

Quantum Monte Carlo

Jong Ho Choi

University of Chicago

- > `Digits := 15 :`
- > `with(QuantumChemistry) :`

Abstract

Quantum Monte Carlo (QMC) method is a numerical method to calculate energy of quantum system. Monte Carlo method is commonly used method to calculate an average of a variable in large dimensional space (e.g. a physical variable in a classical phase space). Metropolis-Hasting algorithm make sampling of a variable to follow specific probability distribution possible. In quantum mechanics, Monte Carlo process can be used, but the variable (energy) and the probability distribution depends on wavefunctions. Variational Monte Carlo (VMC) method is simpler method that approximated trial wavefunctions are used for the Monte Carlo sampling. This method has a systematic error due to the wavefunctions. The more accurate method is Diffusion Monte Carlo (DMC) method where we sample energies from the exact ground state wave functions. QMC method can be used for complex system like transition metals. As an example, the transition energy of Cu atom and Cu⁺ ion result of Caffarel et. al. is presented and compared with other method: Hartree-Fock, Coupled-cluster, and Density functional theory. Application uses the Maple Quantum Chemistry Toolbox.

I. Introduction

Monte Carlo method is a numeric method is effective way to calculate the variables in a space of large dimensions. The method is based on generating a large number of random samples, and then using statistical analysis to estimate the behavior or outcome of interest. The Monte Carlo method is particularly useful for systems that are too complex to model analytically or have too many variables to be solved directly.

In statistical mechanics, the ensemble average of a variable $\langle A \rangle$ is calculated from integrating the product of A and the probability density $P(\mathbf{r}, \mathbf{p})$ in $6N$ -dimensional phase space.

$$\langle A \rangle = \iint A(\mathbf{r}, \mathbf{p}) P(\mathbf{r}, \mathbf{p}) d\mathbf{r}^N d\mathbf{p}^N$$

This high-dimensional integral is impossible to do, so numerical methods are often used. Monte Carlo method is one way. If we can sample A_i to have probability distribution $P(\mathbf{r}, \mathbf{p})$, we can estimate $\langle A \rangle$.

$$\langle A \rangle = \frac{1}{M} \sum_{i=1}^M A_i$$

So how can we sample the variable to follow a particular distribution?

Metropolis-Hastings algorithm makes this possible by focusing on the relative difference in probability, not the absolute probability density.

Assume that i and f are the points in the $6N$ -dimensional phase space in NVT ensemble. Although the

partition function ($Q = \int \int \exp[-E(\mathbf{r}, \mathbf{p})] d\mathbf{r}^N d\mathbf{p}^N$) is needed to calculate the exact probability density, but the relative probability ($P_i \sim \exp[-\beta E_i], P_f \sim \exp[-\beta E_f]$) is easy to know. In Metropolis–Hastings algorithm, *accept/deny process* for a point moving randomly in the phase space makes sampling points with the relative probability possible. Let i is the present state, and the f is the possible following state. The acceptance probability $P_{acc}(f, i)$ is defined as

$$P_{acc}(f, i) = \min \left[1, \frac{\exp[-\beta E_f]}{\exp[-\beta E_i]} \right].$$

If accepted, state f becomes the new state, and if denied, state i is sampled once again. If this process is repeated long enough, each point in space will be sampled by a relative probability.

Same sampling process can be applied for quantum systems. The energy can be written as

$$E = \frac{\int \Phi^* H \Phi d\mathbf{r}}{\int \Phi^* \Phi d\mathbf{r}} = \frac{\int |\Phi(\mathbf{r})|^2 E_{loc}(\mathbf{r}) d\mathbf{r}}{\int |\Phi(\mathbf{r})|^2 d\mathbf{r}}; \quad E_{loc}(\mathbf{r}) = \frac{H\Phi(\mathbf{r})}{\Phi(\mathbf{r})}$$

$$E = \int E_{loc}(\mathbf{r}) P(\mathbf{r}) d\mathbf{r}; \quad P(\mathbf{r}) = \frac{|\Phi(\mathbf{r})|^2}{\int |\Phi(\mathbf{r})|^2 d\mathbf{r}}$$

As the energy is multi-dimensional integral of a variable $E_{loc}(\mathbf{r})$ multiplied by the probability distribution $P(\mathbf{r})$, the energy can be calculated by the Monte Carlo average by using the Metropolis algorithm.

$$E = \frac{1}{M} \sum_{i=1}^M E_{loc}(\mathbf{r}_i)$$

To do this, one need to sample the local energy with probability distribution $P(\mathbf{r})$. We can simply use an approximated wave functions (variational quantum Monte Carlo), but also we can sample from the exact ground state wave function by Diffusion Quantum Monte Carlo method.

