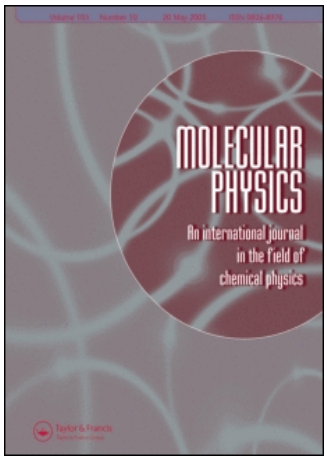

Molecular Physics
An International Journal in the Field of Chemical Physics
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713395160

Virial-bias Monte Carlo methods

Mihaly Mezei [a]
[a] Department of Chemistry, Hunter College of the CUNY. New York, N.Y.. U.S.A.

Online Publication Date: 10 April 1983
To cite this Article: Mezei, Mihaly , (1983) 'Virial-bias Monte Carlo methods', Molecular Physics, 48:5, 1075 - 1082
To link to this article: DOI: 10.1080/00268978300100761
URL: http://dx.doi.org/10.1080/00268978300100761

MOLECULAR PHYSICS, 1983, VOL. 48, No. 5, 1075–1082

Virial-bias Monte Carlo methods
Efficient sampling in the (T, P,N) ensemble

by MIHALY MEZEI
Department of Chemistry, Hunter College of the CUNY,
New York, N.Y. 10021, U.S.A.

(*Received* 5 *August* 1982 ; *accepted* 25 *October* 1982)

A new sampling technique is proposed for the isothermal–isobaric (T, P, N) ensemble Monte Carlo computer simulation. The new method selects the volume perturbations by using the partial derivative of the volume with respect to the internal energy. Test calculations on the Lennard-Jones fluid show significant improvements over the conventionally used method.

1. INTRODUCTION AND BACKGROUND

Computer simulation of liquids by the Metropolis method has recently become a frequently used tool of the liquid state theorist. The method was originally introduced in the (T, V, N) ensemble [1], but it can be generalized directly for use in the (T, P, N) [2, 3] or (T, V, μ) [4, 5] ensembles. The (T, P, N) ensemble method was formalized by Wood and used for hard spheres and Lennard-Jones particles. It was also applied successfully to molecular liquids by Owicki and Scheraga [6] and Jorgensen [7]. The convergence characteristics of the method were studied recently on liquid water by Jorgensen [8].

The long runs necessary for achieving sufficiently converged results prompted several improvements in the sampling techniques. The framework of possible modifications of the Metropolis method was given by Hastings [9]. Generally, use of some detailed information about the system allows one to replace a uniform sampling distribution for a given parameter by a well defined new distribution. The choice of the parameter and the detailed information to be used characterizes the method. For example, in the gradient-bias methods [10–12], the displacement vectors of a given particle are sampled from a distribution that gives higher weight to move along the direction of the force and the cavity-bias (T, V, μ) method attempts insertion of a new particle into cavities only [13].

The purpose of the present paper is to propose an improved sampling distribution for the volume change in the (T, P, N) ensemble simulation. Model calculations are presented for the Lennard-Jones fluid to demonstrate the improvements achieved.

2. THEORY

Simulation in the (T, P, N) ensemble consists of two different steps : perturbation of a particle and perturbation of the volume. The perturbation of a

particle is done the same way as in the (T, V, N) ensemble and is not discussed here in detail. In the original realization of the (T, P, N) ensemble simulation (in the following : TPN) the volume change δV is chosen uniformly from the interval

$$-\Delta V \leqslant \delta V \leqslant \Delta V \tag{1}$$

and the new configuration is accepted with probability

$$P_{\text{acc}} = \min\{1, \exp[(U_o - U_o - P\delta V + kTN(\ln V_n - \ln V_o))/kT]\}. \tag{2}$$

Here N is the number of particles in the system, k is the Boltzmann factor, T is the absolute temperature, P is the external pressure specified and U is the configurational internal energy of a configuration. The subscripts o and n refer to the configurations before and after the perturbation, respectively. The last term in the exponential is equivalent to a factor of $(V_n/V_o)^N$. It enters the formalism through the transformation to scaled coordinates that avoids integration over changing volume.

