Comment on "Frequency dependence and equilibration of the specific heat of glass-forming liquids"

Elijah Flenner and Grzegorz Szamel
Department of Chemistry, Colorado State University, Fort Collins, Colorado 80525, USA
(Received 7 September 2004; published 19 August 2005)

In a recent publication C. C. Yu and H. M. Carruzzo [Phys. Rev. E **69**, 051201 (2004)] determined that the minimum sampling time to calculate the specific heat in a supercooled liquid using energy fluctuations is on the order of 10^3 α relaxation times, which is much longer than the sampling time used in most simulations. We demonstrate how the specific heat can be calculated from simulations run to 15 α relaxation times, and use statistical arguments to explain the systematic deviation of the specific heat calculated from a simulation of finite length from the true expectation value of the specific heat.

DOI: 10.1103/PhysRevE.72.023201 PACS number(s): 65.20.+w, 02.70.Ns, 05.20.−y

The minimum sampling time needed to calculate the specific heat in a computer simulation of a supercooled liquid was determined by Yu and Carruzzo [1] to be $10^2-10^3\alpha$ relaxation times. To determine the minimum sampling time needed to calculate the specific heat, they divided the simulation into blocks of size Δt_b, and calculated the specific heat using energy fluctuations from each block and found an average value of the specific heat for a block length Δt_b. The minimum sampling time was the value of Δt_b when the averaged specific heat vs Δt_b appeared to level off. In this Comment, we will show that the behavior of the block averaged specific heat versus block size can be explained using statistical arguments, and show how the specific heat can be calculated in simulations that are run for less than 15 α relaxation times.

The specific heat C can be calculated from a constant volume simulation N steps long using the equation

$$C_N = \frac{\langle U^2\rangle - \langle U\rangle^2}{k_B T^2}, \tag{1}$$

where $\langle\cdot\rangle$ denotes the thermal average, U is the potential energy, k_B is Boltzmann's constant, and T is the temperature. The specific heat calculated from Eq. (1) is a measure of the variance of the energy distribution. As described in many books on statistics [2], the expectation value of the variance σ_n^2 using n independent values underestimates the true value of the variance σ_∞^2, $\sigma_n^2=\sigma_\infty^2(1-1/n)$. Thus the average expectation value of the specific heat calculated for a simulation of finite length will be less than the true expectation value of the specific heat and depends on the effective number of statistically independent measurements of the energy which are made in the simulation. If we know the effective number of statistically independent measurements n made in a simulation of length N, then we can calculate the true expectation value of the specific heat

$$C_\infty = C_N\left(\frac{n}{n-1}\right). \tag{2}$$

In this Comment we will describe how these concepts can be used to calculate the true expectation value of the specific heat for simulations much shorter than 10^3 α relaxation times, and provide an explanation of the behavior of the block averaged specific heat versus block size observed in Ref. [1]. Also, we give a description of the error analysis that incorporates the effective number of statistically independent measurements.

To calculate the specific heat, we used the trajectories generated by a large (up to 6×10^8 time steps) Brownian dynamics simulation, which is described in detail in Ref. [3]. Briefly, we performed simulations of the Kob-Anderson [4] binary Lennard-Jones mixture at temperatures $T=0.44$, 0.45, 0.47, 0.50, 0.55, 0.60, 0.80, and 1.0. Several production runs were performed at each temperature after a long equilibration run. Care was taken to ensure equilibrium. The results presented in this Comment are at temperatures in which equilibrium is possible to achieve on a modern computer. For reference, the mode coupling transition temperature $T_c=0.435$ [4]. The α relaxation time τ_α is defined as the time when the incoherent intermediate scattering function ($q=7.25$) for the larger particles is equal to e^{-1}. While the simulated system is different from the one in Ref. [1], we believe that the conclusions of this work are system independent.

The calculation of the true specific heat [Eq. (2)] depends on knowing the effective number of statistically independent measurements of the energy. This problem was solved for Monte Carlo simulations by Müller-Krumbhaar and Binder [5]. They determined that the number of effectively independent measurements n from N correlated measurements taken at intervals Δt is approximately $n=N/[2\tau/\Delta t+1]$, where

$$\tau = \int_0^\infty dt \frac{\langle U(t)U(0)\rangle - \langle U(0)\rangle^2}{\langle U^2(0)\rangle - \langle U(0)\rangle^2} \tag{3}$$

is the integrated energy correlation time, and $U(t)$ is the potential energy at time t. We assumed that the number of effectively independent measurements in a Brownian dynamics simulation can be calculated in the same manner. Since the simulations are of finite length, the integral is truncated at $t_{\max}$, which must be chosen such that the integrand in Eq. (3) has decayed to almost zero at $t_{\max}$. This is not a problem for simulations performed beyond one α relaxation time.

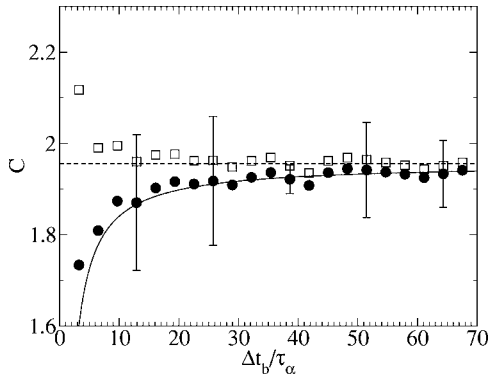

FIG. 1. Block averaged specific heat (•) and the corrected block averaged specific heat (□) compared to $C_\infty(1-1/n)$ (solid line) for $T=0.5$. The correction is given by Eq. (2) where the number of effectively independent measurements $n=\Delta t_b/[2\tau+\Delta t]$, τ is given by Eq. (3) and Δt is the time between stored configurations. C_∞ (dashed line) is calculated using Eq. (2) where C_N is determined using the whole simulation. τ_α is the α relaxation time.

