We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS BOOK CITATION INDEX INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

6

Developing a Systematic Approach for *Ab Initio* Path-Integral Simulations

Kin-Yiu Wong
BioMaPS Institute for Quantitative Biology
Rutgers, The State University of New Jersey
USA

1. Introduction

An ultimate level of theory in molecular simulations [e.g., molecular dynamics (MD) and Monte Carlo (MC) simulations], which can accurately reproduce or even predict many experimental values, should be *ab initio* path integral. In *ab initio* path-integral simulations, both electrons and nuclei are treated quantum mechanically and adiabatically. No empirical parameter is involved, other than those fundamental physical constants (e.g., electronic mass and Planck's constant). The only *inherent* approximations are the Born-Oppenheimer approximation (to decouple internuclear dynamics from electronic motions) and the ergodicity in MD simulations or the importance samplings in MC simulations (to partly integrate the entire phase space). Consequently, correlation energy among electrons, anharmonic zero-point motions and tunnelling effects in nuclei, and isotope effects can all be incorporated in the simulations. Proper consideration of the electronic and internuclear quantum effects, even just partially, can be critical to compare computed values with state-of-the-art experiments, e.g., (I) hydrogen adsorption in carbon nanotechnology (Tanaka, Kanoh et al. 2005; Kowalczyk, Gauden et al. 2007; Kowalczyk, Gauden et al. 2008); (II) electronic redistributions and isotope effects (Wong and Gao 2007; Wong and Gao 2008; Wong, Richard et al. 2009; Gao and Wong 2008) on biochemical reactions in protein (Wong and Gao 2007; Wong and Gao 2011; Wu and Wong 2009; Warshel, Olsson et al. 2006; Gao, Major et al. 2008; Major, Heroux et al. 2009) and RNA enzymes (Wong, Lee et al. 2011; Wong, Gu et al. 2012).

However, owing to the extraordinarily high computational cost, *ab initio* path-integral simulations are thus far not practical even for modest size molecules, and are limited to only some relatively simpler or smaller molecular systems, e.g., thirty-two water molecules, and malonaldehyde [$CH_2(CHO)_2$]. Nevertheless, the unique information and invaluable insight for a molecular system, which can be provided perhaps only from *ab initio* path-integral simulations, have already been recognized in a number of pure computational publications in some high-profile journals, e.g., *Nature*, *Science*, and *Physical Review Letters*, etc (Marx and Parrinello 1995; Tuckerman, Marx et al. 1997; Marx, Tuckerman et al. 1999; Tuckerman and Marx 2001; Tuckerman, Marx et al. 2002; Ohta, Ohta et al. 2004; Hayashi, Shiga et al. 2006; Paesani, Iuchi et al. 2007).

In this chapter, after quickly going over the fundamental physical laws tailoring MD simulations, we (wongky@biomaps.rutgers.edu; kiniu@alumni.cuhk.net) discuss a new

theoretical method that combines our novel systematic free-energy expansion approach, based on Zwanzig's free-energy perturbation theory, with our recently developed automated integration-free path-integral method, based on Kleinert's variational perturbation theory, (Wong and Gao 2007; Wong and Gao 2008; Wong, Richard et al. 2009; Wong, Gu et al. 2012) to perform *ab initio* path-integral simulations for realistic macromolecules at an affordable computational cost. Since in this new method, we can progressively choose computationally affordable levels of theory, now important physical quantities, e.g., free-energy barrier, change of binding energy, pK_a value, and isotope effect, can all be computed at an *ab initio* path-integral level. Therefore, we anticipate this new systematic approach will become an essential computational tool to catch up with or even predict experimental results for breaking down subtle mechanisms underlying a variety of molecular systems in Life and Materials Sciences.

2. Fundamental physical laws governing molecular dynamics simulations

In this section, we lay the theoretical foundation for molecular dynamics (MD) simulations.

2.1 Molecular Schrödinger equation

Ever since quantum mechanics was constructed in the 1920s, solving the non-relativistic time-independent Schrödinger equation for a system of nuclei and electrons has become an essential step to understand *every single* detail of atomic or molecular properties (Kleppner and Jackiw 2000). The non-relativistic time-independent Schrödinger equation for a molecular system (the molecular Schrödinger equation) is:

$$\hat{H}_{mole}\Psi_n = E_n\Psi_n, \tag{1}$$

where $\hat{H}_{mole}$ is the complete (non-relativistic) molecular Hamiltonian, Ψ_n and E_n are an energy eigenfunction (or wave function) and an energy eigenvalue at an eigenstate n, respectively. In contrast to the (intra)nuclear or nucleon Hamiltonian (Dean 2007), the complete molecular Hamiltonian (Hehre, Radom et al. 1986; Szabo and Ostlund 1996; Kohn 1999; Pople 1999; Helgaker, Jørgensen et al. 2000; Springborg 2000) for N_n nuclei and N_e electrons can fortunately be written in an analytic closed form (thanks to the inverse square-distance proportionality in Coulomb's electrostatic force law):

$$\hat{H}_{mole} = \sum_{j}^{N_n} -\frac{1}{2M_j}\nabla_j^2 + \sum_{j<j'}^{N_n}\frac{Z_{j'}Z_j}{x_{jj'}} - \frac{1}{2}\sum_{i}^{N_e}\nabla_i^2 - \sum_{j}^{N_n}\sum_{i}^{N_e}\frac{Z_j}{r_{ij}} + \sum_{i<i'}^{N_e}\frac{1}{r_{ii'}}. \tag{2}$$

In Eq. (2), the units are atomic units, M_j is the mass ratio of nucleus j to electron, and Z_j is the atomic number of nucleus j. The Laplacian operators ∇_j^2 and ∇_i^2 denote the second order differentiation with respect to the coordinates of the jth nucleus and the ith electron, respectively. The first term in Eq. (2) represents the kinetic energy operator for nuclei; the second term is the Coulomb repulsion between nuclei; the third term is the operator for the kinetic energy of electrons; the fourth and fifth terms indicate the Coulomb attraction between electrons and nuclei, and the repulsion between electrons, respectively. The distance between the jth and the j'th nuclei is $x_{jj'}$; the separation between the ith and the i'th electrons is $r_{ii'}$; the distance between the jth nucleus and the ith electrons is r_{ij}.

2.2 Central quantity in quantum thermodynamics: Quantum partition function

Once the energy eigenvalues or the quantized energy spectrum in Eq. (1) are calculated, it is straightforward to obtain a central physical quantity in thermodynamics, i.e., the quantum canonical partition function Q_{qm} (McQuarrie 2000), by the following summation of the Boltzmann energy distribution:

$$Q_{qm} = \sum_n \exp(-\beta E_n), \tag{3}$$

where $\beta = 1/k_B T$, k_B is Boltzmann's constant, and T is temperature. All standard thermodynamic quantities for a system of nuclei and electrons, e.g., free energy, internal energy, entropy, pressure, etc., can be derived from it. In Eq. (3), the lowest energy level E_0, which is often called the ground state energy or zero-point energy (ZPE), is usually the dominant energy level contributing to the partition function. Further, by virtue of Heisenberg's uncertainty principle, the ZPE is always larger than the minimum value of potential energy because a particle can never be at rest anywhere in a given potential or a particle with a particular momentum can be everywhere in a given potential.

2.3 Origin of potential energy surface: Born-Oppenheimer approximation

Unfortunately, even though *all* physics and chemistry of a (time-independent) molecular system is essentially in the molecular Schrödinger equation [Eq. (1)], it can be exactly solved only for simplest one-electron atoms or ions. For other systems, approximations must be introduced to calculate numerical solutions with the aid of computers. The most common and perhaps the mildest approximation is the Born-Oppenheimer approximation (Born and Oppenheimer 1927; Hirschfelder and Meath 1967; Kolos 1970; Ballhausen and Hansen 1972; Hehre, Radom et al. 1986; Szabo and Ostlund 1996; Helgaker, Jørgensen et al. 2000; Springborg 2000; Mielke, Peterson et al. 2003). It decouples internuclear motions from electrons so that nuclei effectively move on a potential energy surface (PES) obtained by solving the electronic part of Schrödinger equation.

This approximation is based on the fact that an electron is much lighter than any nucleus (e.g., a proton, the lightest nucleus, is about 1840 times heavier than an electron). Nuclei move, consequently, much slowlier. As a result, from the electronic perspective, for a given set of nuclear positions, electrons adjust their positions 'instantly' before nuclei have a chance to move. On the other hand, from the standpoint of nuclei, electrons are moving so fast that their effects on nuclei are averaged out over the electronic wave functions. Mathematically, to simplify the molecular Hamiltonian, we first solve the electronic part of the Schrödinger equation for a particular set of nuclear configurations $\{x_j\}$. The electronic part of the complete molecular Hamiltonian [Eq. (2)] is called electronic Hamiltonian:

$$\hat{H}_{elec} = -\frac{1}{2}\sum_i^{N_e} \nabla_i^2 - \sum_j^{N_n}\sum_i^{N_e} \frac{Z_j}{r_{ij}} + \sum_{i<i'}^{N_e} \frac{1}{r_{ii'}}. \tag{4}$$

With this electronic Hamiltonian, we can obtain the electronic energy E_{elec} from the corresponding electronic Schrödinger equation:

$$\hat{H}_{elec}\psi_{elec} = E_{elec}\left(\left\{x_j\right\}\right)\psi_{elec}, \tag{5}$$

where ψ_{elec} is the electronic wave function. Note that the electronic energy $E_{elec}\left(\left\{x_j\right\}\right)$ depends parametrically on the nuclear positions $\left\{x_j\right\}$. With this electronic energy, the molecular Hamiltonian in Eq. (2) can be simplified as follows:

$$\begin{aligned}\hat{H}_{mole} &\approx \sum_{j}^{N_n} -\frac{1}{2M_j}\nabla_j^2 + \sum_{j<j'}^{N_n}\frac{Z_{j'}Z_j}{x_{jj'}} + \left\langle -\frac{1}{2}\sum_{i}^{N_e}\nabla_i^2 - \sum_{j}^{N_n}\sum_{i}^{N_e}\frac{Z_j}{r_{ij}} + \sum_{i<i'}^{N_e}\frac{1}{r_{ii'}}\right\rangle \\ &= \sum_{j}^{N_n} -\frac{1}{2M_j}\nabla_j^2 + \left[\sum_{j<j'}^{N_n}\frac{Z_{j'}Z_j}{x_{jj'}} + E_{elec}\left(\left\{x_j\right\}\right)\right] \\ &= \sum_{j}^{N_n} -\frac{1}{2M_j}\nabla_j^2 + V\left(\left\{x_j\right\}\right),\end{aligned} \tag{6}$$

where $\langle\cdots\rangle$ signifies the average over electronic wave functions or the expectation value. In Eq. (6), V is defined as the sum of the nuclear repulsion energy and electronic energy, which effectively turns out to be the internuclear potential energy function as a consequence of the Born-Oppenheimer approximation:

$$V\left(\left\{x_j\right\}\right) \equiv \sum_{j<j'}^{N_n}\frac{Z_{j'}Z_j}{x_{jj'}} + E_{elec}\left(\left\{x_j\right\}\right). \tag{7}$$

There are many systematic and rigorous theories in electronic structure calculations to derive the internuclear potential energy from first principles (i.e., besides the universal fundamental constants in physics, there is no other empirical parameter involved in the calculations), e.g., Hartree-Fock theory, configuration interaction method, Møller-Plesset perturbation theory, coupled cluster approach, and Kohn-Sham density functional theory. All these quantum mechanical (QM) approaches for electronic structure calculations are often known as *ab initio* methods (Hehre, Radom et al. 1986; Szabo and Ostlund 1996; Kohn 1999; Pople 1999; Helgaker, Jørgensen et al. 2000; Springborg 2000).

