HOKKAIDO UNIVERSITY

Title	Efficient sampling for ab initio Monte Carlo simulation of molecular clusters using an interpolated potential energy surface
Author(s)	Nakayama, Akira; Seki, Nanami; Taketsugu, Tetsuya
Citation	Journal of chemical physics, 130(2), 024107-1-024107-10 https://doi.org/10.1063/1.3055910
Issue Date	2009-01-13
Doc URL	https://hdl.handle.net/2115/56631
Rights	Copyright 2009 American Institute of Physics. This article may be downloaded for personal use only. Any other use requires prior permission of the author and the American Institute of Physics. The following article appeared in The Journal of Chemical Physics 130, 024107 (2009) and may be found at http://dx.doi.org/10.1063/1.3055910.
Type	journal article
File Information	JCP_130_024107_1.3055910.pdf



Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

AIP | The Journal of Chemical Physics

Efficient sampling for ab initio Monte Carlo simulation of molecular clusters using an interpolated potential energy surface

Akira Nakayama, Nanami Seki, and Tetsuya Taketsugu

Citation: The Journal of Chemical Physics **130**, 024107 (2009); doi: 10.1063/1.3055910
View online: http://dx.doi.org/10.1063/1.3055910
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/130/2?ver=pdfcov
Published by the AIP Publishing

Articles you may be interested in

A new four-dimensional ab initio potential energy surface for N2O–He and vibrational band origin shifts for the N2O–He N clusters with N = 1–40
J. Chem. Phys. **137**, 104311 (2012); 10.1063/1.4749248

Efficient ab initio path integral hybrid Monte Carlo based on the fourth-order Trotter expansion: Application to fluoride ion-water cluster
J. Chem. Phys. **132**, 144108 (2010); 10.1063/1.3367724

Interpolating moving least-squares methods for fitting potential energy surfaces: Computing high-density potential energy surface data from low-density ab initio data points
J. Chem. Phys. **126**, 184108 (2007); 10.1063/1.2730798

Hydrogen-bond assisted enormous broadening of infrared spectra of phenol-water cationic cluster: An ab initio mixed quantum-classical study
J. Chem. Phys. **126**, 074304 (2007); 10.1063/1.2434778

Potential energy surfaces for He n Ne + ions: ab initio and diatomics-in-molecule results
J. Chem. Phys. **120**, 7456 (2004); 10.1063/1.1665271

AIP | Journal of Applied Physics



Journal of Applied Physics is pleased to announce André Anders as its new Editor-in-Chief

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 133.50.96.34 On: Thu, 31 Jul 2014 05:38:21

Efficient sampling for *ab initio* Monte Carlo simulation of molecular clusters using an interpolated potential energy surface

Akira Nakayama,[a] Nanami Seki, and Tetsuya Taketsugu
Division of Chemistry, Graduate School of Science, Hokkaido University, Sapporo 060-0810, Japan

(Received 1 October 2008; accepted 3 December 2008; published online 13 January 2009)

An approach is developed to enhance sampling for *ab initio* Monte Carlo and *ab initio* path integral Monte Carlo calculations of molecular clusters by utilizing an approximate potential as a guide to move in the configuration space more efficiently. The interpolated potential energy obtained by the moving least-squares method is used as an approximate potential, and this scheme is applied to a water molecule and small protonated water clusters (H_3O^+,$H_5O_2^+$). It is found that the statistical errors are reduced by almost a factor of 3 in most calculations, which translates into a reduction in the computational cost by an order of magnitude. We also provide an automatic scheme where the *ab initio* data obtained during the simulation is added to the reference data set of interpolation dynamically, which further speeds up the convergence. © *2009 American Institute of Physics*.
[DOI: 10.1063/1.3055910]

I. INTRODUCTION

Simulations of complex molecular systems are generally performed by either the molecular dynamics (MD) or the Monte Carlo (MC) method.[1] With the rapid increase in computational power and refinement of the algorithm, they have proven to be useful tools for many areas of chemistry from drug design to materials science. These simulations require potential energy function as input, and those utilizing classical force fields are very efficient for large systems and good statistics for ensemble averages of physical properties can be obtained. However, such classical force fields are not appropriate in many cases. In particular, it is difficult to construct accurate empirical potential energy function for systems which undergo chemical reactions, where covalent bonds are formed or broken. Also, electronic polarization effects, which have a significant role in aqueous solutions, are poorly described in these potentials. In such cases, *ab initio* calculations are appropriate. The computational cost of *ab initio* calculations is usually higher than that of classical force fields by several orders of magnitude, and yet due to the methodological and computational advances *ab initio* MD or *ab initio* MC method, which solves electronic Schrödinger equation for nuclear potential energy (or its derivatives) as needed, is now being widely used to investigate static and also dynamic properties of molecular systems from first principles.

Although *ab initio* calculations are routinely performed for small molecules, they are still highly demanding when combined with statistical simulations since one needs to obtain the statistical averages over molecular configurations, which could be 10^5–10^6 configurations. Most of the computational time is spent on the numerical evaluation of *ab initio* potential (or its derivatives), thus simulations using sophisticated electronic structure methods are still limited to small system sizes. There is always a demand for efficient schemes in *ab initio* simulations, either by MD or MC.

A MC simulation is efficient in classical force fields, especially when the total potential is represented by a sum of the additive interactions. In each MC step, the pair interactions only between the "chosen" atom and other atoms are needed to evaluate, i.e., it is not necessary to evaluate the potential of the whole system. If one employs the standard approach for *ab initio* MC calculations, *ab initio* potential energy must be calculated in each MC step even if only one atom is chosen to move, which does not seem to be efficient compared to simulations with classical force fields. In contrast to *ab initio* MC calculations, *ab initio* MD has the advantage that the whole molecular configuration is updated in each time step. However, the evaluation of the *ab initio* potential energy gradients requires additional computational cost, which could be several times more than that of energy calculation. Furthermore, the module for the analytical energy gradients is not available for some highly correlated methods such as CCSD(T).

In this paper, we demonstrate that sampling efficiency for *ab initio* MC calculations is realized by utilizing an approximate potential (AP) as a guide to move in the configuration space more effectively. An auxiliary Markov chain is generated on this AP, and as an AP we adopt an interpolated potential energy surface that is constructed by the moving least-squares method.[2–4] This scheme is applied to a water molecule and small protonated water clusters (H_3O^+,$H_5O_2^+$), and the effectiveness of the present approach is analyzed in detail. Since quantum effects of nuclei are important for hydrogen atoms, we also perform *ab initio* path integral MC (PIMC) calculations for these systems, and the efficiency of the method is again discussed. In this work, potential energy at the reference point for interpolation is obtained from *ab initio* calculations at the same level of that used in the simulation, and this data set is prepared prior to the simulation. Since *ab initio* potential energy evaluated

[a]Electronic mail: akira-n@sci.hokudai.ac.jp.

 © 2009 American Institute of Physics

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 133.50.96.34 On: Thu, 31 Jul 2014 05:38:21

during the simulation can be used as a reference data as well, we also propose a novel scheme in which the AP is updated *dynamically* during the simulation.

