JOURNAL OF COMPUTATIONAL PHYSICS 21, 227–237 (1976)

Monte Carlo Simulation of Molecular Fluids on a Minicomputer

J. M. HAILE,* D. LITCHINSKY, R. MCPHERSON, C. G. GRAY

Physics Department, University of Guelph, Guelph, Ontario

AND

K. E. GUBBINS

Chemical Engineering Department, University of Florida, Gainesville, Florida 32611

Received December 4, 1975

The Monte Carlo simulation of molecular liquids using a minicomputer is described. The liquids studied are composed of nonspherical molecules interacting with dipole, quadrupole and anisotropic overlap forces. Results are reported for the angular pair distribution function, the molecular centers distribution function, the configuration energy, the mean squared force, and the mean squared torque, for a system of 128 strongly quadrupolar molecules. The results agree very well with previous Monte Carlo results for such a system obtained on a CDC 6600.

1. INTRODUCTION

The computer simulation of fluids by Monte Carlo and molecular dynamics provides essentially exact results for the model fluid under consideration. Thus, these methods have become of primary importance in the study of fluids, complementing the traditional approach of comparing fluid theory predictions with experimental results. Unfortunately, the use of computer simulations has become of major importance at a time when the availability of large machines for performing such studies seems to be greatly reduced. Hence, we felt it worthwhile to determine the feasibility of programming a minicomputer for such simulation work.

This paper reports success in programming a NOVA 2 minicomputer with 32K 16 bit-words of memory and external disc storage for performing Monte Carlo calculations on systems of 128 nonspherical molecules. While execution time is slow on this particular machine (a sequence of 10^5 configurations for 128 linear molecules requires about 125 hr of computation), we feel the yield justifies the

* Permanent address: Chemical Engineering Department, University of Florida, Gainesville, Florida.

effort since from one run we obtain nine equilibrium property values, the value of the center-to-center pair correlation function, $g_c(r_{12})$, and values for the angular pair correlation function $g(\mathbf{r}_{12}\omega_1\omega_2)$ at five specific orientations. Furthermore, there are other minicomputers on the market which should execute the present program faster by a factor of two or three. We have found the NOVA 2 to be reliable over days of continuous operation and results obtained are comparable, within the statistical precision, with results previously obtained from a similar program executed on a CDC 6600.

In the sections which follow we describe: (a) the Monte Carlo problem for molecular fluids, (b) the NOVA 2 and peripheral configuration employed, (c) the Fortran program developed, (d) preliminary results obtained, (e) conclusions, including options for improving execution time.

2. Monte Carlo Method for Molecular Fluids

The Monte Carlo procedure employed here is the same, in principle, as that for systems of spherical molecules and is well documented elsewhere [1, 2]. The primary additional consideration is the orientation of the nonspherical molecules. In general, properties of fluids of nonspherical molecules depend on molecular orientation as well as the relative distances between molecules. The Monte Carlo program must, therefore, maintain the location and orientation of each molecule in the system. The Monte Carlo procedure for simple molecules is modified, then, by reorienting the nonspherical molecule in addition to its relocation when a new system configuration is proposed.

The potential models which we have studied for linear molecules are of the form:

$$u(\mathbf{r}_{12}\omega_1\omega_2) = u_{LJ}(r_{12}) + u_a(\mathbf{r}_{12}\omega_1\omega_2), \tag{1}$$

where u is the full pair potential between molecules 1 and 2 which are separated by a distance r_{12}, having orientations ω_1 and ω_2, respectively; u_{LJ} is the Lennard–Jones pair potential:

$$u_{LJ}(r_{12}) = 4\epsilon[(\sigma/r_{12})^{12} - (\sigma/r_{12})^6], \tag{2}$$

u_a is the anisotropic part of the potential.