In contrast, a complete empirical method to determine an internuclear potential energy surface is to parameterize the potential energy as an analytic function without treating electronic degrees of freedom. This type of approach is referred to as molecular mechanical (MM) method and the empirical potential energy is called force-field energy. Comparing to *ab initio* approach, MM methods are computationally much less expensive and can be applied to describe equilibrium properties in macromolecular systems involving over tens of thousands of heavy atoms (Hagler, Huler et al. 1974; Brooks, Bruccoleri et al. 1983; Weiner, Kollman et al. 1984; Jorgensen and Tirado-Rives 1988; Mayo, Olafson et al. 1990). But for the process involving electronic redistributions (e.g., electronic transfer, chemical bond breaking or forming, etc.), MM force field is often unable to describe it. Later, a hybrid approach called combined QM/MM method has emerged to synthesize the efficiency of MM force field with the accuracy of QM calculations (Field, Bash et al. 1990; Gao and Truhlar 2002). For the rest of this chapter, discussions are limited to the Born-Oppenheimer approximation, which adiabatically decouples nuclear and electronic degrees of freedom.

2.4 Classical free-energy profile vs classical potential of mean force

In practice, quantum effects on internuclear motions are much smaller than those on the electronic part. In many applications, the internuclear quantum effects are insignificant and could even be neglected. Thus, the eigenenergy spectrum E_n in Eq. (1) would become continuous. Given an internuclear potential V, the quantum canonical partition function in Eq. (3) consequently reduces to the classical canonical partition function as:

$$Q_{cl} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{dx^{3N_n}dp^{3N_n}}{h^{3N_n}} \exp\left\{-\beta\left[\left(\sum_j^{3N_n}\frac{p_j^2}{2M_j}\right)+V\left(\{x_j\}\right)\right]\right\}, \tag{8}$$

where h is Planck's constant and p is the momenta associated with the nuclear coordinates x. Subsequently, the classical free energy G_{cl} of a molecular system can be expressed in terms of the classical partition function Q_{cl} as follows:

$$G_{cl} = -k_B T \ln Q_{cl} = -k_B T \ln \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{dx^{3N_n}dp^{3N_n}}{h^{3N_n}} \exp\left\{-\beta\left[\left(\sum_j^{3N_n}\frac{p_j^2}{2M_j}\right)+V\left(\{x_j\}\right)\right]\right\}. \tag{9}$$

Note that the partition function and the free energy defined above are 'state' functions, which is independent of any nuclear coordinate and momentum (as we integrate out the entire phase space). Given a particular $3N_n$ -degree-of-freedom molecular system described by a particular potential energy function V at particular temperature, the partition function and the free energy are *constants*.

On the other hand, of significant interest in simulating a many-body biochemical or physical event is to examine how the free energy of a molecular system *varies* during the event. Conventionally, we first predetermine a coordinate which should be able to describe the event of interest from the start to the end. Next, we generate a free energy profile, which is an energy function of that predetermined coordinate, to investigate how the profile changes during the event. In fact, such a kind of free-energy profile can also be termed as potential energy of ensemble-average or mean force (Kirkwood 1935). Reasons are given below.

The free energy profile of a molecular system as a function of a predetermined coordinate of interest z can be written as follows:

$$\begin{aligned} G_z(z) &= -k_B T \ln\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{dx^{3N_n}dp^{3N_n}}{h^{3N_n}} \lambda_z \delta(x_z - z) \exp\left\{-\beta\left[\left(\sum_j^{3N_n}\frac{p_j^2}{2M_j}\right)+V\left(\{x_j\}\right)\right]\right\}\right] \\ &= -k_B T \ln\left[\int_{-\infty}^{\infty} dx^{3N_n-1} \exp\left\{-\beta\left[V\left(\{x_j\}, x_z = z\right)\right]\right\}\right] + C, \end{aligned} \tag{10}$$

where δ is Dirac delta function, λ_z is the thermal de Broglie wavelength for the degree of freedom along z-direction, and C is a normalization factor dependent on the inverse of the thermal de Broglie wavelengths for all degrees of freedom (the wavelength is a function of the nuclear mass M_j, and temperature T). C should be a constant during the biochemical or physical event of our interest. The integrand in the final configurational integral of Eq. (10)

is basically the probability density of the molecular system as a function of z. In practice, it is rare to determine the value of C because what we often care about is the free-energy difference at various values of z.

Notably, by taking the negative derivative of $G_z(z)$, i.e., $-dG_z(z)/dz$, we obtain the average force over all ensembles or over all degrees of freedom, which is called the mean force (Kirkwood 1935), based on the ensemble average definition in Eq. (16):

$$\begin{aligned}
&\langle \mathbf{F}(z) \rangle \\
&= \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{dx^{3N_n} dp^{3N_n}}{h^{3N_n}} \lambda_z \delta(x_z - z)\left[-\bar{\nabla}_{x_z} V(\{x_j\})\right]\Big|_{x_z=z} \exp\left\{-\beta\left[\left(\sum_j^{3N_n} \frac{p_j^2}{2M_j}\right) + V(\{x_j\})\right]\right\}}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{dx^{3N_n} dp^{3N_n}}{h^{3N_n}} \lambda_z \delta(x_z - z) \exp\left\{-\beta\left[\left(\sum_j^{3N_n} \frac{p_j^2}{2M_j}\right) + V(\{x_j\})\right]\right\}} \\
&= \frac{\int_{-\infty}^{\infty} dx^{3N_n-1}\left[-\bar{\nabla}_z V(\{x_j\}, x_z = z)\right] \exp\left\{-\beta\left[V(\{x_j\}, x_z = z)\right]\right\}}{\int_{-\infty}^{\infty} dx^{3N_n-1} \exp\left\{-\beta\left[V(\{x_j\}, x_z = z)\right]\right\}} = \frac{-dG_z(z)}{dz}\hat{z}.
\end{aligned} \tag{11}$$

Thus $G_z(z)$, the free energy profile as a function of a predetermined coordinate, is also called the potential of mean force (PMF) (Kirkwood 1935).

However, please note that if the predetermined coordinate of interest is *not* a linear combination of rectilinear coordinates, or in other words, if it is a curvilinear coordinate, then PMF is oftentimes *not* exactly equal to free-energy profile. Not only the Jacobian-determinant contribution makes their difference (Ruiz-Montero, Frenkel et al. 1997; Hénin, Fiorin et al. 2010), but also in a forthcoming paper, we will show that actually change of domains with respect to the coordinate of interest can also contribute to the free-energy profile, i.e., the Leibnizian contribution (Flanders 1973).

In addition, we will also show that according to differential geometry and general relativity, once we realize the equivalence between orthogonal covariant and contravariant vectors (Arfken and Weber 2001), then the Jacobian scale factor for a predetermined curvilinear coordinate of interest, q_ξ, can be proved to be (in contravariant space):

$$h_{q_\xi} = \left|\bar{\nabla} q_\xi\right|^{-1} \tag{12}$$

and the unit vector for q_ξ can be proved as (in contravariant space):

$$\hat{q}_\xi = \bar{\nabla} q_\xi \Big/ \left|\bar{\nabla} q_\xi\right| \tag{13}$$

In Eq. (12) and (13), q_ξ must belong to at least one *complete* set of curvilinear coordinates, hypothetically. In general, unless we explicitly define the rest of the *complete* curvilinear coordinates, the sole definition of q_ξ is *not* sufficient to make the PMF be unique. But, we will show the free-energy profile does *not* suffer from this uniqueness problem. In fact, if we restrict ourselves to a *complete* set of curvilinear coordinates in which q_ξ is orthogonal to the

rest of coordinates, then the PMF will be unique and its relation with the free-energy profile can be proved as follows (den Otter 2000), after using Eq. (12) and Eq. (13):

$$\frac{dG_\xi(\xi_0)}{d\xi_0} = \left\langle \vec{\nabla} V \cdot \left(\frac{\vec{\nabla} q_\xi}{\left|\vec{\nabla} q_\xi\right|^2} \right) \right\rangle_{q_\xi = \xi_0} - k_B T \left\langle \vec{\nabla} \cdot \left(\frac{\vec{\nabla} q_\xi}{\left|\vec{\nabla} q_\xi\right|^2} \right) \right\rangle_{q_\xi = \xi_0} \tag{14}$$

In Eq. (14), the Leibnizian contribution is nil, the first term on RHS is the mean force for q_ξ, the second term is the Jacobian contribution, and $\langle \cdots \rangle_{q_\xi = \xi_0}$ is the ensemble average over all configurations with $q_\xi = \xi_0$.

Finally, the Fixman potential (Fixman 1974), which corrects the velocity-bias in constrained MD, will also be presented with correct dependence on mass in our forthcoming paper.

2.5 Simulating classical thermodynamics: Molecular dynamics simulations

By assuming the molecular system of our interest is ergodic, molecular dynamics (MD) simulation techniques can be employed to compute the ensemble average of a physical quantity. In essence, MD simulations is numerically solving, integrating or propagating the Newtonian equations of motion, one-time-step by one-time-step. Given an internuclear potential V (regardless of using QM, MM, or hybrid QM/MM to construct), the motion or trajectory of a nucleus j as a function of time t is governed by Newton's second law:

$$-\vec{\nabla}_j V\left(\left\{x_j(t)\right\}\right) + \left(\textbf{Extended Forces}\right) = M_j \frac{d^2 \vec{x}_j}{dt^2}. \tag{15}$$

Note the extended forces in Eq. (15) are *essential* for having *canonical* ensemble (constant temperature) instead of *microcanonical* ensemble (constant energy) in MD simulations (Hünenberger 2005). In the ergodic hypothesis (Lebowitz and Penrose 1973; Cogswell 1999) [the dynamical version of ergodic theory was first proposed by Birkhoff (Birkhoff 1931), in which Liouville's theorem was applied to ensure the ensemble distribution in phase-space is invariant with time], if the simulation time for propagating the trajectory $x_j(t)$ of the nucleus j is *infinitely long*, the ensemble average of a physical quantity $\mathbf{f}\left(\{x\},\{p\}\right)$ (which can be either a scalar or a vector) over the *entire* phase space, i.e.,

$$\langle \mathbf{f} \rangle = \frac{1}{Q_{cl}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{dx^{3N_n} dp^{3N_n}}{h^{3N_n}} \, \mathbf{f}\left(\{x\},\{p\}\right) \exp\left\{ -\beta \left[\left(\sum_j^{3N_n} \frac{p_j^2}{2M_j} \right) + V\left(\{x_j\}\right) \right] \right\}, \tag{16}$$

is equal to the time average in MD simulations:

$$\langle \mathbf{f} \rangle = \lim_{t_f \to \infty} \frac{1}{t_f} \int_0^{t_f} \mathbf{f}\left(\{x(t)\},\{p(t)\}\right) dt. \tag{17}$$

In other words, longer MD simulation time allows us to sample more phase space for computing the corresponding ensemble average, which in turn could be in higher accuracy.

