QUANTUM MONTE CARLO FOR ELECTRONIC STRUCTURE

ANOUAR BENALI, PH.D
Computational Scientist
Computational Science Division
Aerial view of Argonne National Laboratory

Theory and Computing Sciences Building
This drawing depicts the historic event on Dec. 2, 1942, when a group of 49 scientists led by Enrico Fermi created the world's first controlled, self-sustaining nuclear chain reaction underneath the University of Chicago's Stagg Field football stadium.
MAJOR SCIENTIFIC USER FACILITIES AT ARGONNE

- Advanced Photon Source
- Argonne Tandem Linear Accelerator System
- Center for Nanoscale Materials
- Argonne Leadership Computing Facility
- Electron Microscopy Center

IBM MIRA
COMPUTATIONAL SCIENCE AT THE ARGONNE LEADERSHIP COMPUTING FACILITY

**Mira** (Production) – IBM BlueGene/Q
- 49,152 nodes / 786,432 cores
- 768 TB of memory
- Peak flop rate: 10 PF
- Linpack flop rate: 8.1 PF

**Theta** (06/17) – Cray/ Intel Xeon Phi (Knights’ Landing – KNL)
- 3624 nodes / 231,936 cores
- 735 TB of memory
- Peak flop rate: 9.65 PF
COMPUTATIONAL SCIENCE AT THE ARGONNE LEADERSHIP COMPUTING FACILITY

One of the Most Advanced Supercomputers Ever Designed
A collaboration with ANL, Intel and Cray to accelerate discovery & innovation

1st exascale machine in the world
(unless the Chinese beat us.. Again..)

Source: Argonne National Laboratory and Intel. *Versus ANL’s current biggest system name MIRA (19PFs and 3.9MW)
Other names and brands may be claimed as the property of others.
Outline

Lecture 1

• Brief introduction to Electronic Structure
• Monte Carlo Simulation
• Quantum Monte Carlo (QMC) Simulations
  - Variational Monte Carlo (VMC)
  - Diffusion Monte Carlo (DMC)
  - Trial WaveFunction
  - Advanced trial wavefunctions
Outline

Lecture 3

• QMC implementation in QMCPACK
• Performance and readiness to Exascale
Outline

Lecture 5

• Examples on solids and molecules
• Tutorial
  • The Case of Beryllium dimer from HF to CI
  • What is the benefit of DMC.
• Discussion
Outline

Lecture 5

Please Download and try to install the following codes by following instructions on web pages:

QMCPACK: https://github.com/QMCPACK/qmcpack
Pyscf : https://github.com/sunqm/pyscf
Quantum Package: https://github.com/LCPQ/quantum_package

QMCPACK Documentation: qmcpack.org
ELECTRONIC STRUCTURE
First-Principles Computational Physics/Chemistry

**Chemistry:**
- chemical reactions
- molecular properties
- bonding patterns/formation
- solvation properties
- surfaces

**Schrodinger Equation**
Theory of electrons and ions

\[ i\hbar \frac{\partial}{\partial t} \Psi = \hat{H}\Psi \]

**Exotic electronic phases:**
- Superconductivity
- Magnetism
- Charge ordering

**Plasmas:**
- equilibrium prop.
- kinetics
- instabilities
- shocks

**Thermodynamics:**
- crystal structures
- phase transitions
- strength
- defects and dislocations

**Optical and Transport Properties:**
- spectra
- conductivity
- viscosity
- diffusion
Born-Oppenheimer Approximation

\( i\hbar \partial_t \Psi(\vec{r},t) = \hat{H}\Psi(\vec{r},t) \)

\[
\hat{H} = -\frac{\hbar^2}{2m} \sum_i \vec{\nabla}_i^2 + \sum_{i<j} \frac{e^2}{|\vec{r}_i - \vec{r}_j|} - \frac{\hbar^2}{2M} \sum_i \vec{\nabla}_i^2 + \sum_{i<j} \frac{Z_i e^2}{|\vec{R}_i - \vec{R}_j|}
\]

\( m_{\text{elec}} \ll m_{\text{prot}} \Rightarrow \) they move on different time scales

\[
\Psi(\vec{r},\vec{R}) = \sum_k \xi_k(\vec{R}) \sum_q \psi_{k,q}(\vec{r},\vec{R})
\]

BO leads to independent solutions for electrons and ions (different approximations) !!!

\[
\left[ -\frac{\hbar^2}{2m} \vec{\nabla}_i^2 + V_{\text{xc}}(\vec{r}_i,\vec{R}) + V_{\text{coul}}(\vec{r}_i,\vec{R}) \right] \psi_q(\vec{r}) = E_{q,\text{BO}} \psi_q(\vec{r})
\]

If finite gap between ground state and 1st excited state, assume electrons remain in the ground state! (Electrons at T=0 K !)

\[
\left[ -\frac{\hbar^2}{2m} \vec{\nabla}_r^2 + V_{\text{coul}}(\vec{R}) + E_{\text{BO}}(\vec{R}) \right] \xi(\vec{R}) = E \xi(\vec{R})
\]

Solve ionic problem !
(VERY) APPROXIMATE SOLUTION

- Hartree Fock

Simplest approximate theories for solving the many-body Hamiltonian.

Wavefunction is given by a single Slater determinant of N spin-orbitals

\[ \Psi = \begin{vmatrix} \psi_1(x_1) & \psi_1(x_2) & \cdots & \psi_1(x_N) \\ \psi_2(x_1) & \psi_2(x_2) & \cdots & \psi_2(x_N) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_N(x_1) & \psi_N(x_2) & \cdots & \psi_N(x_N) \end{vmatrix} , \]

\[ \Psi(x_1, x_2, \ldots, x_i, \ldots, x_j, \ldots, x_N) = -\Psi(x_1, x_2, \ldots, x_j, \ldots, x_i, \ldots, x_N) . \]

x: spatial and spin variable

Anti-symmetric
- Hartree Fock

We choose $\Psi$ the to be an orthonormal set. We introduce a Lagrange multiplier $\epsilon_i$ to impose the condition that the $\Psi$ are normalised, and minimise with respect to the $\Psi$

$$\frac{\delta}{\delta \psi} \left[ \langle \hat{H} \rangle - \sum_j \epsilon_j \int |\psi_j|^2 \, dr \right] = 0 .$$

A simplification of the expressions for the orbitals $\Psi$ results, resulting to a set of one-electron equations of the form

$$-\frac{1}{2} \nabla^2 \psi_i(r) + V_{ion}(r) \psi_i(r) + U(r) \psi_i(r) = \epsilon_i \psi_i(r) ,$$

where $U(r)$ is a non-local potential and the local ionic potential is denoted by $V_{ion}(r)$. The one-electron equations resemble single-particle Schrödinger equations.

$$\epsilon_i \psi_i(r) = \left( -\frac{1}{2} \nabla^2 + V_{ion}(r) \right) \psi_i(r) + \sum_j \int dr' \frac{|\psi_j(r')|^2}{|r - r'|} \psi_i(r) - \sum_j \delta_{\sigma_i \sigma_j} \int dr' \frac{\psi_j^*(r') \psi_i(r')}{|r - r'|} \psi_j(r) .$$

Kinetic  | Electron-ion potential  | Electrostatic potential (self interaction)  | Exchange term (correction for self interaction)
- Hartree Fock

By assuming a single-determinant form for the wavefunction, neglects correlation between electrons.

The electrons are subject to an *average* non-local potential arising from the other electrons, which can lead to a poor description of the electronic structure.

Although qualitatively correct in many materials and compounds, Hartree-Fock theory is insufficiently accurate to make accurate quantitative predictions.
**LESS** APPROXIMATE SOLUTION

- Density Functional Theory (DFT)

The Hohenberg-Kohn theorem states that if N interacting electrons move in an external potential $V_{\text{ext}}$, the ground-state electron density minimises the functional

$$E[n] = F[n] + \int n(r)V_{\text{ext}}(r)\,dr$$

Where $F$ is a universal functional of $n$ and the minimum value of the functional $E$ is $E_0$ the exact ground-state electronic energy.

Levy gave a very simple proof of the Hohenberg and Kohn theorem leading to

$$\hat{F} = \sum_i -\frac{1}{2} \nabla^2_i + \frac{1}{2} \sum_{i \neq j} \frac{1}{|r_i - r_j|} .$$

Considering an N-electron ground state wavefunction $\Psi_0$ which yields a density $n_0$ and minimises $\langle \Psi | \hat{F} | \Psi \rangle$, then from the definition of the functional

$$E[n(r)] = F[n(r)] + \int n(r)V_{\text{ext}}(r)\,dr = \langle \Psi | F + V_{\text{ext}} | \Psi \rangle .$$

$E[n(r)]$ obeys the variational principle. (first part of the demonstration, second part shows $E[n_0(r)] = E_0$.)
Density Functional Theory (DFT)

Kohn and Sham derived a coupled set of differential equations enabling the ground state density \( n_0(r) \) to be found.

Kohn and Sham separated \( F[n(r)] \) into distinct parts, so that the functional \( E \) becomes

\[
E[n(r)] = T_s[n(r)] + \frac{1}{2} \int \int \frac{n(r)n(r')}{|r - r'|} drdr' + E_{XC}[n(r)] + \int n(r)V_{ext}(r)dr ,
\]

where \( T_s[n(r)] \) is defined as the kinetic energy of a non-interacting electron gas with density \( n(r) \) (not for real system)

\[
T_s[n(r)] = -\frac{1}{2} \sum_{i=1}^{N} \int \psi_i^*(r) \nabla^2 \psi_i(r) dr ,
\]
- Density Functional Theory (DFT)

Introducing a normalisation constraint on the electron density, and rewriting everything in terms of an effective potential

\[
\frac{\delta T_F[n(r)]}{\delta n(r)} + V_{\text{eff}}(r) = \mu ,
\]

Where

\[
V_{\text{eff}}(r) = V_{\text{ext}}(r) + \int \frac{n(r')}{|r - r'|} dr' + V_{XC}(r)
\]

and

\[
V_{XC}(r) = \frac{\delta E_{XC}[n(r)]}{\delta n(r)} .
\]

To find the ground state energy, \(E_0\), and the ground state density, \(n_0\), the one electron Schrödinger equation

\[
\left( -\frac{1}{2} \nabla_i^2 + V_{\text{eff}}(r) - \epsilon_i \right) \psi_i(r) = 0
\]

Solved self-consistently with

\[
n(r) = \sum_{i=1}^{N} |\psi_i(r)|^2 ,
\]

Exact Formalism! BUT \(E_{XC}\) unknown!!
DFT is a powerful, formally exact solution to the Schrödinger equation.

