

On Stochastic Error and Computational Efficiency of the Markov Chain Monte Carlo Method

Jun Li^{1,*}, Philippe Vignal^{1,2}, Shuyu Sun³ and Victor M. Calo^{1,3}

¹ Center for Numerical Porous Media, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia.

² Material Science and Engineering, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia.

³ Applied Mathematics and Computational Science, Earth Sciences and Engineering, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia.

Received 11 June 2013; Accepted (in revised version) 28 February 2014

Available online 12 June 2014

Abstract. In Markov Chain Monte Carlo (MCMC) simulations, thermal equilibria quantities are estimated by ensemble average over a sample set containing a large number of correlated samples. These samples are selected in accordance with the probability distribution function, known from the partition function of equilibrium state. As the stochastic error of the simulation results is significant, it is desirable to understand the variance of the estimation by ensemble average, which depends on the sample size (i.e., the total number of samples in the set) and the sampling interval (i.e., cycle number between two consecutive samples). Although large sample sizes reduce the variance, they increase the computational cost of the simulation. For a given CPU time, the sample size can be reduced greatly by increasing the sampling interval, while having the corresponding increase in variance be negligible if the original sampling interval is very small. In this work, we report a few general rules that relate the variance with the sample size and the sampling interval. These results are observed and confirmed numerically. These variance rules are derived for the MCMC method but are also valid for the correlated samples obtained using other Monte Carlo methods. The main contribution of this work includes the theoretical proof of these numerical observations and the set of assumptions that lead to them.

AMS subject classifications: 76T99, 82B05, 82B80, 62M05, 65C40

Key words: Phase coexistence, Gibbs ensemble, molecular simulation, Markov Chain Monte Carlo method, variance estimation, blocking method.

*Corresponding author. *Email addresses:* lijun04@gmail.com (J. Li), philippe.vignal@kaust.edu.sa (P. Vignal), shuyu.sun@kaust.edu.sa (S. Sun), victor.calo@kaust.edu.sa (V. M. Calo)

1 Introduction

The Monte Carlo method has successfully been applied to a wide variety of applications, which include the solution of integral equations by the Markov Chain Monte Carlo (MCMC) method [1], the Boltzmann equation by the Direct Simulation Monte Carlo (DSMC) method [2] and stochastic partial differential equations by a multilevel Monte Carlo method [3]. We focus our discussion on the MCMC method. An essential part of many scientific problems is to evaluate an integral in a high-dimensional space \vec{X} with the integrand containing a weighting function $f(\vec{X})$ (probability distribution function of the configuration \vec{X}) which is large in some area but close to zero almost everywhere else. The computational cost of evaluating the integral by conventional quadrature schemes is prohibitive since it demands a large number of quadrature points inside a high-dimensional space. This integral can be estimated by the average value of the integrand over a large number of configurations sampled inside the domain randomly, independently and uniformly, using the Monte Carlo (MC) method. Metropolis and Ulam [4] (see [5]) dubbed this simulation method *Monte Carlo* since it uses a large number of random fractions generated by a computer. The accuracy of the MC method can be improved by using the importance sampling scheme [6], which generates configurations non-uniformly but according to an artificially selected probability density function $g(\vec{X})$, which is close to $f(\vec{X})$, so that more probability mass is assigned to those configurations with higher probability [5–7]. In order to ensure the sampled configurations remain independent, the process demands the primitive function $G(\vec{X})$ of $g(\vec{X})$ and its inverse function $\vec{X}(G)$. Unfortunately, it is not feasible to find such $g(\vec{X})$ in most applications of interest. Rather than generating independent configurations, the Metropolis method [1], which still uses the importance sampling idea, generates (possibly) correlated configurations from the original $f(\vec{X})$ by a Markov chain. The Markov chain makes the algorithm simple and universal. This method is known as MCMC method [7]. Since the samples are correlated with each other, the variance of MCMC simulations with the same sample size is larger than the variance of the MC simulations using independent configurations. Additionally, the variance of MCMC simulations usually depends on the sampling interval.

The use of averages is common in scientific studies and many quantities related to thermal equilibria are averaged properties, measured in real experiments over large numbers of particles and long time intervals. If the ergodic hypothesis applies to the system at the molecular level [5], we can compute those quantities by ensemble averaging instead of time averaging using the probability distribution function $f(\vec{X})$, known from the partition function of the equilibrium state, an idea stemming from statistical mechanics. The MCMC method is a powerful tool based on ensemble averaging idea that can be used to calculate the quantities related to the thermal equilibrium state.

A system with fixed particle number N , volume V , and temperature T can be described by a canonical ensemble (constant- NVT), with the probability distribution function containing only the coordinates of the N particles as independent variables. This description is valid for systems where the quantities of interest only depend explicitly

on the location of all the particles. MCMC simulations of this system apply a random sequence of displacements to randomly selected particles. This random selection of particles and displacements is known as a trial move. The sample sequence that it forms generates a (correlated) Markov chain. The correlation degree of this sequence depends on the maximal random displacement applied, that is, the step size that determines the acceptance rate of the trial move.

Most real experiments are carried out under conditions of controlled pressure and temperature. Thus, the isobaric-isothermal ensemble (constant- NPT) is widely used in MCMC simulations where the particle location and the volume of the system are randomly modified to visit all possible configurations according to their respective probabilities. Here, the step size of the volume-changing trial move also influences the correlation degree of the successive configurations.

In adsorption studies where the chemical potential μ is fixed, instead of the particle number N , the grand-canonical ensemble (constant- μVT) is used to calculate the average particle number. The corresponding MCMC method includes a displacement trial move, as well as a trial insertion and removal of particles, with a step size usually fixed to one particle. That is, only one particle is tentatively inserted or removed from the volume each time. The acceptance ratio of particle insertion and removal is very small and thus results in a high-correlation degree of the related successive configurations. This correlation degree cannot be reduced because the step size is already the minimal divisible unit, one particle.

For the simulation of coexisting phases, important in many engineering applications, the MCMC algorithms based on the traditional ensembles described above suffer some important drawbacks. For example, limited computational resources imply that the number of particles used to represent the phase-coexistence system is relatively small. Thus, a large fraction of all particles used reside in the vicinity of the interface between phases. This induces a bias towards the interfacial properties when ensemble averages are computed, rather than including a balanced representation of the bulk phases.

In the literature several improvements to the traditional sampling have been proposed. In [8], a Gibbs- NVT MCMC method, where the total particle number, total volume, and temperature are fixed, was proposed to alleviate these algorithmic restrictions. This Gibbs- NVT scheme combines NVT , NPT and μVT ensembles for simulating coexisting phases. This combination skillfully avoids the interface predominance by introducing two subsystems modeled as separate boxes. This model allows particles to swap from one phase (box) to the other, while neglecting the potential energy between particles from different phases. Additionally, volume exchanges are allowed between the two boxes while the total volume is conserved. The acceptance ratio of particle swap is very small, as was the case for the grand-canonical ensemble simulation. This limitation can be particularly severe when the density of the dense phase is relatively very high and becomes important when modeling deposition and separation of dense liquids and solids. This drawback is avoided in the Gibbs-Duhem integration method [9–11]. Nevertheless, this integration scheme needs the initial point on the coexistence curve, and thus relies

on the use of another method that can provide this initial point. If one of the coexisting phases is a crystal, the method proposed in [12] improves the acceptance probability of exchanging particles.

The MCMC method based on Gibbs ensemble has successfully been applied to problems related to water systems [13], as well as oil production and processing [14–20]. In these applications, the solubility of hydrogen sulfide and other corrosive components in the gas-hydrocarbon mixtures is important data. Nevertheless, this solubility is poorly understood due to the lack of experimental results. In Gibbs-*NVT* ensemble simulations of two coexisting phases, there are three kinds of trial moves: particle displacement, volume exchange, and particle swap. In order to reduce the variance of the simulation results by decreasing the correlation degree of configurations, we adjust the step size for the first two trial moves. A discussion of the relationship between the variance and the step size of particle displacement is given in [5] but it is usually difficult to obtain a general rule for such a relationship. Recently [21], the liquid-vapor coexistence of methane was simulated by the Gibbs-*NVT* MCMC method. Then, the variation of mole fraction with pressure in a two-component system at a phase coexistence state was studied with the Gibbs-*NPT* MCMC method proposed in [22], where the total particle number, pressure, and temperature are fixed.

When Markov chain evolution is used for Monte Carlo simulations, it is not advisable to sample the system for the quantities of interest after each cycle, namely each trial move. Saving a large number of samples to reduce the stochastic noise contained in the samples requires a large amount of memory if the correlation is high; instead, the system is sampled at intervals (sampling interval). The larger the sampling interval is, the smaller the correlation degree of the collected samples will be. The same applies to the variance with fixed sample size (i.e., the total number of sampled cycles). The computational time is almost proportional to the product of the number of samples collected and the sampling interval. Thus, increasing the sampling interval either increases the CPU time when keeping the number of samples constant, or increases the variance of the results when keeping the CPU time constant. Nevertheless, our simulation results show that a good trade-off between the CPU time and memory usage can be achieved. In this paper, we describe the Gibbs-*NVT* MCMC method and employ it to model the coexisting phases of a Lennard-Jones (L-J) fluid. To make the problem tractable for the following theoretical analysis, we analyze the influence of the sampling interval and sample size on the variance of the simulation results on an idealized fluid, rather than the L-J fluid system. Finally, a general theoretical analysis is proposed to justify and prove some of the empirical observations and rules proposed.

