Efficient Path Integral Molecular Dynamics

Ethan Brown, Timon Hilker, and Emily Shiavone[1]

[1]*Department of Physics, University of Illinois at Urbana-Champaign,*

1110 W. Green St. , Urbana, IL 61801-3080, USA

(Dated: May 16, 2011)

As initially formulated, Path Integral Molecular Dynamics provides very poor efficiency during simulation. In order to improve this efficiency, we examine various variable transformations, which decouple systematic frequency modes, and thermostats, which move the system from the microcanonical to the canonical ensemble. To compare these various methods, we focus on few particles in a harmonic trap both with and without interactions.

CONTENTS

I. INTRODUCTION

A. Motivation

Often times simulations of atomic systems treat the nuclei of the atoms as point particles and make use of the Born-Oppenheimer approximation. The Born-Oppenheimer approximation exploits the fact that the nucleus of an atom is many times larger than electrons, and thus treats the nuclei as having fixed electronic degrees of freedom. A better approximation might use an ab initio technique such as Density Functional Theory (DFT) or Quantum Monte Carlo (QMC) to relax this condition. However, these approximations normally do not account for thermal expansion of a crystal lattice and can be insufficient when describing systems involving atoms with small nuclei[4]. Conventional ab initio molecular dynamics may take into account such thermal effects, but it still treats the nuclei as classical particles, thus losing often essential quantum mechanical behavior. Thus our attention is focused on the ab initio path integral molecular dynamics method which accounts for all these effects from first principles[7]. For simplicity we are working without the DFT calculation so far and use fixed potential interaction (i.e. free particles or Lennard-Jones).

B. Path Integrals

Feynman path integrals turn a quantum mechanical problem into a series of classical problems. The *action* is defined as

$$\int L dt \tag{1.1}$$

where L is the Lagrangian. Classically, a particle follows a single path from its starting point x to a new position x', corresponding to the path of least action[5]. Using the path integral representation of quantum mechanics the particle, the particle samples all paths from x to x' simultaneously, where each path has a given amplitude. The probability that a particle at position x will later be observed at position x' is the square modulus of the sum of these interfering amplitudes.

C. Path Integral Molecular Dynamics

The canonical one particle partition function can be written as,

$$Z(\beta) = Tr(e^{-\beta H}) = \int dx \left\langle x | e^{-\beta H} | x \right\rangle \tag{1.2}$$

The Hamiltonian in the exponent is the sum of two non-commutative operators, the kinetic and potential energy. Therefore, to separate the term into the product of two non-coupled exponentials, the Trotter expansion must be used

$$e^{\lambda(A+B)} = lim_{P\to\infty}[e^{\frac{\lambda}{2P}A}e^{\frac{\lambda}{P}B}e^{\frac{\lambda}{2P}A}] \tag{1.3}$$

This expression is only exact in the limit P goes to infinity. The error in the Trotter expansion is order τ^2 where $\tau \equiv \beta/P$. With this approximation, one is lead to the following expression for the partition function,

$$Z_P(\beta)) = (\frac{mP}{2\pi\beta\hbar^2})^{\frac{P}{2}} \int dx_1 \cdots dx_P exp\{-\sum_{i=1}^{P}[\frac{mP}{2\beta\hbar^2}(x_{i+1}-x_i)^2 + \frac{\beta}{P}\phi(x_i)]\}_{x_{P+1}=x_1} \tag{1.4}$$

This expression is known as the *discretized path integral*. This Hamiltonian looks precisely like a set of P coupled oscillators in some external potential U_{eff} where,

$$U_{eff}(x_1,\ldots,x_p) = \sum_{i=1}^{P}[\frac{1}{2}\frac{mP}{\hbar^2}(x_{i+1}-x_i)^2 + \frac{1}{P}\phi(x_i)]_{x_{P+1}=x_1} \tag{1.5}$$

However, in order to perform a molecular dynamics simulation on this Hamiltonian, we must add in conjugate momenta. Doing so makes the partition function,

$$Z_P(\beta)) = N \int dp_1 \cdots dp_P \int dx_1 \cdots dx_P exp\{-\beta[\sum_{i=1}^{P}\frac{p_i^2}{2\tilde{m}_i} + U_{eff}(x_1\ldots x_P)]\} \tag{1.6}$$

where N is a constant. Notice that since these additional degrees of freedom are quadratic and decoupled (gaussian), and thus merely change the overall normalization factor, not the overall dynamics[6]. The effective Hamiltonian becomes,

$$H(p,x) = \sum_{i=1}^{P}\frac{p_i^2}{2\bar{m}_i} + U_{eff}(x_1,\ldots,x_P)$$

$$H(p,x) = \sum_{i=1}^{P}[\frac{p_i^2}{2\tilde{m}_i} + \frac{1}{2}m\omega_p^2(x_{i+1}-x_i)^2 + \frac{1}{P}\phi(x_i)]_{x_{P+1}=x_P} \tag{1.7}$$

We are now ready to perform molecular dynamics simulations. There are now P "beads" per quantum mechanical particle. Each bead can be treated as a classical particle in a molecular dynamics simulation with the interaction Ueff. Note that every bead interacts only with the two beads that are nearest neighbors in imaginary time. Hence the one particle is represented by a necklace with P beds. To do the MD simultation, we simply implement the usual Velocity Verlet algorithm acting on each bead of every particle,

$$x(\Delta t) = x(0) + \Delta t\frac{p(0)}{m} + \frac{\Delta t^2}{2m}F(x(0))$$

$$p(\Delta t) = p(0) + \Delta t\frac{p(0)}{2}[F(x(0)) + F(x(\Delta t))] \tag{1.8}$$

In practice there are some issues that need to be addressed so that the MD simulation actually runs well. These issues will be addressed in the following sections.

II. VARIABLE TRANSFORMATIONS

Recall the Hamiltonian,

$$H(p,x) = \sum_{i=1}^{P} \left[\frac{p_i^2}{2\tilde{m}_i} + \frac{1}{2}m\omega_p^2(x_{i+1} - x_i)^2 + \frac{1}{P}\phi(x_i) \right]_{x_{P+1}=x_P} \tag{2.1}$$

where $\omega_P = \frac{\sqrt{P}}{\beta\hbar}$.