DFT has come to prominence over the last decade as a method *potentially* capable of very accurate results at low cost.
ALMOST EXACT SOLUTION

- Configuration Interaction (CI)

- Conceptually simplest methods for solving the many-body Hamiltonian.
- Theoretically elegant
- In principle exact
- (Relatively) simple to implement
- In practice \textit{full} CI Extremely expensive

An \textit{exact} many-body wavefunction, \( \Psi \), may be written as a linear combination of Slater determinants, \( D_k \),

\[ \Psi = \sum_{k=0}^{\infty} c_k D_k \]

Where \( D_k \) is the fully span the Hilbert space of the wavefunction. The determinants can be any complete set of \( N \)-electron antisymmetric functions but are typically constructed from Hartree-Fock orbitals such that is the ground-state Hartree-Fock determinant.

The complexity of the methods scales as (\( N \) number of electrons and \( M \) the number of basis functions).

\[ k_{\text{max}} = \frac{M!}{N!(M-N)!} \]
First-Principles Electronic Structure Theory

$$\left[ -\frac{\hbar^2}{2m} \nabla^2 + V_{e-e}(\vec{r},\{\vec{R}\}) + V_{e-l}(\vec{r},\{\vec{R}\}) \right] \psi_q(\vec{r}) = E^\text{BO}_q \psi_q(\vec{r})$$

Only input:
• $N\alpha, R\alpha, Z\alpha$

Accuracy (eV)

- exact
- 0.001-0.01
- 0.05-0.1
- 0.5+

System size (Nb electrons)

- Cl
- CC
- QMC $N^2-4$
- DFT $N^3$
- TB $N^3$

Density Functional Theory
+ Mean-field cost, well developed codes, easy to learn
- Hard to improve, accuracy is unknown, fails for correlated problems, Not HPC

Quantum Chemistry Methods
+ Very accurate and robust, black box codes.
- High cost (> $N^5$), very hard for solids, complicated theory, Not HPC.

Quantum Monte Carlo
+ Accurate, systematically improvable, Trivially parallel HPC intensive, scales as $N^{2-4}$.
- High computer cost, No forces, problems at high Z (No spin-orbit coupling).
MONTE CARLO SIMULATIONS

Monte Carlo methods were first developed as a method for estimating integrals that could not be evaluated analytically.
Markov chain MC or Random Walk

- Markov chain is a random walk through phase space:
  \[ s_1 \rightarrow s_2 \rightarrow s_3 \rightarrow s_4 \rightarrow \ldots \]
  Here “s” is the state of the system.
- ALL QMC is some type of Markov process. VMC is the simplest.
- The transition probability is \( P(s_n \rightarrow s_{n+1}) \) a stochastic matrix
  \[ P(s \rightarrow s') \geq 0 \quad \sum_s P(s \rightarrow s') = 1 \]
- In a Markov chain, the distribution of \( s_{n+1} \) depends only on \( s_n \) (by definition). A drunkard has no memory!
- Let \( f_n(s) \) be the probability after “n” steps. It evolves according to a “master equation.”
  \[ f_{n+1}(s') = \sum_s f_n(s) P(s \rightarrow s') \]
  \[ f_{n+1} = P f_n \]
- The stationary states are eigenfunctions of \( P \): \( P f = \lambda f \)
In physics and thermodynamics, the ergodic hypothesis says that, over long periods of time, the time spent by a system in some region of the phase space of microstates (s) with the same energy is proportional to the volume of this region, i.e., that all accessible microstates are equiprobable over a long period of time.

Statistically, for this random process, the time average of one sequence of events is the same as the ensemble average.

=> For a Markov Chain, as one increases the steps, there exists a positive probability measure at step n that is independent of probability distribution at initial step 0
Metropolis algorithm

Three key concepts:

1. Sample by using an ergodic random walk.
2. Determine equilibrium state by using detailed balance.
3. Achieve detailed balance by using rejections.

Equilibrium probability Distribution

**Detailed balance** \( \pi(s) P(s \rightarrow s') = \pi(s') P(s' \rightarrow s) \).

*Rate balance from \(s\) to \(s'\).*

Put \(\pi(s)\) into the master equation. *(Or sum above Eq. on \(s\).)*

\[
\sum_{s} \pi(s) P(s \rightarrow s') = \pi(s') \sum_{s} P(s' \rightarrow s) = \pi(s')
\]

- Hence, \(\pi(s)\) is an eigenfunction.
- If \(P(s \rightarrow s')\) is ergodic, \(\pi(s)\) is unique steady state solution.

Stationary distribution
Rejection Method

Metropolis achieves detailed balance by *rejecting* moves.

*General Approach:*
1. Choose distribution to sample, e.g., $\pi(s) = \exp[-\beta H(s)]/Z$
2. Impose detailed balance on transition: $K(s \rightarrow s') = K(s' \rightarrow s)$
   where $K(s \rightarrow s') = \pi(s) P(s \rightarrow s')$
   *(probability of being at $s$) * (probability of going to $s'$).*
3. Break up transition probability into sampling and acceptance:
   $P(s \rightarrow s') = T(s \rightarrow s') A(s \rightarrow s')$
   *(probability of generating $s'$ from $s$) * (probability of accepting move)*

The optimal acceptance probability that gives detailed balance is:

$$A(s \rightarrow s') = \min[1, \frac{T(s' \rightarrow s) \pi(s')}{T(s \rightarrow s') \pi(s)}] = \min[1, \frac{\pi(s')}{\pi(s)}]$$

Normalization of $\pi(s)$ is not needed or used!  

If $T$ is constant!
The “Classic” Metropolis method

*Metropolis-Rosenbluth\(^2\) -Teller\(^2\) (1953) method is:*

- Move from \(s\) to \(s'\) with probability \(T(s \rightarrow s') = \text{constant}\)
- Accept with move with probability:

\[
A(s \rightarrow s') = \min [ 1, \exp (- (E(s') - E(s))/k_B T ) ]
\]

- Repeat many times
- Given ergodicity, the distribution of \(s\) will be the canonical distribution: \(\pi(s) = \exp(-E(s)/k_B T)/Z\)
- **Convergence is guaranteed but the rate is not!**
How to sample

$$S_{\text{new}} = S_{\text{old}} + \Delta \cdot (\text{sprng} - 0.5)$$

Uniform distribution in a cube of side “$\Delta$”.

**Note:** It is more efficient to *move one particle at a time* because only the energy of that particle comes in and the *acceptance ratio will be larger.*

$$A(s \rightarrow s') = \exp[-\beta(V(s') - V(s))]$$

$$= \exp[-\beta \sum_{j \neq i} (v(\mathbf{r}_i' - \mathbf{r}_j) - v(\mathbf{r}_i - \mathbf{r}_j))]$$

*For $V$ with cut-off range, difference is local.*
call \texttt{initstate}(s\_old) \quad \text{Initialize the state}

E\_old = \textbf{action}(s\_old)

\textbf{LOOP}\{ 
 
\quad \text{call \texttt{sample}(s\_old,s\_new,T\_new,1)} \quad \text{Sample snew}

\quad E\_new = \textbf{action}(s\_new) \quad \text{Trial action}

\quad \text{call \texttt{sample}(s\_new,s\_old,T\_old,0)} \quad \text{Find prob. of going backward}

\quad A=\exp(-E\_new+E\_old) \frac{T\_old}{T\_new} \quad \text{Acceptance prob.}

\quad \text{if}(A\gt\text{sprng}()) \{ \quad \text{Accept the move}

\quad \quad \text{\quad s\_old=s\_new} \quad \text{Collect statistics}

\quad \quad \text{\quad E\_old=E\_new}

\quad \quad \text{\quad naccept=naccept+1} \}

\quad \text{call \texttt{averages}(s\_old)} \}


Estimated Errors

• In what sense do we calculate exact properties? Answer: if we average long enough the error goes to zero, the errors of the simulation are controlled.

• Next, how accurate is the estimate of the exact value?
  – Simulation results without error bars are only suggestive.
    • Without error bars one has no idea of its significance.
    • You should understand formulas and be able to make an “eye-ball” estimate.

• Error bar: the estimated error in the estimated mean.
  – Error estimates based on Gauss’ Central Limit Theorem.
  – Average of statistical processes has normal (Gaussian) distribution.
  – Error bars: square root of the variance of the distribution divided by the number of uncorrelated steps.
QUANTUM MONTE CARLO SIMULATIONS
Variational Monte Carlo

The variational principle of quantum mechanics may be derived by expanding a normalized trial wavefunction, $\Psi_T$, in terms of the exact normalized eigenstates of the Hamiltonian,

$$\psi_T = \sum_{i=0}^{\infty} c_i \psi_i$$

where the expansion coefficients, $c_i$, are normalised

$$\sum_{i=0}^{\infty} |c_i|^2 = 1$$

The expectation of the many-body Hamiltonian, $\hat{H}$, may be evaluated

$$\langle \psi_T | \hat{H} | \psi_T \rangle = \langle \sum_i c_i \psi_i | \hat{H} | \sum_j c_j \psi_j \rangle$$

The expectation value of a trial wavefunction with the Hamiltonian must therefore be greater than or equal to the true ground state energy.

$$\sum_i |c_i|^2 \epsilon_i$$

where

$$\epsilon_i = \langle \psi_i | \hat{H} | \psi_i \rangle$$

Variational calculations depend crucially on the form of trial wavefunction used. By selecting trial wavefunctions on physically motivated grounds, accurate wavefunctions may be obtained.
Before the expectation value of a trial wavefunction with the many-body Hamiltonian may be computed, the integral must be transformed into a form suitable for Monte Carlo integration.

