Developments of stochastic perturbation theory and perturbative corrections of model space quantum Monte Carlo

Ladoczki, Bence

(Degree)
博士（計算科学）

(Date of Degree)
2020-03-25

(Date of Publication)
2021-03-01

(Resource Type)
doctoral thesis

(Report Number)
甲第7788号

(URL)
https://hdl.handle.net/20.500.14094/D1007788

Doctoral Thesis

Developments of stochastic perturbation theory and perturbative corrections of model space quantum Monte Carlo

統計的摂動論の開発とモデル空間量子モンテカルロ法における摂動補正に関する研究

Bence Ladoczki

ラドーツキ　ベンツェ

Graduate School of System Informatics

Kobe University

A thesis submitted for PhD. degree

Jan. 2020

Contents

Notations

Notation	Purpose	Explanation
N_b	Booster weights	The number of integer walkers placed on the reference determinant
N_i	Initiator threshold	Threshold of the initiator adaptation
N_w^i	Number of walkers on determinant i	Walker population of determinant i
N_m	Number of steps for accumulation	Number of micro steps for which the perturbative corrections are accumulated
τ	Calculation length in a.u.	The imaginary time for which the calculation runs
τ_1	Integration start step in a.u.	Imaginary time from which the integration commences
τ_2	Integration termination step in a.u.	Imaginary time at which the integration terminates
$\Delta\tau$	Time step in a.u.	Imaginary time for discrete sampling
σ	Standard deviation	Standard deviation of a data set
σ_e	Standard error	Standard error of a data set
Ψ_0	Zeroth order wave function	Zeroth order wave function in perturbation theory
H_0	Zeroth order Hamiltonian	Zeroth order Hamiltonian in perturbation theory
$\hat{V}$	Perturbation operator	Perturbation operator in perturbation theory
$\hat{\Omega}$	Wave operator	Wave operator in perturbation theory
χ	Residual wave function	Residual wave function in the new sampling method
E_{var}	Variational energy	Variational energy
$E^{(2)}$	Second order correction	Second order Epstein-Nesbet correction
$E^{(3)}$	Third order correction	Third order Epstein-Nesbet correction

1 General Introduction

> Science has explained nothing; the more we know the more fantastic the world becomes and the profounder the surrounding darkness.
>
> *Aldous Huxley*

Comprehending the characteristics and peculiarities of our nature is a fundamental desire of human beings. The exploration of an object called the "wave function" can bring us closer to the understanding of certain phenomenas in the quantum world. A wave function in quantum physics is a mathematical description of the quantum state of an isolated quantum system.

For a quantum system having more than a few electrons the many-body Hamiltonian is very complex, and the only feasible way for its treatment is to use some kind of approximation scheme. A natural approximation is to assume that each electron moves in an average potential due to the other electrons.

In computational physics and chemistry, the Hartree-Fock (HF) method is a method of approximation for the determination of the wave function and the energy of a quantum many-body system in a stationary state. The HF approximation is essentially a simplification of the electron-electron repulsion term by introducing the Fock operator to simplify the many-electron problem with the use of the Hartree-Fock potential (or the "field" seen by the $i$the electron). The Fock operator depends on its eigenfunctions therefore the equations are solved iteratively starting from an initial guess. When the "field" does not change anymore, consistency (hence the name self-consistent-field (SCF)) is reached and one obtains the HF wave functions that does not incorporate electronic correlation.

Dynamic correlation is understood as the instantaneous correlation in electron motions due to their mutual repulsion. On the other hand, nondynamic correlation reflects the effect of near degeneracies and other imperfections of the single-configuration model. Neglect of electron correlation can lead to large deviations from experimental results. A number of approaches to this inadequacy, collectively called post-Hartree-Fock methods, have been devised to include electron correlation to the multi-electron wave function. Among these density matrix renormalization group (DMRG)[24, 25, 26, 27, 28, 29, 30, 31], restricted active space (RAS)[32, 33], generalized active space[34] and SplitGas[35] are actively developed to treat complex electronic structures.

The following thesis is centered around one class of post-Hartree-Fock methods, the quantum Monte Carlo (QMC) methods. The large number of recent publications[1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 13, 35, 16, 17, 18, 19, 20, 21, 22, 23, 10, 36, 37] dealing either with applications or further development of some aspects of this method is a clear indication that the scientific community has realized the power and versatility of Monte Carlo Simulations.

The name Monte Carlo arises from the fact that this method uses random numbers similar to those coming out from roulette games. But producing up to 10^{10} such numbers in Monte Carlo in this way would consume a lot of money and time; thus computers are used instead. In a Monte Carlo simulation usually one tries to follow the "time-propagation" of a model for which a change, or growth does not proceed in some rigorously predefined fashion but rather in a stochastic manner, which depends on a on a sequence of random numbers which is generated during the simulation.

The main goal of these researches is to solve the Schrödinger equation to recover the correlation energy of a quantum many-body system. The correlation energy is defined as the difference between the exact nonrelativistic energy and the HF energy obtained in the limit that the basis set approaches completeness.

In a given basis set, the exact wave function along with the correlation energy can be obtained by constructing the Full Configuration Interaction (FCI) matrix and calculating its eigenvectors. The problem with this approach is that one has to diagonalize a matrix in a space whose size grows very rapidly with the system size under consideration. The FCI wave function is basically expanded in a set of Slater determinants. The FCI wave function expanded on the basis of n-tuply excited Slater determinants has the following form:

$$|\Psi\rangle_{\mathrm{FCI}} = c_0|\Phi_0\rangle + \sum_{a,r} c_a^r|\Phi_a^r\rangle + \sum_{r<a,b<s} c_{ab}^{rs}|\Phi_{ab}^{rs}\rangle + \sum_{r<a,b<s,c<t} c_{abc}^{rst}|\Phi_{abc}^{rst}\rangle + ... \tag{1.1}$$

If there are N occupied orbitals out of the 2K spin orbitals, the number of n-tuply excited determinants can be expressed using the following formula:

$$\mathrm{N}_{\text{n-tuple}} = \binom{N}{n}\binom{2K-N}{n} \tag{1.2}$$

Clearly, incorporating all of these excitations in the FCI wave function leads to extreme memory requirements. Although there are efficient algorithms to diagonalize large matrices the applicability of the FCI method is limited to a few dozens of correlated orbitals. This is mainly because even modern supercomputers have limited memory resources and these matrix diagonalization algorithms require at least two vectors whose size is proportional to the Hilbert space size to be stored in memory. As a result, memory will always be a limiting resource for traditional CI matrix diagonalization based approaches. QMC methods address this problem with a new approach, namely that instead of storing large vectors they sample the desired quantities using a random walk in the space of determinants.

One of the most important QMC methods that has become a significant tool in electronic structure theory for obtaining high-accuracy properties of challenging systems is the Full Configuration Interaction Quantum Monte Carlo (FCIQMC) method[25].

The FCIQMC algorithm samples the FCI wave function using walker dynamics that will be introduced in the following sections. Starting from an initial state (usually chosen to be the HF solution) the final wave function is obtained by simulating the imaginary time evolution equation. If multiple states are chosen for the initial wave function one has to diagonalize in this space (model space) to obtain the initial state vectors.

Model Space Quantum Monte Carlo (MSQMC)[29, 30] was introduced by Ten-no in 2014 and this method has been further developed in the following years[41]. In MSQMC the sampled determinantal space is partitioned into a small space (P-space) to be treated deterministically and the rest of the space is treated stochastically (Q-space). Following the determination of the transfer matrix it is possible to calculate the effective Hamiltonian matrix of the system and obtain its eigenvalue for the correlation energy.

The main challenge both for FCIQMC and MSQMC is that the number of walkers required to represent the sampled wave function explodes as one tries to increase the size of the active space. In order to reduce computation costs initiator approximation was introduced[42] to truncate the sampled space and eliminate determinants with very small walker populations. The initiator approximation differentiates between initiator determinants (large walker population) and non-initiator determinants (small walker population). Initiator determinants are allowed to spawn freely just as in the original algorithm and non-initiator determinants are only allowed to spawn onto occupied determinants. The effect of this truncation imposed by these rules vanishes in the limit of infinite walkers.

Although the same rules can be applied in MSQMC, in this thesis I present a different approach to truncated the sampled Hilbert space. In the first chapter, a formulation for stochastic perturbation theory is given. Furthermore, I investigate the size-consistency of the energy in truncated excitation spaces. The CI expansion of the wave function exhibits size-inconsistency errors if higher excitations are neglected. In this thesis, the effects of such truncations are investigated both theoretically and numerically.

The size-inconsistency error arises due to terms with erroneous scaling behaviour. These terms cancel in the FCI space and the cancellation does not occur in truncated CI spaces. In the second part of this chapter I present the required modification in the walker dynamics that will lead to size-consistent energies in every order of perturbation. Based on the perturbative expansion of the wave function I propose a semi-stochastic adaptation of the original method that obtains the deterministic first order wave function from the references wave function to sample the higher order terms with reduced stochastic noise. Furthermore in this part, I show that the first order wave function can be subtracted from the total wave function and the working equation for the residual space is presented in the current thesis as well. In many cases, especially for weakly correlated systems, the second order perturbation energy recovers $\sim$80% of the correlation energy therefore the first order wave function encodes a large proportion of the effects the Hartree-Fock approximation is incapable of describing.

In our calculations the single excitations do not interact with the reference due to the Brilloiun condition and the first order wave function only consists of double excitations. Nevertheless the size of the interacting space can potentially become prohibitive as the system size increases and circumspection is required when using such a semi-stochastic algorithm. No research has been carried out on strongly correlated systems to examine the accuracy of second order perturbation energy, but I expect that the recovery for such systems is less satisfactory.

The second part of the current work is centered around Epstein-Nesbet type corrections of the previously mentioned initiator approximation.

The idea behind this approach originates from selected CI (SCI) approaches[45]. In SCI the wave function is calculated as a result of a two-stage process. First, in the variational stage determinants are generated and the zeroth-order energy is obtained from the diagonalization in this subspace. In the second stage this energy is corrected using perturbation theory.

The wave function, when sampled using the initiator approximation is similar to the variational wave function in SCI, therefore it is possible to correct the initiator error just as in SCI. The second order correction for the initiator error has been introduced by N. S. Blunt in 2018[43] and it has been successfully applied to the homogeneous electron gas and other molecular systems. In this thesis, I formulate the third order Epstein-Nesbet correction of the initiator error and carry out detailed investigations on its accuracy.

As it was been pointed out in N. S. Blunt's paper, the stochastic noise of these corrections is significantly larger than that of the energy. In order to obtain unbiased corrections one has to use two independent walker populations and accumulate the contributions only when removals on both replicas occur. The difficulty of decreasing the standard deviation arises from the fact that as the active space grows in size so does the size of the secondary space and these removal events become rarer and rarer. The problem of sever noise in the corrections was tackled by the introduction of preconditioners and multiple spawning events previously[24].

In this thesis another approach is presented to improve the efficiency of the sampling by collecting samples of several subsequent steps and calculating the corrections from more terms. The mathematical formula for the reduction in stochastic noise is also given in this chapter. Finally the results of several applied calculations are presented to demonstrate the applicability of the initiator adaptation supplemented with second and third order Epstein-Nesbet corrections.

> "In general, the result of a measurement is an approximation of the value of the specific quantity subject to measurement, that is the measurand, and thus the result is complete only when accompanied by a quantitative statement of its uncertainty." – M. P. Nightingale - Basics, Quantum Monte Carlo and Statistical Mechanics.

In the numerical calculations the stochastic noise is demonstrated using the following metrics:

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2} \tag{1.3}$$

$$\sigma_e = \frac{1}{\sqrt{N}}\sigma \tag{1.4}$$

Where σ is the standard deviation and σ_e is the standard error (error bar) of the data set containing N data points and $\bar{x}$ is the average. The main difference between σ and σ_e is that while it is possible to reduce σ_e, σ is stays on the same value when increasing the number of calculations and σ_e approaches zero as the number of calculations tends to infinity. If the distribution is well-behaved and the number of samples is sufficiently large, the Central Limit Theorem can be applied and the probability of being wrong by 1 error bar is 35% and the probability of being wrong by 2 error bars is 5%.

Bibliography

[1] W. M. C. Foulkes, L. Mitas, R. J. Needs, G. Rajagopal Rev. Mod. Phys. **73**, 33 (2001).

[2] Y. Ohtsuka, S. Nagase Chem. Phys. Letters **485**, 367-370 (2010).

[3] N. S. Blunt, S. D. Smart, G. H. Booth, A. Alavi J. Chem. Phys. **143**, 134117 (2015).

[4] R. E. Thomas, Q. Sun, A. Alavi, G. H. Booth J. Chem. Theory Comput. **11**, 5316-5325 (2015).

[5] G. H. Booth, D. Cleland, A. Alavi, D. P. Tew J. Chem. Phys. **137**, 164112 (2012).

[6] S. Y. Willow, S. Hirata J. Chem. Phys. **140**, 024111 (2014).

[7] N. S. Blunt, T. W. Rogers, J. S. Spencer, W. M. C. Foulkes Phys. Rev. B **89**, 245124 (2014).

[8] N. S. Blunt, A. Alavi, G. H. Booth Phys. Rev. Lett. **115**, 050603 (2015).

[9] T. Zhang, F. A. Evangelista J. Chem. Theory Comput. **12**, 4326?4337 (2016).

[10] N. S. Blunt, E. Neuscamman J. Chem. Theory Comput. **15**, 178?189 (2019).

[11] G. H. Booth, A. Grüneis, G. Kresse, A. Alavi Nature **493**, 365?370 (2013).

[12] G. H. Booth, S. D. Smart A. Alavi Molecular Physics **112:14**, 1855-1869 (2014).

[13] G. Li Manni, S. D. Smart, A. Alavi J. Chem. Theory Comput. **12**, 1245-1258 (2016).

[14] A. A. Holmes, N. M. Tubman, C. J. Umrigar J. Chem. Theory Comput. **12**, 3674 (2016).

[15] S. Sharma, A. Holmes, G. Jeanmairet, A. Alavi, C. J. Umrigar J. Chem. Theory Comput. **13**, 1595 (2017).

[16] J. J. Shepherd, G. H. Booth, A. Alavi J. Chem. Phys. **136**, 244101 (2012).

[17] J. A. F. Kersten, G. H. Booth, A. Alavi J. Chem. Phys. **145**, 054117 (2016).

[18] G. H. Booth, G. K.-L. Chan J. Chem. Phys. **137**, 191102 (2012).

[19] G. Jeanmairet, S. Sharma, A. Alavi, J. Chem. Phys. **146**, 044107 (2017).

[20] A. J. W. Thom, Phys. Rev. Lett. **105**, 263004 (2010).

[21] J. S. Spencer, A. J. W. Thom, J. Chem. Phys. **144**, 084108 (2016).

[22] C. J. C. Scott, A. J. W. Thom, J. Chem. Phys. **147**, 124105 (2017).

[23] W. A. Vigor, J. S. Spencer, M. J. Bearpark, A. J. W. Thom J. Chem. Phys. **147**, 094110 (2016).

[24] S. R. White, Phys. Rev. B **48**, 10345 (1993).

[25] S. Wouters, W. Poelmans, P. W. Ayers, D. Van Neck Comput. Phys. Commun. **185**, 1501 (2014).

[26] S. Sharma, G. K.-L. Chan J. Chem. Phys. **136**, 124121 (2012).

[27] D. Zgid, M. Nooijen J. Chem. Phys. **128**, 014107 (2008).

[28] G. Moritz, B. A. Hess, M. Reiher J. Chem. Phys. **122**, 024107 (2008).

[29] O. Legeza, J. Sólyom, M. Reiher Phys. Rev. B **68**, 195116 (2003).

[30] Y. Kurashige, T. Yanai, J. Chem. Phys **130**, 234114 (2009).

[31] R. Olivares-Amaya, W. Hu, N. Nakatani, S. Sharma, J. Yang, G. K.-L. Chan J. Chem. Phys **142**, 034102 (2015).

[32] J. Olsen, B. O. Roos, P. Jørgensen, H. J. A. Jensen, J. Chem. Phys **89**, 2185 (1988).

[33] P. A. Malmqvist, A. Rendell, B. O. Roos J. Chem. Phys **94**, 5477 (1990).

[34] D. Ma, G. L. Manni, L. Gagliardi J. Chem. Phys **135**, 044128 (2011).

[35] G. Li Manni D. Ma F. Aquilante, J. Olsen, L. Gagliardi, J. Chem. Theory Comput. **9**, 3375 (2013).

[36] J. E. Deustua , I. Magoulas , J. Shen, P. Piecuch J. Chem. Phys. **149**, 151101 (2018).

[37] P. K. Samanta, N. S. Blunt, G. H. Booth J. Chem. Theory Comput. **14**, 3532?3546 (2018).

[38] G. H. Booth, A. J. W. Thom, A. Alavi, J. Chem. Phys. **131**, 054106 (2009).

[39] S. Ten-no, J. Chem. Phys. **138**, 164126 (2013).

[40] Y. Ohtsuka, S. Ten-no, J. Chem. Phys. **143**, 214107 (2015).

[41] S. Ten-no, J. Chem. Phys. **147**, 244107 (2017).

[42] D. Cleland, G. H. Booth, A. Alavi, J. Chem. Phys. **132**, 041103 (2010).

[43] N. S. Blunt J. Chem. Phys. **148**, 221101 (2018).

[44] N. S. Blunt, A. J. W. Thom, C. J. C. Scott J. Chem. Phys. **15**, 3537-3551 (2019).

[45] Y. Garniron, A. Scemama, E. Giner, M. Caffarel, P.-F. Loos J. Chem. Phys. **149**, 064103 (2018).

2 Stochastic perturbation theory in truncated configuration spaces

2.1 Introduction

Obtaining accurate solutions of the many-body Schrödinger equation while maintaining moderate computational costs has been one of the main challenges especially in condensed matter physics and quantum chemistry for decades. The exact many-body wave function can be, in principle obtained by the diagonalization of the full configuration interaction (FCI) matrix in a given Hilbert space. However, the dimension of the problem grows combinatorially with the size of the system, which limits the applicability of the method to small number of correlated electrons. Certainly, exact diagonalization is not an option when one wants to treat medium-sized molecules. As a solution, selected configuration interaction (SCI) methods[3, 4, 5, 6, 7] were proposed long time ago, and has been used for a guide function with improved nodes in diffusion Monte Carlo.[8, 9] SCI has renewed attention and has been improved in recent years.[10, 11, 19, 13, 12, 12, 13, 14, 15, 16, 17, 18] Other modern quantum techniques, such as for example density matrix renormalization group (DMRG),[20, 21, 22] have been developed to tackle problems in large Hilbert spaces that were out of reach in the past. Recently, a full coupled-cluster reduction (FCCR) approach has been introduced to demonstrate that selected coupled-cluster is promising for a balanced treatment of dynamic and nondynamic correlation effects on the same footing.[23]

Another approach of importance in recent years is FCI quantum Monte Carlo (FCIQMC),[25, 1] that aims at solving the Schrödinger equation by employing a walker distribution to represent the CI coefficients in a determinantal basis. The imaginary-time evolution (ITE) of the wave function,

$$\frac{d|\psi\rangle}{d\tau} = -(\hat{H} - S)|\psi\rangle, \tag{2.1}$$

is simulated by QMC comprising of spawning, diagonal death-cloning, and annihilation events with the energy shift S used to control the walker population in FCIQMC.

The Fermion sign problem[27, 28] is mitigated by annihilating walkers with different signs on the same determinants. Model space QMC (MSQMC)[29, 30] is an extension of FCIQMC to calculate excited states utilizing effective Hamiltonian formalism. A novel state-selective partitioning[2] has been proposed to treat multiple solutions simultaneously at the cost of a single MSQMC run with coupled-electron pair approximation (CEPA)-type corrections to reduce the size-inconsistency error of the initiator approximation. Open-source QMC software packages including the extensions of FCIQMC are ready available[32] for calculations and further developments.[33]

The sign problem becomes very sever at low walker occupations as the system size increases. This is because determinants with few walkers with fluctuating signs can accidentally give birth to progeny walkers that cause sign-incoherent noise. The initiator approach[1] has been introduced to overcome this difficulty, achieved by a restriction in the spawning step using initiator determinants with occupation exceeding a prescribed threshold, *i.e.* a progeny spawned from a non-initiator determinant cannot survive if it turns out to be on an unoccupied determinant. The initiator approach enables a wide range of applications, but yet is essentially a truncation in the Hilbert space and it leads to size-inconsistency errors.[30] In a recent study[21], with the introduction of second-order Epstein-Nesbet corrections the initiator error was improved in a natural and inexpensive manner. Several semi-stochastic methods have been also proposed.[35, 36]

A stochastic algorithm of many body perturbation theory (MBPT) in arbitrary order has been discussed previously.[2] Although deriving the algebraic expressions of higher order perturbation energies is tedious, stochastic algorithms provide convenient ways for their estimates. These stochastic algorithms are reminiscent of the higher-order MBPT algorithm[40] based on FCI.[40] Nevertheless, the main obstacle in practical calculations is that the number of walkers explodes for higher order terms.

A truncation in configuration space can potentially bridle this explosive growth of the walkers. However, as well-known truncated CI methods provide size-inconsistent energies as higher order excitations are necessary[44] for the cancellation of unlinked terms in the perturbative expansion.

In this thesis I acquaint the reader with the convenience of Antisymmetrized Goldstone Diagrams (ASG) that will be used to express certain important terms in the perturbation series. For the sake of easy comprehension, I limit my discourses to the ITE of the third order wave function and explain the size-inconsistency problem of the fourth order perturbative energy. When all the perturbative terms are sampled the results are free from flaws and size-consistent energies are obtained in every order.

On the contrary, when four fold excitations are excluded from the sampling unpleasant imperfections appear and it would be advantageous to sample a different type of perturbation series that is free from such errors. The research of the current chapter focuses on this problem and proposes a stochastic algorithm based on the linked-cluster expansion to provide size-consistent energies in every order. With some modifications in the original algorithm it is possible to implement the sampling of the linearized coupled cluster (LCC) wave function. Besides the implementation I performed several calculations using our new methodology to demonstrate the correctness and the usability of stochastic perturbation theory. One of the most intriguing questions was how the accuracy changes with the excitation level of the sampled space. Accuracy in this case might be loosely defined and I mainly focus on the size-inconsistency problem to compare the CI and the LCC expansion. Size consistency is a very important property as it has been harangued by the research community. One of the main reasons for this is that in the thermodynamic limit the energy particle obtained by a method exhibiting size-inconsistency error (for example configuration interaction with singles and doubles (CISD)) approaches to zero, which is clearly unphysical. On the other hand, size-consistent expansions, such as couple cluster expansion provide non-zero energies per particle even at the thermodynamic limit.

I also investigate a semi-stochastic algorithm which is radically different from the previous works[41, 42]. It reduces stochastic noise by using a deterministic subspace, and treating the complementary space using walker dynamics. I express the first order wave function using perturbation theory to be treated deterministically as the fragment represented by this term thought to constitute a very important part of the total wave function. The size of the first order space might grow prohibitively large with increasing number of active orbitals and electrons, however contrary to expectations, the first order wave function and the second order energy can be calculated in an inexpensive manner.

When these quantities are available the ITE commences from an initial wave function that is much more accurate than the HF reference. A very similar idea was mentioned by N. Blunt et. al [24]. In some of the i-FCIQMC calculations the reference wave function is chosen to be the CISD wave function instead of a single Slater determinant.

In my work, my approach is somewhat different as I express the ITE of the residual part of the wave function to sample the full FCI wave function as a sum of the first order wave function, the residual part and the HF reference. In the current thesis, I derive the necessary formulations for the variation of the residual wave function and provide a discourse on how to implement the individual terms based on the original MSQMC sampling algorithm. In the numerical results section of this chapter I demonstrate how the standard deviation changes when the residual sampling is applied and I also present calculation results on molecular systems that were previously investigated using stochastic perturbation theory. As a result of a reference wave function, which is more accurate than the HF reference the calculation converges upon the FCI wave function in shorter imaginary time.

I find it important to mention that the number of walkers required to accurately represent the converged wave function does not change, the residual sampling is solely a more effective way to treat

the first order part of the total wave function in a deterministic manner and therefore eliminate the necessity for the QMC steps that would sample this part of the total wave function. The CI coefficients are determined from the deterministic formula and walkers are distributed on the determinants proportional to their coefficient. The exponential growth of walkers is observed at the initial part of the ITE just as in all other projector based QMC methods and the only way to overcome this bottleneck is to introduce certain approximations. Although the first order wave function is stored on memory and there is no need for the sampling of its coefficients the size of the residual space can become a limiting factor. The problem of numerous determinants with low walker population arises and it would be expedient to introduce the initiator adaption into the residual sampling as well.

2.2 Theory

2.2.1 Perturbative expansion based on CI

Perturbation theory has been in the focus of research for decades as it provides a systematically improvable way to solve the many body problem. One can find a solution to the one determinant approximation (HF problem), look for means to describe the neglected dynamic correlation effects and arrive at the perturbation expansion of the wave function.

Many body perturbation theory can be used to progressively approximate the FCI solution by adding higher and higher terms to the series. The perturbative series might converge slowly or even diverge and in many cases one has to calculate many high order terms to extract meaningful results out of the calculations. Evaluating formulae for higher order energy corrections in perturbation theory is no easy task with large memory requirements. Therefore, I introduce a stochastic algorithm based on the imaginary time evolution (ITE) to sample these components. The following derivation is based on a single reference (SR) approach for the wavefunction,

$$\psi(\tau) = \psi_0 + \psi_Q(\tau), \tag{2.2}$$

that is, the zeroth order wavefunction ψ_0 is a fixed single Slater determinant typically chosen to be the Hartree-Fock (HF), and the rest of the wavefunction $\psi_Q(\tau)$ is stochastically sampled using the following ITE:

$$\frac{d\psi}{d\tau} = -\hat{Q}(\hat{H} - E)\psi, \tag{2.3}$$

where $\hat{Q} = 1 - \hat{P}$, and $\hat{P}$ is the projector onto $|\psi_0\rangle$. To formulate stochastic perturbation theory, I employ the Møller-Plesset (MP) partitioning:

$$\hat{V} = \hat{H} - \hat{H}_0 = \hat{H} - \hat{F}, \tag{2.4}$$

where $\hat{F}$ is the Fock operator. The ITE of the nth order perturbative component can be obtained by substituting the expansions of the wave function and energy to equate the contributions in the same order in the perturbation as shown in the literature[2],

$$\begin{aligned}
\frac{d\psi^{(1)}}{d\tau} &= -(\hat{H}_0 - E^{(0)})\psi^{(1)} - \hat{Q}\hat{V}\psi_0 && (2.5)\\
\frac{d\psi^{(2)}}{d\tau} &= -(\hat{H}_0 - E^{(0)})\psi^{(2)} - \hat{Q}(\hat{V} - E^{(1)})\psi^{(1)} && (2.6)\\
\frac{d\psi^{(n)}}{d\tau} &= -(\hat{H}_0 - E^{(0)})\psi^{(n)} - \hat{Q}(\hat{V} - E^{(1)})\psi^{(n-1)} \\
&+ \sum_{i=2}^{n-1} E^{(i)}\psi^{(n-i)}. && (2.7)
\end{aligned}$$

I represent each of the perturbed wavefunction in the above equations using a distribution of signed walkers, and the population of ψ_0 is identified with the so-called booster weight[29], N_b. The wave function is now discretized into an ensemble of signed walkers, each of which is assigned to a particular Slater determinant. The variation of $\psi^{(n)}$ in a time step $\delta\tau$ is expressed as

$$\delta\psi^{(n)} = \delta\psi^{(n)}_{\text{death}} + \delta\psi^{(n)}_{\text{spawn}} + \delta\psi^{(n)}_{\text{trans}}, \tag{2.8}$$

where the contributions of the death/cloning, spawning, and transcription steps are

$$\begin{aligned}
\delta\psi^{(n)}_{\text{death}} &= -\delta\tau(\hat{H}_0 - E^{(0)})\psi^{(n)}, && (2.9)\\
\delta\psi^{(n)}_{\text{spawn}} &= -\delta\tau\hat{Q}\hat{H}_{\,\mathrm{x}}\psi^{(n-1)}, && (2.10)\\
\delta\psi^{(n)}_{\text{trans}} &= -\delta\tau[(\hat{H}_{\,\mathrm{d}} - \hat{H}_0 - E^{(1)})\psi^{(n-1)} \\
&+ \sum_{i=2}^{n-1} E^{(i)}\psi^{(n-i)}], && (2.11)
\end{aligned}$$

and $\hat{V} = \hat{H}_{\mathrm{x}} + \hat{H}_{\mathrm{d}} - \hat{H}_0$ with the diagonal and off-diagonal Hamiltonian components

$$\hat{H}_{\mathrm{d}} = \sum_I |I\rangle\langle I|\hat{H}|I\rangle\langle I|, \tag{2.12}$$

$$\hat{H}_{\mathrm{x}} = \hat{H} - \hat{H}_{\mathrm{d}}, \tag{2.13}$$

defined by the determinants $\{I\}$.