3. Zwanzig's free-energy perturbation theory

Owing to the Boltzmann exponential energy distribution, one of the major difficulties in computing a converged free-energy profile or potential of mean force [Eq. (10)] via MD and MC sampling techniques is that it takes longer simulation time or runs more MC steps to have enough higher-energy samples. Yet, many interesting biochemical or physical molecular properties could be in higher-energy regions, e.g., the transition state during protein folding or biochemical reaction.

In practice, in order for having effective samplings on both the lower-energy (e.g., reactant state) and higher-energy regions (e.g., transition state), Zwanzig's free-energy perturbation (Zwanzig 1954) [which is also referred to as statistical-mechanical perturbation theory (McQuarrie 2000)] has been extensively applied. The feature of the perturbation is relating the change of free energy between *two* systems (both have the same number of degrees of freedom) by an ensemble average taken in only *one* of the two systems. This can be illustrated by first writing the classical free energy G_{cl} corresponding to the partition function in Eq. (8) as follows:

$$G_{cl} = -k_B T \ln Q_{cl} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{dx^{3N_n} dp^{3N_n}}{h^{3N_n}} \exp(-\beta E), \tag{18}$$

where E is the energy at a point $\left(\{p_j\},\{x_j\}\right)$ in the phase space, i.e.,

$$E = E\left(\{p_j\},\{x_j\}\right) = \left(\sum_j^{3N_n} \frac{p_j^2}{2M_j}\right) + V\left(\{x_j\}\right). \tag{19}$$

Next, we rewrite Eq. (18) as:

$$\begin{aligned} G_{cl} &= -k_B T \ln \left[\frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{dx^{3N_n} dp^{3N_n}}{h^{3N_n}} \exp\left[-\beta\left(E - E_0\right)\right] e^{-\beta E_0}}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{dx^{3N_n} dp^{3N_n}}{h^{3N_n}} e^{-\beta E_0}} \right] - k_B T \ln \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{dx^{3N_n} dp^{3N_n}}{h^{3N_n}} e^{-\beta E_0} \\ &= -k_B T \ln \left\langle \exp\left[-\beta\left(E - E_0\right)\right] \right\rangle_0 + G_0, \end{aligned} \tag{20}$$

where G_0 is the free energy of the reference system, E_0 is the energy at a point in the phase space of the reference system, and $\langle \cdots \rangle_0$ is an ensemble average for the reference system. From Eq. (20), we obtain Zwanzig's free-energy perturbation (Zwanzig 1954):

$$G - G_0 = -k_B T \ln \left\langle \exp\left[-\beta\left(E - E_0\right)\right] \right\rangle_0. \tag{21}$$

As a result, by taking the advantage of the perturbation [Eq. (21)], we can readily have enough samples in higher-energy regions in a reference frame where their original high potential energy values intentionally get lowered. Afterwards, the corrected free energy can straightforwardly be recovered by taking the average of the exponential factor $\exp\left[-\beta\left(E - E_0\right)\right]$ over the ensembles sampled in the reference system. This is exactly the idea behind many enhanced sampling methods, such as the umbrella sampling technique.

4. Systematic *ab initio* molecular dynamics approach: Free-energy expansion method as a series of covariance tensors

A fundamental key to have successful molecular simulations is the accuracy of internuclear potential for describing atomic motions during biochemical or physical events. By exploiting Zwanzig's free-energy perturbation (FEP) theory, we are developing a new rigorous method to systematically obtain accurate free-energy profiles, in which the internuclear potential energy is effectively computed at a high-level *ab initio* theory. Our new method is a systematic free-energy expansion (FEE) in terms of a series of covariance tensors. The new expansion will enable us to have a free-energy profile at a level as high as the coupled cluster theory at an affordable computational cost, which is currently known as the gold standard but unreachable level of theory for free-energy simulations. The focus of our FEE method will be on the difference of free energy calculated by two different internuclear potential. Furthermore, in contrast to Car-Parrinello MD (CPMD) which is limited to potential energy derived from DFT (Car and Parrinello 1985), our method is independent of how the potential energy functions being constructed. Therefore, by combining it with our novel automated integration-free path-integral (AIF-PI) method together (See Section 5; Wong and Gao 2007; Wong and Gao 2008; Wong, Richard et al. 2009; Wong, Gu et al. 2012), we will also be able to compute free-energy barriers, changes of binding energy, pK_a values, and isotope effects at an *ab initio* path integral level (see Section 6).

Let's begin with the FEP theory. From Eq. (21), the free energy difference between using lower-level (LL) and higher-level (HL) *ab initio* methods can be expressed as:

$$G_{HL} - G_{LL} = \Delta G = -k_B T \ln\left\langle \exp\left[-\beta\left(E_{HL} - E_{LL}\right)\right]\right\rangle_{LL}. \tag{22}$$

Next we expand the ensemble average in Eq. (22) and sum up the prefactors into a series of cumulants:

$$G_{HL} - G_{LL} = \Delta G = -\frac{1}{\beta}\ln\left\langle \exp\left[-\beta\ \Delta E\right]\right\rangle_{LL} = -\frac{1}{\beta}\ln\left\{\exp\left[\sum_{n=1}^{\infty}\left\langle\left(-\beta\ \Delta E\right)^n\right\rangle_{LL,c}\right]\right\}, \tag{23}$$

where

$$\Delta E = E_{HL} - E_{LL}, \tag{24}$$

$\langle\cdots\rangle_{LL,c}$ is a cumulant, and n is the order of a cumulant. In his original 1954 paper (Zwanzig 1954), Zwanzig showed that the cumulant expansion is fast converging when the change of energy ΔE in the ensemble is reasonably small relative to the inverse of β. However, in terms of computational cost, this cumulant expansion does not provide an advantage for correcting lower-level free energy. This is because the time required for calculating the cumulant average $\langle\cdots\rangle_{LL,c}$ with computer is basically as much as the time needed to directly compute the higher-level free energy G_{HL}, regardless of whether the perturbation ΔE is big.

In order to ease up this situation, in a forthcoming paper we will prove that each cumulant can be further expanded as a Taylor series expansion fluctuating about the ensemble average position $\bar{\mathbf{x}}_{LL}$ in the form:

$$\left\langle \left(-\beta\ \Delta E\right)^n \right\rangle_{LL} = \left\langle f_n(\mathbf{x}) \right\rangle_{LL} = f_n(\bar{\mathbf{x}}_{LL}) + \frac{1}{2!}\left[\hat{\mathbf{D}}^2 f_n(\bar{\mathbf{x}}_{LL})\right]\mathrm{cov}\left(\mathbf{x}^{T_p}, \mathbf{x}\right) + \cdots \tag{25}$$

where T_p is transpose, $f_n(\mathbf{x}) = \left(-\beta\ \Delta E\right)^n$, $\mathbf{x}$ is a position vector of $3N$ Cartesian coordinates of the system, $\hat{\mathbf{D}}^n$ is the nth-order tensor operator for differentiation with respect to the $3N$ coordinates (e.g., $\hat{\mathbf{D}} f_n(\bar{\mathbf{x}}_{LL})$ is the gradient and $\hat{\mathbf{D}}^2 f_n(\bar{\mathbf{x}}_{LL})$ is the Hessian matrix), and $\mathrm{cov}\left(\mathbf{x}^{T_p}, \mathbf{x}\right)$ is the covariance matrix. The higher order terms in Eq. (25) involve higher order covariance tensors. Note that the term associated with the gradient is not shown in Eq. (25) because the first order central moment, i.e., $\left\langle \mathbf{x} - \bar{\mathbf{x}}_{LL} \right\rangle_{LL}$, is always zero by definition.

By combining Eq. (25) with Eq. (23), we have enough equations to systematically approach the exact value of high-level free energy at a reduced computational cost. The number of calculations involving E_{HL} is now considerably decreased to only a single-point energy calculation at $\bar{\mathbf{x}}_{LL}$ for the zeroth order correction, and merely a normal-mode frequency analysis at $\bar{\mathbf{x}}_{LL}$ for the second order correction.

To increase the converging property for the expansion in Eq. (25) as well as to overcome the problem of multi-model probability distribution, we can further generalize the FEE method by considering a decomposition of the ensemble average into subgroups by clustering methods. The clustering scheme will be determined in a way such that the FEE expansion is converged up to the second order correction in each group or each cluster. Please note that, in the limit that the number of clusters becomes as many as the number of ensembles, the formalism reduces back to the original ensemble average, and inclusion of only the zeroth order term in Eq. (25) is able to return us back the exact result of Eq. (23).

Cumulant	Tensor	ΔG (kcal/mol)	Error
1st	0th	−170.601	0.316
	1st	−170.601	0.316
	2nd	−170.680	0.237
	3rd	−170.680	0.237
	∞	−170.680	0.237
2nd	0th	−170.601	0.316
	1st	−170.601	0.316
	2nd	−170.875	0.042
	3rd	−170.877	0.040
	∞	−170.883	0.034
6th	∞	−170.917	0.000

Table 1. Free-Energy correction ΔG for H_2O from HF/6-31G(d) to MP2/6-311G(d,p).

Since single-point energy calculations and a normal-mode frequency analysis at high-level electronic structure calculations are actually very common in literature (which are often used for minimized structures, though), we anticipate this new free-energy expansion method would be particularly useful for coupling accurate results from high-level *ab initio* theory with computational efficiency of lower-level samplings in free-energy calculations.

The preliminary results using this new systematic FEE method, i.e., Eq. (23), are very encouraging. Table 1 shows the free energy correction ΔG for a single water molecule from HF/6-31G(d) to MP2/6-311G(d,p). Even just up to the first cumulant at the zeroth order correction, the computed error is in the order of magnitude ~0.1 kcal/mol. The first cumulant is basically converged as soon as the second order correction is included.

5. Simulating quantum thermodynamics: Feynman's path integral

All the above discussions on simulating internuclear thermodynamics are limited to classical mechanics (regardless of using QM, MM, hybrid QM/MM to construct potential energy). However, the real world is described by quantum mechanics, including nuclei. In some important applications of Life and Materials Sciences, such as hydrogen adsorption in carbon nanotechnology, the transport mechanism of hydrated hydroxide ions in aqueous solution, and kinetic isotope effects on a proton-transfer reaction, actually internuclear quantum-statistical effects (e.g., quantization of vibration and quantum tunneling) are not negligible. A popular choice for incorporating such internuclear quantum-statistical effects in the conventional molecular dynamics (MD) or Monte Carlo (MC) simulations (Tanaka, Kanoh et al. 2005; Warshel, Olsson et al. 2006; Kowalczyk, Gauden et al. 2007; Gao, Major et al. 2008; Kowalczyk, Gauden et al. 2008; Major, Heroux et al. 2009; Wong, Gu et al. 2012) is using Feynman's path integral (Feynman 1948; Feynman 1966; Kleinert 2004; Brown 2005; Feynman, Hibbs et al. 2005).