II. Theory

There are two most important Quantum Monte Carlo methods: Variational quantum Monte Carlo (VMC) and Diffusion quantum Monte Carlo (DMC). The VMC is more efficient and simpler than the DMC, but also less accurate.

Variational Quantum Monte Carlo

As told in previous section, the VMC energy can be written as

$$E_{VMC} = \int E_{loc}(\mathbf{r}) P(\mathbf{r}) d\mathbf{r} = \frac{1}{M} \sum_{i=1}^M E_{loc}(\mathbf{r}_i)$$

$$E_{loc}(\mathbf{r}) = \frac{H\Phi_T(\mathbf{r})}{\Phi_T(\mathbf{r})} \quad \text{and} \quad P(\mathbf{r}) = \frac{|\Phi_T(\mathbf{r})|^2}{\int |\Phi_T(\mathbf{r})|^2 d\mathbf{r}}$$

Here, the trial function $\Phi_T(\mathbf{r})$ is the approximated wavefunction. The most commonly used function is Jastrow-Slater form, a product of one or several Slater determinants ψ_n^{SD} and a Jastrow correlation

function e^U .

$$\Phi_T(\mathbf{r}) = e^U \cdot \sum_n (c_n \psi_n^{SD})$$

The Slater determinants are usually orbital calculated by Hartree-Fock or DFT, and the U of the Jastrow function depend on the electron-nucleus and electron-electron distance.

The VMC method has two type of errors: a statistical uncertainty due to the finite number of sampling ($V[E_{VMC}] \sim \frac{1}{M}$), and a systematic error due to approximated wavefunction.

Diffusion Quantum Monte Carlo

The DMC method start with the time-dependent Schrödinger equation with imaginary time $\tau = it$.

$$\frac{\partial}{\partial \tau} \psi(\tau) = -H \psi(\tau)$$

Let the $\psi(\tau)$ is the solution of this equation and $\{\varphi_i\}$ are eigenfunctions of the Hamiltonian.

$$H|\varphi_i\rangle = E_i|\varphi_i\rangle$$

$$|\psi(\tau)\rangle = \sum_i \left(\langle \varphi_i | \psi(0) \rangle e^{-E_i \tau} |\varphi_i\rangle \right)$$

For the long time limit, the higher energy of the state, the faster the contribution to the wavefunction decays so that $\psi(\tau)$ converges to the exact ground state φ_0 .

We introduced Green's function $G(\mathbf{R}_f, \mathbf{R}_i; t) = \langle \mathbf{R}_f | \exp[-(\hat{H} - E_{ref})t] | \mathbf{R}_i \rangle$. Where E_{ref} is the reference ground state energy. $\psi(t)$ can be expressed as

$$\psi(\mathbf{R}_f; t) = \int G(\mathbf{R}_f, \mathbf{R}_i; t) \psi_G(\mathbf{R}_i) d\mathbf{R}_i$$

$\psi_G(R)$ is guide wavefunction which approximates the ground state wavefunction. This function can be the same Jastrow-Slater function as VMC, but the role of the function is not the same.