It is important to remark that the ITE of the nth order perturbative wavefunction includes contributions from lower order wavefunctions. In this work, I perform *booster-weight-constant* simulations, although N_{b} can be used for a population control in a *walker-number-constant* simulation. A large N_{b} is needed to avoid sign problems for strongly correlated systems by increasing the number of walkers. In this work, I employ N_{b} ranging from 250 to 2.000 which is usually sufficient for weakly correlated systems. Once the walker populations are equilibrated, one can calculate the $(2n)$th and $(2n+1)$th order instantaneous energies using the Wigner rule. However, $\psi^{(n)}$ and $\psi^{(n-1)}$ in these energy formulas appearing on both bra and ket sides provide biased energy estimates when single-set of walker populations are employed. Although this bias can be eliminated using independent replica[17] walker populations, I employ the simple mixed estimator in this particular work:

$$E^{(n)} = \langle\psi_0|\hat{V}|\psi^{(n-1)}\rangle. \tag{2.14}$$

Throughout the ITE the walker population fluctuates and so does the energy in Equation 2.14. This noise can be reduced by integrating the instantaneous energy value over an imaginary time interval:

$$E^{(n)} = \langle E^{(n)}(\tau)\rangle_\tau = \frac{1}{\tau_2 - \tau_1}\int_{\tau_1}^{\tau_2} E^{(n)}(\tau)d\tau \tag{2.15}$$

Note that the stochastic algorithm in this subsection is based on the CI expansion.

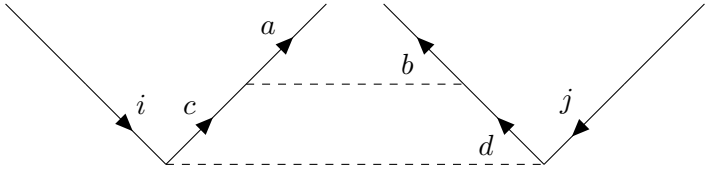
2.2.2 Diagrammatic discussion of the ITE

Before I proceed to the necessary modifications that are required to be implemented in the walker dynamics, a short recapitulation of some MBPT related diagrams and theorems may be necessary. The diagrammatic representation of perturbation theory provide a very convenient way to visualize the effect of operators on Slater determinants and to express complicated algebraic formulas. In the following discussion I use antisymmetrized Goldstone (ASG) diagrams[38] to introduce some of the important concepts of perturbation theory focusing on the ITE of the perturbative terms.

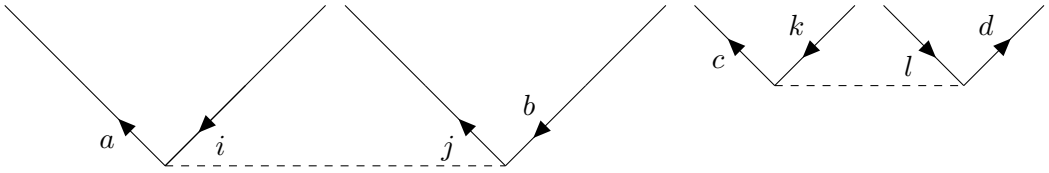
To understand the relationship between the terms in the perturbation series, it is important to clarify what the terms *connected*, *disconnected*, *linked* and *unlinked* diagrams mean.

Connected diagrams are diagrams with no separated parts (its vertexes are connected with interaction lines). Otherwise the diagram is called disconnected. We call diagrams unlinked if it is disconnected and it has closed parts. Connected diagrams or diagrams with no closed parts are called linked.

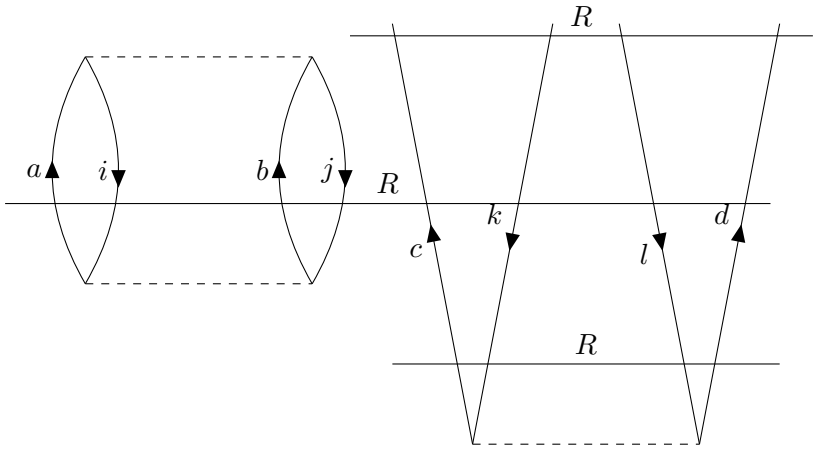
For example this diagram is *connected*:



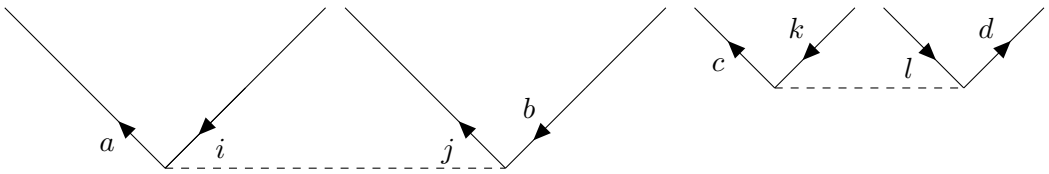
This diagram is *disconnected*:



This diagram is disconnected and has closed parts, therefore *unlinked*:



And this diagram although disconnected but as it has no closed parts is called *linked*:

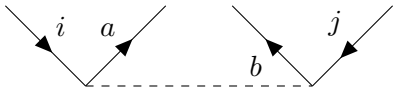


With the use of the perturbation operator in its normal ordered form ($\hat{W}$) and the resolvent operator $\hat{R}_0$ we can express the first order wave function:

$$|\psi^{(1)}\rangle = \hat{R}_0 \hat{W}|0\rangle \quad \hat{W} = \hat{V} - E^{(1)} \quad \hat{R}_0 = \frac{\hat{Q}}{E_0 - \hat{H}_0} \tag{2.16}$$

On the following figures, starting from the reference, I present the corresponding ASGs to illustrate $\psi^{(1)}$ and $\psi^{(2)}$ with the repeated application of $\hat{W}$.

Considering the diagrammatic expression of the normal ordered two-body operator, the first order wave function is expressed as:



, and I use the convention for the resolvent that in open diagrams (wave function diagrams) there is one line above the highest vertex and there are lines between each pair of successive vertices.

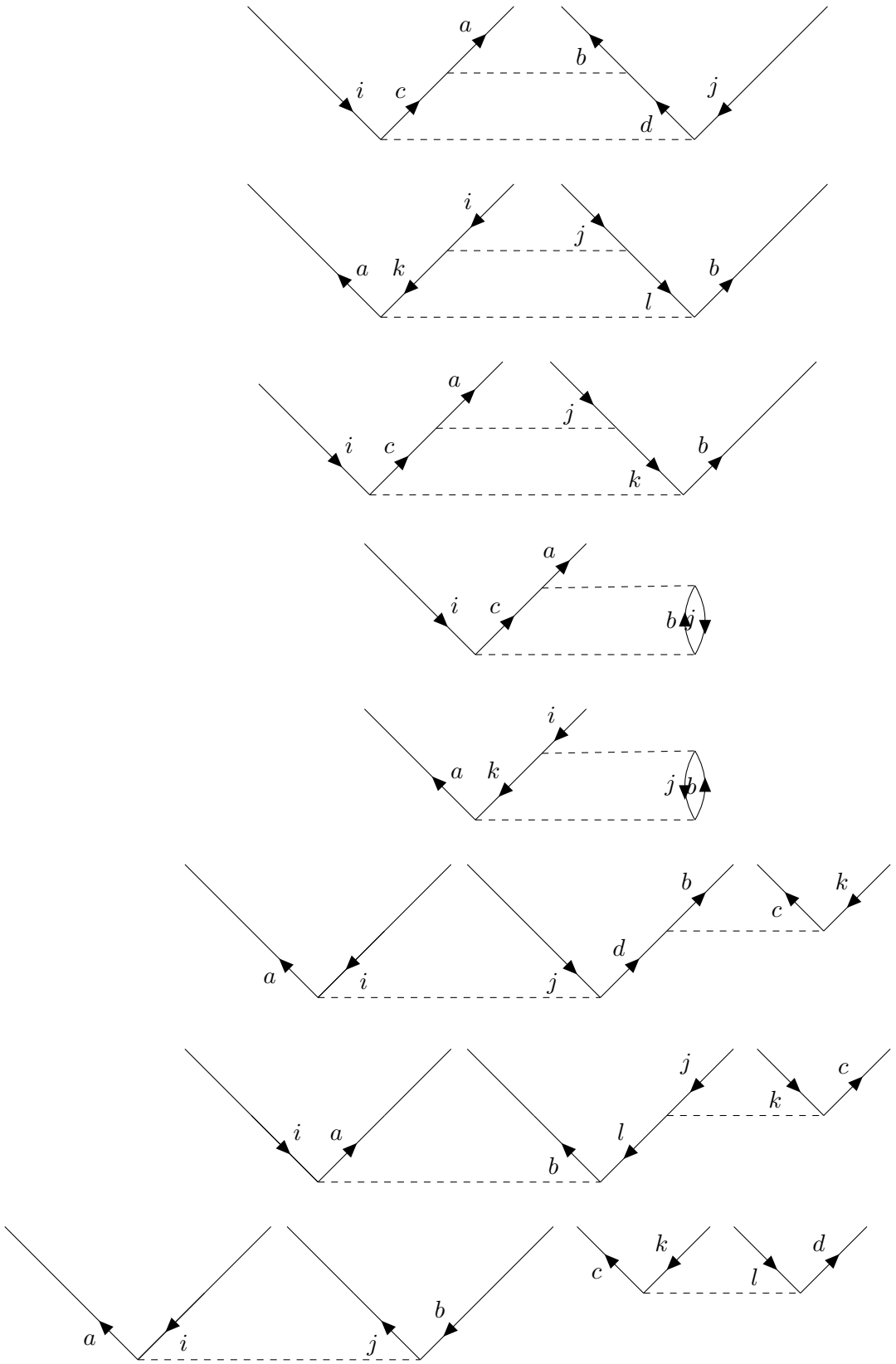
Evaluating the diagram yields:

$$|\psi^{(1)}\rangle = \frac{1}{4} \sum_{abij} \frac{\langle ab||ij\rangle}{\epsilon_{ij}^{ab}} |\phi_{ij}^{ab}\rangle \tag{2.17}$$

Proceeding to the second order by applying $\hat{W}$ and the resolvent again yields:

$$|\psi^{(2)}\rangle = \hat{R}_0 \hat{W} \hat{R}_0 \hat{W} |0\rangle \tag{2.18}$$

This corresponds to the following 8 eight diagrams (omitting the resolvent for clarity):

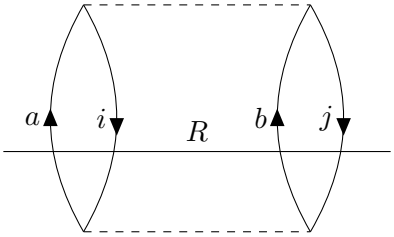


The third order wave function is the first in the perturbation series where the renormalization term appears:

$$|\psi^{(3)}\rangle = \hat{R}_0\hat{W}\hat{R}_0\hat{W}\hat{R}_0\hat{W}|0\rangle - \hat{R}_0E^{(2)}|\psi^{(1)}\rangle \tag{2.19}$$

The linked-diagram theorem states that the wave function is given by a sum of linked diagrams only, provided that all EPV terms are included in the summation for each diagram. In order to satisfy this, the unlinked diagrams in the left term of Equation 2.19. have to cancel with the renormalization term ($\hat{R}_0E^{(2)}|\psi^{(1)}\rangle$). The factorization theorem can be applied to show the cancellation. It states, that a product of two connected open diagrams is equal to a sum of disconnected diagrams, each made up of the two diagrams of the original product with all possible time orderings of the vertices of one of the disconnected part relative to those of the other.

First of all, let's express $\hat{R}_0E^{(2)}|\psi^{(1)}\rangle$ in terms of the previously introduced ASGs. The product consists of $E^{(2)}$:



This can be evaluated as:

$$E^{(2)} = \frac{1}{4}\sum_{abij}\frac{\langle ij||ab\rangle\langle ab||ij\rangle}{\epsilon_{ij}^{ab}} \tag{2.20}$$

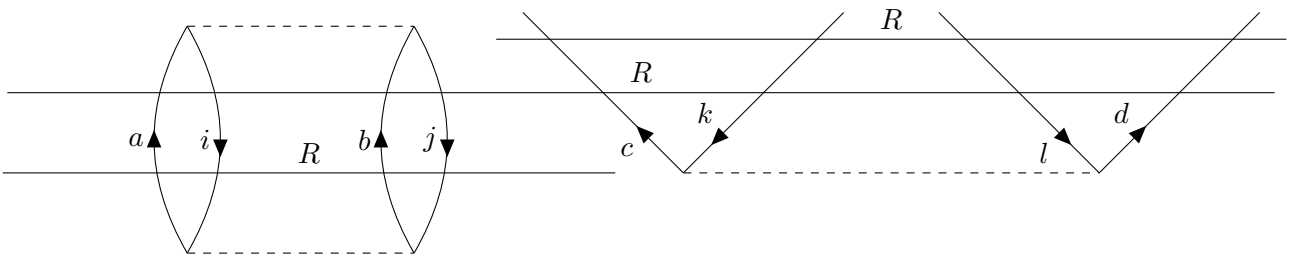
Therefore the renormalization term can be expressed as:

$$\hat{R}_0E^{(2)}|\psi^{(1)}\rangle = \frac{1}{16}\sum_{abijcdkl}\frac{\langle ij||ab\rangle\langle ab||ij\rangle\langle cd||kl\rangle}{\epsilon_{ij}^{ab}\epsilon_{kl}^{cd}\epsilon_{kl}^{cd}}|\phi_{kl}^{cd}\rangle \tag{2.21}$$

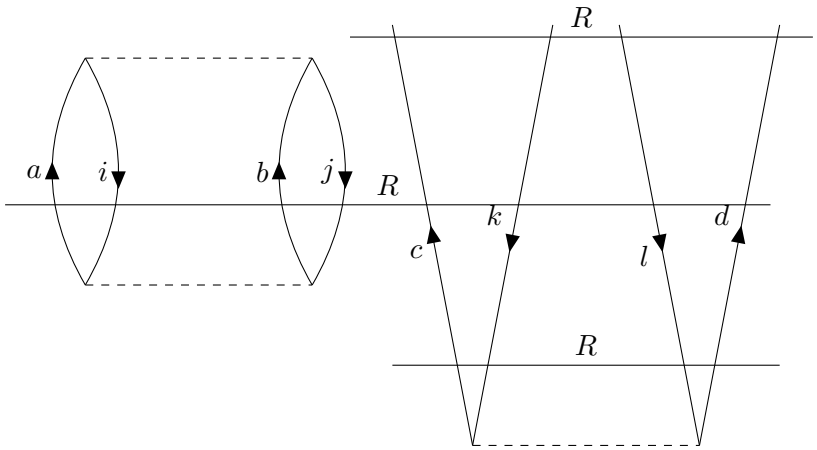
The algebraic manifestation of the factorization theorem (Hugenholtz 1957, Frantz and Mills 1960, Brandow 1977) can be employed here, yielding:

$$\hat{R}_0E^{(2)}|\psi^{(1)}\rangle = \frac{1}{16}\sum_{abijcdkl}\left(\frac{\langle ij||ab\rangle\langle ab||ij\rangle\langle cd||kl\rangle}{\epsilon_{ij}^{ab}\epsilon_{ijkl}^{abcd}\epsilon_{kl}^{cd}} + \frac{\langle ij||ab\rangle\langle ab||ij\rangle\langle cd||kl\rangle}{\epsilon_{kl}^{cd}\epsilon_{ijkl}^{abcd}\epsilon_{kl}^{cd}}\right)|\phi_{kl}^{cd}\rangle \tag{2.22}$$

Using ASGs, the first term can be expressed as:



And the diagrammatic expression of the second term is:



The disconnected quadruply excited diagram in $|\psi^{(2)}\rangle$ can be contracted in two possibly ways with $\hat{W}$ and these contractions will produce the aforementioned two diagrams. Besides, if we take a look at the ITE of $|\psi^{(3)}\rangle$:

$$\frac{d}{d\tau}|\psi^{(3)}\rangle = -(\hat{H}_0 - E^{(0)})|\psi^{(3)}\rangle - \hat{Q}\hat{W}|\psi^{(2)}\rangle + E^{(2)}|\psi^{(1)}\rangle \tag{2.23}$$

, we can conclude that the transcription step, that is responsible for sampling the third term in the ITE expansion samples unlinked contribution in the wave function and these terms with opposite sign will cancel with the terms generated from the spawning step.

The CI based perturbative expansion leads to a size-inconsistency error when I truncate the Hilbert space due to the sampling of unlinked terms. It is also possible to demonstrate the cancellation of these terms based on the following wave operator terminology.

The variation of the spawning step in Equation (2.10) is $\delta\psi^{(3)}_{\text{spawn}} = -\delta\tau\hat{Q}\hat{H}_{\text{x}}\psi^{(2)}$. Representing $\psi^{(n)}$ in the wave operator form,

$$\psi^{(n)} = \Omega^{(n)}\psi_0, \tag{2.24}$$

the factorization theorem[39] tell us that the quadruple excitation in $\hat{\Omega}^{(2)}$ can be expressed as the square of the 1st order wave operator,

$$\Omega^{(2)}_{\text{Q}} = \frac{1}{2}(\Omega^{(1)}_{\text{D}})^2. \tag{2.25}$$

Therefore, the contractions between $\hat{H}_{\text{x}}$ and $\Omega^{(1)}_{\text{D}}$ in $-\frac{1}{2}\delta\tau\hat{Q}\hat{H}_{\text{x}}(\Omega^{(1)}_{\text{D}})^2\psi_0$ give rise to the unlinked term $-\delta\tau E^{(2)}\psi^{(1)}$, and this is cancelled by the corresponding term $\delta\tau E^{(2)}\psi^{(1)}$ of the transcription step in Equation (2.11). This is the linked cluster theorem (LCT) of the ITE. However, LCT does not hold if quadruple excitations are excluded from the pertrubative expansion. If quadruple excitations are discarded from the spawning step no unlinked terms are generated as $\Omega^{(2)}_{\text{Q}}$ is not sampled and $\delta\tau E^{(2)}\psi^{(1)}$ of the transcription step contaminates the expression with erroneous scaling. In other words contribution of the unlinked term to $E^{(4)}$ exhibits an unphysical quadratic scaling $(E^{(2)})^2$ with system size.

In order to solve this problem, I introduce a linked modification to the stochastic perturbation algorithm, *i.e.* the operator product in the spawning step is replaced by a connected product, and the unlinked terms are discarded from the transcription step,

$$\delta\psi^{(n)}_{\text{spawn}} \quad \Rightarrow \quad -\delta\tau\hat{Q}[\hat{H}_{\text{x}}, \hat{\Omega}^{(n-1)}]\psi_0, \tag{2.26}$$

$$\delta\psi^{(n)}_{\text{trans}} \quad \Rightarrow \quad -\delta\tau(\hat{H}_{\text{d}} - \hat{H}_0 - E^{(1)})\psi^{(n-1)}. \tag{2.27}$$

The commutator form in the spawning step guarantees the exclusion of disconnected products.

The following discussion acquaints the reader with the relationship between couple cluster theory and the previous derivation. The coupled-cluster choice of the wave operator:

$$\psi(\tau) = e^{\hat{T}(\tau)}\psi_0 \tag{2.28}$$

immediately manifests the extensivity property. Using the couple-cluster *ansatz* the ITE equation can be written as:

$$\frac{\psi(\tau)}{d\tau} = -\hat{Q}e^{\hat{T}(\tau)}\hat{H}e^{-\hat{T}(\tau)}\psi_0 \tag{2.29}$$

The right hand side of the equation can be transformed using the Baker-Campbell-Hausdorff formula:

$$e^{-\hat{B}}\hat{A}e^{\hat{B}} = \hat{A} + [\hat{A},\hat{B}] + \frac{1}{2}[[\hat{A},\hat{B}],\hat{B}] + \frac{1}{3!}[[[\hat{A},\hat{B}],\hat{B}],\hat{B}] + ... \tag{2.30}$$

. For the sake of simplicity double and higher commutators are neglected to arrive at the working equation:

$$\frac{\psi(\tau)_{\mathrm{LCC}}}{d\tau} = -\hat{Q}(\hat{H} + [\hat{H},\hat{T}(\tau)])\psi_0 \tag{2.31}$$

An order by order expansion enables us to express the ITE of the nth order term of the linearized coupled-cluster wave function:

$$\frac{\psi^{(n)}_{\mathrm{LCC}}(\tau)}{d\tau} = -(\hat{H}_0 - E^{(0)})\psi^{(n)}_{\mathrm{LCC}} - \hat{Q}[\hat{H}_{\mathrm{x}},\hat{T}^{(n-1)}]\psi_0 - (\hat{H}_{\mathrm{d}} - \hat{H}_0 - E^{(1)})\psi^{(n-1)}_{\mathrm{LCC}} \tag{2.32}$$

From this derivation it is clear that if the preceding modifications are implemented in the original stochastic perturbation theory algorithm then the perturbative terms of the linearized coupled-cluster wave function can be calculated. I therefore call this scheme "perturbative expansion based on LCC" to distinguish it from the previous one based on the CI expansion.

Although the linearization of the ITE at Equation 2.31. produces easily programable equations it has its limitations, namely that no matter how high the excitation level in the calculation, the results will not converge to the FCI value though implementing even double commutators (e.g. $\frac{1}{2}[[\hat{H},\hat{T}],\hat{T}]$) is extremely laborious due to operator products in the ITE that would necessitate to use of replica sampling to avoid sampling biases. The imperfections of the LCC expansion as higher order excitations are incorporated are presented in the numerical results section.

2.3 Computational details

Based on the original MSQMC implementation, the algorithm to sample arbitrary order perturbative terms in Slater determinants space was implemented in our program suite using the FORTRAN programing language. The important input parameters to run a stochastic perturbation theory calculation are presented in the following table:

Parameter	**Explanation**
NCORE	Number of core orbitals
IINT	Integration start step
NMAC	Number of macro steps
NMIC	Number of micro steps
NBOST	Number of booster weights
MMP	Maximum order of perturbation to be calculated
NSEED	Number of seeds
ISEED	The value of the seed if NSEED == 1
NACT	Number of active orbitals
NDORM0	Number of dormant orbitals in the P-space
NACT0	Number of active orbitals in the P-space
DTAU	Imaginary time step
NCI	Maximum allowed excitation level
NCC	CI or LCC sampling

Table 2.1: Input parameters of the software to run stochastic perturbation theory calculations.

Currently, the software supports two operational modes and the user can switch between these modes using the input parameter NCC. In default sampling mode (NCC=0) the walker population evolves according to the CI expansion and stochastic MBPT energies are calculated. There is an additional parameter (NCI) to control the excitation level of the sampled space. In case this parameter is provided in the input file all determinants with an excitation level exceeding this threshold are discarded. Users can sample the perturbative components of LCC wave function by setting NCC = 1. When this happens, the spawning events are initiated from connected Hamiltonian matrix elements and no unlinked terms are sampled in the transcription step.

In the numerical results section most of the calculations were carried out following some initial preparatory calculations to explore the optimal combination of parameters that would provide the most accurate results while consuming the least computational power. In most of these cases I found that there are two factors that has significant influence on the calculation results.

First of all the choice of N_b is extremely important as this has an effect on the number of walkers at equilibrium state and in some cases the walker population is unable to converge upon the FCI solution if N_b is too small. Obviously, the larger N_b is the more CPU the calculation will consume, therefore finding the optimal value for N_b is crucial.

Secondly, the length of the calculation is also a very important parameter. Even in a converged state the walker population exhibits some kind of fluctuations and it is no easy task to determine the exact value for the energy if this fluctuation is too large. I found that these parameters depend on the system under consideration, namely some converge in shorter imaginary time than others. The length of the calculation can be changed with DTAU, NMIC and NMAC and the total length of the calculation is the product of these values. As soon as the desired accuracy is attained, there is no need to increase the value of these parameters and it is advised to keep them as small as possible. I found that the calculation is extremely sensitive to the change in DTAU as this value has a direct influence on the slope of the convergence. The larger this value is the faster the calculation converges, however one

must not choose a too large value for this as DTAU is used in the spawning step. The amplitude of the spawned walkers is proportional to the time step on if this value is large the noise of the calculation grows significantly.

2.4 Results and discussions

2.4.1 Convergence of stochastic MBPT

I first investigate the convergence of stochastic MBPT for small molecules using cc-pVDZ basis set. For most of the cases, where the stochastic noise is negligible, a single calculation is enough to demonstrate the correctness of the stochastic perturbation theory. In some cases, I used several calculations with different seeds to average the results to obtain reliable data.

Table 2.2 lists stochastic perturbation energies of N_2 through 7th order compared to those of deterministic values. The bond length is taken from Ref. [43] and core orbitals are frozen. It is shown that all of the stochastic energies are in accord with the deterministic ones within the standard error.

Table 2.2: Stochastic perturbation energies in mE_h for N_2 ($r_{NN} = 1.12013$Å) in comparison with the deterministic result. The standard errors for the last digits are indicated in the parentheses for the stochastic results. The booster weight, 1,000, on the HF determinant is used with 3 different seeds. The integration of the instantaneous energy is started at 100 a.u. with the total simulation length, 350 a.u. The HF energy is -108.9496E_h

	$E^{(n)}$	
n	deterministic	stochastic
2	-313.18	-313.2(1)
3	7.23	7.3(1)
4	-25.06	-24.9(2)
5	6.17	6.2(1)
6	-4.50	-4.5(2)
7	1.26	1.2(2)

In the limit of infinite number of walkers it is possible to reproduce the deterministic perturbative energies. With increasing order of perturbation the number of walkers required to represent a term in the perturbation series grows significantly, therefore the addition of high order terms can cause serious computational overhead. For the calculation I performed on the nitrogen molecule the perturbation series converges and with the addition of a few terms chemical accuracy can be attained. In less favorable cases, the perturbation series might converge slowly or diverge and stochastic perturbation theory would require tremendous amount of walkers to reproduce near FCI values. I those cases the current approach might not be the best choice and one has to resort to different post-Hartree-Fock approaches. Also observe that this calculation was carried out in a non-truncated space, and excitations of every possible levels were generated. Therefore the dimension of the high order terms is much larger than that of the lower order ones. Nevertheless, they recover a much less amount of the correlation energy and this is one of the reasons for the necessity of a very large Hilbert space to obtain accurate properties using the CI expansion.

The objective of the ensuing calculation was to ascertain the usability and the correctness of the stochastic algorithm based on LCC. I implemented the necessary modifications for the sampling of the LCC wave function using stochastic perturbation theory. A modification in the transcription step to sample the terms defined by Equation 2.27. was applied and I was able to calculate the LCC wave function using the previously describe walker dynamics. Figure 2.1. shows the imaginary-time evolution of perturbative energies based on the LCC expansion through 8th order for nitrogen and water molecules. The energies in the SD space gradually approach the LCCSD energies with increasing order of perturbation. Figure 2.1. tells us that even in these simple cases (relatively small basis set and weak dynamic correlation) nothing can be said of the order of the successive terms in the perturbation series. For water $E^{(2)}$ is higher than $E^{(3)}$, whereas for nitrogen $E^{(3)}$ is higher than $E^{(2)}$.

2.4.2 Comparision of CI and LCC

What follows here is a detailed investigation in connexion with the accuracy of the different approaches. The main motivation for this calculation was to find an answer to the question whether it is expedient to use the CI based wave function expansion or the LCC based expansion when one has limited computational resources and the only option to perform a large scale calculation is to introduce some kind of truncation in the walker dynamics.

The necessity for truncation might not seem evident for these test cases where the active space is not large and the dynamic correlation is not significant, however our future goal is to extend the usability of our methods to strongly correlated systems with numerous active orbitals. Table 2.3. presents the total numbers of walkers and deviations of the perturbation energies through the 8th order based on the CI and LCC expansions for H_2O and N_2 taking the FCI energies as references. The configuration space is limited to the maximum excitation level L_{ex} with respect to the HF reference. Note LCC in the SD space ($L_{\mathrm{ex}} = 2$) provides fairly accurate results almost comparable to the CI ones in the SDTQ space ($L_{\mathrm{ex}} = 4$). In this case, the number of walkers is significantly smaller when using the LCC type equations. For most cases that the accuracy of the LCCSD series compared to that of the CISDTQ method is higher or becomes it might even become more accurate for larger systems, since the latter still contains a size-inconsistency error that increases with the system size.