2 The Markov Chain Monte Carlo method

Let the following integral define the expected value of A [5]:

$$\langle A \rangle = \frac{\int_{\Omega_{\vec{X}}} f(\vec{X}) A(\vec{X}) d\vec{X}}{\int_{\Omega_{\vec{X}}} f(\vec{X}) d\vec{X}}, \tag{2.1}$$

where \vec{X} is a high-dimensional vector and the formulas of $A(\vec{X})$ and $f(\vec{X})$ are given. To compute $\langle A \rangle$, it is convenient to use the MCMC method to generate correlated configurations \vec{X}_i after each cycle with a probability density proportional to $f(\vec{X})$. Unknown constant coefficients contained in $f(\vec{X})$ are canceled in the MCMC computation process. The system is sampled at intervals during the simulation and the configuration \vec{X}_j at each sampled cycle is used to estimate the expected value $\langle A \rangle$ by the average value $\bar{A} = \frac{1}{n} \sum_{j=1}^n A(\vec{X}_j)$ over n samples.

2.1 Basic algorithm of MCMC method

The algorithm of MCMC method [1] for solving the general integral (2.1) can be summarized as follows:

1. Initialization of configuration \vec{X} ;
2. For each cycle:
 - (a) Apply trial move changing \vec{X} to \vec{X}' ;
 - (b) Apply acceptance criterion to the new \vec{X}' ;
3. Sample the system at regular intervals (after every d cycles);
4. Stop after getting sufficient samples for analysis.

The initial configuration can be selected randomly from within the domain $\Omega_{\vec{X}}$ of the definition of the configuration space. The Markov chain is generated by randomly modifying the current configuration \vec{X} into \vec{X}' using the trial move algorithm.

The algorithm outlined in steps 1 to 4 should satisfy the ergodicity and time-reversal conditions. The ergodicity condition requires that from the current configuration \vec{X} it is possible to visit any $\vec{X}' \in \Omega_{\vec{X}}$ by a limited number of trial moves. The time-reversal condition requires that the probability for the current configuration to change back to its previous state is larger than zero. The probability density of the trial move event ($\vec{X} \rightarrow \vec{X}'$) is denoted by $\alpha(\vec{X} \rightarrow \vec{X}')$. Any new configuration \vec{X}' generated in step 2a will be accepted or rejected in step 2b based on the following acceptance criterion: \vec{X}' is accepted if the random number distributed uniformly inside $[0,1]$, Rf , is less than

$$\frac{\alpha(\vec{X}' \rightarrow \vec{X})f(\vec{X}')}{\alpha(\vec{X} \rightarrow \vec{X}')f(\vec{X})},$$

or rejected otherwise. This means that the acceptance probability is equal to

$$acc(\vec{X} \rightarrow \vec{X}') = \min \left[1, \frac{\alpha(\vec{X}' \rightarrow \vec{X})f(\vec{X}')}{\alpha(\vec{X} \rightarrow \vec{X}')f(\vec{X})} \right].$$

This selection for the acceptance probability is based on the detailed balance condition for the equilibrium state, which can be stated as

$$f(\vec{X})\alpha(\vec{X} \rightarrow \vec{X}')acc(\vec{X} \rightarrow \vec{X}') = f(\vec{X}')\alpha(\vec{X}' \rightarrow \vec{X})acc(\vec{X}' \rightarrow \vec{X}),$$

and also on the fact that

$$\frac{\min[1, \beta]}{\min[1, \beta^{-1}]} \equiv \beta.$$

The algorithm can be simplified significantly by using symmetric trial moves such that the probability density of the trial move from \vec{X} to \vec{X}' is equal to the probability density of the reverse move, that is, $\alpha(\vec{X} \rightarrow \vec{X}') = \alpha(\vec{X}' \rightarrow \vec{X})$. The detailed balance condition is a sufficient but not a necessary requirement, while in [23] the weaker “balance condition” was shown to be a necessary and sufficient requirement.

Samples are collected in step 3 after the simulation has reached the statistical steady state, that is, after an initial transitional period. The quantities of interest are estimated from samples collected every d cycles.

2.2 MCMC algorithm based on Gibbs-NVT ensemble

We discuss single component systems and assume that each molecule is modeled as a single particle. In the Gibbs-NVT ensemble [8], as described in [5], the probability density distribution function f and the related partition function Q_G are expressed as

$$f(N_1, V_1, \vec{S}_1, \vec{S}_2) = \frac{V_1^{N_1} (V - V_1)^{N - N_1} \exp[-\beta(U_1 + U_2)]}{Q_G(N, V, T) V \lambda^{3N} N_1! (N - N_1)!}, \quad (2.2)$$

and

$$Q_G(N, V, T) = \sum_{N_1=0}^N \int_0^V \int f(N_1, V_1, \vec{S}_1, \vec{S}_2) d\vec{S}_1 d\vec{S}_2 dV_1, \quad (2.3)$$

where T is the fixed temperature of both boxes, V is the fixed total volume, V_1 is the volume occupied by box 1, N is the fixed total particle number, N_1 is the particle number inside box 1, \vec{S}_1 and \vec{S}_2 are high-dimensional vectors that contain the normalized positions \vec{s}_i of all particles inside boxes 1 and 2, respectively, where the normalization parameters are each of the box sizes, which are $V_1^{1/3}$ and $(V - V_1)^{1/3}$, $\lambda = h / \sqrt{2\pi m / \beta}$ is the thermal de Broglie wavelength, h is the Planck constant, m is the molecular mass, $\beta = 1 / (k_B T)$, k_B is the Boltzmann constant, and $U_1 = U_1(\vec{S}_1, V_1)$ is the total potential energy of box 1, namely a summation of pair potential energy u_{ij} contributed by particles i and j inside box 1. The probability density distribution function, given in Eq. (2.2), and the related partition function, given in Eq. (2.3), are obtained after completing the integration with respect to the momentum variables. Here, the configuration \vec{X} consists of N_1 , V_1 , \vec{S}_1 and \vec{S}_2 . In general, we can take U_1 as a function $U_1(N_1, V_1, \vec{S}_1, \vec{S}_2)$, although it only depends on N_1 , V_1 , and \vec{S}_1 . Since we have formula (2.2) for $f(N_1, V_1, \vec{S}_1, \vec{S}_2)$, the expected

value $\langle U_1 \rangle$ can be defined as $\langle A \rangle$ using Eq. (2.1). Similarly, we can define the expected values of $\langle U_2 \rangle$, $\langle p_1 \rangle$, $\langle p_2 \rangle$, $\langle V_1 \rangle$, $\langle V - V_1 \rangle$, $\langle \rho_1 \rangle$ and $\langle \rho_2 \rangle$ using the following definitions of their transient values as functions of N_1 , V_1 , \vec{S}_1 and \vec{S}_2 . In the MCMC simulations, the successive configurations \vec{X}_i are generated by a Markov chain according to $f(N_1, V_1, \vec{S}_1, \vec{S}_2)$, and the samples of the quantities of interest can be determined from the configurations \vec{X}_j at the sampled cycles.

The parameter $Q_G(N, V, T)V\lambda^{3N}$ in the denominator of Eq. (2.2) is constant and avoided in the MCMC applications since only the ratio $f(\vec{X}')/f(\vec{X})$ is computed to determine $acc(\vec{X} \rightarrow \vec{X}')$, as discussed in Section 2.1. During the simulation process, N_1 , V_1 , \vec{S}_1 and \vec{S}_2 are randomly selected in each cycle and tentatively changed by the corresponding *symmetric* trial moves (see Fig. 1, where Δs and ΔV are the corresponding step sizes). We compute $acc(\vec{X} \rightarrow \vec{X}') = \min [1, f(\vec{X}')/f(\vec{X})]$ using the following formula to avoid the evaluation of $Q_G(N, V, T)V\lambda^{3N}$:

$$f(N_1, V_1, \vec{S}_1, \vec{S}_2) \propto \frac{V_1^{N_1} (V - V_1)^{N - N_1} \exp[-\beta(U_1 + U_2)]}{N_1!(N - N_1!)}. \quad (2.4)$$

For Lennard-Jones (L-J) fluids, we have:

$$u_{ij} = u_{L-J}(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right], \quad (2.5)$$

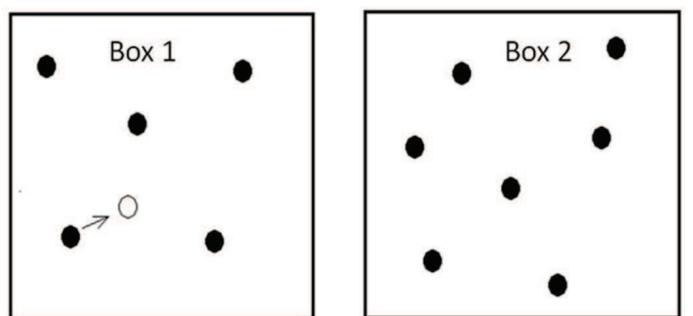
where ϵ is the depth of the potential well, σ is the finite distance at which the pair potential energy is zero, and $r = |\vec{r}_i - \vec{r}_j|$, where \vec{r}_i are the coordinates of particle i , computed using the normalized \vec{s}_i as well as the size of the box concerned. To simplify our computations, we replace Eq. (2.5) by a truncated potential such that

$$u_{ij} = u^{\text{cut}}(r) = \begin{cases} u_{L-J}(r), & r \leq r_c; \\ 0, & r > r_c. \end{cases} \quad (2.6)$$

