LA-UR- 09-03502

Approved for public release;
distribution is unlimited.

Title: Optimized Nested Markov Chain Monte Carlo Sampling: Theory

Author(s):
Joshua D. Coe
Thomas D. Sewell
M. Sam Shaw

Intended for: Shock Compression of Condensed Matter 2009

Los Alamos
NATIONAL LABORATORY
EST 1943

OPTIMIZED NESTED MARKOV CHAIN MONTE CARLO SAMPLING: THEORY

Joshua D. Coe[1], **Thomas D. Sewell**[2], **M. Sam Shaw**[1]

[1] *Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545*
[2] *Department of Chemistry, University of Missouri-Columbia, MO 65211-7600*

Abstract. Metropolis Monte Carlo sampling of a reference potential is used to build a Markov chain in the isothermal-isobaric ensemble. At the endpoints of the chain, the energy is reevaluated at a different level of approximation (the "full" energy) and a composite move encompassing all of the intervening steps is accepted on the basis of a modified Metropolis criterion. By manipulating the thermodynamic variables characterizing the reference system we maximize the average acceptance probability of composite moves, lengthening significantly the random walk made between consecutive evaluations of the full energy at a fixed acceptance probability. This provides maximally decorrelated samples of the full potential, thereby lowering the total number required to build ensemble averages of a given variance. The efficiency of the method is illustrated using model potentials appropriate to molecular fluids at high pressure. Implications for *ab initio* or density functional theory (DFT) treatment are discussed.

Keywords: Monte Carlo, statistical mechanics, dense fluids
PACS: 05.10.Ln,02.70.Tt,05.20.Jj,05.70.Ce

INTRODUCTION

In traditional Monte Carlo (MC) simulation, configuration space is sampled *via* a sequence of elementary steps (single-particle displacements and sample-wide changes such as volume adjustments) punctuated by acceptance tests requiring evaluation of the full system potential energy. The cost of these calculations can be reduced from $O(N^2)$ to $O(N)$ in N-particle systems adequately modeled with pair potentials, but no comparable reduction is possible when accuracy constraints mandate energies calculated self-consistently using density functional theory (DFT) or *ab initio* wavefunction techniques. If analytic gradients are available at the electronic structure level of choice, molecular dynamics (MD) simulation possesses a considerable advantage over MC due to the collective nature of MD steps. Here we outline a thermodynamically optimized variant [1,2] of the Nested Markov Chain Monte Carlo (N(MC)^2) method [3,4] that enables use of many-particle composite MC steps while retaining practical acceptance probabilities for trial moves. The method is illustrated for model fluids at high temperature and pressure, and its application to sampling with a DFT potential is given in a companion proceedings article [5].

NESTED MARKOV CHAIN MONTE CARLO SAMPLING

The N(MC)^2 method uses elementary steps taken with an approximate reference potential to build composite steps taken with the potential of interest (the "full" potential), thereby facilitating collection of partially decorrelated samples of the latter. In the event that the full potential is of higher cost and quality, this procedure offers significant reduction in computational expense. If elementary steps taken on the reference potential are accepted according to the standard Metropolis criterion,

$$\alpha_{ij}^{(0)} = \min\left(1, e^{W_j^{(0)} - W_i^{(0)}}\right), \tag{1}$$

where $\alpha_{ij}^{(0)}$ represents the acceptance probability of a trial step from state i to state j and W represents the thermodynamic weight appropriate to the isothermal-isobaric (NPT) ensemble

$$W_j^{(0)} = -\beta(U_j^{(0)} + PV_j), \tag{2}$$

then the acceptance probability of composite steps taken on the full potential is

$$\alpha_{ij} = \min\left(1, e^{\delta W_j - \delta W_i}\right). \tag{3}$$

We have distinguished reference system quantities from those of the full system by a superscripted zero, and defined the difference in reference and full system weights as $\delta W_j = W_j - W_j^{(0)}$. If the differences in δW at distinct configurations are denoted $\Delta W \equiv \delta W_j - \delta W_i$, (3) can be reexpressed as

$$\alpha_{ij} = \begin{cases} e^{\Delta W}, & \Delta W < 0 \\ 1, & \Delta W \geq 0 \end{cases}. \tag{4}$$

