Sign Learning Kink based (SiLK) Quantum Monte Carlo Applied to H_2O, N_2, F_2 Systems

Xiaoyao Ma
Doctoral Dissertation Committees:
Mark Jarrell
Juana Moreno
Randall W.Hall
Phillip T. Sprunger
Ambar N. Sengupta
Bala "Ramu" Ramachandran
Frank Löffler

March 8, 2020

Outline
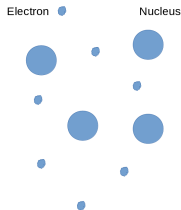
Goal

Correlated electronic systems (the Coulomb interaction energy of electrons is not weak!)

- Magnetic systems: Coulomb interaction between electrons leads a variety of spin ordering patterns(ferromagnetism, antiferromagnetism)
- Catalytic properties of transition metal oxides (industrially important!)
- Energy materials

Introduction

$$\hat{H}_{all} = -\sum_{i=1}^{N} \frac{1}{2} \nabla_i^2 - \sum_{I=1}^{N_{nuclei}} \frac{1}{2M_I} \nabla_I^2 - \sum_{i=1}^{N} \sum_{I=1}^{N_{nuclei}} Q_I / r_{iI}$$

$$+ \sum_{i=1}^{N} \sum_{j=i+1}^{N} \frac{1}{r_{ij}} + \sum_{I=1}^{N_{nuclei}} \sum_{J=1}^{N_{nuclei}} \frac{Q_I Q_J}{R_{IJ}}$$

Introduction

The mass of nucleus is much heavier than the mass of electron
Born Oppenheimer Approximation

$$\hat{H} = -\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N} \sum_{I=1}^{N_{nuclei}} Q_I/r_{iI} + \sum_{i=1}^{N} \sum_{j=i+1}^{N} \frac{1}{r_{ij}}$$

Introduction

We combine the one-electron term (including its kinetic energy and interaction to the nuclei) into a single operator $\hat{h}(i)$ for the ith electron

$$\hat{h}(i) = -\frac{1}{2}\nabla_i^2 - \sum_{I=1}^{N_{nuclei}} Q_I/r_{iI}$$

Then Hamiltonian operator can be written as

$$\hat{H} = \sum_{i=1}^{N} \hat{h}(i) + \sum_{i=1}^{N} \sum_{j=i+1}^{N} \frac{1}{r_{ij}}$$

Hartree Fock Approximation

HF equation:

$$\hat{f}(i)\chi(x_i) = \epsilon\chi(x_i)$$

$$\hat{f}(i) = \hat{h}(i) + v^{HF}(i)$$

Hartree Fock potential $v^{HF}(i)$ is the average potential experienced by ith electron due to the presence of the other electrons.

Procedure: Initial guess of spin orbitals, calculate HF potential v^{HF}, solving HF equation for a new set of spin orbitals, calculate v^{HF} ...until self consistency is reached.

$$E_{corr} = E_{exact} - E_{HF}$$

- E_{corr} is the correlation energy
- E_{exact} is the exact energy
- E_{HF} is the Hartree Fock energy

Different post-HF methods can catch different amount of E_{corr}

Introduction

- Full Configuration Interaction (FCI): all excitations from Hartree Fock (HF) state
- Singles and Doubles (SD): single and double excitations from HF
- Singles, Double, and Triples (SDT): single, double and triple excitations from HF

Introduction
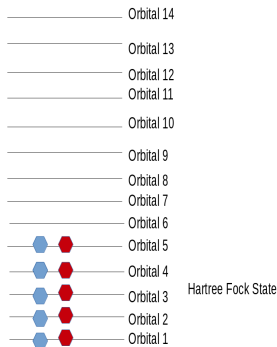
- FCI
- Coupled Cluster Singles Doubles (CCSD)
- Coupled Cluster Singles, Doubles and Triples (CCSDT)
- Coupled Cluster Singles, Doubles plus perturbative Triples (CCSD(T))

$$|\psi_{ccsd}\rangle = e^{\hat{T}_1+\hat{T}_2}|\psi_{HF}\rangle$$

$$|\psi_{ccsdt}\rangle = e^{\hat{T}_1+\hat{T}_2+\hat{T}_3}|\psi_{HF}\rangle$$

Methods	Scaling	Accuracy (Hartree)
HF	$\mathcal{O}\left[N^3\right]$	10^{-1}
MBPT(2)	$\mathcal{O}\left[N^4\right]$	10^{-2}
CCSD	$\mathcal{O}\left[N^6\right]$	10^{-3}
CCSD(T)	$\mathcal{O}\left[N^7\right]$	10^{-4}
CCSDT	$\mathcal{O}\left[N^8\right]$	10^{-4}
FCI	$\mathcal{O}\left[e^N\right]$	0

Table: Assuming (N=M), M: number of spacial orbitals, N: number of electrons.

