$  independent pairs  $ij$  by harmonic springs of energy  $u(r_{ij}) = \frac{1}{2}r_{ij}^2$ . The density of states, entropy, and configurational temperature, for this system are exactly known functions of  $U$ :

$$W(U) \propto U^{(3N/2)-1} \quad (20a)$$

$$S(U) = k_B \left( \frac{3N}{2} - 1 \right) \ln U + \text{constant} \quad (20b)$$

$$\beta(U) = \left( \frac{3N}{2} - 1 \right) U^{-1} \quad (20c)$$

where the  $\mathcal{O}(1/N)$  terms are evident. Figure 2 shows the results of STMD carried out with the (normally omitted) Metropolisation step. We used a flatness criterion at each stage that the minimum entry in the ‘visits’ histogram shall be at least 95% of the mean of all the nonzero entries, a spacing  $\delta U = 0.2$ , a reference temperature  $T_0 = 1$ , and restricted interest to the range  $0.2 < U \leq 5$ . The timestep was  $\Delta t = 0.1$ . The STMD method typically converged with each iteration taking less than  $10^6$  molecular dynamics steps. The exact results are well reproduced, confirming the consistency between the measured microcanonical entropy function  $S(U)$  and its derivative (whether numerical or analytical). Interestingly, the acceptance rate for the Metropolisation step (19) was seen to decrease from about 0.9 at  $U = 4.5$  to around 0.5 at  $U = 5.0$ , suggesting that this might be an issue in STMD.

Our final observation concerns the very interesting suggestion of Rathore et al. [54], to estimate  $\beta(U)$  from the configurational temperature  $\beta_c(U)$ , calculated and averaged at each tabulated energy, and use this in the algorithm instead. Their suggestion, in fact, was to numerically integrate  $\beta_c(U)$  according to eqn (15b), and

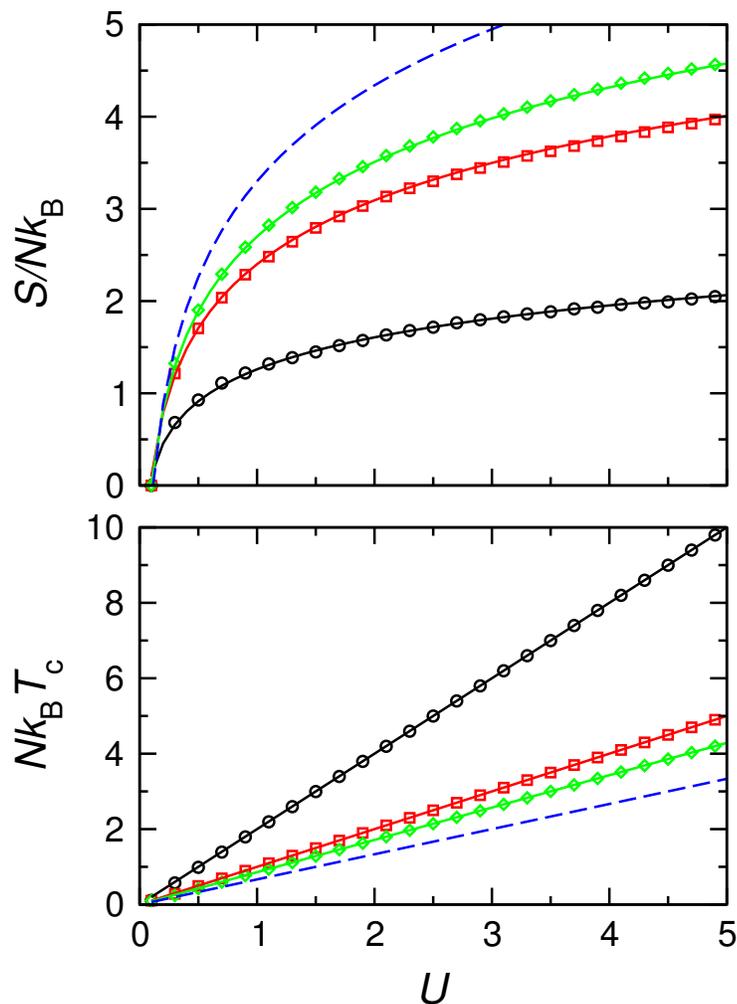


Figure 2. STMD simulations of independent harmonic oscillators in 3D:  $N = 1$  (circles, black),  $N = 2$  (squares, red),  $N = 3$  (diamonds, green). Above: entropy  $S(U)$ , normalized by  $N$ . Below: microcanonical average configurational temperature  $k_B T_c = 1/\beta_c$ , multiplied by  $N$ . Also shown are the exact functions, including finite-size terms (solid lines), and the  $N \rightarrow \infty$  limits (dashed lines, blue).