Let $f(R, t) = \psi(R; t)\psi_G(R)$, and the equation can be reorganized as

$$f(\mathbf{R}_f; t) = \int \tilde{G}(\mathbf{R}_f, \mathbf{R}_i; t) \psi_G(\mathbf{R}_i)^2 d\mathbf{R}_i$$

$$\tilde{G}(\mathbf{R}_f, \mathbf{R}_i; t) = \frac{\psi_G(\mathbf{R}_f) G(\mathbf{R}_f, \mathbf{R}_i; t)}{\psi_G(\mathbf{R}_i)}$$

As addressed before, in long enough time limits, $\psi(R; t)$ converges to the ground state wavefunction, so

$f_\infty(R) = \lim_{t \rightarrow \infty} f(R; t) \sim \varphi_0(R)\psi_G(R_i)$. For short enough timestep $\Delta t = \frac{t}{M}$, $f_\infty(R)$ can be expressed as

$$f_\infty(R) = \lim_{M \rightarrow \infty} \int dR_1 dR_2 \dots dR_M \tilde{G}(R, R_M; \Delta t) \tilde{G}(R_M, R_{M-1}; \Delta t) \dots \tilde{G}(R_2, R_1; \Delta t) \psi_G(R_1)^2$$

From the definition of $G(\mathbf{R}_f, \mathbf{R}_i; t) = \langle \mathbf{R}_f | \exp(-(\hat{H} - E_{ref})t) | \mathbf{R}_i \rangle$ with ignoring higher order term in Δt , the Green's function is proportional to following expression.

$$G(R_f, R_i; \Delta t) \approx \frac{1}{(2\pi\Delta t)^{3N/2}} \exp\left[-\frac{(R_f - R_i)^2}{2\Delta t}\right] \exp\left[-\left(\frac{V(R_f) + V(R_i)}{2} - E_{ref}\right)\Delta t\right]$$

$$\tilde{G}(R_f, R_i; \Delta t) \approx \frac{1}{(2\pi\Delta t)^{3N/2}} \exp\left[-\frac{(R_f - R_i - \mathbf{v}(R_i)\Delta t)^2}{2\Delta t}\right] \exp\left[-\left(\frac{E_{loc}(R_f) + E_{loc}(R_i)}{2} - E_{ref}\right)\Delta t\right]$$

where $\mathbf{v}(R_i) = \frac{\nabla\psi_G(R_i)}{\psi_G(R_i)}$, $E_{loc}(R_i) = \frac{H\psi_G(R_i)}{\psi_G(R_i)}$

Generally, \tilde{G} is not a normalized probability density, so a stochastic process from this equation is not trivial. Nevertheless, we can write \tilde{G} as a product of a stochastic part P and a weight part W.

$$\tilde{G}(R_f, R_i; \Delta t) = P(R_f, R_i) W(R_f, R_i)$$

$$P(R_f, R_i) = \frac{1}{(2\pi\Delta t)^{3N/2}} \exp\left[-\frac{(R_f - R_i - \mathbf{v}(R_i)\Delta t)^2}{2\Delta t}\right]$$

$$W(R_f, R_i) = \exp\left[-\left(\frac{E_{loc}(R_f) + E_{loc}(R_i)}{2} - E_{ref}\right)\Delta t\right]$$

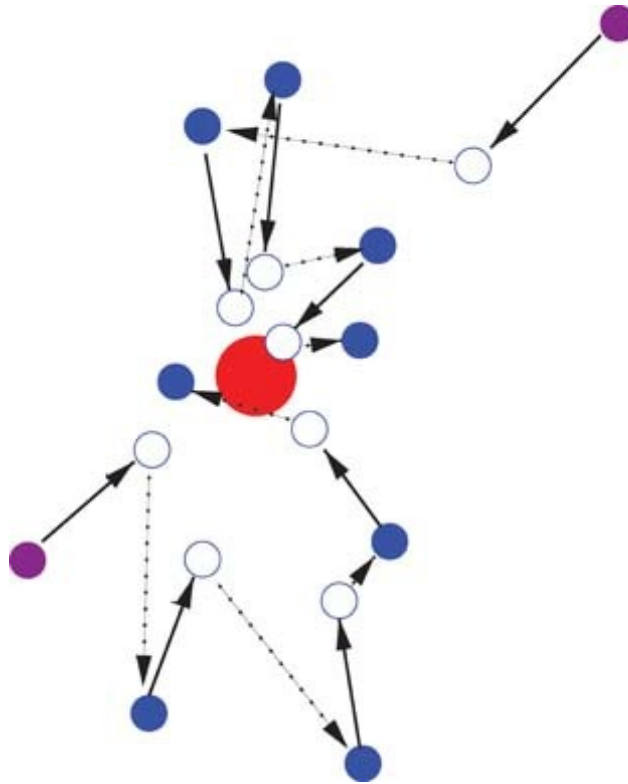


Figure 1

A random walk process visualized for a Helium atom. The walker consists of the two electron positions. This image is from Ref [2]