The anisotropic potentials which have been considered include dipole, quadrupole and anisotropic overlap models, given by (3)–(5):

$$u_{\mu u} = (\mu^2/r_{12}^3)(s_1s_2c - 2c_1c_2), \tag{3}$$

$$u_{QQ} = (3Q^2/4r_{12}^5)[1 - 5(c_1{}^2 + c_2{}^2 + 3c_1{}^2c_2{}^2) + 2(s_1s_2c - 4c_1c_2)^2] \tag{4}$$

$$u_{ov} = 4 \in \delta(\sigma/r)^{12}(3c_1{}^2 + 3c_2{}^2 - 2). \tag{5}$$

Here $c_i = \cos\ \theta_i$, $s_i = \sin\ \theta_i$, $c = \cos\ \phi$, where θ and ϕ are the Euler angles relative to a frame whose polar axis is along the line of molecular centers; μ and Q are the dipole and quadrupole moment, respectively, and δ is a dimensionless overlap parameter.

The minimum number of variables required to specify the location and orientation of a molecule depends, of course, on its shape. For linear molecules, the only shape considered here, the minimum number is five, e.g., three components of a location vector for the molecular center of mass, $\mathbf{r}$, and two Euler angles $\omega = \{\alpha, \beta\}$, between a body-fixed axis and the space-fixed reference frame. If the orientations are actually specified by Euler angles $\{\alpha, \beta\}$, determination of the relative angles θ, ϕ for use in Eqs. (3–5) is computationally slow. Alternatively, if the orientations of N molecules are given by the relative angles $\{\theta, \phi\}$, then $\frac{1}{2}(N)$ $(N - 1)$ pairs of angles must be stored throughout the calculation. As a compromise on speed and storage, we specify the orientation of a molecule by the direction cosines of the molecular axis, from which the cosines of the relative angles $\{\theta, \phi\}$ may be determined by simple dot product relations. Thus, we actually store six variables for each linear molecule, the three components for the center of mass and the three direction cosines for the molecular axis (only two of which are independent).

Another consideration in studying nonspherical molecules is how to determine property values from the simulation. There are two possible methods: (a) use the Monte Carlo calculation to form a histogram of the full angular pair correlation function $g(\mathbf{r}_{12}\omega_1\omega_2)$. Properties can then be obtained from the standard statistical mechanical relations which involve numerical integration over the $g(\mathbf{r}_{12}\omega_1\omega_2)$ histogram. (b) The second method uses the Monte Carlo procedure to directly perform the integrations for the properties. Having determined properties by both methods, we find method (b) to be superior to method (a). Accuracy in the determination of properties by method (a) requires small angular increments in ω_1 and ω_2 which, in turn, requires large amounts of computer storage and makes for long execution times. The second method is simpler to accomplish, more accurate and is readily performed on small computers. The method (a) does have the advantage that only one long Monte Carlo is run from which property values may be estimated.

3. Description of the Minicomputer System

There are a number of reasons for using minicomputers for scientific calculations. Large computers have the advantage where word size, speed and input–output (IO) complexity are concerned; however, where cost is concerned the minicomputer has a definite advantage. For example, Kottler and McGill [3] quote a cost advantage of between 80:1 and 100:1 for a Data General 800 system with hardware floating point multiply/divide when compared to an IBM 360/75. These figures are dependent on the type of program being compared but serve to illustrate the point. This cost advantage must be weighted against increased computing time, the reliability of the minicomputer over long operating periods and less sophistication. Particularly in situations where pure number crunching capability with high CPU attention is required, such as Monte Carlo calculations, the cost advantage of the minicomputer is expected to be high, although no detailed costing was done in the present application.

The computer used here was a Data General Nova 2/10 with optional 32K 16 bit-words of core storage and 1000*n* sec memory cycle time. The CPU included the following options: power monitor, auto restart, auto program load, real time clock, hardware multiply/divide, high performance hardware floating point processor. The following peripherals were serviced by the CPU: (a) two moving head disc units with controller for a total 2.5 megawords of storage, (b) serial matrix line printer (165 cps), (c) fast paper tape reader and punch, (d) teletype unit.