The Trotter expansion leads to an approximation that approaches a quantum system as the number of beads goes to infinity. However, as the number of beads, P, is increased ω_P scales linearly with the number of beads. Thus the entire second term scales linearly with the number of beads. Comparatively the first term does not depend on P and the external potential term varies inversely with P. As a result, the forces due to the harmonic term, which couples all of the beads, dominate. This strong coupling restricts the motion of the necklace as well as the size of the time step that may be used. These challenges are addressed with a change of variables and a time scale iteration algorithm that incorporates two different time scales[6].

In order to avoid these issues, we wish to uncouple the harmonic term through a variable transformation. In this project we investigate the effects of two different approaches: the Staging transformation and the normal mode transformation.

A. Staging Transformation

The first change of variables discussed is known as a *staging transformation*,

$$x_1 = u_1$$
$$x_i = u_i + \frac{i-1}{i}x_{i+1} + \frac{1}{i}u_1 \tag{2.2}$$

where $i = 2, \ldots, P$.

With this transformation the masses are no longer uniform. Instead they become,

$$m_1 = 0$$
$$m_k = \frac{k}{k-1}m \tag{2.3}$$

where $k = 2, \ldots, P$.

so that the new masses become $\tilde{m}_1 = m$ and $\tilde{m}_k = \frac{k}{k-1}m$.

Furthermore, this transformation leads to the following equations of motion,

$$\dot{u}_i = \frac{p_i}{\tilde{m}_i}$$
$$\dot{p}_i = -\frac{\partial U_{eff}^{stage}}{\partial u_i} = -m_i\omega_P^2 u_i - \frac{1}{P}\frac{\partial \phi}{\partial u_i} \tag{2.4}$$

Finally using the chain rule we arrive upon the following forces,

$$\frac{1}{P}\frac{\partial \phi}{\partial u_1} = \frac{1}{P}\sum_{i=1}^{P}\frac{\partial \phi}{\partial x_i}$$

$$\frac{\frac{1}{P}\partial \phi}{\partial u_i} = \frac{1}{P}[\frac{\partial \phi}{\partial x_i} + \frac{i-2}{i-1}\frac{\partial \phi}{\partial x_{i-1}}] \qquad (2.5)$$

B.   Normal Mode Transformation

Another change of variables that may be used to uncouple the harmonic term is the normal mode transformation[6]. This method may either be implemented with a decomposition into Fourier modes. Such a change corresponds to

$$x_k = \sqrt{P}\sum_{l=1}^{P}O_{kl}^{T}u_l \qquad (2.6)$$

where O_{kl}^{T} is an orthogonal matrix that diagonalizes a matrix of the form

$$\begin{bmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 & -1 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & & \vdots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & \ddots & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 & -1 \\ -1 & 0 & 0 & 0 & 0 & 0 & -1 & 2 \end{bmatrix} \qquad (2.7)$$

For this transformation, the equations of motion are the same as before

$$\dot{u}_i = \frac{p_i}{\tilde{m}_i}$$

$$\dot{p}_i = -\frac{\partial U_{eff}^{normal}}{\partial u_i} = -m_i\omega_P^2 u_i - \frac{1}{P}\frac{\partial \phi}{\partial u_i} \qquad (2.8)$$

However, now the masses are given by,

$$m_k = m\lambda_k \qquad (2.9)$$

where λ_k are the eigenvalues of O_{kl}^{T}. Finally, the forces are given by,

$$\frac{1}{P}\frac{\partial \phi}{\partial u_1} = \frac{1}{P}\sum_{l=1}^{P}\frac{\partial \phi}{\partial x_l}$$

$$\frac{1}{P}\frac{\partial \phi}{\partial u_k} = \frac{1}{\sqrt{P}}\sum_{l=1}^{P}\frac{\partial \phi}{\partial x_l}O_{kl}^{T} \qquad (2.10)$$

III. THERMOSTATS

As written in the introduction, our molecular dynamics simulation is in the constant energy micro-canonical ensemble. To sample the partition function at a constant temperature canonical ensemble a thermostat is usually required. Note that necessarily this is unphysical since one is adding degrees of freedom to the system which are not truly present and changing the conserved quantity (the Hamiltonian). Nevertheless, there has been much success with correctly reproducing the correct canonical distribution through thermostatting. We compare two methods: The Nose-Hoover thermostat deterministically introduces new degrees of freedom that lead to a constant temperature. It is considered the *gold standard* of thermostatting. The Langevin thermostat makes use of the Langevin Equation to mimic a heat bath and it thus stochastic. It employs the Einstein relation between friction and the random force to set the temperature.

A. Nose-Hoover Thermostat

The Nose-Hoover thermostat centers on an extended system method[2]. The basic Nose-Hoover thermostat introduces one additional degree of freedom, $\tilde{s}$, such that the Lagrangian becomes,

$$L_{Nose}(\tilde{r}, \dot{\tilde{r}}, \tilde{s}, \dot{\tilde{s}}) = \frac{1}{2}\sum_{i=1}^{N}m_i\tilde{s}^2\dot{\tilde{r}}_i^2 - U(\tilde{r}) + \frac{1}{2}Q\dot{\tilde{s}}^2 - gk_BT_0 ln(\tilde{s}) \qquad (3.1)$$

where the denotes extended system variables. Now our conserved quantity is simply the original Hamiltonian plus the energy of the added Nose-Hoover degrees of freedom. It follows that the equations of motion are,

$$\ddot{\tilde{r}}_i = \frac{F_i}{m_i\tilde{s}^2} - \frac{2\dot{\tilde{s}}\dot{\tilde{r}}_i}{\tilde{s}}$$

$$\ddot{\tilde{s}}_i = \frac{\sum_{i=1}^{N}m_i\dot{\tilde{r}}_i^2\tilde{s}^2 - gk_BT_0}{Q\tilde{s}} \qquad (3.2)$$

The original Nose-Hoover thermostat is only valid for ergodic systems. For application to a non-ergodic system, a Nose-Hoover chain must be used[3]. This method introduces additional degrees of freedom that couple to the original thermostat in order to shake up the system. The length of this chain is a variable which must be determined through testing.

For path integral molecular dynamics this amounts to coupling each bead on each particle with a 1D Nose-Hoover chain. Though this may seem like a massive overhead just to achieve a canonical distribution, the time required to thermalize is usually still much less than the time required for force calculations[6].

