Advances in Quantum Monte Carlo
Advances in Quantum Monte Carlo

Shigenori Tanaka, Editor
Kobe University
Kobe, Hyogo, Japan

Stuart M. Rothstein, Editor
Brock University
St. Catharines, Ontario, Canada

William A. Lester, Jr., Editor
University of California, Berkeley
Berkeley, California, United States

Sponsored by the
ACS Division of Physical Chemistry

American Chemical Society, Washington, DC
Distributed in print by Oxford University Press, Inc.
Foreword

The ACS Symposium Series was first published in 1974 to provide a mechanism for publishing symposia quickly in book form. The purpose of the series is to publish timely, comprehensive books developed from the ACS sponsored symposia based on current scientific research. Occasionally, books are developed from symposia sponsored by other organizations when the topic is of keen interest to the chemistry audience.

Before agreeing to publish a book, the proposed table of contents is reviewed for appropriate and comprehensive coverage and for interest to the audience. Some papers may be excluded to better focus the book; others may be added to provide comprehensiveness. When appropriate, overview or introductory chapters are added. Drafts of chapters are peer-reviewed prior to final acceptance or rejection, and manuscripts are prepared in camera-ready format.

As a rule, only original research papers and original review papers are included in the volumes. Verbatim reproductions of previous published papers are not accepted.

ACS Books Department
Preface

The chapters in this monograph are contributions from the Advances in Quantum Monte Carlo symposium held at Pacifichem 2010, International Chemical Congress of Pacific Basin Societies. The symposium was dedicated to celebrate the career of James B. Anderson, a notable researcher in the field.

Quantum Monte Carlo provides an ab initio solution to the Schroedinger equation by performing a random walk through configuration space in imaginary time. Benchmark calculations suggest that its most commonly-used variant, “fixed-node” diffusion Monte Carlo, estimates energies with an accuracy comparable to that of high-level coupled-cluster calculations. These two methods, each having advantages and disadvantages, are complementary “gold-standards” of quantum chemistry.

There are challenges facing researchers in the field, several of which are addressed in the chapters in this monograph. These include improving the accuracy and precision of quantum Monte Carlo calculations; understanding the exchange nodes and utilizing the simulated electron distribution; extending the method to large and/or experimentally-challenging systems; and developing hybrid molecular mechanics/dynamics and Monte Carlo algorithms.

We extend special thanks to the Physical Chemistry Division of the American Chemical Society and Silicon Graphics, Inc. that sponsored the symposium and provided partial financial support to offset travel expenses of some of its participants.

Shigenori Tanaka
Kobe, Hyogo, Japan

Stuart M. Rothstein
St. Catharines, Ontario, Canada

William A. Lester, Jr.
Berkeley, California, United States
Editors’ Biographies

Shigenori Tanaka

Shigenori Tanaka received his B.S. degree in physics from University of Tokyo in 1982 and his Ph.D. degree in physics from University of Tokyo (S. Ichimaru) in 1986. After a Young Research Fellow of Japan Society for the Promotion of Science and serving two years as a Research Associate at University of Tokyo, he joined in 1989 Toshiba Research and Development Center. As a Visiting Associate, Tanaka also did research at the California Institute of Technology (R.A. Marcus) from 1995–1996. Since 2004, he is a professor at Kobe University.

Stuart M. Rothstein

Stuart M. Rothstein received his B.S. degree in chemistry from the University of Illinois in 1964 and his Ph.D. degree in chemistry from the University of Michigan (S.M. Blinder) in 1968. After a one year post-doc at Johns Hopkins University (H.J. Silverstone) and serving one year as an Assistant Professor at Swarthmore College, in 1970 he joined the Department of Chemistry at Brock University. While on sabbatical leaves, Rothstein has done research at Uppsala University (P.O. Löwdin), the University of Southern California (H.S. Taylor), the University of California at Berkeley (W.A. Lester, Jr.), and Kobe University (S. Tanaka).

William A. Lester, Jr.