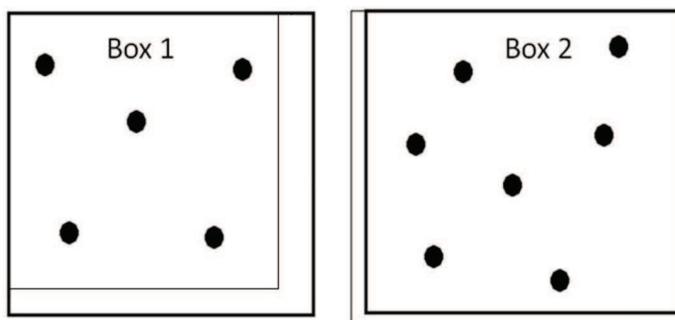
An explicit summation of u_{ij} under periodic boundary conditions takes into consideration the infinite periodic images of all particles. Additionally, a correction term due to the contributions beyond the cutoff distance r_c is added to determine the total potential energy for each box. Taking box 1 as an example, the correction for the total energy U_1 is [5]

$$U_1^{\text{tail}} = \frac{8\pi N_1^2}{3V_1} \epsilon \sigma^3 \left[\frac{1}{3} \left(\frac{\sigma}{r_{c,1}} \right)^9 - \left(\frac{\sigma}{r_{c,1}} \right)^3 \right], \quad (2.7)$$

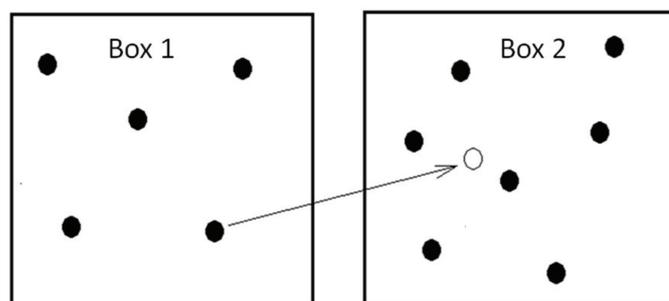
where $r_{c,1}$ is the cutoff distance for box 1. We use $r_{c,1} = 0.45V_1^{1/3}$ and $r_{c,2} = 0.45(V - V_1)^{1/3}$, which implies that boxes with different volumes have different cutoff distances. The total potential energy after the tentative trial move at each cycle is computed to determine $f(\vec{X}')$. The transient pressure, which is computed only at the sampled cycles, can



(a) Particle displacement: $\vec{s}'_i = \vec{s}_i + \Delta s(Rf_1 - 0.5, Rf_2 - 0.5, Rf_3 - 0.5)$



(b) Volume exchange: $\ln \frac{V'_1}{V - V'_1} = \ln \frac{V_1}{V - V_1} + \Delta V(Rf_4 - 0.5)$, $V'_2 = V - V'_1$



(c) Particle swap between boxes: $N'_1 = N_1 \pm 1$, $N'_2 = N - N'_1$

Figure 1: Schematic model for trial moves with the Gibbs-NVT MCMC method.

be calculated using the following definition [5]:

$$p_1 = \frac{N_1 k_B T}{V_1} + \frac{1}{3V_1} \frac{1}{2} \sum'_{i,j,\vec{n}} \left(-\frac{du}{dr} r \right), \quad (2.8)$$

where the factor 1/2 is used to correct for double counting of the pair-wise contributions and \vec{n} is a vector of three integers ranging from $(-\infty, \infty)$ through which we can represent

the contributions by the infinite periodic particle images. The truncation of Eq. (2.6) at the cutoff distance is also applied to the explicit summation of Eq. (2.8) to limit the number of effective pairs. The correction for pressure p_1 due to truncation is [5]

$$p_1^{\text{tail}} = \frac{16\pi N_1^2}{3V_1^2} \epsilon \sigma^3 \left[\frac{2}{3} \left(\frac{\sigma}{r_{c,1}} \right)^9 - \left(\frac{\sigma}{r_{c,1}} \right)^3 \right]. \quad (2.9)$$

In MCMC simulations, it is convenient to use non-dimensional quantities. The resulting non-dimensional system is defined by the following normalized quantities: number density $\rho_1^* = \sigma^3 \rho_1 = \sigma^3 N_1 / V_1$, pressure $p_1^* = p_1 \sigma^3 / \epsilon$, temperature $T^* = T k_B / \epsilon$, and energy $u_{ij}^* = u_{ij} / \epsilon$.

2.3 MCMC simulations using Gibbs-NVT ensemble

We designed the MCMC code according to the above algorithm based on Gibbs-NVT ensemble and ran the simulations for the phase-coexistence study of a L-J fluid on a Dell workstation (Dell T7500 running Ubuntu 12.04, Intel(R) processor Xeon(R) CPU X5650 @ 2.67GHz, RAM: 47GB). The cutoff distance for the two boxes is fixed at 45% (smaller than a half) of the corresponding box size, which is modified after each accepted volume exchange trial move. One thousand particles are used in our simulations and the initial normalized density of the two boxes is $\rho_{\text{init}}^* = 0.3$, unless otherwise stated.

In each cycle, a trial move is applied. It is selected randomly out of three possible cases (displacement move, volume exchange, particle swap, see Fig. 1) that are assigned different probabilities. The probability for selecting the displacement trial move is 0.9, 0.01 for volume exchange and 0.09 for particle swap. After a transitional period (about $L_{\text{init}} = 2 \times 10^7$ cycles for the current simulations), we sample the system every 50 cycles ($d = 50$).

The initial values of Δs and ΔV are chosen to be 0.1 (see Fig. 3). In order to have the acceptance ratios of the related trial moves be close to user-defined values, the step sizes are modified by an adaptive algorithm (see the source code mentioned in the preface of [5]) using the collected information. These step sizes are reset at the end of each $L_{\text{adjust}} = 5 \times 10^5$ cycles. The adaptive procedure used ensures that, by the completion of the initial L_{init} cycles, the step sizes of the different trial moves are such that the acceptance ratios of those trial moves are approximately equal to the predetermined value (e.g., 0.5 in the current simulations). Once the transitional L_{init} cycles are executed, the step sizes are kept fixed for the remainder of the simulation. Fixing the step sizes ensures the symmetry of the following trial moves.

For $T^* = 0.9$, Fig. 2 shows that the normalized density, volume, and pressure of the two boxes are converged after the predetermined $L_{\text{init}} = 2 \times 10^7$ cycles. Before the L_{init} cycles are complete, the step sizes Δs and ΔV are adjusted, the related achieved acceptance ratios are changed correspondingly, and finally approach the predetermined value of 0.5 as shown in Fig. 3 (left). After L_{init} cycles, the step sizes are fixed to their latest values and

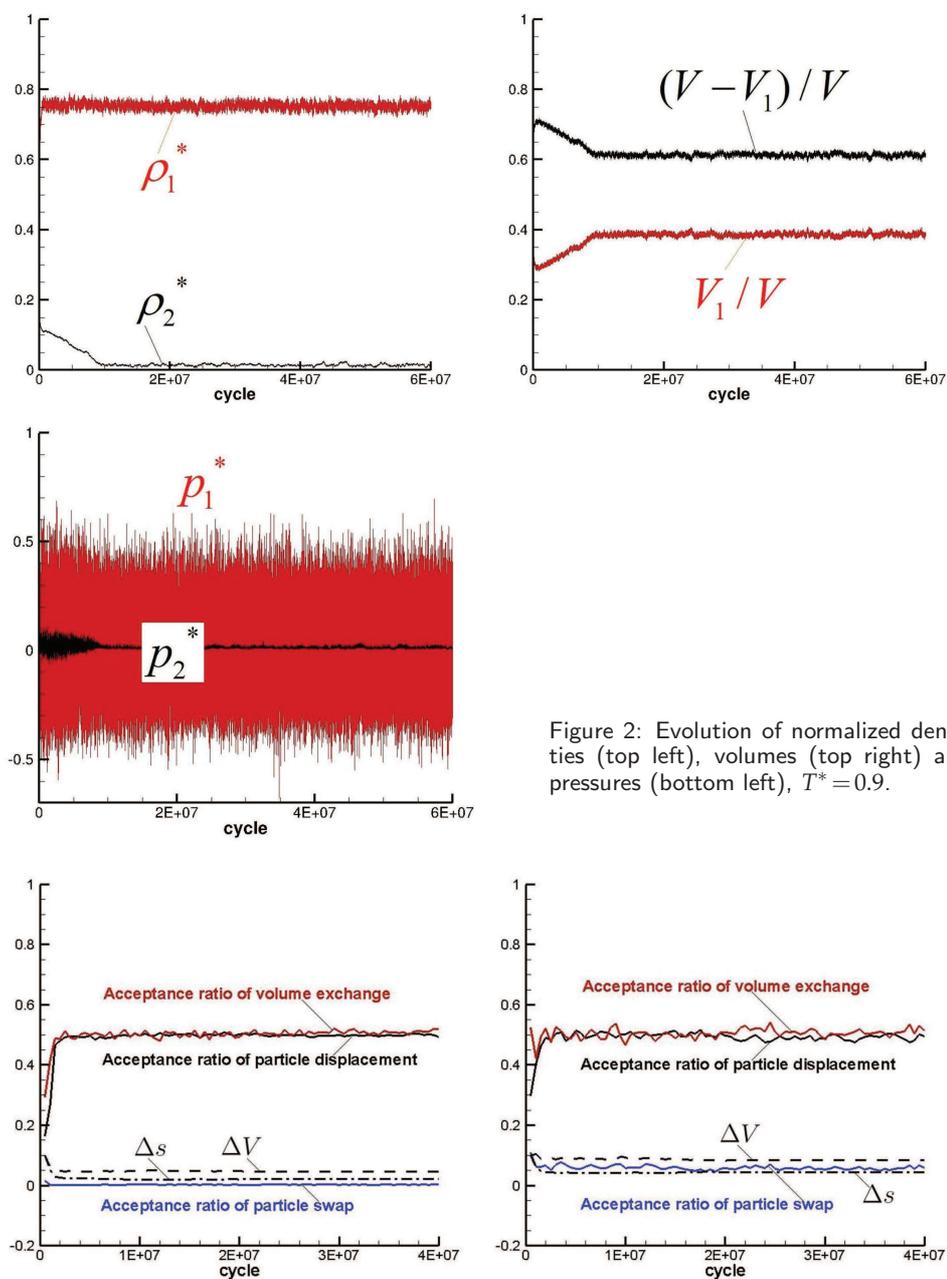


Figure 2: Evolution of normalized densities (top left), volumes (top right) and pressures (bottom left), $T^* = 0.9$.