Trial wavefunctions, $\Psi_T$, are dependent on the set of $N$ electron positions, $R = \{r_1, r_2, \ldots, r_N\}$. The expectation value is given by

$$E = \frac{\int \psi_T^* \hat{H} \psi_T dR}{\int \psi_T^* \psi_T dR},$$

which may be rewritten in an importance sampled form in terms of the probability density:

$$\sigma^2 = \frac{\langle \psi_T | (\hat{H} - E_V)^2 | \psi_T \rangle}{\langle \psi_T | \psi_T \rangle} = \langle \hat{H}^2 \rangle - E_V^2.$$

**Variational Principle:** Rigorous upper bound to the energy (= only for exact wfn)

$$E_V[\Psi_T] \geq E_{\text{exact}}$$

**Zero-Variance Principle:** (Only!!!) Exact solution has zero variance

$$E_V \xrightarrow{\sigma^2 \to 0} E_{\text{exact}}$$
for generation = 1 \cdots N_{MC} do
    for walker = 1 \cdots N_w do
        let \( \mathbf{R} = \{\mathbf{r}_1 \ldots \mathbf{r}_N\} \)
        for particle \( i = 1 \ldots N \) do
            set \( \mathbf{r}_i' = \mathbf{r}_i + \delta \)
            let \( \mathbf{R}' = \{\mathbf{r}_1 \ldots \mathbf{r}_i' \ldots \mathbf{r}_N\} \)
            ratio \( \rho = \frac{\Psi_T(\mathbf{R}')}{\Psi_T(\mathbf{R})} \)
            if \( \mathbf{r} \to \mathbf{r}' \) is accepted then
                update state of a walker
            end if
        end for
        Compute \( E_L = \hat{H} \Psi_T(\mathbf{R}) / \Psi_T(\mathbf{R}) \)
    end for
end for
Quantum Monte Carlo Simulations
Diffusion Monte Carlo (DMC)
Diffusion Monte Carlo
(Projector Monte Carlo)

Imaginary time Schrödinger equation \((\tau = it)\)

\[
\frac{\partial |\psi\rangle}{\partial \tau} = -\hat{H} |\psi\rangle,
\]

The state \(|\psi\rangle\) is expanded in eigenstates of the Hamiltonian \(|\psi\rangle = \sum_{i=0}^{\infty} c_i |\phi_i\rangle\), where \(\hat{H} |\phi_i\rangle = \epsilon_i |\phi_i\rangle\)

A formal solution is

\[
|\psi(\tau_1 + \delta \tau)\rangle = e^{-\hat{H} \delta \tau} |\psi(\tau_1)\rangle
\]

where the state \(|\psi\rangle\) evolves from imaginary time \(\tau_1\) to a later time \(\delta \tau\).

If the initial state is expanded in energy ordered eigenstates,

\[
|\psi(\delta \tau)\rangle = \sum_{i=0}^{\infty} c_i e^{-\epsilon_i \delta \tau} |\phi_i\rangle
\]
Any initial state $|\psi\rangle$, that is not orthogonal to the ground state $|\phi_0\rangle$, will evolve to the ground state in the long time limit:

$$\lim_{\tau \to \infty} |\psi(\tau)\rangle = c_0 e^{-\epsilon_0 \tau} |\phi_0\rangle$$

In the DMC method the imaginary time evolution results in excited states decaying exponentially fast, whereas in the VMC method any excited state contributions remain and contribute to the VMC energy.

In position space:

$$\lim_{\tau \to \infty} |\psi(R, \tau)\rangle = c_0 e^{-\epsilon_0 \tau} \phi_0(R)$$

Density of diffusing particles

Branching term: potential-dependent increase or decrease in the particle density.

W. M. C. Foulkes et al., Rev. Mod. Phys. 73, 33 (2001).
Diffusion Monte Carlo
(Projector Monte Carlo)

\[- \frac{\delta \psi(R, \tau)}{\delta \tau} = \left[ \sum_{i=1}^{N} -\frac{1}{2} \nabla_i^2 \psi(R, \tau) \right] + (V(R) - E_T) \psi(R, \tau)\]

Can be solved with MC method but leads to very inefficient algorithm.

The potential $V(R)$ is unbounded in coulombic systems and hence the rate term, $(V(R) - E_T)$, can diverge. Large fluctuations in the particle density leads to impractically large statistical errors.

Linear scaling for bosons (ground state can be made real)

Exponential scaling for fermions

Does not account for the anti fermionic nature of electrons; Exchange of any 2 electrons imposes a change of sign in the wavefunction.
Importance sampling
(trial wavefunction)

We introduce a guiding/trial function $\psi_G(R)$, which closely approximates the ground state.

$$f(R, \tau) = \psi_G(R)\psi(R, \tau)$$

Previous equation becomes

$$-\frac{\delta f(R, \tau)}{\delta \tau} = \sum_{i=1}^{N} -\frac{1}{2} \nabla^2_i f(R, \tau) - \nabla \left[ \frac{\nabla \psi(R)}{\psi(R)} f(R, \tau) \right] + (E_L(R) - E_T) f(R, \tau)$$

- diffusion
- branching
- drift

$E_T$ is a trial energy introduced to maintain normalization of the projected solution at large $\tau$

$E_L$ is a local energy $E_L(R) = \frac{\hat{H}\psi(R)_G}{\psi(R)_G}$ And $E = \frac{1}{N} \sum E_L(R_i)$

Determination of the local energy is one of the most computationally costly operations performed in QMC calculations. Application of the Hamiltonian to the trial wavefunction requires computation of the second derivatives of the wavefunction and the calculation of the electron-electron and electron-ion potentials.
The fixed-node approximation

electron : fermion
(anti-symmetric wavefunction)

The fixed-node approximation

\[ f^{FN}(R, \infty) = \psi_g(R)\psi_0^{FN}(R) \]
From Important sampling, evolves as:
- Diffusion
- Drift
- Branch

**ensemble**
Diffusion Monte Carlo

\[ \text{for generation} = 1 \cdots N_{MC} \text{ do} \]
\[ \text{for walker} = 1 \cdots N_{w} \text{ do} \]

\[ \text{let } R = \{r_1 \ldots r_N\} \]
\[ \text{for particle } i = 1 \cdots N \text{ do} \]
\[ \text{set } r'_i = r_i + \delta \]
\[ \text{let } R' = \{r_1 \ldots r'_i \ldots r_N\} \]
\[ \text{ratio } \rho = \frac{\Psi_T(R')}{\Psi_T(R)} \]
\[ \text{if } r \rightarrow r' \text{ is accepted then} \]
\[ \text{update state of a walker} \]
\[ \text{end if} \]
\[ \text{end for}\{\text{particle}\} \]

\[ \text{Compute } E_I = \hat{H} \frac{\Psi_T(R)}{\Psi_T(R)} \]
\[ \text{Reweight and branch walkers} \]
\[ \text{Update } E_T \]
\[ \text{end for}\{\text{walker}\} \]
\[ \text{end for}\{\text{generation}\} \]
Variational Monte Carlo

for generation = 1 \cdots N_{MC} do
  for walker = 1 \cdots N_w do
    let \( \mathbf{R} = \{\mathbf{r}_1 \ldots \mathbf{r}_N\} \)
    for particle \( i = 1 \ldots N \) do
      set \( \mathbf{r}_i' = \mathbf{r}_i + \delta \)
      let \( \mathbf{R}' = \{\mathbf{r}_1 \ldots \mathbf{r}_i' \ldots \mathbf{r}_N\} \)
      ratio \( \rho = \Psi_T(\mathbf{R}')/\Psi_T(\mathbf{R}) \)
      if \( \mathbf{r} \rightarrow \mathbf{r}' \) is accepted then
        update state of a walker
      end if
    end for
    Compute \( E_L = \hat{H} \Psi_T(\mathbf{R})/\Psi_T(\mathbf{R}) \)
  end for
end for

No Branch

useDrift="yes"
\[
\delta = \tau \nabla_i \ln \Psi_T + \delta_{RN}
\]

useDrift="no"
\[
\delta = \delta_{RN}
\]
Quantum Monte Carlo Simulations
Trial Wavefunction
What do we use for the trial function?

- **Formal requirements** antisymmetry, continuity, finite variance.
- **Mean field approaches** Slater determinant,
- **Expand in basis sets** e.g. multideterminants, CI.
- **Local energy argument**
  - Control singularity at small $r$, or $r_{12}$. Leads to cusp condition.
  - Behavior at large $r$, plasmons, van der Waals (dispersion) interaction
- **Feynman-Kacs formula** gives connection to local energy
  - Jastrow, Backflow, three-body interactions
Flexibility of the form of the trial wavefunction.

In early studies of bosonic $^4$He by McMillan the wavefunction was taken to be a Jastrow or two-body correlation function

$$
\psi = \exp \left[ -\sum_{i<j}^N u(r_{ij}) \right].
$$

$u$ chosen to minimise the energy of the system under consideration, by:

- Increase the probability of particles being at a distance that minimises their interaction energy.

Applied to fermionic systems by multiplying a determinantal wavefunction by a two-body or higher body correlation functions. (more efficient than CI-based approaches.)

The trial wavefunction satisfies as many known properties of the exact wavefunction as possible.

- Anti-symmetric with respect to the exchange of any two electrons.
- Electron-electron and electron-nucleus coalescence. (``cusp conditions`` and are a constraint on the derivatives of the wavefunction).
For particle-particle coalescence, it may be shown that

\[ \frac{d\Psi}{dr} \bigg|_{r=0} = \zeta \Psi|_{r=0}, \]

where \( \Psi \) is the many-body wavefunction and \( r \) is either an electron-electron or electron-ion separation. \( \zeta \) is \(-Z\) for coalescence of an electron with a nucleus of charge \( Z \), and 1/2 or 1/4 for spin parallel and anti-parallel electron-electron coalescence respectively.

and the complete trial wavefunction is then taken to be

\[ \Psi_T(R) = J(R)\Psi_{AS}(R) = e^{J_1 + J_2 + \ldots} \sum_k^{M} C_k D_k^\uparrow(\varphi) D_k^\downarrow(\varphi) \]

**Correlation (Jastrow)**

\[
J_1 = \sum_i^N \sum_{l=1}^{N_{\text{occ}}} u_1(|r_i - r_l|)
\]

\[
J_2 = \sum_{i \neq j}^N u_2(|r_i - r_j|)
\]

**Anti-symmetric function (Pauli principle)**

\[
D_k^\sigma = \begin{pmatrix}
\varphi_1(r_1) & \ldots & \varphi_1(r_{N^\sigma}) \\
\vdots & \ddots & \vdots \\
\varphi_{N^\sigma}(r_1) & \ldots & \varphi_{N^\sigma}(r_{N^\sigma})
\end{pmatrix}
\]

\[
\varphi_i = \sum_{l=N_b}^{l=N_b} C_l^i \Phi_l
\]

Single-particle orbitals
Example of applications
APPLICATION TO MATERIAL SCIENCE AND QUANTUM CHEMISTRY

Stability of TMO (Ti$_4$O$_7$ / TiO$_2$, NiO etc.)

