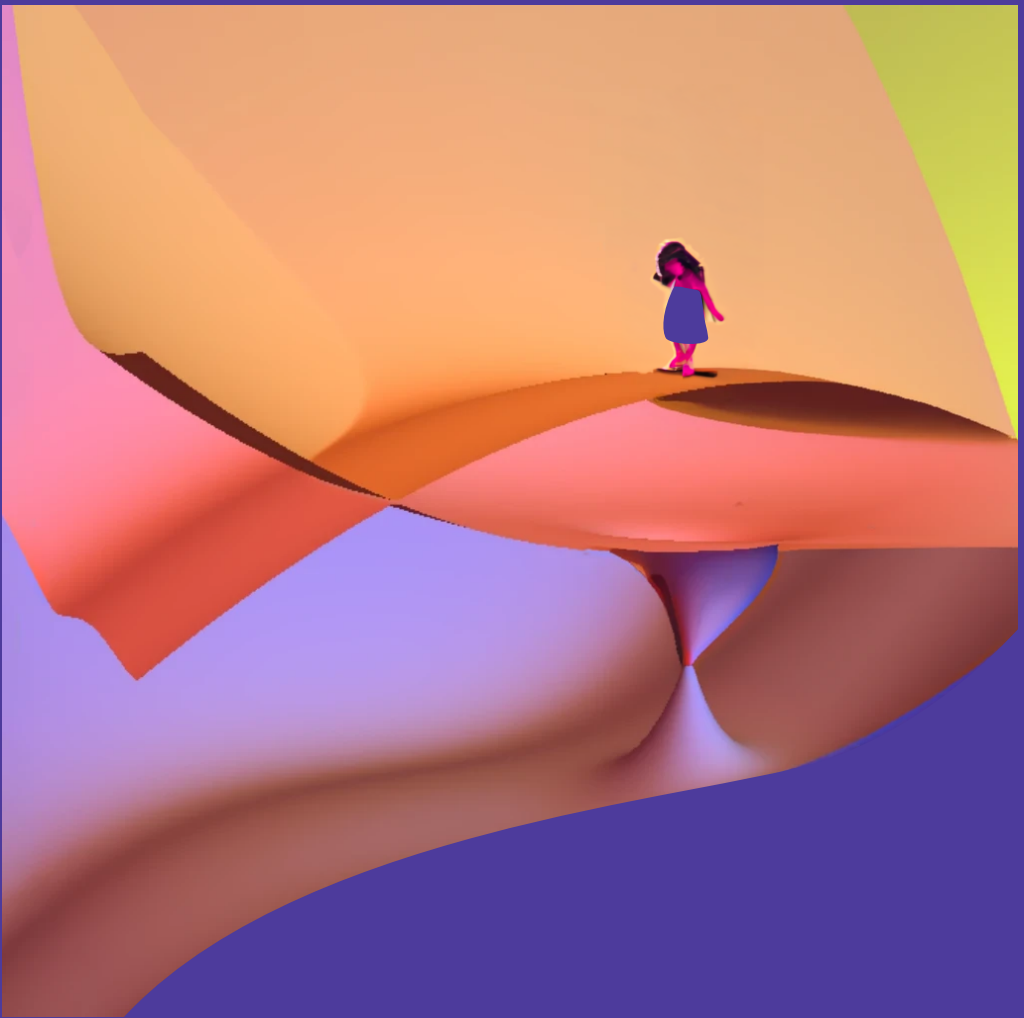# From vertical excitations towards excited-state relaxation: a journey with quantum Monte Carlo

Alice Cuzzocrea

FROM VERTICAL EXCITATIONS TOWARDS EXCITED-STATE RELAXATION: A JOURNEY WITH QUANTUM MONTE CARLO

Alice Cuzzocrea

Graduation Committee:

Prof. dr. J. L. Herek.	University of Twente, Chairman/Secretary
Prof. dr. C. Filippi	University of Twente, Supervisor
Prof. dr. W. J. Briels	University of Twente
Prof. dr. G. H. L. A. Brocks	University of Twente
Dr. ir. B. Ensing	University of Amsterdam
Prof. dr. G. J. Kroes	Leiden University
Prof. dr. J. Neugebauer	University of Münster