We can estimate the ground state energy from weighted random walk with metropolis algorithm, w . First, we start from position R_1 with weight $w_1 = 1$. Sample the second position R_2 with acceptance probability $P_{acc}(R_2, R_1)$. The acceptance probability is

$$P_{acc}(R_f, R_i) = \min \left[1, \frac{P(R_i, R_f) \Psi_G(R_f)^2}{P(R_f, R_i) \Psi_G(R_i)^2} \right]$$

Give the second weight ($w_2 = W(R_2, R_1)w_1$). Repeat this steps and it will reach to the stationary state.

To improve sampling quality, we use N_w independent walkers. From the M steps after the equilibration walk, the ground state energy E_0 can be obtained by followings.

$$E_0 = \frac{\sum_{m=1}^M \sum_{n=1}^{N_w} w_{m,n} E_{loc}(R_{m,n})}{\sum_{m=1}^M \sum_{n=1}^{N_w} w_{m,n}}$$

As independent walkers can be used for the process, DMC can be optimized easily with massively multi-processor computing. The Monte Carlo steps of each walker can be run independently and the trajectories are not needed to be shared.

Note that the probability density which this process is sampling is

$$f_\infty(R) = \frac{\Psi_G(R) \varphi_0(R)}{\int \Psi_G(R) \varphi_0(R) dR}$$

Fermion Sign Problem and Fixed-Node approximation

As shown above the DMC sample the energy with probability that is proportional to $\Psi_G(R) \varphi_0(R)$. This function works only when the sign of $\Psi_G(R)$ and $\varphi_0(R)$ are same. Since electron is fermion, the electronic wavefunction can be both positive and negative, and the sign of the function changed at the nodal surface. Therefore, if $\Psi_G(R)$ and $\varphi_0(R)$ have different node position, distribution $f_\infty(R)$ cannot be treated as probability. To solve this problem, Fixed-node DMC (FN-DMC) method is presented. This methods assume that the exact wavefunction ($\varphi_0(R)$) have the same node with the guide function ($\Psi_G(R) \varphi_0$). The error of this assumption depends on how accurate the position of nodes in the guide function are. So, it is important to use a more accurate guide function to reduce this error.

III. Applications

Transition Metal Compound

The calculation of electronic structure is often difficult when transition metals are involved. [2]

Caffarel et al. presented the effect of electron correlation in the copper atom using all-electron fixed-node diffusion Monte Carlo method. [6]

A "walker" is defined as the set of position of the N electrons, $\mathbf{R} = (r_1, r_2, \dots, r_N)$. About 100 walkers were used for the FN-DMC. Each walker is propagated as a Langevin-type move.

$$\mathbf{R}_{new} = \mathbf{R}_{old} + \mathbf{v}(\mathbf{R}_{old})\Delta t + \eta\sqrt{\Delta t}$$

Where the drift vector $\mathbf{v}(\mathbf{R}_{old}) = \frac{\nabla \Psi_G(\mathbf{R}_i)}{\Psi_G(\mathbf{R}_i)}$ and η is a 3N-dimensional Gaussian random vector. To sample the space effectively, Metropolis algorithm is used and the acceptance of each move is defined as

$$P_{acc} = \min \left[1, \frac{P(\mathbf{R}_{old} \rightarrow \mathbf{R}_{new}) \Psi_G(\mathbf{R}_{new})^2}{P(\mathbf{R}_{new} \rightarrow \mathbf{R}_{old}) \Psi_G(\mathbf{R}_{old})^2} \right].$$

$$\text{Where } P(\mathbf{R}', \mathbf{R}) \sim \exp \left[-\frac{(\mathbf{R}' - \mathbf{R} - \mathbf{v}(\mathbf{R})\Delta t)^2}{2\Delta t} \right]$$