Properties of van der Waals bi-layered materials

Energy barriers in catalysis (Pt surfaces)

Excited States in Molecules

Benchmarking new DFT methods

Metal Organic Framework (Barrier energies)

FIG. 1. Energy barriers in kcal/mol for ethylene hydrogenation with E(TD)-E(R) in red, and E(P)-E(TD) in blue; the error bars on the DMC values are ± one standard deviation.
Advanced Trial Wavefunction
Ideally, all trial wavefunctions should be multi-determinant

\[ \Psi_T(\vec{r}) = e^{i\theta(\vec{r})} \sum_k c_k \sum_q d_{k,q} D_{k,q}^{\uparrow}(\vec{r}^{\uparrow}) D_{k,q}^{\downarrow}(\vec{r}^{\downarrow}) \]

**Multi-Slater Determinant Jastrow**

\[ CSF_k = \sum_q d_{k,q} D_{k,q}^{\uparrow}(\vec{r}^{\uparrow}) D_{k,q}^{\downarrow}(\vec{r}^{\downarrow}) \]

- Configuration State Function (CSF): Space and spin adapted linear combination of Slater determinants.
  - \( \{d_q\} \) are determined by symmetry
- Optimize linear coefficients with VMC.
  - Systematic reduction of fixed-node error!
- Limited applicability in extended systems.
  - Expansion is not size-consistent!
- Recover both static and dynamic correlation
  - Jastrow and linear coefficients must be optimize together to obtain optimal results.
  - Efficient and robust optimization algorithms are crucial!
USING MULTIDETERMINANTS

Strategy: Increase the complexity of the trial wavefunction using a multideterminant expansion (MCSCF)

G1 test set

CONFIGURATION INTERACTION

Virtual orbitals

Occupied orbitals

Ground State

55
CONFIGURATION INTERACTION

Virtual orbitals

Occupied orbitals

Ground State

Single excitations

Single excitations
CONFIGURATION INTERACTION

Virtual orbitals

Occupied orbitals

Ground State  Double excitations  Double excitations  Double excitations  Double excitations
CONFIGURATION INTERACTION

Virtual orbitals

Occupied orbitals

Ground State  Triple excitations  Triple excitations  Triple excitations

Triple excitations
And so on…
And so on…

Not all the determinants have the same weight/importance and their contributions to the energy become smaller and smaller as the order of excitations increases.
And so on…

Not all the determinants have the same weight/importance and their contributions to the energy become smaller and smaller as the order of excitations increases.

Wouldn’t it be great to just select the determinants with the largest contributions??!!
Iterative perturbation calculations of ground and excited state energies from multiconfigurational zeroth-order wavefunctions

B. Huron
Department de Physique Stellaire, Observatoire de Meudon, 92 France

J. P. Malrieu
Laboratoire de Chimie Quantique, 15 rue P. et M. Curie, Paris 75, France

P. Rancurel
Department de Radioastronomie Spatiale, Observatoire de Meudon, 92 France
(Received 30 July 1971)

A method is proposed to calculate the effect of configuration interaction by a Rayleigh–Schrödinger perturbation expansion when starting from a multiconfigurational wavefunction. It is shown that a careless choice of \( H_0 \) may lead to absurd transition energies between two states, at the first orders of the perturbation, even when the perturbation converges for both states. A barycentric definition of \( H_0 \) is proposed, which ensures the cancellation of common diagrams in the calculated transition energies. A practical iterative procedure is defined which allows a progressive improvement of the unperturbed wavefunction \( \Psi_0 \), the CI matrix restricted to a subspace \( S \) of strongly interacting determinants is diagonalized. The desired eigenvector \( \Psi \) of this matrix is perturbed by the determinants which do not belong to \( S \). The most important determinants in \( \Psi \) are added to \( S \), etc. The energy thus obtained after the second-order correction is compared with the ordinary perturbation series where \( \Psi \) is a single determinant. For the ground state, this procedure includes, besides the whole second-order correction, the most important terms of the third and fourth orders. The question of orthogonality of excited states is discussed. This technique, hereafter called CIPSI, has been tested on the ground and several excited states of \( \text{H}_2 \), \( \text{Ne} \), and \( \text{MgO} \), showing both a rapid convergence of the calculated transition energy and the importance of correlation effects on transition energy.
CIPSI CONVERGENCE

1. Define a reference wave function:
   \[ |\Psi\rangle = \sum_{i \in D} c_i |i\rangle \]
   \[ E_{va} = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \]

2. Generate external Determinants
   All single and double excitations

3. Second order perturbative contribution of each determinant |\alpha \rangle
   \[ \Delta E = \frac{\langle \Psi | \hat{H} | \alpha \rangle \langle \alpha | \hat{H} | \Psi \rangle}{E_{va} - \langle \alpha | \hat{H} | \alpha \rangle} \]

4. Select |\alpha \rangle’s with largest \Delta E_\alpha and add them to Determinant space (D)

5. Diagonalize \hat{H} in D then update |\Psi \rangle and E_{var}

6. Iterate until reaching convergence.

- Convergence at 0.1 mHa with 1.2M determinants

- Selected CI (CIPSI) systematically reduces the energy
- Selecting all determinants leads rigorously to the FCI energy
TRUNCATING THE NUMBER OF DETERMINANTS

- $f(E + PT_2)$ converges within 0.1 mHa with ~2100 determinants

3. Second order perturbative contribution of each determinant $|\alpha\rangle$

$$\Delta E = \frac{\langle \psi | \hat{H} | \alpha \rangle \langle \alpha | \hat{H} | \psi \rangle}{E_{\alpha} - \langle \alpha | \hat{H} | \alpha \rangle}$$

4. Select $|\alpha\rangle$'s with largest $\Delta E_\alpha$ and add them to Determinant space (D)

5. Diagonalize $\hat{H}$ in D then update $|\Psi\rangle$ and $E_{\text{var}}$

6. Iterate until reaching convergence.

At anytime

$$E_{PT}^2 = \sum_\alpha \Delta E_\alpha$$

estimates the distance to the exact FCI solution

$f(E + PT_2)$ converges faster reducing significantly the number of determinant
DMC energy is systematically reduced with CIPSI wavefunctions. Number of determinants needs to be reduced to be useful in general. Only 232 VMC optimized determinant coefficients are sufficient to converge within 3mHa.
WATER MOLECULE (GROUND STATE)

![Graph showing energy versus basis set]

<table>
<thead>
<tr>
<th>Basis set</th>
<th>FCI size</th>
<th># dets used in DMC</th>
<th>$E_n^{\text{ref}}$</th>
<th>FCI, Almora-Diaz[27]</th>
<th>Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>cc-pCVDZ</td>
<td>$\sim 10^{10}$</td>
<td>172 256</td>
<td>-76.282136</td>
<td>-76.282865</td>
<td>0.0007</td>
</tr>
<tr>
<td>cc-pCVTZ</td>
<td>$\sim 2.10^{14}$</td>
<td>640 426</td>
<td>-76.388287</td>
<td>-76.390158</td>
<td>0.0018</td>
</tr>
<tr>
<td>cc-pCVQZ</td>
<td>$\sim 2.10^{17}$</td>
<td>666 927</td>
<td>-76.419324</td>
<td>-76.421148</td>
<td>0.0018</td>
</tr>
<tr>
<td>cc-pCV5Z</td>
<td>$\sim 7.10^{19}$</td>
<td>1 423 377</td>
<td>-76.428550</td>
<td>-76.431105</td>
<td>0.0025</td>
</tr>
</tbody>
</table>

Caffarel et al., J. Chem. Phys. 144, 151103 (2016); http://dx.doi.org/10.1063/1.4947093
MULTI-DETERMINANT RSDFT

- Multideterminants improve electron-electron cusp (affects nodal surface). 2 body-Jastrow improves electron-electron cusp without changing the nodal surface (affects variance).

\[ J^2 = \sum_i^e \sum_{j>i}^e u_{ab}(|r_i - r_j|) \]

J2 dn

High Energy determinants only to correct for the Cusp. 2-Body Jastrows in QMC already do some of that work.
\( \hat{V}_{ee} \) Can be decomposed into a long range and a short range component.

We can use an error function null on \( x=0 \) and equals 1 when \( x \to \infty \).

We can rewrite in function of a parameter \( \mu \) (inverse to the cutoff radius):

\[
\hat{w}_{ee}^{lr,\mu}(r) = \frac{\text{erf}(\mu r)}{r}, \quad \hat{w}_{ee}^{sr,\mu}(r) = 1 - \hat{w}_{ee}^{lr,\mu}(r) = \frac{1 - \text{erf}(\mu r)}{r}
\]

Within DFT formalism:

\[
E_0 = \langle \Psi^{lr,\mu} | \hat{T}_e + \hat{W}_{ee}^{lr,\mu} + \hat{V}_{ne} | \Psi^{lr,\mu} \rangle + E_{\text{hxc}}^{sr,\mu}[n_{\Psi^{lr,\mu}}]
\]

sCI: recovers static correlation

DFT (LDA or PBE)
APPLICATIONS TO WATER

Benali et al. in preparation
QUESTION?!
Lecture III

Quantum Monte Carlo Simulations Implemented in QMCPACK
QMCPACK

- Ken Esler (Formerly UIUC)
- Jeremy McMinis (Formerly UIUC)
- Raymond Clay (LLNL)
- Others (postdocs)

Funding and support:
- DOE BES: Predictive Theory and Modeling for Materials and Chemical Science program
- Intel PCC at ANL and SNL
- DOE ASCR – ECP: QMCPACK: A framework for predictive and systematically improvable quantum-mechanics based simulations of materials
- DOE BES – CMS: Center for Predictive Simulation of Functional Materials

GOAL: USE EFFICIENTLY THE NEW GENERATION OF HPC TO SOLVE LARGE PROBLEMS

- Monte Carlo Methods are inherently parallel. But we cannot rely anymore simply on the MC algorithm for scalability on HPC to achieve performance.

- HPCs are bigger and bigger. However, the number of nodes is shrinking while the number of cores per node is increasing very fast
  
  - 16 cores / node (IBM BGQ – 2012)
  - ~60 cores / node (Intel KNL – 2016)
  - ~XXX things / node (Aurora – 2021)

- HPCs come with a number of hardware features that need to be taken into account when developing your code (Vectorization, Intrinsics, prefetching, atomic, etc…)
Quad FPU (QPU)
DMA unit
List-based prefetcher
TM (Transactional Memory)
SE (Speculative Execution)
Wakeup-Unit
Scalable Atomic Operations

Instruction Extensions (QPX) to PowerISA
4-wide double precision FPU SIMD (BG/L,P are 2-wide) usable as:
- scalar FPU
- 4-wide FPU SIMD
- 2-wide complex arithmetic SIMD

Attached to AXU port of A2 core – A2 issues one instruction/cycle to AXU
8 concurrent floating point operations (FMA) + load + store
- 6 stage pipeline
Permute instructions to reorganize vector data supports a multitude of data alignments

4R/2W register file
32x32 bytes per thread
32B (256 bits) data path to/from L1 cache
ARCHITECTURE OF QMCPACK ON CLUSTERS OF SMP

QMCPACK utilizes OpenMP/CUDA to optimize memory usage and to take advantage of the growing number of cores per SMP node or stream multiprocessors per GPU.