The essence of Feynman's path integral is to transform the Schrödinger *differential* equation to become an *integral* equation. As a result, the many-body path integrations can be carried out by the conventional MD or MC sampling techniques. In addition, the quantum canonical partition function can be directly obtained with no need to compute individual energy eigenvalues.

5.1 Kleinert's variational perturbation theory for centroid density of path integrals

Kleinert's variational perturbation (KP) theory (Kleinert 2004) for the centroid density (Gillan 1987; Gillan 1987; Voth 1996; Ramírez, López-Ciudad et al. 1998; Ramírez and López-Ciudad 1999; Feynman, Hibbs et al. 2005) of Feynman path integrals (Feynman 1948; Feynman 1966; Kleinert 2004; Brown 2005; Feynman, Hibbs et al. 2005) provides a complete theoretical foundation for developing non-stochastic methods to systematically incorporate internuclear quantum-statistical effects in condensed phase systems. Similar to the complementary interplay between the rapidly growing quantum Monte Carlo simulations (Anderson 1975; Grossman and Mitas 2005; Lester and Salomon-Ferrer 2006; Wagner, Bajdich et al. 2009) and the well-established *ab initio* or density-functional theories (DFT) for electronic structure calculations (Hehre, Radom et al. 1986; Szabo and Ostlund 1996; Kohn 1999; Pople 1999; Helgaker, Jørgensen et al. 2000; Springborg 2000), non-stochastic path-integral methods can complement the conventional Fourier or discretized path-integral Monte-Carlo (PIMC) (MacKeown 1985; Coalson 1986; Ceperley 1995; Mielke and Truhlar 2001; Sauer 2001) and molecular dynamics (PIMD) (Cao and Voth 1994; Voth 1996) simulations which have been widely used in condensed phases.

To simplify the illustration of the essence of Kleinert's variational perturbation theory, we now consider a one-particle one-dimensional system. For a one-particle one-dimensional system, the classical canonical partition function in Eq. (8) reduces to become:

$$Q_{cl} = \sqrt{\frac{Mk_BT}{2\pi\hbar^2}} \int_{-\infty}^{\infty} e^{-\beta V(x_0)} dx_0. \tag{26}$$

The traditional way to obtain the quantum canonical partition function, i.e., Eq. (3), is to solve the internuclear Schrödinger equation to get the individual energy eigenvalues. But in the path-integral (PI) formulation, we do not know the individual energy eigenvalues for obtaining the quantum partition function. This is because the PI representation of the quantum partition function can be written in terms of the centroid effective potential W as a classical configuration integral (Gillan 1987; Gillan 1987; Voth 1996; Ramírez, López-Ciudad et al. 1998; Ramírez and López-Ciudad 1999; Kleinert 2004; Feynman, Hibbs et al. 2005):

$$Q_{qm} = \sum_n \exp(-\beta E_n) = \sqrt{\frac{Mk_BT}{2\pi\hbar^2}} \int_{-\infty}^{\infty} e^{-\beta W(x_0)} dx_0. \tag{27}$$

Given the centroid potential $W(x_0)$, thermodynamic and quantum dynamic quantities can be accurately determined, including molecular spectroscopy of quantum fluids and the rate constant of chemical and enzymatic reactions. The mass-dependent nature of $W(x_0)$ is also of particular interest because isotope effects can be obtained, and it has been applied to carbon nanotubes (Tanaka, Kanoh et al. 2005; Kowalczyk, Gauden et al. 2007; Kowalczyk, Gauden et al. 2008), and biochemical reactions in protein (Warshel, Olsson et al. 2006; Gao, Major et al. 2008; Major, Heroux et al. 2009) and RNA enzymes (Wong, Gu et al. 2012).

The centroid potential $W(x_0)$ in Eq. (27) is defined as follows (Gillan 1987; Gillan 1987; Voth 1996; Ramírez, López-Ciudad et al. 1998; Ramírez and López-Ciudad 1999; Kleinert 2004; Feynman, Hibbs et al. 2005):

$$W(x_0) = -k_BT \ln\left[\sqrt{\frac{2\pi\hbar^2}{Mk_BT}} \oint \mathrm{D}[x(\tau)]\,\delta(\bar{x} - x_0) \exp\{-\mathcal{A}[x(\tau)]/\hbar\}\right], \tag{28}$$

where τ is a real number and represents the component for pure imaginary time in path integral, $x(\tau)$ describes a path in space-time, $\oint \mathrm{D}x(\tau)\delta(\bar{x} - x_0)$ denotes a summation over *all* possible closed paths in which $\bar{x}$ is equal to x_0 (i.e., a functional integration), and $\bar{x}$ is the time-average position, called 'centroid'

$$\bar{x} \equiv \frac{1}{\beta\hbar} \int_0^{\beta\hbar} x(\tau) d\tau \,. \tag{29}$$

In Eq. (28), $\mathcal{A}$ is the quantum-statistical action:

$$\mathcal{A}[x(\tau)] = \int_0^{\beta\hbar} d\tau \left\{\frac{M}{2}\dot{x}(\tau)^2 + V[x(\tau)]\right\}, \tag{30}$$

where $V(x)$ is the original potential energy of the system. Generalization of Eq. (28) to a multi-dimensional system is straightforward (Kleinert 2004; Feynman, Hibbs et al. 2005).

A number of non-stochastic approaches have been developed to approximately estimate the centroid potential. For example, Feynman and Hibbs described a first-order cumulant expansion by introducing a Gaussian smearing function in a free-particle reference frame to yield an upper bound on the centroid potential (Feynman, Hibbs et al. 2005). This was subsequently modified by Doll and Myers (DM) by using a Gaussian width associated with the angular frequency at the minimum of the original potential (Doll and Myers 1979). Mielke and Truhlar employed a free-particle reference state and approximated the sum over paths by a minimal set of paths constrained for a harmonic oscillator. The action integral is obtained by using the three-point trapezoidal rule for the potential to yield the displaced-point path integral (DPPI) centroid potential (Mielke and Truhlar 2001).

A closely related theoretical approach to the KP theory is the variational method independently introduced by Giachetti and Tognetti (Giachetti and Tognetti 1985), and by Feynman and Kleinert (hereafter labeled as GTFK) (Feynman and Kleinert 1986), which formally corresponds to the first order approximation in the KP theory, i.e., KP1. The GTFK approach is a variational method that adopts a harmonic reference state by variationally optimizing the angular frequency. This variational method has been applied to a variety of systems, including quantum dynamic processes in condensed phases (e.g., water and helium). Although the original GTFK approach is among the most accurate approximate methods for estimating the path-integral centroid potential in many applications (Mielke and Truhlar 2001), significant errors can exist in situations in which quantum effects are dominant, especially at low temperatures. Higher order perturbations of KP theory can significantly and systematically improve computational accuracy over the KP1 results.(Kleinert 2004; Wong and Gao 2007; Wong and Gao 2008; Wong, Richard et al. 2009; Wong, Gu et al. 2012)

In essence, what Kleinert's variational perturbation (KP) theory does is to systematically builds up anharmonic corrections to the harmonic centroid potential calculated in a harmonic reference state characterized by a trial angular frequency Ω (Kleinert 2004). Given the reference, or trial harmonic action:

$$A_{\Omega}^{x_0} = \int_0^{\beta\hbar} d\tau \left\{ \frac{M}{2}\dot{x}(\tau)^2 + \frac{1}{2}M\Omega^2\left[x(\tau) - x_0\right]^2 \right\}. \tag{31}$$

the centroid potential $W(x_0)$ in Eq. (28) can be expressed as a path integral of the harmonic action which is perturbed by the anharmonicity of the original potential:

$$e^{-\beta W(x_0)} = \sqrt{\frac{2\pi\hbar^2}{Mk_BT}} \oint \mathrm{D}x(\tau)\,\delta(\bar{x} - x_0)\, e^{-A_{\Omega}^{x_0}/\hbar} e^{-\left(A - A_{\Omega}^{x_0}\right)/\hbar} = Q_{\Omega}^{x_0} \left\langle e^{-\left(A - A_{\Omega}^{x_0}\right)/\hbar} \right\rangle_{\Omega}^{x_0}, \tag{32}$$

where $Q_{\Omega}^{x_0}$ is the local harmonic partition function given as follows:

$$Q_{\Omega}^{x_0} = \sqrt{\frac{2\pi\hbar^2}{Mk_BT}} \oint \mathrm{D}x(\tau)\,\delta(\bar{x} - x_0)\, e^{-A_{\Omega}^{x_0}/\hbar} = \frac{\beta\hbar\Omega/2}{\sinh(\beta\hbar\Omega/2)}, \tag{33}$$

and $\langle\cdots\rangle_{\Omega}^{x_0}$ is the expectation value over all closed paths of the action in Eq. (31):

$$\left\langle e^{-F\left[x(\tau)\right]/\hbar}\right\rangle_{\Omega}^{x_0}=\frac{1}{Q_{\Omega}^{x_0}}\sqrt{\frac{2\pi\hbar^2}{Mk_BT}}\oint \mathrm{D}x(\tau)\,\delta\left(\overline{x}-x_0\right)e^{-F\left[x(\tau)\right]/\hbar}e^{-A_{\Omega}^{x_0}/\hbar}, \tag{34}$$

In Eq. (34), $F\left[x(\tau)\right]$ denotes an arbitrary functional. It is of interest to note that Eq. (32) is similar to the starting point of Zwanzig's free-energy perturbation (Section 3), which has been extensively used in free-energy calculations through Monte Carlo and molecular dynamics simulations. Their difference is one is for ordinary ensemble average, while another one is for closed-path average, i.e., *functional* average.