The mean acceptance probability $\bar{\alpha}$ of composite steps can be expressed analytically [1] in the limit of their being built from sequences of O elementary steps steps long enough to decorrelate consecutive evaluations of the full potential. If O_{corr} is the correlation length of the potentials (assumed to be roughly the same for both reference and full),

$$\lim_{O \to O_{corr}} \bar{\alpha} \equiv \bar{A} = \left\langle \left\langle \frac{\alpha_{ij} e^{\delta W_j}}{e^{\delta W_i}} \right\rangle \right\rangle_0, \tag{5}$$

where nested brackets imply double averaging over initial and final states i and j and the subscripted '0' indicates that configurations used to build the averages are sampled from the reference distribution. Because no acceptance test is applied in its construction, $\bar{A}$ is an *a priori* estimate of mean acceptance probability for steps taken in an actual N(MC)^2 simulation.

The mean acceptance probability in (5) is a function of reference and full system temperature and pressure through the thermodynamic weights given in (2) and (3), *i.e.* $\bar{A} \equiv \bar{A}(P^{(0)}, T^{(0)}, P, T)$. The full system conditions are those matched to experiment, meaning that $(P^{(0)}, T^{(0)})$ can be varied to maximize $\bar{A}$. Because $\bar{A}$ is built from sampling the distribution of reference states, however, varying $(P^{(0)}, T^{(0)})$ directly is practically infeasible; instead we vary (P,T) to yield $\bar{A}_{max}$ at optimal values (P_{opt}, T_{opt}), then return the full system variables to their original values (P', T') and apply the same transformation to the reference variables (whose original values are P' and T' as well):

$$\begin{aligned} P_{opt}^{(0)} &= P' + (P' - P_{opt}) \\ T_{opt}^{(0)} &= T' + (T' - T_{opt}) \end{aligned}. \tag{6}$$

POTENTIAL ENERGY FUNCTIONS

For purposes of illustration, the optimized N(MC)^2 method is implemented here using diatomic model fluids in which the full interaction between a pair of molecules i and j is a sum over sites a and b of exponential-6 site-site interactions,

$$\varphi(r_{ab}) = \frac{\varepsilon}{\alpha - 6}\left(6e^{\alpha[1-(r_{ab}/r_0)]} - \frac{\alpha}{(r_{ab}/r_0)^6}\right) \tag{7}$$

$$\varphi_{ij} = \sum_{a=1}^{2}\sum_{b=1}^{2}\varphi(r_{ab}). \tag{8}$$

The three parameters appearing in (7) were obtained by fitting pair energies calculated using the Perdew-Burke-Erzenhof (PBE) exchange and correlation functionals [6] paired with a 6-31G* [7] basis set at three fiducial configurations and over a range of center-of-mass distances. Potential parameters obtained in this fashion were the following:

$$\begin{aligned} \varepsilon &= 34.156\ \mathrm{K} \\ r_0 &= 4.037\ \text{Angstrom}\,. \\ \alpha &= 12.29 \end{aligned} \tag{9}$$

The molecular bond length l was reduced from its value of 1.10 Å in the full potential to generate a set of reference potentials at l=1.05-0.90 Å. The anisotropic character of these potentials is illustrated in Fig. 1, along with the behavior of a purely spherical potential and three of the four configuration types used to parameterize (7).

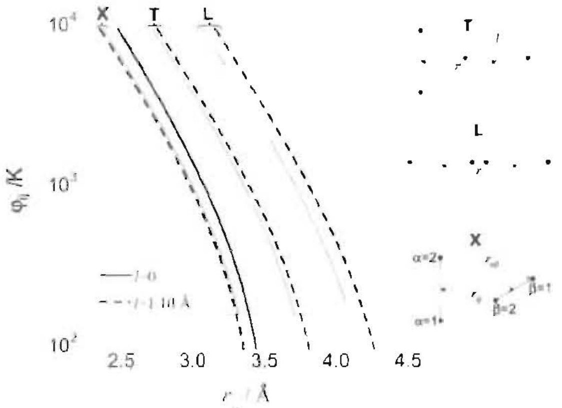

Figure 1. Illustration of the potential defined by (7)-(9) for the three fiducial pair configurations illustrated on the right. Curves are drawn for a purely spherical potential l=0), the full potential (l=1.10 Å) and one of the reference potentials (l=1.00 Å).