The improved sampling is developed by analogy with the improved displacement techniques mentioned above. For displacements, the force acting on a particle was used to obtain a sampling distribution for the particle displacements. For volume perturbation, the quantity that is the analogue of the the force acting on a particle is

$$F_{\text{v}} = \frac{\partial}{\partial V}(U + PV - kTN \ln V) = \frac{\partial U}{\partial V} + P - kTN/V. \tag{3}$$

For systems with pairwise additive potentials u, the volume derivative can be derived in the same way as the pressure equation is obtained [14] :

$$\frac{\partial U}{\partial V} = 1 \Big/ \left[3 \sum_{i<j}^{N} (\mathbf{r}_i - \mathbf{r}_j) \cdot \boldsymbol{\Delta}_i u(\mathbf{r}_i, \mathbf{r}_j) \right] + 2\pi\rho u(R_{\text{c}}) g(R_{\text{c}})/3, \tag{4}$$

where R_{c} is the cut-off applied of the potential, ρ is the liquid density and $g(R)$ is the radial distribution function. The second term in (4) is due to the finite, discontinuous cut-off usually applied to the potential.

In the choice of the sampling distribution for δV, prior experience in sampling the coordinate perturbation can give useful guidance. Two sampling distributions were tried. The first was the direct analogy of the force-bias method of Pangali, Rao and Berne [11] and the second followed the ideas of Rossky, Doll and Friedman [12]. In the first case, that we call exponential virial-bias (T, P, N) method (in the following : EVB/TPN), the sampling distribution was chosen as

$$\exp[\lambda \delta V F_{\text{v}}/kT]/n(V), \tag{5}$$

where $n(V)$ is a normalization constant :

$$n(V) = 2kT \sinh(\lambda F_{\text{v}} \Delta V/kT)/\lambda F_{\text{v}} \tag{6}$$

assuming that δV is restricted to a finite interval (see (1)) and λ is an empirical parameter. It was shown that for small ΔV, $\lambda = 0{\cdot}5$ is optimal. Random numbers from the distribution (5) can be sampled using the following expression :

$$\delta V = kT \ln[\exp(-\lambda \Delta V F_{\text{v}}/kT) + \xi n(V) \lambda F_{\text{v}}/kT], \tag{7}$$

where ξ is a (pseudo) random number with uniform distribution between 0 and 1.

When δV is sampled from (5), the acceptance expression has to be modified :

$$P_{\text{acc}} = \min\{1, \exp[(U_o - U_n - P\delta V + kTN(\ln V_n - \ln V_o))/kT] * \times [n(V_o)/n(V_n)] * \exp[-\lambda(F^o{}_v + F^n{}_v)\delta V/kT]\}. \quad (8)$$

The second sampling distribution for δV, that we call brownian virial-bias (T, P, N) method (in the following : BVB/TPN), samples δV from the distribution

$$(4A\pi)^{-1/2} \exp(-\delta V^2/4A) + AF_v/kT, \quad (9)$$

where A is a constant related to the stepsize parameter ΔV [12] :

$$A = \Delta V^2/6. \quad (10)$$

In the expression (9) the second term is a deterministic term and thus one only has to be concerned with sampling from the gaussian distribution. A convenient way is to approximate it as a sum of uniformly distributed random variables.

For the BVB/TPN, the acceptance expression is

$$P_{\text{acc}} = \min\{1, \exp[(U_o - U_n - P\delta V + kTN(\ln V_n - \ln V_o))/kT] * \times \exp[-((\delta V + F^n{}_v A/kT)^2 - (\delta V - F^o{}_v A/kT)^2)/(4A)]\}. \quad (11)$$

3. Calculations and results

Test calculations have been performed on the Lennard-Jones fluid at reduced density $\rho^* = 0{\cdot}8$ and reduced temperature $T^* = 0{\cdot}75$. A cut-off of $R_c = 2{\cdot}5\sigma$ was used on the potential. Previous calculations [15] gave the reduced configurational internal energy as $U^* = -5{\cdot}78 \pm 0{\cdot}01$ and the reduced pressure $P^* = -0{\cdot}32 \pm 0{\cdot}1$. These results have been corrected for the cut-off [14] and thus correspond to the infinite range Lennard-Jones potential.

The system studied contained 100 particles under face-centred cubic periodic boundary conditions, where the smallest image to image distance is $5{\cdot}6\sigma$. The maximum particle displacement was set to $0{\cdot}15\sigma$ for all runs and the particles were perturbed by the force-bias technique [11] with $\lambda = 0{\cdot}5$. The particles to be perturbed were selected by the shuffled-cyclic procedure [16]. Volume perturbations were performed after every 50 particle perturbations. A relatively high value was chosen, partly to obtain good statistics on the volume perturbations characteristics and partly because the use of the force-bias particle displacement in general calls for more frequent volume perturbations. The system was first equilibrated in the (T, V, N) ensemble at the exact density, resulting $U^* = -5{\cdot}775$ and $P^* = -0{\cdot}281$, which fall well within the error limits of the previous study and thus verify most parts of the computer program used.