The energies obtained using the CI based expansion rigorously converge to the respective FCI limits, while the those of the LCC series exhibit residual errors due to the linearization of the cluster operator. This approximation introduces an error with respect to FCI that does not disappear even when higher excitations are incorporated. When one wants to calculate more accurate results and correct these imperfections the linearization has to be abandoned. However, the expressions of nested commutators from the exponential ansatz are not straightforward to expand and I omit the derivation of programable equations in the current work.

N_2[1]	CI		LCC	
L_{ex}	ΔE	N_{w}	ΔE	N_{w}
2	-36.5(2)	2.7×10^4	-3.1(3)	2.6×10^4
3	-25.8(2)	1.2×10^5	12.4(1)	1.2×10^5
4	-2.4(3)	7.5×10^5	9.6(3)	4.8×10^5
5	-1.5(2)	2.9×10^6	9.4(2)	1.3×10^6
6	-0.5(3)	6.8×10^6	9.8(3)	1.9×10^6
H_2O[2]	CI		LCC	
L_{ex}	ΔE	N_{w}	ΔE	N_{w}
2	-11.4(2)	1.2×10^4	-0.7(2)	1.2×10^4
3	-8.7(1)	6.4×10^4	3.1(2)	6.0×10^4
4	-0.8(1)	3.2×10^5	3.1(3)	2.3×10^5
5	-0.3(3)	9.5×10^5	2.6(3)	4.8×10^5
6	0.0(2)	1.7×10^6	3.2(3)	6.3×10^5

Table 2.3: The errors of the perturbative energies (ΔE) in mE_{h} and total numbers of walkers (N_{w}) of the perturbative series through 8th order based on the CI and LCC expansions for H_2O and N_2. L_{ex} refers to the maximum excitation level of the configuration space with respect to the HF reference. The energy is obtained by taking the average of calculations of 3 different seeds. The simulation time is 100 a.u. and the integration of the mixed estimator is turned on at 60 a.u. imaginary time. 1,000 and 2,000 of booster weights are employed for the calculations of H_2O and N_2, respectively.

The favorable scaling properties of the CC ansatz have been harangued by the scientific community and the results of Table 2.3. can serve as a proof for such opinions. The perturbation series based on the CI expansion of the wave function is size-consistent as long as every possible excited determinants

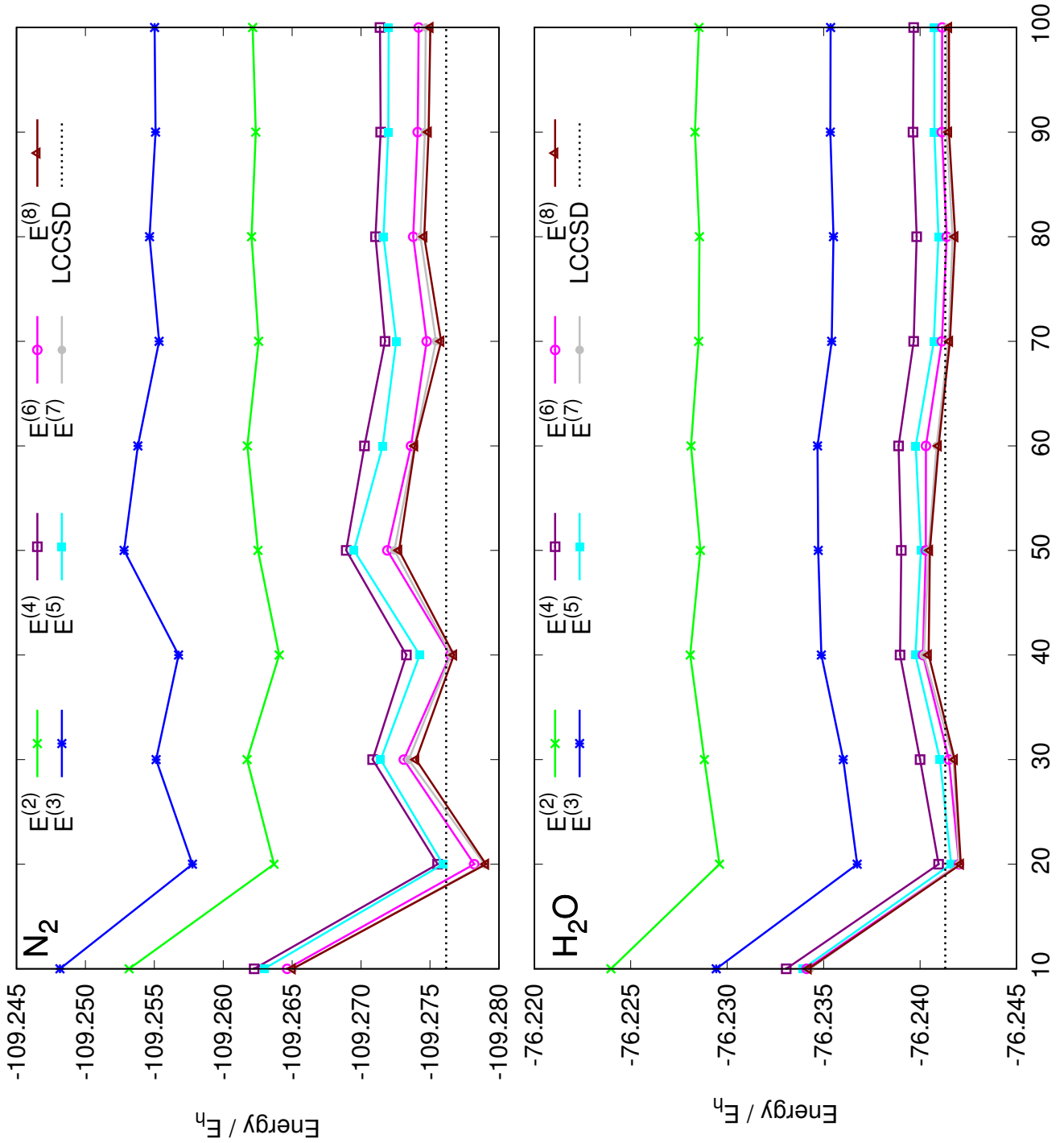


Figure 2.1: Imaginary-time evolutions of perturbative series based on the LCC expansions in the singles and doubles space for the water and nitrogen molecules in cc-pVDZ basis (frozen core). The geometries are taken Ref. [43]. 1,000 of booster weight is used for the duration of the simulation 100 a.u. The dashed lines denote LCCSD energies from deterministic calculations.

are incorporated as no terms with erroneous scaling remain in the series. The CC expansion on the other hand expresses the wave function in such a form that is intrinsically free from such flaws. Therefore, my conjecture is that the stochastic sampling of the CC expansion will have better scaling properties than that of a CI based expansion.

2.4.3 Size-(in)consistency with respect to the excitation level

Size-consistency is a property of a certain computational method. The method is said to be size-consistent if the energies of two systems A and B and of the combined system AB with A and B very far apart, computed using the same method, satisfy

$$E(AB) = E(A) + E(B) \tag{2.33}$$

. Size-consistency seems like a modes requirement of a computational method, however not every approach satisfies this criteria. HF and perturbation theory in a non-truncated spaces do while truncated CI does not. The simplest example is CISD which is only size consistent for systems at most two active electrons. This issue becomes a problem for large systems because this lack of size consistency leads to the wrong limiting behavior of the correlation energy. For larger system the omission of higher excitations introduces such a constraint on the wave function that it fails to provide meaningful properties in the limit of infinite number of particles:

$$\lim_{N\to\infty} \frac{E_{corr}}{N} = 0 \tag{2.34}$$

In other words, the errors of size-inconsistent methods grow severely with system size and the correlation energy per particle vanishes. This implies that the CI expansion based stochastic perturbation theory can only be applied to small molecules when a truncation is applied in the excitation space. For macroscopic system such as a crystal this approximation would fail to provide meaningful results unless the FCI space is employed, which is clearly not a feasible option for large systems. In order to study this issue I perform further investigations regarding the connection between excitation levels and size-consistency.

Figure 2.2 shows the size-inconsistency error of perturbation energies for noninteracting Ne_2 in limited configuration spaces. 1,000 booster weights are used for the simulation of 200 a.u. All results are obtained using 3 seeds. The spawning step is modified such that progenies are allowed to be spawned on certain excitations with respect to the HF reference. It is shown that the CISDT series leads to the size-inconsistent $E^{(4)}$, and the use of the SDQ space provides the correct scaling behavior.

With this in mind, the size-consistency in $E^{(4)}$ can be restored in two possible ways, i.e. (i) use the connected contribution for the spawning within the SD space and restrict the transcription step so that unlinked terms are not sampled, or (ii) preserve the transcription step in the CI form (Equation (2.7)) and utilize the SD(T)Q space to keep sampling the unlinked contributions in the principal term of E(4). Note that triples are unrelated to the size-consistency of the fourth order energy, and CISDQ corresponds to the SDQ-MBPT(4) model.[44] For the CI based expansion the only option to make a certain order of perturbative term in the series size-consistent is to increase the allowed level of excitation until the exact expression is sampled by the ITE.

From Figure 2.2. it is clear that up until $E^{(3)}$ single and double excitations are sufficient to express the exact expression, therefore there is no size-inconsistency error bellow E(4). Thus perturbation theory exhibits no size-inconsistency error neither when the CC expansion nor when the CI expansion is used unless the order of perturbation exceeds 3 in energy or 2 in the wave function.

Because the energy of a macroscopic system is an extensive thermodynamic property, it must be proportional to the number of particles; thus CI based stochastic perturbation theory is not a satisfactory procedure for treating large systems. Besides, size-inconsistency is not only a serious problem when one wants to calculate properties of large molecular systems, but also when studying the dissociation of a molecule one should use a method which is equally good, in a certain sense, for the intact molecule and the resulting fragments.

2.4.4 Polyacenes

In the ensuing section I turn my attention to the acene series and investigate the accuracy of the previously proposed stochastic perturbation theory algorithms.

Polyacenes, linear benzoid hydrocarbons, are acyclic variants of cycloacenes. Due to their technological potential and intrinsic values as models for more complex conjugated molecules, they have been subjected to a large number of experimental and theoretical investigations.

Quantum chemists have been interested in polyacenes $[(C_4H_2)_n]$ for a very long time. Polyacenes, linear polybenzenoid hydrocarbons, are acyclic variants of cycloacenes. The band gap of polyacenes is of considerable interest due to their possible use as semiconducting materials in field-effect transistors. It is pointed out that in polyacene the Fermi surface lies at the edge of the Brillouin zone, but that an accidental degeneracy between the valance and conduction bands makes it metallic nonetheless.

Moreover, this class of materials has the possibility of possessing interesting condensed phases including high-temperature superconductivity and ferromagnetism. Polyacene has been shown to be an unusual one-dimensional conductor with novel electronic structure. They behave rather differently than other quasi-one-dimensional conductors that have been studied to date. Due to their technological potential and their intrinsic vale as models for more complex conjugated molecules they have been subjected to a large number of experimental and theoretical investigations.

Band theory arguments proposed by Kertesz and Hoffmann [45] and by Chapman and Kivelson [46] showed that superconductivity could arise from the unusually small band gaps in polyacenes. Nearly two decades ago they have predicted that polyacenes might have conducting properties [46]. The low band-gaps and bond equalization [47] of linear polyacenes endow them with properties suitable for electronic device application and potential for non-linear optical properties in suitably modified forms [48].

It is well-known that as polyacenes increase in size the singlet triplet energy gap becomes increasingly small. Large polyacenes were predicted to behave as one-dimensional organic conductors with a zero band gap. The get an accurate picture of of the ground state and the first excited state electron correlation should be accounted and high-level ab initio methods should be used. Studies [49] have already discovered that at the RB3LYP/6-31G(d) level of theory, the wave function for as small an oligoacene as hexacene (and all longer oligoacenes) becomes unstable. Shrinking energy gap between the ground and first excited states with increasing number of rings makes time propagator based methods converge slower. On the other hand, the series provides an ideal model to investigate size-consistency effects in perturbative stochastic calculations.

The DMRG energies of the molecules composed of rings from 2 to 12 are available in the literature.[50] In the following calculations I use the same geometry and STO-3G basis correlating the complete π valence space.

Figure 2.3 shows the errors of the CI and CC based fourth order perturbation energies with respect to the full MP4 energies for the acene series. LCCSD clearly exhibits the least error, and will be useful to calculate properties for which size-consistency is inevitable.[44] The perturbative expansion of the wave function combined with QMC sampling implies the perturbative components represented by arrays of walkers in each order. Walkers residing on a certain determinant appertaining to the CI coefficient of this determinant. Consequently, the number of required walkers to represent the component grows with increasing the perturbation order. This becomes prohibitive as spawning on determinants of high excitation levels are allowed.

Figure 2.4 depicts the total numbers of walkers representing the consecutive perturbative wave function components. As expected, the sampling in the SDQ space requires significantly large number of walkers compared to those in the SD space. As a result, considering the computational cost of the two size-consistent series through the fourth order, the stochastic perturbation of LCCSD with the operation

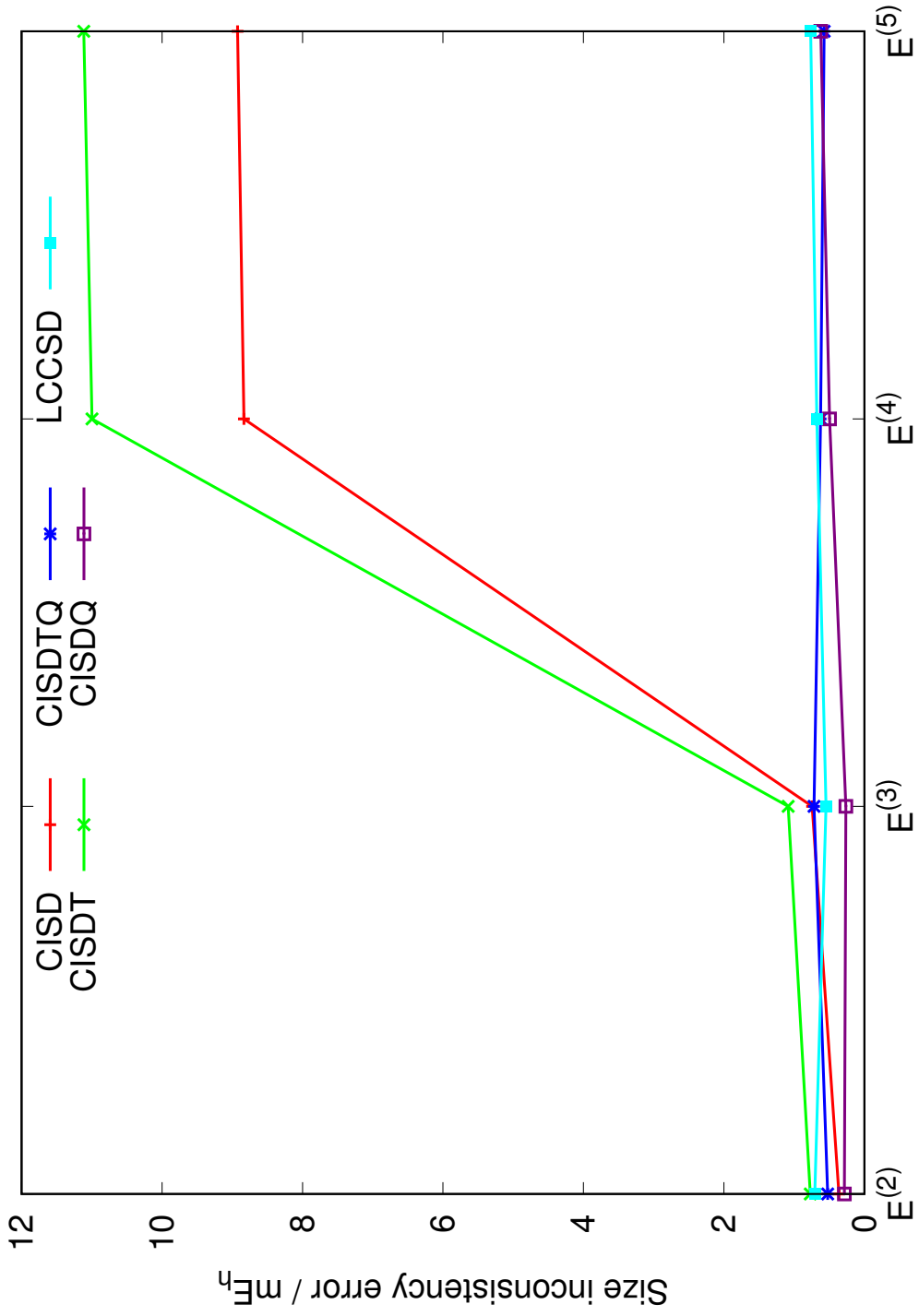


Figure 2.2: Size-inconsistency error for Ne and noninteracting Ne_2 using different excitation levels in the CI and LCC based expansions. The size-inconsistency error is calculated as the difference between the energy of noninteracting two atoms and twice the energy of a single atom.

space comparable to the CISD one is always more economical than CISDQ. Note that the stochastic LCCSD series is rigorously size-consistent even in higher orders.

2.4.5 Residual sampling

I finally introduce a semi-stochastic algorithm that improves the convergence of stochastic perturbation theory. The exact nth order perturbative wave function can be constructed as

$$(E_0 - \hat{H}_0)|\psi^{(n)}\rangle = \hat{Q}\hat{V}|\psi^{(n-1)}\rangle - \sum_{k=0}^{n-1} E^{(n-k)}|\psi^{(k)}\rangle. \tag{2.35}$$

A lower order wave functions possesses large walker populations on determinants in a limited configuration space. Therefore, the use of the deterministic expressions of the lower order components could potentially reduce stochastic noise while preserving computational costs.

In the current work, I employ the deterministic expression for the first order wave function,

$$|\psi^{(1)}_{\mathrm{DET}}\rangle = \frac{1}{E_0 - \hat{H}_0}\hat{Q}\hat{V}|\psi_0\rangle = \frac{N_{\mathrm{boost}}}{4}\sum_{abij}\frac{\langle ab||ij\rangle}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b}|\Phi^{ab}_{ij}\rangle \tag{2.36}$$

which is plugged into the ITE of $|\psi^{(2)}\rangle$ to calculate $E^{(3)}$.

The standard deviation of the stochastic and semi-stochastic algorithms with respect to the number of independent calculations using different seeds is presented in Figure 2.5. Clearly, the semi-stochastic sampling provides more reliable results even with fewer number of seeds.

Based on this idea, it is possible to derive a working equation to efficiently sample the residual space that is the sum of all perturbative wave functions above first order:

$$\chi = \sum_{n=2}^{\infty}\psi^{(i)} = \psi - \psi^{(1)}_{\mathrm{DET}} \tag{2.37}$$

Next, I derive the ITE of the residual space (χ) using the ITE of the total wave function and omitting the terms that do not change with imaginary time:

$$\frac{d\chi}{d\tau} = -\hat{Q}(\hat{H}_0 + \hat{V} - E)(\chi + \psi^{(1)}_{\mathrm{DET}} + \Psi_0), \tag{2.38}$$

To proceed further, I use the definition of the first order wave function from Equation 2.36.:

$$\frac{d\chi}{d\tau} = -\hat{Q}(\hat{H}_0 + \hat{V} - E)(\chi + \frac{1}{E_0 - \hat{H}_0}\hat{Q}\hat{V}|\psi_0\rangle + \Psi_0) \tag{2.39}$$

Further rearanging the terms yields:

$$\frac{d\chi}{d\tau} = -\hat{Q}(\hat{H} - E)\chi(\tau) + \hat{Q}(-\hat{H}_0 - \hat{V} + E + E_0 - E_0)(\frac{1}{E_0 - \hat{H}_0}\hat{Q}\hat{V}|\psi_0\rangle + \Psi_0) \tag{2.40}$$

The reference wave function can be eliminated from the ITE:

$$\frac{d\chi}{d\tau} = -\hat{Q}(\hat{H} - E)\chi(\tau) + \hat{Q}(E_0 - \hat{H}_0 - \hat{V} + E - E_0)(\frac{1}{E_0 - \hat{H}_0}\hat{Q}\hat{V}|\psi_0\rangle + \Psi_0) \tag{2.41}$$

And I obtain the final version of the ITE of the residual wave function, χ:

$$\frac{d\chi(\tau)}{d\tau} = -\hat{Q}(\hat{H} - E)\chi(\tau) - \hat{Q}(\hat{V} - \bar{E})\psi^{(1)}_{\mathrm{DET}}, \tag{2.42}$$

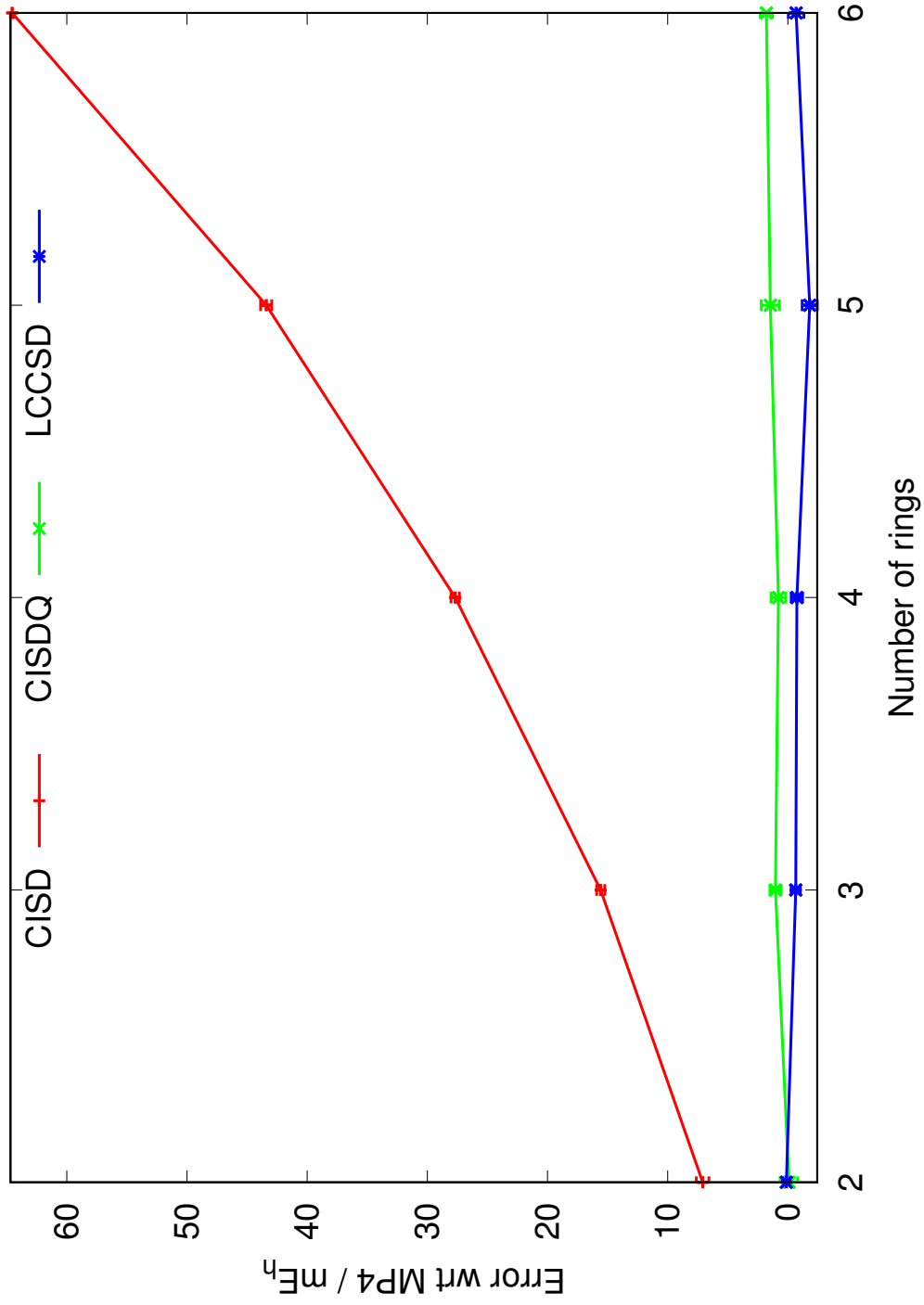


Figure 2.3: The error of the fourth order perturbative energies for the acene series (from Naphthalene to Hexacene) using CI and CC equations. Even though CISDQ is a significantly larger space, it provides less accurate results compared to LCCSD. 500 booster weights are used for the simulation of 20 a.u.

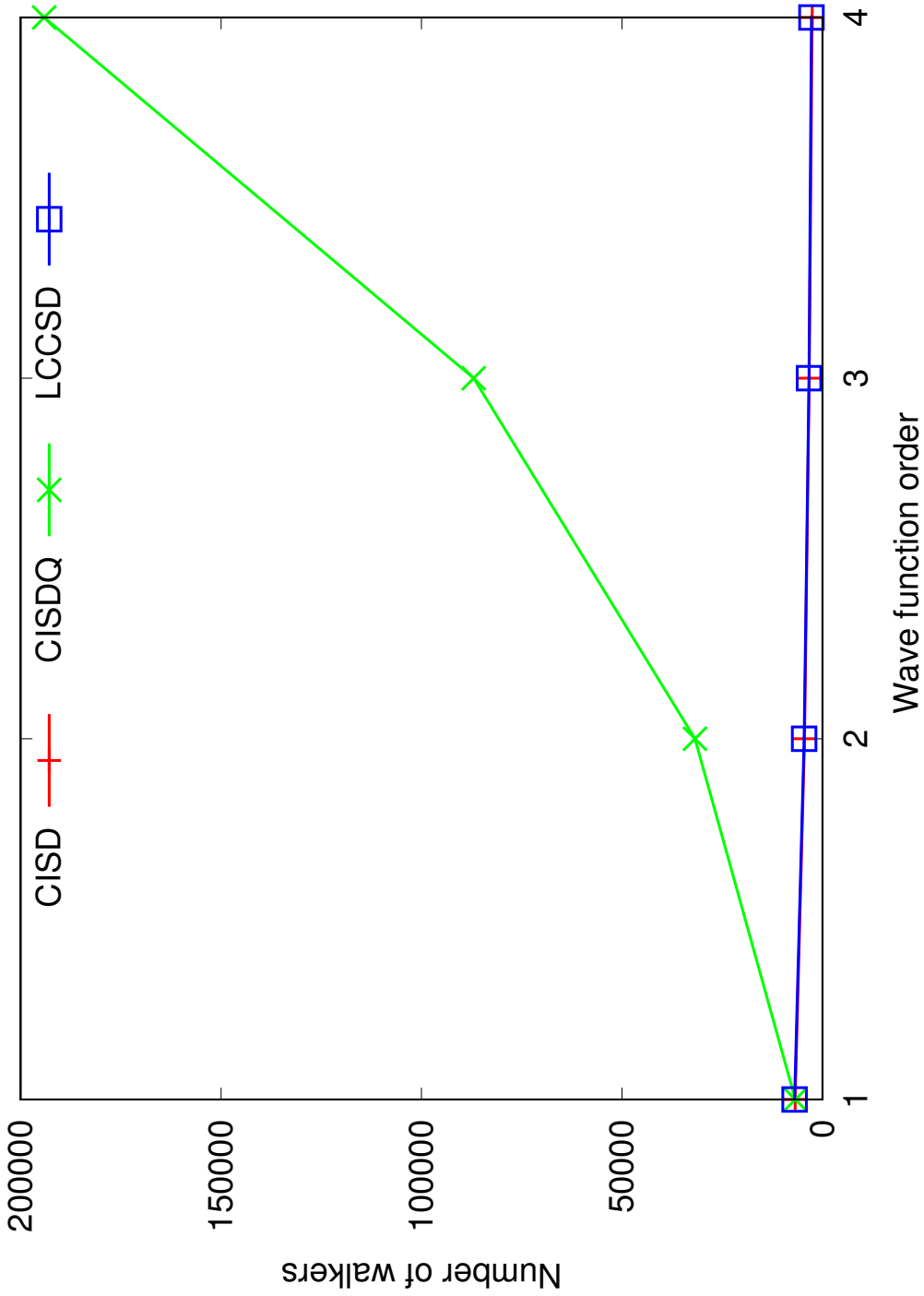


Figure 2.4: Number of walkers in equilibrium state on each perturbative term for a Tetracene molecule in minimal basis. As expected, the number of walkers required to represent the respective perturbative components grows as the sampling is carried out in larger space. 500 booster weights are used for the simulation of 20 a.u.

where $\bar{E} = E - E_0$. This differs from the usual semi-stochastic algorithm that discriminates determinants in the deterministic and stochastic spaces.[41, 42]

In order to arrive at the programable equations for the sampling of the residual wave function I express the Hamiltonian matrix using its diagonal and off-diagonal components just as in the preceding discussion. In the discretized time representation the variation of $\chi(\tau)$ can be expressed as:

$$\delta\chi(\tau) = \delta\psi_{\text{res-s}} + \delta\psi_{\text{res-d/c}} + \delta\psi_{\text{res-t}}, \tag{2.43}$$

, where the individual terms define the walker evolution of χ:

$$\delta\psi_{\text{res-s}} = -\delta\tau\hat{Q}H_x\chi(\tau) - \delta\tau\hat{Q}H_x\psi^{(1)}_{\text{DET}} \tag{2.44}$$

$$\delta\psi_{\text{res-d/c}} = -\delta\tau(H_d - E)\chi(\tau) \tag{2.45}$$

$$\delta\psi_{\text{res-t}} = -\delta\tau(H_d - H_0 - E - E_0)\psi^{(1)}_{\text{DET}} \tag{2.46}$$

Note that there is no data dependency between these step, therefore they can be executed in a completely parallel fashion.

