

Stochastic multi-reference perturbation theory with application to linearized coupled cluster method

Guillaume Jeanmairet and Sandeep Sharma

*Max Planck Institute for Solid State Research,
Heisenbergstraße 1, 70569 Stuttgart, Germany*

Ali Alavi

*Max Planck Institute for Solid State Research,
Heisenbergstraße 1, 70569 Stuttgart, Germany and
Department of Chemistry, University of Cambridge,
Lensfield Road, Cambridge CB2 1EW, United Kingdom*

Abstract

In this article we report a stochastic evaluation of the recently proposed LCC multireference perturbation theory [Sharma S., and Alavi A., *J. Chem. Phys.* **143**, 102815, (2015)]. In that method both the zeroth order and first order wavefunctions are sampled stochastically by propagating simultaneously two populations of signed walkers. The sampling of the zeroth order wavefunction follows a set of stochastic processes identical to the one used in the FCIQMC method. The computation of the perturbation requires to introduce new rules for the stochastic process. In particular it involves a source term coming from the zeroth order wavefunction. With this fully stochastic method, it will become possible to treat simultaneously large active spaces to account for the static correlation and to recover the dynamical correlation using perturbation theory. The second order response energy is also computed stochastically, the different contribution being sampled according to the spawning probability between the zeroth and first order wavefunction.

To illustrate the method, we study a few organic molecules, going from carbon dimer to aromatic molecules. We computed singlet-triplet gaps of benzene and m-xylylene. For m-xylylene which has proved difficult for standard CASSCF+PT we found a gap in good agreement with the experimental values.

I. INTRO

One of the significant challenges in quantum chemistry is the description of electronic systems that simultaneously display chemically-relevant static and dynamical electron correlation. Static correlation often occurs in open shell systems and gives rise to long-ranged, highly entangled, many-electron wavefunctions involving electronic orbitals close to the Fermi energy. Dynamical correlation, on the other hand, correlates electrons on a short-length scale, and whose descriptions requires the single-particle spectrum to extend over many energy scales. There are many systems of practical interest that fall into this class: we can mention transition metal clusters that are found in many protein active sites and that play a key role in a wide number of important biological processes such as photosynthesis or respiration[1]. Because of the multideterminantal nature of the wave-function they can prove hard to study with single-reference approaches such as the widely used density functional theory[2].

Unfortunately the computational requirements to describe the wavefunction of such systems from an exact perspective is utterly daunting. In an ideal scenario, one would allow for a "full" correlation treatment in a large basis: in other words all electrons (even those not close to the Fermi energy) are simultaneously correlated over the entire basis. Such a treatment (full CI), which would amount to the exact solution of the Schrodinger equation in the given basis, is generally out of reach for sufficient numbers of electrons, owing the combinatorial explosion of the Hilbert space of a many-particle system. To overcome this limitation the calculation can be carried by considering only a meaningful subset of the available configurations, the most popular choice is to use a complete active space (CAS) wave function[3–5]. This approach allows to tackle bigger systems than the FCI one however it suffers for the same exponential scaling problem. Different methods have been proposed to allow the use of larger active spaces such as the restricted active space (RAS)[6, 7], generalized active space (GAS)[8] and SplitGas[9] approaches. With modern techniques such as Density Matrix Renormalization Group[10–18] or the Full CI Quantum Monte Carlo[19–23] technique one can treat very large Hilbert spaces, corresponding of up 30-40 electrons in 30-40 orbitals, but even such techniques struggle to handle systems with many hundreds of electrons correlating in hundreds or even thousands of orbitals, which is certainly necessary to able to treat as we move to even intermediate-size molecules.

Approximations in the correlation treatment are therefore necessary, and given the energetic divisions between static and dynamical correlation, multi-reference perturbation theory (MRPT) is a natural starting point, and which leads to a variety of "active-space" methods. Full treatment of correlation among a predefined set of active orbitals (usually chosen to be around the Fermi energy) with a given number of electrons leads to the reference Hamiltonian, followed by a perturbation treatment of the remaining terms in the Hamiltonian which arise from the remaining terms in the Hamiltonian. In this spirit, one can mention CASPT2[24, 25], NEVPT2[26, 27] and MRCI methods[28], as well as

the linearised coupled cluster (LCC) developed recently by two of us[29]. In the perspective of dealing with large systems it is worth noticing that a second-order perturbation theory approach based on the generalized active space self-consistent-field has also recently been proposed[30].

All of these multireference perturbation theories involve a deterministic resolution of the perturbation equations, and the cost of the calculation of the perturbation by itself quickly becomes intractable with the number of core and virtual orbitals. Although this can be dealt by making a further approximation known as internal contraction, which comes at the cost of computing the reduced density matrices (RDM) of the active space up to fourth order, which is itself a significant challenge for large active spaces.

A stochastic resolution of the perturbation equation seems to be a promising direction to reduce the cost of the calculation and to avoid the use of internal contraction. Surprisingly, only a few attempts to implement a stochastic resolution of perturbation theories have been proposed, among which we can mention the stochastic evaluation of MP2 energies by Hirata and collaborators, MC-MP2. This approach is based on the rewriting of second order perturbation equations into the sum of two 13-dimensional integrals thanks to Laplace transform. Those integrals are then evaluated through Monte-Carlo integration[31–33]. Another approach to solve MP2 equations is to express its contributions in term of graph that describe set of connected Slater determinants, and then to stochastically sample those graphs[34]. However those two approaches involve a single reference zero order wavefunction and to our knowledge stochastic resolution of multireference perturbation theories have not been proposed yet.

The purpose of this paper is to show how a stochastic treatment of multi-reference perturbation theory can be implemented within the FCIQMC methodology, namely a walker-based method to solve for the response wavefunction of perturbation theory. Response theory differs from the eigenvalue problem of diagonalisation in that involves the solution of a linear system of equations (the response equations). Since the FCIQMC technique was developed as a ground-state eigenvalue solver, it requires to be modified and generalized in order to handle this new setting. We show how this can be done for the MRLCC method, although it can be similarly implemented for other flavors of multi-reference perturbation theory. Importantly, the resulting method which we term LCCQMC, is a fully uncontracted method, and is therefore potentially more accurate than the internally contracted approximations.

In our new method the sampling of zeroth and first order wavefunctions are done simultaneously. The population on the zero order wavefunction follows the standard FCIQMC rules while the population dynamics of the response wavefunction contains a source term that depends on the population of the zero order wavefunction. This is practically done by allowing walkers on one replica to spawn new walkers on another replica. The structure of the rest of this article is the following, in the next part we

will recall the governing equations of MRPT, with a particular emphasis on the MRLCC method. We also recall some basics of the FCIQMC methods before showing how the MRLCC can be expressed in the “FCIQMC language”, i.e. as a stochastic propagation of a signed walkers population. In the third part a description of the algorithmic of the QMC-LCC method that has been implemented in the `neci` code is given, we also discuss some important technical points. We then illustrate the potential of the method by applying it to some organic molecules. We first tested the method by studying the carbon dimer with systematically more refined basis set, going from cc-pVDZ to cc-pVQZ. Afterwards we turned our attention to the evaluation of Singlet-Triplet gaps; first in the case of the benzene molecule which has a singlet for ground state, and then in the case the m-xylylene diradical that admits a triplet as its ground state.

II. THEORY

A. LCC Perturbation theory

The essence of quantum-mechanical perturbation theories is to split the total Hamiltonian \hat{H} , into the sum of a simpler Hamiltonian \hat{H}_0 and a perturbation operator \hat{V} [35],

$$\hat{H} = \hat{H}_0 + \hat{V}, \quad (1)$$

with

$$\hat{V} = \hat{H} - \hat{H}_0. \quad (2)$$

The zero order order energy and wavefunction, E_0 and $|\Psi_0\rangle$ are solution of the following eigenproblem,

$$\hat{H}_0 |\Psi_0\rangle = E_0 |\Psi_0\rangle, \quad (3)$$

which is assumed to be, and is generally, possible to solve exactly.

In multireference perturbation theories, the zeroth order wavefunction is expressed as a linear combination of Slater determinants $|D_i\rangle$,

$$|\Psi_0\rangle = \sum_i c_i |D_i\rangle, \quad (4)$$

where the expansion set of determinants $\{|D_i\rangle\}$ is chosen to recover most of the correlation. Usually the set $\{|D_i\rangle\}$ is a CASCI space, it then contains all interactions between active electrons.