William A. Lester, Jr. was awarded a B.S. degree in 1958 and a M.S. degree in 1959 in chemistry from the University of Chicago. He received his Ph.D. in chemistry in 1964 from The Catholic University of America while employed at the National Bureau of Standards. He has held appointments at the Theoretical Chemistry Institute, University of Wisconsin, the Department of Chemistry, UW Madison, and the IBM San Jose Research Laboratory. He was Director of the U.S. National Resource for Computation in Chemistry and Professor of Chemistry at the University of California, Berkeley. He is presently Professor of the Graduate School, Department of Chemistry, University of California, Berkeley.
Correlated Sampling for Energy Differences in Diffusion Quantum Monte Carlo

James B. Anderson*

Department of Chemistry, Pennsylvania State University, University Park, Pennsylvania 16802
*E-mail: jba@psu.edu

We describe a method using correlated sampling in fixed-node diffusion QMC calculations to determine differences in energies for similar molecular systems. The method is based on an earlier scheme for calculating corrections to trial wave functions applied to two or more systems with correlated calculations in fixed-node QMC. The result is a significant reduction in the statistical error in comparing the energies of similar molecular structures.

Introduction

The statistical error inherent in Quantum Monte Carlo calculations leads to difficulties in calculating differences in energies for similar systems. Analytic variational calculations, for example, are free of this problem. In variational QMC the statistical error in energy differences may be reduced with use of correlated sampling, but correlated sampling in diffusion QMC has been problematical. We report here a new method which is successful for correlated sampling in diffusion QMC and produces accurate energy differences for similar systems at the diffusion QMC level.

We have described previously several versions of a quantum Monte Carlo method for the direct calculation of corrections to trial wave functions (1, 2). This method gives - for fixed nodes - the difference δ between a true wave function Ψ and a trial wave function Ψ0, as well as the difference between the true energy E and the expectation value of the energy E_var for the trial wave function. The statistical or sampling errors associated with the Monte Carlo procedures as well as any systematic errors occur only in the corrections. Thus very accurate wave
functions and energies may be corrected with very simple calculations. The nodes are unchanged and the wave functions and energies for these systems are corrected to the fixed-node values - those corresponding to the exact solutions for the fixed nodes of the trial wave functions.

We have also described previously some correlated diffusion QMC calculations for similar structures of the H$_3^+$ ion (3). These took advantage of scaled length and energy parameters, but were only mildly successful because the initial correlation of electron positions was diminished as the calculations proceeded.

In the work reported here we have combined the latest version of the correction method (2) with the correlated sampling in diffusion QMC (3). We have carried out correction calculations for cases of two similar systems with correlated sampling of electron positions, moves, and multiplications wherever possible. Since the two systems are slightly different, the initial similarity in walker positions and weights is not maintained exactly, but the correction calculation is fairly short, and the systems remain nearly similar over the length of the calculation.

**Theoretical Basis**

The derivation of the latest version of our correction scheme is given in Ref. (2). We give a brief description here. The basic equation is that for QMC with importance sampling (4–7) where \( f = \Psi \Psi_0 \), the product of the true wave function and a trial wave function, and \( E_{\text{ref}} \) is a reference energy,

\[
\frac{\partial \mathcal{F}}{\partial \tau} = \frac{1}{2} \nabla^2 f - \nabla \cdot \left( \nabla \ln \Psi_0 \right) \left( \frac{H \Psi_0}{\Psi_0} - E_{\text{ref}} \right) f
\]  

Making use of the difference \( \delta = \Psi - \Psi_0 \) and defining another function \( g = \delta \Psi_0 = (\Psi - \Psi_0) \Psi_0 \) we obtain

\[
\frac{\partial g}{\partial \tau} = \frac{1}{2} \nabla^2 g - \nabla \cdot (g \nabla \ln \Psi_0) - \left( \frac{H \Psi_0}{\Psi_0} - E_{\text{ref}} \right) g + \left[ \left( \frac{H \Psi_0}{\Psi_0} - E_{\text{ref}} \right) \Psi_0^2 \right]
\]  

The term \( \frac{H \Psi_0}{\Psi_0} \) is the local energy \( E_{\text{loc}} \) for the trial wave function. The last term in Eq. (2) is a distributed source term \( S \), which may be positive or negative. It is convenient to introduce the expectation value of the energy \( E_{\text{var}} \) for the trial function and write \( S \) as a collection of terms

\[
S = S_p + S_n + S_q ,
\]

where
and where the subscript $p$ indicates a region $E_{\text{loc}} < E_{\text{var}}$ and a positive particle feed, the subscript $n$ indicates a region of $E_{\text{loc}} > E_{\text{var}}$ and a negative particle feed, and the subscript $q$ indicates an additional particle feed, normally negative.