When integrating the Nose-Hoover equations of motion (usually through use of the Liouville operator), it is convenient to factorize the action further than the traditional Trotter expansion. To do so, we use the Suzuki-Yoshida scheme which applies a specific numerical weight to each term in the factorization where the sum of all the weights is necessarily equal to 1. This weight depends on the order of the scheme chosen, where a higher order usually results in a higher accuracy. Once an order is chosen, the weights can be determined numerically, though we simply quote the results from previous work[6]. We also determine the optimal order through testing.

Finally, one could imagine thermalization acting on a faster time scale than e.g. long-range interaction. Because of this, it is often convenient to use a smaller time step when integrating the Nose-Hoover equations of motion. This integration step is then looped over until the total time equals the time step for the rest of the simulation, $\delta t = \Delta t / n$. Again, the optimal factor must be determined through testing.

B. Langevin Thermostat

An alternative to the deterministic Nose Hoover thermostat is the stochastic Langevin thermostat[1]. This thermostat modifies the forces by adding a friction term and a thermal noise term,

$$
\begin{aligned}
\frac{\partial}{\partial t} \tilde{q}_i^{(k)} &= \frac{\tilde{p}_i^{(k)}}{m_i} \\
\frac{\partial}{\partial t} \tilde{p}_i^{(k)} &= -m_i \omega_k^2 \tilde{q}_i^{(k)} - \gamma^{(k)} \tilde{p}_i^{(k)} + \sqrt{\frac{2 m_i \gamma^{(k)}}{\beta_n}} \xi_i^{(k)}(t)
\end{aligned}
\tag{3.3}
$$

where $\gamma^{(k)}$ are the friction coefficients and $\xi_i^{(k)}(t)$ is a normally-distributed random force. Here our conserved quantity changes from the total energy of the system to the total energy minus the the accumulated heat absorbed by the thermostat. The Langevin thermostat ensures canonical dynamics no matter the value of $\gamma^{(k)}$, though better values should lead to faster convergence. The optimal values of $\gamma^{(k)}$ can be determined recursively[1]. However, for our purposes, this made little difference to the efficiency of the simulation. Thus we let each $\gamma^{(k)}$ be unity.

IV. NON-INTERACTING RESULTS

Below we discuss all our results for both non-interacting and interacting systems. Note that for all simulations we are in 1D and contained in a harmonic well external potential.

A. Nose-Hoover Testing

As mentioned previously, in order to fairly compare the Nose-Hoover thermostat, it was necessary to find the optimal parameters for the length of the Nose-Hoover chain, the order of the Suzuki-Yoshida expansion, and the number of discretized time steps required for proper thermalization. In order to do so, we compare convergence efficiencies for a single particle in a harmonic trap for a series of various parameters. Specifically, we look at a

low temperature, $\beta = 10.0$, and a high temperature, $\beta = 1.0$, in order to find settings which are suitable for a range of temperatures.
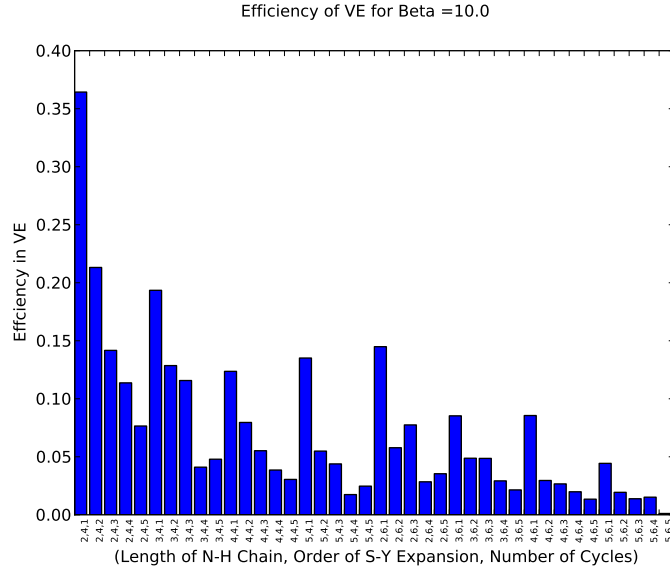
Efficiency of VE for Beta =10.0

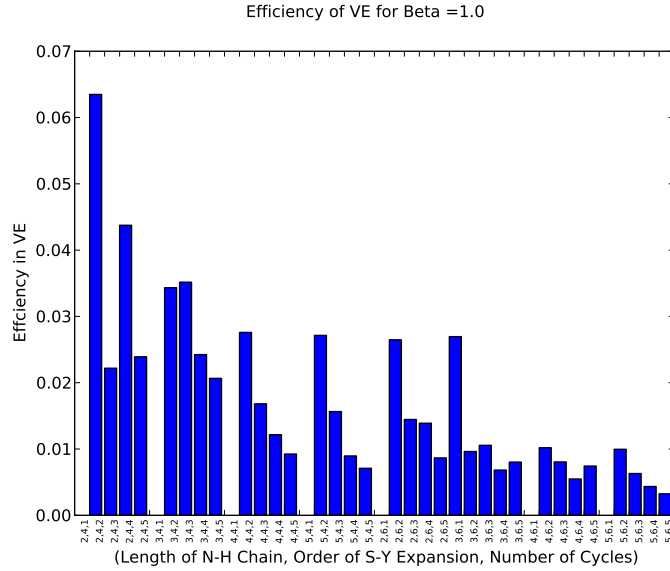FIG. 1. $\beta = 10.0$.

Efficiency of VE for Beta =1.0

FIG. 2. $\beta = 1.0$.

As one can clearly see, at low temperatures, the settings (2,4,1) are most efficient. This corresponds to a Nose-Hoover length of 2, a Suzuki-Yoshida expansion of order 4, and a single thermalization step. However, at high temperatures these settings result in zero efficiency. This implies that a single thermalization step is not adequate for proper coupling to the heat bath. Thus, we settle upon the second set of settings (2,4,2) for the remainder of our methods comparison. Note that for all comparisons we choose the virial energy estimator

which proved more reliable than the primitive estimator.