The computer was operated under the Data General real time operating system, RDOS revision 3.02, which handles task scheduling and system maintenance. The executive remained core resident and occupied about 3K words so that 29K word locations were available for computation. The software available included Data Generals' Fortran IV, Fortran V, Basic, and Algol. The Fortran V compiler was chosen as the medium for compilation. This compiler is a code optimizing compiler, i.e., redundant operations are eliminated and floating point operations are optimized for effective use of the floating point hardware. This leads to a very compact coding of the source program.

4. Description of Fortran Program

The major time-consuming calculation in Monte Carlo programs is evaluation of the system energy U, which is assumed to obey pairwise additivity, i.e.,

$$U = \sum_{i<j}^{N} \sum u(\mathbf{r}_{ij}\omega_i\omega_j). \tag{6}$$

A substantial saving in execution time is realized by taking advantage of the fact that when one particle k, in the system is moved, only $(N - 1)$ pair energies change. Then, if the energy of the "old" configuration (the configuration prior to the move) is known, the energy of the "new" configuration can be found as the sum of the old energy plus the change in energy due to movement of the kth particle, i.e., Eq. (6) may be written:

$$U^{\text{new}} = U^{\text{old}} + \Delta U_k\,, \tag{7}$$

where

$$\Delta U_k = \sum_{j \neq k}^{N} u(\mathbf{r}_{kj}^{\text{new}} \omega_k^{\text{new}} \omega_j) - \sum_{j \neq k}^{N} u(\mathbf{r}_{kj}^{\text{old}} \omega_k^{\text{old}} \omega_j). \tag{8}$$

In the usual Monte Carlo calculation each of the pair energies $u(\mathbf{r}_{ij}\omega_i\omega_j)$ for the starting configuration is calculated and stored and the total system energy is formed by Eq. (6). Subsequently, for each configuration generated, the system energy is determined by equations (7) and (8), i.e., when the kth particle is moved, each of the $(N - 1)$ pair energies $u(\mathbf{r}_{kj}^{\text{new}}\omega_k^{\text{new}}\omega_j)$ are calculated, the corresponding $u(\mathbf{r}_{kj}^{\text{old}}\omega_k^{\text{old}}\omega_j)$ are obtained from storage, and equations (7) and (8) are solved for U^{new}. If the new configuration is accepted, the pair energies $u(r_{kj}^{\text{old}}\omega_k^{\text{old}}\omega_j)$ in storage are updated to the new values $u(\mathbf{r}_{kj}^{\text{new}}\omega_k^{\text{new}}\omega_j)$. This obviously decreases the amount of calculation required, but places a heavy demand on core storage. The matrix of pair energies is symmetric, i.e.,

$$u(\mathbf{r}_{ij}\omega_i\omega_j) = u(\mathbf{r}_{ji}\omega_j\omega_i), \tag{9}$$

even so, $\frac{1}{2}N(N - 1)$ numbers must be stored just for the energy. If other properties are also being calculated, similar storage must be provided for each of them.

We avoid this large storage requirement in the program for the NOVA by calculating both sums in Eq. (8) for each configuration generated, thus storing none of the individual pair properties. Execution time is increased accordingly. Alternatively, one could store the pair energy matrix on disc and transfer it into core column by column as it is needed. When a move is accepted, however, a major portion of the matrix must be brought through core in order to update the changed elements from $u(\mathbf{r}_{kj}^{\text{old}}\omega_k^{\text{old}}\omega_j)$ to $u(\mathbf{r}_{kj}^{\text{new}}\omega_k^{\text{new}}\omega_j)$. On the NOVA 2 system we found this transfer to be slower than recalculating the elements of the matrix as indicated above. Hence, our calculation does not depend upon core–disc transfers during execution. The evaluation of both sums in Eq. (8) for each configuration is the only true modification we have made in the Monte Carlo program to transfer its use from a CDC to minicomputer environment.