This paper is organized as follows: Sec. II presents details of the AP method in connection with the multiple Markov chain (MMC) method, followed by the dynamical update scheme of the AP. Then we briefly describe the interpolating moving least-squares (IMLS) method for construction of the AP. Also we review the path integral MC method for the quantum treatment of nuclei. In Sec. III the results obtained by this method are presented, along with the detailed analysis of the efficiency of the method. Section IV contains the conclusions of this work.

II. METHODOLOGIES

A. Approximate potential method

The MMC methods[5] are originally used to enhance sampling in which Markov chains running at a higher temperature are used to promote transition among different regions of high probability density. The AP method exploits this idea and uses an auxiliary Markov chain that moves on the approximate and computationally less expensive potential. In this method, a MC importance sampling is performed on this AP, and after a fixed number of steps, the difference between the original and AP is used as a criterion to restore sampling with respect to the original potential. With this method, correlation in the Markov chain on the original potential is significantly reduced. Since the AP is chosen such that the computational cost for evaluation is much less expensive than the original one, significant computational saving can be attained. Iftimie *et al.*[6,7] utilized this idea in the *ab initio* simulation of proton transfer reactions and used molecular mechanics potentials as an AP. Gelb[8] applied this scheme to Lennard-Jones fluids and showed that the speedups by almost a factor of 4 are obtained. Bandyopadhyay[9] showed that this method can be used to determine stationary points on the potential energy surface of molecular systems.

In the AP method, the key to efficient sampling is finding a suitable AP. It is desirable that the AP is close to the original potential within a region of configuration space accessible at a given temperature. This results in a high acceptance ratio and ensures an effective sampling. When the AP is identical to the original potential, the acceptance ratio is always one. On the other hand, if the AP is appreciably different from the original one, the Markov chain on the AP may spend long periods of time in the region that is less significant for the system. In some cases, this could lead to less efficient sampling than the standard MC method.

We briefly describe the AP method for a canonical ensemble of classical system, whose probability density is given as

$$\rho(\mathbf{R}) = \frac{1}{Z}\exp[-\beta V(\mathbf{R})], \tag{1}$$

where $\beta = 1/k_B T$ is the reciprocal temperature in units of the Boltzmann constant. Here $\mathbf{R}$ is the configuration of the system and Z is the partition function. The extension to quantum systems using imaginary time path integral formulation, which will be described below, is straightforward. The procedure for the AP method is as follows. Suppose that the current configuration of the system is $\mathbf{R}^{(i)}$. An auxiliary Markov chain is generated from $\mathbf{R}^{(i)}$ with the probability density proportional to $\exp[-\beta V_{\mathrm{AP}}(\mathbf{R})]$ where V_{AP} is the AP. After a fixed number of updates on the AP, one attempts to move the system configuration $\mathbf{R}^{(i)}$ to the final configuration of the auxiliary Markov chain ($\mathbf{R}^{(j)}$) with probability

$$\min[1,\exp\{-\beta[(V(\mathbf{R}^{(j)}) - V_{\mathrm{AP}}(\mathbf{R}^{(j)})) - (V(\mathbf{R}^{(i)}) - V_{\mathrm{AP}}(\mathbf{R}^{(i)}))]\}]. \tag{2}$$

One can see that the transition from $\mathbf{R}^{(i)}$ to $\mathbf{R}^{(j)}$ is performed based on the difference between the original and APs. With this criterion the microscopic reversibility for the system configuration

$$\rho(\mathbf{R}^{(i)})P(\mathbf{R}^{(i)} \to \mathbf{R}^{(j)}) = \rho(\mathbf{R}^{(j)})P(\mathbf{R}^{(j)} \to \mathbf{R}^{(i)}) \tag{3}$$

is satisfied,[6,8] where P is the transition probability. The acceptance probability given in Eq. (2) is closely related to that of the parallel tempering method[5,10,11] (also known as the replica exchange method[12]). If the above move is accepted, another auxiliary Markov chain is generated from $\mathbf{R}^{(j)}$, and if the move is rejected, the current configuration of the system goes back to the original configuration $\mathbf{R}^{(i)}$ and another attempt is made. It is noted that the limiting distribution for the system does not depend on the form of the AP, provided that the updating scheme on the AP is ergodic. In the current implementation of *ab initio* MC, *ab initio* calculation is performed at the final configuration of the auxiliary Markov chain on the AP.

In this study, we employ the interpolated potential energy surface as an AP and the reference points used for interpolation are obtained prior to the MC run. Considering that *ab initio* calculations are performed during the simulation, this information can be also used as the reference points, and the interpolated surface should be more accurate if they are added to the existing data set. Thus, we also propose a novel approach, in which the AP is updated *dynamically* during the simulation. The proposed scheme is as follows. Suppose that the current configuration of the system is $\mathbf{R}^{(i)}$ and the number of reference points for interpolation is N_d. The interpolated surface constructed from these reference points is denoted by $V_{\mathrm{AP}}^{(N_d)}(\mathbf{R})$. An auxiliary Markov chain is generated from $\mathbf{R}^{(i)}$ using the potential $V_{\mathrm{AP}}^{(N_d)}(\mathbf{R})$, and an *ab initio* potential energy is calculated at the final configuration of the auxiliary Markov chain ($\mathbf{R}^{(j)}$). The move of the system configuration from $\mathbf{R}^{(i)}$ to $\mathbf{R}^{(j)}$ is determined based on the probability

$$\min[1,\exp\{-\beta[(V(\mathbf{R}^{(j)}) - V_{\mathrm{AP}}^{(N_d)}(\mathbf{R}^{(j)})) - (V(\mathbf{R}^{(i)}) - V_{\mathrm{AP}}^{(N_d)}(\mathbf{R}^{(i)}))]\}]. \tag{4}$$

Regardless of whether the trial move is accepted or rejected, *ab initio* potential energy is added to the data set and the new interpolated surface $V_{\mathrm{AP}}^{(N_d+1)}(\mathbf{R})$ is created. Then a new auxiliary Markov chain is started from the current configuration of the system $\mathbf{R}^{(k)}$ (if accepted $\mathbf{R}^{(k)}=\mathbf{R}^{(j)}$, if not $\mathbf{R}^{(k)}=\mathbf{R}^{(i)}$) using the potential $V_{\mathrm{AP}}^{(N_d+1)}(\mathbf{R})$, and after a fixed number of updates, the move of the system configuration from $\mathbf{R}^{(k)}$ to

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 133.50.96.34 On: Thu, 31 Jul 2014 05:38:21

$\mathbf{R}^{(l)}$ (which is the final configuration of the auxiliary Markov chain) is attempted on the basis of the probability

$$\min[1,\exp\{-\beta[(V(\mathbf{R}^{(l)})-V_{\mathrm{AP}}^{(N_d+1)}(\mathbf{R}^{(l)}))-(V(\mathbf{R}^{(k)})-V_{\mathrm{AP}}^{(N_d+1)}(\mathbf{R}^{(k)}))]\}]. \quad (5)$$