Figure 3: Evolution of acceptance ratios and step sizes, $T^* = 0.9$ (left) and $T^* = 1.25$ (right).

the related acceptance ratios fluctuate about 0.5 as desired. Fig. 3 (left) also shows that the acceptance ratio of particle swap between boxes is only about 0.0026 for $T^* = 0.9$ because of the very high density of box 1 (see Fig. 2). This situation only worsens as density increases. As discussed in the introduction, this acceptance ratio cannot be improved

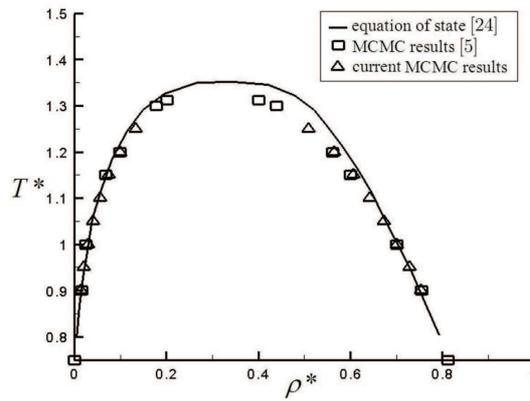


Figure 4: Phase diagram of a Lennard-Jones fluid.

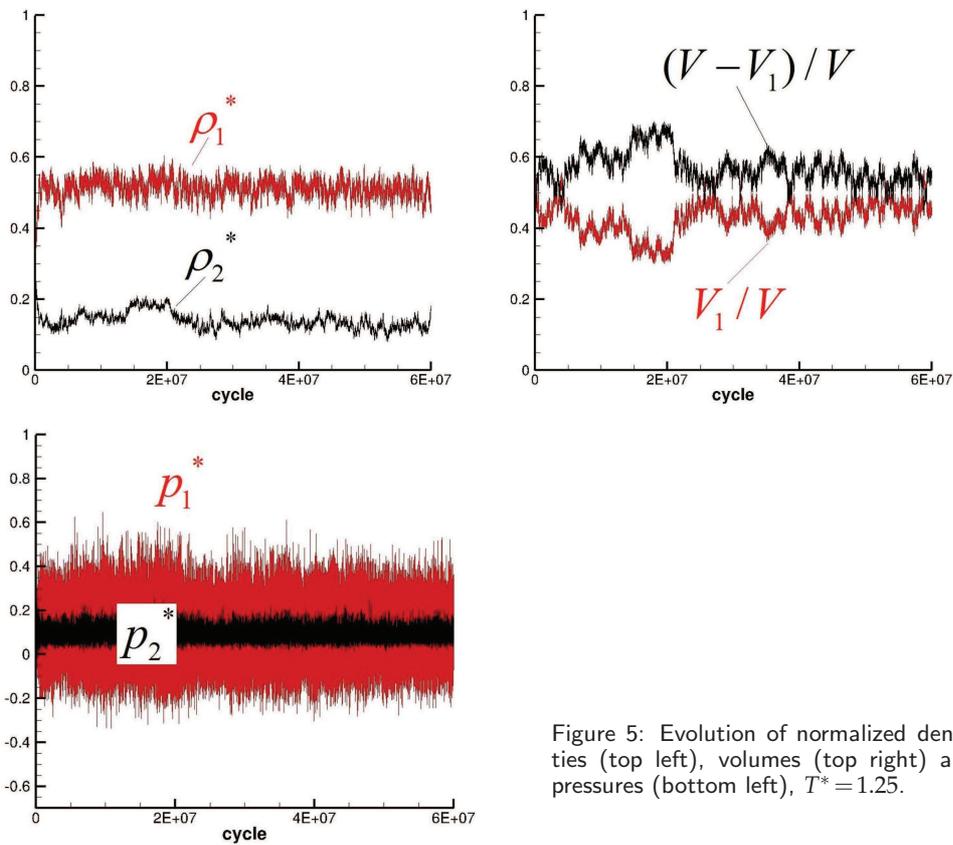


Figure 5: Evolution of normalized densities (top left), volumes (top right) and pressures (bottom left), $T^* = 1.25$.

when T^* is fixed, even though it results in a high-correlation degree of the successive samples. If instead T^* is increased from 0.9 to 1.25, the acceptance ratio of particle swap is increased to about 0.06, as shown in Fig. 3 (right) because the density of the dense phase is decreased. The results of ρ^* for different values of T^* are shown in Fig. 4. They

include a comparison with results computed using the equation of state presented in [24] and MCMC simulations [5].

As shown in Fig. 2, the statistical noise of the simulation results is larger in the dense-phase box than in the lower-density box. A similar observation is made in [5]. For example, the simulation results of $T^* = 1.25$ with the same initial density value of 0.3 are shown in Fig. 5, where we observe that the intensity difference of statistical noise of the two phases is reduced by decreasing the density difference.

3 Blocking method for estimating the variance

In the following discussion, we replace A of Eq. (2.1) by x , as used in the blocking method described in [25] to represent the sampled quantities of interest, including pressure, number density, volume, and total potential energy of each box. In MCMC simulations, each sample x_i is a measurement of a random variable x with an exact but unknown probability distribution, from which we define the expected value $\langle x \rangle$. We use the average value $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ to estimate $\langle x \rangle$. This estimation is then unbiased as $\langle \bar{x} \rangle = \langle x \rangle$. If the measurements can be taken as independent, the variance $\sigma^2(\bar{x})$ of the estimation using \bar{x} is inversely proportional to the size n of the sample set. But, if they are correlated, the variance then also depends on the sampling interval d between two successive samples.

In the blocking method, the following transformation is employed to decrease the sample size till $n' = 2$

$$\begin{cases} x'_i = (x_{2i-1} + x_{2i})/2, \\ n' = n/2. \end{cases} \quad (3.1)$$

After each blocking step, we get a new value for

$$\frac{c'_0}{n'-1} = \frac{1}{(n'-1)n'} \sum_{i=1}^{n'} (x'_i - \bar{x}')^2, \quad (3.2)$$

which increases during the blocking process and approximates $\sigma^2(\bar{x})$ if convergence is achieved. $\frac{c'_0}{n'-1}$ denotes the value we compute in practice. The value at the convergence point is used to estimate the variance of the average value and this estimation is unbiased [25]. If the blocking process does not converge, the largest value during the blocking process is a lower bound of the variance [25]. Convergence happens if the sample set covers a span which is several times larger than the maximal correlation interval τ , so that the "blocking" variables x'_i at the convergence point are independent Gaussian variables. The subtlety of the blocking method is to decrease the correlation degree of the new sample set $(x'_i)_{i=1, \dots, n'}$ making the correlated functions $\gamma'_{i,j} \equiv \langle x'_i x'_j \rangle - \langle x'_i \rangle \langle x'_j \rangle$, $i \neq j$ tend to zero.

The definition of $\sigma^2(\bar{x})$ is given in Eq. (5.2) using the correlation function $\gamma_{i,j}$. An alternative scheme to estimate $\sigma^2(\bar{x})$ is to directly select an estimator for $\gamma_{i,j}$. This selection needs to be done carefully since the most obvious estimator for $\gamma_{i,j}$ is a biased one, as its

expected value is not exactly equal to $\gamma_{i,j}$ [25]. As shown in [25], the estimator of $\sigma^2(\bar{x})$ using $\frac{c'_0}{n'-1}$ is unbiased since the expected value of $\frac{c'_0}{n'-1}$ at the convergence point is equal to $\sigma^2(\bar{x})$. Additionally, the blocking method is more efficient than many other estimators of $\sigma^2(\bar{x})$ [25].

4 Influence of simulation parameters on the variance

We take the set of samples after each cycle as the full sample set. If the trial move at the current cycle is accepted, the current sample is different from the previous one. If the trial move is instead rejected, the configuration remains unchanged and the current sample is the same as the previous one. The repeated samples induce a high correlation degree in the sample set, and are reasonable from a statistical point of view. Unfortunately, they only contain little useful information. The lower the correlation degree is, the smaller the variance with a given sample size will be. Instead of sampling after each cycle, we could, for example, add a sample to the set after each d cycles. The new sample set will be referred to as coarse sample set, which is a subset of the full sample set. We can reduce the correlation degree of the coarse sample set by increasing d . The total number of samples in this coarse set is denoted by n . In MCMC simulations, only the coarse sample set is stored, and the memory or disk usage can be reduced significantly by having d be much larger than one. The average value and the corresponding variance are calculated using the coarse sample set.