We consider the integrals of each of the terms in Eq. (2) over all accessible space. Thus we define the integrals

$$I_g = \int g dV, \quad I_p = \int S_p dV, \quad I_n = \int S_n dV, \quad I_q = \int \Psi_0^2 dV,$$

and Eq. (2) becomes

$$\frac{\partial I_g}{\partial \tau} = \int \frac{1}{2} \nabla^2 g dV - \int \nabla \cdot (g \nabla \ln \Psi_0) dV$$

$$- \int (E_{\text{loc}} - E_{\text{ref}}) g dV + I_p + I_n - (E_{\text{var}} - E_{\text{ref}}) I_q.$$

Since the diffusion and drift terms merely move particles within the volume and the drift term prevents their crossing a nodal surface, they make no contribution to changes in $I_g$, and they are zero and may be eliminated. The multiplication term containing $(E_{\text{loc}} - E_{\text{ref}})$ is applicable to each particle fed to the system. An average growth factor $\overline{f_p}, \overline{f_n}, \overline{f_q}$ for each type of particle during its lifetime in the system may be combined with the feed terms. Thus we obtain

$$\frac{\partial A_g}{\partial \tau} = I_p \overline{f_p} + I_n \overline{f_n} - (E_{\text{var}} - E_{\text{ref}}) I_q \overline{f_q}.$$

Since particles fed at any point in the system tend to the same distribution with increased time in the system, those of sufficient age may be cancelled in equal weights, positive with negative, regardless of their locations. For a steady state and for a complete cancellation of positive and negative particles at a fixed age we have $\partial I_g/\partial \tau = 0$, and the energy $E$ is equal to $E_{\text{ref}}$ which is given by (after rearrangement)

$$E = E_{\text{ref}} = E_{\text{var}} - \frac{I_p \overline{f_p} + I_n \overline{f_n}}{I_q \overline{f_q}}.$$
To determine the energy one then needs to obtain only the ratios of the integrals \( I_p/I_q \) and \( I_n/I_q \) rather than the individual values. Also needed are the values of \( \overline{f_p}, \overline{f_n} \) and \( \overline{f_q} \) as well as \( E_{\text{var}} \).

The ratios of the integrals may be determined by numerical integration, typically by Metropolis sampling of \( \Psi_0 \) with accumulation of average values of \( I_p/I_q \) and \( I_n/I_q \). Their definition in terms of \( E_{\text{var}} \) ensures that the absolute values of these ratios be equal and uncertainty in their relative values is eliminated. Since an accurate value of \( E_{\text{var}} \) is required, an analytic integration to determine \( E_{\text{var}} \) is preferred. The values of \( \overline{f_p}, \overline{f_n} \) and \( \overline{f_q} \) are determined for sample feed particles obtained in the Metropolis integrations. These are selected with probabilities proportional to (in addition to \( \Psi_0 \)) the absolute values of the quantities \( -(E_{\text{loc}} - E_{\text{var}}) \) for \( p \), \( -(E_{\text{loc}} - E_{\text{var}}) \) for \( n \), and unity for \( q \). The particles are subjected to diffusion, drift, and multiplication (weight increase or decrease) for a period (age) sufficiently long to produce no further change in their average weights. Positive particles of type \( p \), fed in regions of \( E_{\text{loc}} < E_{\text{var}} \), give values of \( \overline{f_p} > 1 \). Negative particles of type \( n \), fed in regions of \( E_{\text{loc}} < E_{\text{var}} \), give values of \( \overline{f_n} < 1 \). Particles of type \( q \) normally give a value \( \overline{f_q} \) very close to unity.

The energy \( E \) is given by the known value of \( E_{\text{var}} \) and a correction term. When \( \Psi_0 \) is a good approximation to the true wave function, the correction term is small and any error in the correction term is correspondingly small. As \( \Psi_0 \) approaches the true wave function the ratios \( I_p/I_q \) and \( I_n/I_q \) approach zero and the values of \( \overline{f_p}, \overline{f_n} \) and \( \overline{f_q} \) approach unity.