Question: Why we care chemical accuracy?

Arrhenius equation

$$aA + bB \rightarrow cC, Rate = -\frac{1}{a}\frac{d[A]}{dt} = -\frac{1}{b}\frac{d[B]}{dt} = \frac{1}{c}\frac{d[C]}{dt}$$
$$Rate = k[A]^s[B]^t$$

[A] and [B] are the concentrations of the reactants, s and t are the reaction order, k is a rate constant.

$$k_0 = Ae^{-\frac{E_a}{RT}}, k = Ae^{-\frac{E_a+\Delta}{RT}} = k_0 e^{-\frac{\Delta}{RT}}$$

T=300K

- $e^{-\frac{\Delta}{RT}}$=5.14×10^{-8} , Δ=10 kcal/mol
- $e^{-\frac{\Delta}{RT}}$=0.18669 , Δ=1 kcal/mol
- $e^{-\frac{\Delta}{RT}}$=0.99895 , Δ= 0.627 cal/mol

T=1K

- $e^{-\frac{\Delta}{RT}}$=0, Δ=10 kcal/mol
- $e^{-\frac{\Delta}{RT}}$=0, Δ=1 kcal/mol
- $e^{-\frac{\Delta}{RT}}$=0.72910 , Δ= 0.627 cal/mol

Quantum Monte Carlo

The partition function Q:

$$Q = Tr\left(e^{-\beta H}\right)$$

We discretize the β by P times , $\beta/P = \Delta\tau$

$$\exp(-\Delta\tau H) \approx 1 - \Delta\tau H, \quad \Delta\tau \ll 1$$

$$e^{-\beta H} = \left(e^{-\beta H/P}\right)^P = \left(e^{-\Delta\tau H}\right)^P \approx (1 - \Delta\tau H)^P + O(\Delta\tau).$$

$$1 = \sum_{j_i} |\alpha_{j_i}\rangle\langle\alpha_{j_i}|.$$

$$
\begin{aligned}
Q(P) &= \sum_{j_1, j_2, \dots j_P} \langle\alpha_{j_1}| \exp(-\Delta\tau H)|\alpha_{j_2}\rangle \langle\alpha_{j_2}| \exp(-\Delta\tau H)|\alpha_{j_3}\rangle \cdots \\
&\quad \langle\alpha_{j_P}| \exp(-\Delta\tau H)|\alpha_{j_1}\rangle
\end{aligned}
$$

Hamiltonian dimension (N)=10^6, P= 10^9, Phase space $N^P = 10^{6*10^9}$.
Questions: How to perform QMC when not all weights are postive? How do we sample the important states?

The Minus Sign Problem

- In Mapping of Quantum to Classical system

$$Q = Tr\left(e^{-\beta H}\right) = \sum_c p_c$$

$$<A> = \frac{Tr\left[A \exp\left(-\beta H\right)\right]}{Tr\left[\exp\left(-\beta H\right)\right]} = \frac{\sum_c A_c p_c}{\sum_c p_c}$$

- Sign Problem if $p_c < 0$
  - Can not interpret p_c as probabilities
  - Appears in simulation of fermionic system due to fermion exchange

Sampling Absolute Probability

Strategy: perform simulation using absolute values ($|p_c|$) and keep track of the sign

$$
\begin{aligned}
< A > &= \frac{\sum_c A_c p_c}{\sum_c p_c} \\
&= \frac{\frac{\sum_c A_c \, Sign(p_c)|p_c|}{\sum_c |p_c|}}{\frac{\sum_c Sign(p_c)|p_c|}{\sum_c |p_c|}}
\end{aligned}
$$

$$
< A > = \frac{< A \cdot Sign(p_c) >_{|p_c|}}{< Sign(p_c) >_{|p_c|}}
$$
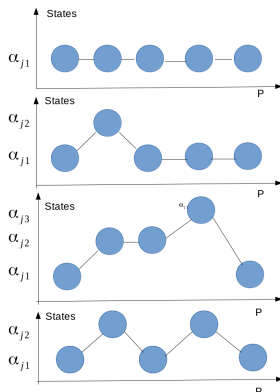
Kink definition

Figure: Kink definition. We define a matrix element $\left\langle \alpha_{j_1} | \exp(-\frac{\beta}{P}H) | \alpha_{j_2} \right\rangle$ with $\alpha_{j_1} \neq \alpha_{j_2}$ as a kink.