FROM VERTICAL EXCITATIONS TOWARDS EXCITED-STATE RELAXATION: A JOURNEY WITH QUANTUM MONTE CARLO

DISSERTATION

to obtain
the degree of doctor at the University of Twente,
on the authority of the rector magnificus,
Prof. dr. ir. A. Veldkamp,
on account of the decision of the graduation committee,
to be publicly defended
on Friday 16th of December 2022 at 14:45

by

Alice Cuzzocrea

born on July 31, 1993
in Melito di Porto Salvo (R.C), Italy

This doctoral dissertation is approved by:

Prof. dr. C. Filippi   Promotor

Contents

Chapter 1

Introduction

The study of photo-physical and photo-chemical processes induced by light-matter interactions attracts great interest because of their role in many natural systems [1,2] and their relevance in a variety of technological applications [3–5]. Examples of such processes are the photo-isomerization of the retinal chromophore [6–8], which initiates vision, and the light-harvesting step that is key to photosynthesis and central to the functioning of photovoltaic devices [9–13]. In this context, understanding photo-driven processes has the double aim of elucidating the fundamental physical mechanisms active in natural photosystems and of tuning the analogous mechanisms in artificial devices for solar-energy technology, photo-medical applications, or bio-imaging. For example, computational studies can aid the design of more efficient and sustainable solar panels [14,15] and directly contribute to pressing problems like the quest for renewable energies.

In the recent past, significant advances have been made in developing theoretical and computational tools to understand light-induced phenomena [16], which require addressing very different scales with different degrees of approximation. Photo-driven processes in complex materials are in fact inherently multi-scale problems as they require the quantum mechanical treatment of the photo-active site, the inclusion of environmental effects, and the account of coupled electron-nuclear dynamics to follow the photo-induced process in time [17–19]. Here, we focus on the quantum mechanical scale, which entails the numerical solution of the Schrödinger equation. The great effort devoted by the electronic-structure community to the computation of ground-state solutions of the same equation has resulted in a number of both reliable and affordable methods which, when combined with a more approximate description of an embedding environment, allow us for example to follow the ground-state evolution of chemically active sites in large protein complexes via hybrid quantum-in-classical simulations [20,21].

The overall picture becomes considerably more complicated when moving to the excited-state description of a system. Not only do we need to adapt our methods to find higher eigenstates of the Hamiltonian but also describe much more intricate "correlated" portions of a potential energy surface (PES) that include, for example, conical intersections with lower-lying states. In this scenario, it is not surprising that a multitude of techniques populates the literature on excited states, each offering a

different degree of compromise between accuracy and computational cost to address larger system sizes [22–25].

In this work, we mainly focus on one family of approaches, namely, real-space quantum Monte Carlo (QMC) methods, which are characterized by the use of Monte Carlo algorithms to solve the Schrödinger equation stochastically [26–30]. On the spectrum of electronic structure methods, they fall more on the accurate than on the cheap side of computations. However, with the increased availability of computational resources, their appeal lies both in the favorable scaling with system size (a mere N^4 with the number of electrons N) and the intrinsic ease of parallelization. Furthermore, while traditionally employed for the computation of ground-state properties (mainly energies), QMC methods have recently undergone significant developments, and their use for the study of excited-state properties appears nowadays not only feasible but, in several cases, even a favorable choice [31–33].

In the following, we will give a short overview of recent advances in QMC, especially those connected with the computation of excited states on which we further build in this thesis. Moreover, we will describe the current open challenges and contextualize the achievements in the broader picture of electronic structure methods. We will focus on the two most widely used flavors of real-space QMC, namely, variational (VMC) and diffusion Monte Carlo (DMC). In the first method, the square of a given trial wave function is sampled to estimate the integrals corresponding to the expectation values of operators on the given wave function. Of course, the results will depend on the choice of wave function and can be improved by variationally optimizing the wave function as well as increasing the complexity of its functional form. In DMC, starting from a given wave function, one stochastically projects it to a better solution of the Schrödinger equation. However, the so-called fixed-node approximation, introduced to deal with fermions, forces the solution to have the same nodal surface as the given wave function. As a consequence, the choice of the starting wave function also matters in DMC.