If we expand the exponential functional in Eq. (32) and sum up the prefactors into an exponential series of cumulants, then the *n*th-order approximation, $W_n^{\Omega}(x_0)$, to the centroid potential $W(x_0)$ can be written as follows (Kleinert 2004):

$$\begin{aligned}e^{-\beta W_n^{\Omega}(x_0)}=Q_{\Omega}^{x_0}\exp\Bigg\{&-\frac{1}{\hbar}\int_0^{\beta\hbar}d\tau\left\langle A_{\mathrm{int}}^{x_0}\right\rangle_{\Omega,c}^{x_0}+\frac{1}{2!\hbar^2}\int_0^{\beta\hbar}d\tau_1\int_0^{\beta\hbar}d\tau_2\left\langle A_{\mathrm{int}}^{x_0}\left[x(\tau_1)\right]A_{\mathrm{int}}^{x_0}\left[x(\tau_2)\right]\right\rangle_{\Omega,c}^{x_0}\\&+\cdots+\left\{\prod_{j=1}^{n}\int_0^{\beta\hbar}d\tau_j\right\}\frac{(-1)^n}{n!\hbar^n}\left\langle\prod_{k=1}^{n}A_{\mathrm{int}}^{x_0}\left[x(\tau_k)\right]\right\rangle_{\Omega,c}^{x_0}\Bigg\},\end{aligned} \tag{35}$$

where $A_{\mathrm{int}}^{x_0}=A-A_{\Omega}^{x_0}$ is the so-called inter-*action*, representing the perturbation to the harmonic reference state, $\left\langle\cdots\right\rangle_{\Omega,c}^{x_0}$ is a cumulant which can be written in terms of expectation values $\left\langle\cdots\right\rangle_{\Omega}^{x_0}$ by the cumulant expansion (Zwanzig 1954; Kubo 1962; Kleinert 2004), e.g.,

$$\left\langle A_{\mathrm{int}}^{x_0}\left[x(\tau)\right]\right\rangle_{\Omega,c}^{x_0}\equiv\left\langle A_{\mathrm{int}}^{x_0}\left[x(\tau)\right]\right\rangle_{\Omega}^{x_0}, \tag{36}$$

$$\left\langle A_{\mathrm{int}}^{x_0}\left[x(\tau_1)\right]A_{\mathrm{int}}^{x_0}\left[x(\tau_2)\right]\right\rangle_{\Omega,c}^{x_0}\equiv\left\langle A_{\mathrm{int}}^{x_0}\left[x(\tau_1)\right]A_{\mathrm{int}}^{x_0}\left[x(\tau_2)\right]\right\rangle_{\Omega}^{x_0}-\left\{\left\langle A_{\mathrm{int}}^{x_0}\left[x(\tau)\right]\right\rangle_{\Omega}^{x_0}\right\}^2, \tag{37}$$

$$\begin{aligned}\left\langle A_{\mathrm{int}}^{x_0}\left[x(\tau_1)\right]A_{\mathrm{int}}^{x_0}\left[x(\tau_2)\right]A_{\mathrm{int}}^{x_0}\left[x(\tau_3)\right]\right\rangle_{\Omega,c}^{x_0}\equiv&\left\langle A_{\mathrm{int}}^{x_0}\left[x(\tau_1)\right]A_{\mathrm{int}}^{x_0}\left[x(\tau_2)\right]A_{\mathrm{int}}^{x_0}\left[x(\tau_3)\right]\right\rangle_{\Omega}^{x_0}\\&-3\left\langle A_{\mathrm{int}}^{x_0}\left[x(\tau_1)\right]A_{\mathrm{int}}^{x_0}\left[x(\tau_2)\right]\right\rangle_{\Omega}^{x_0}\left\langle A_{\mathrm{int}}^{x_0}\left[x(\tau)\right]\right\rangle_{\Omega}^{x_0}+2\left\{\left\langle A_{\mathrm{int}}^{x_0}\left[x(\tau)\right]\right\rangle_{\Omega}^{x_0}\right\}^3\end{aligned} \tag{38}$$

More importantly, Kleinert and co-workers derived a math equation for expressing the expectation value $\left\{\prod_{j=1}^{n}\int_0^{\beta\hbar}d\tau_j\right\}\left\langle\prod_{k=1}^{n}F_k\left[x(\tau_k)\right]\right\rangle_{\Omega}^{x_0}$ from the *functional-integral* form to be in terms of Gaussian smearing convolution integrals (*ordinary* integrals) (Kleinert 2004):

$$\begin{aligned}\left\{\prod_{j=1}^{n}\int_0^{\beta\hbar}d\tau_j\right\}\left\langle\prod_{k=1}^{n}F_k\left[x(\tau_k)\right]\right\rangle_{\Omega}^{x_0}=&\left\{\prod_{j=1}^{n}\int_0^{\beta\hbar}d\tau_j\right\}\left\{\prod_{k=1}^{n}\int_{-\infty}^{\infty}dx_kF_k(x_k)\right\}\\&\times\frac{1}{\sqrt{(2\pi)^n\,\mathbf{Det}\left[a_{\tau_k\tau_{k'}}^2(\Omega)\right]}}\exp\left\{-\frac{1}{2}\sum_{\substack{k=1\\k'=1}}^{n}\left(x_k-x_0\right)a_{\tau_k\tau_{k'}}^{-2}(\Omega)\left(x_{k'}-x_0\right)\right\},\end{aligned} \tag{39}$$

where $\mathbf{Det}\left[a^2_{\tau_k\tau_{k'}}(\Omega)\right]$ is the determinant of the $n\times n$**-matrix** consisting of the Gaussian width $a^2_{\tau_k\tau_{k'}}(\Omega)$, $a^{-2}_{\tau_k\tau_{k'}}(\Omega)$ is an element of the inverse matrix of $a^2_{\tau_k\tau_{k'}}(\Omega)$, and the Gaussian width is a function of the trial frequency Ω:

$$a^2_{\tau\tau'}(\Omega)=\frac{1}{\beta M\Omega^2}\left\{\frac{\beta\hbar\Omega}{2}\frac{\cosh\left[\left(|\tau-\tau'|-\beta\hbar/2\right)\Omega\right]}{\sinh\left(\beta\hbar\Omega/2\right)}-1\right\}. \tag{40}$$

After using these smearing potentials given in Eq. (39), the *n*th-order Kleinert variational perturbation (KP*n*) approximation, $W_n^{\Omega}(x_0)$, shown in Eq. (35) as *functional* integrals, can now be written in terms of *ordinary* integrals as follows (Kleinert 2004):

$$\begin{aligned}&W_n^{\Omega}(x_0)\\&=-k_BT\ln Q_{\Omega}^{x_0}+\frac{k_BT}{\hbar}\int_0^{\beta\hbar}d\tau\left\langle V_{\rm int}^{x_0}\left[x(\tau_1)\right]\right\rangle_{\Omega}^{x_0}-\frac{k_BT}{2!\hbar^2}\int_0^{\beta\hbar}d\tau_1\int_0^{\beta\hbar}d\tau_2\left\langle V_{\rm int}^{x_0}\left[x(\tau_1)\right]V_{\rm int}^{x_0}\left[x(\tau_2)\right]\right\rangle_{\Omega,c}^{x_0}\\&\quad+\cdots+k_BT\frac{(-1)^{n+1}}{n!\hbar^n}\left\{\prod_{j=1}^{n}\int_0^{\beta\hbar}d\tau_j\right\}\left\langle\prod_{k=1}^{n}V_{\rm int}^{x_0}\left[x(\tau_k)\right]\right\rangle_{\Omega,c}^{x_0},\end{aligned} \tag{41}$$

where $V_{\rm int}^{x_0}\left[x(\tau)\right]=V\left[x(\tau)\right]-\frac{1}{2}M\Omega^2\left[x(\tau)-x_0\right]^2$ (the kinetic energy terms in Eq. (30) and Eq. (31) cancel each other out).

As *n* tends to infinity, $W_n^{\Omega}(x_0)$ approaches the exact value of the centroid potential $W(x_0)$ in Eq. (28), which is independent of the trial Ω. But the truncated sum in Eq. (41) does depend on Ω, and the optimal choice of this trial frequency at a given order of KP expansion and at a particular centroid position x_0 is determined by the least-dependence of $W_n^{x_0}(\Omega)$ on Ω itself. This is the so-called frequency of least dependence, which provides a variational approach to determine the optimal value of Ω, $\Omega_{\mathbf{opt},n}(x_0)$ (Kleinert 2004).

Of particular interest is the special case when $n=1$, which turns out to be identical to the original GTFK variational approach. An important property of KP1 or the GTFK variational approach is that there is a definite upper bound for the computed $W_1^{\Omega}(x_0)$ by virtue of the Jensen-Peierls inequality, i.e., from Eq. (32) and (35):

$$e^{-\beta W(x_0)}=Q_{\Omega}^{x_0}\left\langle\exp\left(-\frac{A-A_{\Omega}^{x_0}}{\hbar}\right)\right\rangle_{\Omega}^{x_0}\geq Q_{\Omega}^{x_0}\exp\left\langle-\frac{A-A_{\Omega}^{x_0}}{\hbar}\right\rangle_{\Omega}^{x_0}=e^{-\beta W_1^{\Omega}(x_0)}. \tag{42}$$

Note that by choosing $\Omega=0$ (i.e., the reference state is for a free particle), KP1 or GTFK (Giachetti and Tognetti 1985; Feynman and Kleinert 1986) reduces to the Feynman-Hibbs approach (Feynman, Hibbs et al. 2005). For higher orders of *n*, unfortunately, it is not guaranteed that a minimum of $W_n^{x_0}(\Omega)$ actually exists as a function of Ω. In this case, the least dependent Ω is obtained from the condition that the next derivative of $W_n^{x_0}(\Omega)$ with respect to Ω is set to zero. Consequently, Ω is considered as a variational parameter in the Kleinert perturbation theory such that $W_n^{x_0}\left[\Omega_{\mathbf{opt},n}(x_0)\right]$ is least-dependent on Ω.

This variational criterion relies on the uniformly and exponentially convergent property demonstrated from the KP theory. Kleinert and coworkers proved that his theory exhibits this property in several strong anharmonic-coupling systems. More importantly, this remarkably fast convergent property can also be observed even for computing the *electronic* ground state energy of a hydrogen atom (3 degrees of freedom). The ground state energy was determined by calculating the electronic centroid potential at the zero-temperature limit. The accuracies of the first three orders of the KP theory for a hydrogen atom are 85%, 95%, and 98%, respectively (Kleinert 2004).

In practice, for odd n, there is typically a minimum point in Ω, but due to the alternating sign of the cumulants in Eq. (41), there is usually *no* minimum in Ω for even n. Nevertheless, the frequency of least-dependence for an even order perturbation in n can be determined by locating the inflexion point, i.e., the zero-value of the second derivative of $W_n^{x_0}(\Omega)$ with respect to Ω. Since the KP expansion is uniformly and exponentially converged, Kleinert has demonstrated that the least-dependent plateau in $W_n^{x_0}(\Omega)$, which is characterized by a minimum point for odd n or by an inflexion point for even n, grows larger and larger with increasing orders of n (Kleinert 2004).

5.2 Automated integration-free path-integral method

An especially attractive feature of Eq. (41) is that the if the real system potential is expressed as a series of polynomials or Gaussians, then analytic expressions of Eq. (41) can be obtained, making the computation extremely efficient because the time-demanding Monte Carlo samplings for multi-dimensional numerical integrations could be avoided. Hereafter, the level of calculations up to nth order KP expansion for an mth-order-polynomial potential is denoted as KPn/Pm. For other potentials, KPn theory still involves elaborate n-dimensional space-time ($2n$ degrees of freedom) smearing integrals in Eq. (39). The intricacy of the smearing integrals increases tremendously for multidimensional potentials, where Ω becomes a $3N \times 3N$ matrix Ω_{ij} for N nuclei. This complexity is a major factor limiting applications of the KP theory beyond KP1, the original FK approach.