Again, the *ab initio* potential energy is added to the data set and the new interpolated surface $V_{\mathrm{AP}}^{(N_d+2)}(\mathbf{R})$ is obtained. In this way, the AP is dynamically improved during the simulation and high efficiency can be obtained. Note that the acceptance criterion given in Eqs. (4) and (5) leaves the Boltzmann distribution for the system invariant and correct limiting distribution is obtained, since the weaker (balance) condition for the system variables given in Eq. (3) is satisfied.[13]

With this method, one can start the simulation without reference points (or with a small number of reference points) and update the AP dynamically by accumulating the reference data. Of course, one cannot continue this process until the end of the simulation because the data set becomes extremely large. Thus, the maximum number of reference points should be set in advance. A large number of reference points always lead to an accurate potential, but it is offset by the expensive computational cost of evaluating the interpolated potential. Once the maximum number of reference points is reached, the sampling becomes the same as that of the original AP method and ensures ergodicity (provided that the sampling on the AP is ergodic).

B. The IMLS approach

The AP energy surface is constructed by the interpolation method using the moving least-squares approach.[2–4] The modified Shepard interpolation method[14] is widely used for construction of potential energy surface, and it is usually more efficient than the IMLS method. The modified Shepard interpolation method requires the derivatives normally up to the second order to obtain a reliable potential energy. The *ab initio* calculations of these derivatives at reference points are computationally expensive for some highly sophisticated electronic structure methods. The IMLS method, on the other hand, yields a sufficiently accurate potential only with the potential energy. Of course, one can use the derivatives to improve the accuracy of it,[15] but we do not take into account them. The IMLS is relatively expensive since one needs to solve a weighted-least-squares equation at each evaluation point, but still much less expensive than *ab initio* calculations by several orders of magnitude.

The interpolated potential energy at a nuclear configuration $\mathbf{Z}$ is represented by a linear combination of linearly independent basis functions $b_i(\mathbf{Z})$ as

$$V(\mathbf{Z})=\sum_{i=1}^{M} a_i b_i(\mathbf{Z}), \quad (6)$$

where M is the total number of basis functions. The coefficients a_i are obtained by minimizing the following weighted deviations:

$$D[V(\mathbf{Z})]=\sum_{i=1}^{N_d} w_i(\mathbf{Z})\left[\sum_{j=1}^{M} a_j b_j(\mathbf{Z}^{(i)})-V_{\mathrm{ref}}(\mathbf{Z}^{(i)})\right]^2. \quad (7)$$

Here, N_d is the number of reference data points and $V_{\mathrm{ref}}(\mathbf{Z}^{(i)})$ is the energy value at point $\mathbf{Z}^{(i)}$ in the data set. The weight function $w_i(\mathbf{Z})$ decays with distance $|\mathbf{Z}-\mathbf{Z}^{(i)}|$ so that the closer data points have a larger weight than the more distant ones. Note that the coefficients a_i are implicitly dependent on $\mathbf{Z}$ because of the weight function. In this study, the coordinates are taken to be the reciprocal internuclear distances (i.e., $Z_k=1/r_{ij}$), which is commonly used in the context of the IMLS interpolations and also in the modified Shepard interpolations.

The minimization condition with respect to a_i leads to the following equation:

$$\mathbf{B}^T\mathbf{W}\mathbf{B}\mathbf{a}=\mathbf{B}^T\mathbf{W}\mathbf{V}, \quad (8)$$

where $\mathbf{a}=(a_1,a_2,\ldots,a_M)^T$, $\mathbf{V}=(V(\mathbf{Z}^{(1)}),V(\mathbf{Z}^{(2)}),\ldots,V(\mathbf{Z}^{(N_d)}))^T$, and $\mathbf{W}$ is a diagonal matrix whose elements are $W_{ij}=w_i(\mathbf{Z})\delta_{ij}$. $\mathbf{B}$ is composed of values of the basis functions at $\mathbf{Z}^{(i)}$ and given by

$$\mathbf{B}=\begin{pmatrix} b_1(\mathbf{Z}^{(1)}) & b_2(\mathbf{Z}^{(1)}) & \cdots & b_M(\mathbf{Z}^{(1)}) \\ b_1(\mathbf{Z}^{(2)}) & b_2(\mathbf{Z}^{(2)}) & \cdots & b_M(\mathbf{Z}^{(2)}) \\ \cdots & \cdots & \cdots & \cdots \\ b_1(\mathbf{Z}^{(N_d)}) & b_2(\mathbf{Z}^{(N_d)}) & \cdots & b_M(\mathbf{Z}^{(N_d)}) \end{pmatrix}. \quad (9)$$

The interpolated potential for a given configuration $\mathbf{Z}$ is obtained by solving the normal equation Eq. (8), which is implemented by using the singular value decomposition.

In this work, we use polynomials up to the second order of each component of $\mathbf{Z}$ as the basis functions. For a system of N atoms, the number of atomic pairs is $N_I=N(N-1)/2$, thus there are $1+N_I+N_I(N_I+1)/2$ basis functions for polynomials up to the second order. The weight function of the form

$$w_i(\mathbf{Z})=\frac{1}{|\mathbf{Z}-\mathbf{Z}^{(i)}|^2} \quad (10)$$

is employed throughout this study. The identical particle symmetry can be enforced by extending the data set to include all possible permutations of each data point, but for simplicity we do not take this into account in the present work.

C. Path integral Monte Carlo method

The imaginary time path integral formulation of quantum statistics provides a conceptual and computationally practical route for studying the quantum nature of systems at thermal equilibrium.[16] Based on Feynman's notion that a quantum system can be mapped onto a classical model of interacting "polymers" with path integrals, the path integral MC method has proven to be extremely useful for studying finite temperature properties of many-particle systems. Quantum effects such as zero-point motion and tunneling as well as thermal fluctuations are included rigorously in this method, and it is now seeing widespread use in a variety of applications to chemical systems. A combination of path integral techniques with *ab initio* calculations has been real-

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 133.50.96.34 On: Thu, 31 Jul 2014 05:38:21

ized for systems where the parameterized potential (force fields) is not available, and it is implemented in some suite of codes available. Most of the calculations are done with density functional theory employing a plane-wave basis set, in particular, in the context of the Car–Parrinello approach.[17–19] The wave function based method, such as MP2, is now also used to evaluate the potential energy,[20,21] for which the density functional theory is not sufficiently accurate.