Wave function form. The most commonly adopted wave function in QMC is the so-called Jastrow-Slater one, where a Jastrow correlation factor $\mathcal{J}$ that accounts for dynamical correlations is multiplied by an expansion of Slater determinants expressed over single-electron orbitals that embody static correlation:

$$\Psi = \mathcal{J} \sum_{i=1}^{N_{\text{det}}} c_i D_i \,, \tag{1.1}$$

where N_{det} is the number of determinants, and D_i and c_i are the determinants and their coefficients, respectively. The Jastrow factor explicitly depends on the interparticle distances and, commonly, has an exponential form. The choice of the determinantal part can be very diverse, and the search for novel functional forms is an open field of research. Possible choices to identify the most important determinantal contributions span from the simple Hartree-Fock (HF) determinant to selected configuration interaction (sCI) expansions. Most recently, neural networks have also been employed in QMC as sophisticated determinantal components with multi-electron

orbitals for ground and excited states [34–38]. Relative to the problem of interest, one will have to compromise accuracy (many determinants or alternative forms) and computational cost (few parameters).

In excited-state calculations with Jastrow-Slater wave functions, the choice of the wave function has two main objectives:

  i) finding the smallest determinantal set that gives accurate results;

  ii) ensure a balanced description of the multiple states involved.

Traditionally, a reasonable compromise was found to be the use of a complete active space (CAS) expansion, which includes all possible excitations of a given number of electrons in a small set of active orbitals. Using a shared active orbital space for all states ensures the balance, and the relevant orbitals are chosen based on the nature of the excitation in the system of interest and, ultimately, on the chemical intuition of the person performing the calculation. Despite the many successes of this protocol [39–45], various limitations remain even after fully optimizing the resulting Jastrow-Slater wave function in QMC. The construction can suffer from a poor chemical choice of the active space resulting in the exclusion of critical determinantal contributions. Furthermore, a complete active space may contain many non-relevant determinants which slow down the QMC computation without improving the results. Finally, given the exponential scaling with the number of orbitals, the size of the active space must be chosen smaller and smaller as the size of the system grows.

Recently, a new protocol has been established that bases the construction of the wave function on the CI perturbatively selected iteratively (CIPSI) method [31–33, 46–50]. With this sCI method, we replace chemical intuition with an automatic selection of relevant determinants based on their second-order perturbation energy contribution, and we obtain a compact wave function that is systematically improvable as we increase the size of the expansion. To balance the multiple states, we ensure that the expansions of the states involved are of similar quality, using as indicator the fact that their CI variances or their second-order perturbation energy contribution (i.e. an estimate of the error with respect to the full CI limit) are matched [32, 33]. With this protocol, as we show in Chapter 4, we can reach high accuracies with relatively compact wave functions making it possible to push the QMC computation of excited states to increasingly large molecular sizes, which are, for instance, challenging for high-level coupled cluster methods.

Optimization of the wave function. After the choice of the form of the wave function, the second fundamental challenge in a QMC calculation is the variational optimization of the trial wave function. The optimization step is the core computation in VMC and also provides the starting wave function for a subsequent DMC calculation. To this aim, one typically minimizes the energy of the system, but for excited states, other variational principles such as variance minimization are possible candidates, as we explore in Chapter 3.

In optimizing the wave function, one aims at reaching convergence to the optimal parameters as quickly as possible. The optimization of the energy requires the com-

putation of the gradients and, possibly, of more complex quantities (e.g. Hessian), which can be used to speed up the convergence. Since all integrals are computed stochastically, the choice of minimization method in QMC must be a compromise between speed and robustness to the noise. For example, the stochastic gradient descent optimizers used in deep learning algorithms are very robust to noise when employed in VMC, but require many iterations to converge since they only use the information of the gradients. On the other hand, in the Newton method [51] or the so-called linear method [52,53], the number of iterations becomes smaller, but longer Monte Carlo runs are needed at each iteration to sample the quantities of interest. Indeed, despite the efforts to formulate these algorithms in terms of QMC estimators characterized by reduced variance, these quantities remain very noisy. The reason is that the computation of the Hessian elements, for instance, includes functions that are rather different from the sampled wave function (e.g. the Hamiltonian acting on the wave function derivatives) and are, therefore, more severely affected by noise. In this thesis, we mainly employ the stochastic reconfiguration (SR) method [54, 55], which is often sufficiently fast and robust to noise since it goes beyond a gradient-only approach, using a poor man's Hessian related to the overlap of the derivatives of the wave function. We illustrate the strengths and limitations of this approach when it comes to the convergence of interatomic forces during the variational optimization in Chapter 6.