In the above simulation of a L-J fluid with $T^*=0.9$ in Fig. 2, we observed more statistical noise in box 1, which has the denser phase. Fig. 6 (left) also shows that the variances, estimated by the blocking method, of the normalized density and pressure of box 1, are larger than those of box 2 (the final wild fluctuation is due to numerical instabilities when n' becomes very small). Their volume variances are the same since the total volume V is fixed, which is consistent with the data shown in Fig. 2. The relative differences in variance of the number density and the pressure between the two boxes are reduced in Fig. 6 (right) compared to these in Fig. 6 (left) due to the increase of T^* , which is consistent with the comparison between Fig. 5 and Fig. 2.

Now, we discuss the variances of the simulation results of $T^*=1.25$. Fig. 5 implies that $\sigma^2(\rho_1)$ is larger than $\sigma^2(\rho_2)$ but Fig. 6 (right) shows that the variance $\sigma^2(\bar{\rho}_1)$ of $\bar{\rho}_1$ of the dense phase is smaller than $\sigma^2(\bar{\rho}_2)$, which is different from the observation of $T^*=0.9$ where the dense phase has a larger variance. Fig. 6 (right) also shows that $\sigma^2(\bar{p}_1)$ is smaller than $\sigma^2(\bar{p}_2)$, although $\sigma^2(p_1)$ is larger than $\sigma^2(p_2)$ as shown in Fig. 5 where the variation of p_1 ranges from about -0.2 to 0.4, and p_2 varies from about 0.45 to 0.55. Eq. (5.2) gives the definition of the variance $\sigma^2(\bar{x})$ of \bar{x} as a summation of the correlation functions $\gamma_{i,j}$, which can be replaced by γ_t where $t=|i-j|$ is the interval between the two samples of x_i and x_j . Although $\gamma_0(p_1) > \gamma_0(p_2)$, as $\gamma_0(p_1) = \sigma^2(p_1)$ and $\gamma_0(p_2) = \sigma^2(p_2)$, the decay speed of $\gamma_t(p_1)$ with the increase of t is much faster than that of $\gamma_t(p_2)$, as shown in Fig. 5, where p_1 has periodic fluctuations of scales larger than those observed in the fluctuations

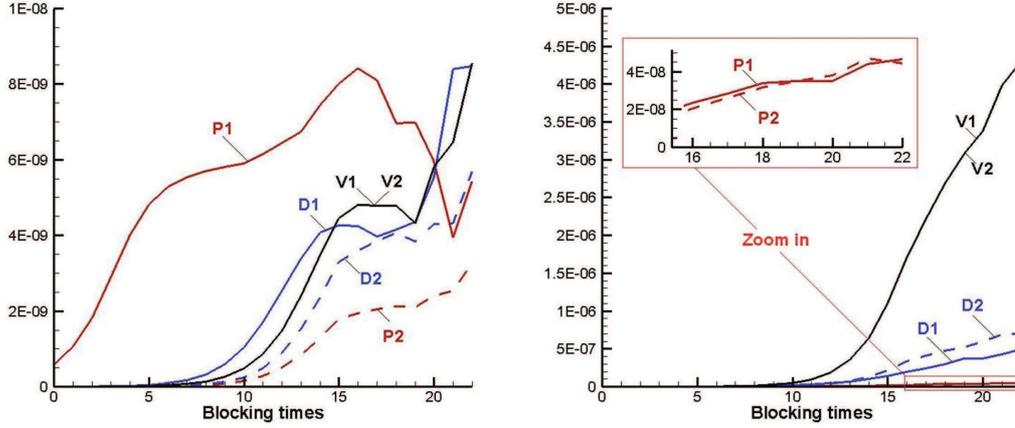


Figure 6: Variance estimates by the blocking method, $T^* = 0.9$ (left) and $T^* = 1.25$ (right), $n = 2^{25}$ and $d = 50$, P1 and P2 are the variances of average pressures in boxes 1 and 2, respectively, D1 and D2 are the variances of average densities in boxes 1 and 2, respectively, and V1 and V2 are the variances of average volumes in boxes 1 and 2, respectively.

of p_1 . Thus, $\sigma^2(\bar{p}_1)$ can be smaller than $\sigma^2(\bar{\rho}_1)$ even though $\sigma^2(p_1) > \sigma^2(\rho_1)$ according to Eq. (5.2). A similar interpretation applies to the observation of $\sigma^2(\bar{\rho}_1) < \sigma^2(\bar{\rho}_2)$.

The CPU time is proportional to the total cycle times L_{total} , which is almost equal to $n \times d$ ($L_{\text{total}} = L_{\text{init}} + n \times d$, but the cycle times L_{init} before convergence is negligible). We discuss the influence of n and d on the variance in what follows. The rules that we obtain are expected to be independent of the particular MCMC simulation used to generate the correlated sample set. Therefore, an ideal system, which is simpler than the L-J system and makes the simulation more efficient, is used in the following simulations.

In the ideal system, the total particle number N is equal to 12, and particle coordinates only take integral numbers $s_i = \pm 1$ as in the Ising model. The probability distribution function becomes

$$f(N_1, \vec{S}_1, \vec{S}_2) \propto \frac{\exp[-(U_1 + U_2)]}{N_1!(N - N_1)!}, \quad (4.1)$$

where the total potential energy $U_1 = -J \sum_{\substack{i=1, \dots, N_1 \\ i < j \leq N_1}} s_i s_j$ is a summation over all pairs (the periodic images of the particles are neglected here) located inside box 1 (the same for U_2), and $J = 0.1$, so that the acceptance ratio is not too small. For this model, we only need the trial move of particle swap and the spin trial move which randomly selects a particle and changes the sign of s_i . The properties of the two boxes are equivalent, thus the correlation degree of their sample sets is the same. We use the energy samples of U_1 and U_2 as x_i to compute $\frac{c_0'}{n'-1}$ using Eq. (3.2). The MCMC simulation results show that the blocking processes of the sample sets of U_1 and U_2 are very similar as shown in Fig. 7. During the blocking process, only the evolution of $\frac{c_0'}{n'-1}$ at the initial stage provides useful information relevant to the approaching process to the variance. As n' shrinks, the evolution of the blocking process becomes unstable, leading to wild fluctuations that can be arbitrar-

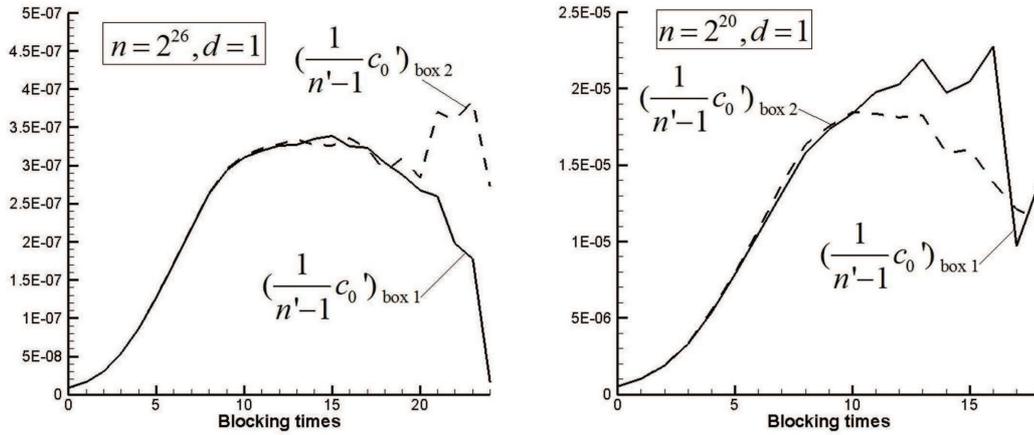


Figure 7: Blocking process of energy sample sets from the two boxes of the ideal system.

ily large, either increasing or decreasing the computed value. These oscillations are due to numerical instabilities and only serve to bound the trustworthy region of the blocking computation. These instabilities do not cause any problem if $\frac{c'_0}{n'-1}$ converges before losing stability, as seen in Fig. 7 (left). Thus, the value at the convergence point can be used to estimate the corresponding variance. But, in some cases where $\frac{c'_0}{n'-1}$ cannot converge before losing stability (see Fig. 7 (right)), it is difficult to judge where the separation point of the two stages is located. That is, the lower bound of the variance, namely the largest value before losing stability, is unknown. When using two sample sets with the same correlation degree, their initial stages should be the same while their final stages are random, which makes the separation point of the two stages easy to find. As shown in Fig. 7 (left) using $n = 2^{26}$ samples, the two curves overlap with each other and deviate after blocking 14 times which is the separation point. As it converges before the separation point, the variance for these two sample sets is about 3.3×10^{-7} . In Fig. 7 (right), while using 2^{20} samples, the two curves overlap with each other before blocking 10 times (the separation point) but are still not converged. This gives us the lower bound of the variance, which is about 1.84×10^{-5} . This is the largest value achieved before losing stability. Using different sample sets with similar correlation degrees simplifies the computation of the lower bound of the variance. Nevertheless, in real MCMC simulations, this would incur in prohibitive computational demands in terms of memory usage and CPU time. Thus, as shown in Fig. 7, we propose to use the first maximal point in the blocking process as the separation point and to estimate the lower bound of the variance. This observation is justified by the fact that $\frac{c'_0}{n'-1}$ is a theoretically non-decreasing quantity, while the oscillations shown in Figs. 6 and 7 can be justified by the loss of stability in the blocking computation.