In the path integral MC calculations, the expectation value of an operator $\hat{A}$ corresponding to an observable at thermal equilibrium can be expressed as

$$\langle \hat{A} \rangle = \frac{1}{Z} \prod_{j=1}^{N} \left(\frac{m_j P}{2\pi\beta\hbar^2} \right)^{3P/2} \times \int d\mathbf{R}_1 \int d\mathbf{R}_2 \cdots \int d\mathbf{R}_P A(\mathbf{R}_1) \times \exp\left[-\sum_{j=1}^{N} \frac{m_j P}{2\beta\hbar^2} \sum_{i=1}^{P} |\mathbf{r}_i^{(j)} - \mathbf{r}_{i+1}^{(j)}|^2 - \frac{\beta}{P} \sum_{i=1}^{P} V(\mathbf{R}_i) \right], \tag{11}$$

where P is the number of time slices and $Z=\mathrm{Tr}(e^{-\beta\hat{H}})$ is the canonical partition function. Here we assume that the system of interest consists of N atoms with the Cartesian positions $\mathbf{R}=\{\mathbf{r}^{(1)},\mathbf{r}^{(2)},\ldots,\mathbf{r}^{(N)}\}$. In the above equation, $\hat{A}$ is assumed to be a function of the position operator and cyclic periodic boundary condition $\mathbf{R}_{P+1}=\mathbf{R}_1$ is imposed. A MC sampling is performed according to the distribution function given by

$$w(\mathbf{R};\beta) = \exp\left[-\sum_{j=1}^{N} \frac{m_j P}{2\beta\hbar^2} \sum_{i=1}^{P} |\mathbf{r}_i^{(j)} - \mathbf{r}_{i+1}^{(j)}|^2 - \frac{\beta}{P} \sum_{i=1}^{P} V(\mathbf{R}_i) \right]. \tag{12}$$

The path integral MD scheme can be also employed to obtain the statistical average of the above integral, and by combining variable transformation to normal mode coordinates with Nosé–Hoover chain thermostats and multiple time scale integration,[22] it has been proven that the PIMD method is nearly as efficient as the staging PIMC approach. The staging PIMC algorithm[23,24] is highly efficient as a sampling method in PIMC, thus we employ this technique in a MC sampling on the AP.

When applying the AP method to PIMC simulations, the acceptance probability for the update of the system configuration from $\mathbf{R}$ to $\mathbf{R}'$ is given by

$$\min\left[1, \exp\left\{ -\frac{\beta}{P} \left[\sum_{i=1}^{P} (V(\mathbf{R}_i') - V_{\mathrm{AP}}(\mathbf{R}_i')) - \sum_{i=1}^{P} (V(\mathbf{R}_i) - V_{\mathrm{AP}}(\mathbf{R}_i)) \right] \right\} \right]. \tag{13}$$

Note that the harmonic spring terms cancel out. Therefore, the acceptance criterion is based on the difference between the *ab initio* and AP at the current and trial configurations.

TABLE I. Statistical error (in kelvin) and efficiency for a total of 1 000 000 steps on a water molecule at 300 K with the SPC/F2 model. The statistical error is estimated by block averages. The efficiency is relative to that of the standard calculation. The acceptance ratio is also shown.

(N_s,N_p)	σ_V	Efficiency	Acceptance ratio
(5,20)	0.73	5.0	0.976
(5,100)	0.71	5.2	0.979
(10,20)	0.52	9.8	0.976
(10,100)	0.56	8.6	0.979
(20,20)	0.41	15.6	0.976
(20,100)	0.39	17.1	0.979
Standard	1.63	1.0	···

III. RESULTS AND DISCUSSION

A. Pilot calculations on a water molecule with the SPC/F2 model

In order to demonstrate the efficiency of our method, pilot calculations are performed for a water molecule at 300 K with the simple point-charge/flexible 2 (SPC/F2) potential model,[25] instead of performing "on-the-fly" *ab initio* calculations. The error estimates and the adjustments of parameters using this model can be a guide to the following on-the-fly *ab initio* MC and *ab initio* PIMC calculations.

We first present the results of classical MC simulations. The reference points for interpolation were sampled by classical MD calculations at a constant temperature of 300 K with the massive Nosé–Hoover thermostat chains. The time step of 1.0 fs was employed, and all configurations and potential energies along the trajectory were stored as the reference data. This time step is not appropriate from the viewpoint of the energy conservation, but it can explore the configuration space more quickly. A total of 100 reference points were stored.

Table I shows the statistical error of the potential energy for MC simulations, in which a total of 1 000 000 steps are taken. The statistical errors are estimated by block averages. Since the intramolecular potential of the SPC/F2 model is harmonic, the potential energy at $T=300$ K is estimated to be $\langle V \rangle = 450$ K. The simple single-particle Metropolis MC algorithm was employed in the Markov chain on the AP and also in the standard calculations for comparison, and the step size was adjusted to produce an acceptance ratio of approximately 0.5. Note that in the AP method a fixed number of steps N_S on the AP are taken in a single step on the original potential, thus it takes much longer time to perform the calculations than the standard Metropolis MC for the same number of MC steps. However, in the following *ab initio* MC calculations, *ab initio* calculation takes much longer time than the IMLS potential evaluation by orders of magnitude. Thus, it is appropriate to discuss the effectiveness using the number of potential evaluations. In classical MC calculations, a single *ab initio* calculation is performed in one MC step. In the process of IMLS, a total of 100 reference points are stored and the nearest N_p points from an evaluation point are used for the fitting. Using large N_p leads to higher accuracy of the AP.

The efficiency relative to that of the standard method,

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 133.50.96.34 On: Thu, 31 Jul 2014 05:38:21

TABLE II. Average of potential $\langle V \rangle$ and kinetic energies $\langle K \rangle$ and these statistical errors (in kelvin) for a total of 50 000 000 potential calculations on a water molecule at 300 K with the SPC/F2 model. L is the length of the path segment in the staging algorithm. The statistical error is estimated by block averages, and the efficiency is the relative to that of $L=16$.

L	$\langle V \rangle$	σ_V	Efficiency	$\langle K \rangle$	σ_K	Efficiency
2	2883.03	1.01	2.5	3755.03	3.20	3.5
4	2883.96	0.81	3.9	3752.22	2.93	4.2
6	2883.10	0.79	4.1	3757.21	3.10	3.7
8	2884.06	0.87	3.4	3757.20	3.07	3.8
12	2880.04	1.10	2.1	3756.54	4.55	1.7
16	2883.33	1.60	1.0	3746.00	5.98	1.0

which is proportional to the inverse of the square of the statistical error, is also displayed in Table I. As seen in the table, with $N_s=10$, $N_p=20$, the statistical error is reduced by almost a factor of 3, and an increase in efficiency by about an order of magnitude is observed. The relatively large statistical error for calculations with $N_s=5$ reflects the fact that there remains some correlation in the Markov chain on the system potential.

Next, we investigate the performance of the AP method in PIMC calculations. We first determine the optimal length of the path segment in the staging algorithm in the context of *ab initio* PIMC calculations, in order to make a proper comparison with the AP method. As mentioned earlier, most of the simulation time in the on-the-fly *ab initio* approach is spent on the *ab initio* part, thus it is natural to discuss the efficiency of the methodologies based on the number of *ab initio* calculations. In the staging algorithm, the path segment with length L is clipped out and positions of L-1 "beads" are updated in a single step, which requires L-1 potential evaluations. The larger L can provide a faster move in configuration space, but it gives less probability that the trial move is accepted.