The first term involves off-diagonal Hamiltonian matrix elements and is responsible for spawning events both from the residual wave function and the first order wave function. The second term is essentially the same as the death-cloning step of the previous two algorithms. Finally, the third term contains diagonal contributions from the Hamiltonian matrix therefore it can be implemented using the transcription step of the CI or the LCC based expansion. The annihilation and the merging step of the walkers are unchanged. The imaginary time dependent total energy ($E_{tot}(\tau)$) is calculated as the sum of the second order energy and the residual energy (E_{res}) at the time step τ:

$$E_{tot}(\tau) = E^{(2)} + E_{res} = E^{(2)} + \langle\psi_0|\hat{H}|\chi(\tau)\rangle \tag{2.47}$$

The integration of the energy is carried out in an identical fashion:

$$E_{tot} = \langle E_{tot}(\tau)\rangle_\tau = \frac{1}{\tau_2 - \tau_1}\int_{\tau_1}^{\tau_2} E_{tot}(\tau)d\tau \tag{2.48}$$

As for many weakly correlated systems, the first order perturbative wave function constitutes a fairly accurate representation of the correlated wave function, significant increases in efficiency are expected relying on its aid. However, in the initial step of the calculation $\psi^{(1)}_{\text{DET}}$ needs to be calculated and this might be a bottleneck for large systems. As soon as the first order wave function is available the second order energy can be calculated and the ITE commences from the second order energy instead of the Hartree-Fock energy. As a result, a more accurate energy is used in the walker dynamics and also the fact that the spawning steps that were originally necessary to sample the coefficients of $\psi^{(1)}_{\text{DET}}$ are completely omitted can speed up the calculation. These effects are investigated in details in the numerical results section.

In every time step (*macro-step*) the variation of the wave function is calculated in the following four steps:

1. Walkers are generated as a result of two off-diagonal spawning processes. Every walker on both χ and $\psi^{(1)}_{\text{DET}}$ attempts to spawn a walker by calculating H_{ij} where i is the index of the parent determinant and j is the index of the progeny determinant. This steps samples the term $\delta\psi_{\text{res-s}}$.
2. Every walker is either cloned or killed after calculating the difference between the diagonal matrix element (H_d) of the Slater determinant they reside on and $E_{tot}(\tau)$ to obtain the probability of this event. This steps samples the term $\delta\psi_{\text{d/c}}$.

3. Walkers are generated as a result of a diagonal spawning precess from $\psi^{(1)}_{\text{DET}}$. This steps samples the term $\delta\psi_{\text{res-t}}$.

4. The list of newly spawned walkers are merged into the list of parent walkers and walkers with opposite signs are annihilated.

Readers, who wish to understand the walker evolution in details, can refer to Figure 2.6.

When sampling $\delta\psi_{\text{res-t}}$, the diagonal matrix elements of the first order wave function are referenced, therefore I found it expedient to store these on memory and since for a given parameter space the dimension of $\psi^{(1)}_{\text{DET}}$ is fixed this should not cause significant overhead.

In the initial phase of the calculation walkers are distributed to the reference determinant and the first order wave function. The number of walkers are proportional to the coefficient calculated from the perturbative expansion and the ITE commences. In this phase, the spawning events generate tremendous amount of walkers albeit much less than in the algorithm with stochastic first order wave function. As a result, an exponential growth of walkers is observed and one needs to exercise prudence and allocate sufficient memory for the calculation to survive.

A necessary condition for convergence is non-zero overlap between the reference wave function and the target state. If this condition is satisfied, the walker population should equilibrate and, when using no approximations, the FCI wave function is obtained. The importance sampling employed in this approach chooses the most important determinants prior to the less important ones. These determinants are usually closer in energy to E_{tot}, therefore they die with less probability than the determinants that are sampled in the latter part of the ITE. In equilibrium state, the number of spawning events are roughly equal to the number of death events and the number of walkers does not grow anymore. The number of walkers attained in equilibrium state is strongly dependent on the initial number of walkers employed and this influences both calculation costs and accuracy.

On Figure 2.7. I carried out an investigation to shed light on the connexion between the number of initial walkers (N_b) and the accuracy of the results. The objective of this calculation was a water molecule using cc-pVDZ basis for which the FCI energy is known. On the upper plot I demonstrate the total number of walkers during the ITE for 5 a.u. using three different N_b values. Explosive walker growth until 1.5 a.u. is observed in every cases and the walker population seem to reach equilibrium after the number of walkers exceeds 10^7. Clearly the calculation with $N_b = 400$ exhibits the least accuracy and even after the integration commenced the stochastic noise distorts the energy and it not easy to extract meaningful results from this calculation. However, when increase N_b to 1000 the ITE relaxes on the residual wave function with much less stochastic noise and the FCI energy can be obtained in a very short imaginary time. Further increasing N_b to 30.000 smoothens the convergence curve significantly, though the required CPU time also increases therefore I only recommend using extremely large N_b values when computational resources permit. For this weekly correlated system, the second order perturbation energy $E^{(2)}$ is only $\sim$10 mEh above the FCI energy and the perturbation series is well-behaved. This implies fast convergence when the first order wave function is treated deterministically and this was indeed observed as in 5 a.u. the FCI wave function was converged upon. In more complicated cases in which the perturbation series converges slowly or diverges residual sampling might posses less promising convergence properties.

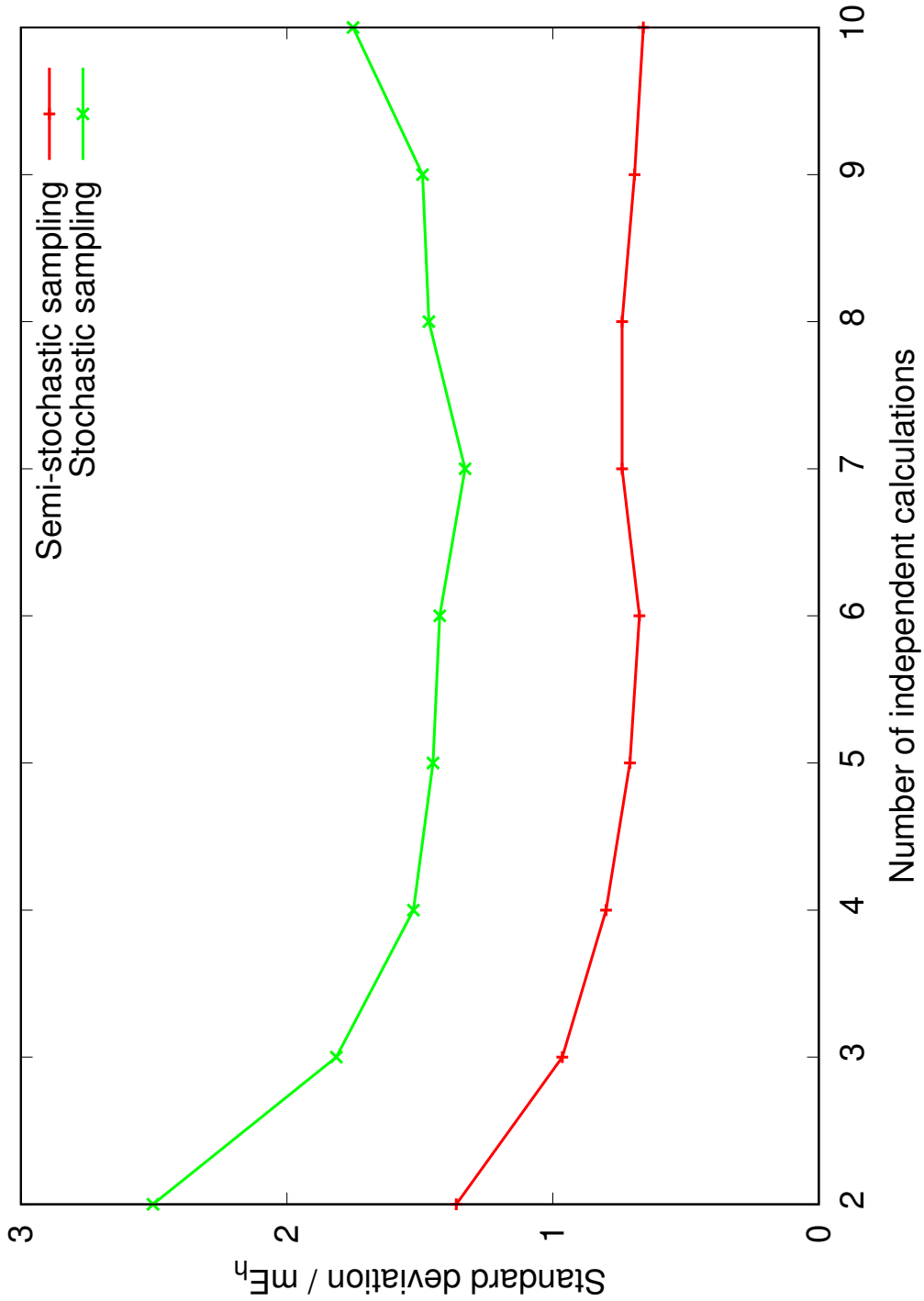


Figure 2.5: The standard deviation of $E^{(3)}$ for Ne in cc-pVDZ with respect to the number of independent calculations using different seeds. The integer part of $\psi^{(1)}$ is for simplicity with $N_{\text{boost}} = 500$ for the short simulations of 5 a.u.

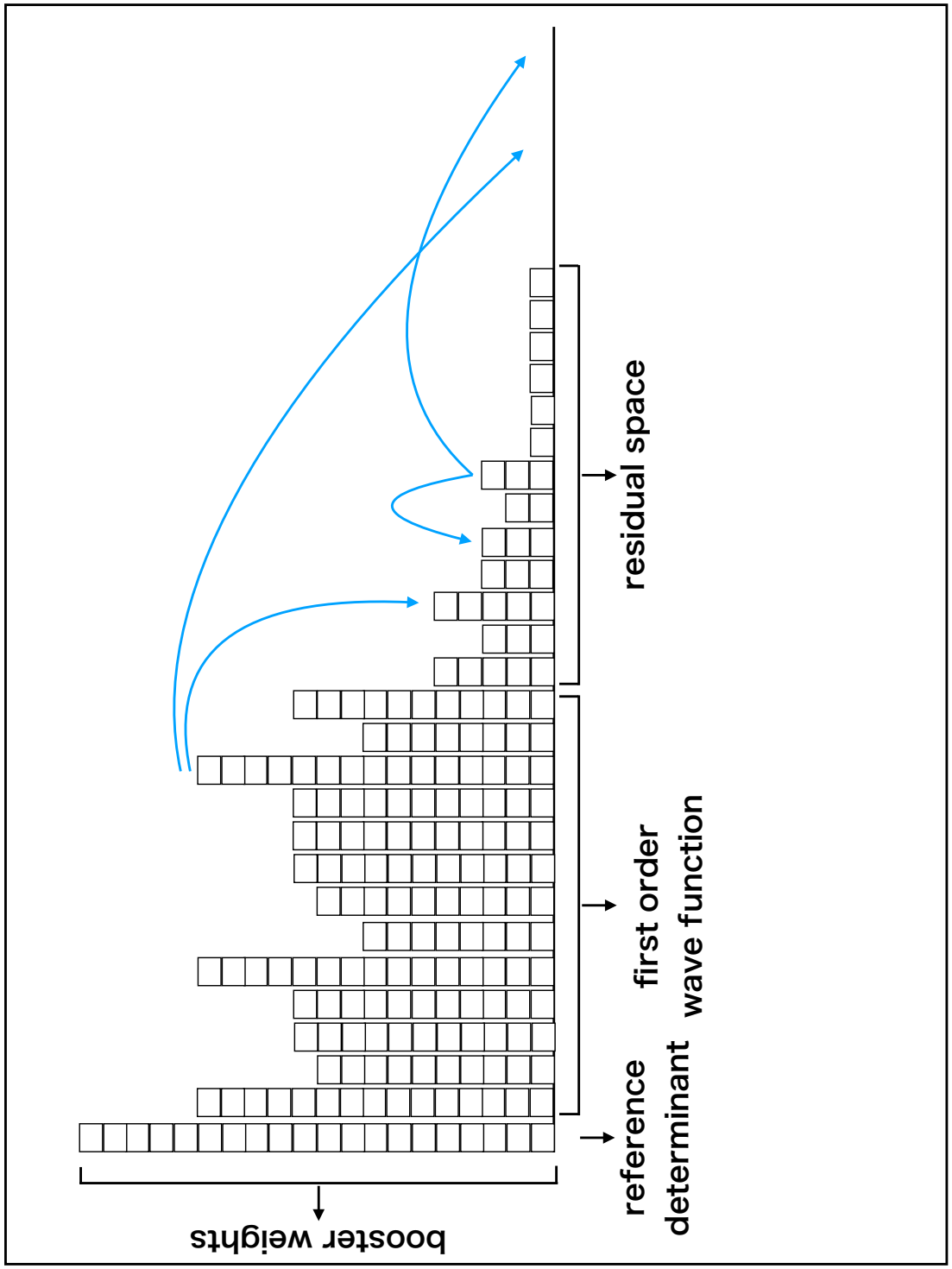


Figure 2.6: Schematic representation of the residual sampling algorithm in which $\Psi^{(1)}$ is treated deterministically and the rest of the space (residual space) is sampled using walker dynamics.

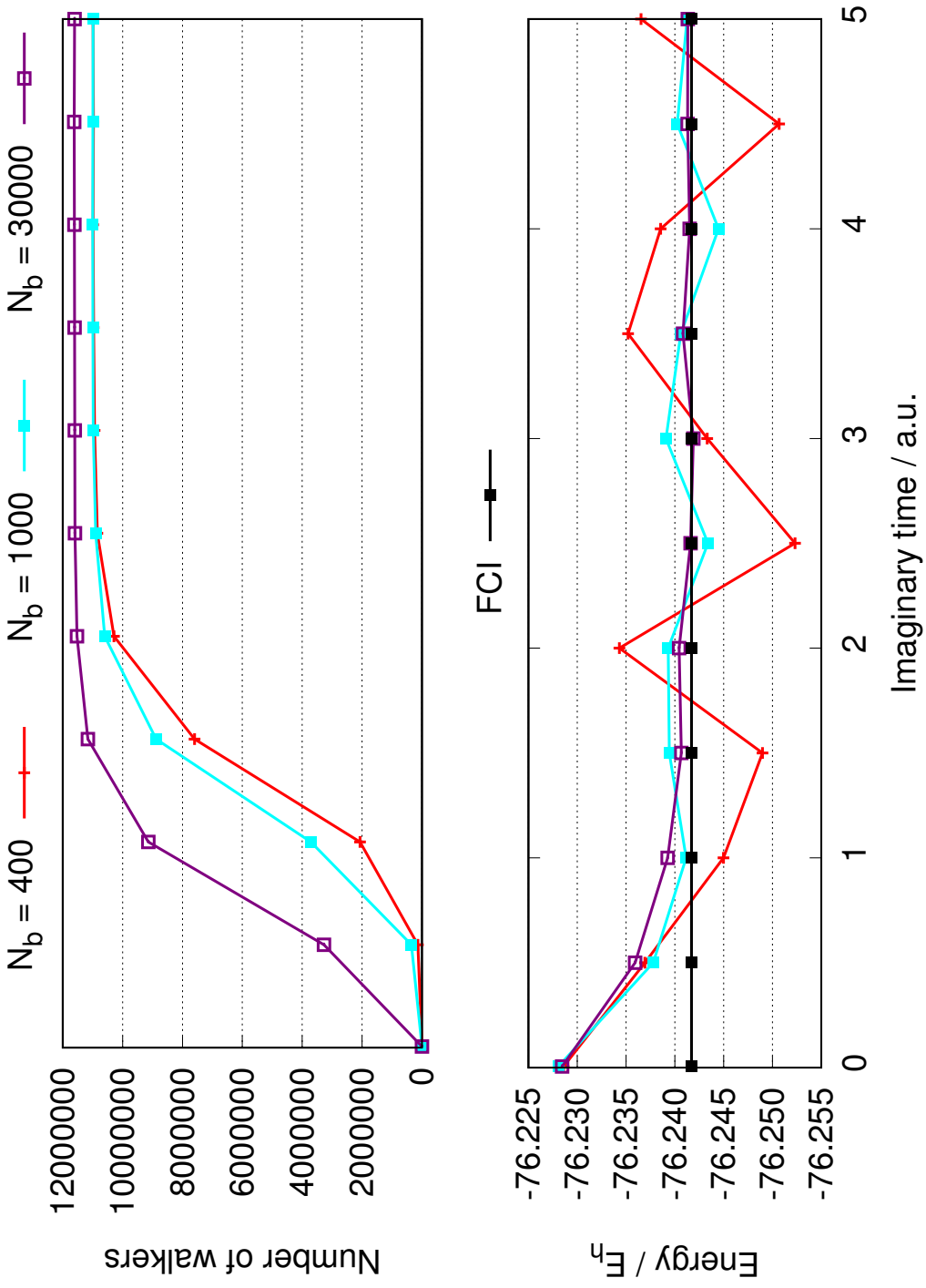


Figure 2.7: Walker growth and energy convergence with different N_b values using residual sampling. The calculation was carried out on an H_2O molecule with $\tau = 5$ a.u. and $\tau_1 = 2.5$ a.u. using the cc-pVDZ basis set. The frozen-core approximation was used for this calculation.

2.5 Conclusion

I may summarize the various aspects of my findings as follows.

1. I have found that by simulating the imaginary time evolution of the perturbation wave functions it is possible to provide stochastic MBPT corrections to the zeroth order wave function. My findings also show that in order to reach FCI accuracy the number of walkers required to express these terms is proportional to the FCI space size. When one wants to decrease computation costs, it is straightforward to truncate the space, allowing only certain levels of excited determinants to be spawned.

2. I have carried out an ASG based investigation to trace the origin of the size-inconsistency error in truncated Hilbert spaces when using a CI based perturbative expansion. The unlinked terms and their cancellation at the third order wave function was observed.

3. A CC based approach to ensure the correct scaling of the method has been proposed. This linear CC variant turned out to be a useful tool to obtain LCC wave functions of arbitrary order of excitations. Using the CC expansion, I was able to obtain more accurate fourth order energies than with the CI based variant for the acene series.

4. Finally, I proposed a modification in the sampling scheme that can improve the convergence of the original algorithm. The residual sampling seems to perform remarkably well for molecular systems for which the first order wave function along with the reference wave function is a good approximation to the exact solution of the many-body problem.

Within the confines of the stochastic perturbation theory, there is much work remaining before it will be routinely applicable to large molecules. Some of this work involves the efficient calculation paths that take every technological advantage from implementation on parallel computing architectures to robust samplings. Other aspects will involve the capability of accurately describing electronic structures at stretched geometries in the complete basis set limit, in which multi-reference and explicitly correlated F12[51] treatments are essential.

Bibliography

[1] B. Huron, J. P. Malrieu, P. Rencurel, J. Chem. Phys. **58**, 5745-5759 (1973).

[2] R. J. Buenker, S. D. Peyerimhoff, Theor. Chim. Acta **35**, 33-58 (1974).

[3] S. Evangelisti, J.-P. Daudey, J.-P. Malrieu, Chem. Phys. **75**, 91 (1983).

[4] R. J. Harrison, J. Chem. Phys. **94**, 5021 (1991).

[5] A. Povill, J. Rubio, F. Illas, Theor. Chim. Acta **82**, 229 (1992).

[6] E. Giner, A. Scemama, M. Caffarel, Can. J. Chem. **91**, 879 (2013).

[7] E. Giner, A. Scemama, M. Caffarel, J. Chem. Phys. **142**, 044115 (2015).

[8] J. B. Schriber, F. A. Evangelista, J. Chem. Phys. **144**, 161106 (2016).

[9] N. M. Tubman, J. Lee, T. Y. Takeshita, M. Head-Gordon, K. B. Whaley, J. Chem. Phys. **145**, 044112 (2016).

[10] A. A. Holmes, N. M. Tubman, C. J. Umrigar J. Chem. Theory Comput. **12**, 3674 (2016).

[11] W. Liu, M. R. Hoffman, J. Chem. Theory Comput. **12**, 1169 (2016).

[12] J. B. Schriber, F. A. Evangelista, J. Chem. Theory Comput. **13**, 5354 (2017).

[13] J. B. Schriber, K. P. Hannon, C. Li, F. A. Evangelista, J. Chem. Theory Comput. **14**, 6295 (2018).

[14] Y. Garniron, A. Scemama, E. Giner, M. Caffarel, P.-F. Loos, J. Chem. Phys. **149**, 064103 (2018).

[15] Y. Garniron, et al. J. Chem. Theory Comput. **15**, 3591 (2019).

[16] A. D. Chien, A. A. Holmes, M. Otten, C. J. Umrigar, S. Sharma, P. M. Zimmerman J. Phys. Chem. A **122**, 2714 (2018).

[17] P.-F. Loos, A. Scemama, A. Blondel, Y. Garniron, M. Caffarel, D. Jacquemin, J. Chem. Theory Comput. **14**, 4360 (2018).

[18] P.-F. Loos, M. Boggio-Pasqua, A. Scemama, M. Caffarel, D. Jacquemin, J. Chem. Theory Comput. **15**, 1939 (2019).

[19] M. Dash, S. Moroni, A. Scemama, C. Filippi, J. Chem. Theory Comput. **8**, 14 (2018).

[20] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).

[21] S. R. White, R. L. Martin, J. Chem. Phys. **110**, 4127 (1999).

[22] G. K.-L. Chan, S. Sharma, Annu. Rev. Phys. Chem. **62**, 465-481 (2011).

[23] E. Xu, M. Uejima, S. L. Ten-no, Phys. Rev. Lett. **121**, 113001 (2018).

[24] N. S. Blunt, A. J. W. Thom, C. J. C. Scott J. Chem. Phys. **15**, 3537-3551 (2019).

[25] G. H. Booth, A. J. W. Thom, A. Alavi, J. Chem. Phys. **131**, 054106 (2009).

[26] D. Cleland, G. H. Booth, A. Alavi, J. Chem. Phys. **132**, 041103 (2010).

[27] J. S. Spencer, N. S. Blunt, W. M. C. Foulkes J. Chem. Phys. **136**, 054110 (2012).

[28] D. M. Cleland, G. H. Booth, A. Alavi, J. Chem. Phys. **134**, 024112 (2011).

[29] S. Ten-no, J. Chem. Phys. **138**, 164126 (2013).

[30] Y. Ohtsuka, S. Ten-no, J. Chem. Phys. **143**, 214107 (2015).

[31] S. Ten-no, J. Chem. Phys. **147**, 244107 (2017).

[32] J. S. Spencer, N. S. Blunt, S. Choi, J. Etrych, M.-A. Filip, W. M. C. Foulkes, R. S. T. Franklin, W. J. Handley, F. D. Malone, V. A. Neufeld, J. Chem. Theory Comput. **3**, 15 (2019).

[33] A. J. W. Thom, V. A. Neufeld J. Chem. Theory Comput. **1**, 15 (2019).

[34] N. S. Blunt J. Chem. Phys. **148**, 221101 (2018).

[35] S. Sharma, A. Holmes, G. Jeanmairet, A. Alavi, C. J. Umrigar, J. Chem. Theory Comput. **13**, 1595 (2017).

[36] Y. Garniron, A. Scemama, P.-F. Loos, M. Caffarel, J. Chem. Phys. **147**, 034101 (2017).

[37] C. Overy, G. H. Booth, N. S. Blunt, J. J. Shepherd, D. Cleland, A. Alavi, J. Chem. Phys. **141**, 244117 (2015).

[38] B. H. Brandow Rev. Mod. Phys. **39**, 771 (1967).

[39] I. Lindgren, J. Morrison, Atomic Many-Body Theory Springer Series in Chemical Physics, Vol. 13; (1982).

[40] N.C. Handy, P. J. Knowles, K. Somasundram, Theor. Chim. Acta **68**, 87 (1985).

[41] F. R. Petruzielo, A. A. Holmes, H. J. Changlani, M. P. Nightingale, C. J. Umrigar, Phys. Rev. Lett. **109**, 230201 (2012).

[42] N. S. Blunt, S. D. Smart, J. A. F. Kersten, J. S. Spencer, G. H. Booth, A. Alavi, J. Chem. Phys. **142**, 184107 (2015).

[43] M. L. Leininger, W. D. Allen, H. F. Schaefer III, and C. D. Sherrill, J. Chem. Phys. **112**, 9213 (2000).

[44] R. J. Bartlett, I. Shavitt, G. D. Purvis, J. Chem. Phys. **71**, 281 (1979).

[45] M. Kertesz, R. Hoffmann, Solid State Comun. 47, 97 (1983).

[46] S. Kivelson, O. L. Chapman, Phys. Rev. B 28, 7236 (1983).

[47] K. N. Houk, P. S. Lee, M. Nendle J. Org. Chem. 66, 5517-5521 (2001).

[48] Gorman, C. G., Marder, S. R. Proc. Natl. Acad. Sci. U.S.A. 90, 11297 (1993).

[49] Michael Bendikov, Hieu M. Duong, Kyle Starkey, K. N. Houk, Emily A. Carter, Fred Wudl J. Am. Chem. Soc. 126 , 7416 - 7417 (2004).

[50] J. Hachmann, J. J. Dorando, G. K.-L. Chan, J. Chem. Phys. **127**, 134309 (2007).

[51] A. Grüneis, S. Hirata, Y.-y. Ohnishi, S. Ten-no, J. Chem. Phys. **146**, 080901 (2017).

[52] P.-F. Loos, B. Pradines, A. Scemama, J. Toulouse, E. Ginner, J. Phys. Chem. Lett. **10**, 2931 (2019).

3 Epstein-Nesbet corrections of the initiator error

3.1 Introduction

Monte Carlo computer simulations are now a standard tool in scientific fields such as condensed-matter physics, including surface physics and applied physics problem (metallurgy, diffusion, and segregation, etc.), polymer statistics and field theory. The range of different physical phenomena which can be explored using Monte Carlo methods is exceedingly broad. With the increasing ability of this method to deal with quantum-mechanical problems such as spin systems or many-fermion systems, it will become useful for other questions in the fields of elementary-particle and nuclear physics as well. QMC methods are a vital tool in the study of many-body systems, regularly providing the most accurate and reliable results for correlated fermionic systems of interest.

MSQMC, along with FCIQMC has became a tool for the simulation of the underlying imaginary-time Schrödinger equation of the interacting Hamiltonian. In the limit of infinite walkers MSQMC and FCIQMC provide accurate FCI energies for systems that were out of reach in the past due to limited memory resources of the Neumann architecture based computers.

The number of possible excitations (and so does the number of walkers required to represent the wave function) grows combinatorially with system size and becomes intractable when one wants to calculate more than a few dozens of active orbitals. During the initial stages of the calculation the number of walkers grow very rapidly and it reaches a plateau upon convergence. The height of the plateau sets the minimum memory requirements of such a calculation.

The main reason for the large number of walkers in equilibrium state can be traced back to determinants with very low occupation. In other words, a large portion of the determinants have one or at most two walkers residing on them and although they hardly contribute to the energy they consume significant amount of memory.

In order to solve this problem the initiator adaptation was introduced in 2010 by Ali Alavi[1]. This scheme is in a sense analogous to the branching Monte Carlo algorithm where the time average of quantities of interest are calculated as a product of fluctuating weights. Usually, there are only a few dominant terms and the small terms are equally expensive to compute, but play no significant role in the average. In branching Monte Carlo this solved by performing multiple simultaneous walks. One evolves a collection of walkers from one generation to the next, and the key idea is to eliminate as the process progresses the light-weight walkers that produce relatively small contributions to the time average. To keep the number of walkers reasonably constant, heavy-weight walkers are duplicated and the clones are subsequently evolved quasi-independently.