Formalism

$$
Q(P) = \sum_{j_1} \left[\left\langle \alpha_{j_1} | e^{-\frac{\beta}{P}H} | \alpha_{j_1} \right\rangle \right]^P +
$$

$$
\sum_{j_1 j_2} \sum_{n=0}^{P-2} \left[\left\langle \alpha_{j_1} | e^{-\frac{\beta}{P}H} | \alpha_{j_1} \right\rangle \right]^n \left[\left\langle \alpha_{j_2} | e^{-\frac{\beta}{P}H} | \alpha_{j_2} \right\rangle \right]^{P-2-n} \times
$$

$$
\left[\left\langle \alpha_{j_1} | e^{-\frac{\beta}{P}H} | \alpha_{j_2} \right\rangle \right]^2 + \cdots.
$$

$$
x_j \approx \langle \alpha_j | 1 - \Delta\tau H | \alpha_j \rangle
$$

$$
x_j = 1 - \Delta\tau H_{\alpha_j,\alpha_j}
$$

$$
t_{ij} = \left\langle \alpha_i | e^{-\Delta\tau H} | \alpha_j \right\rangle = -\Delta\tau \langle \alpha_i | H | \alpha_j \rangle = -\Delta\tau H_{\alpha_i,\alpha_j}
$$

t_{ij} is defined as a kink

Formalism

$$Q = \sum_{j_1} [x_{j_1}]^P + \sum_{j_1, j_2} \sum_{n=0}^{P-2} [x_{j_1}]^n [x_{j_2}]^{P-2-n} [t_{j_1 j_2}]^2 + \cdots$$
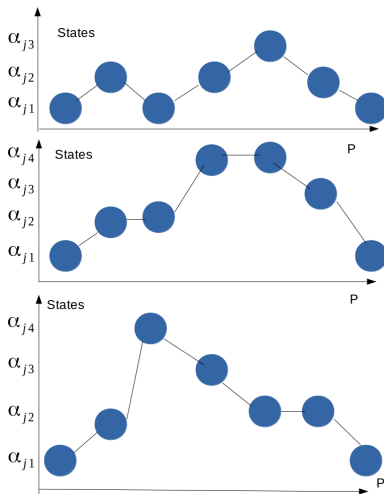
So we can write Q as summation of terms with zero, two, three ... kinks.

$$Q = Q_0 + Q_2 + Q_3 + \cdots + Q_P$$

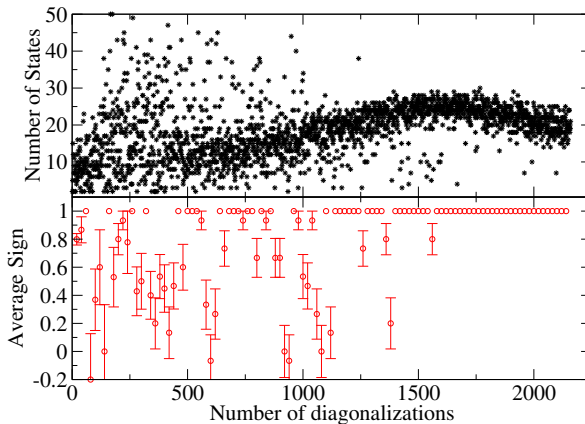$$Q(P) = \sum_{n=1}^{P} \left(\prod_{i=1}^{n} \sum_{j_i} \right) \rho_n(\{\alpha_j\})$$

In this equation, n represents the number of kinks, ρ_n is the weight corresponding to n kinks and compose of a set of states $\{\alpha_j\}$.
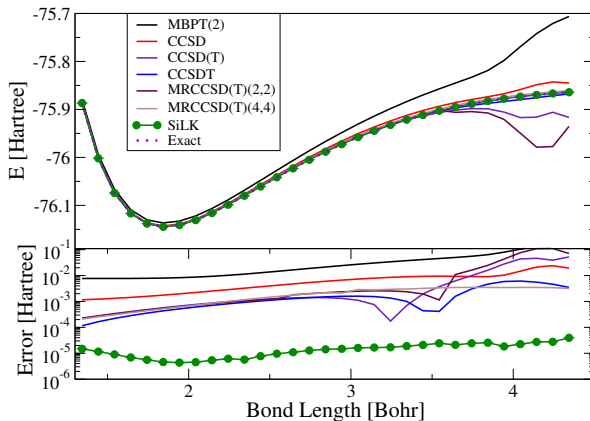
Formalism: Improving the basis