The staging PIMC calculations were performed with $P=24$, where the well-converged results with respect to the number of beads are achieved at $T=300$ K, employing different L for a fixed total number of potential energy calculations (50 000 000 calculations). Table II shows the average of potential and kinetic energies, along with its statistical errors and efficiencies. Note that the number of MC steps is different for each L since the potential energies are evaluated L-1 times in a single step (i.e., 50 000 000/(L-1) PIMC steps are taken for length L). As seen in the table, the best performance is obtained with $L=4$ or 6 for energy estimators. Hereafter, we employ $L=6$ for comparison with the AP method. (In general, the performance with large L would be better than this example since this system is strongly bounded. If a system is more floppy, the optimal length would be larger.)

In order to generate reference points for the IMLS method in PIMC calculations, MD calculations were performed in the same way as in the classical MC calculations, except that the temperature was set much higher than the system temperature, in order to cover the broader configuration space for quantum simulation, where higher energy region is frequently sampled due to quantum zero-point motion. The temperature was set according to the scaled temperature[26] given by

$$T' = \frac{\hbar\omega}{2k_B}\left[\tanh\left(\frac{\hbar\omega}{2k_BT}\right)\right]^{-1}, \tag{14}$$

which becomes approximately $T'=2700$ K when ω is set to the OH vibrational frequency of a water molecule. The MD calculation at a constant temperature of 2700 K with a time step of 1.0 fs was performed and 100 reference points are stored in total.

Table III shows the potential and kinetic energies, along with these statistical errors and relative efficiencies for a PIMC run of 1 000 000 steps. The staging algorithm with path length $L=6$ was employed and N_s steps were taken in the Markov chain on the AP in a single MC step. For comparison, the standard PIMC simulation with the staging algorithm employing $L=6$ was performed with the same total number of potential calculations (it corresponds to $1\,000\,000\times P/(L\text{-}1)$ MC steps). It is shown that an increase in efficiency by roughly a factor of 6 is obtained with at least $N_s=48$, leading to a significant reduction in the computational cost.

TABLE III. Average of potential $\langle V \rangle$ and kinetic energies $\langle K \rangle$ and these statistical errors (in kelvin) for a total of 1 000 000 steps on a water molecule at 300 K with the SPC/F2 model. N_s is the number of steps on the AP. The statistical error is estimated by block averages, and the efficiency is relative to that of the standard calculation.

N_s	$\langle V \rangle$	σ_V	Efficiency	$\langle K \rangle$	σ_K	Efficiency	Acceptance ratio
24	2883.71	1.51	3.8	3763.36	6.82	2.2	0.835
48	2883.50	1.22	5.8	3756.12	5.24	3.7	0.835
72	2883.37	1.28	5.3	3758.12	3.96	6.4	0.835
96	2883.14	1.12	6.9	3750.73	4.30	5.5	0.835
Standard	2881.74	2.94	1.0	3747.35	10.05	1.0	···

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 133.50.96.34 On: Thu, 31 Jul 2014 05:38:21

TABLE IV. Average of geometrical parameters and energies along with these statistical errors and relative efficiencies for *ab initio* classical MC (1 000 000 steps) and *ab initio* PIMC (100 000 steps) calculations on a H_2O molecule. AP is the result of calculation with the AP. The equilibrium geometrical parameters are $R_{OH}=0.9640$ Å and $\theta=104.00$ degrees.

	Classical (AP)	Classical (standard)	Quantum (AP)	Quantum (standard)
$\langle R_{OH}\rangle$/Å	0.9669	···	0.9817	···
$\langle\theta\rangle$/deg	103.95	···	103.79	···
$\langle V\rangle$/K	454.0	452.4	3180.1	3170.7
σ_V/K	0.6	1.9	3.4	10.5
Efficiency	8.9	1.0	9.9	1.0
$\langle K\rangle$/K	···	···	3960.8	4077.6
σ_K/K	···	···	23.4	73.3
Efficiency	···	···	9.8	1.0

B. *Ab initio* simulations on H_2O

On-the-fly *ab initio* MC calculations were performed on a H_2O molecule. The *ab initio* calculations were performed at the MP2 level of theory using the MOLPRO2006 package.[27] The segmented DZP basis set augmented with diffuse functions[28] was used throughout this paper. The reference points were generated in the same way as in the previous subsection, where the *ab initio* MD calculation was carried out and 100 reference points were stored in total. The *ab initio* MC calculation with $N_s=10$ and $N_p=100$ was performed for a total of 1 000 000 steps.

Table IV shows the obtained average of the geometrical parameters and potential energies along with these statistical errors and relative efficiencies. The result of the standard MC calculation is also displayed for comparison. The zero of the potential energy corresponds to the energy at the equilibrium structure.

The *ab initio* PIMC calculations were also carried out for a total of 100 000 steps. Reference points were generated by the *ab initio* MD simulation at $T=2700$ K as described above. A MC sampling on the AP was performed with $N_s=96$ and $N_p=100$ using the staging algorithm with path length $L=6$.

The obtained results are summarized in Table IV, along with the result of the standard PIMC method using the staging algorithm with $L=6$. Since we employed the primitive form for the kinetic energy estimator, the large statistical error for kinetic energy was observed in comparison to the potential energy. The virial estimator is preferable for the kinetic energy, but since it requires potential energy derivatives, we have not employed it here. The acceptance ratio for the update of the system configuration is around 0.91 with a current set of parameters, which is also a measure of the accuracy of the IMLS potential.

In this example on H_2O, an increase in efficiency by roughly an order of magnitude is obtained for both the classical and quantum simulations. It is noted that the above AP simulations take only around 1%–2% more in CPU time than that of the corresponding standard calculations.

C. *Ab initio* simulations on H_3O^+

The hydronium cation H_3O^+ has been studied extensively in relation to Eigen's proton transfer mechanism in water. This cation is a floppy molecule and involves the Walden inversion between pyramidal structures in the C_{3v} symmetry. Previous calculations by *ab initio* path integral MD showed that the Walden inversion is enhanced significantly by the inclusion of quantum effects of nuclei.[20]

Before presenting the results of *ab initio* simulations, we examined the accuracy of *ab initio* calculations for a quantitative description of the potential energy surface, particularly with regard to the Walden inversion. Table V summarized the results of *ab initio* calculations employing various electronic structure methods and basis sets. As can be seen, the barrier height between the pyramidal structures is only a few kcal/mol and in order to obtain reliable potential energy, the electronic correlation should be taken into account carefully. It is noted that DFT calculations tend to underestimate the barrier height, which is often seen in the hydrogen-bonded systems. Our result is in good agreement with highly correlated method (CCSD(T)) with a large basis set and also with the

TABLE V. Equilibrium (C_{3v}) and transition state (D_{3h}) geometrical parameters and relative energies between the two structures. The experimental data are taken from Ref. 29.