**Introducing Correlation**

To calculate the difference in diffusion QMC energies for two systems A and B the variational energies \( E_{\text{varA}} \) and \( E_{\text{varB}} \) for the two systems and their difference are determined in correlated variational QMC calculations using Metropolis sampling. Electron positions may be identical for the two systems or scaled in one way or another to maintain similarity. The sampling may be carried out with a value (or weight) function \( W \) given by \( W_A + W_B \), the sum of squares of the two trial wave functions. In this case the energy for each system is given by the weighted average of its local energies using weights proportional to \( W_A \) or \( W_B \). If the two systems are nearly the same the statistical error in the difference \( E_{\text{varA}} - E_{\text{varB}} \) will be much smaller than the statistical error in the individual values \( E_{\text{varA}} \) and \( E_{\text{varB}} \). The values of \( I_p, I_n, \) and \( I_q \) as well as sample configurations of type \( p, n, \) and \( q \) are retained.

The next step is a correction calculation for each of the trial wave functions carried out as described in the section above using the correlated type \( p, n, \) and \( q \) samples to determine values of \( \overline{f_p}, \overline{f_n} \) and \( \overline{f_q} \) for each of the systems. For these correlated diffusion QMC calculations the initial electron positions correspond but the initial weights as found in the variational calculations may vary for the two
systems. The drift terms and the multiplication terms may also vary, but if the systems are similar the differences will be small. The diffusion terms and the points of multiplication are matched for the two systems, so that electron movements and multiplications are correlated as much as possible. Since the initial configurations \(p\) and \(q\) are deviations from typical configurations, their local energies also deviate, and the changes in weights and contributions to the values of \(\bar{f}_p\) and \(\bar{f}_n\) are large at the start of a run. To avoid time-step error the time-step size must be very small at the start. As the run proceeds the walkers approach normal behavior and the values of \(\bar{f}_p\) and \(\bar{f}_n\) approach their asymptotic values.

A Simple Test: Neon vs. Neon (moved slightly)

To test the method we calculated the difference in diffusion QMC energies for two systems A and B which are known to have the same energy: a neon atom in each case (non-interacting) with the first located at the origin of coordinates and the second displaced a short distance from the origin. The displacement was 0.02 bohr. Each had a trial function consisting of a fairly good SCF function and a Schmidt-Moscowitz (\(\delta\)) Jastrow function centered on its nucleus. The two trial functions were identical except for positioning. In the variational calculation the sampling value for walker moves was taken as the sum of the squares of the trial functions. The results are shown in Table 1.

<table>
<thead>
<tr>
<th>Table 1. Correlated DQMC: Ne vs. Ne shifted 0.02 bohr</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Energies (hartrees)</strong></td>
</tr>
<tr>
<td>Ne #1</td>
</tr>
<tr>
<td>(E_{\text{var}})</td>
</tr>
<tr>
<td>(\pm 0.000 090)</td>
</tr>
<tr>
<td>(E_{\text{cor}})</td>
</tr>
<tr>
<td>(\pm 0.001 100)</td>
</tr>
<tr>
<td>(E_{\text{cor}}) difference</td>
</tr>
<tr>
<td>(\pm 0.000 077)</td>
</tr>
<tr>
<td>(E_{\text{exact}})</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

\(^{a}\) Reference (9).

It may be seen that the uncertainty in the difference for variational energies is 16 microhartrees and that for diffusion energies is 77 microhartrees, while the corresponding uncertainties in the individual values are about 90 and 1100 microhartrees, respectively. The 77-microhartree error in the calculated difference in diffusion energies is about the same as the indicated statistical error in that quantity. This is the desired result. That is, the known difference in
DQMC energies (zero) is reproduced (as 0.000 115 ± 0.000 077 hartrees) despite uncertainties of 0.001 100 hartrees in the individual values.

Comparison of Two H₂ Molecules

A second example, for which the energy difference is known nearly exactly, is the case of two ground-state H₂ molecules with slightly different internuclear distances, the first with \( R = 1.4000 \) bohr and the other with \( R = 1.4011 \) bohr. In this case the wave function is nodeless. The trial function was a high quality single-determinant function with an optimized Jastrow term. The time-step and distance parameters for the diffusion calculations were scaled to place electrons in geometrically similar positions – in effect by placing the nuclei in identical positions in the scaled systems. The results are shown, together with those of accurate analytic variational calculations by Cencek and Szalewicz (10) for comparison, in Table 2.