B. Comparison of Methods

To study the non-interacting system we examine only a single particle in one dimension in a harmonic trap. (Studying more particles without interaction would prove unnecessary.) Below we plot convergence in the number of time slices for each method combination. Note that we are showing the scaled virial energy estimator. It is being scaled by the analytical classical energy, making 0 the classical energy. The black line represents the exact quantum mechanical energy found through exact diagonalization. For all simulations, we used 1000000 sweeps with a time step of 0.01 and a block size of 10000.
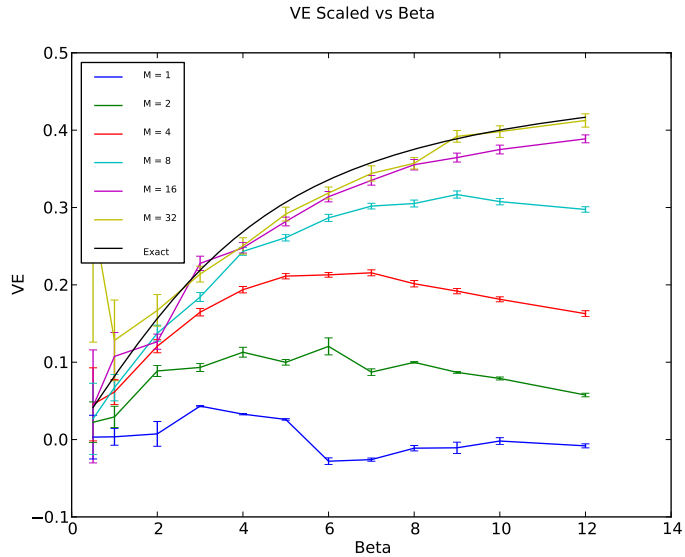
FIG. 3. No Transformation, Nose-Hoover.

It is difficult to tell how each method performs relative to the others from these plots. However, it does appear that methods using the Langevin thermostat are better converged as well as are those using some variable transformation. For all the methods we notice an anomalous gap between the converged result and the exact quantum mechanical result around $\beta = 4.0$. It has yet to be confirmed, but this feature is likely a cause of the non-ergodicity of the harmonic trap. As mentioned in a previous section, any sort of thermostatting is in essence unphysical since it adds degrees of freedom not necessarily found in the actual system. What is important, however is that all the methods are converging to approximately the same result making their direct comparison a viable option.

For a better comparison of the methods we must look at the relative simulation durations and errors. Using these two numbers, we can then compute the efficiency of each simulation, $\zeta \sim 1/(T\sigma^2)$.

This plot demonstrates the each method combination converges to approximately the same result. Here the number of times slices is 32. Thus, a comparison of the methods must look at second order results such as the
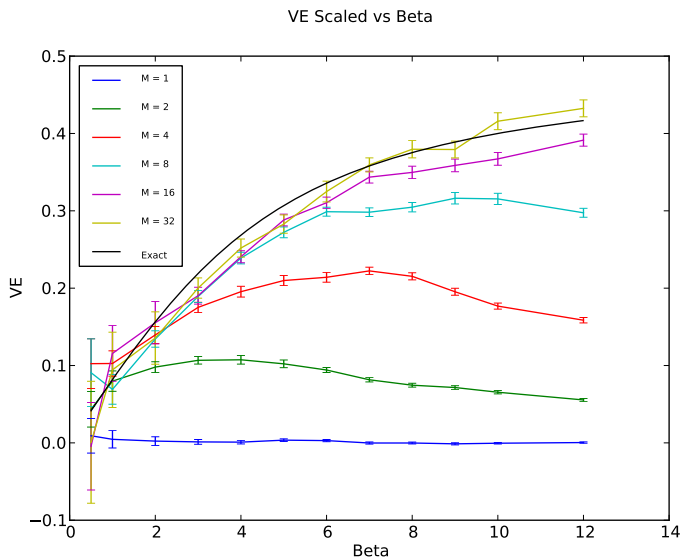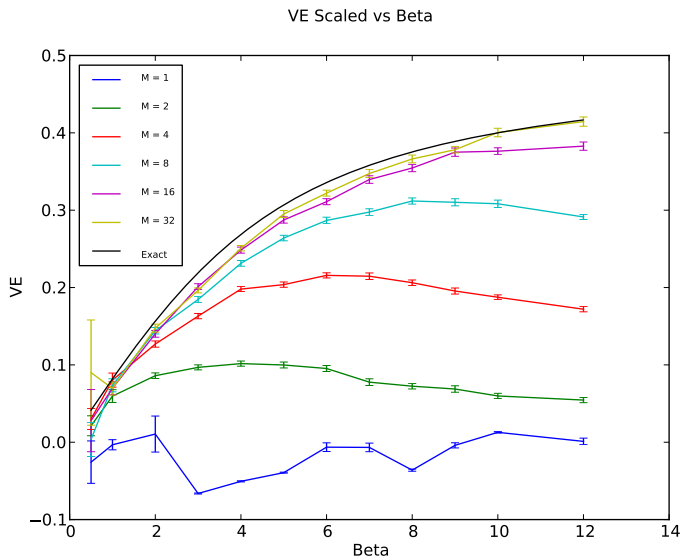
FIG. 4. No Transformation, Langevin.

FIG. 5. Staging, Nose-Hoover.

error and subsequently, the efficiency.

Errors between the methods are fairly similar for most temperatures. However, at high temperatures the simulations without a variable transformation perform worse. This is because, as mentioned previously, without a transformation all modes are coupled. At high temperatures this causes the particles to oscillate more wildly.