However the choice of \hat{H}_0 is not unique since several operators will admit the CASCI wavefunction as an eigenvector. Among the most popular ones we can mention the Fock operator used in CASPT[24, 36], the Dyll[37] Hamiltonian used in NEVPT[26, 27] and the excitation conserving Hamiltonian

of Fink[38, 39] used in the recently proposed MPS-LCC theory[29, 40]. Since MRLCC seems to outperformed other methods of similar cost, this is the one that is used in this study.

If we split the orbitals into an active set where the orbital occupancy can be 0, 1 or 2, a core set where the orbitals are doubly occupied and a virtual set of empty orbitals then the total Hamilton, \hat{H} and Fink's Hamiltonian, \hat{H}_0 can be expressed in second quantization as,

$$\hat{H} = \sum_{ij} t_{ij} a_i^\dagger a_j + \sum_{ijkl} \langle ij|kl \rangle a_i^\dagger a_j^\dagger a_l a_k \quad (5)$$

$$\hat{H}_0 = \sum_{\substack{ij; \\ \Delta n=(0,0,0)}} t_{ij} a_i^\dagger a_j + \sum_{\substack{ijkl; \\ \Delta n=(0,0,0)}} \langle ij|kl \rangle a_i^\dagger a_j^\dagger a_l a_k \quad (6)$$

where i, j, k, l refer to any orbitals and Δn denotes the change in the total number of electrons between the three subsets of orbitals. The only operators belonging to \hat{H}_0 are the ones that do not transfer electrons between the three subsets.

The successive correction ($|\Psi_m\rangle$) to the zeroth order wavefunction can be computed by using the following equation,

$$(\hat{H}_0 - E_0) |\Psi_m\rangle = -Q \left(\hat{V} |\Psi_{m-1}\rangle - \sum_{k=1}^{m-1} E_k |\Psi_{m-k}\rangle \right), \quad (7)$$

where Q is the projector onto the orthogonal space of the zeroth order wavefunction. Those sets of equation can be solved sequentially to compute the m^{th} order of the wavefunction, $|\Psi_m\rangle$. Once $|\Psi_m\rangle$ is known the $2m$ and $2m + 1$ energies can be computed thanks to Wigner's rules:

$$E_{2m} = \langle \Psi_{m-1} | V | \Psi_m \rangle - \sum_{k=1}^m \sum_{j=1}^{m-1} E_{2m-k-j} \langle \Psi_k | \Psi_j \rangle, \quad (8)$$

$$E_{2m+1} = \langle \Psi_m | V | \Psi_m \rangle - \sum_{k=1}^m \sum_{j=1}^m E_{2m+1-k-j} \langle \Psi_k | \Psi_j \rangle. \quad (9)$$

Note that with the definition of the zeroth order Hamiltonian given in Eq.6, the first order energy $E_1 = \langle \Psi_0 | V | \Psi_0 \rangle$ vanishes.

In a recent paper[29], both the 0 order wavefunction and the successive corrections were expressed as matrix product states (MPS) and computed deterministically by functional minimization. Here we proposed an alternative approach where both the zero order and the perturbation wavefunctions are sampled stochastically in the Fock space, the perturbation second-order energy is also evaluated stochastically.

To sample the zero order wavefunction and energy we used the standard FCIQMC approach restricted to the CAS space. We will recall here the main points of this approach.

B. Elements of FCIQMC

FCIQMC is a method which aims at stochastically minimizing the energy of a ground state wavefunction expressed as a CASCI (or Full-CI) expansion. The wavefunction can be expressed as a linear combination of determinants belonging to the CAS-CI space.

$$|\Psi\rangle = \sum_i c_i |D_i\rangle. \quad (10)$$

Formally, the idea is to find the ground state of the Hamiltonian operator \hat{H} , by integrating the imaginary time Schrodinger equation (ITSE),

$$\frac{\partial |\Psi\rangle}{\partial \tau} = -\hat{H} |\Psi\rangle. \quad (11)$$

The discretization of Eq.11 with a time step $\Delta\tau$ leads to the following evolution equation

$$|\Psi(t + \Delta\tau)\rangle = \left(\mathbb{1} - \Delta\tau (\hat{H} - S\mathbb{1}) \right) |\Psi(t)\rangle. \quad (12)$$

S is a shift parameter used to control the walkers population and $\mathbb{1}$ is the identity operator. Thus starting from a guess wavefunction, for instance the Hartree Fock determinant, the ground state can be reached by repetitively applying the following projector .

$$\hat{P} = \mathbb{1} - \Delta\tau (\hat{H} - S\mathbb{1}), \quad (13)$$

To circumvent the prohibitive storage of the full CI vector, in FCIQMC this projection operation is realized stochastically such as the proper projection is recovered on average. To do so the coefficient of the expansion in Eq.10 are sampled by a population of signed walkers. Each of those carries a signed weight and is located on a Slater determinant. The total signed sum of the walkers residing on the same determinant can be interpreted as an instantaneous measure of its weight c_i . The walker population evolves through a set of stochastic processes that mimic the projector of Eq.13:

1. A cloning/death step, in which the walker population on each determinant is increased/reduced with a probability $(H_{ii} - S) \Delta\tau$. S is a shift parameter that is used to control the total walker population.
2. A spawning step. For each walker on a determinant $|D_i\rangle$ a singly or doubly connected determinant $|D_j\rangle$ is generated with a probability p_{gen}^{ij} . A signed child is actually generated on the determinant $|D_j\rangle$ with a spawning probability

$$p_{spawn}^{ij} = \frac{|H_{ij}| \Delta\tau}{p_{gen}^{ij}}. \quad (14)$$

The sign of the newly spawned walker is the same as the sign of the parent if $H_{ij} > 0$, it is of opposite sign otherwise.

3. Each pairs of negative and positive newly spawned walkers lying on the same determinant are removed during an Annihilation step. This avoid the growth of an infinite noise due to the so-called sign problem.

We propose here a modification of the FCIQMC algorithm in order to stochastically sample simultaneously the zeroth order wavefunction and the successive order of the perturbation of Eq.7. Even if in principle any order of perturbation can be reached by this technique, in this article we will only consider the calculation of the first order correction to the wave function, that is given through Eq.7 by

$$|\Psi_1\rangle = \left(\hat{H}_0 - E_0\right)^{-1} Q\hat{V} |\Psi_0\rangle \quad (15)$$

In that case the problem is simpler since the Hilbert space on which the zeroth and first order wavefunctions are expanded is limited to the CASSD space. Moreover this space can be expressed as a direct sum of two subspaces $\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1$, where \mathcal{H}_0 correspond to the CAS space and \mathcal{H}_1 is its orthogonal complement which contains all the determinants that are single or double excitations from the ones belonging to \mathcal{H}_0 . Applying \hat{V} to $|\Psi_0\rangle$ only generates determinants on \mathcal{H}_1 . There is thus no need to ensure the orthogonality of the two wavefunctions and the Q projector operator in Eq.15 can be dropped. Of course higher order wavefunctions would contain determinants that are higher order excitations from the CASCI space. Orthogonalization with respect to $|\Psi_0\rangle$ and to the lower order perturbation wavefunctions would also be required.