- Walkers within a MPI task is distributed among the cores in CPU or the SMs in GPU.
- Big common data is shared by all the walkers like wave function coefficients.

```plaintext
1: for MC generation = 1 · · · M do
2:     for walker = 1 · · · N_w do
3:         let R = {r_1, . . . , r_N}
4:     for particle i = 1 · · · N do
5:         set r_i = r_i + δ
6:         let R’ = {r_1, . . . , r_i, . . . , r_N}
7:         ratio ρ = ψ_T(R’)/ψ_T(R)
8:            (One Body, Two Body, 3D B-Spline) derivatives ∇ψ_T, ∇ψ_T
9:            (One Body, Two Body, 3D B-Spline)
10:     if r → r’ is accepted then
11:         update state of a walker (Inverse Update)
12:     end if
13: end for {particle}
14: local energy E_T = F[ψ_T(R’)/ψ_T(R)]
15: (One Body, Two Body, 3D B-Spline)
16: reweight and branch walkers
17: end for {walker}
18: end for {MC generation}
```

Since 2012, we estimate a total speed up of ~37X (From algorithms and Machine performance)
ROUTE TO EXASCALE

**QMCPACK**  A framework for predictive and systematically improvable quantum-mechanics based simulations of materials


**Intel® Parallel Computing Centers**

**QMCPACK**  Threading and Vectorization of Quantum Monte Carlo Codes for Hyperscale Applications

**Anouar Benali**(1), Ye Luo(1)
Luke Shulenburger(3)
Jeongnim Kim(5), Amrita Mathuriya(5)

Extending Moore's Law Computing with Quantum Monte Carlo
Anouar Benali, Argonne National Laboratory
Aurora ESP 2017
Materials Science
What we learned (From Optimization on IBM BGQ):
- Straight forward application of QPX, SSE or Assembly brings 1.5~2x speedup
- Better algorithm + Structures of Array (SoA) doubles the performance due to better memory bandwidth
- Portable algorithm are possible but portable implementation is the challenge.

What we do (Intel PCC / Exascale Computing Program)
- Move from Double Precision to Mixed Precision
- Move from Array of Structure (AoS) to SoA to AoSoA (Tiling)
- Using nested threading at the walker level to reduce the time to solution (In connexion to previous Talk).
- Rewrite completely Distance table function
- Generate a Micro-QMC (Miniapp) to experiment with algorithms within the constraints of QMCPACK
Profiling of QMCPACK on BG Q

- Ti-S79, DMC (1896 e⁻)

### Experiment Aggregate Metrics

<table>
<thead>
<tr>
<th>Scope (Original Algo)</th>
<th>WALL (usec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment Aggregate Metrics</td>
<td>3.06e+09 100%</td>
</tr>
<tr>
<td>SpinWaitTaskSwitchBGQ</td>
<td>5.79e+08 18.9%</td>
</tr>
<tr>
<td>eval_multi_USBspline_3d_s_vgh</td>
<td>4.54e+08 14.8%</td>
</tr>
<tr>
<td>zmv5vfe</td>
<td>2.42e+08 7.9%</td>
</tr>
<tr>
<td>zgeruvfe</td>
<td>2.25e+08 7.3%</td>
</tr>
<tr>
<td>SymmetricDTD::moveOnSphere</td>
<td>1.50e+08 4.9%</td>
</tr>
<tr>
<td>eval_multi_USBspline_3d_s</td>
<td>1.47e+08 4.8%</td>
</tr>
<tr>
<td>sched_yield</td>
<td>1.01e+08 3.3%</td>
</tr>
<tr>
<td>PAMI_Context_trylock_advancev</td>
<td>8.70e+07 2.8%</td>
</tr>
<tr>
<td>SymmetricDTD::move</td>
<td>8.62e+07 2.8%</td>
</tr>
<tr>
<td>LPQHI Basis</td>
<td>7.50e+07 2.5%</td>
</tr>
<tr>
<td>floor</td>
<td>6.06e+07 2.0%</td>
</tr>
</tbody>
</table>

Spline: 19.6%
BLAS/LAPACK: 15.2%
Spline Algorithm

Phase 0: original Esler’s algorithm has 4 nested loops (x,y,z,spline)
Spline Algorithm

Phase 0: original Esler’s algorithm has 4 nested loops \((x,y,z,\text{spline})\)

Phase 1: the innermost loops spline and \(z\) are interchanged. Loop \(z\) is completely unrolled. \(a(x)\cdot b(y)\) is computed only once. Cross-platform
Spline Algorithm

Phase 0: original Esler’s algorithm has 4 nested loops \((x,y,z,\text{spline})\)

Phase 1: the innermost loops spline and \(z\) are interchanged. Loop \(z\) is completely unrolled. \(a(x)\cdot b(y)\) is computed only once. Cross-platform

Phase 2: The spline loop is unrolled 4 times and vectorization is applied. Fused multiply-add (IBM QPX) instructions are used.
**Spline Algorithm**

- **Phase 0**: original Esler’s algorithm has 4 nested loops \((x, y, z, \text{spline})\)

- **Phase 1**: the innermost loops \(\text{spline}\) and \(z\) are interchanged. Loop \(z\) is completely unrolled. \(a(x) \cdot b(y)\) is computed only once.

- **Phase 2**: The spline loop is unrolled 4 times and vectorization is applied. Fused multiply-add (IBM QPX) instructions are used.

- **Phase 2p**: Add the data pre-fetch built-in functions to speed up accessing the spline coefficients table.

---

**Phase 0**

```c
for x = x_0, \ldots, x_0+3 do
  for y = y_0, \ldots, y_0+3 do
    for z = z_0, \ldots, z_0+3 do
      for spline = 1, \ldots, N_{\text{spline}} do
        abc = a(x) \cdot b(y) \cdot c(z)
        val(spline) += \text{coef}(x, y, z, \text{spline}) \cdot abc
      end for
    end for
  end for
end for
```

**Phase 1**

```c
ab = a(x) \cdot b(y)
for spline = 1, \ldots, N_{\text{spline}} do
  val(spline) += ab \cdot [c(z_0) \cdot \text{coef}(x, y, z_0, \text{spline}) + c(z_0+1) \cdot \text{coef}(x, y, z_0+1, \text{spline}) + c(z_0+2) \cdot \text{coef}(x, y, z_0+2, \text{spline}) + c(z_0+3) \cdot \text{coef}(x, y, z_0+3, \text{spline})]
end for
```

**Phase 2**

```c
ab = a(x) \cdot b(y)
for spline = 1, 5, 9, \ldots, N_{\text{spline}} do
  val_{\text{vec}}(spline) += ab \cdot [c(z_0) \cdot \text{coef}_{\text{vec}}(x, y, z_0, \text{spline}) + c(z_0+1) \cdot \text{coef}_{\text{vec}}(x, y, z_0+1, \text{spline}) + c(z_0+2) \cdot \text{coef}_{\text{vec}}(x, y, z_0+2, \text{spline}) + c(z_0+3) \cdot \text{coef}_{\text{vec}}(x, y, z_0+3, \text{spline})]
end for
```

**Phase 2p**

```c
remainder
```
Profiling of the Spline Algorithm

- Ti – S79 - 948e x 2, Single precision
- VMC, 256x4t, 40 steps

<table>
<thead>
<tr>
<th>time(s)</th>
<th>3d_s</th>
<th>%</th>
<th>s_vgh</th>
<th>%</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original.noQPX</td>
<td>316</td>
<td>13.3%</td>
<td>176</td>
<td>7.4%</td>
<td>2370</td>
</tr>
<tr>
<td>New.noQPX</td>
<td>220</td>
<td>9.8%</td>
<td>137</td>
<td>6.1%</td>
<td>2240</td>
</tr>
<tr>
<td>New.QPX</td>
<td>79</td>
<td>3.9%</td>
<td>86</td>
<td>4.2%</td>
<td>2050</td>
</tr>
</tbody>
</table>

- Speedup ~4 folds for Value
- Speedup ~2 folds for Value-Gradient-Hessian
Profiling of QMCPACK on BG Q

- Ti-S79, DMC

<table>
<thead>
<tr>
<th>Scope (QPX)</th>
<th>WALL (usec)</th>
<th>Experiment Aggregate Metrics</th>
<th>2.15e+09</th>
<th>100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>SpinWaitTaskSwitchBGQ</td>
<td>3.85e+08 17.9%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>zmv5vfe</td>
<td>2.45e+08 11.4%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>zgeruvfe</td>
<td>2.43e+08 11.3%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>eval_multi_UBspline_3d_s_vgh</td>
<td>1.07e+08 5.0%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SymmetricDTD::moveOnSphere</td>
<td>7.33e+07 3.4%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAMI::Interface::Context::advance</td>
<td>7.23e+07 3.4%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAMI_Context_trylock_advancev</td>
<td>6.90e+07 3.2%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>__sched_yield</td>
<td>6.42e+07 3.0%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SymmetricDTD::move</td>
<td>4.98e+07 2.3%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TwoBodyJastrowOrbital::ratio</td>
<td>4.48e+07 2.1%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAMI::Device::Interface::BaseDevice::advance</td>
<td>4.19e+07 1.9%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scope (Original Algo)</th>
<th>WALL (usec)</th>
<th>Experiment Aggregate Metrics</th>
<th>3.06e+09</th>
<th>100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>SpinWaitTaskSwitchBGQ</td>
<td>5.79e+08 18.9%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>eval_multi_UBspline_3d_s_vgh</td>
<td>4.54e+08 14.8%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>zmv5vfe</td>
<td>2.42e+08 7.9%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>zgeruvfe</td>
<td>2.25e+08 7.3%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SymmetricDTD::moveOnSphere</td>
<td>1.50e+08 4.9%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>eval_multi_UBspline_3d_s</td>
<td>1.47e+08 4.8%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sched_yield</td>
<td>1.01e+08 3.3%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAMI_Context_trylock_advancev</td>
<td>8.70e+07 2.8%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SymmetricDTD::move</td>
<td>8.62e+07 2.8%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LPQHIBasis</td>
<td>7.50e+07 2.5%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>floor</td>
<td>6.06e+07 2.0%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Spline:** 19.6% to 6.6%  
**BLAS/LAPACK:** 15.2% to 22.7%
DOUBLE TO SINGLE PRECISION

- Gain performance not only on KNL but also on BG/Q.

- Small core counts: 20% faster on KNL, gained from computing.

- Full node: 55% faster, gained from both computing and memory BW.