RESULTS AND DISCUSSION

Table 1 provides a summary of results at each point in the procedure. $\bar{A}$ was built from (5) by sampling the reference distribution at $(P^{(0)}=P', T^{(0)}=T')$, then optimized as a function of (P,T) to yield $\bar{A}_{\max}(P',T',P_{\mathrm{opt}},T_{\mathrm{opt}})$. These *a priori* estimates are to be compared with acceptance probabilities computed *a posteriori* from actual N(MC)^2 simulations carried out at the thermodynamic conditions indicated in the functional dependence of $\bar{A}_{\mathrm{MC}}$. Note that MC simulations performed using $(P_{\mathrm{opt}}, T_{\mathrm{opt}})$ or $(P^{(0)}_{\mathrm{opt}}, T^{(0)}_{\mathrm{opt}})$ have significantly higher acceptance probabilities than those performed at the original conditions $(P=P^{(0)}=P', T=T^{(0)}=T')$. Discrepancies between *a priori* and *a posteriori* estimates most likely reflect use of incompletely decorrelated samples in estimating the former.

TABLE 1. Summary of parameters and results for each stage of the procedure described in the text. Uncertainties in the last digit are recorded in parenthesis.

l /Å	1.05	1.00	0.90
T' /K	728	728	728
P' /GPa	4.84	4.84	4.84
T_{opt} /K	756	796	867
P_{opt} /K	5.08	5.35	6.08
$T^{(0)}_{\mathrm{opt}}$/K	700	660	589
$P^{(0)}_{\mathrm{opt}}$/GPa	4.60	4.33	3.60
$\bar{A}(P',T',P',T')$	0.367(5)	0.069(3)	0.001(0)
$\bar{A}_{\mathrm{MC}}(P',T',P',T')$	0.355(6)	0.084(3)	0.006(1)
$\bar{A}_{\max}(P',T',P_{\mathrm{opt}},T_{\mathrm{opt}})$	0.512(5)	0.195(5)	0.024(2)
$\bar{A}_{\mathrm{MC}}(P',T',P_{\mathrm{opt}},T_{\mathrm{opt}})$	0.553(7)	0.279(9)	0.030(5)
$\bar{A}_{\mathrm{MC}}(P^{(0)}_{\mathrm{opt}},T^{(0)}_{\mathrm{opt}},P',T')$	0.548(6)	0.251(7)	0.021(5)

It is important to note that estimates shown in Table 1 are for composite steps in the limit of O_{corr}, and that smaller O will yield potentially much higher acceptance probabilities. Figure 2 provides a direct comparison of $\bar{A}_{\mathrm{MC}}$ for N(MC)^2 simulations using optimized (black) and unoptimized (gray, red online) conditions as a function of O, the number of reference system steps used to build a single composite step made in the full system; $\bar{\alpha}_{ij}$ increases (often substantially) at all values of O for every potential.

From Eqn. (4) it is clear that a Dirac $\delta(0)$ distribution of ΔW would yield unit acceptance probability of composite trial steps; therefore, the principal aim of optimization is to shift the mean of the actual ΔW distribution as close to zero as possible and minimize its width. Figure 3 confirms

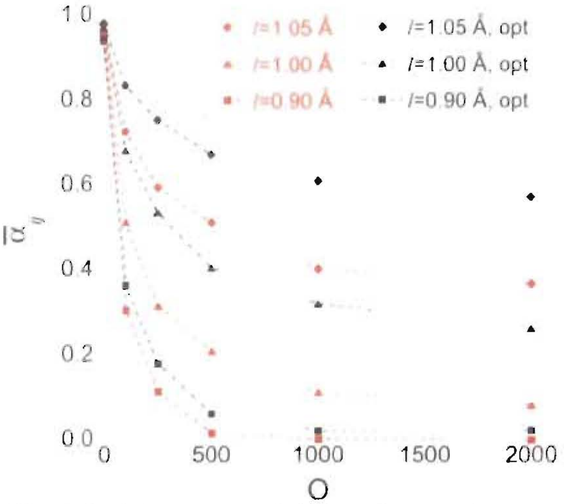

Figure 2. Acceptance probability of composite steps prior to (gray, red online) and following (black) thermodynamic optimization, as a function of composite step size O. Statistical uncertainties are smaller than the symbol size.