Research effort is continuously being devoted to improve existing optimization methods, for example, by constructing a hybrid stochastic gradient and linear-method optimizer [56] or by rewriting the Hamiltonian in a transcorrelated fashion for the optimization of the linear coefficients [57]. Similarly, in Chapter 3, we develop alternative approximate Hessian expressions when optimizing the variance of the energy with the Newton method. We also note that, for the various optimization approaches in the literature, low-memory implementations specific to QMC have been proposed, which avoid the additional bottleneck of having to store large matrices [58–60].

When optimizing excited-state wave functions, the optimization can be done in a state-average or state-specific mode. The state-average procedure in QMC optimizes Jastrow-Slater wave functions of multiple states, which share the Jastrow and orbital parameters and have different linear coefficients to ensure orthogonality among the states [39, 61] as also done in Chapter 3. Such an approach avoids the collapse of higher-energy states but is variationally less flexible in situations where orbital relaxation is required, rendering state-specific approaches more appealing. Among state-specific procedures, minimizing the variance instead of the energy would seem particularly suitable for excited states since, for exact wave functions, the variance has a minimum for every state. Various methods based on this variational principle have been recently applied to excited states [49,62]. However, as we demonstrate in Chapter 3, the use of variance minimization is, in general, problematic due to the possible presence of very shallow or no barrier in the landscape of the variance for approximate wave functions [61]. Alternatively, one can proceed with energy minimization imposing orthogonality among the states using a penalty method [37, 63–65]. This state-specific approach appears very promising since it can benefit both from well-tested energy minimization schemes, such as the SR method, and from the in-

dependent optimization of all parameters, only having to impose an orthogonality constraint [63, 66].

Nuclear forces. With the availability of robust QMC protocols to build and optimize the electronic wave function, one ingredient is still needed to turn QMC into an internally-consistent method, namely, the optimization of nuclear geometries with QMC forces. This has recently become possible thanks to the formalism introduced in Ref. [67,68], which enables the computation of complete sets of energy derivatives in VMC at a similar cost per Monte Carlo step as the computation of the energy itself. This formalism was successfully employed within our research group, allowing us to optimize all variational parameters in large determinantal expansions. Using CIPSI wave functions, we managed to determine the accurate optimal VMC geometries of several small and medium-size molecules in the ground and excited states [31,32,46].

Therefore, in principle, we should now be in a position to follow an excited-state relaxation in QMC performing Born-Oppenheimer molecular dynamics over the time scale of few picoseconds. To date, however, this has not been done because of one main remaining hurdle, namely, the lack of energy conservation due to the stochastic nature of the interatomic forces. While the increase in computational power might mitigate this issue by enabling longer Monte Carlo sampling, such a brute-force approach is bound to fail as the size of the system grows or the time scales needed are longer. The noise in the QMC forces has been "exploited" within a Langevin scheme to define a temperature and compute thermodynamic averages [69–71], but the direct use of QMC forces in an actual molecular dynamics remains unexplored. In this thesis, we take on this challenging task and analyze in depth several issues related to the use of noisy forces in dynamics. In particular, in Chapter 5, we propose various strategies to deal with the impact of the statistical noise on molecular dynamics simulations, at least over the time scales of a few picoseconds. We then conclude in Chapter 6 by investigating other surprising consequences of the use of QMC forces in molecular dynamics, for instance, those due to their slow convergence during the variational optimization.

As a final remark on nuclear forces in QMC, we note that the computation of forces within DMC is a current topic of research [72–76] and the expression for unbiased energy derivatives of the fixed-node DMC energy has been recently worked out but only applied so far to model systems [77].