Method	Basis set	R_{OH}/Å (C_{3v})	θ/deg (C_{3v})	R_{OH}/Å (D_{3h})	ΔE/kcal/mol
MP2	aug-cc-pVDZ	0.9831	111.03	0.9756	2.221
	aug-cc-pVTZ	0.9795	111.49	0.9716	2.185
	aug-cc-pVQZ	0.9770	111.74	0.9696	1.944
	aug-cc-pV5Z	0.9764	111.88	0.9692	1.874
CCSD(T)	aug-cc-pVDZ	0.9832	110.95	0.9755	2.344
	aug-cc-pVTZ	0.9793	111.40	0.9713	2.296
	aug-cc-pVQZ	0.9766	111.70	0.9692	2.020
	aug-cc-pV5Z	0.9760	111.83	0.9688	1.953
DFT(B3LYP)	aug-cc-pVDZ	0.9830	112.51	0.9771	1.380
	aug-cc-pVTZ	0.9806	112.75	0.9743	1.479
	aug-cc-pVQZ	0.9794	112.82	0.9733	1.405
	aug-cc-pV5Z	0.9792	112.87	0.9731	1.382
MP2[a]	DZP+diffuse[a]	0.9820	110.98	0.9737	1.863
Expt.		0.9758	111.3	0.9609	1.92

[a]The *ab initio* MC and PIMC calculations were performed with this method and basis set.

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 133.50.96.34 On: Thu, 31 Jul 2014 05:38:21

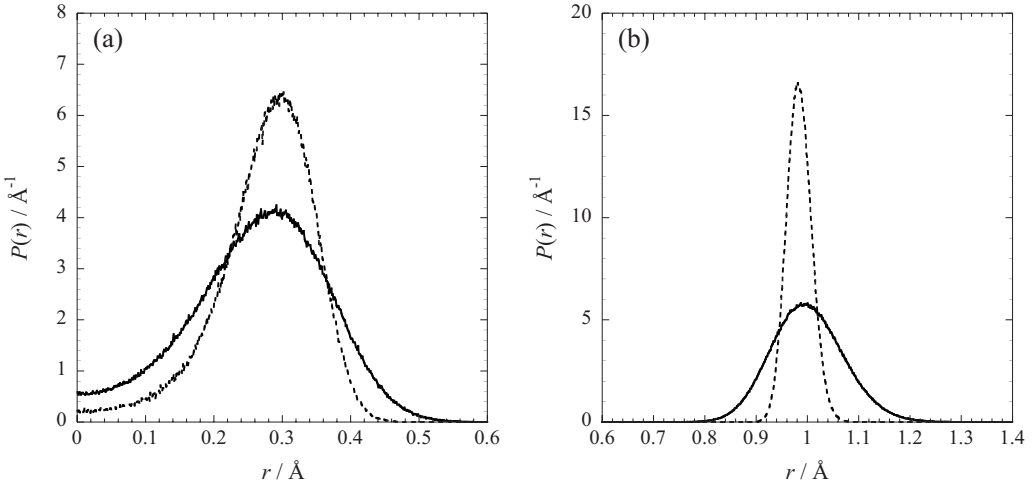


FIG. 1. Distribution functions of the distance between (a) the O atom and H–H–H plane and (b) the O and H atoms. The solid and dashed lines are the results of PIMC and classical MC calculations, respectively.

experimental results[29] for both the geometrical parameters and barrier height.

The *ab initio* MC and PIMC calculations were performed with $N_s=20$ and $N_s=96$ for a total of 1 000 000 and 100 000 steps, respectively. A total of 500 reference points were generated by the MD simulation at $T=2700$ K, and the nearest $N_p=100$ points were used for the fitting. Figure 1 shows the distribution function of the distance between the O atom and H–H–H plane and between the O and H atoms. The increased value at zero for quantum simulation in Fig. 1(a) reflects the fact that the inversion is enhanced significantly by the quantum treatment of nuclei.

Again, in order to demonstrate the effectiveness of our method, the standard MC and staging PIMC calculations were performed for 1 000 000 and $100\,000\times P/(L\text{-}1)$ steps, respectively. Table VI summarizes the obtained geometrical parameters and energies, along with these statistical errors and relative efficiencies. As seen in the table, an increase in efficiency by an order of magnitude is obtained for the expectation values of energy. The acceptance ratio for the update of the system configuration was ~0.81 and ~0.82 for classical MC and PIMC simulations, respectively. It is noteworthy that simulation with the acceptance ratio of around 0.8 still gives a significant reduction in computational cost. One can improve this rate further by increasing reference points. With the current set of parameters, the computational cost is increased by only 1%–2% compared to the standard calculation.

TABLE VI. Average of geometrical parameters and energies, along with these statistical errors (in kelvin) and relative efficiencies for *ab initio* simulations on H_3O^+. AP is the result of calculation with the AP.

	Classical (AP)	Classical (standard)	Quantum (AP)	Quantum (standard)
$\langle R_{OH}\rangle$/Å	0.9845		1.0007	
$\langle\theta\rangle$/deg	112.08		112.31	
$\langle V\rangle$/K	928.2	932.9	5194.5	5183.5
σ_V/K	1.1	3.8	4.7	13.7
Efficiency	12.9	1.0	8.4	1.0
$\langle K\rangle$/K	···	···	5826.1	6035.8
σ_K/K	···	···	28.4	93.4
Efficency			10.8	1.0

D. *Ab initio* simulations on $H_5O_2^+$

Our final application is to a protonated water dimer $H_5O_2^+$. This molecule is known as the Zundel cation and has been studied extensively over the past 40 years both experimentally and theoretically. It is motivated by its role in proton solvation and transfer processes in aqueous solutions. High-level *ab initio* calculations predict that the global minimum structure is in the C_2 symmetry,[30] where the shared proton is between and equidistant from the O atoms of water monomers, as illustrated in Fig. 2. Various quantum calculations including variational calculations,[31] diffusion Monte Carlo,[31] and path integral MD calculations[17,21] have been reported, which use *ab initio* potentials obtained either by directly using on-the-fly approach or after fitting to an analytical functional form.

Since we have discussed classical MC calculations extensively thus far, only PIMC calculations for $H_5O_2^+$ are presented in this subsection. The results of classical MC calculations are given in the next subsection in order to demonstrate the update scheme of the AP method.

The *ab initio* PIMC calculations were performed with $N_s=198$ for a total of 50 000 steps. A total of 3 000 reference points were prepared as a reference data set, and the nearest $N_p=500$ points were used for the fitting. Figure 3 shows the

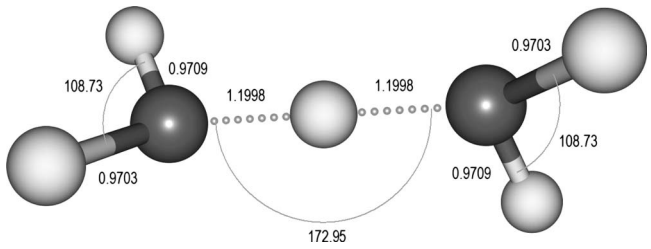


FIG. 2. Equilibrium geometry of $H_5O_2^+$ in the C_2 symmetry. Interatomic distances and bond angles are given in Å and degrees, respectively.