The weight is given by

$$W(\mathbf{R}', \mathbf{R}) = \exp \left[-\left(\frac{E_{loc}(\mathbf{R}') + E_{loc}(\mathbf{R})}{2} - E_{ref} \right) \Delta t \right]$$

$$\text{Where } E_{loc} = \frac{H\Psi_G(\mathbf{R}_i)}{\Psi_G(\mathbf{R}_i)}$$

In the long time limit, the probability distribution of walkers are proportional to $\Psi_G(\mathbf{R})\varphi_0(\mathbf{R})$. A Choice of guide wavefunction $\Psi_G(\mathbf{R})$ is important in FN-DMC. The standard form is

$$\Psi_G(\mathbf{R}) = \exp \left[\sum_{i,j} U(r_i, r_j, r_{ij}) \right] D^+(\mathbf{R}) D^-(\mathbf{R})$$

Where $D^\pm(\mathbf{R})$ are SCF determinant from one-particle orbitals and $\exp \left[\sum_{i,j} U(r_i, r_j, r_{ij}) \right]$ is Jastrow term taking into account n-body effect. The standard minimal form was chosen for the

$$U(r_i, r_j, r_{ij}) = \frac{ar_{ij}}{(1 + br_{ij})} - p(r_i + r_j).$$

Fixed-node DMC method entails four types of possible errors. The most fundamental error is "fixed-node" error which is due to the approximation that the solution of the Schrödinger equation and the guide wavefunction will have same node. From the previous researches this error is generally small. [6] The

second error is short time error, which appears when Δt is not small enough. This error can be eliminated by conducting same calculation with different Δt and extrapolate the result to the zero Δt limit. The third error is population control error. Sometimes, the results of DMC depend on the number of walkers although the number of walkers supposed to affect only the efficiency of the sample, not the results. To consider this error, the calculation should be done with different number of walkers. The last error is the statistical uncertainty that necessarily accompanies the Monte Carlo method, which is inversely proportional to the number of Monte Carlo step. To reduce the error, large enough Monte Carlo steps should be done.

First, the energies of three lowest electronic states (2S , 2D , 2P) of Cu atom and the ground state (1S) of Cu^+ ion from Hartree-Fock, CCSD(T), and FN-DMC are compared in the Table 1.

	$E(HF)$	$E(CCSD(T))$	$E(FN - DMC)$
$\text{Cu}(^2S)$	-1638.9637	-1640.3971	-1640.411 \pm 0.0005
$\text{Cu}(^2D)$	-1638.9528	-1640.3276	-1640.342 \pm 0.0004
$\text{Cu}(^2P)$	-1638.8508	-1640.2619	-1640.273 \pm 0.0005
$\text{Cu}^+(^1S)$	-1638.7276	-1640.1204	-1640.137 \pm 0.0006

Table 1
All-electron total energies and correlation energies of the low-lying states of copper and its cation using HF, CCSD(T), and FN-DMC. Energies in Hartree. Reprint Table II of Ref [6]

The result shows that both CCSD(T) and FN-DMC energies are lower than the HF energies, so the correlation energy is restored for both methods. In addition, the FN-DMC energy is lower than the CCSD(T) energy. Therefore, the authors concluded that DMC is a better way to capture the correlation energy.[6].

Many people think that estimating the differences between energy levels is more important than estimating the energy itself. Table 2 shows the transition energies from HF, CCSD(T), and FN-DMC methods. The Hartree-Fock reproduced the transition energy from 2S to 2D badly compared to the results from other methods. This is because the occupation of 3d shell changes during the transition, so the electronic correlation plays a major role in the process. [6] For other transition, where the occupancy of 3d shell does not change, HF results is much better, but still deviate from the other results. The FN-DMC results are overlapping the CCSD(T) results within the statistical errors ($\approx 1500 \text{ cm}^{-1}$).