The zero order wavefunction follows an ITDSE similar to Eq. 11,

$$\frac{\partial |\Psi_0\rangle}{\partial \tau} = -\hat{H}_0 |\Psi_0\rangle, \quad (16)$$

it thus can be solved by using the standard FCIQMC algorithm, with the exception that new generated determinants will be restricted to the one accessible by applying \hat{H}_0 on the current wave-function. This will effectively restrict the Hilbert space to the CAS space, and correspond to freezing the core and virtual orbitals. The computation of $|\Psi_1\rangle$ as defined in Eq.15 is less straightforward. Indeed in FCIQMC we do not have access to a proper description of the zeroth order wavefunction. It is thus not possible to compute $|\Psi_1\rangle$ by using a projection approach. As an alternative we also decide to sample $|\Psi_1\rangle$ stochastically. However as opposed to $|\Psi_0\rangle$ the perturbation wavefunctions are not solution of an ITDSE. We introduced the following hierarchy of differential equation (DE).

$$\frac{\partial |\Psi_m\rangle}{\partial \tau} = -\left(\hat{H}_0 - E_0\right) |\Psi_m\rangle - QV \left(|\Psi_{m-1}\rangle - \sum_{k=1}^{m-1} E_k |\Psi_{m-k}\rangle \right). \quad (17)$$

If the left hand side of Eq.17 cancels out, we recover the expression of Eq.7 for $|\Psi_m\rangle$, in other words the successive correction wavefunctions are stationary solution of these DE. Note that propagating this

equation is actually going to reach a steady state, that is equal to $|\Psi_m\rangle$ because the $(\hat{H}_0 - E_0)$ matrix is positive definite. In particular the DE for the first order perturbation is

$$\frac{\partial |\Psi_1\rangle}{\partial \tau} = -(\hat{H}_0 - E_0)|\Psi_1\rangle - V|\Psi_0\rangle. \quad (18)$$

This equation is similar to the one used for $|\Psi_0\rangle$, with the addition of a source term due to the second term of the right hand side. In the next section we will show how the simultaneous solving of Eq.16 and Eq.18 has been implemented in the NECI program.

III. IMPLEMENTATION

To simultaneously sample the zeroth and first order wavefunctions, we use the multi-replica technique[41]. A first replica, labeled 0, is sampling the 0 order wavefunctions by propagating the ITDSE of Eq.16 while another one, labeled 1, is sampling the first order perturbation. We first start with a small amount of walkers on a reference determinant, typically the Hartree-Fock one, on replica 0 and no walkers on replica 1.

In replica 0 new walkers are spawned by applying one and two electrons operators that belong to \hat{H}_0 , thus only determinants that belong to the CAS space are generated. Because the sampling of $|\Psi_0\rangle$ is equivalent to a standard FCIQMC sampling in a CASCI, we can use all the optimizations and approximations that have been introduced in previous publications such as initiator approximation[20, 42] or the semi-stochastic approximation[43].

Once the population on replica 0 is equilibrated we start to sample $|\Psi_1\rangle$. This equilibration of the zeroth order wavefunction can be monitored by looking at the variational energy for this replica and checking that it corresponds to the CASCI energy. At this point we attempt multiple spawning from each walker of replica 0. In addition to the excitation that belongs to \hat{H}_0 we also generate excitation belonging to \hat{V} . This thus generates determinants that belong to the external space, \mathcal{H}_1 . The spawning probability of those $|\Psi_0\rangle$ to $|\Psi_1\rangle$ walkers follows the expression of Eq.14. However those walkers are spawned on replica 1 instead of replica 0.

Replica 1 that was initially empty starts getting populated, the walkers coming from replica 0 correspond to the source term in Eq.18.

We emphasize the fact that in replica 0 the population dynamic is not modified by this extra excitation step, and that by construction the first and zeroth order wavefunctions are orthogonal to each other, this obviates the use of orthogonalization techniques that will be required for the higher order perturbation. The walkers on replica 1 are also subjected to a cloning/dying step and a spawning step at each iteration. For the dying step the applied operator is $(\hat{H}_0 - E_0)$, where the E_0 is the

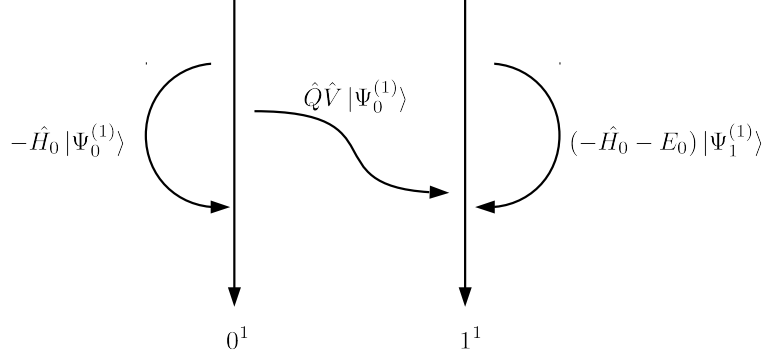


Figure 1: Schematized description of an iteration update of the two replicas. On the left the 0 order wave function, sampled in replica 0 is updated by applying the $\left(\mathbb{1} - \Delta\tau \left(\hat{H}_0 - S\mathbb{1}\right)\right)$ operator. On the right the first order perturbation, in replica 1, is updated by applying the $\left(\mathbb{1} - \Delta\tau \left(\hat{H}_0 - E_0\mathbb{1}\right)\right)$ operator to $|\Psi_1\rangle$ and adding walkers spawned by applying \hat{V} onto $|\Psi_0\rangle$.

projected energy in replica 0. For the spawning step, the applied excitation operator belongs to \hat{H}_0 . The simultaneous sampling of the two wavefunctions is schematized in Fig.1.

Note that with this implementation the timesteps used for $0 \rightarrow 0$, $0 \rightarrow 1$ and $1 \rightarrow 1$ spawning steps, and the 0 and 1 cloning and dying steps are identical. It is the one that has been optimized for replica 0. In practice this timestep is chosen to ensure that the spawning probability in replica 0 is small enough to ensure that the spawning probability is not much bigger than 1 for all $|D_i\rangle$. As can be seen of Eq.14 the spawning probability is inversely proportional to the generation probability p_{gen} to attempt a spawning on $|D_j\rangle$ from $|D_i\rangle$. Because in interesting systems the size of the external space is expected to be much bigger than the one of the active space the probability of generating $|D_i\rangle$, $|D_j\rangle$ pairs is much smaller in case of $0 \rightarrow 1$ and $1 \rightarrow 1$ than in case of $0 \rightarrow 0$ excitation. As a consequence the spawning probabilities of those excitation will be much bigger than $0 \rightarrow 0$ spawning probability for the same timestep, and a single walker would give birth to multiple walkers; such an event is called blooming. In other words, the time step should be smaller in the excitations involving the response functions. This problem is dealt with as follow, the overall time step of the simulation is set by the $0 \rightarrow 0$ dynamic. The first time blooming occurs during the $0 \rightarrow 1$ spawning step we keep track of the biggest bloom and compute the first integer bigger than this, n_{01} . During further step, the time step value will be divided by this n_{01} factor for $0 \rightarrow 1$ spawning, this ensures to have a spawning probability not much bigger than 1. To keep the overall dynamic at the same timestep we have to do n_{01} spawning attempt to $|1\rangle$ for each walkers on $|0\rangle$. We use a similar procedure for the $1 \rightarrow 1$ spawning defining a n_{11} timestep scaling factor. Those n_{01} and n_{11} factors are updated along the run, this prevent any explosion of the replica 1 population.

With this implementation there is no way to control the total number of walkers on $|\Psi_1\rangle$ because

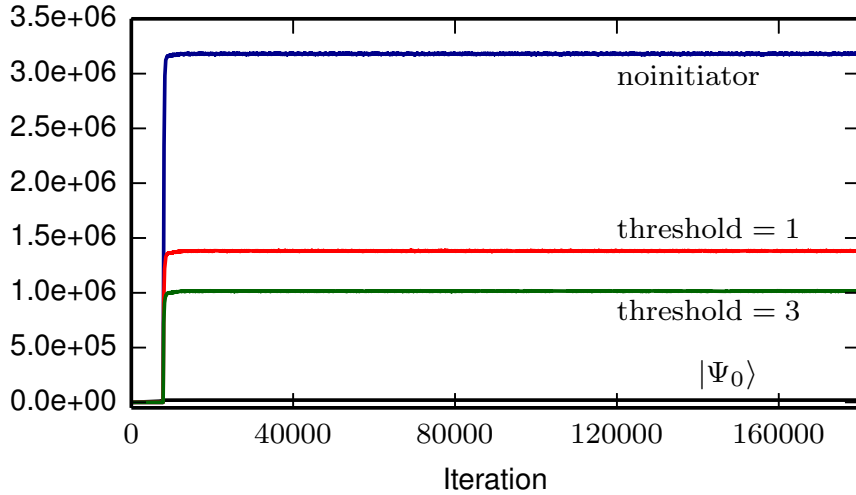


Figure 2: Number of walkers in replica 0 (black), and in replica 1 with an initiator approximation of 3 (green), 1 (red) and no initiator approximation (blue) for the C_2 molecule in the cc-pVQZ basis set.

there is no analogue of the shift control parameter that is used for replica 0. After a few steps the total number of walkers in replica 1 reaches a plateau that is dependent of the system. As a first attempt to control the value of that plateau the initiator approximation has been implemented for walkers that belong to replica 1. We allow the initiator threshold to be different in replica 0 and replica 1. As an illustration we studied the carbon dimer molecule with the cc-pVQZ basis set[44], with \mathcal{H}_0 corresponding to a CAS (8,8) which is the valence space of the molecule. We use a initiator threshold of 3 and a targeted number of walker of 50k for replica 0. In Fig.2 we present the evolution of the number of walker on replica 0 and on replica 1, the different calculations have been run with a initiator criterion of 3, 1 and no initiator approximation on replica 1.