The above plot compares the average time required to complete a single block for each method combination. The two quickest methods are using the Langevin thermostat, one without a transformation and one with the staging transformation. Methods using the normal mode transformation tend to be slower. This is most likely
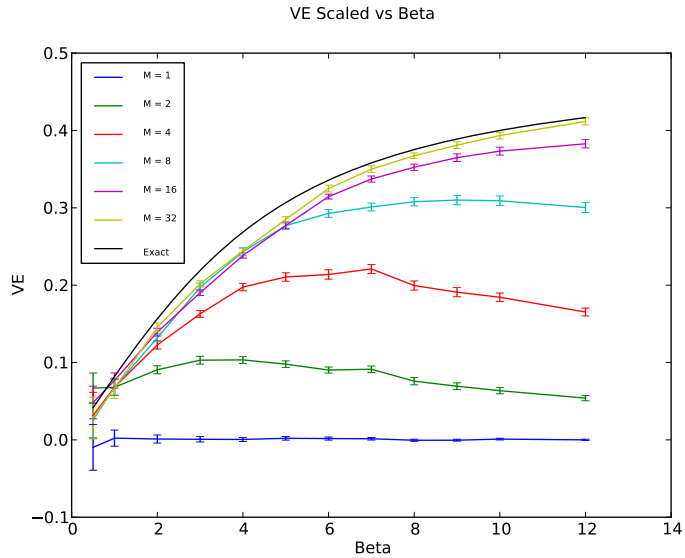
FIG. 6. Staging, Langevin.
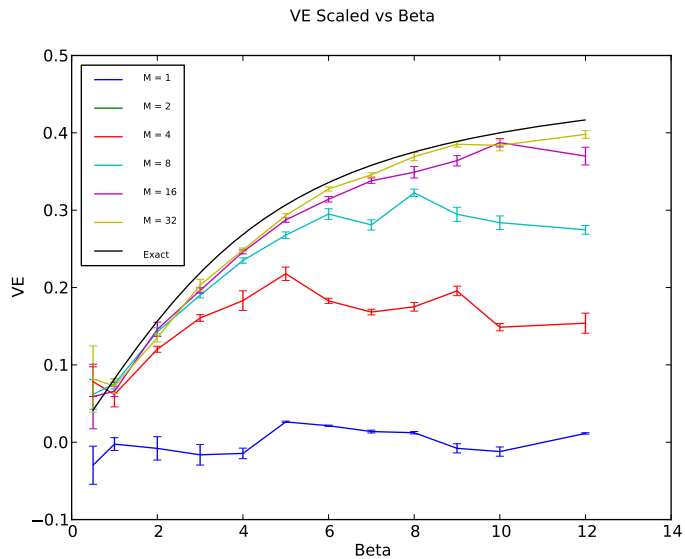
FIG. 7. Normal Mode, Nose-Hoover.

because the orthogonal matrix multiplication required for the method must loop through all time slices. On the other hand, the staging transformation is defined by a recursion relation requiring information on the previous time slice. This discrepancy could be made smaller using Fast Fourier Transforms with is order $Nlog(N)$. However, this is still larger than the order of the staging transformation.

As one can see, the combination of a staging transformation coupled to a Langevin thermostat proved to be the most efficient method for swift convergence. The discrepancies between the other methods are debatable, though it seems that having no transformation at all is never favorable.

VE Scaled vs Beta
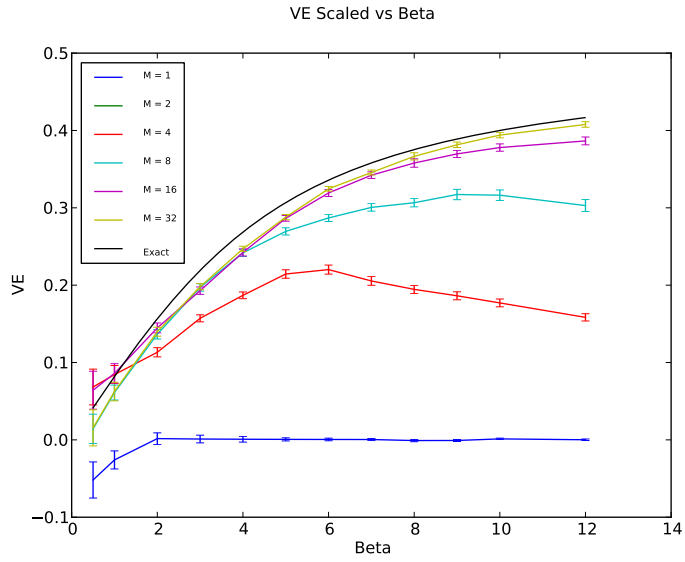
FIG. 8. Normal Mode, Langevin.
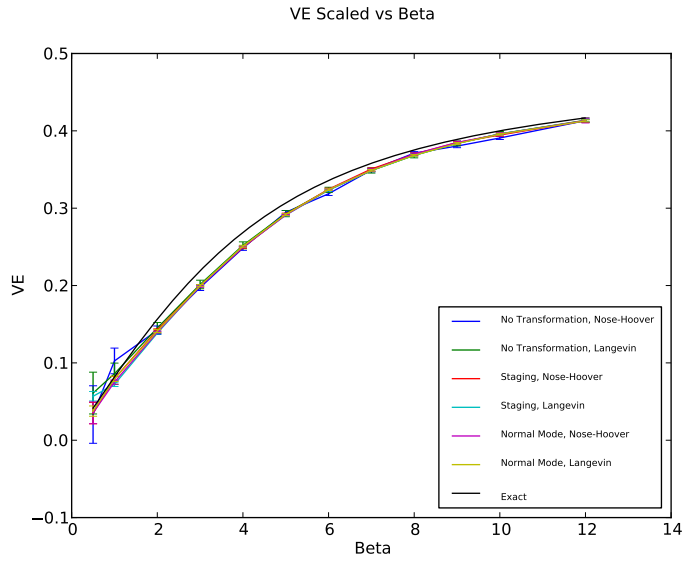
VE Scaled vs Beta

FIG. 9. Consistency Between Methods.