that this is precisely what has been achieved by the themodynamic optimization procedure described above for the reference potential l=1.00 Å and O=1000. The mean value of ΔW moves from -4.97 to -1.69 upon optimization, and the width of this distribution drops from σ=6.21 to σ=3.94.

For contexts in which the accuracy of a DFT or *ab initio* potential would be useful, the optimized N(MC)^2 method provides efficient sampling without having to sacrifice accuracy or efficiency.

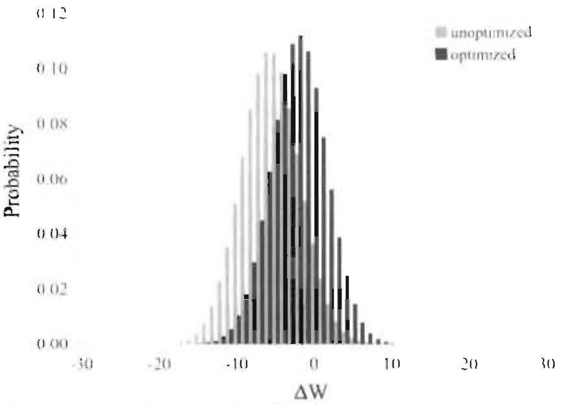

Figure 3. ΔW for l=1.00 Å and O=1000. Upon optimization of $(P^{(0)},T^{(0)})$, the ΔW appearing in (4) narrows and shifts closer to zero, raising the mean acceptance probability $\bar{\alpha}$.

ACKNOWLEDGEMENTS

JDC is a Director's Fellow at Los Alamos National Laboratory (LANL). MSS is supported by the LANL High Explosives Project of the National Nuclear Security Administration (NNSA) Advanced Strategic Computing Program (HE-ASC). TDS is supported by the LANL Laboratory Directed Research and Development (LDRD) Program and by the Army Research Office under Grant No. W911NF-05-1-0265. LANL is operated by Los Alamos National Security L.L.C. under the auspices of the NNSA and the United States Department of Energy, under Contract No. DE-AC52-06NA25396.

REFERENCES

1. Coe, J.D., Sewell, T.D., and Shaw, M.S., "Optimal Sampling Efficiency in Monte Carlo Sampling with an Approximate Potential", J. Chem. Phys. 130, 164104, 2009.
2. Coe, J.D., Sewell, T.D., and Shaw, M.S., "Equilibrium Thermodynamics of Dense Fluid Nitrogenas Described by Nested Markov Chain Monte Carlo Sampling of a Density Functional Theory Potential", J. Chem. Phys., submitted.
3. Iftimie, R., Salahub, D., Wei, D., and Schofield, J., "Using a Classical Potential as an Efficient Importance Function for Sampling with an Ab Initio Potential", J. Chem. Phys. 113, 4852, 2000.
4. Gelb, L.D., "Monte Carlo Simulations Using Sampling From an Approximate Potential", J. Chem. Phys. 18, 7747, 2003.
5. Shaw, M.S., Coe, J.D., and Sewell, T.D., "Optimized Nested Markov Chain Monte Carlo Sampling: Application to the Liquid Nitrogen Hugoniot Using Density Functional Theory", Shock Compression of Condensed Matter – 2009, AIP Conference Proceedings ??, edited by ??, AIP, Melville, NY, 2009, pp ??-??.
6. Perdew, J.P., Burke, K., and Ernzerhof, M., "Generalized Gradient Approximation Made Simple", Phys. Rev. Lett. 77, 3865, 1996.
7. Frisch, M., Pople, J.A., and Binkley, J.S., "Self-Consistent Molecular Orbitals Methods 25. Supplementary Functions for Gaussian Basis Sets", J. Chem. Phys. 80, 3265, 1984.