To render the KP theory feasible for many-body systems with N particles, we decouple the instantaneous normal mode (INM) coordinates $\left\{q^{x_0}\right\}^{3N}$ for a given configuration $\{x_0\}^{3N}$ (Wong and Gao 2007; Wong 2008; Wong and Gao 2008; Wong, Richard et al. 2009; Wong, Gu et al. 2012). Hence the multidimensional V effectively reduces to $3N$ one-dimensional potentials along each normal mode coordinate. Note that INM are naturally decoupled through the second order Taylor expansion of V. The approximation of decoupling the INM coordinates has also been used elsewhere (Stratt 1995; Deng, Ladanyi et al. 2002). This approximation is particularly suited for the KP theory because of the exponential decaying property of the Gaussian convolution integrals in Eq. (39). In the decoupling INM approximation, the total effective centroid potential for N nuclei can be simplified as:

$$W_n^{\Omega}\left(\{x_0\}^{3N}\right) \approx V\left(\{x_0\}^{3N}\right) + \sum_{i=1}^{3N} w_{i,n}^{\Omega}\left(q_i^{x_0}\right), \tag{43}$$

where $w_{i,n}^{\Omega}\left(q_i^{x_0}\right)$ is the centroid potential for normal mode *i*. Although the INM approximation sacrifices some accuracy, in exchange, it allows analyses of quantum mechanical vibration and tunneling, and their separate contributions to the *W*. Positive and negative values of w_i raise (vibration) and lower (tunneling) the original potential, respectively. In practice, real frequencies from the INM analysis often yields positive w_i's in Eq. (43) with dominant contributions from zero-point-energy effects. For imaginary frequencies in the INM, the values of w_i are often negative, due to tunneling contributions.

To obtain analytical expressions for the expectation values in Eq. (41), we use an *m*th order polynomial (P*m*) to approximate or interpolate the potential along q_i. Hereafter, an *m*th order polynomial representation of the original potential energy function obtained with an interpolating step size *q* Å both in the forward and backward directions along the normal mode coordinate at x_0 is denoted as P*m*-*q*A. Note that analytical results for P4 have been used by Kleinert for a quadratic-quartic anharmonic potential and a double-well potential (Kleinert 2004); however, higher order polynomials are needed to achieve the desired accuracy in real systems. We have thus derived the analytical closed forms of Eq. (41) up to P20 (Wong and Gao 2007; Wong 2008; Wong and Gao 2008; Wong, Richard et al. 2009; Wong, Gu et al. 2012). Consequently, the *W* as a function of an arbitrary Ω can be promptly obtained. This provides a convenient way to determine the least dependent Ω value without computing the complicated smearing integrals [Eq. (39)] iteratively for different trial values of Ω by Monte Carlo multi-dimensional numerical integrations. In fact, after the interpolating potential along each instantaneous normal-mode coordinate is determined, there is little computational cost for obtaining the *W*. Thereby, high level *ab initio* or density-functional (DFT) methods can be used to evaluate the potential energy function for *ab initio* path-integral calculations (Wong, Richard et al. 2009; Wong, Gu et al. 2012).

The computational procedure for obtaining the first and second order KP approximations to the centroid potential using our automated integration-free path-integral (AIF-PI) method is summarized below (Wong and Gao 2007; Wong 2008; Wong and Gao 2008; Wong, Richard et al. 2009; Wong, Gu et al. 2012):

1. For each $\{x_0\}^{3N}$, the mass-scaled Hessian matrix is diagonalized to obtain $\left\{q^{x_0}\right\}^{3N}$.
2. The original potential *V* is scanned from the configuration $\{x_0\}^{3N}$ along each $q_i^{x_0}$ for 10 points both in the forward and backward directions to interpolate *V* as P20-0.1A. A step size of 0.1 Å should be a reasonable choice to yield *W* in a few per cent of the exact.
3. After the P20-0.1A interpolations, each $w_{i,n}^{\Omega}(q_i^{x_0})$ as a function of Ω is readily obtained using the analytical expressions of KP1/P20 or KP2/P20. Note that the path integrals for these polynomials have been analytically integrated.
4. The values of $w_{i,n}^{\Omega}(q_i^{x_0})$ are determined by numerically locating the least dependence of $w_{i,n}^{\Omega}(q_i^{x_0})$ on Ω, i.e., zeroing the lowest order derivative of $w_{i,n}^{\Omega}(q_i^{x_0})$ w.r.t. Ω (first derivative for KP1 and usually second derivative for KP2).

The procedure presented above is integration-free and essentially automated (Wong and Gao 2007; Wong 2008; Wong and Gao 2008; Wong, Richard et al. 2009; Wong, Gu et al. 2012). We hope it could be used by non-path-integral experts or experimentalists as a "black-box" for any given system. We are currently developing a formalism to systematically couple instantaneous normal-mode coordinates.

Due to the integration-free feature, our AIF-PI method is computationally efficient such that the potential energy can be evaluated using *ab initio* or density-functional theory (DFT) for performing the so-called *ab initio* path-integral calculations. Consequently, we used DFT to construct the internuclear potential energy function for computing kinetic isotope effects (KIE) on several series of proton transfer reactions in water with the AIF-PI method. These reactions are relevant to biosynthesis of cholesterol. The computed KIE results at the KP2 level are in good agreement with experiment (Wong, Richard et al. 2009). Recently, we also employed the same computational technique to perform *ab initio* path-integral calculations of KIE on some RNA model reactions. Again, as shown in Table 2, the calculated values are in good agreement with experiments (Wong, Gu et al. 2012).

Reaction	KP2		Expt	
	$^{18}k_{\mathrm{Nu}}$	$^{18,34}k_{\mathrm{Lg}}$	$^{18}k_{\mathrm{Nu}}$	$^{18,34}k_{\mathrm{Lg}}$
Native	0.968	1.059	0.981(3)	1.034(4)
S3′	1.043	1.008	1.119(6)	1.0118(3)
S5′	1.042	1.002	1.025(5)	1.0009(1)

Table 2. Calculated primary kinetic isotope effects (KIEs) on 2′ nucleophile ($^{18}k_{\mathrm{Nu}}$) and 5′ leaving ($^{18}k_{\mathrm{Lg}}$ or $^{34}k_{\mathrm{Lg}}$) oxygens for RNA-model reactions using our AIF-PI method based on second order of Kleinert's variational perturbation theory (KP2), along with the most relevant available experimental (Expt) results for comparison. Experimental errors in the last decimal place are given in parenthesis.

Another compelling feature of the AIF-PI method is that it does not suffer the convergence difficulties of PIMC or PIMD simulations at the zero-temperature limit, i.e., absolute zero temperature. At the zero-temperature limit (T = 0 K), in principle, minimizing the centroid effective potential with respect to the nuclear positions can give us two important physical quantities: the exact value of the eigenenergy for zero-point motion (i.e., the zero-point energy ZPE or the ground state energy) and the exact expectation values of the nuclear positions at the ground state (Ramírez, López-Ciudad et al. 1998; Ramírez and López-Ciudad 1999), i.e.,

$$\lim_{T\to 0} W_{\min}\left(x_{\min}\right) = E_0, \tag{44}$$

and

$$x_{\min} = \left\langle \psi_0 \middle| x \middle| \psi_0 \right\rangle, \tag{45}$$

where x is the position operator, and $x_{\min}$ and $W_{\min}\left(x_{\min}\right)$ are, respectively, the coordinate and value at the (global) minimum of the centroid potential. In Eq. (44) and (45), ψ_0 is the nuclear ground state wave function and E_0 is the lowest eigenvalue of the Hamiltonian, i.e., the zero-point energy. In a forthcoming paper, we will have a rigorous proof showing that in fact at absolute zero temperature, there is *only one* stationary and minimum point in centroid potential, which is true even for any many-body systems. Hence, our recently derived analytical zero-temperature-limit results provide a convenient way to compute these two important physical quantities without solving the Schrödinger

equation (Wong 2008; Wong and Gao 2008), e.g., see Table 3. Together with the accurate low-lying excitation energies (Ramírez and López-Ciudad 2001) which could be obtained by the frequency analysis of the Hessian matrix at the sole minimum point at absolute zero temperature (including tunneling splitting), potentially one day our AIF-PI method could replace MC or MD *simulations* to have highly *reproducible* and *precise* free-energy *calculations* for many-body systems.

Molecule	Quantum	Harmonic	KP1	KP2
HCl	4.231	4.274	4.253	4.234
HF	5.732	5.793	5.762	5.736
H_2	6.193	6.284	6.238	6.202

Table 3. Ground state energy values (kcal/mol) for hydrogen chloride, hydrogen fluoride, and hydrogen molecules from the Morse potential using the harmonic-oscillator approximation, and our AIF-PI method based on first and second orders of the Kleinert's variational perturbation theory (KP1 and KP2).

Born-Oppenheimer Approximation	
Electronic Schrödinger equation *Ab initio* molecular orbital theory	**Internuclear** Schrödinger equation Systematic internuclear thermodynamics theory
Most molecular properties of interest are at **low lying electronic energy states**	All thermodynamic properties virtually can be derived from **quantum partition functions**
Hartree-Fock (HF) theory	**Kleinert's variational perturbation theory** for centroid effective potential
Independent electron (single-electron) approximation	**Decoupled instantaneous normal coordinate** approximation (DINCA)
Roothaan and Hall expressed the Fock operator in terms of **basis functions** for solving HF equations in matrix algebra self-consistently (SCF)	We propose interpolating potential energy functions to *m*th order **polynomials** in which analytic results of path-integration can be derived
Explain chemical properties in terms of **frontier occupied** and **unoccupied** molecular orbitals	Quantum effects from **vibration** and **tunneling** are separated and quantified in one mathematical framework
Post Hartree-Fock method to include **correlation energy by systematically couple single-electron orbitals**	Work out a formalism to **systematically couple instantaneous normal coordinates**

Table 4. Comparison (1) between Kleinert's variational perturbation (KP) theory and Hartree-Fock (HF) theory, (2) between our decoupled instantaneous normal coordinate approximation and independent electron approximation, and (3) between our integration-free path-integral results for polynomials in the KP theory and Roothann-Hall basis function approach for HF theory.

Finally, we make a quite interesting table (Table 4) to compare the traditional *ab initio* molecular orbital theory for electronic structure calculations with our systematic approach for computing internuclear quantum effects. In short, the rigor and the spirit of both types of methods is the same. We first breakdown or dissect a complicated many-body problem into many one-body problems. Then we identify which one bodies are more important. Next we couple back those important one bodies to systematically approach the exact.

6. Systematic *ab initio* path-integral free-energy expansion approach

In order to systematically refine a classical free-energy profile to become ultimate quantum free-energy profile, in which both electrons and nuclei are treated quantum mechanically and adiabatically, we are developing a systematic *ab initio* path-integral free-energy expansion (SAI-PI-FEE; $\psi\pi\phi$) approach. In this $\psi\pi\phi$ approach, we combine our novel free-energy expansion (FEE) method (Section 4) with our automated integration-free path-integral (AIF-PI) method (Section 5.2) such that we can perform *ab initio* path-integral simulations for realistic molecular systems. The key of this combination is that first we realize the quantum partition function can be computed as a classical configuration shown in Eq. (27), then now in Eq. (23), we treat the ΔE as:

$$\Delta E = W - V, \tag{46}$$

where V is the original internuclear potential and W is the centroid potential. So once we get the accurate value of W using our AIF-PI method, we can go ahead using our FEE method to systematically upgrade the level of our classical free-energy profile to an *ab initio* path-integral level, in which zero-point energy and tunnelling effects in nuclei, and isotope effects could all be incorporated.