V. INTERACTING RESULTS

Now we turned to the more interesting case of interacting quantum particle. We chose particles trapped in a harmonic potential with a Lennard Jones interaction ($\epsilon = 1, r_0 = 1$) . To really do ab initio calculations one would need to implement a running DFT calculation at each time slice. This could be added to the code but we worked with a fixed potential for simplicity.
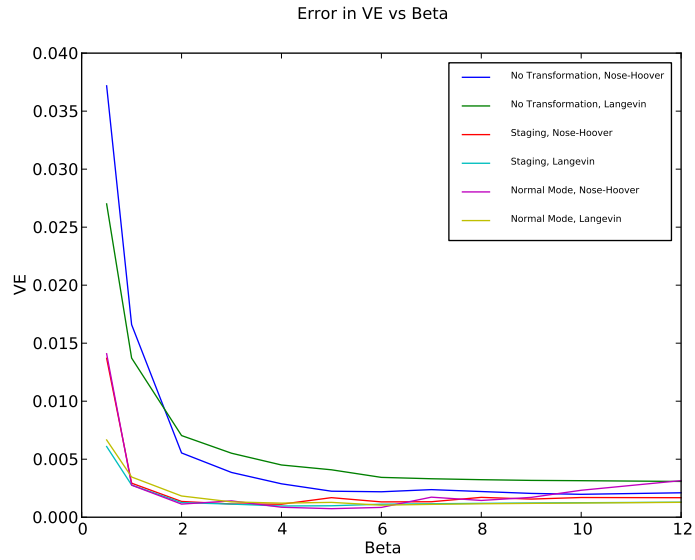
FIG. 10. Error of Methods.
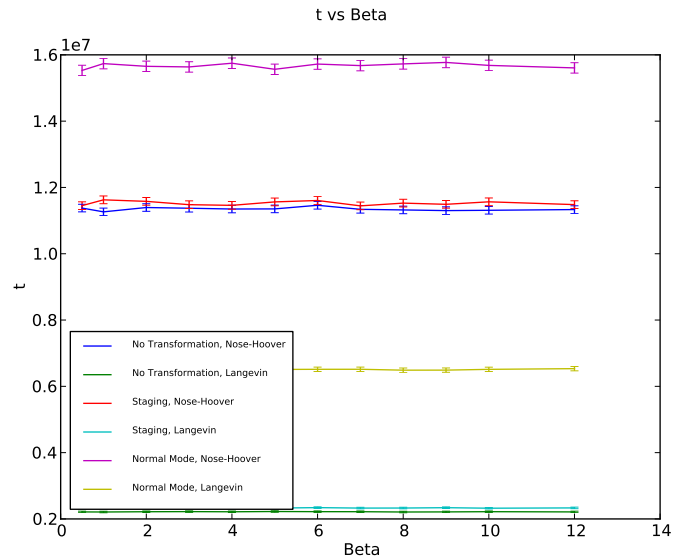
FIG. 11. Time per Block of Methods.

A. Two-particle Check

To have an idea if the results we are getting from the simulation actually approach the correct quantum mechanical result, we started with only two particles. The plot shows again the total energy (by the Virial Energy Estimator VE) at different temperatures beta and with different numbers of beads per particle. For comparison we added the classical result from statistical mechanics (green) obtained by numerical integration of the partition function and the quantum result (blue) obtained by direct diagonalization of a discretized Hamiltonian. As in the non-interacting case the lines with different number beads of per particle interpolate
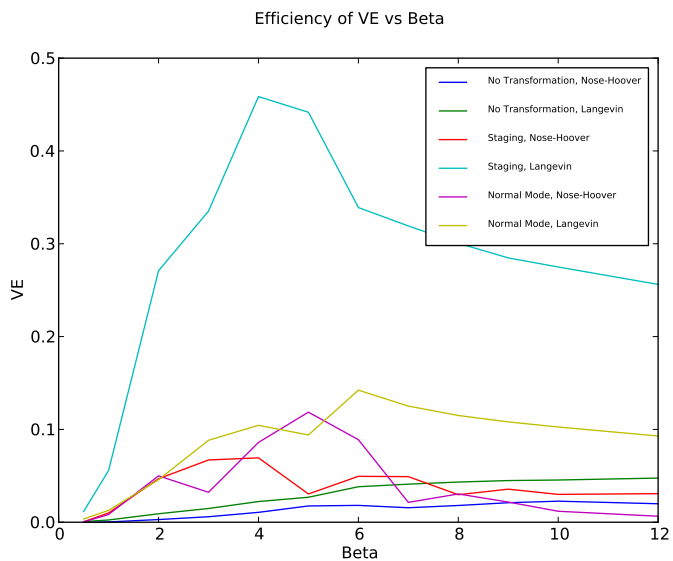
FIG. 12. Efficiency of Methods.

between these two limits. But convergence to the quantum behavior is much slower. Even for 2 times 64 beads, there is still a noticeable offset from the quantum result. For only one bead per particle, which should be a result classical Molecular Dynamics simulation we do not understand yet the reason for the offset from the result by integration of the partition function.
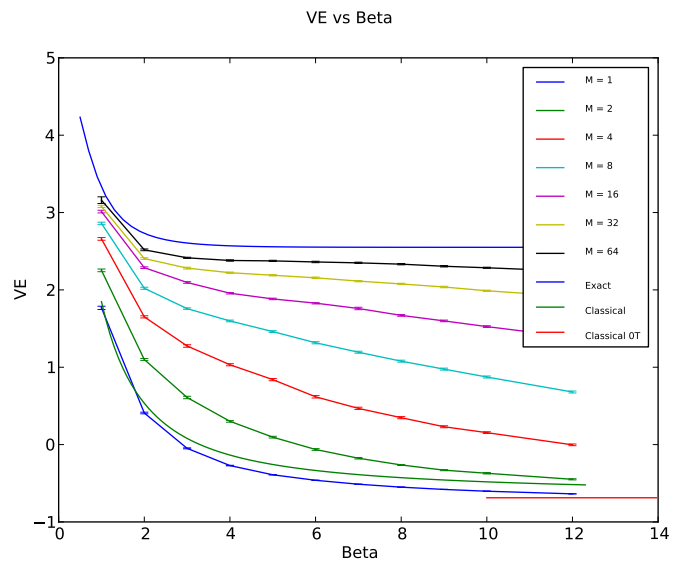

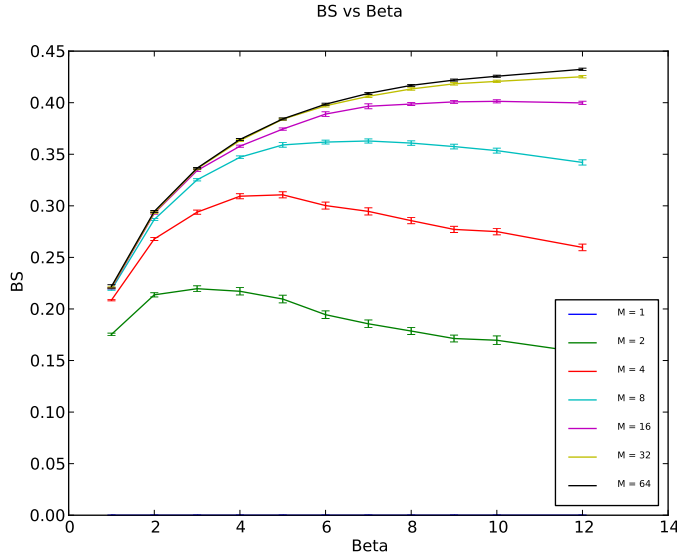FIG. 13. Virial Energy Estimator.