Figure 1 shows a simplified schematic flow diagram of the program used on the NOVA 2. The Monte Carlo calculation is done by the main program labeled 1, which is supported by several subroutines, the two principal ones being named INITIAL and ENERGY. Subroutine ENERGY consumes the major portion of execution time as it forms the sums analogous to Eq. (8) for all properties of interest. Also indicated in the figure is the fact that every 30 min or so, corresponding to M number of moves, the program dumps current particle location vectors, direction cosines and property values onto disc file 2. Disc file 1 holds the starting configuration's location vectors and direction cosines. The periodic data dump serves as a back-up against possible power failure and allows

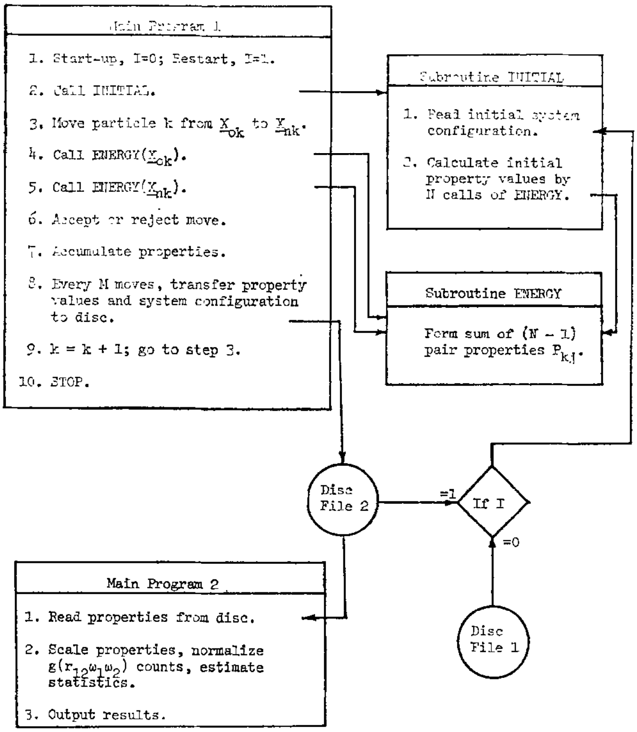
Main Program 1
1. Start-up, I=0; Restart, I=1.
2. Call INITIAL.
3. Move particle k from $\underline{X}_{ok}$ to $\underline{X}_{nk}$.
4. Call ENERGY($\underline{X}_{ok}$).
5. Call ENERGY($\underline{X}_{nk}$).
6. Accept or reject move.
7. Accumulate properties.
8. Every M moves, transfer property values and system configuration to disc.
9. k = k + 1; go to step 3.
10. STOP.
Subroutine INITIAL
1. Read initial system configuration.
2. Calculate initial property values by N calls of ENERGY.
Subroutine ENERGY
Form sum of (N – 1) pair properties P_{kj}.
Disc File 2
=1
If I
=0
Disc File 1
Main Program 2
1. Read properties from disc.
2. Scale properties, normalize $g(r_{12}\omega_1\omega_2)$ counts, estimate statistics.
3. Output results.

FIG. 1. Simplified Schematic Flow Diagram of Fortran Monte Carlo program developed for NOVA 2.

other system users to interrupt program execution. This entire program resides in core throughout the calculation and required about 27K words of NOVA 2 memory. A further saving of 3–4K words could be attained by making the subroutime INITIAL a separate main program, as it is only used at the initial start (or restart) of a calculation. The main program labeled 2 takes the raw data generated by program 1, reduces it to the desired form and estimates the statistical precision of the results.

Execution time for the Monte Carlo program depends largely on the number of particles in the system, to some extent on the complexity of the potential model under study, and only slightly on the number of properties being calculated. These findings are summarized in Table I which compares the number of configurations generated per hour by the Fortran program for 64 and 128 particle systems for the simple Lennard–Jones fluid and the Lennard–Jones plus quadrupole model of Eqs. (1) and (4). It can be seen from the table that doubling the number of particles roughly doubles execution time, whereas changing the number of proper-