Table 1 displays the variance for different combinations of n and d . The rows of Table 1 correspond to a fixed sampling interval, d , which implies that the correlation degree for

Table 1: Variance of Markov Chain Monte Carlo simulation results with different sample size n and sampling interval d .

	$n=2^{26}$	$n=2^{24}$	$n=2^{22}$	$n=2^{20}$
$d=2^0$	3.3×10^{-7}	1.3×10^{-6}	Not converged	Not converged
$d=2^2$	8.2×10^{-8}	3.3×10^{-7}	1.4×10^{-6}	Not converged
$d=2^4$	2.3×10^{-8}	9.1×10^{-8}	3.6×10^{-7}	1.5×10^{-6}
$d=2^6$	/	4.3×10^{-8}	1.7×10^{-7}	6.8×10^{-7}

*Note: "Not converged" refers to simulations where the blocking process becomes unstable before achieving a definite maximum, as shown in Fig. 7 (right).

the coarse sample set is also fixed. As observed in Table 1, for fixed d , the variance is almost inversely proportional to the sample size n . This feature is well known for independent sample sets but deserves further theoretical analysis for a general sample set. The CPU time is almost proportional to $n \times d$ as mentioned before. Thus, the variance with fixed d is also almost inversely proportional to the CPU time and so, it is the most rewarding choice for reducing the variance to increase n in view of CPU time. Using $V(n,d)$ as the variance at (n,d) , we observed in Table 1 that:

$$\frac{V(n_2,d)}{V(n_1,d)} = \frac{n_1}{n_2}. \quad (4.2)$$

Although increasing n is an efficient choice for reducing the variance in view of CPU time, it has an onerous cost for memory or disk usage. For $d=1$, the variance becomes very small only if n is very large which makes the memory requirement unacceptable. In order to reduce the variance while keeping the memory or disk usage low, we decrease the correlation degree of the coarse sample set by increasing d . For $n=2^{24}$, the variance decreases to about a quarter of the previous value when d increases from one to four, which is almost as efficient as increasing n in view of CPU time, also increased four times. But, when d increases from 16 to 64 with the CPU time being increased four times again, variance is reduced to 0.47 times the previous value, instead of 0.25 times as the ideal value, from 9.1×10^{-8} to 4.3×10^{-8} . This is wasteful with regard to CPU time, because we can choose to increase n from 2^{24} to 2^{26} while fixing d at 16, with CPU time increasing by four times too, but with the variance decreasing to about 0.25 times the previous value (9.1×10^{-8} to 2.3×10^{-8}), as already pointed out above. The following theoretical analysis can further prove that

$$\frac{d_1}{d_2} < \frac{V(n,d_2)}{V(n,d_1)} \leq 1, \quad (4.3)$$

where $d_2 > d_1$. The equality holds when the samples of the coarse sample set of d_1 are already independent and so the correlation degree of the coarse sample set cannot be further reduced by increasing d .

Usually, we also want to know how to reduce the variance for a given CPU time, namely $n \times d$. In the case of small d , the magnitude of the variance is more dependent

on the CPU time. The larger the CPU time is, the smaller the variance will be. Given the same CPU time, the larger the sample size n is, the smaller the variance will be. In the case of large d though, the magnitude of the variance depends more on the sample size n and in the limit case, becomes independent of d . In fact, these rules are nothing new compared with Eqs. (4.2)-(4.3), from which we have that

$$\frac{V(n_2, d_2)}{V(n_1, d_1)} = \frac{n_1 V(n_2, d_2)}{n_2 V(n_2, d_1)}, \tag{4.4}$$

such that

$$\frac{n_1 d_1}{n_2 d_2} < \frac{V(n_2, d_2)}{V(n_1, d_1)} \leq \frac{n_1}{n_2}. \tag{4.5}$$

For $n_1 \times d_1 = n_2 \times d_2$ corresponding to the same CPU time, Eq. (4.5) can be replaced by a special form using a new variable $V'(\text{CPUtime}, d)$, such that

$$1 < \frac{V'(\text{CPUtime}, d_2)}{V'(\text{CPUtime}, d_1)} \leq \frac{n_1}{n_2}, \tag{4.6}$$

where $d_2 > d_1$, as required in Eq. (4.3).

The maximal correlation interval τ of the full sample set can be estimated by the sampling interval d and the blocking times before convergence, because the 'blocking' variables x'_i at the convergence point are independent Gaussian variables [25]. For the case of Fig. 7 (left) with $n = 2^{26}$ and $d = 1$, the blocking process converges after blocking about 11 times. The estimation of τ is thus $2^{11} \times d = 2^{11}$. In Table 2, we present the estimates of τ for different sampling intervals d and sample sizes n , reported in Table 1. The data shows that when the blocking processes converge, different n and d lead to similar estimates of τ , with a value close to 2^{11} . This is to be expected as we are using different d and n to sample the same random experiment, where the correlation degree of the full sample set is fixed. For the variance analysis of Table 1, all coarse sample sets satisfy the conditions of $d \ll \tau$ and $nd \gg \tau$, which are required in the following theoretical analysis of the relationship between variance and the sampling parameters d and n .

Table 2: Blocking times before convergence used by different coarse sample sets from the same random experiment.

	$n = 2^{26}$	$n = 2^{24}$	$n = 2^{22}$	$n = 2^{20}$
$d = 2^0$	11	11	Not converged	Not converged
$d = 2^2$	9	8	9	Not converged
$d = 2^4$	8	7	7	7
$d = 2^6$	/	5	5	5

5 Theoretical analysis

In Section 4, we describe some empirical rules between the variance and the sample size n and sampling interval d , namely Eqs. (4.2)-(4.3). These rules are independent of the blocking method used to calculate the variance and reflect the underlying feature of the statistical rules, which are independent of the Monte Carlo methods used to generate the correlated samples. The theoretical analysis in this section justifies these rules.

Let x_1, x_2, \dots, x_n be the full sample set of a random variable x in a MCMC simulation at thermal equilibrium, which has the following features [25]:

$$\begin{cases} \langle x_i \rangle = \langle x_{i+t} \rangle, & \forall t, \\ \langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle = \langle x_l x_m \rangle - \langle x_l \rangle \langle x_m \rangle, & |i-j| = |l-m|, \end{cases} \quad (5.1)$$

where $\langle \dots \rangle$ denotes the expected value with respect to these exact but unknown probability distributions. In MCMC simulations, we estimate the expected value $\langle x \rangle$ by the average quantity $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$. The variance of \bar{x} is [25]:

$$\sigma^2(\bar{x}) = \langle \bar{x}^2 \rangle - \langle \bar{x} \rangle^2 = \frac{1}{n^2} \sum_{i,j=1}^n \gamma_{i,j} = \frac{1}{n} \left[\gamma_0 + 2 \sum_{t=1}^{n-1} \left(1 - \frac{t}{n} \right) \gamma_t \right], \quad (5.2)$$

where $\gamma_{i,j} = \langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle$ and $\gamma_t \equiv \gamma_{i,j}, t = |i-j|$. The variance $\sigma^2(\bar{x})$ of the average value \bar{x} differs from the variance $\sigma^2(x)$ of the random variable x itself. We have $\sigma^2(x) = \gamma_{i,i} = \gamma_0$, which is a fixed value for a given random variable x , with a fixed probability distribution, while $\sigma^2(\bar{x})$ depends on n and d . We define the maximal correlation interval τ for the full sample set as $\gamma_t \approx 0$, where $t > \tau$ (see Fig. 8). In MCMC simulations, it is reasonable to assume that

$$\frac{1}{n^2} \sum_{\substack{i,j=1 \\ i \neq j}}^n \gamma_{i,j} = \sigma^2(\bar{x}) - \frac{1}{n} \sigma^2(x) \geq 0, \quad (5.3)$$

where the equality holds when the samples are independent of each other. Fig. 8 shows some representative results of $\gamma_{i,j}$ in usual MCMC simulations. Fig. 8 (left) shows the results of a high-correlation sample set compared to Fig. 8 (middle). In the limit case where all samples are independent, $\gamma_{i,j}$ is equal to a constant $\sigma^2(x)$ for $i = j$ and zero otherwise, as shown in Fig. 8 (right). We use these schematic models only to show the contour distributions, the monotone interval and the location of maximal value. These models make it easy to understand the following linear interpolation scheme.

Theorem 5.1. *In MCMC simulations, the correlation degree of the full sample set is given, thus τ is fixed. For a general coarse sample set with fixed d , the variance $\sigma^2(\bar{x})$ is inversely proportional to n if $n \times d \gg \tau$.*

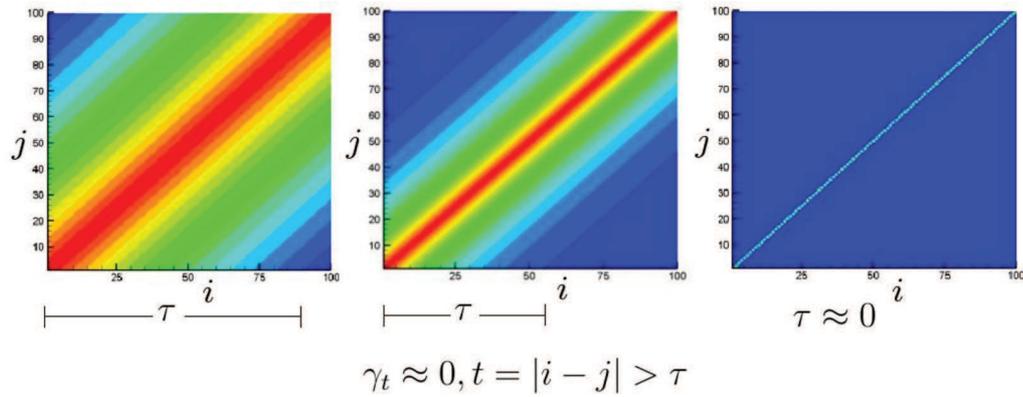


Figure 8: γ_t of representative sample sets with different correlation degrees.