FIG. 14. Bead Spread.

B. Five-particle Study

After we have been convinced that our code converges towards the exact quantum limit for large numbers of beads, we did again the comparison of the different transformations and thermostats. To have little more interesting system, we chose 5 particles with Lennard-Jones interaction in a harmonic trap.

1. Comparison of Methods

The first plot shows that all methods converge to the same result. This is for 32 beads per particle (thus 160 beads total), where we are still quite far away from the quantum limit. Still it should be possible to analyze the efficiency and draw some results from it that are also true for larger numbers of beads.

So we look at the CPU-time it takes to run the different methods (with the same time step, same number of steps) and at the statistical error (uncertainty). One can see again that the Nose-Hoover thermostat really slows the calculation because of all the additional degrees of freedom. This effect should reduce however for increasing number of interacting particles because the force/energy calculation is the only part that scales as N^2 and will dominate for large N, so that the extra time of the Nose Hoover thermostat will be less relevant for the efficiency. So far we could not find an advantage of the Nose-Hoover over the Langevin thermostat by comparing the statistical error. As expected one can also see that the statistical error of the no-transformation calculations is a lot higher. Thus, a transformation is important to achieve fast convergence.

Finally we calculated the efficiency. For the case of 5 particles the Langevin Thermostat thermostat with the Staging Transformation clearly outperforms all other combinations. So this result agrees with the non-interacting case.
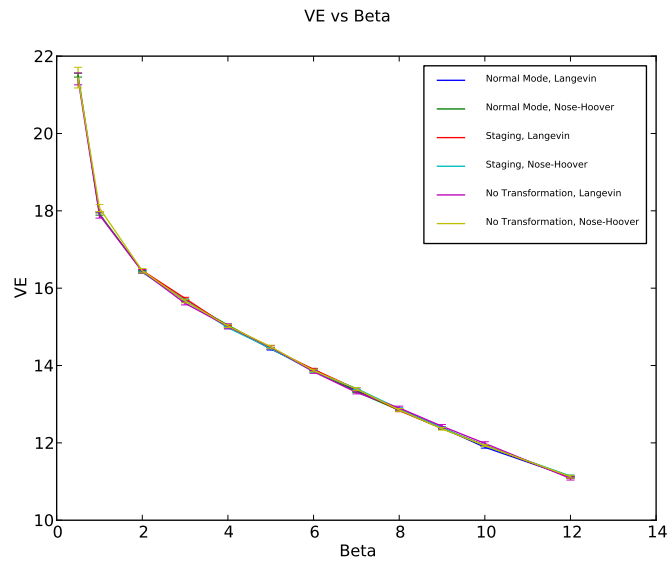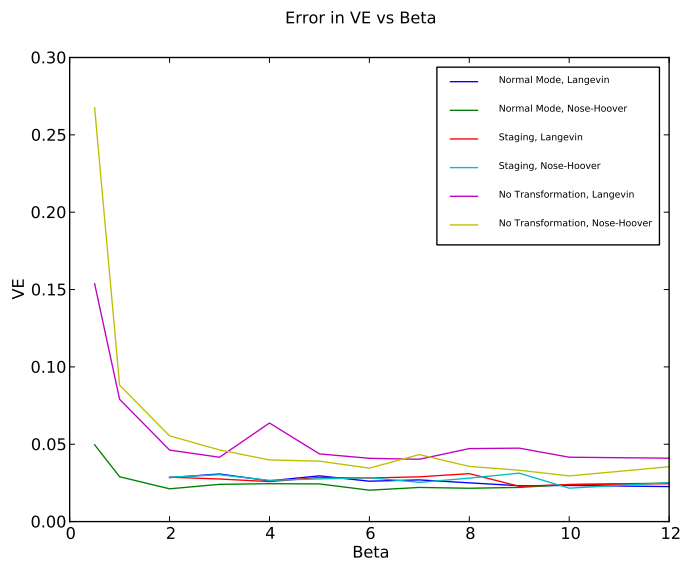
FIG. 15. Consistency Between Methods.

FIG. 16. Error of Methods.

2. Supplementary Observables

We also looked at some other observables like the density and the pair correlation function.

The density shows how the five particle line up in the one dimensional trap for low temperature. One in the middle and two on each side (Plot only show half the trap for $x > 0$). One can see how the system thermally expands with increasing temperature due to the non-harmonic parts of the Lennard-Jones interaction. The density was calculated via a histogram of the bead centroids.

The pair correlation (also calculated from the bead centroids) shows the same behavior of localized particles
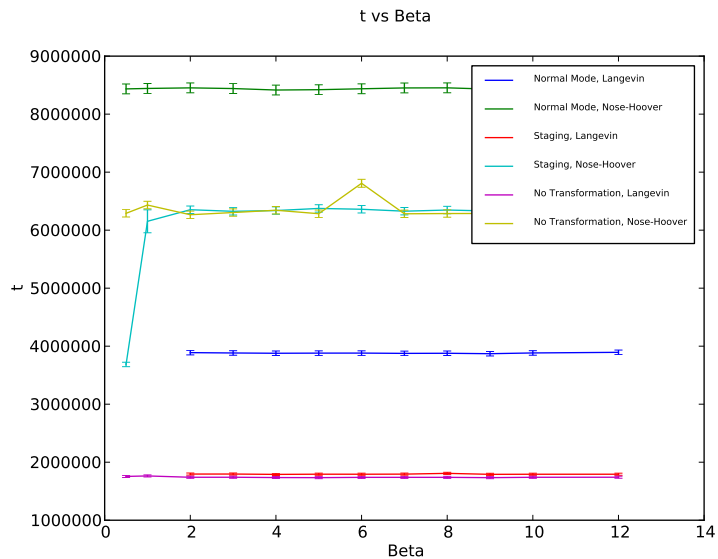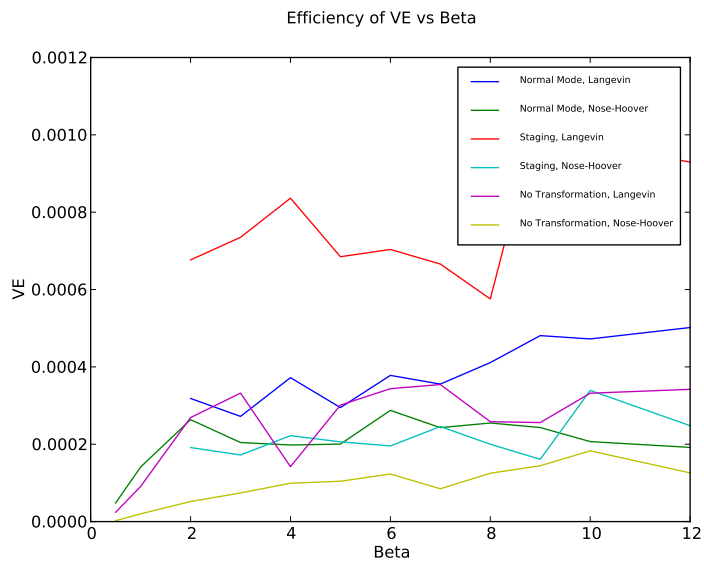
FIG. 17. Time per Block of Methods.