Looking at Fig.2 it can be seen that with no initiator approximation on $|\Psi_1\rangle$, the number of walkers grows to more than 60 times the number of walkers in the reference. When using the most moderate initiator criteria of 1, this number is already reduced by more than a factor of 2, it can be further decreased by increasing the initiator threshold. However going from a threshold of 1 to 3 only reduced the total number of walkers by roughly 30 %. This is still not satisfying since there is no way to know *a priori* what the number of walkers on $|\Psi_1\rangle$ is going to be. The cost of the calculation cannot be known before running it; for this reason we describe hereafter a way to control the population on replica 1 independently from the initiator threshold.

In Eq.15 it can be seen that $|\Psi_1\rangle$ scales linearly with the perturbation \hat{V} . We start by making the assumption that the number of walkers on $|\Psi_1\rangle$ also depends linearly on the perturbation and thus scale down the perturbation by a real prefactor α , that is typically small. This is done in practice by multiplying the matrix elements H_{ij} by this factor when the spawning probability from a determinant

$|D_i\rangle$ on \mathcal{H}_0 to a determinant $|D_i\rangle$ on \mathcal{H}_1 is computed. This allows us to tune more easily the total number of walkers on replica, also the relation between the value of the plateau and the α remains unknown. The number of walkers does not strictly scales linearly with the value of α because of the offdiagonal spawning in replica 1 and because of the initiator approximation. To circumvent this problem we implemented a dynamic updating of α in order to reach a target number of walker on $|\Psi_1\rangle$. The simulation is started with a small α , typically 10^{-2} ; after the a few thousand step of equilibration, if the number of walkers on replica 1 is not included within a 3% treshold of the target, we update alpha to a new value α' ,

$$\alpha' = \alpha \left(\gamma + (1 - \gamma) \frac{N_t}{N} \right) \quad (19)$$

, where N_t is the target number of walkers, N is the current number of walkers, and γ is a dumping parameter to prevent too drastic a change of α . We use typically $\gamma = 0.5$.

Having implemented this way of controlling the population in replica 1 we now will study the influence of the number of walkers used to sample the response function on the second order energy. The second order correction energy is expressed as

$$E_2 = \frac{\langle \Psi_0 | \hat{V} | \Psi_1 \rangle}{\langle \Psi_0 | \Psi_0 \rangle} = \frac{\sum_{i \in \mathcal{H}_0} \sum_{j \in \mathcal{H}_1} c_i c_j \langle D_i | \hat{V} | D_j \rangle}{\sum_{i \in \mathcal{H}_0} c_i^2}. \quad (20)$$

In the FCIQMC framework this can be rewritten as a function of the walker population on each determinants involved. In the LCC perturbation theory $\langle D_i | \hat{V} | D_j \rangle = \langle D_i | \hat{H} | D_j \rangle = H_{ij}$

$$E_2 \approx \langle \tilde{E}_2 \rangle = \frac{\langle \Psi_0 | \hat{V} | \Psi_1 \rangle}{\langle \Psi_0 | \Psi_0 \rangle} = \frac{\langle \sum_{i \in \mathcal{H}_0} \sum_{j \in \mathcal{H}_1} N_i N_j H_{ij} \rangle}{\langle \sum_{i \in \mathcal{H}_0} N_i^2 \rangle} = \frac{\langle \sum_{j \in \mathcal{H}_1} N_j [\sum_{i \in \mathcal{H}_0} N_i H_{ij}] \rangle}{\langle \sum_{i \in \mathcal{H}_0} N_i^2 \rangle}, \quad (21)$$

where $\langle . \rangle$ denotes the average over multiple timesteps and \tilde{E}_2 is second order energy computed with the instantaneous population in the two replica.

In principle we could exactly compute \tilde{E}_2 by finding all the connections between determinants belonging to \mathcal{H}_0 and \mathcal{H}_1 but this is prohibitive. Instead the connection are found through the spawning process, this strategy has already been used for the calculation of reduced density matrices[41]. When a successful spawning from a determinant $|D_i\rangle$ in replica 0 to a determinant $|D_j\rangle$ in replica 1 occurs, the product of the matrix element and of the number of walkers on $|D_i\rangle$, $N_i H_{ij}$ is communicated along the spawned walker to the processor holding the child determinant $|D_j\rangle$. Once all the spawning attempt have been done, the processor that keep track of the $|D_j\rangle$ walkers population will also contains all the $N_i H_{ij}$ contribution from all the determinants that spawned to $|D_j\rangle$ this iteration. This strategy does not cause any noticeable increase of the computational cost. As the contribution of a $|D_i\rangle, |D_j\rangle$ pair of determinants to the E_2 energy is only taken into account when a successful spawning step is actually happening this contribution should be rescaled by the normalized probability of spawning at

least one child (of any weight) onto $|D_j\rangle$ from $|D_i\rangle$ during the current iteration. More details on how to compute this probability can be found in [41]. To avoid double counting of $C_i C_j$ contribution, it is necessary to carefully check for the rare but still possible case of multiple spawning from the same determinant $|D_i\rangle$ to the same determinant $|D_j\rangle$. Thus the $N_i H_{ij}$ contribution is only communicated to the processor holding $|D_j\rangle$ in the first occurrence of such an event.

An additional difficulty comes from the computation of the denominator of Eq.21. If we just take the square of the number of walkers on a determinant in replica 0, a bias is introduced because $\langle N_i \rangle \langle N_i \rangle \neq \langle N_i^2 \rangle$, the square of the instantaneous value of the walker population is correlated. In the RDM computation this particular problem of the normalization was circumvented by using the trace condition to normalize RDM *a posteriori*. There is no such relation to unbiased the measure of the norm of $|\Psi_0\rangle$ here, instead we use a replica trick. This requires the addition a third replica, labeled 0' that is also sampling the zero order wavefunction. Because the two replica 0 and 0' are uncorrelated we can have an unbiased measure of the norm of $|\Psi_0\rangle$ by $\langle \Psi_0^{(0)} | \Psi_0^{(0')} \rangle = \sum_{i \in \mathcal{H}_0} N_i^{(0)} N_i^{(0')}$. Finally, it is necessary to rescale the $|\Psi_1\rangle$ function by the α factor to compute the E_2 energy.

With this efficient way of computing the second order energy, we can go back to the example of Fig. 2 and examine how the estimation of E_2 converge with the initiator approximation.

In Fig.3 we plotted the second order energy for the Carbon dimer with the cc-pVQZ basis computed with the QMC-LCC framework. The black line correspond to the simulations presented in fig.2 i.e. 50k walkers and a initiator threshold of 3 in replica 0 and an initiator criterion of 1, 3 and no initiator approximation for the first order response function. It can be seen that without approximation, the computed energy is in agreement with the one computed using MPS-LCC, shown in blue for reference. When the initiator approximation is used, we obtain a second order energy that is slightly higher than the correct one. However the estimation remains quite good since the error in the energy is lower than $1 mE_H$ with the two criteria used here.

In Fig.3, the blue curve is obtained with an initiator threshold of 1 and different values of α to constrained the number of walkers. The most interesting feature is that, for the same number of walkers on $|\Psi_1\rangle$ it seems to be better to actually set a smaller initiator threshold and to use the α trick than increasing the initiator threshold. For instance for roughly 1M walkers the value obtained with initiator threshold of 1 and an α value of ≈ 0.73 is $0.16 mE_H$ lower than the one obtained by using an initiator of 3.