	$\Delta E(HF)$	$\Delta E(CCSD(T))$	$\Delta E(FN - DMC)$
$\text{Cu}(^2S \Rightarrow ^2D)$	2392	15254	15144 \pm 1405
$\text{Cu}(^2S \Rightarrow ^2P)$	24779	29673	30288 \pm 1552
$\text{Cu}(^2S) \Rightarrow \text{Cu}^+(^1S)$	51818	60729	60136 \pm 1714

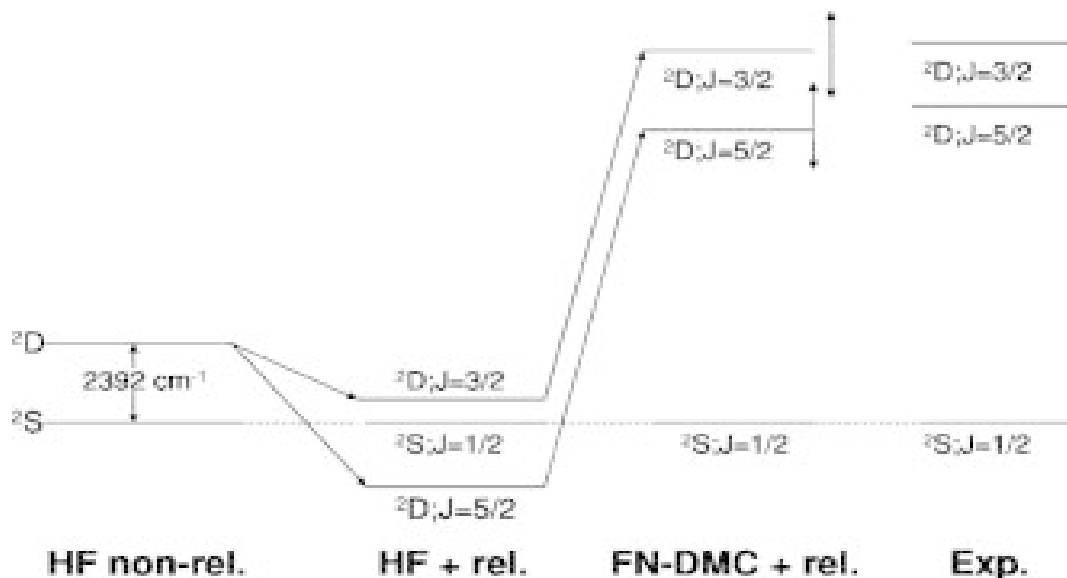
Table 2

Transition energies calculated from HF, CCSD(T), and FN-DMC. Energies in cm^{-1} Reprint Table III of Ref [6]

To describe the electron transition completely, the relativistic contributions needed to be evaluated. The relativistic contribution is calculated at the Dirac-Fock level. Table 3 shows the sum of the relativistic contribution ($\Delta E(DF)$) and non-relativistic transition energies ($\Delta E(CCSD(T))$ and $\Delta E(FN-DMC)$) and the experimental transition energies ($\Delta E(exp)$).

	J	$\Delta E(exp)$	$\Delta E(DF)$ + $\Delta E(CCSD(T))$	$\Delta E(DF) + \Delta E(FN-DMC)$
$Cu(^2S \Rightarrow ^2D)$	$\frac{5}{2}$	11203	11123	10333 ± 1405
	$\frac{3}{2}$	13245	13631	12841 ± 1405
$Cu(^2S \Rightarrow ^2P)$	$\frac{1}{2}$	30535	30529	31122 ± 1552
	$\frac{3}{2}$	30784	30711	31304 ± 1552
$Cu(^2S) \Rightarrow Cu^+ (^1S)$	0	62317	61875	61304 ± 1714

Table 3
The experimental transition energies from Ref [7] and the calculated transition energies with including the relativistic effects. Energies in cm^{-1} Reprint Table V of Ref [6]



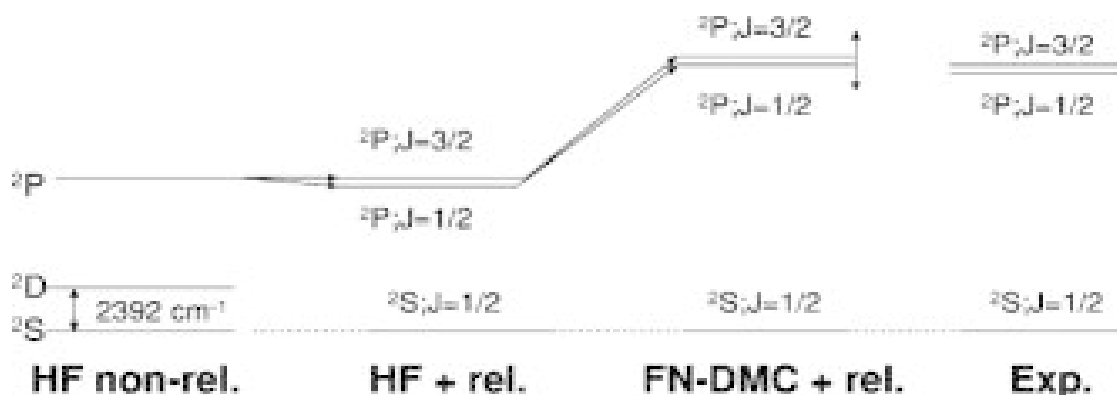


Figure 2

Pictograph represent the energy level of $^2S \Rightarrow ^2D$ (top) and $^2S \Rightarrow ^2P$ (bottom). FIG 1 and 2 of Ref [6]

The results shows that the relativistic contribution is not negligible in this system. The both experiment and CCSD results are in the error range of the FN-DMC results.

Let's compare the ionization energy to a most common DFT calculation.

```
> mol := [{"Cu", 0, 0, 0}]:
> data1 := DensityFunctional(mol, basis = "cc-pVTZ", xc = "B3LYP"):
> data2 := DensityFunctional(mol, basis = "cc-pVTZ", xc = "B3LYP", charge = 1):
> IE_hartree := (data2[e_tot] - data1[e_tot]);
IE_hartree := 0.29019452 (4.1.1)
```

```
> IE_cm := IE_hartree * 219474.63;
IE_cm := 63690.33406224 (4.1.2)
```

The IE from B3LYP/cc-pVTZ is about 64000 cm^{-1} , compared to the CCSD(T) and FN-DMC results (60000 cm^{-1}) the ionization energy is overestimated unlike the underestimating HF result ($\sim 52000 \text{ cm}^{-1}$). This results shows that the for transition metal like Cu, higher level of calculation is required, and FN-DMC can be one option.

IV. Discussion/Conclusions

The quantum Monte Carlo method is a method of calculating the energy of a complex system numerically. This method is classified into an easy and simple but error-embedded VMC method, and a DMC method that requires more resources but enables sampling from exact grand state. In order for the DMC method to sample the ground state, it must always use a guide function with the same sign as the ground state wave function. It's not always possible, so a fixed node assumption is used. The VMC method can be used in systems that are complex to calculate with conventional quantum calculations, such as transition metals. The biggest advantage of this method is that it is easy to optimize for multi-processor computers by using multiple independent workers. Considering the recent development of multi-core CPU and GPGPU clusters, active use of QMC is expected.

References

- (1) F. Jensen, [*Introduction to computational chemistry*](#), (John Wiley & sons, 2017)
- (2) A. Lüchow, [Wiley Interdiscip. Rev. Comput. Mol. Sci. 1\(3\), 388-402. \(2011\)](#). "Quantum Monte Carlo methods"
- (3) W. K. Hastings, [Biometrika, 57, 1, 97–109](#), "Monte Carlo sampling methods using Markov chains and their applications"
- (4) N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, [J. Chem. Phys. 21, 1087 \(1953\)](#), "Equation of state calculations by fast computing machines. The journal of chemical physics"
- (5) J. Toulouse, R. Assaraf, C. J. Umrigar, [Adv. Quant. Chem. 73, 285-314. \(2016\)](#), "Introduction to the variational and diffusion Monte Carlo methods. In Advances in Quantum Chemistry"
- (6) M. Caffarel, J. P. Daudey, J. L. Heully, A. Ramírez-Solís, [J. Chem. Phys. 123, 094102 \(2005\)](#) "Towards accurate all-electron quantum Monte Carlo calculations of transition-metal systems: Spectroscopy of the copper atom"
- (7) C. E. Moore, US Department of Commerce, National Bureau of Standards, (1949), "Atomic energy levels as derived from the analyses of optical spectra (No. 467)"