The table shows the desired effect: the uncertainty in the calculated difference in the DQMC energies (0.040 microhartrees) is much lower than the uncertainties (4.800 and 5.300 microhartrees) in the DQMC energies of the calculations considered separately.

### Table 2. Correlated DQMC: H₂ with \( R = 1.4000 \) bohr vs. H₂ with \( R = 1.4011 \) bohr

<table>
<thead>
<tr>
<th></th>
<th>( R = 1.4000 )</th>
<th>( R = 1.4011 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( E_{\text{var}} ) (VQMC)</td>
<td>–1.173 434 517</td>
<td>–1.173 433 737</td>
</tr>
<tr>
<td>( E_{\text{cor}} ) (DQMC)</td>
<td>–1.174 470 644</td>
<td>–1.174 470 901</td>
</tr>
<tr>
<td>( E_{\text{cor}} ) difference</td>
<td>±0.000 004 800</td>
<td>±0.000 005 300</td>
</tr>
<tr>
<td>( E_{\text{exact}} ) difference(^a)</td>
<td>–0.000 000 257</td>
<td>±0.000 000 040</td>
</tr>
<tr>
<td>( E_{\text{exact}} ) (^a)</td>
<td>–1.174 475 714</td>
<td>–1.174 475 931</td>
</tr>
</tbody>
</table>

\(^a\) Reference (10).

The table shows the desired effect: the uncertainty in the calculated difference in the DQMC energies (0.040 microhartrees) is much lower than the uncertainties (4.800 and 5.300 microhartrees) in the DQMC energies of the calculations considered separately.
Comparison of Two C\textsubscript{10} Rings

A third example, for which the energy difference is not known exactly, is the case of two C\textsubscript{10} rings of slightly different radii. The ten carbon atoms were placed evenly on circles, one with a radius of 2.04 bohr, the other with a radius of 2.08 bohr. The circles were centered on the origin, and the length and time scales of the second were adjusted to obtain geometric and dynamic similarity. The trial wave function was a single determinant function with a Jastrow term, the same for both systems. The effect of geometric scaling was to give different values for the wave functions and local energies for corresponding electron configurations. The results are shown in Table 3.

<table>
<thead>
<tr>
<th>Energies (hartrees)</th>
<th>$R = 2.04$</th>
<th>$R = 2.08$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_{\text{var}}$</td>
<td>$-379.4307$</td>
<td>$-379.4319$</td>
</tr>
<tr>
<td>(VQMC)</td>
<td>$\pm 0.0067$</td>
<td>$\pm 0.0053$</td>
</tr>
<tr>
<td>$E_{\text{cor}}$</td>
<td>$-379.4485$</td>
<td>$-379.4500$</td>
</tr>
<tr>
<td>(DQMC)</td>
<td>$\pm 0.0130$</td>
<td>$\pm 0.0130$</td>
</tr>
<tr>
<td>$E_{\text{cor}}$ difference</td>
<td>$+ 0.0015$</td>
<td>$\pm 0.0013$</td>
</tr>
</tbody>
</table>

Again the results show significantly lower statistical errors for the differences in variational QMC energies and the differences in diffusion QMC energies than for the individual values of these energies. The uncertainty in the diffusion QMC difference is approximately equal to the indicated difference as expected for a minimum in the energy. Additional calculations for slightly different geometries indicate a minimum at a radius of about 2.08 bohr. Independent analytical calculations (MP2, with substantially different trial functions) indicate a minimum in the range of 2.08 to 2.12 bohr.