Contextualizing with other methods. As mentioned above, despite their favorable scaling with the number of electrons, QMC methods are positioned on the computationally heavy side of the spectrum of electronic structure methods and are surely not a mainstream technique for excited-state investigations. On the side of relatively cheap approaches, we find the widely used density functional theory (DFT) and, when it comes to excited states, its time-dependent formulation (TD-DFT). TDDFT is generally quite successful [78–82] but there are well-known problematic cases such as the description of excitations of charge-transfer or multi-reference nature [79,80,83]. Several solutions have been put forward, but the lack of a unifying approach and the dependence of the results on the choice of the exchange-correlation

functional can lead to practical difficulties in using the method.

A more accurate description of excited states is often offered by the so-called multi-reference methods [25], of which perhaps the most popular is the CAS self-consisted field (CAS-SCF) approach and its second-order perturbation theory extensions, which can successfully incorporate static as well as dynamical correlation via perturbation theory. However, as mentioned above, the results may strongly depend on the choice of the active space and, ultimately, on the users' chemical intuition [84]. Moreover, even though these methods are typically computationally more affordable for small systems than a QMC calculation, they scale exponentially with the number of orbitals in the active space, so a good description with a large enough active space becomes quickly prohibitive for medium to large systems.

Currently, the gold standard for the computation of vertical excitation energies is coupled cluster (CC) theory [85]. This class of techniques is quickly developing and, like QMC, has the advantage of being systematically improvable. Indeed, one can start from a version that only includes single and double excitations (CCSD) and gain further accuracy by including, for instance, triple (T) or even quadruple (Q) excitations. While these methods are black-box and relatively easy to use, they scale poorly with system size with CC3 and CCSDT (needed to treat single excitations) and CC4 and CCSDTQ (needed for double excitations) scaling with the number of electrons as N^7, N^8, N^9, and N^{10}, respectively. Consequently, their application to the study of excited states of larger molecules becomes problematic as we will also see in Chapter 4.

Very expensive but highly accurate are methods that aim at approaching the full CI (FCI) limit in the space of all possible determinants in a given orbital basis. While these methods are computationally demanding and, ultimately, scale exponentially, they have seen remarkable developments in recent years. In particular, approaches like FCI-QMC [87] or the aforementioned selected CI (CIPSI) have been developed to accelerate the convergence to the FCI limit by selecting in a stochastic or deterministic manner, respectively, the most important determinants in the complete space. As already discussed, we make here extensive use of the CIPSI technique not to extrapolate to the FCI limit but to generate compact "smart" determinantal components for our QMC wave functions of multiple states.

To conclude, we note that machine learning can also be used as a tool to enhance the exploration of excited-state PESs with the aid of *ab initio* training data. While machine learning is well established for ground-state problems, its use in the context of excited states, although promising, is still relatively new [88, 89].

In conclusion. With this brief overview, we have tried to position QMC methods in relation to other, more familiar electronic structure techniques, underlining their potential as well as current hurdles. For excited states, a wide variety of quantum approaches are available, which should be viewed as complementary tools, given the difficulty of tackling with a single technique all aspects of an excited-state process (e.g. from the primary photo-excitation to the relaxation towards a conical intersection region). This thesis contributes to this field of research by focusing on the development of QMC methods for excited states and, in particular, setting the first

few steps for the use of QMC forces in molecular dynamics simulations. We believe that further developments of the methods are possible and that QMC approaches are an important asset for excited-state calculations both because of their ability to account for static and dynamical correlation and for their natural adaptability to the on-going increase of computational resources.

1.1 Structure of the thesis

This thesis comprises four research chapters preceded by a concise methodological overview. In the Methods Chapter, we introduce the background theory needed for the applications and further developments contained in this thesis. In particular, we illustrate various techniques available for solving the electronic problem, starting with the Hartree-Fock solution and moving to multi-configuration methods such as CASSCF and CIPSI. We then cover real-space quantum Monte Carlo techniques, which are central to this dissertation. We conclude with various aspects related to the treatment of the nuclei, giving a microscopic derivation of Langevin dynamics to introduce the starting equations for the theory developed in Chapter 5.
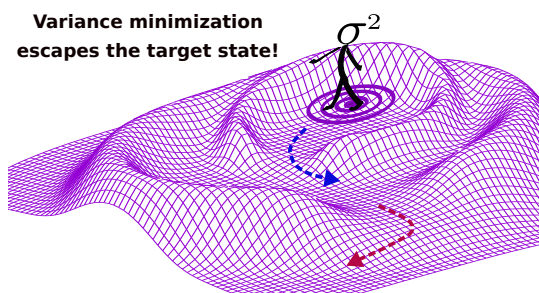
Figure 1.1: Illustrative drawing of our findings regarding variance minimization, where the system loses the target state and converges to states with lower variance, which may lie higher or lower in energy.