To test the ideas presented above, we examined the behavior of the block averaged specific heat versus block size. The block averages are found by dividing the simulation into blocks of equal length Δt_b, and calculating the specific heat for each block using Eq. (1). Then we calculated the average value of the specific heat $\bar{C}$ for each block size Δt_b. Shown in Fig. 1 is the block averaged specific heat (closed symbols) for one of the production runs at $T=0.5$ and $C_{\Delta t_b}=C_\infty(1-1/n)$ (solid line). The error bars are the standard deviation $\sigma=[(1/m-1)\Sigma_{i=1}^{m}(C_i-\bar{C})^2]^{1/2}$ of the block averages, and do not represent the relative error in the specific heat (see below). The large value of σ for small Δt_b is due to a large variation in the specific heat. For larger Δt_b, the variation in the calculated specific heat is less, which results in a smaller σ. For the largest values of Δt_b, m is small thus σ is large. The dashed horizontal line is C_∞ calculated using Eq. (2), where we determined C_N using the whole simulation. Also shown are the block averages corrected for the block size (open symbols). The average integrated energy correlation time is close to 1/2 the average α relaxation time for $T<0.8$, and less than 1/2 the α relaxation time for $T\geqslant 0.8$, which suggests that there is at least one independent measurement of the energy every one α relaxation time. Moreover the temperature dependence of the integrated energy correlation time was the same as the α relaxation time for $T<0.8$.

Equation (2) fits the block averages very well. Also, the corrected block averaged specific heat levels and fluctuates around C_∞ beginning at some time before 15 α relaxation times. If the criterion used in Ref. [1] to determine the minimum sampling time (i.e., leveling off of the C vs Δt_b plot) is applied to the corrected specific heat, then we can conclude that the minimum sampling time needed to calculate the specific heat is on the order of 15 α relaxation times. This is at least one order of magnitude less than the 10^2-10^3 α relaxation times suggested by Ref. [1]. Examination of other temperatures and other runs lead to the same conclusion.

These ideas can be extended [5,6] to estimate the uncertainty in the specific heat ΔC. As shown by Ferrenberg *et al.* [6], for large n the relative statistical error

$$r=\frac{\Delta C}{C}=[(2-3V_U)/n]^{1/2}, \tag{4}$$

where

$$V_U=1-\frac{\langle U^4\rangle-4\langle U^3\rangle\langle U\rangle+6\langle U^2\rangle\langle U\rangle^2-3\langle U\rangle^4}{3(\langle U^2\rangle-\langle U\rangle^2)^2}. \tag{5}$$

It can be shown that V_U is the negative of the ratio of the fourth cumulant to three times the second cumulant squared. Thus, if the probability distribution of the energy is Gaussian, then $V_U=0$. In practice, the energy distribution is not perfectly Gaussian, but $|V_U|<0.01$ for the simulations described in this Comment. Hence, the relative error $r\approx\sqrt{2/n}$. To get the relative error to within 10%, simulations need to be run around 200 α relaxation times. We conclude that the simulation time needed to calculate the specific heat is not orders of magnitude longer than the α relaxation time, but the simulation time needed to calculate the specific heat *accurately* may be orders of magnitude longer than the α relaxation time.

We have shown that the minimum sampling time to calculate the specific heat is on the order of 15 α relaxation times. We used a common definition of the α relaxation time that is similar to the definition used by Yu and Carruzzo so that we could compare our results with their work. Moreover, we have observed that the minimum sampling time has the same temperature dependence as the α relaxation time close to the mode coupling transition temperature. To determine the equilibrium value of the specific heat, the correction given in Eq. (2) must be made, which depends on the number of effectively independent measurements. We used Eq. (3) to determine the integrated energy correlation time, which is then used to determine the number of effectively independent measurements. It should be noted that Eq. (3) assumes exponential decay of the energy auto correlation function that may or may not be valid below the mode coupling transition temperature. Therefore, the method used to obtain the number of effectively independent measurements may need modification for some simulations at some temperatures. However, the method we have outlined is not computationally expensive and it is easy to determine its validity by examining graphs like Fig. 1. Furthermore, the main conclusion is still true. Simulations of finite length underestimates any estimate of the specific heat that is calculated by the use of energy fluctuations, but a correction can be made if the number of effectively independent measurements are known.

[1] C. C. Yu and H. M. Carruzzo, Phys. Rev. E **69**, 051201 (2004).
[2] O. Dunn and V. Clark, *Applied Statistics* (John Wiley & Sons, New York, 1974).
[3] G. Szamel and E. Flenner, Europhys. Lett. **67**, 779 (2004).
[4] W. Kob and H. C. Andersen, Phys. Rev. E **51**, 4626 (1995); **52**, 4134 (1995).
[5] H. Müller-Krumbhaar and K. Binder, J. Stat. Phys. **8**, 1 (1973).
[6] A. M. Ferrenberg, D. P. Landau, and K. Binder, J. Stat. Phys. **64**, 867 (1991).