- Diagonalize Hamiltonian in those population states
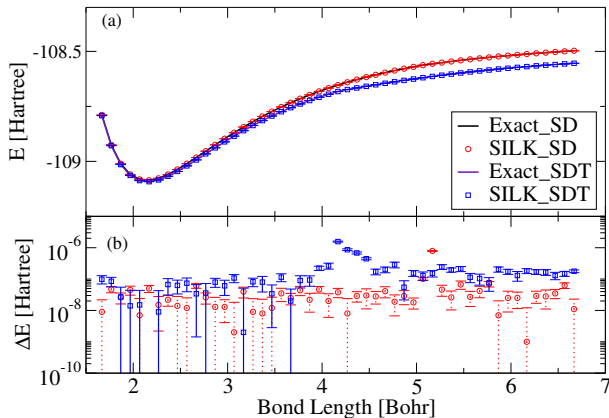- The new states will be linear combination of the old states

Dimension vector space used in FCI= 128829

Water : SiLK VS Coupled Cluster Methods
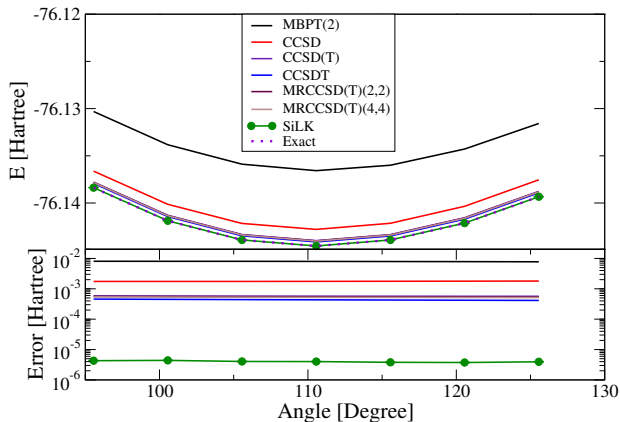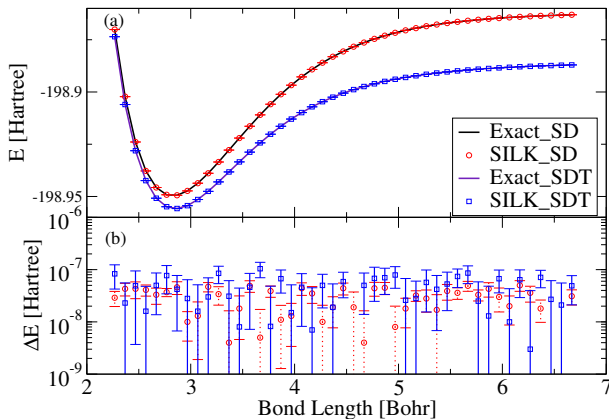
Dimension vector space used in SD= 898
Dimension vector space used in SDT= 14642

Dimension vector space used in SD= 1196
Dimension vector space used in SDT= 24980

Conclusion

- SiLK QMC is feasible for both equilibrium geometry and non-equilibrium geometry
- SiLK QMC indeed addresses the minus sign problem by improving the basis
- SiLK QMC is variational
- SiLK QMC can beat approximate single-reference Coupled Cluster methods for extended bond length

Possible project directions

- Develop a massively parallel implementation
- Develop analytical theories to calculate energy using SiLK optimized wavefunctions
- Add the functionality to allow finite temperature calculations

Back up: Energy conversion

Chemical accuracy:
1 kcal/mol=0.0016 Hartree =4.184 KJ/mol
SiLK accuracy:
10^{-6} Hartree= 0.6275 cal/mol= 2.6255 J/mol

Scaling in detail

General speaking, n: number of occupied spacial orbitals, N: number of unoccupied spacial orbitals .