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 133.50.96.34 On: Thu, 31 Jul 2014 05:38:21

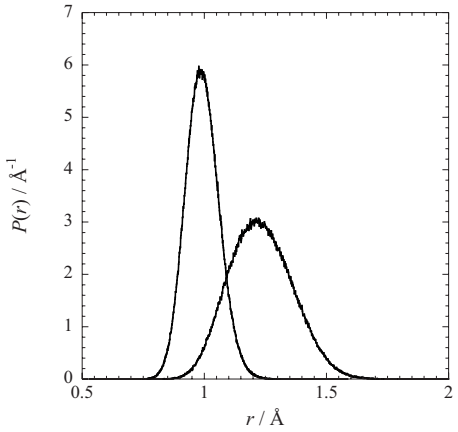


FIG. 3. Distribution function of the six OH distances.

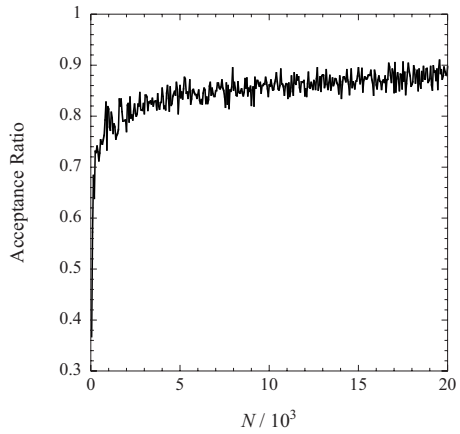


FIG. 4. Average of the acceptance ratio over the 16 independent MC run as a function of the number of steps.

distribution function of the six OH distances. The peak at shorter OH distance corresponds to the four OH bonds in two water molecules, whereas the broader peak at longer distances reflects the distance between the shared H and each of the O atoms. Table VII summarizes the obtained geometrical parameters, energies, and these statistical errors and efficiencies along with those of the standard calculations. Here the four distributions have an average OH bond length of 0.989 Å, which is ~0.02 Å larger than the equilibrium distance, while the average distance between the shared H and O atoms is 1.230 Å, which is ~0.03 Å larger than the equilibrium value. These observations are consistent with the previous calculations.[17,21] An increase in efficiency by roughly an order of magnitude is obtained for the expectation values of energies, and the acceptance ratio was ~0.74. A significant reduction in computational cost is again observed in this simulation.

E. Dynamical update scheme

The proposed dynamical update scheme of the AP is tested on a classical simulation of a $H_5O_2^+$ molecule. The basis set for *ab initio* calculations was changed to cc-pVDZ in this calculation in order to save computer time. A total of 16 independent classical MC run with 20 000 steps were carried out, starting from a data set with 1 000 reference points, in which the parameters N_s=40 and N_p=500 were used. During the simulation, the result of *ab initio* calculation was added to the data set in each MC step. Figure 4 shows the average of the acceptance ratio over the above 16 MC run as a function of the number of steps. As seen in the figure, after an initial sharp rise in the acceptance ratio from ~0.4, a very slow increase was observed, reaching ~0.9 after 20 000 steps. Although the high acceptance ratio leads to efficient sampling for the Markov chain on the system potential, it is not advisable to continue this process until the end of the simulation since the computational effort required for the IMLS process is increased and the tradeoff is made.

Finally, we investigate the efficiency as a function of the acceptance ratio by employing a different number of reference points. Figure 5 shows the relative efficiency for the potential energy as a function of the acceptance ratio for a total of 500 000 MC steps. The value at zero corresponds to the efficiency of the standard method. The relative efficiency is roughly the same as that obtained in the applications presented in this work for a given acceptance ratio. If the acceptance ratio is around 0.5, it will reduce the computational cost by a factor of 5, and the acceptance ratio of 0.7~0.8 will make a reduction by an order of magnitude. It is a somewhat subtle issue to estimate the optimal value of the acceptance ratio since it depends on the ratio between the cost of

TABLE VII. Average of geometrical parameters and energies, along with these statistical errors and relative efficiencies for *ab initio* PIMC simulations on $H_5O_2^+$. $\langle R_{OH}\rangle$ corresponds to the four OH bonds in two water molecules, and $\langle R_{OH^*}\rangle$ reflects the distance between the shared H and each of the O atoms.

	Quantum (AP)	Quantum (standard)
$\langle R_{OH}\rangle$/Å	0.989	···
$\langle R_{OH^*}\rangle$/Å	1.230	···
$\langle V\rangle$/K	8773.4	8695.8
σ_V/K	12.2	43.4
Efficiency	12.7	1.0
$\langle K\rangle$/K	9714.3	9566.7
σ_K/K	49.2	139.4
Efficiency	8.0	1.0

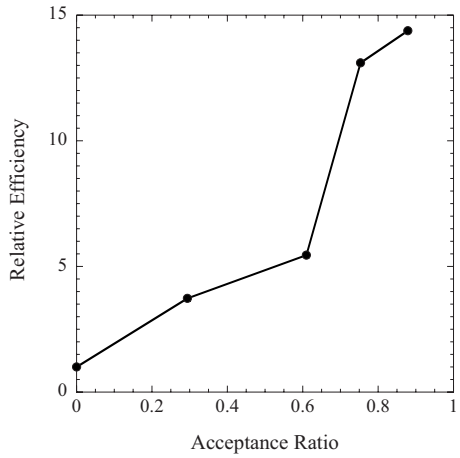


FIG. 5. The relative efficiency as a function of the acceptance ratio for a total of 500 000 MC steps. The value at zero corresponds to the efficiency of the standard method.

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 133.50.96.34 On: Thu, 31 Jul 2014 05:38:21

an *ab initio* calculation and that of performing one cycle of auxiliary Markov walk on the AP, but this figure would be a guide for choosing the number of reference points. In the above application on $H_5O_2^+$, we suggest that one starts a simulation with a small data set and continues the updating scheme until the total reference points reaches ~5 000. After that, the AP method with a fixed number of reference points would be enough. One can, for example, write a program code which monitors the acceptance ratio and stops the updating scheme automatically after the increase is damped.

IV. CONCLUDING REMARKS

In this paper, we demonstrated that the efficiency of *ab initio* MC and *ab initio* path integral MC calculation is enhanced significantly using a suitable AP. The interpolated potential energy obtained by the IMLS method was used as an AP, and this scheme was applied to a water molecule and small protonated water clusters. It was observed that the statistical errors were reduced by almost a factor of 3 in most of the calculations presented in this work, which translates into a reduction of the computational cost by an order of magnitude. We also provided an automatic scheme where the *ab initio* data obtained during the simulation is added to the reference data, which further speeds up the convergence.