The Neon Dimer

The neon dimer represents a special challenge to quantum mechanics in general, and this challenge might be met by correlated diffusion QMC calculations. One should expect significant cancellations in calculations for separated neon atoms (Ne + Ne) vs. the Ne dimer (Ne-Ne) at or near its equilibrium separation distance of about 5.8 bohr. The trial function used was the product of two single atom functions (as used for the Ne atom above) with additional cross terms in the Jastrow function. The results are shown in Table 4.
Table 4. Correlated DQMC: (Ne + Ne) vs. Ne₂ (\( R = 5.8 \) bohr)

<table>
<thead>
<tr>
<th></th>
<th>( Ne + Ne )</th>
<th>( Ne₂ )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( E_{\text{var}} ) (VQMC)</td>
<td>( -257.538 ) 127 ( \pm 0.000 ) 200</td>
<td>( -257.538 ) 261 ( \pm 0.000 ) 200</td>
</tr>
<tr>
<td>( E_{\text{corr}} ) (DQMC)</td>
<td>( -257.845 ) 542 ( \pm 0.000 ) 500</td>
<td>( -257.846 ) 338 ( \pm 0.000 ) 400</td>
</tr>
<tr>
<td>( E_{\text{corr}} ) difference</td>
<td>( -0.000 ) 796 ( \pm 0.000 ) 700</td>
<td></td>
</tr>
</tbody>
</table>

In this case we were successful in demonstrating the effectiveness of correlated variational QMC and obtained a well depth of 0.000134 +/- 0.000009 hartrees, in good agreement with accepted values derived from experiment (11). But, for the correlated diffusion QMC the indicated well depth was 0.000796 +/- 0.000700 hartrees. With this large value of the statistical error, we conclude that a much better trial function is required for this difficult case. We expect to pursue this case further.

Discussion

It is clear that correlated sampling provides significant improvements in the determination of differences in diffusion QMC energies of similar molecular systems. Of course, the accuracy in the differences in calculated energies depends strongly on the degree of similarity of the systems compared. For closely similar systems the difference method can provide (loosely) the derivative in energy with respect to a variable describing the difference in the systems.

We note that the method of correlated sampling in diffusion QMC described here can be applied, in a similar fashion, to the optimization of a trial wave functions. In this case, we compute the difference in diffusion QMC energies for the same system and two similar wave functions. Since the difference in energies depends only on the node locations, the node structure may be optimized, thereby providing a solution to the so-called ‘sign problem of quantum Monte Carlo’. We are encouraged by our early investigations.

Acknowledgments

Support by the National Science Foundation (IGERT Award No. 9987589) and by the Research Computing and Cyber Infrastructure (a unit of ITF) of the Pennsylvania State University is gratefully acknowledged.
References

Chapter 2

Population Control Bias with Applications to Parallel Diffusion Monte Carlo

Jaron T. Krogel*1,2 and David M. Ceperley1,2

1Department of Physics, University of Illinois, Urbana-Champaign, Urbana, Illinois 61801
2National Center for Supercomputing Applications (NCSA), University of Illinois, Urbana-Champaign, Urbana, Illinois 61801
*E-mail: jkrogel2@illinois.edu

The future of scientific computing will be driven by highly distributed parallel machines with millions of compute nodes. In order to take advantage of this already arriving wave of computing capability we must identify and remove the remaining barriers to parallel scaling in the Diffusion Monte Carlo algorithm. To address these scaling issues in a simple way, we propose that a time delay be introduced into the population control feedback. In order to assess this algorithm, we investigate the behavior of population fluctuations and the population control bias (which will emerge into greater relevance with larger physical systems and requirements of higher accuracy) in a model system for both the standard and time delayed DMC algorithms. We then condense our findings into a simple set of recommendations to improve the scaling of DMC while managing the population control bias.

The Future of Parallel DMC

Modern parallel machines are composed of tens of thousands of SMP nodes which will rapidly increase into the millions in coming years. Keeping pace with such developments is crucial to the success of scientific computing. Monte Carlo methods are particularly well suited to make use of this ever increasing computing power since they contain low serial dependency. In Markov Chain Monte Carlo, a function is evaluated along a large chain of randomly generated
particle configurations to obtain its mean. Conceptually, we think of the chain as a random walk, meaning that it is formed by a walker which steps from configuration to configuration. Typically, straightforward replication of Monte Carlo random walks allows near perfect scaling, however a central feature of any Monte Carlo algorithm will limit the practical efficiency gains realized when scaling to millions of processors: the equilibration time of the random walks.