TABLE I

Approximate Number of Configurations Generated per Hour on Nova 2

		$LJ + QQ$		
			No. $g(\mathbf{r}_{12}\omega_1\omega_2)$ tested	
No. particles	LJ	No. properties	1	5
64	3025	1	2153	1846
		8	1912	1671
128	1616	1	1131	997
		8	1003	893

ties calculated has much less effect. Additionally, we find the dipole model of Eqs. (1) and (3) to execute about 2 % faster than the quadrupole fluid, and the overlap model of equations (1) and (5) to be about 2 % faster than the dipole model. These program execution times of the NOVA 2 may be compared with speeds attained on IBM and CDC machines. A Fortran program for simulating the overlap model fluid using 64 particles and calculating only the system internal energy and seven values for the angular pair correlation function generates about 20,000 configurations per hour on an IBM 370/155 and about 106,000 configurations per hour on a CDC 6600.

5. Preliminary Results

The program described above calculates the following canonical averages for systems of up to 128 molecules interacting with potentials of Eqs. (1), (2), and (3) or (4) or (5):

$$\Big\langle \sum_{i<j}\sum u(\mathbf{r}_{ij}\omega_i\omega_j)\Big\rangle, \qquad \Big\langle \Big(\sum_{i<j}\sum u(\mathbf{r}_{ij}\omega_i\omega_j)\Big)^2\Big\rangle,$$

$$\Big\langle \sum_{i<j}\sum r_{ij}u(\mathbf{r}_{ij}\omega_i\omega_j)\Big\rangle, \qquad \Big\langle \sum_{i<j}\sum r_{ij}\frac{\partial u(\mathbf{r}_{ij}\omega_i\omega_j)}{\partial r_{ij}}\Big\rangle,$$

$$\Big\langle \sum_{i<j}\sum r_{ij}^2\frac{\partial u(\mathbf{r}_{ij}\omega_i\omega_j)}{\partial r_{ij}}\Big\rangle, \qquad \Big\langle \sum_{i<j}\sum \nabla_1^2 u_{LJ}(r_{ij})\Big\rangle,$$

$$\Big\langle \sum_{i<j}\sum \cos\gamma_{ij}\Big\rangle, \qquad \Big\langle \sum_{i<j}\sum \cos^2\gamma_{ij}\Big\rangle,$$

where γ_{ij} is the angle between the axes of molecules i and j. From these avarages one may obtain the following equilibrium properties: the configurational internal energy, U^c, the pressure, the constant volume heat capacity, C_v, the Fowler model surface tension γ^{fm}, the Fowler model surface excess energy, E_s^{fm}, the mean squared force on a molecule, $\langle F_1^2 \rangle$, the mean squared torque on a molecule, $\langle \tau_1^2 \rangle$, and the Kirkwood angular correlation parameters G_1 and G_2. In addition, the program determines the center-to-center pair correlation function $g_c(r_{12})$ and the angular pair correlation function $g(\mathbf{r}_{12}\omega_1\omega_2)$ at five specific orientations.

TABLE II

Comparison of NOVA 2 and CDC results for Property values of $LJ + QQ$ model fluid at $kT/\epsilon = 0.719$, $\rho\sigma^3 = 0.80$, $Q/(\epsilon\sigma^5)^{1/2} = 1$

Property	NOVA 2	CDC 6600
$U^c/N\epsilon$	-8.642 ± 0.197	$-8.483^a \pm 0.131$
$\langle F_1^2 \rangle/(\epsilon/\sigma)^2$	744.0 ± 34.0	$772.0^b \pm 15.0$
$\langle \tau_1^2 \rangle/\epsilon^2$	31.2 ± 1.09	$31.0^c \pm 1.5$

[a] See [4].
[b] See [5].
[c] See [6].

In Table II we compare results for linear quadrupoles obtained from the program run on the NOVA 2 with results previously obtained on a CDC 6600 [4–6]. The values for U^c, $\langle F_1^2 \rangle$ and $\langle \tau_1^2 \rangle$ reported in references [4–6] do not include a long-range correction for values of the radial component r greater than the value at which the potential is set to zero, r_{cutoff}. For the sake of consistency in the comparison we have included no long-range correction on the minicomputer values shown in Table II. Further, the values for the mean-squared force $\langle F_1^2 \rangle$ and mean-squared torque $\langle \tau_1^2 \rangle$ reported in [5, 6] were obtained by integration over the Monte Carlo-determined correlation function as described under method (a), in Section 2 above. The values for these two properties from the NOVA run are, therefore, more reliable since they were obtained by method (b) of Section 2.