For the (T, P, N) ensemble calculations, the reduced pressure was set to $-0{\cdot}02$, a value that was obtained by deducing from $-0{\cdot}32$ the contributions to the pressure from $|r_i - r_j| > 2{\cdot}5\sigma$ and adding the correction due to the finite cut-off (see (4)). Starting from the equilibrated configuration, several 100 K long runs (1 K is equivalent to 1000 perturbations) were made in the (T, P, N) ensemble, using the different sampling techniques described.

The efficiency of the sampling can be characterized to some extent by considering the changes in the acceptance ratio, $\langle P_{acc} \rangle$ and the average of the magnitude of the accepted volume changes, $\langle \delta V_a \rangle$. Our 100 K configurations yield these quantities to 1–5 per cent accuracy. It is usually true that increasing $\langle P_{acc} \rangle$ results in a decrease in $\langle \delta V_a \rangle$ and vice versa. If the successive volume values are considered to form a one dimensional random walk, then these two indices should be combined as $\langle P_{acc} \rangle \langle \delta V_a \rangle^2$ [17].

As a final test, three long runs were performed using the ΔV and λ values found optimal for the three methods studied here to compare the convergence of the average volume $\langle V \rangle$ and the isothermal compressibility κ, a fluctuation property :

$$\kappa = (\langle V^2 \rangle - \langle V \rangle^2)/kT^2 \langle V \rangle). \tag{12}$$

Tables 1–3 collect the convergence characteristics of the *TPN*, BVB/*TPN* and EVB/*TPN* runs, respectively. The optimum value for ΔV was found to be 5·0, 5·0 and 6·0–7·0 for the EVB/*TPN*, BVB/*TPN* and *TPN* methods, respectively. For stepsizes near the optimal, $\lambda = 1{\cdot}0$ was found to be the best. However, for larger stepsizes, progressively smaller values appear to be the best. The reverse seems to be true for smaller stepsizes.

Table 4 gives the value of the volume and the isothermal compressibility at various stages of the three long runs. The error bounds given for $\langle V \rangle$ were computed from 50 K block averages with the method of batch-means [18, 2].

The values obtained for $\langle V \rangle$ are well within the respective error bounds, providing a rather good verification of the different sampling codes. The direction of the deviation from the reference value, $125\sigma^3$, is in accord with the deviation of the pressure from the reference value obtained in our (T, V, N) ensemble run.

Table 1. Convergence characteristics of the *TPN* method.

ΔV	$\langle \delta V_a \rangle$	$\langle P_{acc} \rangle$	$\langle P_{acc} \rangle \langle \delta V_a \rangle^2$
4·0	1·58	0·590	1·48
5·0	1·73	0·513	1·54
6·0	1·97	0·438	1·70
7·0	2·08	0·393	1·70
10·0	2·23	0·267	1·33

Table 2. Convergence characteristics of the BVB/*TPN* method.

ΔV	$\langle \delta V_a \rangle$	$\langle P_{acc} \rangle$	$\langle P_{acc} \rangle \langle \delta V_a \rangle^2$
3·0	1·47	0·914	1·96
4·0	1·95	0·817	3·11
5·0	2·27	0·678	3·49
6·0	2·55	0·504	3·28
7·0	2·25	0·346	1·76
10·0	1·53	0·130	0·30

Table 3. Convergence characteristics of the EVB/*TPN* method.