Within the initiator approximation a very similar operation is performed by dividing the determinants into two groups. As explained on Figure 3.1. initiator determinants are allowed to spawn freely, while non-initiator determinants can only spawn onto already occupied determinants.

The initiator threshold is a predefined calculation parameter and determinants with walker whose number of walkers exceeds this threshold are called initiator determinants and the others are called non-initiators. By doing so, a truncation is imposed on the Hamiltonian *i.e.* the Hamiltonian matrix element between non-initiators and unoccupied determinants is set to zero. The implementation of such rules in the walker dynamics leads to significant decrease in the number of walkers as non-initiator determinants are not allowed to further increase the size of the sampled space. In the limit of infinite walkers all determinants become initiators and the error caused by the approximation is

zero. In most calculations the initiator threshold is chosen to be 3. In some cases though even higher thresholds provide surprisingly accurate results when appropriate corrections are applied.[2] Among these corrections are the Davidson:

$$E_D^{(+Q)} = \frac{L_2^{non-i}}{1 + L_2^i} E_{corr} \tag{3.1}$$

, Pople:

$$E_P^{(+Q)} = [1 - \frac{2}{N_e}] \frac{L_2^{non-i}}{1 + L_2^i} E_{corr} \tag{3.2}$$

and Meissner type corrections:

$$E_M^{(+Q)} = [\frac{(N_e - 2)(N_e - 3)}{N_e(N_e - 1)}] \frac{L_2^{non-i}}{1 + L_2^i} E_{corr} \tag{3.3}$$

, with N_e as the number of active electrons and L_2^{non-i} is the L_2 norm of the non-initiator space and L_2^i is the L_2 norm of the initiator space.

As shown on Figure 3.2. the initiator error increases with increasing initiator threshold as less and less determinants can spawn freely and the effect of the truncation becomes severe.

Second order perturbation theory has been successfully applied to non-linearly parameterized wave functions by S. Sharma[18] to improve Variational Monte Carlo (VMC) energies.

The main idea behind this approach is to calculate the wave function using a relatively small MPS bond dimension in the density matrix renormalization group algorithm and subsequently improve the obtained energy with second order perturbation theory.

Stochastic perturbation theory proved to be an efficient alternative to capture most of the correlation energy missing from the zeroth order wave function even for matrix product state (MPS)[19] wave functions for which, compared to linearly parameterized wave functions several difficulties arise. Among these problems is the definition of the zeroth order Hamiltonian and the action of $\hat{Q}$ that ensures that the first order wave function is orthogonal to the zeroth order wave function. Notwithstanding these issues the derivation of the second order energy is possible, albeit using some kind of approximations and the resulting algorithm seems to recover more than 90% of the correlation energy for the C_2 molecule that was missing from the zeroth order state. Along with this work stochastic perturbation theory was successfully applied to the DMRG algorithm to calculate the total energies of Cr_2 molecule to demonstrate the applicability of the method to large active spaces.[20]

Based on these ideas N. S. Blunt[21] proposed a second order perturbative correction to correct the initiator error that appears in i-FCIQMC calculations. A second order correction is indeed possible, and that it can be built primarily from the information discarded in applying initiator criteria and therefore already present in the simulation. Remarkable improvements were achieved only with a slight modification of the algorithm and the correction was proven to correct a large portion (96%-%110) of the initiator error for the Hubbard model, the uniform electron gas and other molecular systems without significantly increasing calculation costs.

In the following section, I start by providing a brief overview of Epstein-Nesbet perturbation theory and I present a derivation to obtain the formula for the third order correction of the initiator error to further improve the accuracy of the results. With the calculation of second and third order perturbative corrections one can obtain energies of near FCI accuracy without having to increase the number of walkers during simulation.

In the numerical results section, the accuracy and applicability of these corrections are demonstrated for several molecular systems. I calculate i-MSQMC energies on the potential energy curve of the N_2

molecule in several points and I compare my results obtained using second and third order perturbation with MP2 and CCSD(T). Moreover, I investigate the convergence of both systematic and random errors in the calculations and propose a new methodology to mitigate the stochastic noise in the calculations.

3.2 Theory

I provide here a brief overview of the underlying framework of Epstein-Nesbet perturbation theory to correct a zeroth order wave function as it is done in selected configuration interaction (SCI) approaches. These are another important class of methods for obtaining FCI-level accuracy in challenging systems by drawing a clear distinction between those electronic configurations which make an important contribution to the total electronic wave function and those which do not. SCI methods appeared and had been researched many years ago[3, 4] and they have seen a particular renewal of interest in recent years.[12, 10, 11, 13] SCI usually consists of the following stages:

1. An initial configuration space is chosen. If only a single root is required then the RHF wave function suffices.
2. In the following variational stage, a subspace of important determinants is generated. These important determinants may be selected upon their energy lowering at second order of perturbation theory or their coefficient in the first-order wave function.
3. Next, the Hamiltonian eigenvalue problem is solved in this subspace to obtain a zeroth-order energy and wave function. The diagonalization of the CI matrix is performed in a subspace of strongly interacting determinants.
4. To improve the zeroth-order energy perturbative correction is calculated, often of the second-order Epstein-Nesbet type to estimate the energy contribution of the neglected configurations.

This procedure should allow a progressive construction of the zeroth-order wave function as its dimension increases. The variational space is expanded until one reaches a stable result and/or the practical/financial limitis of the computer.

In perturbation theory the Hamiltonian operator is partitioned into a zeroth order operator and the perturbation:

$$\hat{H} = \hat{H}_0 + \hat{V} \tag{3.4}$$

When the $\hat{H}_0$ is a simple sum of monoelectronic operators the partitioning is called Moller-Plesset partitioning. On the other hand, when when one employs the Epstein-Nesbet definition[14, 15, 16] of $\hat{H}_0$, namely, that it is constructed exactly in the variational subspace (ν) and the diagonal elements are accumulated from the secondary space $\hat{H}_0$ has the following form:

$$\hat{H}_0 = \sum_{ij\in\nu} H_{ij}|D_i\rangle\langle D_j| + \sum_{a\notin\nu} H_{aa}|D_a\rangle\langle D_a| \tag{3.5}$$

, where determinants in the variational subspaces are denoted by $|D_i\rangle$ and $|D_j\rangle$ and the rest of the space is spanned by determinants denoted by $|D_a\rangle$.

By standard perturbation theory it is possible to express the second order energy that would, in a sense take into account the interaction of the zeroth order wave function with the determinants that are excluded from the variational space:

$$\Delta E_2 = \sum_{a\notin\nu} \frac{(\sum_{i\in\nu} H_{ai}c_i)^2}{E_0 - H_{aa}} \tag{3.6}$$

The PT corrected root provides an approximation to the FCI value and the convergence of the algorithm depends mainly on the choice of the selection algorithm and the selection threshold.

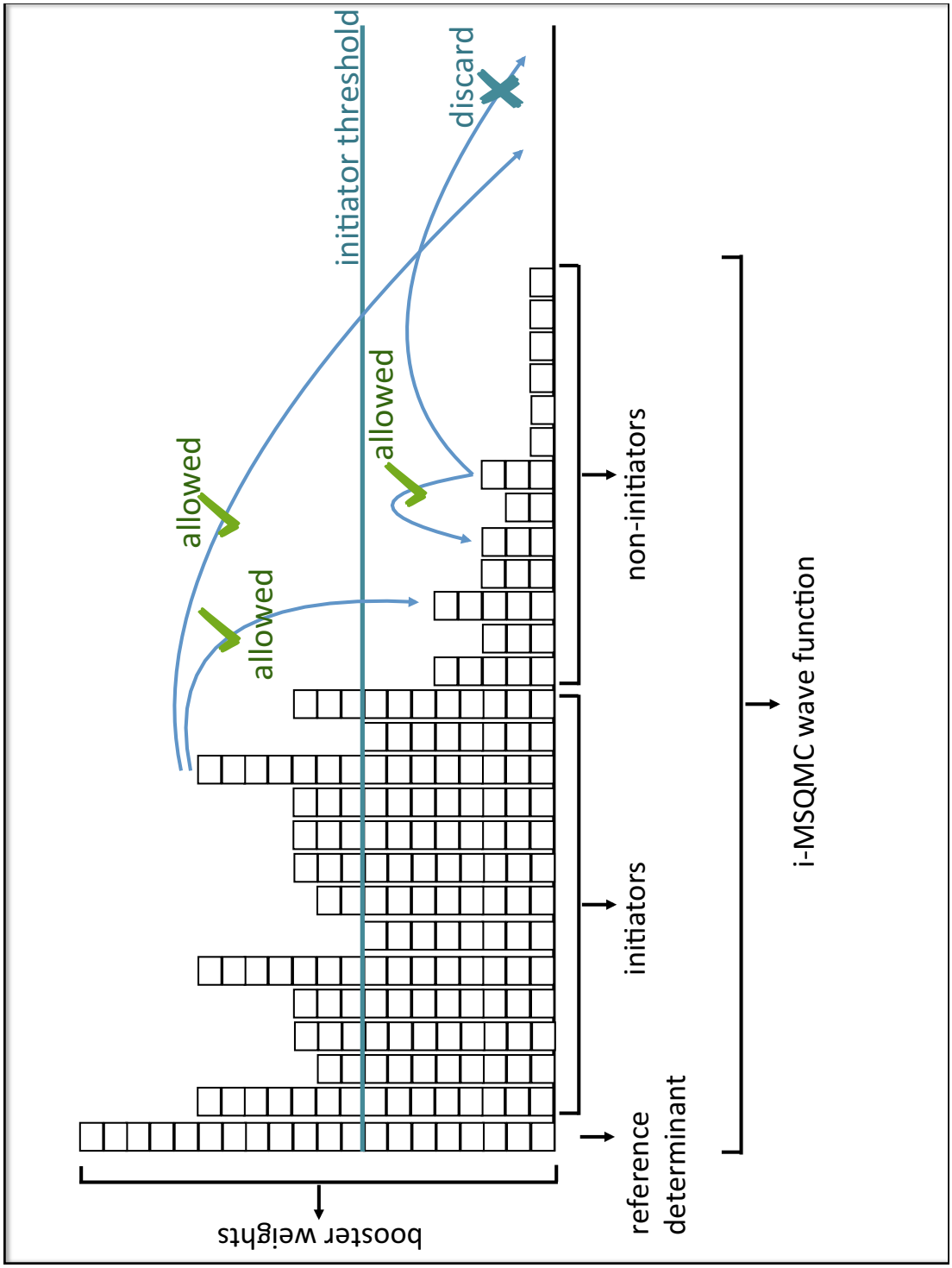


Figure 3.1: Schematic representation of the initiator adaptation in which initiators are allowed to spawned freely and spawning events from non-initiator determinants onto unoccupied determinants are not allowed.

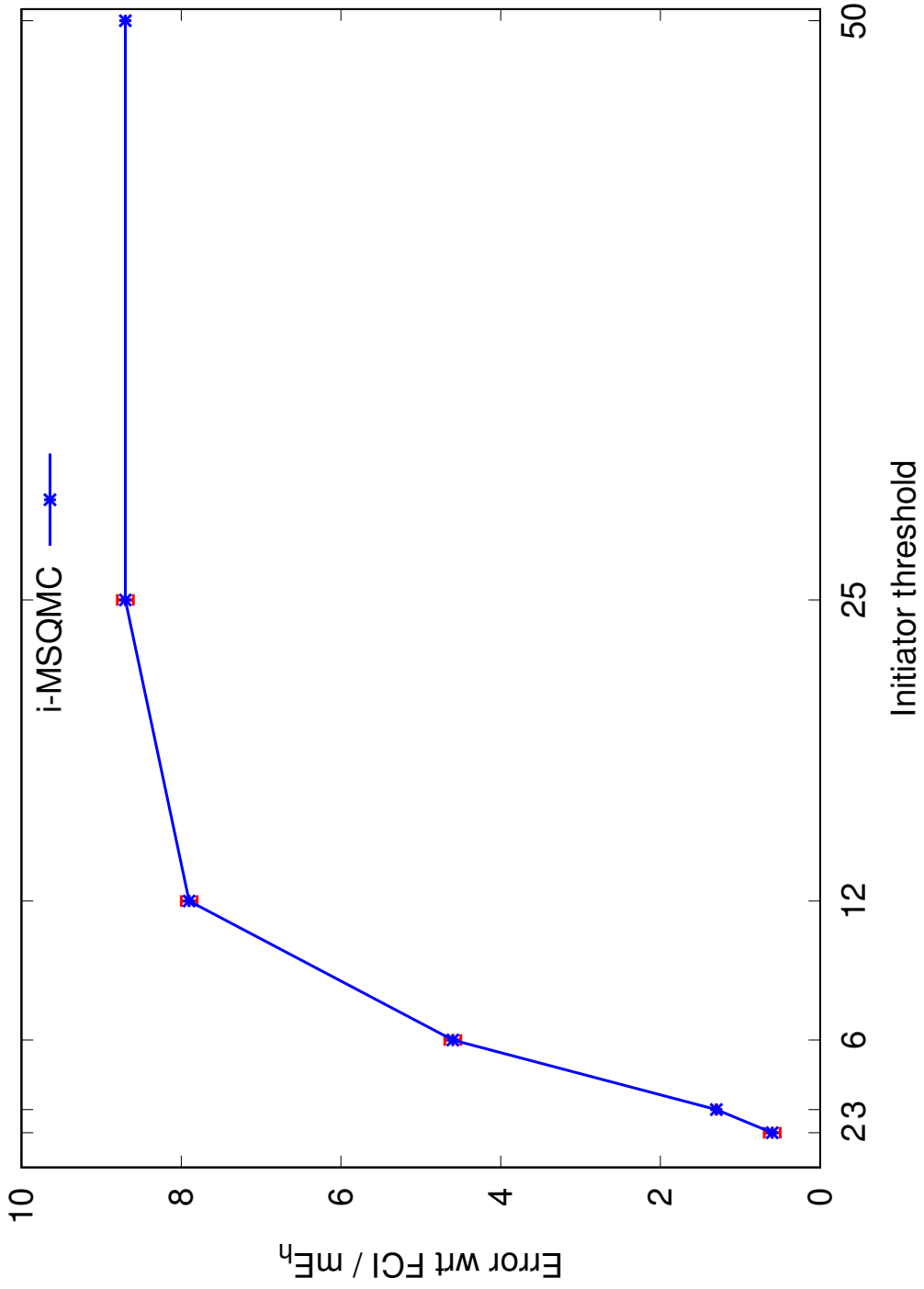


Figure 3.2: The initiator error with respect to the initiator threshold. i-MSQMC calculation was carried out on a water molecule in cc-pVDZ basis with frozen core orbitals using 250 booster weights. The imaginary time length was 1000 a.u. and the integration of the energy from 60 a.u.

3.3 Second and third order energy corrections

The MSQMC algorithm, in the limit of infinity walkers, samples the exact FCI wavefunction, however in order to reduce computational costs approximations (truncations) are inevitable. Initiator-MSQMC (i-MSQMC) truncates the sampled space restricting spawning events from non-initiators. Therefore it is possible to treat the truncated i-MSQMC wavefunction as the variational space in the preceding theoretical introduction.

The variational energy (imaginary time dependent) is available from the i-MSQMC wavefunction ($\psi(\tau)$) and can be calculated using the usual Rayleigh quotient:

$$E_{var}(\tau) = \frac{\langle\psi(\tau)|\hat{H}|\psi(\tau)\rangle}{\langle\psi(\tau)|\psi(\tau)\rangle} \tag{3.7}$$

The second order energy correction can be calculated using the first order wavefunction ($\psi^{(1)}(\tau)$):

$$E^{(2)}(\tau) = \langle\psi(\tau)|\hat{V}|\psi^{(1)}(\tau)\rangle \tag{3.8}$$

The first order wavefunction is obtained from the intreracting space of the variational wavefunction:

$$|\psi^{(1)}(\tau)\rangle = \sum_{a\notin\nu(\tau)} c_a(\tau)|\phi_a\rangle \tag{3.9}$$

The corresponding coefficients ($c_a(\tau)$) are calculated using the variational subspace:

$$|\psi^{(1)}(\tau)\rangle = \sum_{a\notin\nu(\tau)} \frac{\sum_{i\in\nu(\tau)} H_{ia}c_i(\tau)}{E_{var}(\tau) - H_{aa}}|\phi_a\rangle \tag{3.10}$$

, where $c_i(\tau)$ is the imaginary time dependent CI coefficient of the ith determinant in the variational space, calculated using the walker population on the given determinant:

$$c_i(\tau) = \frac{N_i(\tau)}{L_2^{walk}(\tau)} \tag{3.11}$$

, with the $L_2(\tau)$ norm defined as:

$$L_2^{walk}(\tau) = \sqrt{\sum_{j\in\nu(\tau)} N_j(\tau) * N_j(\tau)} \tag{3.12}$$

and H_{ia} and H_{aa} are off-diagonal and diagonal Hamiltonian matrix elements respectively. The variational energy ($E_{var}(\tau)$) can be calculated in many ways. In this work I use the expression of the instantaneous energy introduced above.

The first order wavefunction can be effectively sampled during the spawning step using the walker dynamics. Spawning from non-initiator determinants to unoccupied determinants are discarded. In other words, although these determinants interact with the variational space they are being excluded. This fact can be exploited to calculate an approximate first order wave function to accumulate the second order energy correction.

Using Equation 3.8. and Equation 3.10. the second order energy can be expressed as:

$$E^{(2)}(\tau) = \sum_{a\notin\nu(\tau)} \frac{(\sum_{i\in\nu(\tau)} H_{ia}c_i(\tau))^2}{E_{var}(\tau) - H_{aa}} \tag{3.13}$$

At every micro step, walkers spawned from non-initiator determinants onto unoccupied determinants are discarded and the discarded quantity can be expressed as:

$$S_a(\tau) = -\Delta\tau \sum_{i \in \nu(\tau)} H_{ia} N_i(\tau) \tag{3.14}$$

The instantaneous second order Epstein-Nesbet correction in the initiator adaptation can therefore be computed from the discarded walker population:

$$E^{(2)}(\tau) = \frac{1}{(L_2^{walk}(\tau))^2 (\Delta\tau)^2} \sum_{a \notin \nu(\tau)} \frac{(S_a(\tau))^2}{E_{var}(\tau) - H_{aa}} \tag{3.15}$$

The problem with this approach is that the nominator contains the square of a quantity sampled from the same walker distribution and this known to cause biased results [17]. Here, I use an approach for the sampling of unbiased corrections within the walker dynamics, which requires only small computational overheads. This is achieved via an independent replica population of walkers in the dynamic, sampled alongside the original population and accumulating the correction from contributions from independent populations:

$$E^{(2)}(\tau) = \frac{1}{(L_2^{walk}(\tau))^2 (\Delta\tau)^2} \sum_{a \notin \nu(\tau)} \frac{S_a^{(1)}(\tau) S_a^{(2)}(\tau)}{E_{var}(\tau) - H_{aa}} \tag{3.16}$$

In our implementation, the replicas are propagated with independent random number seeds, spawning, death, and annihilation events are then applied to each population independently and simultanous removal events are registered at every micro step. The two simulations are therefore completely uncorrelated, apart from sharing the same initial conditions, from which they quickly decorrelate and the resulting correction is free from systematic error. Although this effectively doubles the computational resources required per iteration, there are two independent estimates of all quantities of interest, and so by using this additional information, the random errors in these quantities are reduced by a factor of 2. This results in the calculation requiring only half the time. Computational overheads therefore only effectively result from increased memory requirements from the additional replica population. Using the same derivation, I express the third order perturbative energy correction as:

$$E^{(3)}(\tau) = \frac{1}{(L_2^{walk})^2 (\Delta\tau)^2} \sum_{a,b \notin \nu(\tau)} \frac{S_a^{(1)}(\tau) H_{ab} S_b^{(2)}(\tau)}{(E_{var}(\tau) - H_{aa})(E_{var}(\tau) - H_{bb})} \tag{3.17}$$

The value of the instantaneous second and third order corrections fluctuates over the imaginary time evolution, therefore in order to get accurate results the values are integrate over a certain time interval:

$$E^{(2)} = \langle E^{(2)}(\tau) \rangle_\tau = \frac{1}{\tau_2 - \tau_1} \int_{\tau_1}^{\tau_2} E^{(2)}(\tau) d\tau \tag{3.18}$$

$$E^{(3)} = \langle E^{(3)}(\tau) \rangle_\tau = \frac{1}{\tau_2 - \tau_1} \int_{\tau_1}^{\tau_2} E^{(3)}(\tau) d\tau \tag{3.19}$$

3.4 Evaluation of the stochastic error

In this section I wish to shed light on the computational efficiency of the Epstein-Nepsbet corrections and discuss some of the challenges caused mainly by stochastic noise.

According to my discourse in the preceding previous section, in order to obtain unbiased energies one needs to initialize and run two independent calculations. This doubles the required memory space, however as the number of samples is also increased one can attain the same accuracy with less steps using two independent walker populations. Propagator based QMC calculations are usually limited by the amount of available CPU time and not the memory, therefore replica sampling does not put significant constraints on the applicability of the method. As it was pointed out in N. S. Blunt's

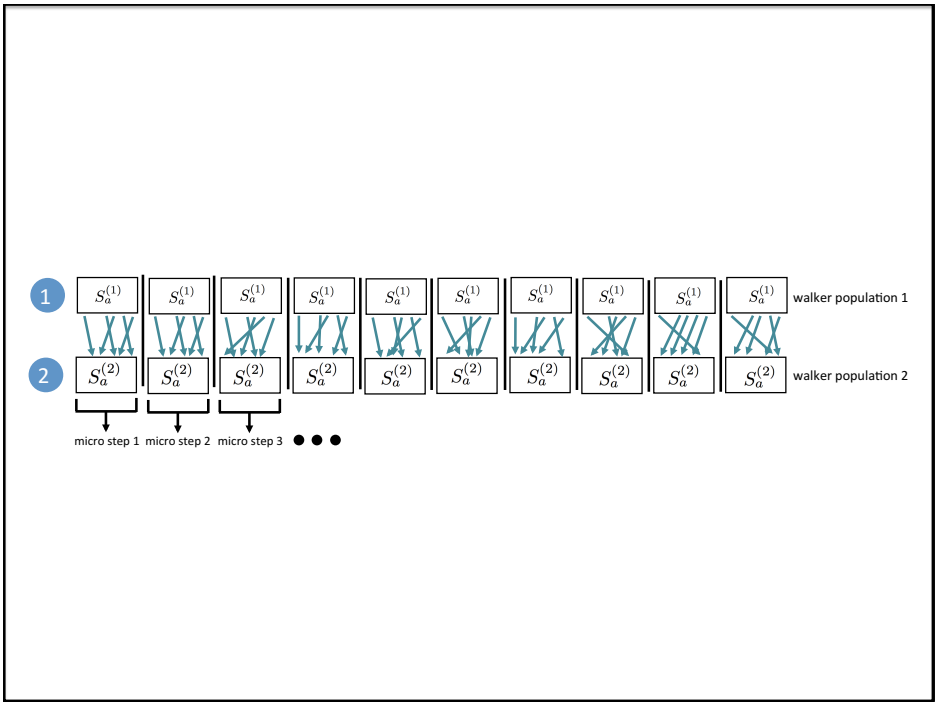


Figure 3.3: Schematic representation of how the second order Epstein-Nesbet correction is accumulated in an i-MSQMC calculation. In every micro step a contribution from S_a is recorded when the same removal event occurs on both replicas.

work[21] the stochastic noise is of $E^{(2)}$ can be much more sever than that of the i-MSQMC energy. This inefficiency is caused by the fact that the in every micro step a contribution from S_a is recorded when the same removal event occurs on both replicas (Figure 3.3.) and these events are extremely rare.

Clearly, when the size of the system under consideration grows, this inefficiency becomes more sever as the frequency of a removal event is proportional to the inverse of the size of the connected space of a given determinant. In our current implementation every walker is allowed to spawn once from a given determinant and this determines the probability of spawning a new walker or removing a walker.

The only way to increase the efficiency of the sampling is to let walkers spawn more than once in a time step. The multi-spawning[22] algorithm achieves this and it provides very accurate energies using larger time steps. This approach substantially reduces statistical noise on perturbative corrections to initiator error, which improves the accuracy of FCIQMC but which can suffer from significant noise in the original scheme. In the following section, I demonstrate how to implement a similar approach into our i-MSQMC framework and sample $E^{(2)}$ and $E^{(3)}$ in a more efficient manner.

3.5 Proposal to mitigate stochastic error

In this section I present an idea to mitigate the stochastic noise of $E^{(2)}$ by collecting samples of S_a from numerous micro steps and accumulating much more samples per one calculation.

In the new sampling algorithm $E^{(2)}$ is calculated from more samples of S_a by storing the removed walkers for multiple micro steps in order to achieve increased number of coinciding indices of S_a. The implementation of the new method is visualized on Figure 3.4.

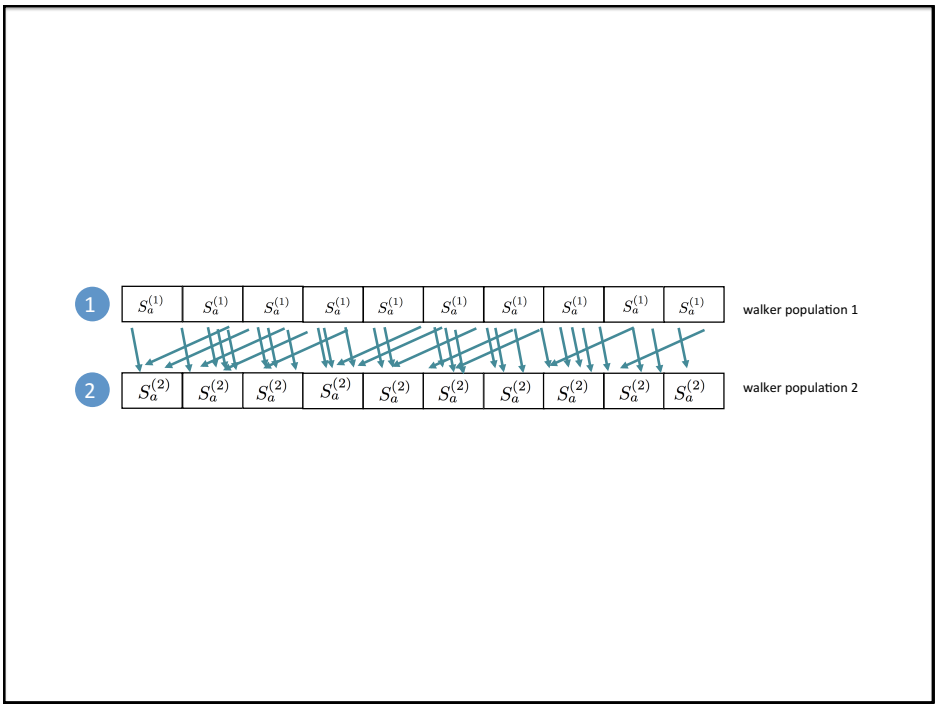


Figure 3.4: Schematic representation of how the second order Epstein-Nesbet correction is accumulated in the improved algorithm. In this case $N_m = 10$.

In Equation 3.20. the number of terms (N_t) in the summation of $E^{(2)}$ depends on the number of micro steps (N_m) for which the samples are collected.

$$E^{(2)}(\tau) = \frac{1}{N_m} \frac{1}{(L_2^{walk})^2 (\Delta\tau)^2} \sum_{a \notin \nu}^{N_t(N_m)} \frac{S_a^{(1)}(\tau) S_a^{(2)}(\tau)}{E_{var}(\tau) - H_{aa}} \tag{3.20}$$

$$E^{(3)}(\tau) = \frac{1}{N_m^2} \frac{1}{(L_2^{walk})^2 (\Delta\tau)^2} \sum_{a,b \notin \nu}^{N_t(N_m)} \frac{S_a^{(1)}(\tau) H_{ab} S_b^{(2)}(\tau)}{(E_{var}(\tau) - H_{aa})(E_{var}(\tau) - H_{bb})} \tag{3.21}$$

In order to understand the connection between N_t and N_m, let me investigate the size of the secondary space (a):

$$a = \{\Psi_{a_1}, \Psi_{a_2} ... \Psi_{a_i}\} \tag{3.22}$$

Ψ_{a_i} appears in a if $p_{rm}^{(i)} > \frac{1}{N_m}$, where $p_{rm}^{(i)}$ is the probability of a removal on Ψ_{a_i} in a given micro step. N_t or the number of terms in the summation of $E^{(2)}$ equals to the dimension of a:

$$N_t = |a| \tag{3.23}$$

and the standard deviation of $E^{(2)}$ is proportional to the inverse square root of the number of samples:

$$\sigma_{E^{(2)}} \sim \frac{1}{\sqrt{N_t(N_m)}} \tag{3.24}$$

In order to demonstrate the effect of accumulating the perturbative correction from numerous steps I performed a calculation on the nitrogen molecule using the cc-pVDZ basis set and I measured the standard deviation of the second order energy with respect to the number of steps I used to accumulate the contributions. My results are presented on Figure 3.5. The initial calculation is essentially identical to the original calculation in which only one set of data is used to calculated $E^{(2)}$. Then, I gradually increased N_m to decrease the standard deviation and to obtain a less noisy result. Initially, when changing N_m from 1 to 10 and then to 100 the deviation decreases significantly, however further increasing N_m does not improve the results in a palpable manner. There is a slight overhead for accumulating more terms, therefore I recommend a reasonable value for N_m when performing actual calculations. Henceforth, I used 5 steps for the accumulation to obtain accurate energies for the N_2 molecule and the chromium dimer.