The bottleneck due to equilibration time arises from the pursuit to simulate ever larger systems. Since the time to obtain new configurations is increased, the equilibration time consumes a larger fraction of the fixed serial run time. A direct way to minimize the ratio of serial equilibration time and total run time is to minimize the number of walkers per SMP node. For many Monte Carlo algorithms it is possible to achieve the exact minimum of just one walker per node; however this is not the case for the standard Diffusion Monte Carlo (DMC) method.

DMC is a Monte Carlo procedure to solve the many-body time independent Schrödinger equation (1, 2). It is currently the most widely used Quantum Monte Carlo algorithm for chemical and condensed matter systems. In the standard algorithm, the population of walkers fluctuates through a birth-death process known as branching and so load balancing challenges arise in the limit of few walkers per node. For example, a set of nodes each running with a single walker will quickly evolve into a situation where nodes contain zero, one, or two or more walkers each. In this circumstance, the overall efficiency can fall since all walkers must complete a step before the trial energy is collected and branching is performed. In typical implementations, current codes use on the approximately one hundred walkers per node to avoid load balancing inefficiency, however this leads to equilibration inefficiency as mentioned above. Though load balancing operations cannot perfectly redistribute the load for arbitrary population sizes, they will have to be performed frequently to prevent the situation from becoming even worse. Since most implementations of load balancing rely on operations within the Message Passing Interface (which impose synchronization at some level), parallel efficiency will further suffer with increasing machine size. At the same time, the branching process must be tightly regulated to prevent chronic shortages of walkers which could leave large numbers of nodes idle.

Demanding smaller fluctuations in the DMC walker population (greater stability) increases a systematic error within DMC as we detail below. However, with the advent of larger machines, a greater level of accuracy will inevitably be pursued, requiring parallel algorithms with greater robustness and less bias. Thus the population control bias, which has largely been neglected in the past, could become increasingly relevant to future applications.

In this work, we explore the relationship between the population control parameters and the stability and accuracy of the standard DMC method. Using a simple and efficient model problem, the simple harmonic oscillator, we obtain empirical formulas describing the behavior of population stability and the population control bias over a range of control parameters. Since we expect population dynamics (and hence stability and bias) to depend more strongly on the control method used than the particular details of the physical system, these results should apply broadly to current DMC practice.
Looking ahead to the rapidly approaching parallel efficiency difficulties outlined above, we move on to a possible solution to the scaling problem. We propose the introduction of a feedback delay in the population control mechanism which removes synchronization requirements from the algorithm and allows load balancing techniques that capitalize on keeping the time averaged load constant rather than the instantaneous load. These features ease the tradeoff between serial and parallel efficiency. As we must be cautious not to introduce unmanageable instability or bias into the simulation, we repeat the model analysis on cases with delayed feedback. These results are then consolidated into a brief set of practical recommendations to improve the parallel efficiency and reduce the population control bias in typical DMC simulations.

Other types of Quantum Monte Carlo, such as Reptation (3) and Path Integral Monte Carlo (4), do not suffer from population control bias. However, those methods have not been systematically applied to systems of many electrons, and in particular, there have been no studies of their relative efficiency, e.g. how their errors scale with computer time and number of electrons. In addition, these methods have other difficulties in scaling to very large number of processors. For these reasons, we limit the discussion to Diffusion Monte Carlo.

Review of the DMC Method

The Diffusion Monte Carlo algorithm (1, 2) can be viewed as a technique to iteratively refine the standard variational method (5). In the variational method, an ansatz for the many-body wavefunction ($\psi_T$) is optimized to be as close as possible to the true ground state, usually by minimizing a combination of its energy moments, $\mu_n = \langle \psi_T | H^n | \psi_T \rangle$. The Hamiltonian operator, $\hat{H}$, is given by

$$\hat{H} = -\frac{\hbar^2}{2m} \nabla^2 + V$$

(1)

The variational energy obtained by integrating the local energy ($E_L = \psi_T^* \hat{H} \psi_T$) over the trial probability density ($f_0 = |\psi_T|^2$) is a strict upper bound on the ground state energy of the system.

$$E_V = \frac{\int dR f_0(R)E_L(R)}{\int dR f_0(R)} \geq E_0$$

(2)

In DMC, the variational probability density is evolved forward in imaginary time through the iterative application of a short time Green’s function.

$$f_{t+\tau}(R') = \int dR G(R',R;\tau)f_t(R)$$

(3)