Figure 2 compares values obtained from the NOVA 2 with previously obtained values [4] for the center-to-center pair correlation function, $g_c(r_{12})$, for a $LJ + QQ$ model fluid. The earlier results were determined for a system of 64 linear molecules from a chain of some 80,000 configurations after the system had reached equilibrium. The NOVA 2 results are from a run of similar length on a system of 128 molecules, hence, the extended range of $g_c(r_{12})$ indicated in the figure.

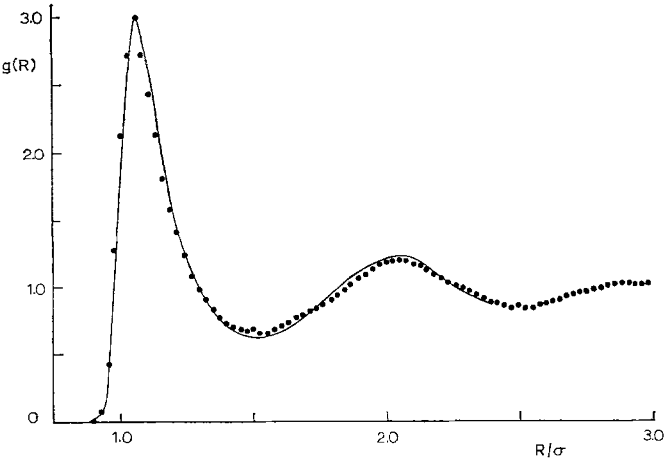

FIG. 2. The center-to-center pair correlation function, $g_c(r_{12})$, comparing CDC (line) [4] and NOVA 2 (points) results, for $LJ + QQ$ at $kT/\epsilon = 0.719$, $\rho\sigma^3 = 0.80$, $Q/(\epsilon\sigma^5)^{1/2} = 1$.

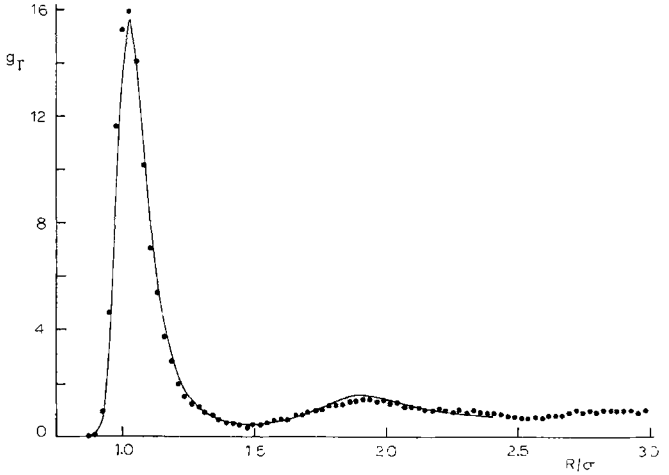

FIG. 3. The angular pair correlation function $g(\mathbf{r}_{12}\omega_1\omega_2)$ for $LJ + QQ$ fluid for molecular pairs in the "T" orientation, comparing CDC (line) [7] and NOVA 2 (points) results at $kT/\epsilon = 0.719$, $\rho\sigma^3 = 0.80$, $Q/(\epsilon\sigma^5)^{1/2} = 1$.

Figure 3 shows a similar comparison for the angular pair correlation function $g(\mathbf{r}_{12}\omega_1\omega_2)$ for the model fluid $LJ + QQ$ for molecular pairs in the "T" configuration, i.e., $\theta_1 = \pi/2$, $\theta_2 = 0$, ϕ_{12} undefined, where θ_1, θ_2, and ϕ_{12} are the relative angles defined under equation (5) above. The earlier results are from a chain of about 410,000 configurations for 64 particles generated on the CDC 6600 [7]. To achieve consistency in the comparison, the same angular increment of ±15° was used in the minicomputer calculation as was used in [7]. The agreement in Fig. 3 is remarkable considering the minicomputer run was only one-fifth as long as the CDC calculation.