- HF $\mathscr{O}\left[(N+n)^3\right]$, 10^{-1} Hartrees
- MBPT(2) $\mathscr{O}\left[n^2N^2\right]$, 10^{-2} Hartrees
- CCSD $\mathscr{O}\left[n^2N^4\right]$, 10^{-3} Hartrees
- CCSD(T) $\mathscr{O}\left[n^2N^4N_{iteration} + n^3N^4\right]$, 10^{-4} Hartrees
- CCSDT $\mathscr{O}\left[n^3N^5\right]$, 10^{-4} Hartrees
- FCI $\mathscr{O}\left[e^N\right]$

$N_{iteration}$: number of iterations to reach convergence

Double-Zeta basis

It contains two basis functions for each atomic orbital. In H_2O, $n < 3$

- Oxygen : $1s, 2s, 2p_0, 2p_{-1}, 2p_{+1}$
- Hydrogen 1 : $1s$
- Hydrogen 2 : $1s$

There are 7 (5+2) orbitals in each basis function. Double-Zeta basis contains 14 orbitals.

Double-Zeta basis

H: 2 ; H: 2 ; O: 4+2×3=10

Multi-reference CCSD(T)

- MRCCSD(T)$_{(2,2)}$
- MRCCSD(T)$_{(4,4)}$

Model space: (m,n) m: the number of orbitals n: the number of electrons
(2,2) has 4 references
(4,4) has 36(6×6) references
Single reference CC only has one reference (Hartree Fock State)
HOMO(H): Highest occupied molecular orbital
LUMO(L): Lowest unoccupied molecular orbital
(2,2): two orbitals [H,L]
(4,4): four orbitals [H-1,H,L,L+1]

HOMO-LUMO

Orbital 14
Orbital 13
Orbital 12
Orbital 11
Orbital 10
Orbital 9
Orbital 8
Orbital 7 ← LUMO+1
Orbital 6 ← LUMO
Orbital 5 ← HOMO
Orbital 4 ← HOMO-1
Orbital 3
Orbital 2
Orbital 1

Hartree Fock State
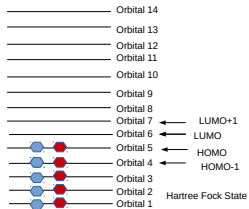
- Triple bond occupied by three pairs of electrons in equilibrium.
- Dissociation of N_2 is bond breaking process.

- SiLK is feasible for finite temperature calculations.
- SiLK can perform geometry optimization in future.
- SiLK has potential to attack larger systems with high chemical accuracy by overcoming the memory bottleneck.

SiLK VS post HF methods

- Higher accuracy
- Variational
- SiLK can treat electron correlation exactly, includes the effects of temperature
- most post HF methods can not properly describe multi-reference electronic correlation effects (SiLK does not suffer problems from (near) degeneracies)

FCIQMC

Using projector P stochastically on a trial state Ψ_T to project out the ground state

$$\lim_{n \to \infty} P^n |\Psi_T\rangle \propto |\Psi_0\rangle$$

$$P_{ij} = \delta_{ij} - \tau(H_{ij} - E_T \delta_{ij})$$

$$\Psi_0 = \sum_i C_i |\Phi_i\rangle$$

$$C_i(\infty) \propto \sum_{n=0}^{\infty} w_i(n\tau)$$

$$
\begin{aligned}
C_i(t + \tau) &= \sum_j P_{ij} C_j(t) \\
&= (1 - (H_{ii} - E_T)) C_i(t) + \sum_{j \neq i} -\tau H_{ij} C_j(t).
\end{aligned}
$$

Correlation energy

Geometry	E_{corr}(Hartree)
R_e	-0.148028
$1.5R_e$	-0.210992
$2R_e$	-0.310076

Data from Chem. Phys. Lett. 95, 386 (1983)

- dynamical correlation energy : neglect electrons avoiding each other (inclusion of interaction that normally would not occur)
- nondynamical correlation energy: inadequacy of a single reference in describing a given molecular state (Bond breaking case)

Size consistence and Variational

Methods	Dynamical	Nondynamical	Size consistent	Variational
HF	None	None	Yes	Yes
MBPT2	Fair	None	Yes	No
CISD	Good	Fair	No	Yes
CCSD	Very Good	Fair	Yes	No
CCSD(T)	Excellent	Good	Yes	No
SiLK	Excellent	Excellent	Yes	Yes

Dynamical: Dynamical correlation
Nondynamical: Nondynamical correlation
CISD: Configuration Interaction Singles and Doubles