ΔV	λ	$\langle \delta V_a \rangle$	$\langle P_{acc} \rangle$	$\langle P_{acc} \rangle \langle \delta V_a \rangle^2$
4·0	0·5	2·05	0·825	3·48
4·0	0·7	2·67	0·863	4·03
4·0	0·8	2·26	0·871	4·44
4·0	0·9	2·35	0·886	4·91
4·0	1·0	2·31	0·871	4·63
4·0	1·1	2·46	0·860	5·19
4·0	1·2	2·52	0·856	5·42
4·0	1·3	2·46	0·841	5·10
5·0	0·4	2·39	0·671	3·83
5·0	0·5	2·47	0·717	4·37
5·0	0·6	2·62	0·723	4·96
5·0	0·7	2·74	0·735	5·52
5·0	0·8	2·84	0·685	5·55
5·0	0·9	2·92	0·714	6·10
5·0	1·0	3·08	0·677	6·42
5·0	1·1	3·05	0·670	6·27
5·0	1·2	3·10	0·646	6·22
6·0	0·5	2·89	0·557	4·65
6·0	0·7	3·08	0·531	5·06
6·0	1·0	3·64	0·478	6·37
7·0	0·5	3·27	0·445	4·78
7·0	0·6	3·22	0·434	4·51
7·0	0·7	3·63	0·368	4·85
7·0	0·8	3·60	0·347	4·51
7·0	0·9	3·78	0·335	4·79
7·0	1·0	3·79	0·323	4·66
10·0	0·1	2·49	0·318	1·97
10·0	0·2	2·79	0·289	2·25
10·0	0·3	2·84	0·268	2·17
10·0	0·4	3·20	0·234	2·40
10·0	0·5	3·36	0·168	1·90
10·0	0·6	3·24	0·158	1·66

The compressibility values still show significant differences although the longer *TPN* run appears to approach the EVB/*TPN* result. The compressibility obtained by the BVB/*TPN* is much lower than the value obtained with the other two methods. However, it was also observed that the 50 K block averages are significantly more correlated for the BVB/*TPN* than for either the *TPN* or the EVB/*TPN*, a fact that can account for the low value.

4. Discussion

The results obtained show that the sampling techniques presented in this paper allow better sampling of the configuration space using the (T, P, N) ensemble. The variations experienced in the values of the isothermal compressibility again highlight the difficulties in getting reliable values for fluctuation

Table 4. Convergence of the volume and the isothermal compressibility κ using different sampling techniques.

No. of conf.	*TPN* κ	*TPN* $\langle V \rangle$	BVB/*TPN* κ	BVB/*TPN* $\langle V \rangle$	EVB/*TPN* κ	EVB/*TPN* $\langle V \rangle$
50 K	0·1396	125·70	0·1246	125·87	0·1284	124·87
100 K	0·1381	125·70	0·1103	126·30	0·1386	125·90
150 K	0·1215	127·45	0·1174	129·19	0·1207	126·59
200 K	0·1298	125·91	0·1150	126·32	0·1202	126·37
250 K	0·1323	126·06	0·1142	126·31	0·1240	126·53
300 K	0·1374	126·24	0·1092	126·40	0·1199	126·58
350 K	0·1409	126·43	0·1044	126·38	0·1236	126·67
400 K	0·1387	126·18	0·1064	126·57	0·1218	126·59
450 K	0·1356	126·22	0·1054	126·58	0·1224	126·37
500 K	0·1312	126·20 ± 1·5	0·1071	126·58 ± 0·8	0·1228	126·21 ± 0·8
550 K	0·1301	126·14				
600 K	0·1345	126·25				
650 K	0·1301	126·14				
700 K	0·1305	126·16				
750 K	0·1281	126·13				
800 K	0·1284	126·20				
850 K	0·1273	126·09				
900 K	0·1286	126·08				
950 K	0·1293	125·95				
1000 K	0·1293	126·03 ± 0·9				

(*a*) Configurations are in units of 1000 particle perturbations (K); (*b*) $\Delta V = 6{\cdot}0$, 5·0 and 5·0 was used for *TPN*, BVB/*TPN* and EVB/*TPN*, respectively; (*c*) $\lambda = 1{\cdot}0$ was used for EVB/*TPN*.

properties. There are significant variations between the performance of EVB/*TPN* with different values and the theoretically derived $\lambda = 0{\cdot}5$ is not the best.

There is an extra expense, however, namely the computation of the virial sum in (4). In comparison with the original method, it is a significant addition, since it requires the forces on each of the particles [19]. However, it was also shown that the computation of the forces is cost effective for the use in the particle perturbation already [11, 12]. We can thus conclude, that if gradient-bias particle perturbations are used, the extra effort is only the updating of the virial sum. This becomes increasingly negligible as the potential gets more complex.

The quantitative assessment of the gains is a rather difficult task, since any criterion applied is necessarily arbitrary to some extent. For stepsizes near their optimal values, the acceptance rate showed a 25–40 per cent improvement for the EVB/*TPN* over the *TPN* and a 15–25 per cent improvement for the BVB/*TPN* over the *TPN*. The average magnitude of the accepted volume changes also improved by 40–50 per cent for the EVB/*TPN* and ~25 percent for the BVB/*TPN*.