In order to rigorously validate our $\psi\pi\phi$ method in a more effective way, the free-energy perturbation (FEP) in the Hamiltonian space will be performed, using the recently derived "universal" probability density function (UPDF), which is defined as follows:

$$P(\Delta E) = K \exp\left\{a\left[b(\Delta E - s) - e^{b(\Delta E - s)}\right]\right\}. \tag{47}$$

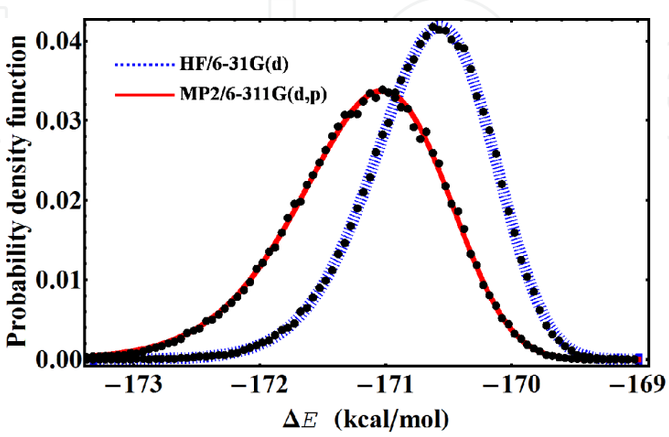


Fig. 1. Free energy perturbation for a water molecule in the Hamiltonian space using the universal probability density function (UPDF).

This UPDF can be used to determine the change of free energy ΔG in Eq. (23), by simply locating the intersection point of two probability density functions (Nanda, Lu et al. 2005; Chipot and Pohorille 2007). In Eq. (47), ΔE is a variable for the difference of the Hamiltonian or energy between two levels of theory, while K, a, b, and s are the fitting parameters. In Figure 1, we demonstrate the simultaneous fitting to the UPDF to determine the change of free-energy for a water molecule from HF/6-31G(d) to MP2/6-311G(d,p). The intersection point of the two probability functions at −170.917 kcal/mol is the best estimate value for the ΔG in Table 1 above.

7. Conclusion and outlook

In this chapter, we (wongky@biomaps.rutgers.edu; kiniu@alumni.cuhk.net) discuss developing the $\psi\pi\phi$ method to systematically generate quantum free-energy profiles at an *ab initio* path-integral level in molecular simulations. Since quantum free energy or partition function is a universal central quantity in thermodynamics of biology, chemistry, and physics, we anticipate our $\psi\pi\phi$ method would be very crucial in both Life and Materials Sciences and wish that it could be used by non-specialists as a black box one day.

8. References

Anderson, J. B. (1975). Random-walk simulation of the Schroedinger equation. Hydrogen ion ($H3^+$). *Journal of Chemical Physics*, Vol. 63, No. 4, pp. 1499-1503.

Arfken, G. B. and H.-J. Weber (2001). *Mathematical methods for physicists*. San Diego, Academic Press. See exercise 2.2.3 for the fundamental definition of an orthogonal unit vector; See section 2.10 for the discussion of the covariant and contravariant vectors.

Ballhausen, C. J. and A. E. Hansen (1972). Electronic spectra. *Annual Review of Physical Chemistry*, Vol. 23, pp. 15-38.

Birkhoff, G. D. (1931). Proof of the Ergodic Theorem. *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 17, No. 12, pp. 656-660.

Born, M. and J. R. Oppenheimer (1927). Zur Quantentheorie der Molekeln (On the Quantum Theory of Molecules). *Annalen der Physik*. 84,pp. 457-484. English translation: *Quantum chemistry : classic scientific papers,* Hettema, H., Ed.; World Scientific: Singapore; London, 2000, pp 1-24.

Brooks, B. R., R. E. Bruccoleri, et al. (1983). CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry*, Vol. 4, No. 2, pp. 187-217.

Brown, L. M. (2005). *Feynman's thesis : a new approach to quantum theory*. Singapore ; Hackensack, NJ, World Scientific.

Cao, J. and G. A. Voth (1994). The formulation of quantum statistical mechanics based on the Feynman path centroid density. IV. Algorithms for centroid molecular dynamics. *Journal of Chemical Physics*, Vol. 101, No. 7, pp. 6168-6183.

Car, R. and M. Parrinello (1985). Unified approach for molecular dynamics and density-functional theory. *Physical Review Letters*, Vol. 55, No. 22, pp. 2471-2474.

Ceperley, D. M. (1995). Path integrals in the theory of condensed helium. *Reviews of Modern Physics*, Vol. 67, No. 2, pp. 279-355.

Chipot, C. and A. Pohorille (2007). Calculating free energy differences using perturbation theory. *Springer Series in Chemical Physics.*, Vol. 86, Free Energy Calculations, pp. 33-75.

Coalson, R. D. (1986). On the connection between Fourier coefficient and Discretized Cartesian path integration. *Journal of Chemical Physics*, Vol. 85, No. 2, pp. 926-936.

Cogswell, K. (1999). Computers and Advanced Mathematics in the Calculus Classroom. *The College Mathematics Journal*, Vol. 30, No. 3, pp. 213-216.

Dean, D. J. (2007). Beyond the nuclear shell model. *Physics Today*, Vol. 60, No. 11, pp. 48-53.

den Otter, W. K. (2000). Thermodynamic integration of the free energy along a reaction coordinate in Cartesian coordinates. *Journal of Chemical Physics*, Vol. 112, No. 17, pp. 7283-7292.

Deng, Y., B. M. Ladanyi, et al. (2002). High-frequency vibrational energy relaxation in liquids: the foundations of instantaneous-pair theory and some generalizations. *Journal of Chemical Physics*, Vol. 117, No. 23, pp. 10752-10767.

Doll, J. D. and L. E. Myers (1979). Semiclassical Monte Carlo methods. *Journal of Chemical Physics*, Vol. 71, No. 7, pp. 2880-2883.

Feynman, R. P. (1948). Space-Time Approach to Non-Relativistic Quantum Mechanics. *Reviews of Modern Physics*, Vol. 20, No. 2, pp. 367-387.

Feynman, R. P. (1966). The development of the space-time view of quantum electrodynamics. *Science*, Vol. 153, No. 3737, pp. 699-708.

Feynman, R. P., A. R. Hibbs, et al. (2005). *Quantum mechanics and path integrals*. Mineola, N.Y., Dover Publications.

Feynman, R. P. and H. Kleinert (1986). Effective classical partition functions. *Physical Review A*, Vol. 34, No. 6, pp. 5080-5084.

Field, M. J., P. A. Bash, et al. (1990). A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations. *Journal of Computational Chemistry*, Vol. 11, No. 6, pp. 700-733.

Fixman, M. (1974). Classical statistical mechanics of constraints: a theorem and application to polymers. *Proceedings of the National Academy of Sciences of the United States of America,* Vol. 71, No. 8, pp. 3050-3053.

Flanders, H. (1973). Differentiation Under the Integral Sign. *The American Mathematical Monthly,* Vol. 80, No. 6, pp. 615-627.

Gao, J., K.-Y. Wong, et al. (2008). Combined QM/MM and path integral simulations of kinetic isotope effects in the proton transfer reaction between nitroethane and acetate ion in water. *Journal of Computational Chemistry,* Vol. 29, No. 4, pp. 514-522.

Gao, J., D. T. Major, et al. (2008). Hybrid quantum and classical methods for computing kinetic isotope effects of chemical reactions in solutions and in enzymes. *Molecular Modeling of Proteins.* A. Kukol, Humana Press: 37-62.

Gao, J. and D. G. Truhlar (2002). Quantum mechanical methods for enzyme kinetics. *Annual Review of Physical Chemistry*, Vol. 53, pp. 467-505.

Giachetti, R. and V. Tognetti (1985). Variational approach to quantum statistical mechanics of nonlinear systems with application to sine-Gordon chains. *Physical Review Letters*, Vol. 55, No. 9, pp. 912-915.

Gillan, M. J. (1987). Quantum-classical crossover of the transition rate in the damped double well. *Journal of Physics C: Solid State Physics*, Vol. 20, No. 24, pp. 3621-3641.

Gillan, M. J. (1987). Quantum simulation of hydrogen in metals. *Physical Review Letters*, Vol. 58, No. 6, pp. 563-566.

Grossman, J. C. and L. Mitas (2005). Efficient Quantum Monte Carlo Energies for Molecular Dynamics Simulations. *Physical Review Letters*, Vol. 94, No. 5, pp. 056403/056401-056403/056404.

Hagler, A. T., E. Huler, et al. (1974). Energy functions for peptides and proteins. I. Derivation of a consistent force field including the hydrogen bond from amide crystals. *Journal of the American Chemical Society*, Vol. 96, No. 17, pp. 5319-5327.

Hayashi, A., M. Shiga, et al. (2006). H/D isotope effect on the dihydrogen bond of $NH4^{+}$.BeH2 by ab initio path integral molecular dynamics simulation. *Journal of Chemical Physics*, Vol. 125, No. 20, pp. 204310.

Hehre, W. J., L. Radom, et al. (1986). *Ab initio molecular orbital theory*. New York, Wiley.

Helgaker, T., P. Jørgensen, et al. (2000). *Molecular electronic-structure theory*. Chichester ; New York, Wiley.

Hénin, J. r., G. Fiorin, et al. (2010). Exploring Multidimensional Free Energy Landscapes Using Time-Dependent Biases on Collective Variables. *Journal of Chemical Theory and Computation*, Vol. 6, No. 1, pp. 35-47.

Hirschfelder, J. O. and W. J. Meath (1967). Nature of intermolecular forces. *Advances in Chemical Physics*, Vol. 12, pp. 3-106.

Hünenberger, P. H. (2005). Thermostat Algorithms for Molecular Dynamics Simulations. *Advances in Polymer Science*, Vol. 173, pp. 105-149.

Jorgensen, W. L. and J. Tirado-Rives (1988). The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society*, Vol. 110, No. 6, pp. 1657-1666.

Kirkwood, J. G. (1935). Statistical mechanics of fluid mixtures. *Journal of Chemical Physics*, Vol. 3, No. 5, pp. 300-313.

Kleinert, H. (2004). *Path integrals in quantum mechanics, statistics, polymer physics, and financial markets*. Singapore; River Edge, NJ, World Scientific. For the quantum mechanical integral equation, see Section 1.9; For the variational perturbation theory, see Chapters 3 and 5.

Kleppner, D. and R. Jackiw (2000). Pathways of discovery: One hundred years of quantum physics. *Science*, Vol. 289, No. 5481, pp. 893-898.

Kohn, W. (1999). Nobel lecture: electronic structure of matter-wave functions and density functionals. *Reviews of Modern Physics*, Vol. 71, No. 5, pp. 1253-1266.

Kolos, W. (1970). Adiabatic approximation and its accuracy. *Advances in Quantum Chemistry*, Vol. 5, pp. 99-133.

Kowalczyk, P., P. A. Gauden, et al. (2008). Cryogenic Separation of Hydrogen Isotopes in Single-Walled Carbon and Boron-Nitride Nanotubes: Insight into the Mechanism of Equilibrium Quantum Sieving in Quasi-One-Dimensional Pores. *Journal of Physical Chemistry B*, Vol. 112, No. 28, pp. 8275-8284.

Kowalczyk, P., P. A. Gauden, et al. (2007). Thermodynamics of Hydrogen Adsorption in Slit-like Carbon Nanopores at 77 K. Classical versus Path-Integral Monte Carlo Simulations. *Langmuir*, Vol. 23, No. 7, pp. 3666-3672.