6. Conclusions

The results shown above indicate that the NOVA 2 is capable of reliable, sustained operation of sufficient magnitude to produce useful results by the Monte Carlo method for simulating fluids. The only limitation found in this feasibility study is that of speed of program execution. This, in turn, imposes a limitation on the class of problems for which one might use a minicomputer. Probably an upper bound on the time one would be willing to invest in a single run is of the order of two weeks of computer time. On the fastest minicomputer available this would yield 5–10 (10^5) configurations of a Monte Carlo sequence, depending on the number of particles, potential model, etc. Thus, one would not consider using present minicomputers for specialized studies which require significantly longer runs than calculation of bulk fluid properties, such as the study of phase transitions. Though the class of problems is restricted, it is still large, and we feel the minicomputer provides a viable alternative which one might seriously consider, especially in light of present economic constraints.

There are a few options which can be considered in connection with one's own particular problem of interest which would improve execution time. We have already mentioned use of faster hardware than is avaialable on the NOVA 2. A second possibility would be a careful machine coding of the part of the program which we refer to as ENERGY (Fig. 1). A third consideration is the shape of the unit cell used in the Monte Carlo procedure. The program described here does *not* use the usual cube-shaped cell but rather a truncated dodecahedron after Wang and Krumhansl [8]. Since our interests include such long range properties as pressure, and since we also wish so obtain as full a description of the distribution functions $g(\mathbf{r}_{12}\omega_1\omega_2)$ as possible, we need a large value for the cutoff distance of the potential, r_{cutoff}. By considering shapes other than the cube, the value for r_{cutoff} may be extended (see reference [8]). Thus, the value for r_{cutoff} for a system of N particles using the truncated dodecahedron is the same as for a system of $N2^{1/2}$ particles using a cube. The program using the truncated dodecahedron with N particles is about 25% slower than the same program using a cube with N particles; whereas a program using a cube with $N2^{1/2}$ particles is about 40% slower than the same program using a cube with N particles. Many users may not be interested in long range effects, however, in which case use of a cubic cell would improve execution time accordingly.

Acknowledgments

We are grateful to S. S. Wang for many useful discussions on the Monte Carlo method and to I. R. McDonald for clarifying several details about Monte Carlo. Two of us (JMH and KEG) would like to thank P. A. Egelstaff and the staff of the Physics Department, University of Guelph for their hospitality during the time when this work was done. This work was supported by grants for the Petroleum Research Fund of the American Chemical Society, the National Science Foundation, the National Research Council of Canada, and the American Gas Associaton.

References

1. W. W. Wood, *in* "Physics of Simple Liquids," (H. N. V. Temperley, J. S. Rowlinson, G. S. Rushbrooke, Eds.), Chap. 5. North-Holland, Amsterdam, 1968.
2. I. R. McDonald and K. Singer, *Q. Rev. Chem. Soc.* **24** (1970), 238.
3. C. F. Kottler and R. McGill, *Instrum. Control Sys.* **46** (1973), 57.
4. S. S. Wang, C. G. Gray, P. A. Egelstaff, and K. E. Gubbins, *Chem. Phys. Lett.* **21** (1973), 123.
5. C. G. Gray, S. S. Wang, and K. E. Gubbins, *Chem. Phys. Lett.* **26** (1974), 610.
6. C. H. Twu, K. E. Gubbins, and C. G. Gray, *Molec. Phys.* **30**, 1607 (1975).
7. S. S. Wang, P. A. Egelstaff, C. G. Gray, and K. E. Gubbins, *Chem. Phys. Lett.* **24** (1974), 453.
8. S. S. Wang and J. A. Krumhansl, *J. Chem. Phys.* **56** (1972), 4287.