In order to improve further the efficiency of our implementation, we notice that the \mathcal{H}_1 space can be split as the orthogonal sum of 8 smaller subspaces corresponding to the 8 subclasses of excitation

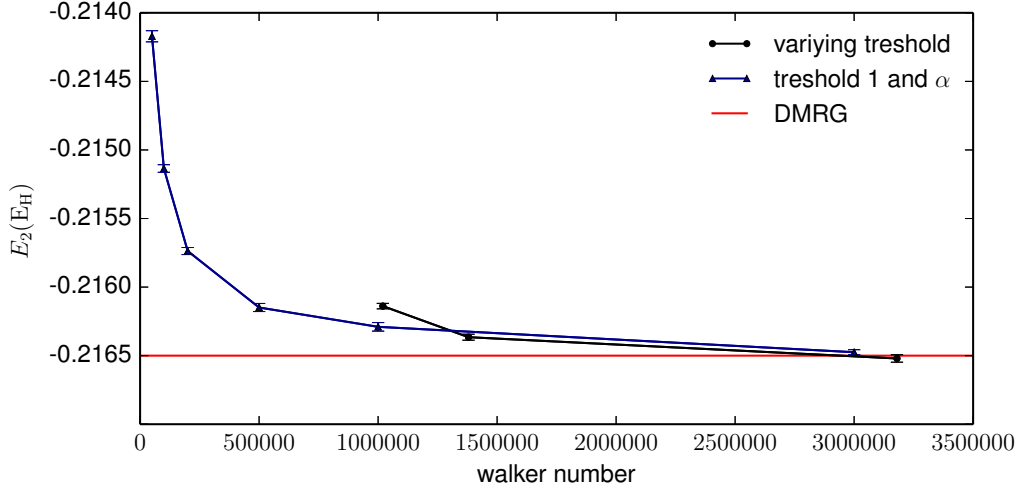


Figure 3: Comparison of the E2 energy with respect to the number of walkers. The black curve is obtained by setting the initiator threshold to different values, respectively 3,1 and no initiator threshold. The blue curve is obtained by setting the initiator threshold to 1 and controlling the number of walkers on $|\Psi_1\rangle$ by using the α controlling parameter. For comparison purpose, the value obtained using DMRG is in red.

defined by Malrieu and collaborators[26, 27].

$$\mathcal{H}_1 = \mathcal{H}_{(-2,2,0)} \oplus \mathcal{H}_{(0,-2,2)} \oplus \mathcal{H}_{(-2,1,1)} \oplus \mathcal{H}_{(-2,0,2)} \oplus \mathcal{H}_{(-1,1,0)} \oplus \mathcal{H}_{(-1,0,1)} \oplus \mathcal{H}_{(-1,-1,2)} \oplus \mathcal{H}_{(0,-1,1)}, \quad (22)$$

where the subscript described the change in number of electrons in the core, active and virtual space with respect the \mathcal{H}_0 determinants. To each of this subspace correspond a subclass of excitation that we can denote $\hat{V}_{(a,b,c)}$, connecting $|\Psi_0\rangle$ to an $\mathcal{H}_{(a,b,c)}$ subspace.

When \hat{H}_0 is applied to a determinant belonging to one of those subspaces, the generated determinants also belong to this subspace. This means that during the dynamic, there are no interactions between walkers belonging to 2 different subclasses. Instead of running a single calculation by applying the full \hat{V} to $|\Psi_0\rangle$ it is possible to run 8 independent simpler calculations using each of 8 different classes of excitations $\hat{V}_{(a,b,c)}$ and finally sum up the 8 different second order energies obtained.

IV. RESULTS

The LCC-QMC method proposed here has been applied to several organic molecules. First we look at the behavior of the method with respect to the size of the inactive space by computing the E_2 energy for the C_2 molecule with different basis set. In each case the active space consist of the valence orbitals of the molecule which correspond to a (8,8) CAS. In addition to the CAS, there are 2 core orbitals

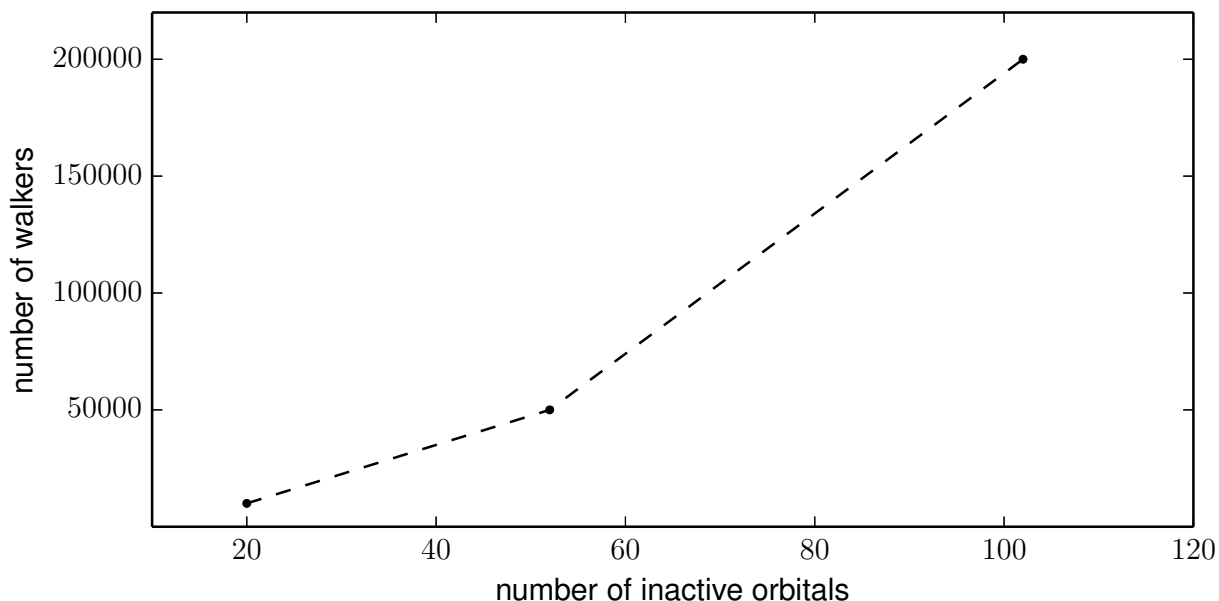


Figure 4: Comparison of the number of walkers that is necessary to use on replica 1 to reach a value within $1 mE_H$ for E_2 with respect to the energy computed by MPS-LCC used as reference. In all case the CAS space is (8,8), the three different basis set used are cc-pVDZ with 20 inactive orbitals, cc-pVTZ with 52 inactive orbitals and cc-pVQZ with 102 inactive orbitals. The dotted line is here for eyes guidance.

and 16, 50 and 100 virtual orbitals respectively in the cc-pVDZ, cc-pVTZ and cc-pVQZ basis sets used here. The CASSCF orbitals have been generated using the Molpro quantum chemistry package[45].

We used 50k walkers and an initiator threshold of 3 in replica 0, this threshold is set to 1 in replica 1. The response wavefunction is sampled by applying the full \hat{V} operator to the $|\Psi_0\rangle$ wavefunction.

To investigate how the cost of the response scale with the size of the inactive space, we increase the number of walkers in replica 1 until the computed value of E_2 agrees within $1 mE_H$ with the same quantity computed deterministically with MPS-LCC using the Block code[46]. The Required number of walkers are represented in Fig.4. From this curve it can be seen that the number of walkers on $|\Psi_1\rangle$ necessary to reach a mE_H precision scales roughly as the square of the number of inactive orbitals.

To test the applicability of the method, we now turn our attention to the computation of benzene singlet-triplet gap.

We used the same geometry than Roos et al[36, 47], i.e. C-C and C-H bond lengths of 1.395 Å and 1.085 Å and an hexagonal symmetry, for the singlet ground state and triplet excited state. The ground state of the benzene molecule is singlet $^1A_{1g}$, and we target the lower excited triplet of symmetry $^3E_{1u}$. We used the Dunning cc-pVDZ basis set, the active space contains 6 electrons and the six valence π orbitals extended with the six second shell π orbitals. We used CASSCF orbitals generated using

Molpro.