- Always: about 70% faster on BG/Q, gained from memory BW.
Single particle orbitals

- Plane wave representations are inefficient
- A real space basis is much more efficient
  - Extensive testing shows that 3D b-splines are a good choice for efficiency
  - Only 64 basis elements are nonzero at each point (versus thousands or more for plane waves)

- \( B_{ijk}(x,y,z) = a_i(x)b_j(y)c_k(z) \)
- Efficient routines exist to convert from a plane wave basis to b-splines with an arbitrary real space spacing using the FFT

\[
\Psi_T = \exp(J)D^\dagger(\{\phi\})D^\dagger(\{\phi\}),
\]

\[
D = \begin{vmatrix}
\phi_1(r_1) & \cdots & \phi_1(r_N) \\
\vdots & \ddots & \vdots \\
\phi_N(r_1) & \cdots & \phi_N(r_N)
\end{vmatrix}
\]

\[
\phi_n(x, y, z) = \sum_{i' = i-1}^{i+2} \sum_{j' = j-1}^{j+2} \sum_{k' = k-1}^{k+2} b_{i',j',k'}^{\pm 3}(x) b_{i,j}^{\pm 3}(y) b_{k}^{\pm 3}(z) p_{i',j',k',n}.
\]
Array of Structure → Structure of Array

- Optimizing the memory access for the vector computing is essential.

- Changes
  - Gradient \((N,3)\) to \((3,N)\), Hessian \((N,3,3)\) to \((6,N)\), symmetric

```c
(a) class BsplineAoS {
  T P[Nx][Ny][Nz][N]; // read only and shared among threads.
  void VGH(T x, T y, T z, T* v, T* g, T* h) {
    //compute the lower-bound index i0, j0, k0
    //compute prefactors using (x-x0,y-y0,z-z0)

    for(int i=0; i<4; ++i)
      for(int j=0; j<4; ++j)
        for(int k=0; k<4; ++k) {
          const T* p=P[i+i0][j+j0][k+k0];
          #pragma omp simd
          for(int n=0; n<N; ++n) {
            v[n] += F[p[n]];  
g[3*n+n0] += Gx(p[n]); 
g[3*n+1+n0] += Gy(p[n]);
            h[9*n+n0] += Hxx(p[n]);
            h[9*n+1+n0] += Hxy(p[n]);
            ...
          }
        }
  }
};

(b) class BsplineSoA {
  T P[Nx][Ny][Nz][N]; // read only and shared among threads.
  void VGH(T x, T y, T z, T* v, T* g, T* h) {
    //compute the lower-bound index i0, j0, k0
    //compute prefactors using (x-x0,y-y0,z-z0)
    T *gx=g[x], *gy=g[y], *gz=g[z];
    T *hxz=h, *hxy=h+yN, *hxx=h+xN;
    T *hyy=h+yN, *hxx=h+xN, *hzz=h+zN;
    for(int i=0; i<4; ++i)
      for(int j=0; j<4; ++j)
        for(int k=0; k<4; ++k) {
          const T* p=P[i+i0][j+j0][k+k0];
          #pragma omp simd
          for(int n=0; n<N; ++n) {
            v[n] += F[p[n]];  
gx[n]  += Gx(p[n]);  
gy[n]  += Gy(p[n]);  
hxx[n] += Hxx(p[n]);
            hxy[n] += Hxy(p[n]);
            ...
          }
        }
  }
};
```

Fig. 4. \( VGH \) using the Gradients and Hessians (a) in AoS memory layout and (b) in SoA memory layout.
AoS → SoA

- Optimizing the memory access for the vector computing is essential.
- Read bandwidth increase from 60-98 GB/s to 238 GB/s.
- Small L1/L2 lead to a speed drop at large N because of writing back to memory.
AoSoA TRANSFORMATION (TILING)

- Improve cache performance and enable nested threading

- AoSoA widely used in stencil methods:
  - N orbitals will be divided into blocks of Nb size.
  - Data blocks are small enough to be kept within last level cache (LLC).

- Performance:
  - Nb needs to be scanned to get the peak performance
  - Nb choice is independent of problem size but depends only on platforms
AoSoA TRANSFORMATION (TILING)

- The performance is recovered for large N.
PERFORMANCE SUMMARY

(NiO-32) Rest DetUpdate SPO-vgl Bspline-v Bspline-vgh J2 DistTable

(NiO-64) Rest DetUpdate SPO-vgl Bspline-v Bspline-vgh J2 DistTable

Figure 1: Strong scaling of NiO-64 benchmark on Trinity at LANL (KNL) and Serrano at SNL (BDW) systems. The performance is normalized by a reference throughput using 64 BDW sockets. Slopes of the ideal-scaling lines are provided in parentheses.

(Embracing a new era of highly efficient and productive quantum Monte Carlo simulations) A. Mathuriya, Y. Luo, A. Benali, L. Shulenburger, R. Clay, J. Kim, accepted: Super Computing 17

(Optimization and parallelization of B-spline based orbital evaluations in QMC on multi/many-core shared memory processors), A. Mathuriya, Y. Luo, A. Benali, L. Shulenburger, J. Kim, International Parallel & Distributed Processing Symposium (Proceedings) 2017

<table>
<thead>
<tr>
<th></th>
<th>NiO-32</th>
<th>NiO-64</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>384</td>
<td>768</td>
</tr>
<tr>
<td>NiD</td>
<td>32</td>
<td>64</td>
</tr>
<tr>
<td>NiD/uint cell</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td># of uint cells</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>Ion types (Z)</td>
<td>Ni(18), O(6)</td>
<td></td>
</tr>
<tr>
<td># of unique SPOs</td>
<td>144</td>
<td>240</td>
</tr>
<tr>
<td>FFT grid</td>
<td>80x80x80</td>
<td>80x80x80</td>
</tr>
<tr>
<td>B-spline (GB)</td>
<td>1.3</td>
<td>2.1</td>
</tr>
</tbody>
</table>
QUESTION?!
Lecture V

Examples and Tutorial
Ti$_4$O$_7$ Magnéli phase (Low temperature ~ 120K)

$E_{\text{diff}} = \sim 4\text{meV/Ti}_4\text{O}_7$ Using LDA +U , LDA +ASIC, HSE06 and All electrons calculations

LDA +U favors the FM state as the ground state while LDA +ASIC and HSE06 favor the AF3 and AF1 over the FM states.

AF1: (+,0,-,0|+,0,-,0)
AF2: (+,0,+|-,0,-,0)
AF3: (+,0,-,0|-,0,+)
FM : (+,+,+,+|+,+,+,+)
QMC Increases the energy differences between phases in Ti$_4$O$_7$

<table>
<thead>
<tr>
<th></th>
<th>( E(\text{FM} - \text{AF3}) ) (eV)</th>
<th>( E(\text{AF1} - \text{AF3}) ) (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMC</td>
<td>(+0.1779 \pm 0.0134)</td>
<td>(+0.07041 \pm 0.01035)</td>
</tr>
<tr>
<td>LDA +U</td>
<td>(+0.607)</td>
<td>(+0.004)</td>
</tr>
<tr>
<td>LDA + ASIC</td>
<td>(+0.118)</td>
<td>(+0.004)</td>
</tr>
</tbody>
</table>

AF1: (+,0,−,0|+,0,−,0)
AF2: (+,0,+,0|−,0,−,0)
AF3: (+,0,−,0|−,0,+,0)
FM: (+,+,+,+,|+,+,+,+)

E: FM, AF1, AF3
O2 ADSORPTION ON GRAPHENE

- The Adsorption and Diffusion of O₂ on a Single Layer Graphene using Diffusion Monte Carlo

Calculated O₂ adsorption energy curve using DMC

vdW interaction plays important role in O₂ transition in A ↔ V
ELLIPISTICINE

- Ellipticine is a planar polycyclic aromatic molecule
- Bind to DNA by non-covalent pi-pi-stacking with the nucleic acid Watson-Crick base pairs
- Binding energy is directly correlated to biological activity of the molecule in cancer treatment
<table>
<thead>
<tr>
<th>Level of theory</th>
<th>( \Delta E_{\text{Bind}} ) (Kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DFT(^1)</td>
<td>+5.2</td>
</tr>
<tr>
<td>vDW-TS(^1)</td>
<td>-46.6</td>
</tr>
<tr>
<td>vDW-TB(^1)</td>
<td>-39.1</td>
</tr>
<tr>
<td>PBE-D3/QZVP(^2)</td>
<td>-35.68 (D2); -32.84 (D2+D3)</td>
</tr>
<tr>
<td>PBE-NL/QZVP(^2)</td>
<td>-39.11 (D2); -36.27 (D2+D3)</td>
</tr>
<tr>
<td>dDsC-PBE/QZ4P(^2)</td>
<td>-40.91 (D2); -38.07 (D2+D3)</td>
</tr>
<tr>
<td>vDW-MB(^1)</td>
<td>-34</td>
</tr>
<tr>
<td>DMC(^3)</td>
<td>-33.6 ± 0.9</td>
</tr>
</tbody>
</table>

Collective many-body van der Waals interactions in molecular systems

Robert A. DiStasio, Jr.*, O. Anatole von Lilienfeld*, and Alexandre Tkatchenko\(^*\)

*Department of Chemistry, Princeton University, Princeton, NJ 08544; *Argonne Leadership Computing Facility, Argonne National Laboratory, Argonne, IL 60439; and *Fritz-Haber Institute der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195 Berlin, Germany

Edited by Peter J. Kollman, The University of Texas at Austin, Austin, TX, and approved July 27, 2012 (received for review May 22, 2012)

Van der Waals (vdW) interactions are ubiquitous in molecules and condensed matter, and play a crucial role in determining the structure, stability, and function for a wide variety of systems. The accurate prediction of these interactions from first principles is a substantial challenge because they are inherently quantum mechanical.