One of the attractive features of the method is that potential energy gradients are not required in the calculations, which enables one to perform *ab initio* simulations where the module for analytical gradients is not available. This will open the door for quantum simulations with highly accurate *ab initio* potential. Another feature is that one can use any form of the MC sampling on the AP. If the system has the multiple minima separated by high energy barriers, it usually exhibits a quasiergodic behavior. This problem can be circumvented by employing the advanced sampling method, such as J-walking,[32] multicanonical sampling,[33] and parallel tempering method.[5,10,11]

Here we should comment on the scalability of the method. The computational cost for evaluating the IMLS potential goes as N_pM^2, where N_p is the number of reference data points used for the fitting and M is the number of basis functions. Therefore, it is important to reduce the number of basis functions. A basis set of polynomials up to a given order grows rapidly with dimension and, furthermore, high order multicoordinate terms sometimes worsen the conditioning of the least-squares matrix. Thus, it is recommended to restrict the order of the basis functions. In this work, we have used a primitive form for the basis functions. Instead, one can employ sophisticated methods such as the high-dimensional model representation[34] to construct the cost-effective basis functions. Also, we have not taken into account the identical particle symmetry in the basis functions. The inclusion of this property will further reduce the computational cost. With regard to the number of reference points, it is of course not advisable to add a new data if close points are already included in the reference data set. Although the number of reference points would grow rapidly with dimension in order to attain a given acceptance ratio, it is noted that not all the reference data are used for the fitting due to the weight function.

Admittedly, our scheme has unfavorable scaling with respect to the number of atoms for the construction of AP, yet the limiting factor of the simulation is still the number of *ab initio* calculations. It is a demanding task to obtain the statistical average over molecular configurations using *ab initio* potential energies. A number of approximate methods have been devised to be applicable to large systems, and our scheme could be used in some of these methodologies. For example, one could apply our method to the fragment molecular orbital method[35] when combined with MC simulations. In this case, each fragment and fragment pair potentials are approximated by the interpolated potentials and the AP for the whole system is constructed from these interpolated fragment potentials. The scalability and the application to large systems will be investigated in future work.

Ab initio simulations have the strength in which one can have direct access to the electronic property of molecules. The detailed understanding of these properties, such as electronic charge or dipole moment, will provide deeper insight into the microscopic solvation structure. We are also planning to apply this scheme to various protonated water clusters $H^+(H_2O)_n$, hydroxylated water clusters $OH^-(H_2O)_n$, and negatively charged water clusters $(H_2O)_n^-$, which are of great interest since they serve as models of solvation of ions and excess electrons in water.

ACKNOWLEDGMENTS

A.N. acknowledges financial support by a Grant-in-Aid for Young Scientists from the Ministry of Education, Culture, Sports, Science and Technology (MEXT, Japan). We thank the anonymous referee for helpful comments for improving the paper.

[1] D. Frenkel and B. Smit, *Understanding Molecular Simulation* (Academic, San Diego, 1996).
[2] T. Ishida and G. C. Schatz, Chem. Phys. Lett. **314**, 369 (1999).
[3] G. G. Maisuradze, D. L. Thompson, A. F. Wagner, and M. Minkoff, J. Chem. Phys. **119**, 10002 (2003).
[4] A. Kawano, Y. Guo, D. L. Thompson, A. F. Wagner, and M. Minkoff, J. Chem. Phys. **120**, 6414 (2004).
[5] C. J. Geyer and E. A. Thompson, J. Am. Stat. Assoc. **90**, 909 (1995).
[6] R. Iftimie, D. Salahub, D. Q. Wei, and J. Schofield, J. Chem. Phys. **113**, 4852 (2000).
[7] R. Iftimie and J. Schofield, J. Chem. Phys. **114**, 6763 (2001).
[8] L. D. Gelb, J. Chem. Phys. **118**, 7747 (2003).
[9] P. Bandyopadhyay, J. Chem. Phys. **122**, 091102 (2005).
[10] A. P. Lyubartsev, A. A. Martsinovski, S. V. Shevkunov, and P. N. Vorontsov-Velyaminov, J. Chem. Phys. **96**, 1776 (1992).
[11] E. Marinari and G. Parisi, Europhys. Lett. **19**, 451 (1992).
[12] K. Hukushima and K. Nemoto, J. Phys. Soc. Jpn. **65**, 1604 (1996).
[13] V. I. Manousiouthakis and M. W. Deem, J. Chem. Phys. **110**, 2753 (1999).
[14] J. Ischtwan and M. A. Collins, J. Chem. Phys. **100**, 8080 (1994).
[15] T. Xie and J. M. Bowman, J. Chem. Phys. **117**, 10487 (2002).
[16] R. P. Feynman, *Statistical Mechanics* (Addison-Wesley, Redwood City, 1972).
[17] M. E. Tuckerman, D. Marx, M. L. Klein, and M. Parrinello, Science **275**, 817 (1997).
[18] D. Marx, M. E. Tuckerman, J. Hutter, and M. Parrinello, Nature (London) **397**, 601 (1999).

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 133.50.96.34 On: Thu, 31 Jul 2014 05:38:21

[19] M. E. Tuckerman, D. Marx, and M. Parrinello, Nature (London) **417**, 925 (2002).
[20] M. Tachikawa and M. Shiga, J. Chem. Phys. **121**, 5985 (2004).
[21] M. Tachikawa and M. Shiga, J. Am. Chem. Soc. **127**, 11908 (2005).
[22] M. E. Tuckerman, B. J. Berne, G. J. Martyna, and M. L. Klein, J. Chem. Phys. **99**, 2796 (1993).
[23] E. L. Pollock and D. M. Ceperley, Phys. Rev. B **30**, 2555 (1984).
[24] M. Sprik, M. L. Klein, and D. Chandler, Phys. Rev. B **31**, 4234 (1985).
[25] J. Lobaugh and G. A. Voth, J. Chem. Phys. **106**, 2400 (1997).
[26] J. P. Bergsma, P. H. Berens, K. R. Wilson, D. R. Fredkin, and E. J. Heller, J. Phys. Chem. **88**, 612 (1984).
[27] H.-J. Werner, P. J. Knowles, R. Lindh *et al.*, MOLPRO, version 2006.1, a package of *ab initio* programs.
[28] See: http://setani.sci.hokudai.ac.jp/sapporo/
[29] D. J. Liu, T. Oka, and T. J. Sears, J. Chem. Phys. **84**, 1312 (1986).
[30] E. F. Valeev and H. F. Schaefer III, J. Chem. Phys. **108**, 7197 (1998).
[31] A. B. McCoy, X. C. Huang, S. Carter, M. Y. Landeweer, and J. M. Bowman, J. Chem. Phys. **122**, 061101 (2005).
[32] D. D. Frantz, D. L. Freeman, and J. D. Doll, J. Chem. Phys. **93**, 2769 (1990).
[33] B. A. Berg and T. Neuhaus, Phys. Lett. B **267**, 249 (1991).
[34] G. Y. Li, C. Rosenthal, and H. Rabitz, J. Phys. Chem. A **105**, 7765 (2001).
[35] K. Kitaura, E. Ikeo, T. Asada, T. Nakano, and M. Uebayasi, Chem. Phys. Lett. **313**, 701 (1999).

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 133.50.96.34 On: Thu, 31 Jul 2014 05:38:21