Kubo, R. (1962). Generalized Cumulant Expansion Method. *Journal of the Physical Society of Japan*, Vol. 17, No. 7, pp. 1100-1120.

Lebowitz, J. L. and O. Penrose (1973). Modern Ergodic Theory. *Physics Today*, Vol. 26, No. 2, pp. 23-29.

Lester, W. A. and R. Salomon-Ferrer (2006). Some recent developments in quantum Monte Carlo for electronic structure: Methods and application to a bio system. *Theochem*, Vol. 771, No. 1-3, pp. 51-54.

MacKeown, P. K. (1985). Evaluation of Feynman path integrals by Monte Carlo methods. *American Journal of Physics*, Vol. 53, No. 9, pp. 880-885.

Major, D. T., A. Heroux, et al. (2009). Differential quantum tunneling contributions in nitroalkane oxidase catalyzed and the uncatalyzed proton transfer reaction. *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 106, No. 49, pp. 20734-20739.

Marx, D. and M. Parrinello (1995). Structural quantum effects and three-center two-electron bonding in $CH5^+$. *Nature (London)*, Vol. 375, No. 6528, pp. 216-218.

Marx, D., M. E. Tuckerman, et al. (1999). Quantum dynamics via adiabatic ab initio centroid molecular dynamics. *Computer Physics Communications*, Vol. 118, No. 2-3, pp. 166-184.

Mayo, S. L., B. D. Olafson, et al. (1990). DREIDING: a generic force field for molecular simulations. *Journal of Physical Chemistry*, Vol. 94, No. 26, pp. 8897-8909.

McQuarrie, D. A. (2000). *Statistical mechanics*. Sausalito, Calif., University Science Books

Mielke, S. L., K. A. Peterson, et al. (2003). $H+H_2$ Thermal Reaction: A Convergence of Theory and Experiment. *Physical Review Letters*, Vol. 91, No. 6, pp. 063201.

Mielke, S. L. and D. G. Truhlar (2001). Displaced-points path integral method for including quantum effects in the Monte Carlo evaluation of free energies. *Journal of Chemical Physics*, Vol. 115, No. 2, pp. 652-662.

Mielke, S. L. and D. G. Truhlar (2001). A new Fourier path integral method, a more general scheme for extrapolation, and comparison of eight path integral methods for the quantum mechanical calculation of free energies. *Journal of Chemical Physics*, Vol. 114, No. 2, pp. 621-630.

Nanda, H., N. Lu, et al. (2005). Using non-Gaussian density functional fits to improve relative free energy calculations. *Journal of Chemical Physics*, Vol. 122, No. 13, pp. 134110/134111-134110/134118.

Ohta, Y., K. Ohta, et al. (2004). Ab initio centroid path integral molecular dynamics: Application to vibrational dynamics of diatomic molecular systems. *Journal of Chemical Physics*, Vol. 120, No. 1, pp. 312-320.

Paesani, F., S. Iuchi, et al. (2007). Quantum effects in liquid water from an ab initio-based polarizable force field. *Journal of Chemical Physics*, Vol. 127, No. 7, pp. 074506.

Pople, J. A. (1999). Nobel lecture: Quantum chemical models. *Reviews of Modern Physics*, Vol. 71, No. 5, pp. 1267-1274.

Ramírez, R. and T. López-Ciudad (1999). The Schroedinger formulation of the Feynman path centroid density. *Journal of Chemical Physics*, Vol. 111, No. 8, pp. 3339-3348.

Ramírez, R., T. López-Ciudad, et al. (1998). Feynman Effective Classical Potential in the Schrödinger Formulation. *Physical Review Letters*. 81,pp. 3303-3306. Comment: Andronico, G.; Branchina, V.; Zappala, D. *Phys. Rev. Lett.* 2002, *88,* 178901; Reply to comment: Ramirez, R.; López-Ciudad, T. *Phys. Rev. Lett.* 2002, *88,* 178902.

Ramírez, R., T. López-Ciudad, et al. (2001). Low lying vibrational excitation energies from equilibrium path integral simulations. *Journal of Chemical Physics,* Vol. 115, No. 1, pp. 103-114.

Ruiz-Montero, M. J., D. Frenkel, et al. (1997). Efficient schemes to compute diffusive barrier crossing rates. *Molecular Physics*, Vol. 90, No. 6, pp. 925-942.

Sauer, T. (2001). The Feynman path goes Monte Carlo. *Fluctuating paths and fields : festschrift dedicated to Hagen Kleinert on the occasion of his 60th birthday*. W. Janke, A. Pelster, H.-J. Schmidt and M. Bachmann. River Edge, NJ, World Scientific,pp. 29-42. And in Los Alamos National Laboratory, Preprint Archive, Physics arXiv:physics/0107010v1 [physics.hist-ph], 1 (2001).

Springborg, M. (2000). *Methods of electronic-structure calculations : from molecules to solids*. Chichester ; New York, Wiley.

Stratt, R. M. (1995). The Instantaneous Normal Modes of Liquids. *Accounts of Chemical Research*, Vol. 28, No. 5, pp. 201-207.

Szabo, A. and N. S. Ostlund (1996). *Modern quantum chemistry : introduction to advanced electronic structure theory*. Mineola, N.Y., Dover Publications

Tanaka, H., H. Kanoh, et al. (2005). Quantum Effects on Hydrogen Isotope Adsorption on Single-Wall Carbon Nanohorns. *Journal of the American Chemical Society*, Vol. 127, No. 20, pp. 7511-7516.

Tuckerman, M. E. and D. Marx (2001). Heavy-Atom Skeleton Quantization and Proton Tunneling in "Intermediate-Barrier" Hydrogen Bonds. *Physical Review Letters*, Vol. 86, No. 21, pp. 4946-4949.

Tuckerman, M. E., D. Marx, et al. (1997). On the quantum nature of the shared proton in hydrogen bonds. *Science*, Vol. 275, No. 5301, pp. 817-820.

Tuckerman, M. E., D. Marx, et al. (2002). The nature and transport mechanism of hydrated hydroxide ions in aqueous solution. *Nature*, Vol. 417, No. 6892, pp. 925-929.

Voth, G. A. (1996). Path-integral centroid methods in quantum statistical mechanics and dynamics. *Advances in Chemical Physics*, Vol. 93, New Methods in Computational Quantum Mechanics, pp. 135-218.

Wagner, L. K., M. Bajdich, et al. (2009). QWalk: A quantum Monte Carlo program for electronic structure. *Journal of Computational Physics*, Vol. 228, No. 9, pp. 3390-3404.

Warshel, A., M. H. M. Olsson, et al. (2006). Computer simulations of isotope effects in enzyme catalysis. *Isotope Effects in Chemistry and Biology*. A. Kohen and H.-H. Limbach. Boca Raton, Taylor & Francis: 621-644.

Weiner, S. J., P. A. Kollman, et al. (1984). A new force field for molecular mechanical simulation of nucleic acids and proteins. *Journal of the American Chemical Society*, Vol. 106, No. 3, pp. 765-784.

Wong, K.-Y. (2008). Simulating biochemical physics with computers: 1. Enzyme catalysis by phosphotriesterase and phosphodiesterase; 2. Integration-free path-integral

method for quantum-statistical calculations. Ph.D. thesis, Minneapolis, University of Minnesota (USA), pp. 342.

Wong, K.-Y. and J. Gao (2007). The Reaction Mechanism of Paraoxon Hydrolysis by Phosphotriesterase from Combined QM/MM Simulations. *Biochemistry,* Vol. 46, No. 46, pp. 13352-13369.

Wong, K.-Y., T.-S. Lee, et al. (2011). Active Participation of the Mg^{2+} Ion in the Reaction Coordinate of RNA Self-Cleavage Catalyzed by the Hammerhead Ribozyme. *Journal of Chemical Theory and Computation.* Vol. 7, No. 1, pp. 1-3.

Wong, K.-Y. and J. Gao (2011). Insight into the phosphodiesterase mechanism from combined QM/MM free energy simulations. *FEBS Journal,* Vol. 278, No. 14, pp. 2579-2595.

Wong, K.-Y. and J. Gao (2007). An automated integration-free path-integral method based on Kleinert's variational perturbation theory. *Journal of Chemical Physics*, Vol. 127, No. 21, pp. 211103.

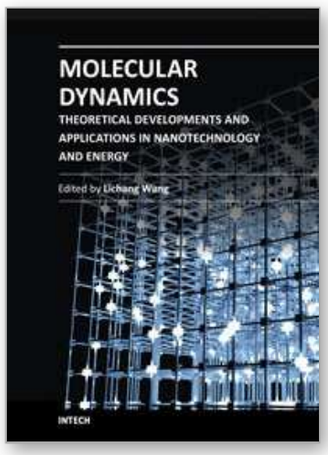
Wong, K.-Y. and J. Gao (2008). Systematic Approach for Computing Zero-Point Energy, Quantum Partition Function, and Tunneling Effect Based on Kleinert's Variational Perturbation Theory. *Journal of Chemical Theory and Computation*, Vol. 4, No. 9, pp. 1409-1422.

Wong, K.-Y., J. P. Richard, et al. (2009). Theoretical Analysis of Kinetic Isotope Effects on Proton Transfer Reactions between Substituted alpha -Methoxystyrenes and Substituted Acetic Acids. *Journal of the American Chemical Society*, Vol. 131, No. 39, pp. 13963-13971.

Wong, K.-Y., H. Gu, et al. (2012). Characterization of the Reaction Path and Transition States for RNA Transphosphorylation Models from Theory and Experiment. *Angewandte Chemie, International Edition*, Vol. 51, No.3, pp. 647-651.

Wu, E. L., K.-Y. Wong, et al. (2009). Determination of the Structure Form of the Fourth Ligand of Zinc in Acutolysin A Using Combined Quantum Mechanical and Molecular Mechanical Simulation. *Journal of Physical Chemistry B,* Vol. 113, No.8, pp. 2477-2485.

Zwanzig, R. W. (1954). High-temperature equation of state by a perturbation method. I. Nonpolar gases. *Journal of Chemical Physics*, Vol. 22, No. 8, pp. 1420-1426.



Molecular Dynamics - Theoretical Developments and Applications in Nanotechnology and Energy

Edited by Prof. Lichang Wang

ISBN 978-953-51-0443-8
Hard cover, 424 pages
Publisher InTech
Published online 05, April, 2012
Published in print edition April, 2012

Molecular Dynamics is a two-volume compendium of the ever-growing applications of molecular dynamics simulations to solve a wider range of scientific and engineering challenges. The contents illustrate the rapid progress on molecular dynamics simulations in many fields of science and technology, such as nanotechnology, energy research, and biology, due to the advances of new dynamics theories and the extraordinary power of today's computers. This first book begins with a general description of underlying theories of molecular dynamics simulations and provides extensive coverage of molecular dynamics simulations in nanotechnology and energy. Coverage of this book includes: Recent advances of molecular dynamics theory Formation and evolution of nanoparticles of up to 106 atoms Diffusion and dissociation of gas and liquid molecules on silicon, metal, or metal organic frameworks Conductivity of ionic species in solid oxides Ion solvation in liquid mixtures Nuclear structures

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen