Quantum Monte Carlo with Coupled-Cluster wave functions

Alessandro Roggero,1,2 Abhishek Mukherjee,3,2 and Francesco Pederiva1,2

1 Physics Department, University of Trento, via Sommarive 14, I-38123 Trento, Italy
2 INFN-TIFPA, Trento Institute for Fundamental Physics and Applications
3 ECT*, Villa Tambosi, I-38123 Villazzano (Trento), Italy

(Dated: May 10, 2014)

We introduce a novel many body method which combines two powerful many body techniques, viz., quantum Monte Carlo and coupled cluster theory. Coupled cluster wave functions are introduced as importance functions in a Monte Carlo method designed for the configuration interaction framework to provide rigorous upper bounds to the ground state energy. We benchmark our method on the homogeneous electron gas in momentum space. The importance function used is the coupled cluster doubles wave function. We show that the computational resources required in our method scale polynomially with system size. Our energy upper bounds are in very good agreement with previous calculations of similar accuracy, and they can be systematically improved by including higher order excitations in the coupled cluster wave function.

Introduction.—Quantum Monte Carlo (QMC) methods have now become standard tools for computations in a wide variety of strongly correlated systems [1] ranging from quantum chemistry [2] to condensed matter [3], and nuclear physics [4]. The most celebrated capability of such methods is to provide very accurate estimates of ground state and thermodynamic properties of many-body systems, while enjoying a very favorable scaling of computational time with system size.

In fermionic systems, QMC methods are plagued by the ‘sign problem’. In principle, the sign problem can be circumvented in the ‘fixed-node’ (FN) approximation with the help of importance sampling with a trial ground-state wave function [5]. The spectacular success of the FN QMC is largely because of the development of high quality trial ground-state wave functions. However, most of these wave functions have forms which are convenient for calculations in coordinate space (r-space).

The lack of accurate and computationally efficient trial ground-state wave functions has, thus far, precluded a wide exploration of these algorithms within the configuration interaction (CI) scheme (see, however, Ref. [4]). Generally, QMC methods within the CI scheme tend to rely on auxiliary fields introduced via the Hubbard-Stratonovich transformation [6]. Much recent interest, however, was sparked by the demonstration that even within the CI scheme it is possible to apply stochastic projection to systems much larger than what would be possible using conventional matrix diagonalization [7].

The development of an efficient ground state QMC algorithm within the CI scheme/momentum space would be of great interest in nuclear physics, where most of the modern interactions are written in non-local forms [8]; and in electronic structure calculations with non-local pseudopotentials [9].

On the other hand, coupled cluster (CC) wave functions provide very accurate and size extensive approximations for the ground state wave function for many physical systems of interest [11]. The CC energies towards accurate QMC calculations in many-body systems [12]. In fact, CC calculations are considered to be the ‘gold standard’ in quantum chemistry [13]. More recently, CC theory has also reemerged as a method of choice in nuclear structure calculations [14].

In this Letter, we introduce a novel scheme to combine these two powerful many-body techniques. We propose to use CC wave functions as importance functions for a projection quantum Monte Carlo algorithm within the configuration interaction scheme, which we call configuration interaction Monte Carlo (CIMC). In CIMC the ground state wave function of a CI Hamiltonian is filtered out by propagating the amplitudes of an initial arbitrary wavefunction via a random walk in the many body Hilbert space spanned by a basis of Slater determinants [15]. The CC wave functions are used to guide this random walk via importance sampling in order to circumvent the sign problem. Our method provides rigorous upper bounds to the ground state energy whose tightness can be systematically improved by including higher order excitations in the CC wave function.

We apply our method to the three dimensional homogeneous electron gas (3DEG) in momentum space. The 3DEG is described by a simple Hamiltonian; it nevertheless encapsulates many of the difficulties associated with modern many-body theories. In particular, it has both the weakly and strongly correlated regimes which can be accessed via a single tunable density parameter, the Wigner-Seitz radius Rs, thus providing an ideal system for benchmarking many body theories [10, 21].

Monte Carlo in Fock space.— The CIMC projection algorithm is discussed in detail in Ref [15]. We describe it briefly for completeness. Consider a general second quantized fermionic CI Hamiltonian which includes only two-body interactions (atomic units will be used throughout)

\[ H = \sum_{i \in S} \varepsilon_i a_i^+ a_i + \sum_{abij \in S} V^{ab}_{ij} a_i^+ a_j a_i a_j, \]  

(1)
where $\alpha_i^d$ creates a particle in the single-particle (sp) state labeled by $i$ ($i$ is a collective label for all sp quantum numbers). The set $S$ of sp states is assumed to be finite and of size $N_s$. The $V_{ij}^{ab}$ are general two-body interaction matrix elements.

For homogeneous systems we can use the plane wave states, with definite momentum and spin, as the sp basis set. The sp energies are $\varepsilon_i = k_i^2/2m$, where $k_i$ is the momentum of the $i$-state, and $m$ is the fermion mass. We include in our sp basis all single-particle states $i$ with $k_i^2 \leq k_{\text{max}}^2$. In principle, cutoff independent results can be obtained by performing successive calculations with increasing $k_{\text{max}}$ and then extrapolating to $k_{\text{max}} \to \infty$.

For the 3DEG, the $V_{ij}^{ab}$ are given by

$$V_{ij}^{ab} = \langle 1 - \delta_{k_i, k_j} \rangle \delta_{k_i+k_j, k_a+k_b} \frac{4\pi}{\Omega} \frac{1}{(k_a - k_i)^2}. \quad (2)$$

The volume $\Omega$ of the simulation box (and hence the unit spacing in momentum) is determined by the density and the number of particles $N$ in the simulation. We ignore the Madelung term because it does not affect the correlation energy.

The many-body Hilbert space is spanned by the set of all $N$-particle Slater determinants constructed from the sp orbitals $i \in S$. We will denote these Slater determinants or ‘configurations’ with the vector notation, $|n\rangle$, where $n \equiv \{n_i\}$ and $n_i = 0, 1$ is the occupation number of the sp orbital $i$ in $n$.

The ground state wave function $\Psi_{gs}$ of $H$ can be projected out by repeated applications of the projection operator $P = 1 - \Delta \tau (H - E_T)$, on some initial wave function $\Psi_0$ which has a non-zero overlap with $\Psi_{gs}$,

$$|\Psi_{gs}\rangle = \lim_{M \to \infty} P^M |\Psi_0\rangle. \quad (3)$$

where the energy shift $E_T$ and the imaginary time step $\Delta \tau$ are related by $\Delta \tau < 2/(E_{\text{max}} - E_T)$, with $E_{\text{max}}$ being the maximal eigenvalue of $H$.

In a Monte Carlo algorithm the wave function at any time step $M$, $\Psi_M$, is represented as an ensemble of configurations. The one time-step projection (single application of $P$) is performed independently and stochastically for each configuration, by interpreting the matrix elements $\langle m|P|n\rangle$ as probabilities.

In general (the 3DEG inclusive), $P$ will have one or more negative off-diagonal matrix elements, which cannot be interpreted as probabilities. In principle, a stochastic evolution can still be carried out by evolving ‘signed’ configurations, but this leads to an exponential decay in the signal to noise ratio with increasing time-step $M$. This is a manifestation of the sign problem. Recently, it was shown that the sign problem can be somewhat moderated by including a configuration-annihilation step in the evolution [22]. Nevertheless, even the latter algorithm has exponential scaling with the system size, though with a reduced exponent.

In CIMC, we circumvent the sign problem with the help of an importance function $\Phi_G$. We define a family of Hamiltonians, $H_\gamma$, whose off-diagonal matrix elements are given by [24, 25],

$$\langle m|H_\gamma|n\rangle = \begin{cases} -\gamma \langle m|H|n\rangle & s(m, n) > 0 \\ \langle m|H|n\rangle & \text{otherwise} \end{cases}, \quad (4)$$

while the diagonal matrix elements are given by,

$$\langle n|H_\gamma|n\rangle = \langle n|H|n\rangle + (1 + \gamma) \sum_{m \neq n} s(m, n). \quad (5)$$

where $s(m, n) = \Phi_G(m)\langle m|H|n\rangle/\Phi_G(n)$.

In addition, we define a new propagator $P_\gamma$ as,

$$\langle m|P_\gamma|n\rangle = 1 - \Delta \tau \Phi_G(m)\langle m|H_\gamma - E_T|n\rangle/\Phi_G(n). \quad (6)$$

The propagator $P_\gamma$, by construction, is free from the sign problem for $\gamma \geq 0$, and filters out the wave function $\Phi_G(n)|\gamma_\gamma\rangle$, where $|\gamma_\gamma\rangle$ is the ground state wave function of $H_\gamma$.

The ground state energy $E_\gamma$ of $H_\gamma$ is an strict upper bound for the ground state energy $E_{gs}$ of the true Hamiltonian $H$ for $\gamma \geq 0$, and this upper bound is tighter than the variational upper bound $\langle \Phi_G|H|\Phi_G \rangle$ [15, 24, 25]. In addition, a linear extrapolation of $E_\gamma$ from any two values of $\gamma$ to $\gamma = -1$ also provides a rigorous upper bound on $E_{gs}$ [27].

We found the extrapolation using $\gamma = 0, 1$ to be best compromise between accuracy and statistical error. Thus, the tightest upper bound for the ground state energy $E_0$ is $E_{CIMC} = 2E_{\gamma=0} - E_{\gamma=1}$.

The simple projection algorithm described above becomes extremely inefficient for large $E_{\text{max}}$, i.e., for large $N$ or $N_s$. In practice, we use a much more efficient algorithm (free from any time-step error) which was proposed in Ref. [22] (see also, Ref. [25]).

**Coupled cluster importance functions.**—A good choice for the importance function $\Phi_G$ is crucial for the success of our method. We need $\Phi_G$ to be sufficiently flexible to include the dominant correlations in the system, and yet be quick to evaluate. In many strongly correlated systems, CC wave functions fulfill the first criterion. Below, we describe a recursive algorithm which can be used to evaluate CC wave functions very efficiently.

The CC wave function can be written as

$$|\Phi_{CC}\rangle = e^T|\Phi_0\rangle \quad (7)$$

where $\Phi_0$ is a model wave function and $T$ is an operator which generates excitations on $\Phi_0$. We choose $\Phi_0 = \Phi_{HF}$, i.e., the Hartree-Fock wave function. Due to momentum conservation, in a homogeneous system the simplest non-trivial approximation for $T$ is $T \equiv T_2 = \sum_{ij,ab} \delta_{ij} \delta_{ab} \alpha_i^\dagger \alpha_j^\dagger \alpha_j \alpha_i$, where sp states $i$ and $j$ (a and $b$) are occupied (unoccupied) in $\Phi_{HF}$. This constitutes the so-called coupled cluster doubles (CCD) approximation for the ground state wave function.
Writing the wave function in terms of the excitations

\[ \Phi_{\text{CCD}}^m \left( p_1 p_2 \cdots p_m \right) = \sum_{\gamma=2}^{m} \sum_{\mu<\nu} \left( - \right)^{\gamma+\mu+\nu} r_{p_\mu p_\nu} \Phi_{\text{CCD}} \left( p_1 p_2 \cdots p_{\mu-1} p_{\mu+1} \cdots p_{\nu-1} p_{\nu+1} \cdots p_m \right) \]

for even \( m > 0 \). In the above equation, \( \Phi_{\text{CCD}}^m \left( p_1 p_2 \cdots p_m \right) = \Phi_{\text{CCD}}(n) \) for \( n = a_{p_1}^i \cdots a_{p_m}^i a_{h_1} \cdots a_{h_m} \Phi_{\text{HF}} \), with \( p_1 < \ldots < p_m \) and \( h_1 < \ldots < h_m \). The wave function has a vanishing component for odd \( m \); and the component for \( m = 0 \) is fixed by our choice of normalization to be 1.

Using Eq. (5) we can easily calculate \( \Phi_{\text{CCD}} \) recursively. The computational effort for a single calculation has a combinatorial scaling with the average number of particle-hole excitations in \( \Phi_G \). However, it does not depend on \( N \) or \( N_s \).

Although for this exploratory study we have used the CCD wave function, it is evident that similar recursive relations can be easily written for other CC type wave functions, e.g., CCSD or CCSDT, which will become necessary for applying our method to inhomogeneous systems or, systems with stronger correlations or many-body interactions.

Results for the homogeneous electron gas.—In Table I we show the CCD energies, calculated using conventional CC theory, along with the corresponding Monte Carlo energies of an 3DEG system with \( N = 14 \) and \( N_s = 342 \), for \( r_s = 0.5, 1.0 \) and 2.0. We see that, for \( r_s = 0.5 \) the CCD energy is very close to the corresponding Monte Carlo energy. But, for \( r_s = 1.0 \) and 2.0 they are, in fact, lower than the corresponding Monte Carlo energies. Since, \( E_{\text{CIMC}} \leq \langle \Phi_{\text{CCD}} | H | \Phi_{\text{CCD}} \rangle \) (see discussion above), this shows, once again, the non-variational nature of the energies obtained from conventional CC theory.

| \( r_s \) | CCD | + CIMC | CCD(1) | + CIMC |
|---|---|---|---|---|
| 0.5 | -0.572682 | -0.5729(3) | -0.659641 | -0.5733(2) |
| 1.0 | -0.506701 | -0.5021(3) | -0.657347 | -0.5025(2) |
| 2.0 | -0.417946 | -0.40317(2) | -0.665071 | -0.4029(3) |

TABLE I. Correlation energies (a.u.) for \( N = 14 \) and \( N_s = 342 \) from conventional CC theory with the CCD and CCD(1) wave functions, along with the corresponding CIMC energies using each as importance functions. The numbers in parenthesis indicate statistical error in the last significant digit.

In principle, the CC amplitudes \( t_{ij}^{ab} \) can be obtained from conventional CC theory. However, solving the CC equations is computationally expensive. Therefore, we investigated the possibility of using less computationally demanding ways of obtaining the \( t_{ij}^{ab} \), while still preserving the structure of the CC wave function.

A simple option is to compute \( t_{ij}^{ab} \) by the second order Møller-Plesset perturbation theory (MP2). If the CC equations are solved iteratively, then this is equivalent to stopping after the first iteration. In Table II we also compare the energies obtained from our CIMC projection using the MP2, denoted by CCD(1).

The CCD(1) amplitudes produce a worse approximation to the ground state wave function as compared to the full CCD amplitudes. Nevertheless, when used as importance functions in our CIMC algorithm, the final estimate for the ground state energy for both cases are very close. The statistical errors are comparable for \( r_s = 0.5 \) and 1.0. For \( r_s = 2.0 \) they are about an order of magnitude lower when the full CCD amplitudes are used.

The above observation is extremely encouraging because it means that, for our purposes, it may not necessary to solve the full CC equations to get reasonable \( t_{ij}^{ab} \) amplitudes. Of course, we do not expect the CCD(1) amplitudes to be satisfactory for more strongly correlated systems. Still, even in those cases one can, presumably, use computationally inexpensive approximations to the full CC equations. For the rest of this work, all the CIMC results have been computed using the CCD(1) amplitudes.

In Fig. 1 we show the CIMC ground state energy estimates for \( N = 14 \) and \( r_s = 0.5, 1.0, 2.0 \) and 3.0 for some of our large basis size calculations. In Refs. 20, 21, and 28, it was suggested that for the 3DEG it might be possible to extrapolate to the \( N_s \to \infty \) limit by employing a linear \( 1/N_s \) dependence of the correlation energy for large but finite \( N_s \). Although, for \( r_s = 0.5 \) and \( N = 14 \) such a linear trend in the correlation energy is visible, for the other values of \( r_s \) shown in the figure, no such trend is evident. The situation is similar for calculations we performed with \( N = 32 \) and 54. Thus, at least up to our largest basis size \( N_s = 2378 \), we cannot safely do an extrapolation to \( N_s \to \infty \) with a reasonably low \( \chi^2 \).

In Table III we show the ground state energy of different \( r_s \) and \( N \) for the largest \( N_s (= 2378) \) calculated by us. For comparison, we also include the energy estimates from Refs. 19 and 21.

The energies in Ref. 19 are calculated using the r-space diffusion Monte Carlo method with an importance function that included backflow correlation on top of the
for a CI Hamiltonian. However, most calculations are performed using the initiator approximation (i-FCIQMC) with adds a bias to the energy estimate. The energies in i-FCIQMC are not necessarily upper bounds to the true ground state energy. Due to the sign problem, the computational resources required in either FCIQMC or i-FCIQMC scale as the size the many body Hilbert space, i.e., they are exponential in the system size, \( N \), and the basis size, \( N_s \).

Our method, instead, provides strict energy upper bounds with a bias due to the fixed-node approximation. Nevertheless, they are believed to be highly accurate.

On the other hand, in Ref. \[21\] the energies are calculated in a finite CI like basis set (as in this work), using the initiator full configuration interaction quantum Monte Carlo method (i-FCIQMC). The \( N_s \rightarrow \infty \) results are obtained by using the so called ‘single point extrapolation’ from much smaller values of \( N_s \) than ours.

Our finite-basis set results are already in good agreement with the other calculations, capturing between 93\% to 99\% of the correlation energy. The energy upper bounds can be systematically improved by including higher order excitations (triples) in the CC wave function and by using larger basis sizes. Possibly, a much faster way to achieve basis set convergence is to use a finite basis renormalized Coulomb interaction \[12\].

Our method shares many similarities with the FCIQMC method. The FCIQMC method, in principle, can provide exact ground state energies for a CI Hamiltonian. However, most calculations are performed using the initiator approximation (i-FCIQMC) with adds a bias to the energy estimate. The energies in i-FCIQMC are not necessarily upper bounds to the true ground state energy. Due to the sign problem, the computational resources required in either FCIQMC or i-FCIQMC scale as the size the many body Hilbert space, i.e., they are exponential in the system size, \( N \), and the basis size, \( N_s \).

Slater-Jastrow wave function. These are strict energy upper bounds with a bias due to the fixed-node approx-
imation. Nevertheless, they are believed to be highly accurate.

On the other hand, in Ref. \[21\] the energies are calculated in a finite CI like basis set (as in this work), using the initiator full configuration interaction quantum Monte Carlo method (i-FCIQMC). The \( N_s \rightarrow \infty \) results are obtained by using the so called ‘single point extrapolation’ from much smaller values of \( N_s \) than ours.

Our finite-basis set results are already in good agreement with the other calculations, capturing between 93\% to 99\% of the correlation energy. The energy upper bounds can be systematically improved by including higher order excitations (triples) in the CC wave function and by using larger basis sizes. Possibly, a much faster way to achieve basis set convergence is to use a finite basis renormalized Coulomb interaction \[12\].

Our method shares many similarities with the FCIQMC method. The FCIQMC method, in principle, can provide exact ground state energies for a CI Hamiltonian. However, most calculations are performed using the initiator approximation (i-FCIQMC) with adds a bias to the energy estimate. The energies in i-FCIQMC are not necessarily upper bounds to the true ground state energy. Due to the sign problem, the computational resources required in either FCIQMC or i-FCIQMC scale as the size the many body Hilbert space, i.e., they are exponential in the system size, \( N \), and the basis size, \( N_s \).