FIG. 18. Efficiency of Methods.

at low temperatures, which thermally expand for higher temperatures. The decrease of the height of the peaks is due to the finite number of particles (without periodic boundaries).

VI. CONCLUSIONS

We are able to use PIMD to model a quantum mechanical system. At high temperatures, PIMD simulations approach a classical result where all of the beads move close together. At lower temperatures, the beads spread
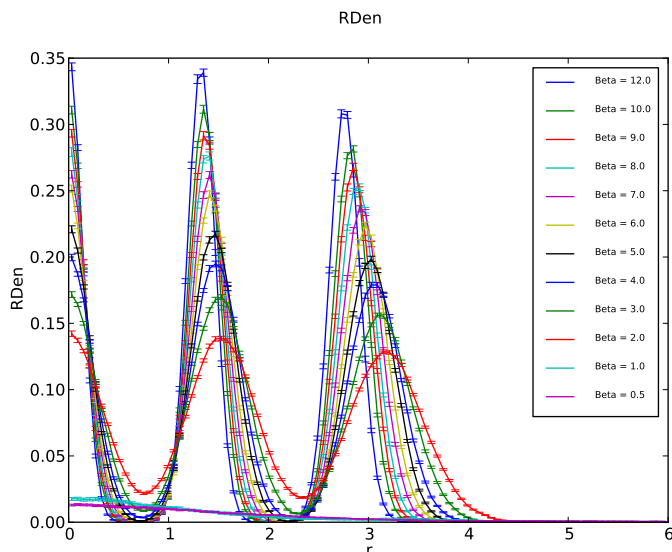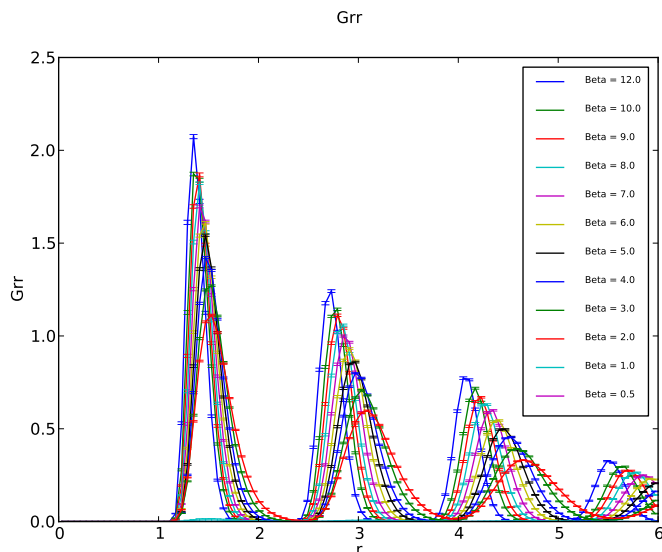
FIG. 19. Radial Density.

FIG. 20. Pair Correlation.

out, corresponding to a quantum mechanical regime. Our simulations clearly showed the relationship between bead spreading and classical or quantum mechanical behavior.

We found that the number of time slices, or beads on the necklace, required for quantum convergence varies depending on the system. Also, the time step for high temperatures needs to be smaller than what we used to run our simulations. However, the focus of our project was on the comparison of different method combinations, so we used the same time step in every simulation.

The best combination of methods is clearly the Langevin thermostat with the staging transformation. This

conclusion is based upon the time that it took to run the simulation and the errors in the values predicted by the simulation. The next best combination is the Langevin thermostat with the normal mode transformation. Even with our optimizations, the Nose-Hoover thermostat in combination with any of the transformations did not compare to the use of the Langevin thermostat with the staging transformation.

In the future, it would be interesting to apply the PIMD approach to an actual system using our optimal settings. Analysis of systems such as the interaction of hydrogen with a metal lattice would benefit from the use of the PIMD method, because it would incorporate quantum mechanical effects due to the small mass of the hydrogen atom. For this application, PIMD would be combined with an ab initio method such as DFT. For further comparisons, PIMD should be compared with other methods such as the path integral Monte Carlo method, looking at efficiency.

---

[1] M. Ceriotti, M. Parrinello, T. E. Markland, and D. E. Manolopoulos. Efficient stochastic thermostatting of path integral molecular dynamics. *J. Chem. Phys.* , 133(12):124104–+, September 2010.

[2] P.H. Hunenberger. *Thermostat Algorithms for Molecular Dynamics Simulations.* Springer Berlin / Heidelberg, 2005.

[3] G. J. Martyna, M. L. Klein, and M. Tuckerman. Nosé-Hoover chains: The canonical ensemble via continuous dynamics. *J. Chem. Phys.* , 97:2635–2643, August 1992.

[4] D. Marx and M. Parrinello. Ab initio path integral molecular dynamics: Basic ideas. *J. Chem. Phys.* , 104:4077–4082, March 1996.

[5] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to Quantum Field Theory.* Westview Publisher, 1995.

[6] Mark E. Tuckerman. *Statistical Mechanics: Theory and Molecular Simulation.* Oxford University Press, 2010.

[7] C. Zhang and A. Michaelides. Quantum nuclear effects on the location of hydrogen above and below the palladium (100) surface. *Surface Science*, 605:689–694, April 2011.