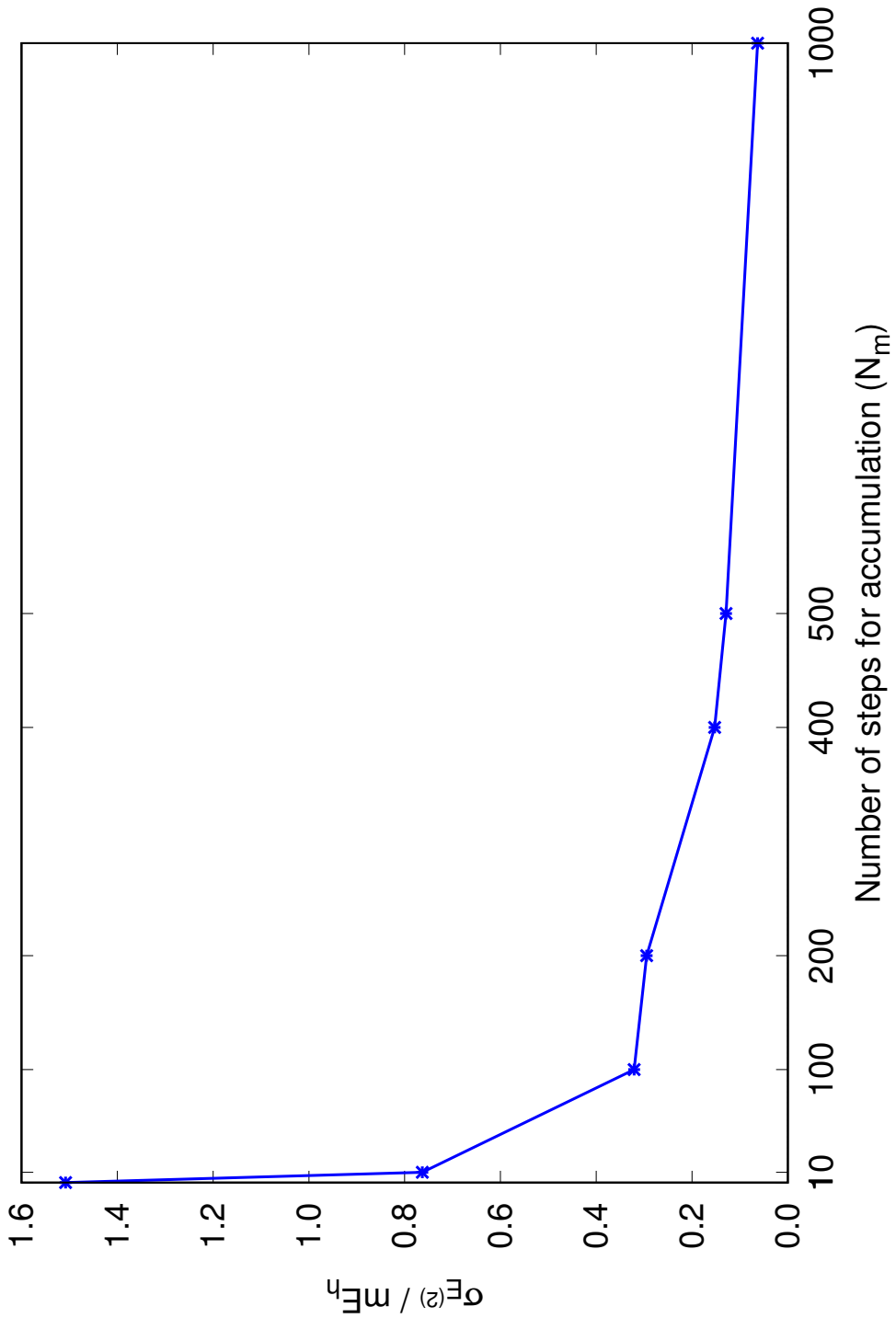


Figure 3.5: The standard deviation of the second order energy vs N_m. The calculation was carried out on a nitrogen molecule with $N_b = 250$, $\tau = 100$ a.u. and $\tau_1 = 60$ a.u. using the cc-pVDZ basis set.

3.6 Investigation on accuracy and memory requirements

When the initiator threshold is fixed, there are two choices to obtain more accurate results. One way is to increase the number of booster weights (e.g. the number of walkers) in order to let more determinants become initiators and converge onto the FCI wave function. On the other hand, based on the derivation in the previous section it is possible to calculate second and third order Epstein-Nesbet type corrections. In the following section a detailed investigation is carried out in connection with the memory requirements of both approaches. The main problem with the first solution is that as the initiator threshold decreases the number of determinants with small occupation grows severely and it could eventually fill up the available memory space without significantly increasing the accuracy of the result.

The number of determinants with various occupation levels with respect to the initiator threshold is presented on Figure 3.6. Clearly, as the initiator threshold is decreased the number of determinants with $N_w = 1$ increases rapidly unlike determinants with larger occupation.

What follows next is a numerical assessment on the water molecule in cc-pVDZ basis focusing on the memory requirements of the various approaches to correct the initiator error.

First of all, a simple comparison was made on Figure 3.7. to evaluate the accuracy of i-MSQMC with varying the initiator threshold. It is worth comparing the right and the left slope of the curve and come to a realization that the initiator error decreases in a much more rapid fashion for large values (small number of determinants) compared to small values (large number of determinants). The most worrying finding was that although the accuracy of the energy does not change significantly change when switching from $I_t = 3$ to $I_t = 2$ the number of sampled determinants grows approximately double.

The Epstein-Nesbet corrections meant to solve this problem and provide much more accurate energies with modest memory requirements by sampling the secondary space whose sampled dimension is smaller than that of the unoccupied space.

In order to understand the reason for implementing an Epstein-Nesbet perturbative correction I have carried out an investigation on Figure 3.8. comparing the number of determinants that are required to attain a certain accuracy. With varying N_t it is possible to control the number of determinants incorporated in an i-MSQMC calculation. This is extremely useful when memory resources are limited and one has to truncate the sampled determinantal space. As it is clear from the results of Figure 3.8. calculating the Epstein-Nesbet perturbative corrections seems to be a much more memory efficient way than increasing the size of the sampled space by chosing smaller initiator thresholds. The memory requirements of the second order correction are the same as the third order as they are calculated from the same set of determinants. For the calculation of Figure 3.8. while i-MSQMC with approximately 2000 determinants exhibits more than 4 mEh error by sampling a few determinants (70) from the secondary space the absolute value of the error of $E^{(3)}$ shrinks below 1 mEh. The same tendency was observed during this calculation as on Figure 3.7. namely that the accuracy of the energy (even the perturbative energies) increases very slowly with the incorporation of determinants of low population.

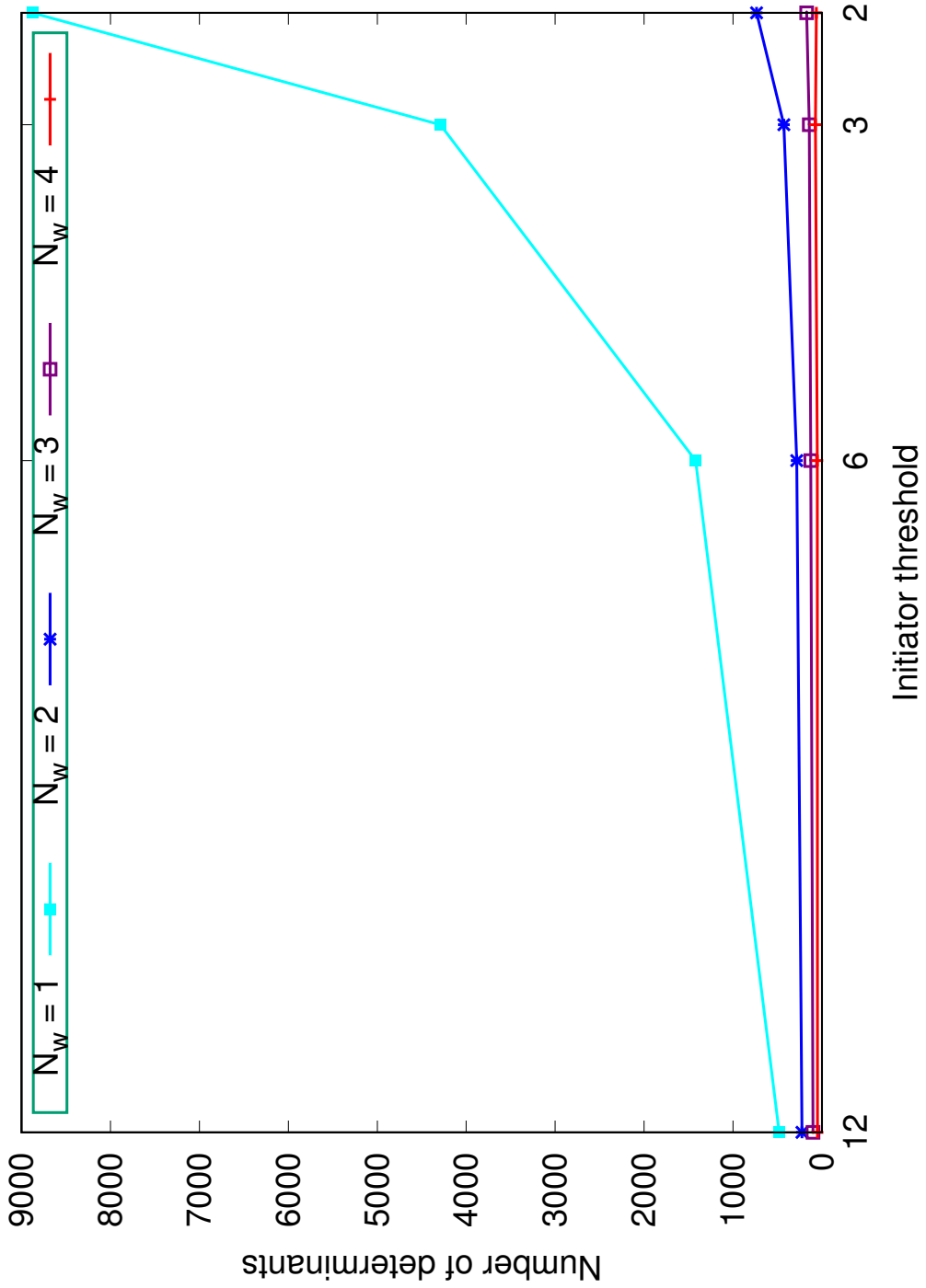


Figure 3.6: The number of determinants of certain (1-4) number of walkers with respect to the initiator threshold. The calculation was carried out on a water molecule with $N_b = 250$, $\tau = 1000$ a.u. and $\tau_1 = 600$ a.u. using the cc-pVDZ basis set.

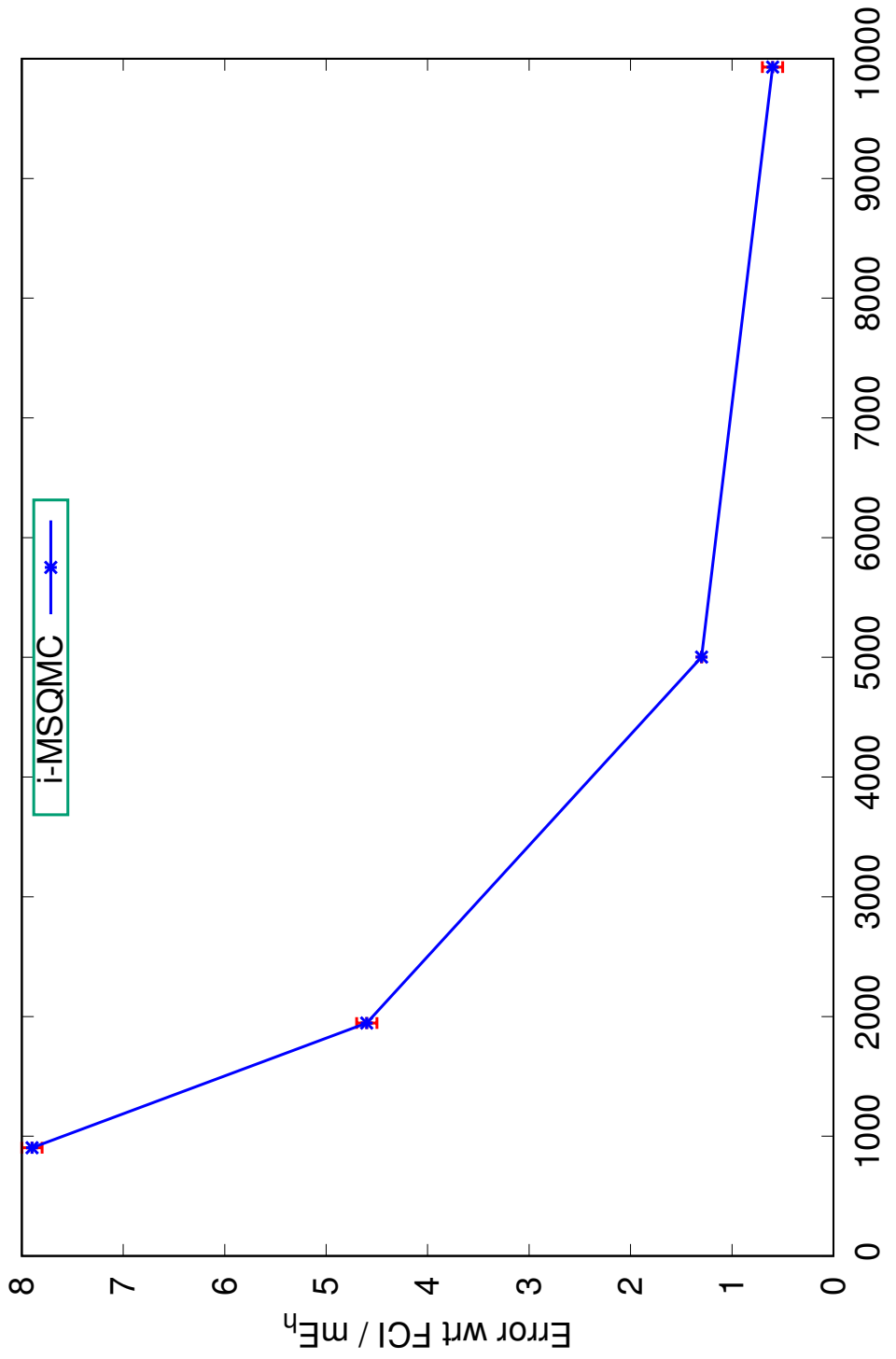


Figure 3.7: The accuracy of an i-MSQMC calculation with respect to the number of determinants sampled by the initiator adaptation. The calculation was carried out on a water molecule with $N_b = 250$, $\tau = 1000$ a.u. and $\tau_1 = 600$ a.u. using the cc-pVDZ basis set.

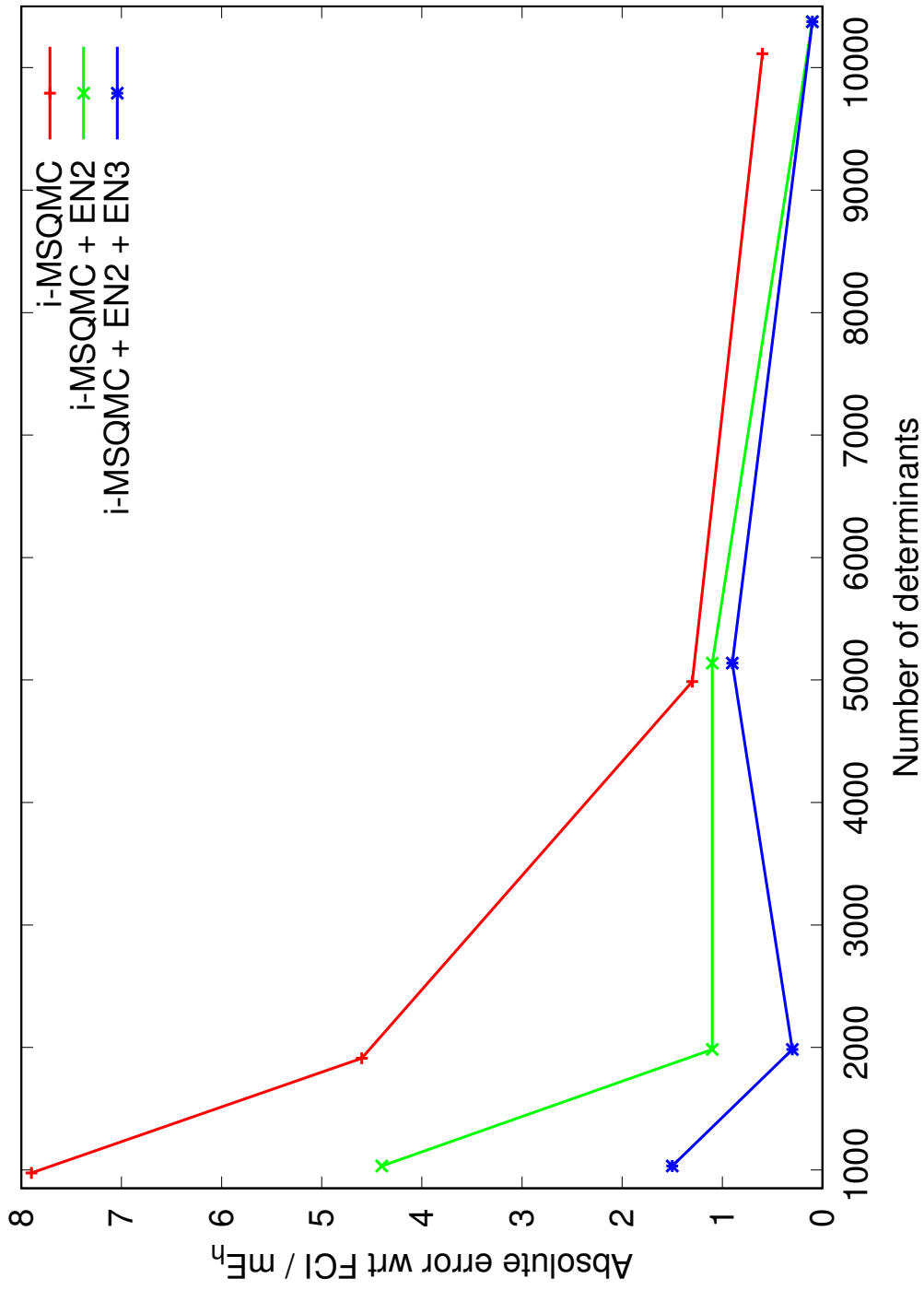


Figure 3.8: Investigation results of the absolute value of the error in energy with respect to the FCI energy. The number of determinants (X-axis) are varied using the initiator threshold in four points (12, 6, 3 and 2). The number of determinants for perturbative results (blue and green lines) are calculated as the sum of the number of determinants in the i-MSQMC wave function and the number of determinants in the secondary space. This calculation was carried out on a water molecule with $N_b = 250$, $\tau = 1000$ a.u. and $\tau_1 = 600$ a.u. using the cc-pVDZ basis set. The frozen-core approximation was used for this calculation.

3.7 Definition of the zeroth order wave function

It might not be apparent for the first sight that the definition of the zeroth order space is loose in a sense that the wave function sampled by the i-MSQMC dynamics changes throughout imaginary time. Even though in a converged state the value of the energy does not change significantly there are determinants constantly leaving and entering the space. This happens because walkers die in the death-cloning step and they get remove while others are newly sampled and are inserted into the sampled wave function. Therefore a constant fluctuations is observed at the end of the calculation and it is not straightforward to define a zeroth order wave function in such a space (the truncation is on the Hamiltonian, not on the space). The main problem with this is that the exact sampling of $E^{(2)}$ and $E^{(3)}$ requires a well-defined zeroth order wave function, otherwise one only gets approximate values of these quantities.

Another deficiency arises from the fact that contributions that should arise from the connection between initiator and non-occupied determinants are never sampled in $E^{(2)}$ and $E^{(3)}$, because these spawning events are allowed and these populations are never removed. In the second article[22] of N. S. Blunt a detailed investigation was given on this and it turned out that it is possible to derive a more rigorous estimation for the second order energy starting from the perturbative definition of the first order wave function.

In the following discourse I present the equations governing this more rigorous approach to the perturbative corrections. Perturbation theory has been applied to DMRG with the following definition of H_0:[20]

$$\hat{H}_0 = \hat{P}E_0\hat{P} + \hat{Q}\hat{H}_d\hat{Q} \tag{3.25}$$

, with the following definition of E_0:

$$\hat{H}_0|\psi_0\rangle = E_0|\psi_0\rangle \tag{3.26}$$

First of all, the second order perturbation energy is usually calculated from the first order wave function as the following expectation value:

$$E^{(2)} = \langle\psi_0|\hat{V}|\psi^{(1)}\rangle. \tag{3.27}$$

where ψ_0 is the zeroth order wave function. It the previous chapter this was chosen to be the HF solution, however in order to calculate the second order correction within the initiator approximation in this case $|\psi_0\rangle$ is the i-MSQMC wave function.

$|\psi^{(1)}\rangle$ is expressed using the following formulae from perturbation theory:

$$|\psi^{(1)}\rangle = \frac{1}{H_0 - E_0}\hat{Q}\hat{V}|\psi_0\rangle \tag{3.28}$$

therefore the second order can be expressed as:

$$E^{(2)} = \langle\psi_0|\hat{V}\frac{1}{H_0 - E_0}\hat{Q}\hat{V}|\psi_0\rangle \tag{3.29}$$

. Next, I substitute the definition of the zeroth order hamiltonian ($\hat{H} = \hat{V} + \hat{H}_0$) into the energy expression and obtain:

$$E^{(2)} = \langle\psi_0|[\hat{H} - \hat{H}_0]\frac{1}{H_0 - E_0}\hat{Q}[\hat{H} - \hat{H}_0]|\psi_0\rangle \tag{3.30}$$

Making use of the fact that $|\psi_0\rangle$ is an eigenfunction of H_0 I arrive at the following equation:

$$E^{(2)} = \langle\psi_0|[\hat{H} - \hat{E}_0]\frac{1}{H_0 - E_0}\hat{Q}[\hat{H} - \hat{E}_0]|\psi_0\rangle \tag{3.31}$$

Treating the diagonal and the off-diagonal elements separately:

$$E^{(2)} = \langle\psi_0|[\hat{H}_x + \hat{H}_d - \hat{E}_0]\frac{1}{H_0 - E_0}\hat{Q}[\hat{H}_x + \hat{H}_d - \hat{E}_0]|\psi_0\rangle \tag{3.32}$$

Using the assumption that $(\hat{H} - E_0)||\psi_0\rangle$ is zero in the space of the occupied determinants, the final formula for $E^{(2)}$ is:

$$E^{(2)} = -\langle\psi_0|\hat{H}_x\frac{1}{H_0 - E_0}\hat{Q}\hat{H}_x|\psi_0\rangle - \langle\psi_0|\hat{H}_x|\psi_0\rangle \tag{3.33}$$

During the ITE, the spawned walkers are proportional to the following quantity:

$$S_i = -\Delta\tau\sum_j\langle D_i|\hat{H}_{ij}|D_j\rangle \tag{3.34}$$

, which can be used to express the second order correction:

$$E^{(2)} = \frac{1}{(\Delta\tau)^2}\sum_i\frac{S_i^{(1)}S_i^{(2)}}{E - H_{ij}} - \sum_{ij}C_iH_{ij}C_j \tag{3.35}$$

. Besides, with the addition of the variational energy $E_{var} = \langle\Psi|H_x + H_d|\Psi\rangle$:

$$E_{var} + E^{(2)} = \frac{1}{(\Delta\tau)^2}\sum_i\frac{S_i^{(1)}S_i^{(2)}}{E - H_{ij}} + \sum_i C_iH_iC_i \tag{3.36}$$

I arrive at a formula more suitable for parallel calculations. Compare this expression, with the original expression in Equation 3.15. There are two main differences I find it useful to mention here. First of all, the summation in the first term runs on the spawned walker array, unlike in Equation 3.15. where only the discarded walkers were considered. Note that the first term contains all the terms of Equation 3.15. therefore the original proposal is also strongly related to perturbation theory though in a loose sense. Secondly, there is an additional term from the zeroth order wave function.

3.8 Computational details

The algorithm to sample $E^{(2)}$ and $E^{(3)}$ from the discarded walker populations from non-initiator determinants was implemented within our MSQMC program suite.

Based on the original MSQMC implementation, the algorithm to sample second and third order Epstein-Nesbet corrections to the initiator approximation was implemented in our program suite using the FORTRAN programing language.

The explanation of parameters that are identical to the ones in the preceding chapter regarding stochastic perturbation theory is omitted from the following table from clarity and I only included the newly introduced parameters.

Parameter	**Explanation**
NIEP	Number of micro steps to accumulate the corrections (N_m)
IE3	Switch on or off the calculation of $E^{(3)}$

Table 3.1: Input parameters of the software to run i-MSQMC calculations with $E^{(2)}$ and $E^{(3)}$.

Regarding our parallel implementation, both spawned and removed determinants are communicated to MPI processes and the same MPI process is allocated to both replicas. As a result of this, the accumulation of $E^{(2)}$ does not interfere with the parallel operation between MPI processes. In other words, fractions of $E^{(2)}$ can be calculated at every MPI process and these fractions are accumulated by the main process at the end of every time step.

The calculation of $E^{(3)}$ is less pleasant and our implementation currently only supports sequential $E^{(3)}$ calculations. The reason for this is that there are two indices in the summation of $E^{(3)}$ which requires massive communication of the removed walkers between MPI processes. This has not yet been implemented and all the calculations with $E^{(3)}$ results are calculated using a single MPI process.

In the following few lines, I compare the computational costs of our original i-MSQMC implementation using MPI to distribute the main walker array to processes with the implementation including the sampling of $E^{(2)}$. The expression of the second order correction contains only one summation index, therefore $E^{(2)}$ can be calculated in a parallel fashion without additional overhead.

3.9 Results and Discussions

What follows here is a detailed explanation regarding the accuracy of i-MSQMC, $E^{(2)}$ and $E^{(3)}$ with respect to FCI. In order to demonstrate the favourable computational costs of MSQMC compared to that of FCI the size of the FCI space is calculated using the following formula:

$$\mathrm{N_{FCI}} = \binom{K}{n_\alpha}\binom{K}{n_\beta} \tag{3.37}$$

, where K is the number of correlated orbitals n_α and n_β are the number of spin up and spin down electrons respectively. One of the goals of these calculations were to ascertain the correctness of my algorithm and to first find optimal input parameters to be used in calculations on realistic systems. There are two very important parameters the user can set. These are the initiator threshold (N_i) and the number of walkers on the reference determinants (N_b). They are thought to be important, because they can significantly influence the both the accuracy of the calculation and calculation costs.

3.9.1 Second and third order Epstein-Nesbet perturbation theory

In the subsequent two calculation results I shed light on the accuracy of the i-MSQMC energy and the perturbative corrections with respect to the initiator threshold for two molecular systems. Neither of these molecules can be regarded as strongly correlated systems therefore the corrections were expected to exhibit high accuracies. In both cases overcorrelation due to $E^{(2)}$ was observed, however calculating $E^{(3)}$ and adding it to obtain the final energy enabled me to obtain near FCI values even at remarkably low walker populations (or with high initiator threshold equivalently).

Firstly, let us turn our attention to Figure 3.9. Water in the cc-PVDZ basis set can serve as weakly correlated system with modest computational requirements to test the correctness of our theory and implementation. For this system there are 23 active orbitals and 8 active electrons, therefore the size of the FCI space is in the order of 10^8.

For H_2O the third order perturbative correction at $N_i = 25$ seems to recover most of the initiator error (∼8mEh) and the final energy is already very near the FCI value. Employing 250 walkers on the reference determinant and setting $N_i = 25$ the number of walkers in equilibrium state were around 2000 and I were able to calculate near FCI energies using $E^{(3)}$. In every point of the curve the standard error (calculated from the standard deviation using the results of 3 independent calculations) is presented using red vertical lines. The standard error for this calculation is in the order of 1mEh for the corrections and in the order of 0.1mEh for the i-MSQMC energy. Note that the accumulation of correction samples was not applied in this case ($N_m = 1$). Pushing the limits of our implementations slightly farther a calculation on the N_2 molecule was performed to the analogy of the preceding application and my results are presented on Figure 3.10. The nitrogen molecule in cc-PVDZ basis requires much more memory as the number of active orbitals is 26 and the number of active electrons is 5 resulting in a FCI space with a size in the order of $\sim 10^{10}$.