Results and Discussion

To accurately compute the nonadditive many-body vdW energy, we begin by performing a self-consistent quantum mechanical calculation to generate the molecular electron density using semilocal density-functional theory (DFT) (23)—a method which accurately

METALLORGANIC FRAMEWORKS (MOF)

Ethylene hydrogenation

Investigated electronic energy barrier in MOFs using DFT and DMC

Discrepancy in the barrier is shown between DMC and DFT including hybrid functionals

C-H ACTIVATION


Reactant + CH₄ → activated complex (activated C-H bond)

Here: Cu-Ni and Cu-Zn complexes (two inequivalent positions for Cu-Ni and Cu-Zn)

Cu-Ni decorated Zr6 node and ligands


Figure 5: Energy differences in kcal/mol for Ni functionalized Zr₆ nodes with E(TS)-E(R) in red for the 2-Ni node, and blue for the 4-Ni node; the error bars on the DMC values are ± one standard deviation.
USING MULTIDETERMINANTS

Strategy: Increase the complexity of the trial wavefunction using a multideterminant expansion (MCSCF)

G1 test set

WATER MOLECULE (GROUND STATE)

\[
\begin{align*}
\text{Basis set} & & \text{FCI size} & & \# \text{ dets used in DMC} & & E_{\text{av}} & & \text{FCI, Almora-Diaz[27]} & & \text{Deviation} \\
\text{cc-pCVDZ} & & \sim 10^{10} & & 172 \ 256 & & -76.282136 & & -76.282865 & & 0.0007 \\
\text{cc-pCVTZ} & & \sim 2.10^{14} & & 640 \ 426 & & -76.388287 & & -76.390158 & & 0.0018 \\
\text{cc-pCVQZ} & & \sim 2.10^{17} & & 666 \ 927 & & -76.419324 & & -76.421148 & & 0.0018 \\
\text{cc-pCV5Z} & & \sim 7.10^{19} & & 1 \ 423 \ 377 & & -76.428550 & & -76.431105 & & 0.0025 \\
\end{align*}
\]
**H₂O MOLECULE (EXCITED STATE)**

**FIG. 2.** Error (in eV) compared to the TBE of Ref. [28] for the three lowest singlet and three lowest triplet exFCI (top) and exDMC (bottom) vertical excitation energies of the water molecule computed with various basis sets. The CBS values (in yellow) are also reported. The error bar corresponds to one standard error.
Tutorial
Quantum Monte Carlo calculation of the binding energy of the beryllium dimer


Michael J. Debije, Melody Keesler, Kevin E. Gasperich, and Kenneth D. Jordan

ABSTRACT

The accurate calculation of the binding energy of the beryllium dimer is a challenging theoretical problem. In this study, the binding energy of Be₂ is calculated using the diffusion Monte Carlo (DMC) method, using single determinant trial wave functions of orbitals obtained from density functional theory calculations overestimate the binding energy, while DMC calculations using Hartree-Fock or CAS(4,8), complete active space trial functions significantly underestimate the binding energy. In order to obtain an accurate value of the binding energy of Be₂ from DMC calculations, it is necessary to employ trial functions that
## TABLE I. Total energies of Be and Be$_2$ and the Be$_2$ dissociation energy computed with DMC using various trial functions.

<table>
<thead>
<tr>
<th>Trial function$^a$</th>
<th>Total energy (a.u.)</th>
<th>Be$^b$</th>
<th>Be$_2$</th>
<th>$D_c$ (cm$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HF/QZ-g</td>
<td>-14.657 30(4)</td>
<td>-29.317 89(6)</td>
<td>724(21)</td>
<td></td>
</tr>
<tr>
<td>LDA/QZ-g</td>
<td>-14.657 21(4)</td>
<td>-29.319 77(7)</td>
<td>1174(25)</td>
<td></td>
</tr>
<tr>
<td>PBE/QZ-g</td>
<td>-14.657 31(5)</td>
<td>-29.319 60(8)</td>
<td>1094(26)</td>
<td></td>
</tr>
<tr>
<td>BLYP/QZ-g</td>
<td>-14.657 25(4)</td>
<td>-29.319 56(8)</td>
<td>1113(26)</td>
<td></td>
</tr>
<tr>
<td>B3LYP/QZ-g</td>
<td>-14.657 27(3)</td>
<td>-29.319 46(8)</td>
<td>1079(23)</td>
<td></td>
</tr>
<tr>
<td>PBE0/QZ-g</td>
<td>-14.657 28(3)</td>
<td>-29.319 07(8)</td>
<td>992(21)</td>
<td></td>
</tr>
<tr>
<td>BH&amp;HLYP/QZ-g</td>
<td>-14.657 26(5)</td>
<td>-29.318 91(7)</td>
<td>966(26)</td>
<td></td>
</tr>
<tr>
<td>BD/QZ-g</td>
<td>-14.657 18(4)</td>
<td>-29.318 72(7)</td>
<td>955(24)</td>
<td></td>
</tr>
<tr>
<td>CAS(4,8)/QZ-fg$^c$</td>
<td>-14.667 23(1)</td>
<td>-29.337 07(3)</td>
<td>573(8)</td>
<td></td>
</tr>
<tr>
<td>CAS(4,16)/QZ-fg$^c$</td>
<td>-14.667 30(1)</td>
<td>-29.338 32(3)</td>
<td>819(8)</td>
<td></td>
</tr>
<tr>
<td>Ext. CAS(4,16)/QZ-fg</td>
<td>-14.667 30(1)</td>
<td>-29.338 41(2)</td>
<td>838(7)</td>
<td></td>
</tr>
<tr>
<td>CAS(4,16)/QZ-g$^c$</td>
<td>-14.667 27(2)</td>
<td>-29.338 38(3)</td>
<td>845(8)</td>
<td></td>
</tr>
<tr>
<td>Ext. CAS(4,16)/QZ-g</td>
<td>-14.667 27(2)</td>
<td>-29.338 45(2)</td>
<td>857(9)</td>
<td></td>
</tr>
<tr>
<td>CI/QZ-g$^c$</td>
<td>-14.667 25(1)</td>
<td>-29.338 48(2)</td>
<td>873(6)</td>
<td></td>
</tr>
<tr>
<td>Ext. CI/QZ-g</td>
<td>-14.667 25(1)</td>
<td>-29.338 64(2)</td>
<td>908(6)</td>
<td></td>
</tr>
<tr>
<td>Experimental$^d$</td>
<td>-14.667 356</td>
<td>-29.338 97</td>
<td>934.9(4)</td>
<td></td>
</tr>
</tbody>
</table>

$^a$QZ refers to the cc-pVQZ basis set. The “-g” and “-fg” indicate, respectively, that the $g$ functions and $f$ and $g$ functions were omitted from the basis sets. Ext. refers to CAS and CI results extrapolated to the full configuration space for the active orbital list as described in the text.

$^b$The DMC energies of the Be atom calculated using various single determinant trial functions should agree. The spread of the energies in the table is the result of statistical errors and the use of a finite (0.001 a.u.) time step.

$^c$0.001 threshold on CI coefficients for retained configurations.

$^d$The experimental $D_c$ value for Be$_2$ is from Ref. 5. The non-relativistic energy of the Be atom is from Ref. 60.

### Binding distance 2.453603 Å

![Graph showing DMC energies of twice the beryllium atom and the dimer for several single-determinant trial wave functions.](image)

FIG. 3. DMC energies of twice the beryllium atom and the dimer for several single-determinant trial wave functions.
Single Determinant Calculations

Using PYSCF

```python
#!/usr/bin/env python
from pyscf import gto
from pyscf import scf, dft, df
mol = gto.Mole()
mol.verbose = 5
mol.atom = '''
Be 0 0 0
Be 0 0 2.453603
'''
mol.unit = 'A'
mol.basis = 'cc-pvtz'
mol.build()

mf = dft.ROKS(mol)
mf.xc = 'B3LYP'

e_scf = mf.kernel()
title = "Be2"
kpts = []
from PyscfToQmcpack import savetoqmcpack
savetoqmcpack(mol, mf, title=title, kpts=kpts)
```
Single Determinant Calculations

Using PYSCF

```python
#!/usr/bin/env python
from pyscf import gto
from pyscf import scf, dft, df
mol = gto.Mole()
mol.verbose = 5
mol.atom = ""'
Be 0 0 0
Be 0 0 2.453603
''
mol.unit='A'
mol.basis = 'cc-pvtz'
mol.build()

mf = dft.ROKS(mol)
mf.xc = 'B3LYP'

e_scf=mf.kernel()
title="Be2"
kpts=[]
from PyscfToQmcpack import savetoqmcpack
savetoqmcpack(mol,mf,title=title,kpts=kpts)
```

converged SCF energy = -29.3240393868394 Ha

Energy_ref = -29.33897 Ha
Converting Input for QMCPACK

convert4qmc -pyscf Be2.h5 --addCusp --production

Generated files:

Be2.qmc.in.xml
Be2.wfj.xml
Be2.structure.xml
QMC run

Jastrow Optimization blocks

VMC Block

DMC Block

To get around the bias of autocorrelation, we group the MC steps into blocks, take the average of the data in the steps of each block, and then finally average the averages in all the blocks.

qmcpack Be2.qmc.in.xml | tee Be2.qmc.out
### Data analysis

**Nexus!!**

```bash
export PYTHONPATH=$QMCPACKPATH/nexus/bin:$PYTHONPATH
```

```
qmca -q ev *.scalar.dat
```

<table>
<thead>
<tr>
<th>Be2 series</th>
<th>LocalEnergy</th>
<th>Variance</th>
<th>ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-29.141364</td>
<td>0.014897</td>
<td>0.879478</td>
</tr>
<tr>
<td>1</td>
<td>-29.178192</td>
<td>0.005026</td>
<td>0.879577</td>
</tr>
<tr>
<td>2</td>
<td>-29.260836</td>
<td>0.001345</td>
<td>0.152701</td>
</tr>
<tr>
<td>3</td>
<td>-29.274219</td>
<td>0.003540</td>
<td>0.180419</td>
</tr>
<tr>
<td>4</td>
<td>-29.273993</td>
<td>0.003910</td>
<td>0.203364</td>
</tr>
<tr>
<td>5</td>
<td>-29.272941</td>
<td>0.000718</td>
<td>0.212846</td>
</tr>
<tr>
<td>6</td>
<td>-29.270059</td>
<td>0.000964</td>
<td>0.225165</td>
</tr>
<tr>
<td>7</td>
<td>-29.276406</td>
<td>0.001558</td>
<td>0.222388</td>
</tr>
<tr>
<td>8</td>
<td>-29.278666</td>
<td>0.001618</td>
<td>0.224764</td>
</tr>
<tr>
<td>9</td>
<td>-29.282798</td>
<td>0.001405</td>
<td>0.231414</td>
</tr>
<tr>
<td>10</td>
<td>-29.281780</td>
<td>0.001759</td>
<td>0.227447</td>
</tr>
<tr>
<td>11</td>
<td>-29.281114</td>
<td>0.001504</td>
<td>0.231435</td>
</tr>
<tr>
<td>12</td>
<td>-29.282288</td>
<td>0.001761</td>
<td>0.227753</td>
</tr>
<tr>
<td>13</td>
<td>-29.277376</td>
<td>0.003869</td>
<td>0.230026</td>
</tr>
<tr>
<td>14</td>
<td>-29.281418</td>
<td>0.001792</td>
<td>0.225809</td>
</tr>
<tr>
<td>15</td>
<td>-29.282496</td>
<td>0.000521</td>
<td>0.234214</td>
</tr>
<tr>
<td>16</td>
<td>-29.319683</td>
<td>0.000462</td>
<td>0.226178</td>
</tr>
</tbody>
</table>
Isolated Atom