For the $|\Psi_0\rangle$ CASCI wavefunction we used an initiator threshold of 3 and 100k walkers. In order to make the calculation more tractable, we run 8 subcalculations corresponding to each of the orthogonal classes of excitation $\hat{V}_{(a,b,c)}$. The initiator threshold on $|\Psi_1\rangle$ is set to 1.5. We start by using 100k walkers for each first order function, but to test the applicability of the methods, this number has been increased until the second order energy agrees with a precision of 1 mE_h with the one predicted deterministically by LCC-MPS. The values obtained for the second order energies for the different classes and the number of walkers it was necessary to use on $|\Psi_1\rangle$ to obtain these values are specified in Table.I. The $(-1, 1, 0)$ class is not contributing since it has no overlap with the $|\Psi_0\rangle$ wavefunction for spatial orbital symmetry reasons.

We can see by looking at this table that the classes of excitations are not equivalent in terms of the number of walkers necessary to converge. The classes involving virtual orbitals require more walkers to reduce the initiator error. In particular the $(-2, 0, 2)$ class is particularly difficult, this is understandable since in this case this the class containing the biggest number of determinants. Considering the Singlet-Triplet gap, the CASSCF value is equals to 4.97 eV while it is reduced to 4.88 eV when the LCC correction is used, the experimental value determined by electron-impact spectroscopy is 4.76 eV[48]. We also computed the singlet-triplet gap with CASPT2 where internal contraction is used only for the subspaces requiring at most the knowledge of the second order reduced density matrix while the other subspaces are left uncontracted and NEVPT2 strongly contracted using Molpro, we obtained an difference of 4.65 eV and 5.03 eV respectively.

We then turned our attention to the computation of the triplet-singlet gap in the m-xylylene diradical. The study of radical organic species and the prediction of their spin properties is of interest because of potential application in the developement of molecule-based magnetic materials[49].

The key parameter for such application is the triplet-singlet gap, thus some effort have been done in order to tune this parameter. Along the experiments it is interesting to predict the value of this singlet triplet gap to guide the synthesis of new promising molecules.

Among the different organic diradicals, the m-xylylene has been used quite often as a benchmark system since it is rather stable and quite well characterized experimentally. The molecule belong to C_{2v} point group, the ground state has been proved to be a triplet by using EPR[50], and it has electronic state 3B_2 . It has been shown by using NIPES that the lowest lying excited state is 1A_1 [51]. This system has been quite extensively studied numerically. For instance Mañeru *et al*[52] carried an extensive study with DFT and several wavefunction methods exploring the effect of the geometries and of the basis set on the predicted gap. They found that the choice of the basis set have low influence on the predicted gap, however as expected the value of the gap is highly dependent of the

Class	Singlet		Triplet	
	E_2	Walkers number	E_2	Walkers number
$(-2, 2, 0)$	-0.02025	500 000	-0.01167	500 000
$(0, -2, 2)$	-0.01716	500 000	-0.01681	500 000
$(-2, 1, 1)$	-0.03330	1000 000	-0.03558	1000 000
$(-2, 0, 2)$	-0.46096	10 000 000	-0.46706	10 000 000
$(-1, 1, 0)$	0.00000	N/A	0.00000	N/A
$(-1, 0, 1)$	-0.21671	1000 000	-0.20215	1000 000
$(-1, -1, 2)$	-0.08402	1000 000	-0.09871	5000 000
$(0, -1, 1)$	-0.00669	500 000	-0.00680	500 000
CASCI	-230.8070		-230.6244	
CASCI+LCC	-231.6461		-231.4667	

Table I: E_2 energies in E_H for the 8 different classes of excitation for the singlet (left) and triplet (right) predicted by MPS-LCC and LCC-QMC. Next to each value is the number of walkers necessary to get this number with the method described in this article. The $(-1, 1, 0)$ class has no overlap with $|\Psi_0\rangle$ because the excitation are forbidden by symmetry.

choice of the functional. On the other hand, all the different wavefunctions methods they used tends to overestimate the singlet triplet gap. Even with more sophisticated basis set their predicted value of the singlet-triplet gap remains quite close to the value of 4092 cm^{-1} previously predicted by Hrovat et al using a CAS-PT2 calculation with a CAS (8,8) and 6-31g* basis set[53].

As the authors states “the difficulty of the wave function-based methods in describing the triplet–singlet gap arises quite unequivocally from dynamical correlation”. They proposed to extend the CAS space as a way to recover the dynamical correlation is improperly taken account, however as they correctly stated this will make the calculation computationally challenging.

As an alternative we decided to examine how a better treatment of dynamical correlation by using the LCC multireference perturbation theory with the same CAS space would improve the prediction of the singlet-triplet gap. We used the equilibrium geometries optimized at the CASSCF(8,8)/6-311++g** level given in supporting information of Mañeru et al, for both the singlet and the triplet. We used the same basis set to run a CASSCF(8,8) calculation to generate the 2 and 4 indexes integrals files with the Molpro quantum package. We computed the E_2 energies for the 8 classes with MPS-LCC using internal contraction as a reference. We start by using a procedure similar to the one used with the benzene molecule which is setting the iniator criteria to 1.5 in for the first order response functions and progressively increasing the number of walker sample it. However except for the two classes $(-2, 2, 0)$ and $(0, -1, 1)$ this procedure has failed, since the obtained value where much higher than the one predicted by MPS-LCC. We thus decided to run calculation with a initiator criterium of 1

Class	Singlet		Triplet	
	E_2	Walkers number	E_2	Walkers number
$(-2, 2, 0)$	-0.00625	500 000	-0.00630	500 000
$(0, -2, 2)$	-0.0324	36 000 000	-0.0319	650 000 000
$(-2, 1, 1)$	-0.0438	50 000 000	-0.0437	40 000 000
$(-2, 0, 2)$	-0.46096 *		-0.46706*	
$(-1, 1, 0)$	0.00000	N/A	0.00000	N/A
$(-1, 0, 1)$	-0.182	13 000 000	-0.180	88 000 000
$(-1, -1, 2)$	-0.08402*		-0.09871*	
$(0, -1, 1)$	-0.0168	500 000	-0.0152	500 000
CASCI	-230.8070		-230.6244	
CASCI+LCC	-231.6461		-231.4667	

Table II: E_2 energies in E_H for the 8 different classes of excitations for the singlet (left) and triplet (right) of the m-xylylene diradical predicted by LCC-QMC. Next to each value is the number of walkers necessary to get this number with the method described in this article. It has not been possible to make the classes $(-2, 0, 2)$ and $(-1, -1, 2)$ converge with LCC-QMC, the E_2 labeled (*) are obtained with MPS-LCC using internal contraction

and no control of the population i.e an α factor of 1. The number obtained and the number of walkers reached on replica 1 are given in table II. With this procedure the obtained number are in agreement with the MPS-LCC results, there are usually a bit more negative since this is a fully uncontracted technique. However this approach is not really practical since the number of walkers reached on $|\Psi_1\rangle$ is generally huge, making the calculation extremely costly. Moreover in the case of the $(-2, 0, 2)$ and $(-1, -1, 2)$ classes it was not possible to do the calculation.

If we retains the values predicted by MPS-LCC for the classes where QMC-LCC cannot be converged, we found a value of 3440 cm^{-1} for the triplet-singlet gap of m-xylylene, a value that agrees well with the experimental value $3358 \pm 70 \text{ cm}^{-1}$. This shows that using the MRLCC perturbation theory improves a lot the description of the dynamical correlation for this system.

Thus as Mañeru *et al* stated, the problem in the description of the m-xylylene was a proper treatment of the dynamical correlation. However this problem is still out of reach for the fully uncontracted LCC-QMC, this indicated clearly that to make LCC-QMC practical it would be necessary to implement some internal contraction treatment of the perturbation classes. However the fact that the most easy classes in the LCC-QMC approach is the one that would require the 3 and 4-RDM to be computed with internal contraction is encouraging.

V. CONCLUSIONS

In this article we described a way to do CASCI+MRLCC calculation that is completely stochastic. The zero order wavefunction is computed with the FCIQMC approach, *i.e* a population of signed walkers which is evolving thanks to a series of stochastic rules, is sampling the CASCI wavefunction and solving the zero order Hamiltonian eigenproblem. Simultaneously, a second population of walkers submitted to a different set of stochastic process is sampling the first order wavefunction, by finding the steady state of a differential equation. The first order wavefunction being a function of the zero order one, we include a source term in the population dynamic sampling $|\Psi_1\rangle$ that depend of the population in $|\Psi_0\rangle$, this requires spawning from one replica to the other, and this is the main originality of the proposed algorithm. We presented different strategy to make the use of this technique practicable, some are directly adapted from the strategies proposed for FCIQMC, such as the initiator approximation and the semi-stochastic approach. We also proposed to scale the source term as a way to control the population in the replica sampling the first order wavefunction. Because the size of the CASCI and of the perturbation are quite different it was also necessary to have different timesteps for the 0 to 0, 0 to 1 and 1 to 1 spawning steps. We illustrate the possibility of the proposed methods on several applications. First we shows that this approach allows to recover the results computed by the deterministic MPS-LCC method on the case of the C_2 molecule. This calculation shows that the cost of the calculation, linked to the number of walkers, scale quadratically with the number of inactive orbitals.

The computation of the singlet-triplet gap of the challenging m-xylylene diradical with the proposed method confirmed the better performance of the MRLCC perturbation theory with respect to the CASPT2 technique. It has been possible to obtain a value of this gap that is in good agreement with the experimental results while the CASPT2 approach using the same active space overestimate this quantity by 20%. However it has also been demonstrated that the proposed algorithm is still not adapted to study such complicated systems since the class of excitation involving two cores electrons and to virtuals holes cannot be computed for the m-xylylene molecule. This problem has been circumvent by computing those classes by using the contracted MPS-LCC technique. This point out the necessity to develop a similar contracted approach with our fully stochastic procedure, this is currently under investigation.

The LCC-QMC approach proposed here would allow to treat both large active space and large inactive space; this associated with the recently demonstrated possibility of using FCIQMC-CASSCF[54] will allow to study systems that are currently out of reach.

Acknowledgments

The calculations made use of the facilities of the Max Planck Society's Rechenzentrum Garching. We are grateful to Pr. Illas and Dr. Daniel Reta for providing us their equilibrium geometries for the m-xylylene molecule.

-
- [1] H. Beinert, R. H. Holm, and E. MEnck, *Science* **277**, 653 (1997), ISSN 0036-8075, 1095-9203, URL <http://science.sciencemag.org.ezproxy.fkf.mpg.de/content/277/5326/653>.
 - [2] F. Neese, *JBIC Journal of Biological Inorganic Chemistry* **11**, 702 (2006), ISSN 0949-8257, 1432-1327, URL <http://link.springer.com/article/10.1007/s00775-006-0138-1>.
 - [3] P. Siegbahn, A. Heiberg, B. Roos, and B. Levy, *Physica Scripta* **21**, 323 (1980), ISSN 1402-4896, URL <http://iopscience.iop.org/1402-4896/21/3-4/014>.
 - [4] B. O. Roos, P. R. Taylor, and P. E. M. Siegbahn, *Chemical Physics* **48**, 157 (1980), ISSN 0301-0104, URL <http://www.sciencedirect.com/science/article/pii/0301010480800450>.
 - [5] P. E. M. Siegbahn, J. Almlf, A. Heiberg, and B. O. Roos, *The Journal of Chemical Physics* **74**, 2384 (1981), ISSN 0021-9606, 1089-7690, URL <http://scitation.aip.org.ezproxy.fkf.mpg.de/content/aip/journal/jcp/74/4/10.1063/1.441359>.
 - [6] J. Olsen, B. O. Roos, P. Jrgensen, and H. J. A. Jensen, *The Journal of Chemical Physics* **89**, 2185 (1988), ISSN 0021-9606, 1089-7690, URL <http://scitation.aip.org.ezproxy.fkf.mpg.de/content/aip/journal/jcp/89/4/10.1063/1.455063>.
 - [7] P. A. Malmqvist, A. Rendell, and B. O. Roos, *The Journal of Physical Chemistry* **94**, 5477 (1990), ISSN 0022-3654, URL <http://dx.doi.org/10.1021/j100377a011>.
 - [8] D. Ma, G. L. Manni, and L. Gagliardi, *The Journal of Chemical Physics* **135**, 044128 (2011), ISSN 0021-9606, 1089-7690, URL <http://scitation.aip.org/content/aip/journal/jcp/135/4/10.1063/1.3611401>.
 - [9] G. Li Manni, D. Ma, F. Aquilante, J. Olsen, and L. Gagliardi, *Journal of Chemical Theory and Computation* **9**, 3375 (2013), ISSN 1549-9618, URL <http://dx.doi.org/10.1021/ct400046n>.
 - [10] S. R. White, *Physical Review Letters* **69**, 2863 (1992), URL <http://link.aps.org/doi/10.1103/PhysRevLett.69.2863>.
 - [11] S. R. White, *Physical Review B* **48**, 10345 (1993), URL <http://link.aps.org/doi/10.1103/PhysRevB.48.10345>.
 - [12] S. R. White and R. L. Martin, *The Journal of Chemical Physics* **110**, 4127 (1999), ISSN 0021-9606, 1089-7690, URL <http://scitation.aip.org.ezproxy.fkf.mpg.de/content/aip/journal/jcp/110/9/10.1063/1.478295>.
 - [13] S. Wouters, W. Poelmans, P. W. Ayers, and D. Van Neck, *Computer Physics Communications* **185**, 1501 (2014), ISSN 0010-4655, URL <http://www.sciencedirect.com/science/article/pii/S0010465514000496>.

- [14] D. Zgid and M. Nooijen, The Journal of Chemical Physics **128**, 014107 (2008), ISSN 0021-9606, 1089-7690, URL <http://scitation.aip.org.ezproxy.fkf.mpg.de/content/aip/journal/jcp/128/1/10.1063/1.2814150>.
- [15] G. Moritz, B. A. Hess, and M. Reiher, The Journal of Chemical Physics **122**, 024107 (2005), ISSN 0021-9606, 1089-7690, URL <http://scitation.aip.org.ezproxy.fkf.mpg.de/content/aip/journal/jcp/122/2/10.1063/1.1824891>.
- [16] O. Legeza and J. Sólyom, Physical Review B **68**, 195116 (2003), URL <http://link.aps.org/doi/10.1103/PhysRevB.68.195116>.
- [17] Y. Kurashige and T. Yanai, The Journal of Chemical Physics **130**, 234114 (2009), ISSN 0021-9606, 1089-7690, URL <http://scitation.aip.org.ezproxy.fkf.mpg.de/content/aip/journal/jcp/130/23/10.1063/1.3152576>.
- [18] R. Olivares-Amaya, W. Hu, N. Nakatani, S. Sharma, J. Yang, and G. K.-L. Chan, The Journal of Chemical Physics **142**, 034102 (2015), ISSN 0021-9606, 1089-7690, URL <http://scitation.aip.org/content/aip/journal/jcp/142/3/10.1063/1.4905329>.
- [19] G. H. Booth, A. J. W. Thom, and A. Alavi, The Journal of Chemical Physics **131**, 054106 (2009), ISSN 0021-9606, 1089-7690, URL <http://scitation.aip.org/content/aip/journal/jcp/131/5/10.1063/1.3193710>.
- [20] D. Cleland, G. H. Booth, C. Overy, and A. Alavi, Journal of Chemical Theory and Computation **8**, 4138 (2012), ISSN 1549-9618, URL <http://dx.doi.org/10.1021/ct300504f>.
- [21] G. H. Booth, S. D. Smart, and A. Alavi, arXiv :1305.6981 [cond-mat, physics :physics] (2013), arXiv :1305.6981, URL <http://arxiv.org/abs/1305.6981>.
- [22] Nature **493**, 365 (2013), ISSN 0028-0836, URL <http://www.nature.com/nature/journal/v493/n7432/abs/nature11770.html>.
- [23] F. R. Petruzielo, A. A. Holmes, H. J. Changlani, M. P. Nightingale, and C. J. Umrigar, Physical Review Letters **109**, 230201 (2012), URL <http://link.aps.org/doi/10.1103/PhysRevLett.109.230201>.
- [24] K. Andersson, P. A. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski, The Journal of Physical Chemistry **94**, 5483 (1990), ISSN 0022-3654, URL <http://dx.doi.org/10.1021/j100377a012>.
- [25] K. Andersson, P.-A. Malmqvist, and B. O. Roos, The Journal of Chemical Physics **96**, 1218 (1992), ISSN 0021-9606, 1089-7690, URL <http://scitation.aip.org.ezproxy.fkf.mpg.de/content/aip/journal/jcp/96/2/10.1063/1.462209>.
- [26] C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger, and J.-P. Malrieu, The Journal of Chemical Physics **114**, 10252 (2001), ISSN 0021-9606, 1089-7690, URL <http://scitation.aip.org/content/aip/journal/jcp/114/23/10.1063/1.1361246>.
- [27] C. Angeli, R. Cimiraglia, and J.-P. Malrieu, The Journal of Chemical Physics **117**, 9138 (2002), ISSN 0021-9606, 1089-7690, URL <http://scitation.aip.org/content/aip/journal/jcp/117/20/10.1063/1.1515317>.
- [28] H.-J. Werner and P. J. Knowles, J. Chem. Phys. **89**, 5803 (1988), URL <http://dx.doi.org/10.1063/1.455556>.
- [29] S. Sharma and A. Alavi, The Journal of Chemical Physics **143**, 102815 (2015), ISSN 0021-9606, 1089-7690,