Our method, instead, provides strict energy upper bounds, the tightness of which can be systematically improved by improving the quality of the importance function. Importantly, the computational cost of our Monte Carlo algorithm nominally grows as \( N^2(N_s-N) \), for both computational time (per Monte Carlo step) and memory requirements. Due to this polynomial scaling, we expect our method to be applicable to much larger systems than those manageable by conventional diagonalization methods or by (i-)FCIQMC. We see evidence of this in
Table III where we compare computational time in our method with that in i-FCIQMC. The statistical error for the two methods are comparable in each case.

Conclusion.— We have introduced a many body technique which combines the power of ground state projection Monte Carlo with coupled cluster theory. As a first benchmark and application we studied the homogeneous electron gas in momentum space for large single particle basis sizes. Our results are already in very good agreement with existing accurate calculations, and they can be systematically improved. In principle, the fixed-node bias can be removed by performing a further projection starting from our CIMC wave functions with the true projector $\mathcal{P}$ and signed configurations using FCIQMC and the semistochastic projector Monte Carlo method [29].

The use of CC wave functions as importance functions in our Monte Carlo algorithm serves dual purposes. On the one hand, given that the CC wave functions in our Monte Carlo algorithm serve dual purposes, the semistochastic projector Monte Carlo method [29]. The statistical error for $\mathcal{P}$ and signed configurations using FCIQMC and the semistochastic projector Monte Carlo method [29].

On the other hand, the energies obtained from conventional CC theory are not variational. However, the energies we obtain from CIMC are rigorous upper bounds for the exact ground state energy. By using CC wave functions in our Monte Carlo we are providing the ‘best’ variational energy one can hope to get using the CC class of wave functions. Given the importance of CC type wave functions in modern quantum many body calculations, the second purpose is as important as the first.

Computations were performed at the Open Facilities at Lawrence Livermore National Laboratory. We would like to thank C. Umrigar for stimulating discussions and careful reading of the manuscript, and Y. Alhassid for comments on the manuscript.

---

1 roggero@science.unitn.it
2 mukherjee@ctstar.eu
3 pederiva@science.unitn.it

[1] M. P. Nightingale and C. J. Umrigar, Quantum Monte Carlo methods in physics and chemistry, Vol. 525 (Springer, 1998).
[2] D. M. Ceperley and L. Mitas, in New Methods in Computational Quantum Mechanics//Advances in Chemical Physics, XCVIII, edited by I. Prigogine and S. A. Rice (John Willey & Sons, NY, 1995); J. B. Anderson, Rev. Comp. Chem. 13, 133 (2007).
[3] K. Binder, ed., The Monte Carlo method in condensed matter physics, Top. Appl. Phys., Vol. 71 (1995); W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Rev. Mod. Phys. 73, 33 (2001).
[4] S. C. Pieper, Nucl. Phys. A751, 516 (2005).
[5] P. J. Reynolds, D. M. Ceperley, B. J. Alder, W. A. Lester Jr, et al., J. Chem. Phys. 77, 5593 (1982).
[6] M. Kolodrubetz and B. K. Clark, Phys. Rev. B 86, 075109 (2012).
[7] G. Sugiyama and S. Koonin, Ann. Phys. 168, 1 (1986); S. Zhang, in Theoretical Methods for Strongly Correlated Electrons, CRM Series in Mathematical Physics, edited by D. Sénéchal, A.-M. Tremblay, and C. Bourbannais (Springer New York, 2004) pp. 39–74; A. Bulgac, J. E. Drut, and P. Magierski, Phys. Rev. A 78, 023625 (2008) J. E. Drut, T. A. Lähde, G. Włazłowski, and P. Magierski, Phys. Rev. A 85, 051601 (2012); J. Carlson, S. Gandolfi, K. E. Schmidt, and S. Zhang, Phys. Rev. A 84, 061602 (2011).
[8] G. H. Booth, A. Grünies, G. Kresse, and A. Alavi, Nature 493, 365 (2013).
[9] U. Van Kolck, Prog. Part. Nucl. Phys. 43, 337 (1999).
[10] W. E. Pickett, Phys. Rev. B 25, 1218 (1982).
E. Neuscamman, C. J. Umrigar, and G. K.-L. Chan,
Phys. Rev. B 85, 045103 (2012)