Efficient tight-binding Monte Carlo structural sampling of complex materials

Parthapratim Biswas
Debye Institute, Utrecht University, Princetonplein 5, 3508 TA Utrecht, the Netherlands

G.T. Barkema
Theoretical Physics, Utrecht University, Princetonplein 5, 3584 CC Utrecht, the Netherlands

Normand Mousseau
Department of Physics and Astronomy and CMSS, Ohio University, Athens, OH 45701, USA

W.F. van der Weg
Debye Institute, Utrecht University, Princetonplein 5, 3508 TA Utrecht, the Netherlands

(December 14, 2022)

While recent work towards the development of tight-binding and ab-initio methods has focused on molecular dynamics, Monte Carlo methods can often lead to better results with relatively little effort. We present here a multi-step Monte Carlo algorithm that makes use of the possibility of quickly evaluating local energies. For the thermalization of a 1000-atom configuration of Si, this algorithm gains about an order of magnitude in speed over standard molecular dynamics. The algorithm can easily be ported to a wide range of materials and can be dynamically optimized for a maximum efficiency.

Monte Carlo techniques are extensively used in statistical physics, but have received relatively little attention in the materials theory community. Most of the recent efforts towards the development of fast tight-binding or ab-initio codes, for example, has been done within the framework of molecular dynamics (MD) [1–3]. Since all atoms move simultaneously within MD, there are limitations to the approaches that can be used. In particular, the most promising order-N MD-based methods [4–6] are only applicable to materials with a well-defined electronic gap, reducing significantly the number of problems that can be studied with such algorithms. Monte Carlo (MC) techniques, on the other hand, allow a much wider range of steps to be used and provide an alternative to MD for a number of problems.

In this Letter, we introduce such a Monte Carlo algorithm. For a 1000-atom tight-binding model, we find a gain in efficiency of an order of magnitude over MD without the need for a perfectly clean electronic gap or any other special consideration. We first describe the core of the algorithm and then present some of the details associated with computing the intermediate steps. Examples are given using a series of amorphous silicon models described by the tight-binding interaction of Kwon et al. [7].

The general problem that we address here is the generation of a set of configurations, defined by the atomic positions \( \vec{X} \equiv \{ \vec{x}_1, \ldots, \vec{x}_N \} \), sampled according to the microcanonical ensemble, i.e., with a probability proportional to their Boltzmann weight:

\[
P(\vec{X}) \sim \exp(-\beta E(\vec{X})), \tag{1}
\]

where \( E(\vec{X}) \) is the total potential energy of configuration \( \vec{X} \) and \( \beta = (k_b T)^{-1} \), the inverse of Boltzmann’s constant \( k_b \) times the temperature \( T \). This sampling can be done using standard MD or MC.

In Monte Carlo, a markovian chain of configurations is generated through a sequence of trial moves or elementary steps. A standard move is obtained by randomly displacing a configuration \( \vec{X}_i \) to a trial position \( \vec{X}_{i+1} = \vec{X}_i + \delta \vec{X} \), shifting one or all atoms in the box. The trial move is then accepted with a probability given by the Metropolis criterion [8]

\[
P_a = \text{Min}[1, \exp(-\beta \delta E)], \tag{2}
\]

where \( \delta E = E(\vec{X}_{i+1}) - E(\vec{X}_i) \) is the energy difference between the trial and the initial positions.

A straightforward implementation of this standard Monte Carlo algorithm requires thus one full energy calculation per trial step, making it usually much slower than a standard MD simulation. It is possible, however, to speed up the MC simulation significantly, especially as the cost of computing the total energy increases faster than linear with the system size. This improvement exploits the fact that the energy difference between two configurations, \( \delta E \), which differ only locally, can be estimated in a much quicker way than the total energy, as discussed below. Using this approximate value, the acceptance ratio, Eq. (2), is modified as follows:

\[
P_a = \text{Min} \left[ 1, \exp \left( -\beta \delta E + \tilde{\beta} \delta \tilde{E} \right) \right]
\times \text{Min} \left[ 1, \exp \left( -\beta \delta E + \tilde{\beta} \delta \tilde{E} \right) \right], \tag{3}
\]

where \( \tilde{\beta} \) is an inverse temperature close to \( \beta \). Note that detailed balance is still strictly obeyed, irrespectively of the quality of the estimate and the value of \( \tilde{\beta} \). This expression can be implemented by first using an accept–reject procedure based on the first factor, followed by an accept–reject procedure based on the second factor only.
if the first one is accepted. The gain in doing this is that
the first factor in this equation requires only an estimate
of the energy difference, which is easily computed, while
the second factor, that involves an expensive total energy
calculation, is computed only if the first one is accepted.
If the estimate is accurate, \( \beta \) can be chosen equal to
and the acceptance probability for the second factor
is close to unity; effectively a full energy calculation is
only required per accepted step rather than per trial step.
Since rejected moves are now cheaper, the trial step size
can be enlarged for an increased efficiency.

A further reduction in the number of full energy calcu-
lations is possible by making a sequence of \( M \) interme-
diate displacement steps, each one accepted or rejected
based on the estimate of the energy difference, followed
by a single accept–reject procedure which requires a full
energy calculation. The resulting acceptance ratio, which
still strictly preserves detailed balance, becomes

\[
P_a = \prod_{i=1}^{M} \left[ \min \left[ 1, \exp \left( -\beta \delta E_i \right) \right] \right]
\times \min \left[ 1, \exp \left( -\beta \delta E - \sum_{i=1}^{M} \beta \delta E_i \right) \right].
\] (4)

These two variables, the number and typical size of the
intermediate trial moves \( M \), can then be optimized for
speed. For large simulation cells, a good choice is to
select a trial step size yielding a local acceptance ratio
around 50%, and to maximize the number of interme-
diate (cheap) trial moves \( M \) under the constraint that
the global acceptance ratio does not fall below 50%. The
resulting \( M \) depends on the accuracy of the energy esti-
mates.

Depending on the system (i.e., metal or insulator, crys-
talline, disordered solid, or liquid), there are different
types of Monte Carlo moves for which approximate en-
ergy differences can be computed efficiently and ac-
curately. In metals, for example, a MC move localized in
real space can be combined with an embedded-atom po-
tential, while a MC move localized in reciprocal space
is more appropriate in conjunction with plane-waves ab-
initio calculations. Here, we present an approach de-
signed for semiconductors and insulators, and apply it
to models of amorphous silicon described by the tight-
binding interaction developed by Goodwin, Skinner and
Pettifor \(^9\) and modified by Kwon et al. \(^8\).

Within the tight-binding formalism, the total energy
can be written as

\[
E = \sum_i \langle \Psi_i | H | \Psi_i \rangle + E_r + E_0 N_a
\] (5)

where the first term represents the electronic contribu-
tion of the energy given by the sum of the occupied eigen-
values of the tight-binding Hamiltonian \( H \), the second term
represents the ion-ion repulsion as well as the correction
from the double counting of the electron-electron inter-
action and \( E_0 \) is a constant energy shift per atom, and
\( N_a \) is the total number of atoms in the system. Details
of the potential and parameters used here can be found
in Ref. \(^6\).

We want to compute the approximate energy differ-
ences between two systems that are close to each other.
Because amorphous silicon is a semiconductor, we keep
the difference between the two configurations localized in
real space. Consider two configurations \( A \) and \( B \) with
total energies \( E_A \) and \( E_B \) respectively, such that the
atomic coordinates differ only in a small region around
a point \( \bar{x}_m \); for instance, in a Monte Carlo approach
with single-atom displacements, it is convenient to select
\( \bar{x}_m = (\bar{x}_i + \bar{x}_1)/2 \), the geometrical center between
the “active” region before and after the displacement \(^9\). A
spherical localization volume is constructed around \( \bar{x}_m \),
large enough to include the displaced atoms or atoms
and its local environment, following the principle of near-
sightedness of an equilibrium system \(^1\) which asserts
that the effects on the deformation on atoms outside the
localization volume is small. As discussed below, for the
model system used here, a sphere including about a hun-
dred atoms around \( \bar{x}_m \) is found to be sufficient.