URL <http://scitation.aip.org/content/aip/journal/jcp/143/10/10.1063/1.4928643>.

- [30] D. Ma, G. Li Manni, J. Olsen, and L. Gagliardi, *Journal of Chemical Theory and Computation* **12**, 3208 (2016), ISSN 1549-9618, URL <http://dx.doi.org/10.1021/acs.jctc.6b00382>.
- [31] S. Y. Willow, K. S. Kim, and S. Hirata, *The Journal of Chemical Physics* **137**, 204122 (2012), ISSN 0021-9606, 1089-7690, URL <http://scitation.aip.org.ezproxy.fkf.mpg.de/content/aip/journal/jcp/137/20/10.1063/1.4768697>.
- [32] S. Y. Willow, M. R. Hermes, K. S. Kim, and S. Hirata, *Journal of Chemical Theory and Computation* **9**, 4396 (2013), ISSN 1549-9618, URL <http://dx.doi.org/10.1021/ct400557z>.
- [33] S. Y. Willow, K. S. Kim, and S. Hirata, *The Journal of Chemical Physics* **138**, 164111 (2013), ISSN 0021-9606, 1089-7690, URL <http://scitation.aip.org.ezproxy.fkf.mpg.de/content/aip/journal/jcp/138/16/10.1063/1.4801862>.
- [34] A. J. W. Thom and A. Alavi, *Physical Review Letters* **99**, 143001 (2007), URL <http://link.aps.org/doi/10.1103/PhysRevLett.99.143001>.
- [35] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular electronic-structure theory* (John Wiley & Sons, Chichester, 2000).
- [36] J. M. O. Matos, B. O. Roos, and P.-A. Malmqvist, *The Journal of Chemical Physics* **86**, 1458 (1987), ISSN 0021-9606, 1089-7690, URL <http://scitation.aip.org.ezproxy.fkf.mpg.de/content/aip/journal/jcp/86/3/10.1063/1.452235>.
- [37] K. G. Dyall, *The Journal of Chemical Physics* **102**, 4909 (1995), ISSN 0021-9606, 1089-7690, URL <http://scitation.aip.org/content/aip/journal/jcp/102/12/10.1063/1.469539>.
- [38] R. F. Fink, *Chemical Physics Letters* **428**, 461 (2006), ISSN 0009-2614, URL <http://www.sciencedirect.com/science/article/pii/S0009261406011043>.
- [39] R. F. Fink, *Chemical Physics* **356**, 39 (2009), ISSN 0301-0104, URL <http://www.sciencedirect.com/science/article/pii/S0301010408004527>.
- [40] S. Sharma, G. Jeanmairet, and A. Alavi, *The Journal of Chemical Physics* **144**, 034103 (2016), ISSN 0021-9606, 1089-7690, URL <http://scitation.aip.org.ezproxy.fkf.mpg.de/content/aip/journal/jcp/144/3/10.1063/1.4939752>.
- [41] C. Overy, G. H. Booth, N. S. Blunt, J. J. Shepherd, D. Cleland, and A. Alavi, *The Journal of Chemical Physics* **141**, 244117 (2014), ISSN 0021-9606, 1089-7690, URL <http://scitation.aip.org/content/aip/journal/jcp/141/24/10.1063/1.4904313>.
- [42] D. Cleland, G. H. Booth, and A. Alavi, *The Journal of Chemical Physics* **132**, 041103 (2010), ISSN 0021-9606, 1089-7690, URL <http://scitation.aip.org/content/aip/journal/jcp/132/4/10.1063/1.3302277>.
- [43] N. S. Blunt, S. D. Smart, J. a. F. Kersten, J. S. Spencer, G. H. Booth, and A. Alavi, *The Journal of Chemical Physics* **142**, 184107 (2015), ISSN 0021-9606, 1089-7690, URL <http://scitation.aip.org.ezproxy.fkf.mpg.de/content/aip/journal/jcp/142/18/10.1063/1.4920975>.
- [44] T. H. Dunning, Jr, *The Journal of Chemical Physics* **90**, 1007 (1989), ISSN 0021-9606, 1089-7690, URL <http://scitation.aip.org.ezproxy.fkf.mpg.de/content/aip/journal/jcp/90/2/10.1063/1.456153>.

- [45] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schütz, Wiley Interdisciplinary Reviews : Computational Molecular Science **2**, 242 (2012), ISSN 1759-0884, URL <http://onlinelibrary.wiley.com.ezproxy.fkf.mpg.de/doi/10.1002/wcms.82/abstract>.
- [46] S. Sharma and G. K.-L. Chan, The Journal of Chemical Physics **136**, 124121 (2012), ISSN 0021-9606, 1089-7690, URL <http://scitation.aip.org.ezproxy.fkf.mpg.de/content/aip/journal/jcp/136/12/10.1063/1.3695642>.
- [47] B. O. Roos, K. Andersson, and M. P. Fülcher, Chemical Physics Letters **192**, 5 (1992), ISSN 0009-2614, URL <http://www.sciencedirect.com/science/article/pii/000926149285419B>.
- [48] J. P. Doering, The Journal of Chemical Physics **51**, 2866 (1969), ISSN 0021-9606, 1089-7690, URL <http://scitation.aip.org.ezproxy.fkf.mpg.de/content/aip/journal/jcp/51/7/10.1063/1.1672424>.
- [49] J. S. Miller and A. J. Epstein, Angewandte Chemie International Edition in English **33**, 385 (1994), ISSN 1521-3773, URL <http://onlinelibrary.wiley.com.ezproxy.fkf.mpg.de/doi/10.1002/anie.199403851/abstract>.
- [50] B. B. Wright and M. S. Platz, Journal of the American Chemical Society **105**, 628 (1983), ISSN 0002-7863, URL <http://dx.doi.org/10.1021/ja00341a057>.
- [51] P. G. Wenthold, J. B. Kim, and W. C. Lineberger, Journal of the American Chemical Society **119**, 1354 (1997), ISSN 0002-7863, URL <http://dx.doi.org/10.1021/ja9623830>.
- [52] D. Reta Mañeru, A. K. Pal, I. d. P. R. Moreira, S. N. Datta, and F. Illas, Journal of Chemical Theory and Computation **10**, 335 (2014), ISSN 1549-9618, URL <http://dx.doi.org/10.1021/ct400883m>.
- [53] D. A. Hrovat, M. A. Murcko, P. M. Lahti, and W. T. Borden, Journal of the Chemical Society, Perkin Transactions 2 pp. 1037–1044 (1998), ISSN 1364-5471, URL <http://pubs.rsc.org/en/content/articlelanding/1998/p2/a706409h>.
- [54] G. Li Manni, S. D. Smart, and A. Alavi, Journal of Chemical Theory and Computation (2016), ISSN 1549-9618, URL <http://dx.doi.org/10.1021/acs.jctc.5b01190>.