Proof. We introduce $Y_b = \{y_{b,i} | y_{b,i} = x_{(i-1)d+1}, i = 1, 2, \dots, n\}$ containing n samples generated once from each d cycles. For an arbitrary d , Eq. (5.2) is modified to define the variance $\sigma^2(\bar{y}_b)$ as

$$\begin{aligned} \sigma^2(\bar{y}_b) &= \frac{1}{n^2} \sum_{\substack{i=(k_i-1)d+1 \\ j=(k_j-1)d+1 \\ k_i, k_j=1, \dots, n}} \gamma_{i,j} \\ &= \frac{1}{n} \left[\gamma_0 + 2 \sum_{t=1}^{n-1} \left(1 - \frac{t}{n}\right) \gamma_{td} \right] \approx \frac{1}{n} \left[\gamma_0 + 2 \sum_{t=1}^{\tau/d} \left(1 - \frac{t}{n}\right) \gamma_{td} \right]. \end{aligned} \tag{5.4}$$

Assuming $nd \gg \tau$, making $\frac{\tau/d}{n} \ll 1$, we conclude that

$$\sigma^2(\bar{y}_b) \approx \frac{1}{n} \left[\gamma_0 + 2 \sum_{t=1}^{\tau/d} \left(1 - \frac{t}{n}\right) \gamma_{td} \right] \approx \frac{1}{n} \left[\gamma_0 + 2 \sum_{t=1}^{\tau/d} \gamma_{td} \right] \tag{5.5}$$

which is inversely proportional to n and consistent with Eq. (4.2). □

Remark 5.1. The relationship between the variance and the sample size in Theorem 5.1 is well-known for an independent sample set but holds for a correlated sample set only if $nd \gg \tau$. This requirement is satisfied in the data presented in Tables 1 and 2. If $d \gg \tau$, making the samples in Y_b independent, $\sigma^2(\bar{y}_b)$ is always inversely proportional to n even if n is small, which can be understood from the definition of Eq. (5.2) where $\gamma_t \equiv 0, t > 0$, for independent sample sets.

Theorem 5.2. Given two sample sets with the same sample size n but different sampling intervals d_1 and d_2 ($d_2 > d_1$), respectively. If $nd_1 \gg \tau$, $nd_2 \gg \tau$ and $d_1, d_2 \ll \tau$, their variances satisfy

$$\frac{d_1}{d_2} < \frac{\sigma^2(n, d_2)}{\sigma^2(n, d_1)} \leq 1.$$

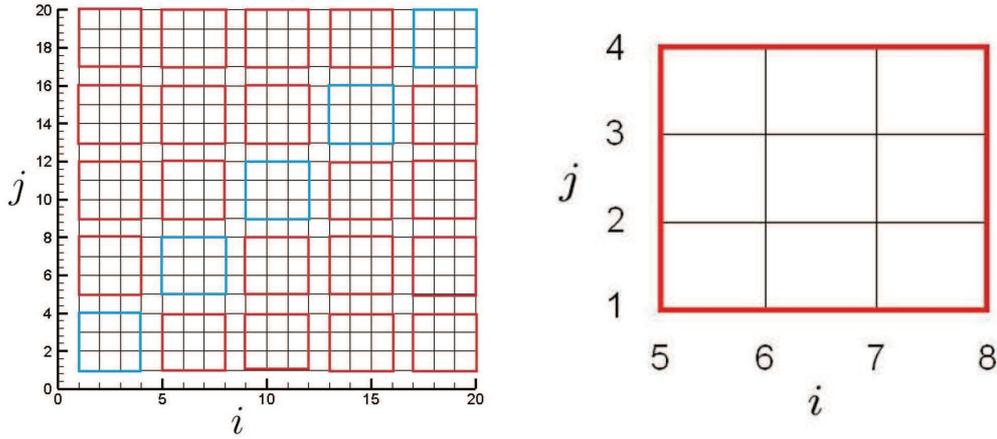


Figure 9: Schematic model for the summations with $d=4$ (left) and linear interpolation model (right).

Proof. We first discuss two sample sets: $Y_a = \{y_{a,i} | y_{a,i} = x_i, i = 1, 2, \dots, nd\}$ containing $n \times d$ samples and $Y_b = \{y_{b,i} | y_{b,i} = x_{(i-1)d+1}, i = 1, 2, \dots, n\}$ containing n samples generated once from each d samples of Y_a . From Eq. (5.2), we have:

$$\begin{cases} \sigma^2(\bar{y}_a) = \frac{1}{(nd)^2} \sum_{i,j=1}^{nd} \gamma_{i,j}, \\ \sigma^2(\bar{y}_b) = \frac{1}{n^2} \sum_{\substack{i=(k_i-1)d+1 \\ j=(k_j-1)d+1 \\ k_i, k_j=1, \dots, n}} \gamma_{i,j}. \end{cases} \quad (5.6)$$

As shown in Fig. 9, $\sum_{i,j=1}^{nd}$ is a summation over all vertexes (without repeating) of small black quadrilaterals, but

$$\sum_{\substack{i=(k_i-1)d+1 \\ j=(k_j-1)d+1 \\ k_i, k_j=1, \dots, n}}$$

is a summation over only the bottom-left vertexes of larger quadrilaterals, which are marked by red and blue colors and have indexes $k_i, k_j \in [1, n]$.

In the area of each blue quadrilateral, Ω_{blue} , centered at the maximal value of γ_t (see Fig. 9), it can be observed that

$$d \sum_{\substack{i=(k_i-1)d+1 \\ j=(k_j-1)d+1 \\ k_i, k_j \in \Omega_{\text{blue}}}} \gamma_{i,j} \leq \sum_{i,j \in \Omega_{\text{blue}}} \gamma_{i,j} < d^2 \sum_{\substack{i=(k_i-1)d+1 \\ j=(k_j-1)d+1 \\ k_i, k_j \in \Omega_{\text{blue}}}} \gamma_{i,j}, \quad (5.7)$$

where the equality holds when $\gamma_t \equiv 0, t > 0$. This can be understood by considering one

of the blue quadrilaterals in Fig. 9 (left), while realizing that the leftmost summation,

$$\sum_{\substack{i=(k_i-1)d+1 \\ j=(k_j-1)d+1 \\ k_i, k_j \in \Omega_{\text{blue}}}} \gamma_{i,j},$$

only contains the bottom-left corner of the blue quadrilateral, namely a maximal value which lies on the diagonal. Multiplying this maximal value by d will be lower or equal to $\sum_{i,j \in \Omega_{\text{blue}}} \gamma_{i,j}$, having d maximal terms and other terms with smaller but still positive values. The second part of the inequality stems from the fact that d^2 is multiplying one maximal term, and this will always be greater than $\sum_{i,j \in \Omega_{\text{blue}}} \gamma_{i,j}$, having d^2 terms, with only d terms taking maximal values.

In the area Ω_{red} of those red quadrilaterals located always at the monotone interval of γ_t , we assume d is much smaller than τ , and thus the linear interpolation is valid in each small local red area of size d . For the representative red quadrilateral shown in Fig. 9 (right) with $k_i = 2$ and $k_j = 1$, we have $\gamma_{5,1} = \gamma_{6,2} = \gamma_{7,3} = \gamma_{8,4}$. According to linear interpolation, we get $\gamma_{6,1} + \gamma_{5,2} = \gamma_{7,2} + \gamma_{6,3} = \gamma_{8,3} + \gamma_{7,4} \approx 2\gamma_{5,1}$, $\gamma_{7,1} + \gamma_{5,3} = \gamma_{8,2} + \gamma_{6,4} \approx 2\gamma_{5,1}$ and $\gamma_{8,1} + \gamma_{5,4} \approx 2\gamma_{5,1}$. Thus, we have the following estimate

$$\sum_{\substack{i=5,\dots,8 \\ j=1,\dots,4}} \gamma_{i,j} \approx 4^2 \sum_{\substack{i=(k_i-1)d+1 \\ j=(k_j-1)d+1 \\ k_i=2, k_j=1}} \gamma_{i,j} = 4^2 \gamma_{5,1}. \tag{5.8}$$

Generally, the following approximation for any arbitrary red quadrilateral is valid:

$$\sum_{i,j \in \Omega_{\text{red}}} \gamma_{i,j} \approx d^2 \sum_{\substack{i=(k_i-1)d+1 \\ j=(k_j-1)d+1 \\ k_i, k_j \in \Omega_{\text{red}}}} \gamma_{i,j}. \tag{5.9}$$