use the resultant estimate of  $S(U)$  in a biased Monte Carlo simulation, thus avoiding the Wang-Landau process. An obvious extension of the idea is to employ this averaged  $\beta_c(U)$  *directly* in an algorithm for solving the equations of motion (17), for instance eqns (18). In principle, this would give a way of generating a flat histogram *without* explicitly measuring the flatness and using it in a feedback loop, which is essentially what the Wang-Landau method does. This approach appears to be closely related to the adaptive biasing force method of generating flat distributions with respect to a chosen order parameter [42]. We have tested this method, on Lennard-Jones clusters of various sizes, and on the bulk Lennard-Jones system employed by Kim et al. [46]. It proves to be possible to obtain good consistency with the  $\beta(U)$  curves generated by STMD, and indeed to generate very broad distributions in  $U$ , but these are sensitive to numerical details of the procedure. A truly *flat* distribution is difficult to reproduce, without the feedback provided by a Wang-Landau-like refinement.

## 4. Conclusions

In this paper we have highlighted some connections between some sampling methods that are intermediate in some sense between Monte Carlo and molecular dynamics. One general conclusion is that it is sometimes helpful to mix a little Monte Carlo into a molecular dynamics algorithm. We have not attempted to create a comprehensive list of methods of this kind, as the field is extremely diverse and fast moving. To keep track of these developments, and avoid re-inventing techniques, close contact between the mathematical, computational, and physical science communities is highly desirable. Such activities, of course, have been promoted over many years by the Centre Européen de Calcul Atomique et Moléculaire (CECAM), whose growth into an international network of nodes has been strongly motivated by Giovanni Ciccotti.

## Acknowledgements

We are happy to acknowledge our debt to Giovanni Ciccotti, who has long pioneered molecular dynamics methods as a viable alternative to Monte Carlo in a variety of circumstances. Indeed, one of us (MPA) is particularly indebted to the inspiration provided by Giovanni at a long CECAM workshop in Orsay, 1978, which gave him the chance to meet the experts in Brownian Dynamics, Smart Monte Carlo, constraint dynamics, and nonequilibrium dynamics.

## References

- [1] E. Cancès, F. Legoll and G. Stoltz, *ESAIM Math. Model. Num. Anal.* **41**, 351 (2007).
- [2] Legoll, Frédéric, Lelièvre, Tony and Stoltz, Gabriel, *ESAIM Proc.* **22**, 217 (2008).
- [3] N. Bou-Rabee and E. Vanden-Eijnden, *Commun. Pure Appl. Math.* **63** (5), 655 (2010).
- [4] D.P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics*, 3rd ed. (Cambridge University Press, Cambridge, 2009).
- [5] D. Frenkel and B. Smit, *Understanding molecular simulation: from algorithms to applications*, 2nd ed. (Academic Press, San Diego, 2002).
- [6] M.P. Allen and D.J. Tildesley, *Computer simulation of liquids*, paperback ed. (Clarendon Press, Oxford, 1989), p. 385.
- [7] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
- [8] W.K. Hastings, *Biometrika* **57**, 97 (1970).
- [9] S. Duane, A.D. Kennedy, B.J. Pendleton and D. Roweth, *Phys. Lett. B* **195**, 216 (1987).
- [10] W.C. Swope, H.C. Andersen, P.H. Berens and K.R. Wilson, *J. Chem. Phys.* **76**, 637 (1982).
- [11] S. Toxvaerd, *Phys. Rev. E* **50**, 2271 (1994).
- [12] H.C. Andersen, *J. Chem. Phys.* **72**, 2384 (1980).
- [13] E.V. Akhmatkaya and S. Reich, *J. Comput. Phys.* **227**, 4934 (2008).
- [14] E.V. Akhmatkaya, N. Bou-Rabee and S. Reich, *J. Comput. Phys.* **228**, 2256 (2009).
- [15] E.V. Akhmatkaya, N. Bou-Rabee and S. Reich, *J. Comput. Phys.* **228**, 7492 (2009).
- [16] P.J. Hoogerbrugge and J.M.V.A. Koelman, *Europhys. Lett.* **19**, 155 (1992).
- [17] J.M.V.A. Koelman and P.J. Hoogerbrugge, *Europhys. Lett.* **21**, 363 (1993).
- [18] R.D. Groot and P.B. Warren, *J. Chem. Phys.* **107**, 4423 (1997).
- [19] H.H. Rugh, *Phys. Rev. Lett.* **78**, 772 (1997).
- [20] B.D. Butler, G. Ayton, O.G. Jepps and D.J. Evans, *J. Chem. Phys.* **109**, 6519 (1998).
- [21] M.P. Allen, *J. Phys. Chem. B* **110**, 3823 (2006).
- [22] B. Hafskjold, C.C. Liew and W. Shinoda, *Molec. Simul.* **30**, 879 (2004).
- [23] A.F. Jakobsen, O.G. Mouritsen and G. Besold, *J. Chem. Phys.* **122**, 204901 (2005).
- [24] P.J. Rossky, J.D. Doll and H.L. Friedman, *J. Chem. Phys.* **69**, 4628 (1978).
- [25] D.L. Ermak, *J. Chem. Phys.* **62**, 4189 (1975).
- [26] A.M. Horowitz, *Phys. Lett. B* **268**, 247 (1991).