bash-4.1$ qmca -q ev *.scalar.dat

<table>
<thead>
<tr>
<th>LocalEnergy</th>
<th>Variance</th>
<th>ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Be series 0</td>
<td>-14.589331</td>
<td>+/- 0.002335</td>
</tr>
<tr>
<td>Be series 1</td>
<td>-14.589560</td>
<td>+/- 0.001405</td>
</tr>
<tr>
<td>Be series 2</td>
<td>-14.639495</td>
<td>+/- 0.001203</td>
</tr>
<tr>
<td>Be series 3</td>
<td>-14.643407</td>
<td>+/- 0.001341</td>
</tr>
<tr>
<td>Be series 4</td>
<td>-14.635569</td>
<td>+/- 0.001453</td>
</tr>
<tr>
<td>Be series 5</td>
<td>-14.642393</td>
<td>+/- 0.001617</td>
</tr>
<tr>
<td>Be series 6</td>
<td>-14.640723</td>
<td>+/- 0.000762</td>
</tr>
<tr>
<td>Be series 7</td>
<td>-14.642467</td>
<td>+/- 0.001487</td>
</tr>
<tr>
<td>Be series 8</td>
<td>-14.640439</td>
<td>+/- 0.001054</td>
</tr>
<tr>
<td>Be series 9</td>
<td>-14.645242</td>
<td>+/- 0.002944</td>
</tr>
<tr>
<td>Be series 10</td>
<td>-14.643115</td>
<td>+/- 0.000930</td>
</tr>
<tr>
<td>Be series 11</td>
<td>-14.643895</td>
<td>+/- 0.000692</td>
</tr>
<tr>
<td>Be series 12</td>
<td>-14.640975</td>
<td>+/- 0.002026</td>
</tr>
<tr>
<td>Be series 13</td>
<td>-14.645640</td>
<td>+/- 0.002781</td>
</tr>
<tr>
<td>Be series 14</td>
<td>-14.641956</td>
<td>+/- 0.001089</td>
</tr>
<tr>
<td>Be series 15</td>
<td>-14.642829</td>
<td>+/- 0.000355</td>
</tr>
<tr>
<td><strong>Be series 16</strong></td>
<td><strong>-14.657254</strong></td>
<td><strong>+/- 0.000149</strong></td>
</tr>
</tbody>
</table>

converged SCF energy = -14.6586947229 Ha

Energy_ref = -14.667356 Ha
Atomization energy:

\[ E_A = E[\text{Be}_2] - 2E[\text{Be}] \]

\[ E_A(\text{B3LYP}) = -29.3240393868394 - 2(-14.6586947229835) \text{ Ha} = -0.0066499408724 \text{ Ha} = -1459.49 \text{ cm}^{-1} \]

\[ E_A(\text{DMC-B3LYP}) = -29.319744 - 2(-14.657254) + \sqrt{0.000445^2 + 0.000149^2} \text{ Ha} = -0.005236 + 0.00047 \text{ Ha} = -1149.1691 \pm 103.15 \text{ cm}^{-1} \]

1 \text{ Ha} = 219474.63 \text{ cm}^{-1}

Energy\_exp (\text{Be}) = -14.667356 \text{ Ha}

Energy\_exp (\text{Be}_2) = -29.33897 \text{ Ha}

Energy\_exp (E_A) = 934.9 \text{ cm}^{-1}
CIPSI

Input File (xyz file):

<table>
<thead>
<tr>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Be 0 0 0</td>
</tr>
<tr>
<td>Be 0 0 2.453603</td>
</tr>
</tbody>
</table>

> qp_create_ezfio_from_xyz Be2.xyz -b cc-pvqz
CIPSI

Running Hartree Fock

> qp_run SCF Be2.ezfio | tee Be2.scf

Energy_ref = -29.33897Ha
CIPSI

Running Hartree Fock

> qp_run SCF Be2.ezfio | tee Be2.scf

Energy: -29.13403775969117 Ha

Energy_ref = -29.33897 Ha
CIPSI

Running Hartree Fock

> qp_run SCF Be2.ezfio | tee Be2.scf

Energy: -29.13403775969117 Ha

Running sFCI

> qp_run fci_zmq Be2.ezfio | tee Be2-fci

Energy_ref = -29.33897 Ha
Running Hartree Fock

> qp_run SCF Be2.ezfio | tee Be2.scf

Energy: -29.13403775969117 Ha

Running sFCI

> qp_run fci_zmq Be2.ezfio | tee Be2-fci

Energy: -29.3349989877026 Ha (E+PT2 ~472k determinants)

Energy_ref = -29.33897Ha
Energy_{ref} = -29.3387\text{Ha}
CIPSI

Running Hartree Fock

> qp_run SCF Be.ezfio | tee Be.scf

Energy: -14.57296815347836 Ha

Running sFCI

> qp_run fci_zmq Be.ezfio | tee Be-fci

Energy: -14.6655503627978 Ha (E=E+PT2 with 21953 determinants)
CIPSI

Energy_ref = -14.667356Ha
CIPSI – Analysis (Be2)

mo_num 168
det_num 944670

-0.890735126548
111100000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000

0.274742200817
111010000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000

0.0928256545581
111001000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000

0.0928224237091
111000100000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000

0.0885417625249
110101000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000

0.0885412483091
110100100000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000

Argonne National Laboratory
Atomization energy:

\[ E_A = E[Be_2] - 2E[Be] \]

\[ E_A(B3LYP) = -29.3240393868394 - 2(-14.6586947229835) \text{ Ha} = -0.0066499408724 \text{ Ha} = -1459.49 \text{ cm}^{-1} \]

\[ E_A(DMC-B3LYP) = -29.319744 - 2(-14.657254) +/\sqrt{(0.000445^2 + 0.000149^2)} \text{ Ha} = -0.005236 +/\ 0.00047 \text{ Ha} = -1149.1691 +/\ 103.15 \text{ cm}^{-1} \]

\[ E_A(CIPSI) = -29.3349989877026 - 2(-14.6655503627978) \text{ Ha} = -0.003898262107 \text{ Ha} = -855.564 \text{ cm}^{-1} \]

\[ E_A(DMC-CIPSI) = -29.338919 - 2(-14.667331) +/\sqrt{(0.000295^2 + 0.000070^2)} \text{ Ha} = -0.004257 +/\ 0.00030 \text{ Ha} = -934.30 +/\ 65.03 \text{ cm}^{-1} \]

\[ \text{Energy}_{exp}(Be) = -14.667356 \text{ Ha} \]
\[ \text{Energy}_{exp}(Be_2) = -29.33897 \text{ Ha} \]
\[ \text{Energy}_{exp}(E_A) = 0.00425972 \text{ Ha} \]
\[ 934.9 \text{ cm}^{-1} \]
### TABLE I. Total energies of Be and Be$_2$ and the Be$_2$ dissociation energy computed with DMC using various trial functions.

<table>
<thead>
<tr>
<th>Trial function$^a$</th>
<th>Be$^b$</th>
<th>Be$_2$</th>
<th>$D_e$ (cm$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HF/QZ-g</td>
<td>-14.657 30(4)</td>
<td>-29.317 89(6)</td>
<td>724(21)</td>
</tr>
<tr>
<td>LDA/QZ-g</td>
<td>-14.657 21(4)</td>
<td>-29.319 77(7)</td>
<td>1174(25)</td>
</tr>
<tr>
<td>PBE/QZ-g</td>
<td>-14.657 31(5)</td>
<td>-29.319 60(8)</td>
<td>1094(26)</td>
</tr>
<tr>
<td>BLYP/QZ-g</td>
<td>-14.657 25(4)</td>
<td>-29.319 56(8)</td>
<td>1113(26)</td>
</tr>
<tr>
<td>B3LYP/QZ-g</td>
<td>-14.657 27(3)</td>
<td>-29.319 46(8)</td>
<td>1079(23)</td>
</tr>
<tr>
<td>PBE0/QZ-g</td>
<td>-14.657 28(3)</td>
<td>-29.319 07(8)</td>
<td>992(21)</td>
</tr>
<tr>
<td>BH&amp;HLYP/QZ-g</td>
<td>-14.657 26(5)</td>
<td>-29.318 91(7)</td>
<td>966(26)</td>
</tr>
<tr>
<td>BD/QZ-g</td>
<td>-14.657 18(4)</td>
<td>-29.318 72(7)</td>
<td>955(24)</td>
</tr>
<tr>
<td>CAS(4,8)/QZ-f$^c$</td>
<td>-14.667 23(1)</td>
<td>-29.337 07(3)</td>
<td>573(8)</td>
</tr>
<tr>
<td>CAS(4,16)/QZ-f$^e$</td>
<td>-14.667 30(1)</td>
<td>-29.338 52(3)</td>
<td>819(8)</td>
</tr>
<tr>
<td>Ext. CAS(4,16)/QZ-f$^g$</td>
<td>-14.667 30(1)</td>
<td>-29.338 41(2)</td>
<td>838(7)</td>
</tr>
<tr>
<td>CAS(4,16)/QZ-g$^c$</td>
<td>-14.667 27(2)</td>
<td>-29.338 38(3)</td>
<td>845(8)</td>
</tr>
<tr>
<td>Ext. CAS(4,16)/QZ-g$^c$</td>
<td>-14.667 27(2)</td>
<td>-29.338 45(2)</td>
<td>857(9)</td>
</tr>
<tr>
<td>CI/QZ-g$^c$</td>
<td>-14.667 25(1)</td>
<td>-29.338 48(2)</td>
<td>873(6)</td>
</tr>
<tr>
<td>Ext. CI/QZ-g$^c$</td>
<td>-14.667 25(1)</td>
<td>-29.338 64(2)</td>
<td>908(6)</td>
</tr>
<tr>
<td>Experimental$^d$</td>
<td>-14.667 356</td>
<td>-29.338 97</td>
<td>934.9(4)</td>
</tr>
</tbody>
</table>

$^a$QZ refers to the cc-pVQZ basis set. The “-g” and “-f$^g$” indicate, respectively, that the $g$ functions and $f$ and $g$ functions were omitted from the basis sets. Ext. refers to CAS and CI results extrapolated to the full configuration space for the active orbital list as described in the text.

$^b$The DMC energies of the Be atom calculated using various single determinant trial functions should agree. The spread of the energies in the table is the result of statistical errors and the use of a finite (0.001 a.u.) time step.

$^c$0.001 threshold on CI coefficients for retained configurations.

$^d$The experimental $D_e$ value for Be$_2$ is from Ref. 5. The non-relativistic energy of the Be atom is from Ref. 60.
QUESTION?!
Acknowledgments

- AB was supported through the Center for Predictive Simulation of Functional Materials, DOE BES Computational Materials Sciences program

- An award of computer time was provided by the Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program. This research used resources of the Argonne Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC02-06CH11357.