According to Eqs. (5.7)-(5.9), we have

$$d \sum_{\substack{i=(k_i-1)d+1 \\ j=(k_j-1)d+1 \\ k_i, k_j \in \Omega_{\text{blue,all}}}} \gamma_{i,j} + d^2 \sum_{\substack{i=(k_i-1)d+1 \\ j=(k_j-1)d+1 \\ k_i, k_j \in \Omega_{\text{red,all}}}} \gamma_{i,j} \leq \sum_{i,j=1}^{nd} \gamma_{i,j} < d^2 \sum_{\substack{i=(k_i-1)d+1 \\ j=(k_j-1)d+1 \\ k_i, k_j=1,\dots,n}} \gamma_{i,j}. \tag{5.10}$$

At this point, we assume Eq. (5.3) is valid and apply it to the sample set Y_b . Considering that $d > 1$ and $d^2 > d$, we get

$$d \sum_{\substack{i=(k_i-1)d+1 \\ j=(k_j-1)d+1 \\ k_i, k_j=1,\dots,n}} \gamma_{i,j} \leq \sum_{i,j=1}^{nd} \gamma_{i,j} < d^2 \sum_{\substack{i=(k_i-1)d+1 \\ j=(k_j-1)d+1 \\ k_i, k_j=1,\dots,n}} \gamma_{i,j}, \tag{5.11}$$

where the equality holds when $\gamma_t \equiv 0, t > 0$. Substituting Eq. (5.11) into Eq. (5.6), we have

$$\frac{\sigma^2(\bar{y}_b)}{d} \leq \sigma^2(\bar{y}_a) < \sigma^2(\bar{y}_b). \quad (5.12)$$

We introduce $Y_c = \{y_{c,i} | y_{c,i} = x_i, i = 1, 2, \dots, n\}$ which contains n samples as Y_b and has the same correlation degree as Y_a . Since $n \gg \tau$ and $nd \gg \tau$ according to the assumption, the conclusion of Eq. (5.5) implies that

$$\frac{\sigma^2(\bar{y}_c)}{\sigma^2(\bar{y}_a)} = \frac{nd}{n} = d. \quad (5.13)$$

Substituting Eq. (5.13) in Eq. (5.12), we get

$$\frac{1}{d} < \frac{\sigma^2(\bar{y}_b)}{\sigma^2(\bar{y}_c)} \leq 1, \quad (5.14)$$

with which the proof is complete. \square

Remark 5.2. Taking the sample set with d_1 in Eq. (4.3) as Y_c and the other as Y_b , we observe that Eq. (4.3) is equivalent to Eq. (5.14) proved here. If the sample set Y_a (namely Y_c) has a high correlation degree, the summation over area Ω_{red} is dominant (see Fig. 8 (left)) and $\frac{\sigma^2(\bar{y}_b)}{\sigma^2(\bar{y}_c)}$ converges to $\frac{1}{d}$ according to Eq. (5.9) which implies $\sigma^2(\bar{y}_a) = \sigma^2(\bar{y}_b)$. In contrast, $\frac{\sigma^2(\bar{y}_b)}{\sigma^2(\bar{y}_c)} = 1$ if the samples in Y_a are independent.

The assumptions of Theorem 5.2 are that $nd_1 \gg \tau, nd_2 \gg \tau$ and $d_1, d_2 \ll \tau$, which are satisfied in the data shown in Tables 1 and 2. In real applications, nd should be much larger than τ since otherwise the variance of the average value is very high, which makes the average value not trustworthy. For the selection of d , we suggest to let d be much larger than 1 to reduce memory usage. In addition, we also suggest to let d be much smaller than τ as otherwise this leads to loss of too much correlated information that can still reduce the variance effectively. The two necessary assumptions can thus be easily satisfied in real applications.

6 Conclusions

The influence of the sample size n and sampling interval d used in MCMC simulations on the variance of the average quantities is analyzed using numerical results and proved theoretically. If $n \times d$ is much larger than the maximal correlation interval τ of the full sample set, the variance of the estimation using a coarse sample set with fixed d is inversely proportional to n and the CPU time. For a given CPU time, the memory or disk usage (namely the sample size) can be reduced greatly by increasing d , while getting a negligible increase in variance if the original d is very small.

In the implementation of the blocking method, the blocking process is subject to increased fluctuations when the sample size n' is reduced; in particular, the fluctuation gets worse when n' approaches two. The current results show that the fluctuation starts near the first maximal point obtained during the blocking process. Additionally, the corresponding maximal value can be used as an estimate of the variance if the blocking process converges, or as a lower bound estimate of the variance if the blocking process does not converge.

Acknowledgments

This work was supported in part by the King Abdullah University of Science and Technology (KAUST) Center for Numerical Porous Media. In addition, S. Sun would also like to acknowledge the support of this study by a research award from King Abdulaziz City for Science and Technology (KACST) through a project entitled "Study of Sulfur Solubility using Thermodynamics Model and Quantum Chemistry".

References

- [1] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, Equation of state calculations by fast computing machines, *J. Chem. Phys.*, 21 (1953), 1087-1092.
- [2] G. A. Bird, *Molecular gas dynamics and the direct simulation of gas flows*, Clarendon Press, Oxford, 1994.
- [3] A. Barth, A. Lang, Multilevel Monte Carlo method with applications to stochastic partial differential equations, *Int. J. Comput. Math.*, 89 (2012), 2479-2498.
- [4] N. Metropolis, The beginning of the Monte Carlo method, *Los Alamos Science*, 12 (1987), 125-130.
- [5] D. Frenkel, B. Smit, *Understanding molecular simulation, from algorithms to applications*, Academic press, 2002.
- [6] A. W. Marshall, The use of multi-stage sampling schemes in Monte Carlo methods, *Symposium on Monte Carlo methods*, Wiley, New York, 1956, 123-140.
- [7] J. S. Liu, *Monte Carlo strategies in scientific computing*, Harvard university, 2001.
- [8] A. Z. Panagiotopoulos, Direct determination of phase coexistence properties of fluids by Monte Carlo simulation in a new ensemble, *Mol. Phys.*, 61 (1987), 813-826.
- [9] D. A. Kofke, Gibbs-Duhem integration: A new method for direct evaluation of phase coexistence by molecular simulations, *Mol. Phys.*, 78 (1993), 1331-1336.
- [10] D. A. Kofke, Direct evaluation of phase coexistence by molecular simulation via integration along the coexistence line, *J. Chem. Phys.*, 98 (1993), 4149-4162.
- [11] R. Agrawal, D. A. Kofke, Solid-fluid coexistence for inverse-power potentials, *Phys. Rev. Lett.*, 74 (1995), 122-125.
- [12] P. Tilwani, D. Wu, Direct simulation of phase coexistence in solids using the Gibbs ensemble: Configuration annealing Monte Carlo, Master's thesis, Department of Chemical Engineering, Colorado School of Mines, Golden, Colorado, USA, 1999.
- [13] J. R. Errington, A. Z. Panagiotopoulos, A fixed point charge model for water optimized to the vapor-liquid coexistence properties, *J. Phys. Chem. B*, 102 (1998), 7470-7475.

- [14] B. Smit, S. Karaborni, J. I. Siepmann, Computer simulation of vapor-liquid phase equilibria of *n*-alkanes, *J. Chem. Phys.*, 102 (1995), 2126-2140.
- [15] M. G. Martin, J. I. Siepmann, Transferable models for phase equilibria 1. United-atom description of *n*-alkanes, *J. Phys. Chem. B.*, 102 (1998), 2569-2577.
- [16] S. K. Nath, F. A. Escobedo, J. J. de Pablo, On the simulation of vapour-liquid equilibria for alkanes, *J. Chem. Phys.*, 108 (1998), 9905-9911.
- [17] J. R. Errington, A. Z. Panagiotopoulos, A new intermolecular potential model for the *n*-alkane homologous series, *J. Phys. Chem. B.*, 103 (1999), 6314-6322.
- [18] J. J. Potoff, J. R. Errington, A. Z. Panagiotopoulos, Molecular simulation of phase equilibria for mixtures of polar and non-polar components, *Mol. Phys.*, 97 (1999), 1073-1083.
- [19] P. Ungerer, V. Lachet, B. Tavitian, Applications of molecular simulation in oil and gas production and processing, *Oil & gas Science and technology - Rev. IFP*, 61 (2006), 387-403.
- [20] M. Hajipour, S. F. Aghamiri, H. Sabzyan, F. Seyedejn-Azad, Extension of the exp-6 model to the simulation of vapor-liquid equilibria of primary alcohols and their mixtures, *Fluid Phase Equilibria*, 301 (2011), 73-79.
- [21] J. Li, S. Sun, V. M. Calo, Monte Carlo molecular simulation of phase-coexistence for oil production and processing, *SPE Reservoir Characterization and Simulation Conference and Exhibition*, 2011, no. 148282.
- [22] A. Z. Panagiotopoulos, N. Quirke, M. R. Stapleton, D. J. Tildesley, Phase equilibria by simulations in the Gibbs ensemble: Alternative derivation, generalization and application to mixtures and membrane equilibria, *Mol. Phys.*, 63 (1988), 527-545.
- [23] V. I. Manousiouthakis, M. W. Deem, Strict detailed balance is unnecessary in Monte Carlo simulation, *J. Chem. Phys.*, 110 (1999), 2753-2756.
- [24] J. J. Nicolas, K. E. Gubbins, W. B. Streett, D. J. Tildesley, Equation of state for the Lennard-Jones fluid, *Mol. Phys.*, 37 (1979), 1429-1454.
- [25] H. Flyvbjerg, H. G. Petersen, Error estimates on averages of correlated data, *J. Chem. Phys.*, 91(1) (1989), 461-466.