This discussion may be written formally. Denoting I
and II as the regions inside and outside the localization
volume, the exact energy difference can be written as

\[
\delta E_{AB} = E_B - E_A
\]
\[
= (E_B^{\text{I}} + E_B^{\text{II}}) - (E_A^{\text{I}} + E_A^{\text{II}})
\]
\[
= (E_B^{\text{I}} - E_A^{\text{I}}) + (E_B^{\text{II}} - E_A^{\text{II}}).
\]

To obtain an estimate of the approximate energy dif-
ference, we assume that the contribution from region II
for both configurations is be the same. This is a good
approximation as long as the radius of the localized vol-
ume is greater than the localization length of the density
matrix of the system. The approximate energy difference
\( \delta E_{AB} \) is therefore simply given by
\( \delta E_{AB} = E_B^{\text{I}} - E_A^{\text{I}} \).
Furthermore, if the localization volume is big enough
and remains stationary during the atomic movement, the
boundary effects too can be neglected. We construct a
local Hamiltonian for the localization region I before and
after the movement with open boundary condition and
calculate the band energies for the corresponding local
Hamiltonian, the difference of which would give us the
approximate change of the band energy in the two con-
figurations. Typically in our calculations we have used vol-
umes with a radius of 7.5 \( \AA \), which roughly corresponds
to a hundred atoms.

We emphasize that although the quality of the approx-
imation is strongly related to the degree of localization,
the algorithm works even if there are states in the gap.
In this case, the error in the estimate increases but will
be corrected for appropriately in the final MC step using
the full energy calculation. In this sense, the algorithm
is much more stable under the presence of defects than a typical order-N MD.

We test this scheme on three sample configurations of $α$-Si with 300, 500 and 1000 atoms, respectively. The initial cells were prepared using the optimized Wooten-Winer-Weaire [14] bond-switching algorithm described in Ref. [13]. This method, which uses a Keating interaction potential [14], produces the best $α$-Si models to date, much larger than the value we used for a 1000-atom simulation. The second method approach, called the multi-local approach, is the Monte Carlo scheme proposed here: multiple local displacements, accepted or rejected based on an estimate of the energy difference, followed by a single full total energy calculation to recover correct sampling of the Boltzmann distribution. Here, the size $M$ of the sequence of local moves and the size of the atomic displacements are chosen so to maximize the diffusion in phase space; the values we used are $M = 300$ and a step size of 0.1 Å.

The standard MD simulation is performed at constant temperature, by rescaling the velocities at every time step. This modifies the trajectory but ensures a faster convergence than a more flexible constraint dynamics, damping considerably the thermal oscillations. In MD, the forces also need to be computed, in addition to the total energy. This requires the evaluation of the eigenvectors of the tight-binding Hamiltonian, a step which is not necessary in MC, and which increases the computational cost as well as the memory requirements significantly. In spite of these additional costs, MD is much faster than standard MC, since it is a second-order scheme which makes use of inertia.

We compare the computational efficiency of these methods in figure 2, which shows the evolution of the total energy as a function of CPU time on a fast workstation (DEC Alpha, 21264 architecture, 667 MHz) for the methods applied to three different size cells. All curves are fitted to the function

$$E(t) - E(0) = a \left( 1 - \exp(-t/\tau) \cos(\omega t - \phi) \right).$$

Here, the initial excess energy is given by $a$, the thermalization time by $\tau$, and the oscillatory behavior in MD by $\omega$ and $\phi$; in the fits to the MC data, we took $\omega = \phi = 0$ since MC inherently does not show oscillatory behavior. Important for us are the thermalization times. The fitted parameters for the multi-local MC approach are $\tau/N^3 = 21, 8.0$ and $6.2$ μs, respectively, for system sizes $N=300, 500$ and 1000 atoms. For MD, we find