Similar peculiarities of $E^{(2)}$ and $E^{(3)}$ were observed as for H_2O. For large initiator thresholds $E^{(2)}$ is lower than the FCI energy and $E^{(3)}$ recovers most of this discrepancy. The optimum point regarding the tradeoff between accuracy and computational costs for this system lies in the regime around $N_i = 6$ and this value will be used henceforth. Employing 250 walkers on the reference determinant and setting $N_i = 6$ the number of walkers in equilibrium state were around 6000 and I were able to calculate near FCI energies using $E^{(3)}$.

In every point of the curve the standard error (calculated from the standard deviation using the results of 3 independent calculations) is presented using red vertical lines. The standard error for this calculation is in the order of 0.5mEh for the corrections and in the order of 0.1mEh for the i-MSQMC energy. Note that the accumulation of correction samples was not applied in this case ($N_m = 1$).

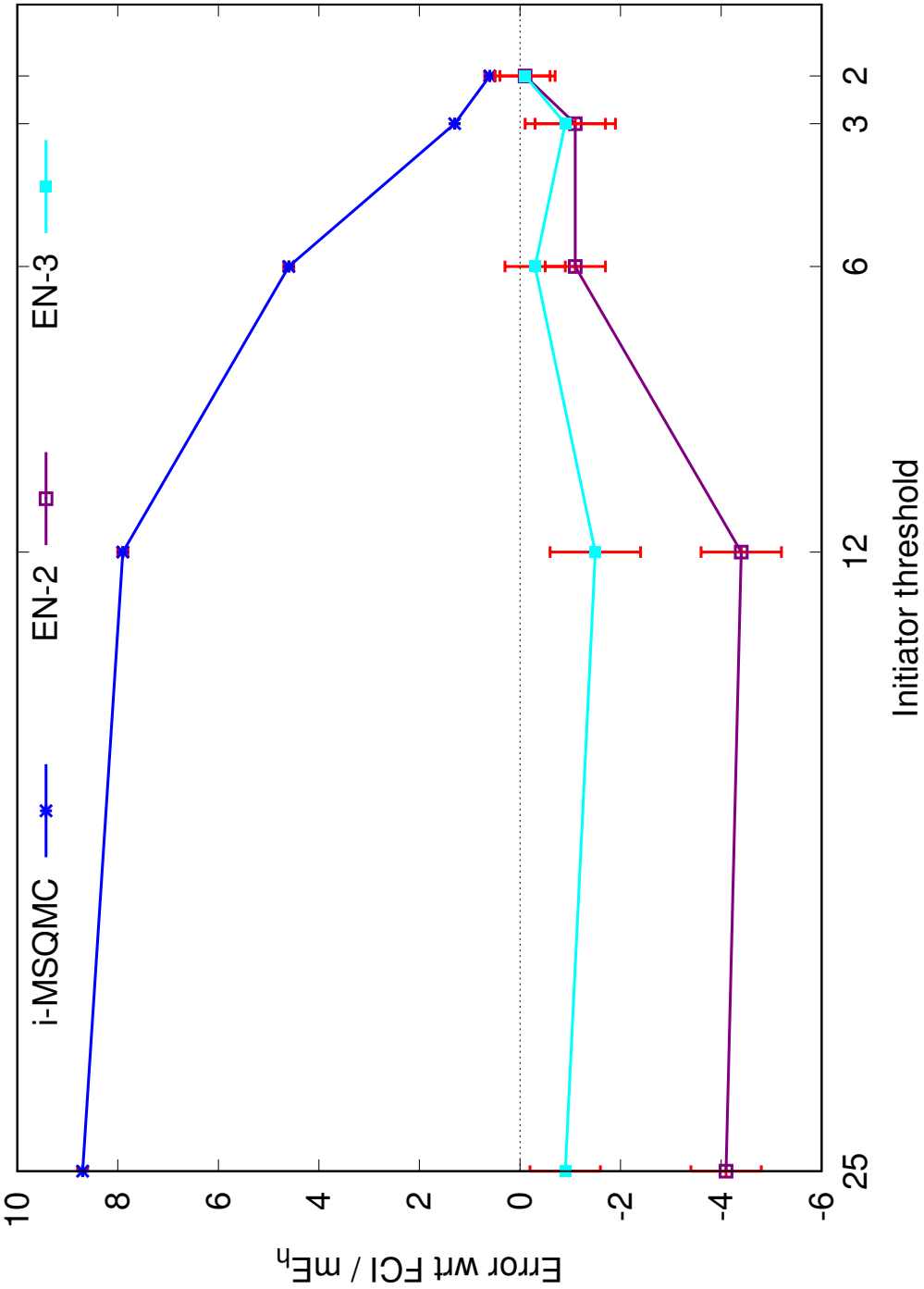


Figure 3.9: The accuracy of i-MSQMC, $E^{(2)}$ and $E^{(3)}$ with respect to FCI while varying the initiator threshold. The calculation was carried out on an H_2O molecule with $N_b = 250$, $N_m = 1$, $\tau = 1000$ a.u. and $\tau_1 = 600$ a.u. using the cc-pVDZ basis set. The frozen-core approximation was used for this calculation.

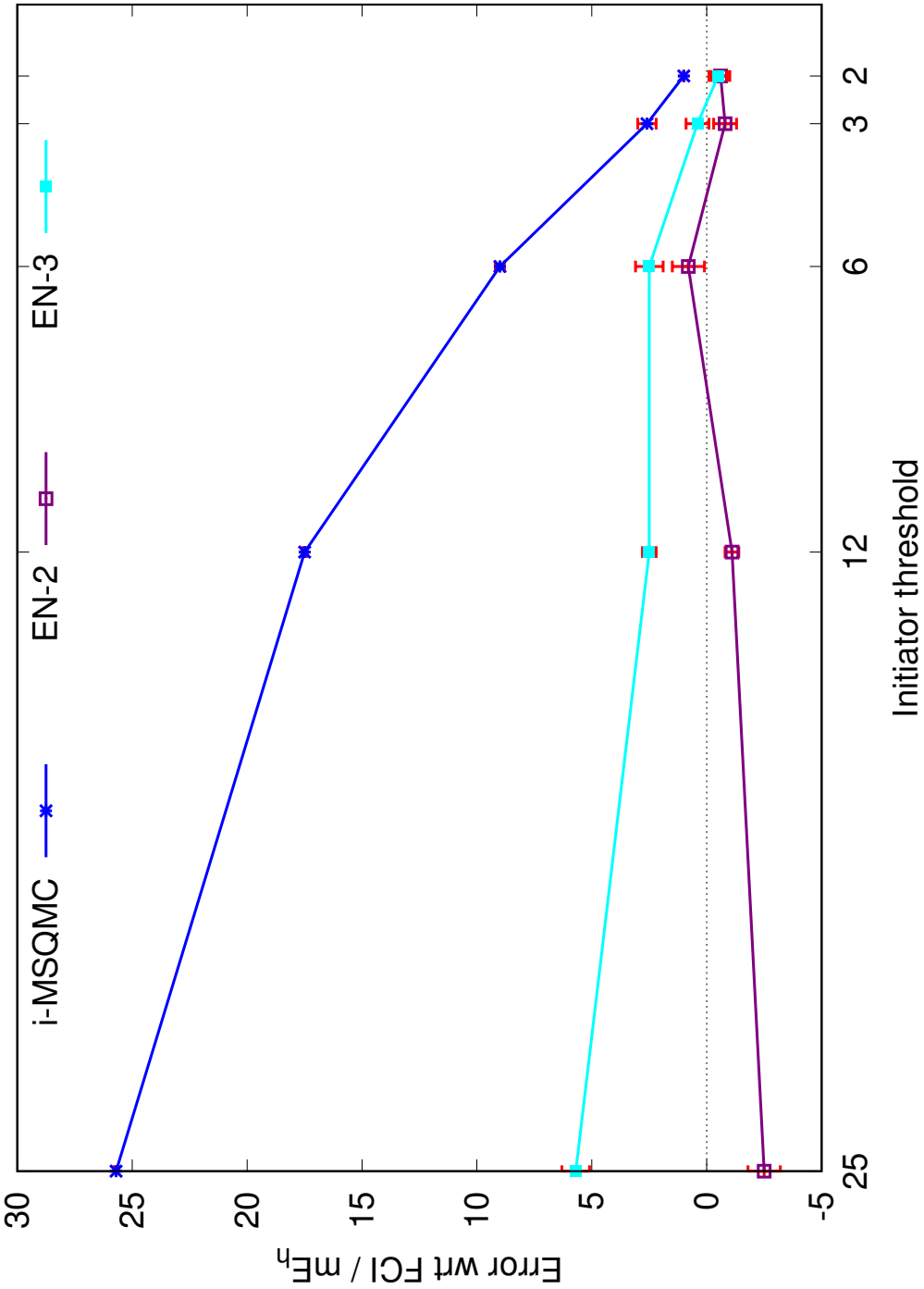


Figure 3.10: The accuracy of i-MSQMC, $E^{(2)}$ and $E^{(3)}$ with respect to FCI while varying the initiator threshold. The calculation was carried out on an N_2 molecule with $r_{NN} = 1.12013$Å, $N_b = 250$, $N_m = 1$, $\tau = 1000$ a.u. and $\tau_1 = 600$ a.u. using the cc-pVDZ basis set. The frozen-core approximation was used for this calculation.

3.9.2 Size-inconsistency error related investigations

The source of the size-inconsistency error has been investigated in our previous work[36]. The initiator approximation truncates the sampled space, therefore it introduces size-inconsistency error similar to those that truncated CI methods exhibit. In order to investigate the size-consistency properties of the initiator adaptation and the perturbation correction we performed several calculations using different initiator thresholds on a neon atom and on two non-interacting neon atoms (i.e. with large inter-atom separation). Our investigations commenced at $N_i = 30$ for which neither the calculation of the atom nor the calculation of the diatomic system incorporated initiator determinants expect for the reference. In this case the initiator is almost as large as 10 mEh and $E^{(2)}$ seems to recover most of this flaw. The number of initiator determinants increases with decreasing N_i reducing the initiator error and the results approach the FCI energy, which is obviously size-consistent. Regarding the computational costs, the calculation of a single energy using one seed with $N_i = 30$ requires approximately 5 times less computational time than the one with $N_i = 3$. Note that the calculation with $N_i = 30$ supplemented with $E^{(2)}$ exhibits similar amount of size-inconsistency error as the calculation with $N_i = 3$ without $E^{(2)}$.

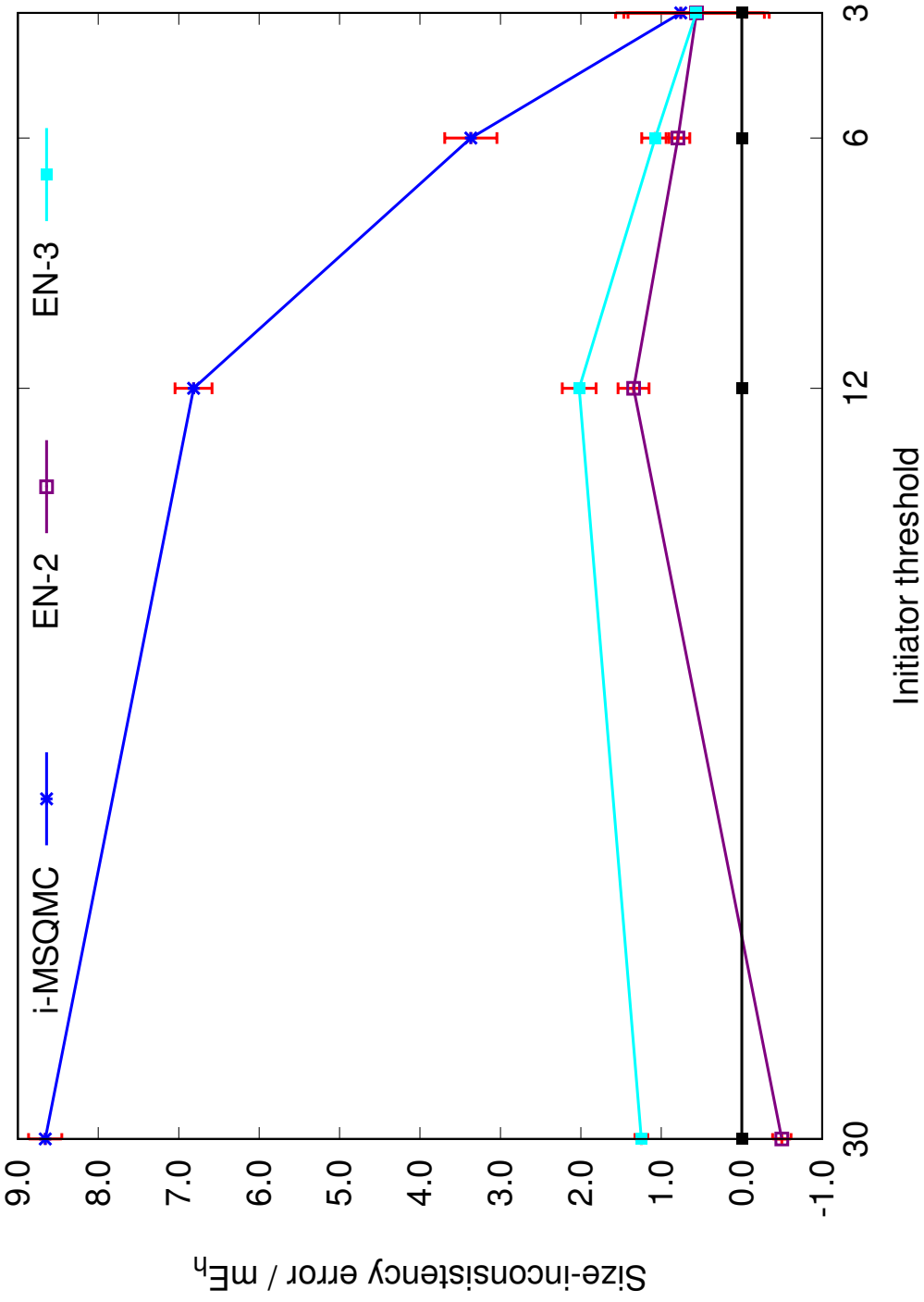


Figure 3.11: Size-inconsistency error with respect to the initiator threshold. This calculation was performed with $N_b = 250$, $N_m = 5$, $\tau = 1000$ a.u. and $\tau_1 = 600$ a.u. using the cc-pVDZ basis set. The frozen-core approximation was used for this calculation. Standard deviation is demonstrated using red vertical lines.

3.9.3 The N_2 molecule near equilibrium

The description of electronic states at non-equilibrium geometries and bond stretching in quantum chemistry is a nontrivial challenge for all approximate correlated methods. The N_2 molecule in equilibrium state possesses a triple bond and the molecule has been under investigation for many years.[23, 24, 25, 26, 27, 28, 29, 30, 31] Bellow, I present calculation results for the potential energy curve of the N_2 molecule in cc-pVDZ basis. The motivation for this calculation can be traced back to the results of Figure 3.10. From that figure it is clear that $E^{(2)}$ and $E^{(3)}$ are very close to the FCI result when $N_i = 6$. Using the approach of accumulating multiple samples to calculate the corrections enabled me to calculate very accurate energies even at low walker populations. For this case samples of 5 steps were collected ($N_m = 5$) and only 250 walkers were placed on the reference determinant initially ($N_b = 250$). Figure 3.12. shows that MSQMC with the initiator approximation fails to provide an accurate picture of the bond dissociation curve. This is mainly due to the fact that the initiator threshold is higher than usual and N_b is also very small.

As FCI values were not available at every points of the curve therefore CCSD(T) results are presented as reference. CCSD(T) believed to provide reliable results with moderate computational costs in the equilibrium geometry region[28]. For this calculation, at every point results of 3 different seeds were averaged and the standard deviation is also presented using red vertical lines. The fact that the standard deviation of the corrections is significantly larger than that of the energy within the initiator adaptation would hinder me from obtaining accurate values for the energy, therefore I chose to accumulate the contributions from the samples of 5 micro steps($N_m = 5$). Using this scheme it was possible to shrink the standard deviation below a value that would not distort the shape of the curve significantly.

As it is usually the case for weakly correlated systems $E^{(2)}$ seems to overcorrelate the results and the curve is below the CCSD(T) curve. On the other hand, the value of the third order correction is positive, larger in absolute value than the second order and it is very close to the CCSD(T) results.

On the upper plot of Figure 3.13. I demonstrate how the total number of walkers changes when the nitrogen atoms are moving apart. On the lower plot, I show how the number of determinants with N_w^i exceeding 10 changes with bond length. Both the total number of walkers and the number of important determinants (e.g. determinants with large number of walkers) increase as the separation between the nitrogen atoms increases. This fact implies that the Hartree-Fock determinant becomes less and less important with increasing bond length and that causes in the extreme case of infinite bond length the breakdown of single-reference methods. All of the methods on Figure 3.12. are single-reference approaches, therefore they only provide a reliable picture of the system near equilibrium geometry. When diatomic molecules are torn apart these methods (including CCSD(T), which in some cases does not even have a solution at stretched geometries) exhibit significant inaccuracies and the definition of the reference wave function requires reconsideration.

The N_2 molecule in cc-pVDZ basis is an ideal molecular system for the verification of the stochastic algorithm supplemented with second and third order perturbative corrections. As a diatomic molecule, it is convenient to set up calculations in non-equilibrium geometries to ascertain the correctness of the algorithm even when the reference determinant is less important than at the equilibrium geometry. My implementation reproduces the CCSD(T) curve in several points of the potential energy curve when the second and the third order corrections are added to the i-MSQMC energy.

In Table 3.2. I present numerical data for the all-electron energies on the dissociation curve of the N_2 molecule. I used a variety of methods to ascertain the shape of the curve and I observed that i-MSQMC with the third order correction provides the most accurate energies regardless of the position on the dissociation curve. At first, I calculated near-equilibrium energies and then I gradually decreased and increased the bond length of the molecule. For every point I averaged the results of 3 independent calculations and calculated the standard deviation along with the energy. MP2 and CCSD are always above the reference CCSD(T) energy and they fail to provide an accurate picture of the bond dissociation.

3.9.4 The chromium dimer

Bellow, I present calculation results for the Cr_2 molecule using the SV basis set[35]. This system has been notoriously one of the most challenging in the quantum chemical community. Almost all methods struggle with its mix of important dynamical and static correlation effects owing to bonding character[34] between both low-lying 3d and more diffuse 4s orbitals, as well as its weak and unusual hextuple bond. I performed a single-point calculation of the chromium dimer, at a bond length of 1.5Å , correlating 24 electrons in an orbital space of 30 Hartree-Fock orbitals, formed from an underlying SV basis. A similar calculation was carried out in 2013 using i-FCIQMC[33] and the this energy (-2086.4212(3) Hartrees) is used as a reference for my work. Although the SV basis will be insufficient to capture the dynamical correlation effects, it should be a sufficient model of the valence space to include the strong static correlation effects of the d-d binding which dominates this bond length. Note that this space is well outside what could be achieved with FCI. A single vector of the size of the space would require around 60,000 terabytes of memory.

Following a detailed investigation on the input parameter space, I found that $N_i = 5$ with $N_b = 200$ are the values with which the ITE of the chromium dimer can be converged upon and I was able to obtain the final energy within a few mEh away from the i-FCIQMC value. For the calculation I used our MPI parallel implementation with 10 MPI process to calculate both the i-MSQMC energy and the second order perturbative correction.

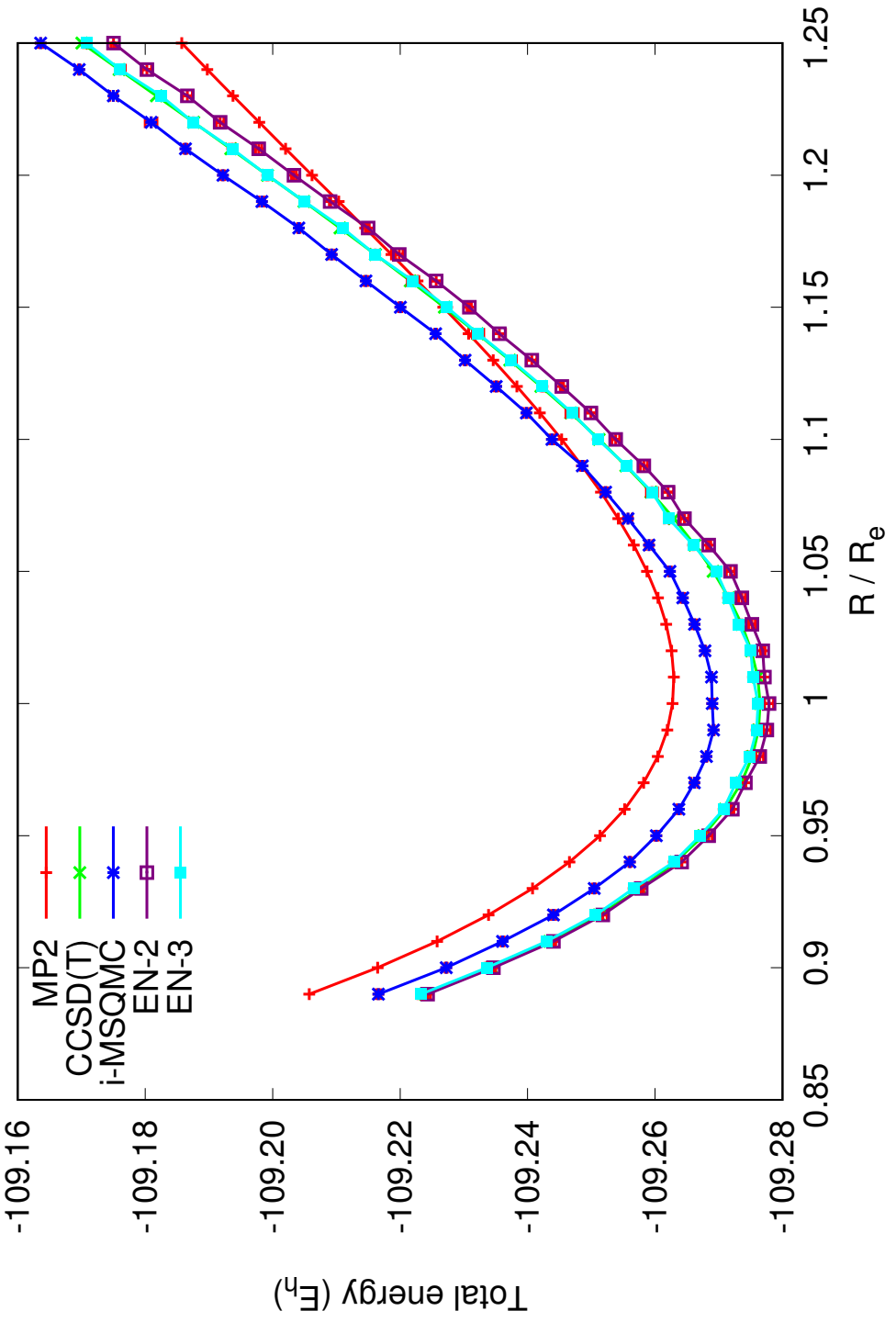


Figure 3.12: This calculation was carried out on a nitrogen molecule with $N_i = 6$, $N_b = 250$, $N_m = 5$, $\tau = 1000$ a.u. and $\tau_1 = 600$ a.u. using the cc-pVDZ basis set. The frozen-core approximation was used for this calculation. The X axis represents the multiplier of the equilibrium bond length ($r_{\mathrm{NN}} = 1.12013$Å). Standard deviation is demonstrated using red vertical lines.

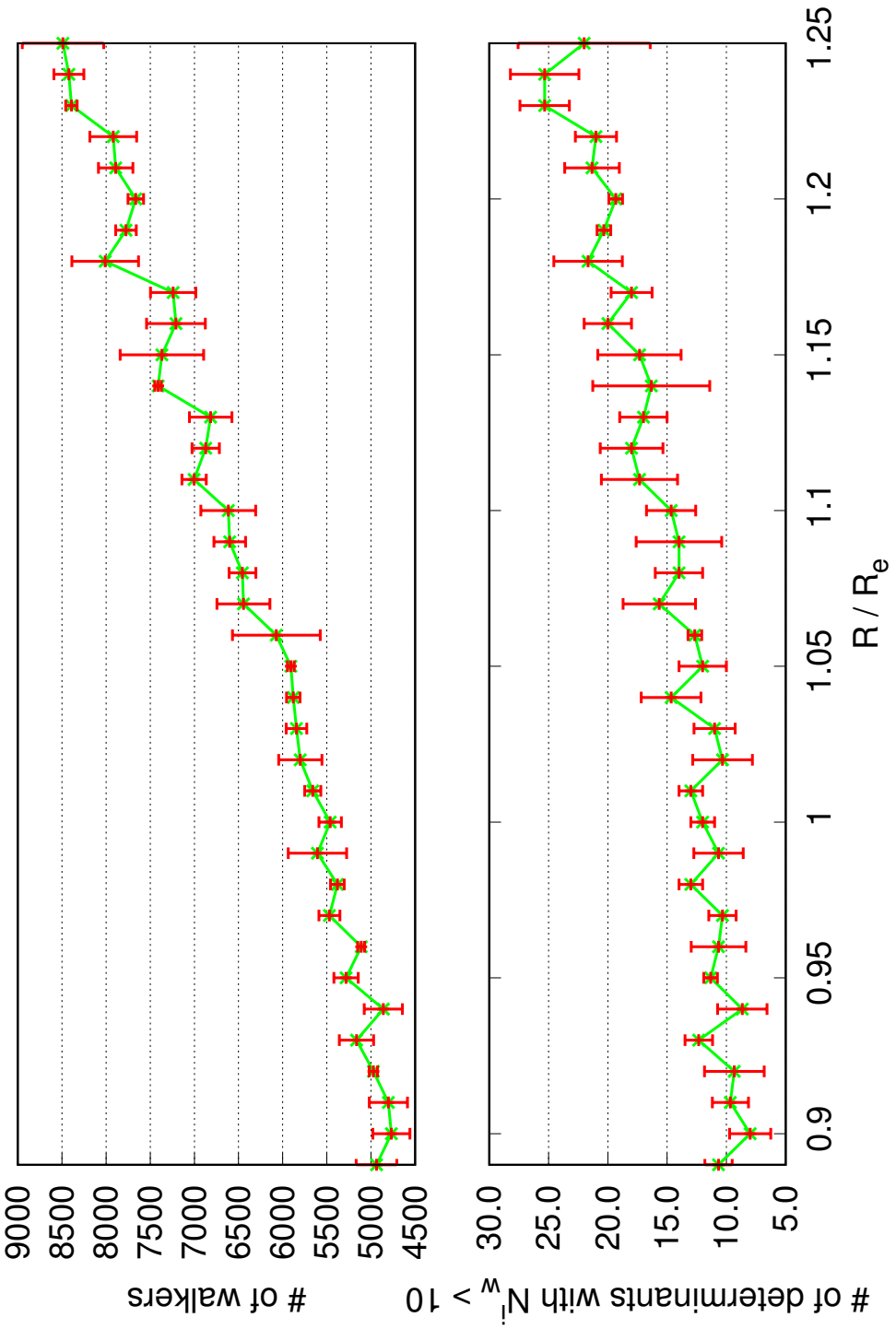


Figure 3.13: Walker growth with respect to bond length. This calculation was carried out on a nitrogen molecule with $N_i = 6$, $N_b = 250$, $N_m = 5$, $\tau = 1000$ a.u. and $\tau_1 = 600$ a.u. using the cc-pVDZ basis set. The frozen-core approximation was used for this calculation. The X axis represents the multiplier of the equilibrium bond length ($r_{\mathrm{NN}} = 1.12013$Å). Standard deviation is demonstrated using red vertical lines.