The comparison of the error bounds on $\langle V \rangle$ permits one quantitative comparison of efficiency since it is known that to decrease the error on a Monte Carlo average by a factor of c one needs a run that is c^2 times longer. Due to

the stronger correlations found in the BVB/*TPN* mentioned earlier, the error bounds obtained by the method of batch means are less reliable, thus comparison will only be done between *TPN* and EVB/*TPN*. It is also worth noting that while the cumulative averages appear to converge rather fast for all techniques tried, the large error bounds obtained by the method of batch means indicate that the batch averages fluctuate rather strongly and thus the actual convergence may be less good than one could infer from the cumulative averages alone.

By comparing the error bounds at 500 K in table 4, one can estimate that $(1{\cdot}5/0{\cdot}8)^2 = 3{\cdot}5$ times longer *TPN* run is required to obtain the same precision. Considering the error bounds on the *TPN* run at 1000 K it appears that this factor is only slightly larger than two. The discrepancy between the two estimates is due to the uncertainty in the error bounds obtained by the method of batch means. Comparison of the combined index $\langle P_{acc}\rangle\langle \delta V_a\rangle^2$ shows a factor of 3·8. Thus, by conservative estimates, the EVB/*TPN* offers a factor of two increase in the sampling efficiency over the *TPN* and the gains may be as large as a factor of 3·8.

It should be also stressed that the volume perturbation frequency was kept constant during all runs. A possible way of realizing the economies offered by the new sampling techniques proposed is to keep the run length the same but reduce the frequency of volume perturbations. The best procedure, though, would optimize the volume perturbation frequency separately for each technique. Values used in the literature vary in a wide range from perturbing the volume at each step (hard spheres [2], Lennard-Jones mixtures [3]) to perturbing at each $3N$ or $5N$ steps (liquid water [6, 7]).

In summary, we can conclude that the two new sampling techniques proposed offer significant improvement over the conventional (T, P, N) ensemble method, with EVB/*TPN* performing better than BVB/*TPN*. The fact that the improvements are rather sensitive to the biasing distribution suggests that the presently achieved gains can be further enhanced by better biasing functions.

This research was supported by NIH Grant No. 5-R01-GM-24914. Fruitful discussions with Professor D. L. Beveridge and Mr. T. R. Vasu are gratefully acknowledged.

References

[1] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E., 1953, *J. chem. Phys.*, **21**, 1087.
[2] Wood, W. W., 1968, *Physics of Simple Liquids*, edited by H. N. V. Temperley, F. S. Rowlinson and G. S. Rushbrooke (North-Holland).
[3] McDonald, I. R., 1972, *Molec. Phys.*, **23**, 41.
[4] Norman, G. E., and Filinov, V. S., 1969, *High Temp. USSR*, **7**, 216.
[5] Adams, D. J., 1975, *Molec. Phys.*, **29**, 311.
[6] Owicki, J. C., and Scheraga, H. A., 1977, *J. Am. chem. Soc.*, **99**, 7403.
[7] Jorgensen, W. L., 1982, *J. chem. Phys.*, **77**, 4156.
[8] Jorgensen, W. L., 1982, *Chem. Phys. Lett.*, **92**, 404.
[9] Hastings, W. K., 1970, *Biometrika*, **57**, 97.
[10] Ceperley, D., Chester, G. V., and Kalos, M. H., 1977, *Phys. Rev.* B, **16**, 3081.
[11] (a) Pangali, C., Rao, M., and Berne, B. J., 1979, *Chem. Phys. Lett.*, **55**, 413. (b) Rao, M., Pangali, C., and Berne, B. J., 1979, *Molec. Phys.*, **37**, 1773.

[12] ROSSKY, P. J., DOLL, J. D., and FRIEDMAN, H. L., 1978, *J. chem. Phys.*, **69**, 4628.
[13] MEZEI, M., 1980, *Molec. Phys.*, **40**, 901.
[14] BARKER, J. A., and HENDERSON, D., 1976, *Rev. mod. Phys.*, **48**, 587.
[15] HANSEN, J. P., and VERLET, L., 1969, *Phys. Rev.*, **184**, 151.
[16] MEZEI, M., 1981, *J. comput. Phys.*, **39**, 128.
[17] KALOS, M. (private communication).
[18] BLACKMAN, R. B., and TUCKEY, J. W., 1958, *The Measurement of Power Spectra* (Dover). BLACKMAN, R. B., 1965, *Data Smoothing and Prediction* (Addison-Wesley).
[19] MEHROTRA, P. K., MEZEI, M., and BEVERIDGE, D. L., 1983, *J. chem. Phys.* (in the press).