- [27] D.L. Ermak and J.A. McCammon, *J. Chem. Phys.* **69**, 1352 (1978).
- [28] J. Rotne and S. Prager, *J. Chem. Phys.* **50**, 4831 (1969).
- [29] M. Fixman, *Macromolecules* **19**, 1204 (1986).
- [30] A.J. Banchio and J.F. Brady, *J. Chem. Phys.* **118**, 10323 (2003).
- [31] C.H. Bennett, *J. Comput. Phys.* **19**, 267 (1975).
- [32] I.C. Lin and M.E. Tuckerman, *J. Phys. Chem. B* **114** (48), 15935 (2010).
- [33] A.P.E. Kunz, H. Liu and W.F. van Gunsteren, *J. Chem. Phys.* **135** (10) (2011).
- [34] S. Michielssens, T.S. van Erp, C. Kutzner, A. Ceulemans and B.L. de Groot, *J. Phys. Chem. B* **116** (29, SI), 8350 (2012).
- [35] L.B. Wright and T.R. Walsh, *Phys. Chem. Chem. Phys.* **15** (13), 4715 (2013).
- [36] B.A. Berg and T. Neuhaus, *Phys. Rev. Lett.* **68**, 9 (1992).
- [37] J. Lee, *Phys. Rev. Lett.* **71**, 211 (1993).
- [38] F.G. Wang and D.P. Landau, *Phys. Rev. E* **64**, 056101 (2001).
- [39] A. Laio and M. Parrinello, *Proc. Nat. Acad. Sci.* **99** (20), 12562 (2002).
- [40] C. Micheletti, A. Laio and M. Parrinello, *Phys. Rev. Lett.* **92** (17) (2004).
- [41] A. Laio and F.L. Gervasio, *Rep. Prog. Phys.* **71** (12), 126601 (2008).
- [42] E. Darve, D. Rodriguez-Gomez and A. Pohorille, *J. Chem. Phys.* **128** (14) (2008).
- [43] U.H.E. Hansmann, Y. Okamoto and F. Eisenmenger, *Chem. Phys. Lett.* **259**, 321 (1996).
- [44] N. Nakajima, H. Nakamura and A. Kidera, *J. Phys. Chem. B* **101**, 817 (1997).
- [45] E.J. Barth, B.B. Laird and B.J. Leimkuhler, *J. Chem. Phys.* **118**, 5759 (2003).
- [46] J. Kim, J.E. Straub and T. Keyes, *Phys. Rev. Lett.* **97**, 050601 (2006).
- [47] J. Kim, J.E. Straub and T. Keyes, *J. Chem. Phys.* **126**, 135101 (2007).
- [48] J.S. Lintuvuori and M.R. Wilson, *J. Chem. Phys.* **132** (22) (2010).
- [49] O.G. Jepps, G. Ayton and D.J. Evans, *Phys. Rev. E* **62**, 4757 (2000).
- [50] G. Rickayzen and J.G. Powles, *J. Chem. Phys.* **114**, 4333 (2001).
- [51] J.G. Powles, G. Rickayzen and D.M. Heyes, *Molec. Phys.* **103**, 1361 (2005).
- [52] K.P. Travis and C. Braga, *Molec. Phys.* **104**, 3735 (2006).
- [53] P.H. Nguyen, E. Mittag, A.E. Torda and G. Stock, *J. Chem. Phys.* **124**, 154107 (2006).
- [54] N. Rathore, T.A.K. IV and J.J. de Pablo, *Biophys. J.* **85**, 3963 (2003).