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure1}
\caption{Accuracy of the $O(1)$ estimate of the energy difference between two nearby configurations as described in the text. The exact energy difference $\delta E$ between two configurations, calculated using a direct diagonalization method, is plotted horizontally. The difference between the exact and estimated energy differences is plotted vertically. The different symbols stand for the system sizes indicated in the figure.}
\end{figure}
The total self-diffusion \( \sum_i (\vec{r}_i(t) - \vec{r}_i(0))^2 \) as a function of CPU time, for system sizes \( N = 300 \) (bottom), \( N = 500 \) (middle) and \( N = 1000 \) (top). The solid curves correspond to MD and the dashed lines to the multi-local MC.

In summary, we have presented here a multi-local Monte Carlo procedure that exploits the fact that good estimations of energy differences between relatively close configurations can often be obtained cheaply. This permits us to construct a two-stage algorithm which first accepts-rejects a large number of moves based on this approximate calculation of the energy and then corrects the accumulated errors with a final accept-reject step using one full energy computation. This method is highly flexible and can easily be adapted to a wide-range of problems including those for which order-N molecular dynamics is inapplicable.

We have also presented an implementation of this algorithm for thermalization of tight-binding \( \alpha \)-Si models. In this case, the multi-local MC outperforms MD in the exploration of phase space by up to an order of magnitude.

NM acknowledges partial support from the National Science Foundation under grant number DMR-9805848. PB acknowledges support by NWO within the priority program “Solar cells in the 21st century”.

[1] E. Kim and Y.H. Lee, Phys. Rev. B 49, 1743 (1994)
[2] L. Colombo, Annual Reviews of Computational Physics, Vol. 4 (edited by D. Stauffer, World Scientific, Singapore, 1996)
[3] S. Goedecker and M. Teter, Phys. Rev. B 51, 9455 (1995)
[4] P. Ordejon, D.A. Drabold, R.M. Martin and M.P. Grumbach, Phys. Rev. B 51, 1456 (1995); U. Stephen and D.A.
[5] X.-P. Li, R.W Nunes and D. Vanderbilt, Phys. Rev. B 47, 10891 (1993); R.W. Nunes and D. Vanderbilt, Phys. Rev. B 50, 17611 (1994)

[6] S. Goedecker, Reviews of Mod. Phys. 71, 1085 (1999)

[7] I. Kwon, R. Biswas, C.Z. Wang, K.M. Ho, and C.M. Soukoulis, Phys. Rev. B 49, 7242 (1994)

[8] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller, J. Chem. Phys. 21, 1087 (1953).

[9] L. Goodwin, A.J. Skinner and D.G. Pettifor, Europhys. Lett. 9, 701-706 (1989).

[10] One aspect that is important for the proof of correctness of the Monte Carlo approach is that the localization region is centered around $\vec{x}_m$, thus independent whether the starting point was $A$ or $B$, i.e., $\delta E_{AB} = -\delta E_{BA}$. If for instance the localization region would have been centered around $A$, this property would not hold any longer, and detailed balance would be violated.

[11] W. Kohn, Phys. Rev. Lett. 76, 17 (1996); V. Heine in Solid State Physics, Vol. 35 (Academic Press).

[12] F. Wooten, K. Winer and D. Weaire, Phys. Rev. Lett. 54, 1392 (1985); F. Wooten and D. Weaire, Solid State Physics 40, 1 (1987).

[13] G.T. Barkema and N. Mousseau, Phys. Rev. B 62, 4985-4990 (2000).

[14] P. N. Keating, Phys. Rev. 145, 637 (1966).

[15] P. Biswas, to be published.

[16] G.T. Barkema and N. Mousseau, Phys. Rev. Lett. 81, 1865 (1998); N. Mousseau and G.T. Barkema, Phys. Rev. B 61, 1898 (2000).