R	0.9	1.0	1.1	1.2	1.25	1.55
RHF	-108.935 348	-108.949 557	-108.893 708	-108.808 349	-108.761 581	-108.494 693
MP2	-109.216 468	-109.262 734	-109.245 329	-109.206 187	-109.185 748	-109.124 038
CCSD	-109.225 022	-109.263 898	-109.234 480	-109.177 241	-109.145 085	-108.973 744
i-MSQMC	-109.227 23(9)	-109.268 9(5)	-109.243 8(4)	-109.192 2(1)	-109.163 6(2)	-109.019 0(9)
EN-2	-109.234 7(5)	-109.277 9(9)	-109.253 83(8)	-109.203 3(5)	-109.175 0(4)	-109.027(1)
EN-3	-109.233 6(4)	-109.276 2(9)	-109.251 01(3)	-109.199 2(4)	-109.170 8(6)	-109.026(1)
CCSD(T)	-109.234 369	-109.276 478	-109.251 234	-109.199 197	-109.170 055	-109.025 388

Table 3.2: All-electron energies of the nitrogen molecule vs bond length (in multiplier of the equilibrium bond length ($r_{\mathrm{NN}} = 1.12013$Å)) as calculated by a variety of methods, using the cc-pVDZ set. All energies are in hartree, and the QMC standard deviations are on the last digit. The calculation parameters are as follows: $N_i = 6$, $N_b = 250$, $N_m = 5$, $\tau = 1000$ a.u. and $\tau_1 = 600$ a.u. The frozen-core approximation was used for this calculation.

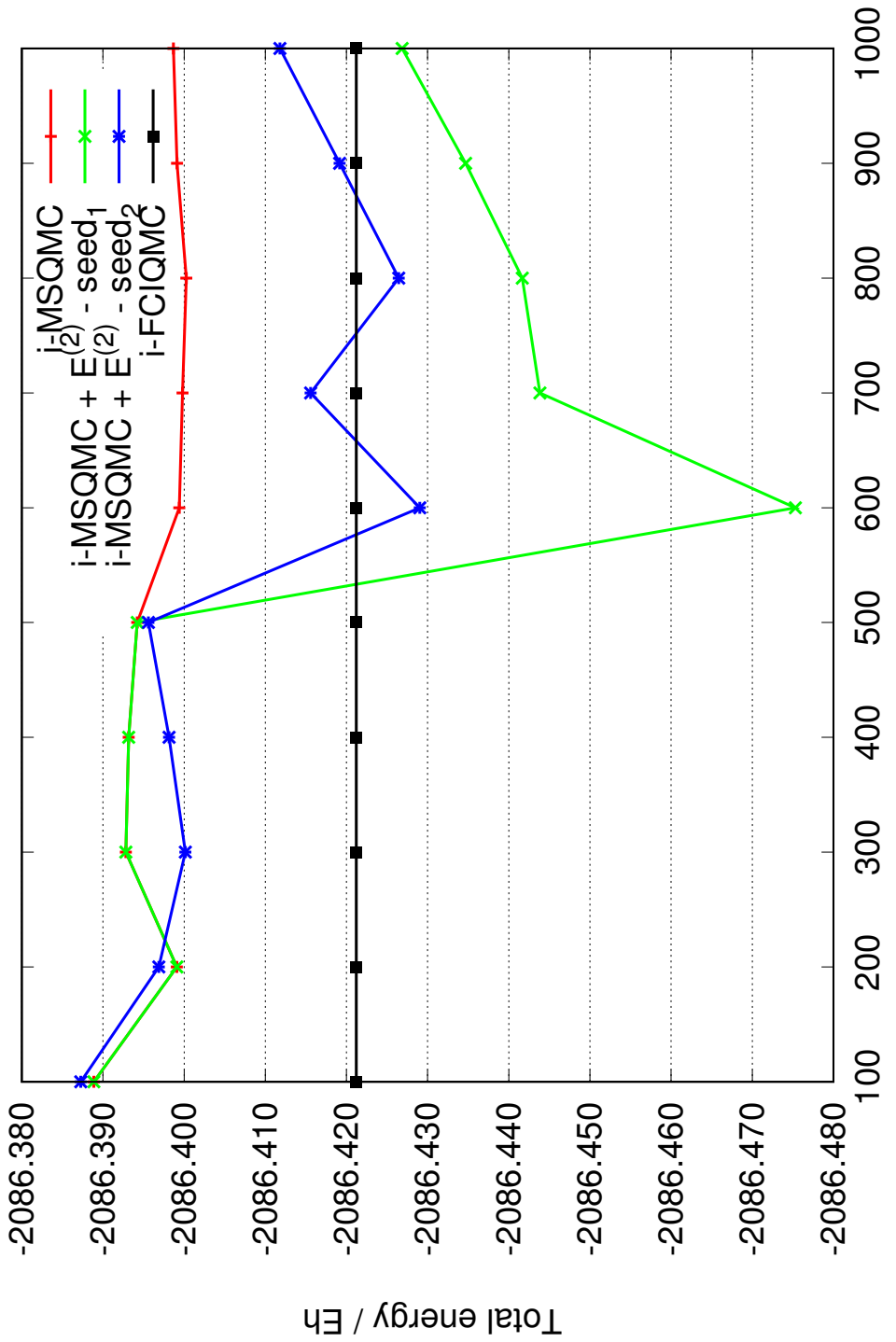


Figure 3.14: Convergence of the i-MSQMC energy and the second order correction. This calculation was carried out on a chromium molecule with $N_i = 5$, $N_m = 5$ $N_b = 200$, $\tau = 1000$ a.u. and $\tau_1 = 600$ a.u. using the SV basis set.

3.9.5 Size-inconsistency error related investigations

The source of the size-inconsistency error has been investigated in my previous work[36]. The initiator approximation truncates the sampled space, therefore it introduces size-inconsistency error similar to those that truncated CI methods exhibit. In order to investigate the size-consistency properties of the initiator adaptation and the perturbation correction I performed several calculations using different initiator thresholds on a neon atom and on two non-interacting neon atoms (i.e. with large inter-atom separation). My investigations on Fig. 3.15. commenced at $N_i = 30$ for which neither the calculation of the atom nor the calculation of the diatomic system incorporated initiator determinants expect for the reference. In this case the initiator is almost as large as 10 mEh and $E^{(2)}$ seems to recover most of this flaw. The number of initiator determinants increases with decreasing N_i reducing the initiator error and the results approach the FCI energy, which is obviously size-consistent. Regarding the computational costs, the calculation of a single energy using one seed with $N_i = 30$ requires approximately 5 times less computational time than the one with $N_i = 3$. Note that the calculation with $N_i = 30$ supplemented with $E^{(2)}$ exhibits similar amount of size-inconsistency error as the calculation with $N_i = 3$ without $E^{(2)}$.

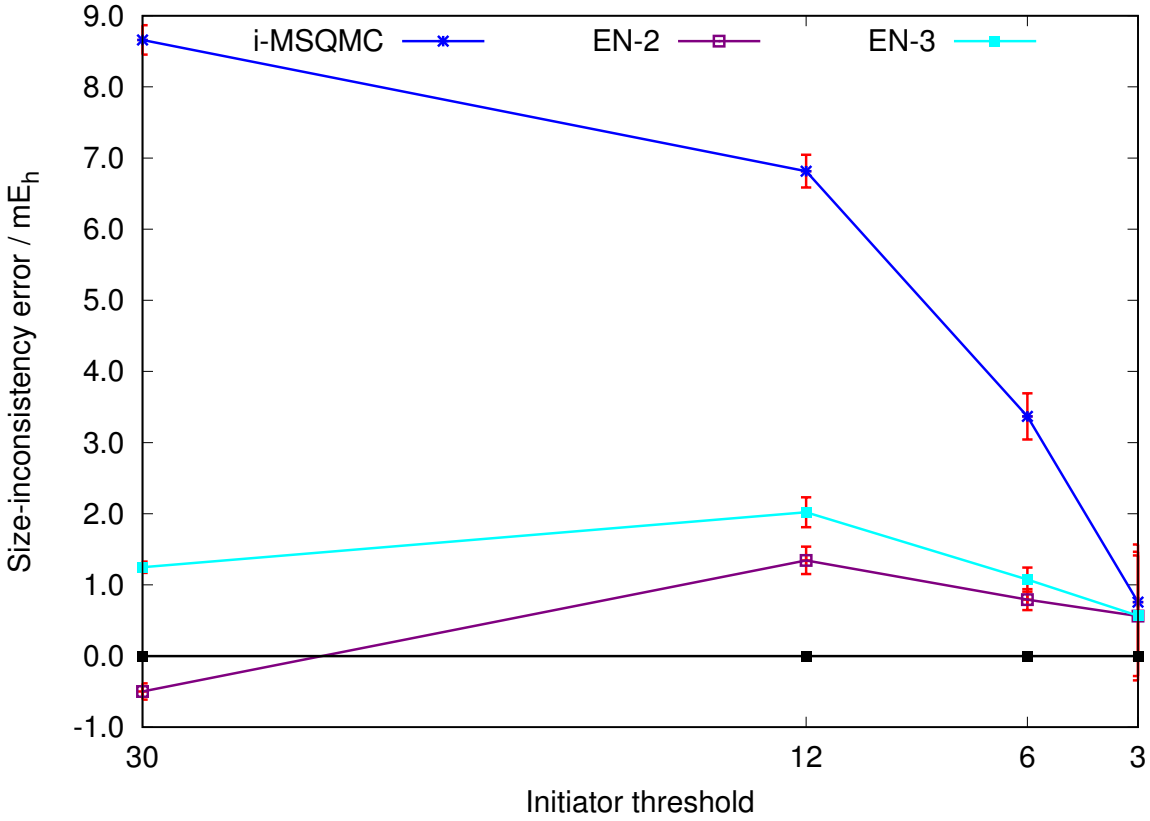


Figure 3.15: Size-inconsistency error with respect to the initiator threshold. This calculation was performed with $N_b = 250$, $N_m = 5$, $\tau = 1000$ a.u. and $\tau_1 = 600$ a.u. using the cc-pVDZ basis set. The frozen-core approximation was used for this calculation. Standard deviation is demonstrated using red vertical lines. The size-inconsistency error is calculated as the difference between twice of the energy of a neon atom and the energy of two non-interacting neon atoms.

3.9.6 Singlet-triplet splitting of the acene series

The non-diminishing singlet-triplet gap of the acene series has been investigated both by experimental chemists and numerical calculations. One of the most accurate description of the ground state and the first excited state was given using DMRG[50]. According to the results in the previous section, second and third order perturbation theory compensates some of the size-inconsistency error introduced by the initiator approximation. On Fig. 3.16. I plot the gap values of acenes from 2 rings up to 6 rings obtained using a very large initiator threshold. In these calculations only the reference determinant acted as initiator. Even with such a crude approximation, perturbation theory was able to reproduced the gap values in a size-consistent manner. In other words, as it is demonstrated on Fig. 3.16., there is a constant error between the DMRG values and our values, however the error does not grow with system size. Consequently, second and third order perturbation theory can be used to calculated properties of large systems for which size-inconsistent methods fail to provide meaningful results.

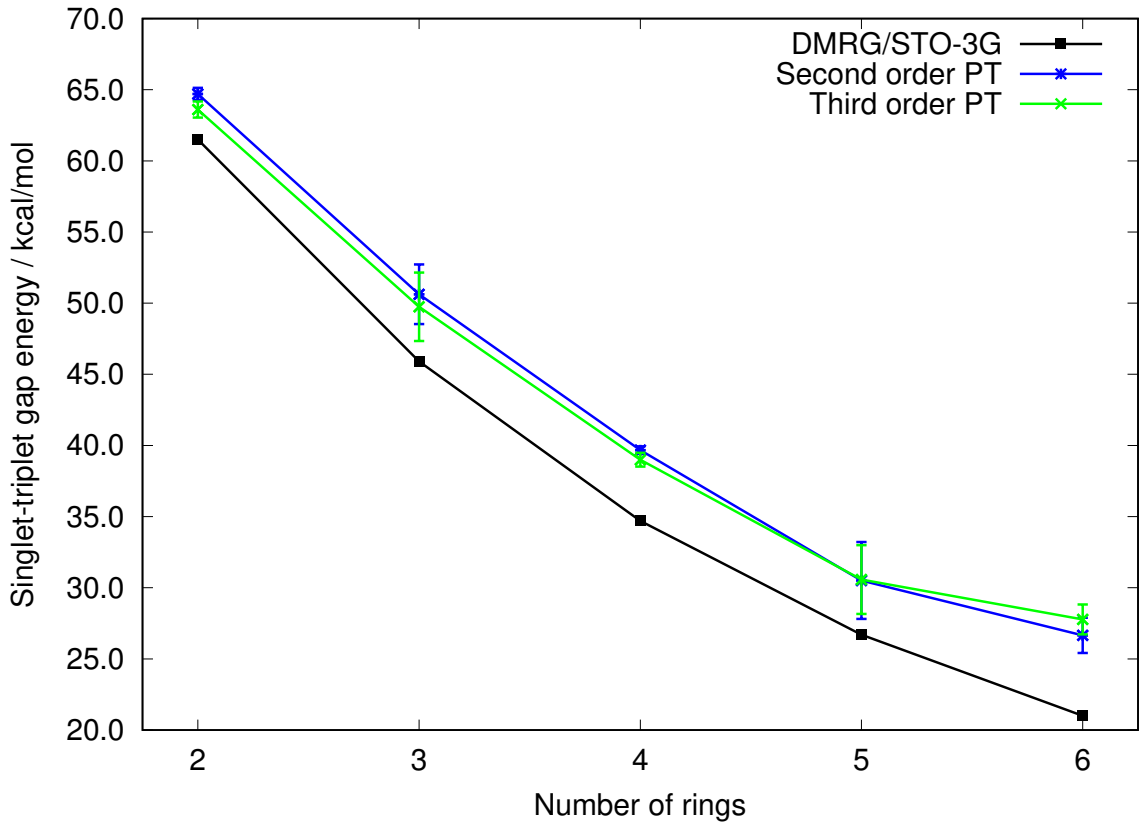


Figure 3.16: Singlet-triplet gap values for the acene series up to hexacene using second and third order Epstein-Nesbet perturbation theory. The calculations was performed with $N_i = 150$, $N_b = 250$, $N_m = 2$, $\tau = 1000$ a.u. and $\tau_1 = 600$ a.u. using the STO-3G basis set correlating the π orbitals. Standard deviation is demonstrated using vertical lines.

3.10 Conclusion

The initiator adaptation of MSQMC provides a memory effective way to sample the wave function in large active spaces in which the exact diagonalization of the FCI matrix is impossible even on supercomputers.

The findings of the current research showed that it is much more efficient to calculate second and third order Epstein-Nesbet corrections when the initiator threshold is large instead of further increasing the number of walkers or lowering the threshold. The dimension of the secondary space in realistic calculations can be extremely large and this could prevent the accurate sampling of these corrections. In order to ameliorate this problem I presented an idea to collect multiple samples and accumulate the corrections less frequently.

In many of the calculations only a small fraction of the Hilbert space was sampled by using a few hundred walkers on the reference determinant and choosing a fairly high initiator threshold.

The results of this research prove the long-existing conjecture that only a few determinants contribute significantly to the correlation energy and in most of the cases there is no necessity for the storage and incorporation of determinants with small CI coefficients.

On the other hand, the way how one excludes these less important contributions is extremely important and if the truncation is done without circumspection it can lead to unphysical results as it was observed in the first part of this thesis.

With the appearance of Monte Carlo methods in quantum chemistry the importance sampling of determinants became a prevalent way to treat extremely large Hilbert spaces by systematically choosing contributions in decreasing order of importance. The convenience of Epstein-Nesbet perturbation theory manifests in the initiator adaptation and seem to recover much of the correlation energy for weakly correlated molecular systems.

Regarding the calculation results of this chapter, $E^{(2)}$ seem to overcorrelate the results for most of the systems and $E^{(3)}$ recovers much of this flaw. The calculation of $E^{(2)}$ is by no means an overhead as it can be calculate in a parallel manner without extra inter-process communication of data. For some systems where $E^{(2)}$ is not accurate enough one can turn on the calculation of $E^{(3)}$, however this slows down the calculation because of the large number of terms that are required in the summation.

Note that for each term in the expression of $E^{(3)}$ the Hamiltonian matrix elements appear and the calculation of these can be costly as they appear in the innermost loop of the whole algorithm.

The extension to higher orders can be carried out using the perturbation formulas that can be found in any advance quantum chemistry textbook. The main problem for these terms is not only the fact that increasing number of terms appear in them but also that they contain products of removed walker populations and the previously mentioned bias issue hinders naive sampling.

There is an ongoing research on the applicability of $E^{(2)}$ and $E^{(3)}$ to strongly correlated systems for which a small active space especially at stretched geometries fails to describe the electronic structure. Increasing the size of the active space limits the applicability of MSQMC mostly because the supercomputers have limited memories and runtimes available to users. The only promising solution is the use of small number of walkers for the simulation and recovering the rest of the correlation energy with the previously describe corrections. Clearly, as the available computational power increases it is desirable to have systematically improvable algorithms to match accuracy with the available machine power.

In this research I

1. Implemented an algorithm to sample $E^{(2)}$ and $E^{(3)}$
2. Gave a proposal using multiple samples to mitigate the stochastic noise that would hinder the accurate sampling of these corrections at low walker populations.

3. Calculated $E^{(2)}$ and $E^{(3)}$ for various molecular systems and compared my results with other ab initio quantum chemistry methods.
4. Reproduced the CCSD(T) potential energy curve of the nitrogen molecule near equilibrium in cc-pVDZ basis with the aid of $E^{(2)}$ and $E^{(3)}$.

$E^{(2)}$ and in the cases, where computational resources allowed the calculation of $E^{(3)}$ corrected much of the initiator error and the obtained total energies were astonishingly near the corresponding FCI values.

The purpose of these numerical calculations was a systematical assessment of these corrections. Upon the whole, the initiator approximation solves the problem of exponential walker growth while introducing a truncation on the Hamiltonian and with the aid of these Epstein-Nesbet type corrections one has the option to calculate remarkably accurate energies for weakly and even strongly correlated molecular systems that were nowhere near reachable before the advent of QMC methods.

First of all, the correctness of the method was verified with calculations in Hilbert spaces of moderate sizes. To further demonstrate the applicability of $E^{(2)}$ and $E^{(3)}$ I calculated the potential energy curve of the N_2 molecule using the cc-pVDZ basis set in several points. $E^{(2)}$ corrected the initiator error for these molecules significantly with almost no computational overhead. With the further addition of $E^{(3)}$ I was able to counterpoise the overcorrelation effect of the second order correction and at each point of the curve my total energies are aligned with the CCSD(T) results.

Finally, it might be expedient to mention possible future researches based on mine. The extension to multi-reference approaches is evidently a direction that needs attention, however the derivation of multi-reference stochastic perturbation theory is no easy task.

Secondly, a wise parallel implementation of the third order correction has not yet been discovered. This is essential when the scope of the calculations exceeds that of the current one and these possibilities are under consideration.

A third option that will require attention in the future is extension of the whole framework for excited states. It would be handy to calculate excitation energies and energy gaps within the current framework and I will endeavor to implement such algorithms in the near future.

Bibliography

[1] D. Cleland, G. H. Booth, A. Alavi, J. Chem. Phys. **132**, 041103 (2010).

[2] S. Ten-no, J. Chem. Phys. **147**, 244107 (2017).

[3] B. Huron, J. P. Malrieu, P. Rencurel, J. Chem. Phys. **58**, 5745-5759 (1973).

[4] R. J. Buenker, S. D. Peyerimhoff, Theor. Chim. Acta **35**, 33-58 (1974).

[5] S. Evangelisti, J.-P. Daudey, J.-P. Malrieu, Chem. Phys. **75**, 91 (1983).

[6] R. J. Harrison, J. Chem. Phys. **94**, 5021 (1991).

[7] A. Povill, J. Rubio, F. Illas, Theor. Chim. Acta **82**, 229 (1992).

[8] E. Giner, A. Scemama, M. Caffarel, Can. J. Chem. **91**, 879 (2013).

[9] E. Giner, A. Scemama, M. Caffarel, J. Chem. Phys. **142**, 044115 (2015).

[10] J. B. Schriber, F. A. Evangelista, J. Chem. Phys. **144**, 161106 (2016).

[11] N. M. Tubman, J. Lee, T. Y. Takeshita, M. Head-Gordon, K. B. Whaley, J. Chem. Phys. **145**, 044112 (2016).

[12] W. Liu, M. R. Hoffman, J. Chem. Theory Comput. **12**, 1169 (2016).

[13] A. A. Holmes, N. M. Tubman, C. J. Umrigar J. Chem. Theory Comput. **12**, 3674 (2016).

[14] P. S. Epstein Phys. Rev. **28**, 695 (1926).

[15] R. K. Nesbet, Proc. R. Soc. A **230**, 312 (1955).

[16] R. K. Nesbet, Proc. R. Soc. A **230**, 922 (1955).

[17] C. Overy, G. H. Booth, N. S. Blunt, J. J. Shepherd, D. Cleland, A. Alavi, J. Chem. Phys. **141**, 244117 (2015).

[18] S. Sharma, e-print arXiv:1803.04341 [cond-mat.str-el] (2018).

[19] G. K.-L. Chan, A. Keselman, N. Nakatani, Z. Li S. R. White J. Chem. Phys. **145**, 014102 (2016).

[20] S. Guo, Z. Li, G. K.-L. Chan J. Chem. Phys. **148**, 221104 (2018).

[21] N. S. Blunt J. Chem. Phys. **148**, 221101 (2018).

[22] N. S. Blunt, A. J. W. Thom, C. J. C. Scott J. Chem. Phys. **15**, 3537-3551 (2019).

[23] M. Musial R. J. Bartlett J. Chem. Phys. **122**, 224?102 (2005).

[24] P. Piecuch, K. Kowalski J. Chem. Phys. **113**, 5644 (2000).

[25] G. K.-L. Chan, M. Kallay, J. Gauss J. Chem. Phys. **121**, 6110 (2004).

[26] W. D. Laidig, P. Saxe, R. J. Bartlett J. Chem. Phys. **86**, 887 (1987).

[27] T. V. Voorhis, M. Head-Gordon J. Chem. Phys. **112**, 5633 (2000).

[28] J. W. Krogh J. Olsen, J. Chem. Phys. **344**, 578 (2001).

[29] J. W. Krogh, J. Olsen J. Chem. Phys. **344**, 578 (2001).

[30] S. R. Gwaltney, E. F. C. Byrd, T. V. Voorhis, M. Head-Gordon Chem. Phys. Lett. **353**, 359 (2002).

[31] H. Larsen, J. Olsen, P. Jørgensen, O. Christiansen J.Chem. Phys. **113**, 6677 (2000).

[32] J. Olsen, P. Jørgensen, H. Koch, A. Balkova, R.J. Bartlett, J.Chem. Phys. **104**, 8007 (1996).

[33] G. H. Booth, S. D. Smart A. Alavi Molecular Physics **112:14**, 1855-1869 (2014).

[34] P. Celani, H. Stoll, H.-J. Werner P.J. Knowles Mol. Phys. **102**, 2369 (2004).

[35] A. Schafer, H. Horn, R. Ahlrichs J. Chem. Phys. **97**, 2571 (1992).

[36] B. Ladóczki, S. L. Ten-no, J. Chem. Phys. **151**, 114113 (2019)

4 General Conclusion

Within the scope of this work, I have carried out a detailed investigation regarding the applicability and computational requirements of the MSQMC method. I have reproduced several deterministic results with high accuracy using the MSQMC implementation developed by our research group.

In Chapter 3. I have investigated the theoretical and numerical aspects of stochastic perturbation theory in truncated configuration spaces. Compared to the configuration interaction ansatz the favourable peculiarities of the couple cluster ansatz have been observed in numerous researches in the past. Naive truncation in the Slater determinant space leads to unexpected results as the system size grows. This is due to the well-known size-inconsistency problem of truncated CI methods. In order to understand the source of this imperfection a recapitulation in connexion with diagrammatic perturbation was given. Stochastic methods always suffer from stochastic noise and in some cases the noise becomes so large that it is impossible to extract meaningful results out of the calculations. Therefore, as long as computation resources allow the users to store certain amount of deterministic data describing the most important part of the object to be sampled semi-stochastic approaches are worth considering.

In Chapter 4. I carried out a detailed investigation on second and third order Epstein-Nesbet perturbation corrections within the initiator adaptation. These corrections can be calculated without consuming significant amount of memory and it was shown that the second order correction can be obtained conveniently using the original parallel implementation. The main issue under investigation with these correction is the amount of stochastic noise they exhibit.

Obviously, by increasing the number of sampled terms the standard deviation of the results can be reduced and based on this idea a new sampling approach was presented in Chapter 4. The necessary modifications to sample these quantities were made in the original MSQMC implementation and the software is ready to calculate very accurate energies. They exhibit near FCI accuracy for weakly correlated systems, nay some calculations results for the second order correction were also presented in the current thesis for notoriously challenging (strongly correlated) systems such as the chromium dimer for example.

The chromium dimer is a benchmark systems to test new computational methods and $E^{(2)}$ seems to be very near to the latest computation results of other methods such as i-FCIQMC or i-MSQMC with CEPA(0) type corrections.

Upon the whole, the importance of MSQMC among the numerical methods of quantum chemistry has been discerned. It provides a convenient way to sample the wave function in Slater determinant space and by introducing systematically improvable truncations to economize computational time and memory it can serve as a tool to solve the many-body Schrödinger equation that I believe to become ubiquitous for the scientific community.

Alas, notwithstanding its ability to converge upon the FCI wave function and recover the correlation energy that the Hartee-Fock approximation fails to deliver the single-determinant description of the wave function is not always accurate enough and several imperfections are observed. The inability to describe such a static correlation can be tackled using multi-reference wave function theories and that would necessitate a new derivation in the MSQMC framework to sample such wave functions.

Due to limitations of resources that were at the disposal of the author of the current thesis no investigation was carried out regarding such possibilities and the topic of multi-reference approaches has been left as a future research topic.

I am exploring the possibility of extending our methods using various approaches. Research in the future will be centered around the incorporation of higher order commutators in the LCC based approach.

Another intriguing direction of our research will include derivations to express the same type second and third order corrections in multi-state effective Hamiltonian approaches. These and similar issues are to be discussed in detail, together with application to realistic molecular systems exceeding 100 active orbitals.

List of Publications

M. Kállay , P. R. Nagy, D. Mester, Z. Rolik, Gy. Samu, J. Csontos, J. Csóka, P. B. Szabó, L. Gyevi-Nagy, B. Hegely, I. Ladjánszki, L. Szegedy, B. Ladóczki and others.

The MRCC program system: accurate quantum chemistry from water to proteins

J. Chem. Phys., In press, (2020).

B. Ladóczki and S. L. Ten-no,

Epstein-Nesbet corrections of the initiator error and their applications

In preparation, (2020).

B. Ladóczki and S. L. Ten-no,

Stochastic perturbation theory in a limited configuration space,

J. Chem. Phys., **151**, 114113 (2019).

B. Ladóczki, C. Fernandez, O. Moya, P. Babarczi, J. Tapolcai and D. Guija

Robust Network Coding in transport networks

IEEE, **9781467371315**, 27-28 (2015).

Z. Rolik, L. Szegedy, I. Ladjánszki, B. Ladóczki, and M. Kállay

An efficient linear-scaling CCSD(T) method based on local natural orbitals

J. Chem. Phys., **139**, 094105 (2013).

Conferences

B. Ladóczki, J. Tapolcai, A. Pasic, ”Monitoring-Flow Based Network Verification and Failure Localization in Software Defined Networks” IEEE, July 7-12 2019, Paris

Ladoczki Bence、Seiichiro L. Ten-no, ”Stochastic perturbation theory in a limited configuration space” 第22回理論化学討論会, 2019年5月27-29日, 北海道大学 学術交流会館

Ladoczki Bence、天能精一郎「イニシエーター近似の摂動補正に関する研究」第12回分子科学討論会, 2018年9月10-13日, 福岡国際会議場

Ladoczki Bence、天能精一郎「確率論的手法とUnlinked図形の寄与に関する研究」第21回理論化学討論会, 2018年5月15日-17日、岡崎コンファレンスセンター

ラドーツキ　ベンツェ、天能精一郎　「確率論的手法に基づいた摂動論における打ち切りの影響に関する研究」、第11回分子科学討論会, 2017年9月15-18日, 東北大学川内北キャンパス

Bence Ladoczki, Seiichiro Ten-no, ”Stochastic perturbation theories based on the MSQMC method”, 第20回理論化学討論会, 2017年5月16-18日, 京都大学吉田キャンパス

B. Ladóczki, C. Fernandez, O. Moya, P. Babarczi, J. Tapolcai, D. Guija, ”Robust Network Coding in transport networks” IEEE, Apr 26 - May 1 2015, Hong Kong

Acknowledgements

The calculations in this work were performed using the Kobe University's Research Computing cluster (titan) and the Kei computer.

My sincere thanks also goes to Professor Ten-no Seiichiro, Professor Motoyuki Uejima, and Professor Takashi Tsuchimochi, who provided me an opportunity to join their team as research student in 2014, and who gave access to the laboratory and computational facilities. Without they precious support it would not be possible to conduct this research.

Doctor Thesis, Kobe University
“Developments of stochastic perturbation theory and perturbative corrections of model space quantum Monte Carlo”, 81 pages
Submitted on January 24, 2020.
The date of publication is printed in cover of repository version published in Kobe University Repository Kernel.