Chapter 3 investigates the optimization procedure that better suits the QMC treatment of excited states. In particular, we compare the use of two variational principles, namely, variance and energy minimization. Using energy minimization, we find accurate excitation energies for two prototypical molecules while, with variance minimization, we encounter severe difficulties. In particular, over long optimization runs, one loses the target state since the optimization finds little or low barrier to escape from a local minimum, and the wave function converges to another state characterized by a lower variance as illustrated in a cartoonish manner in Fig. 1.1. Since we do not know *a priori* which eigenstate of the Hamiltonian corresponds to the global minima of the variance, we conclude that variance optimization is not a suitable procedure for the state-specific treatment of excited states.

In Chapter 4, we employ QMC to establish accurate excitation energies of cyanine dye molecules and explore how to build compact QMC wave functions that offer

a balanced description of the ground and the excited state of interest. To this aim, we generate multi-determinantal components of the QMC wave functions with the CIPSI approach and find that employing expansions with matched CI variances guarantees a good relative description of the states and of the corresponding excitation energies. Following this protocol, we establish new reference data for cyanine dyes of increasing size, reaching systems where other accurate, suitable methods, such as CC3, are not computationally affordable, as illustrated in Fig 1.2.
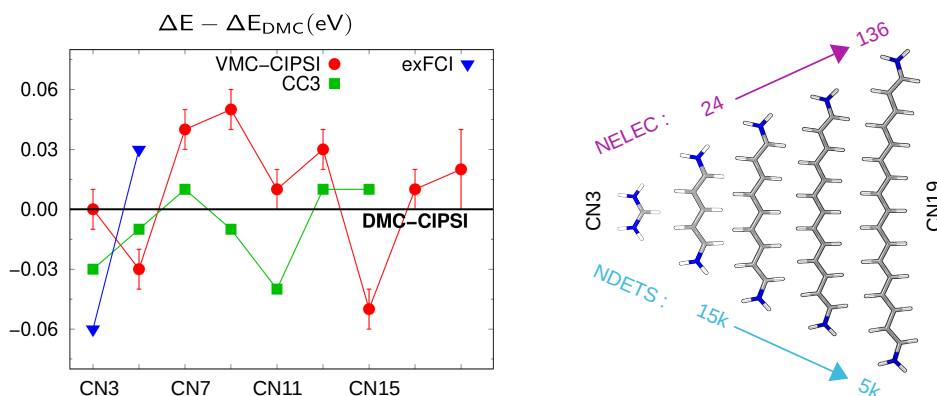
Figure 1.2: Difference in excitation energies between highly-correlated methods such as CC3 and extrapolated FCI (when available) and DMC-CIPSI (left panel) for a series of cyanine dyes. The VMC-CIPSI relative values are also plotted. In the right panel, we show some of the cyanine molecules treated and indicate the size of the determinantal component used in the QMC calculations.

In Chapters 5 and 6, we explore the impact of using noisy forces in molecular dynamics simulations. Our interest is in using QMC forces to dynamically follow the relaxation of a system in the excited state and therefore focus on the time scales of a few picoseconds. As shown in Fig. 1.3 for thiophene, as a prototypical molecule, the use of noisy QMC forces leads to an increase of the total energy in time, unphysical rotations and translations, and an overall corruption of the dynamical path. In Chapter 5, we develop various strategies to stabilize molecular dynamics simulations in which the forces are affected by stochastic noise and, for the sake of efficiency, do so using a classical force field to which we manually add the QMC statistical error on the forces. We find a promising route to ameliorate the problem by performing a fit of the forces in time and, successively, apply a Langevin-like scheme to thermalize the excess noise in the forces.

Having explored a protocol to limit the impact of the noise, in Chapter 6, we evidence additional problems that arise when performing molecular dynamics simulations using actual QMC forces. In particular, a partially converged variational optimization of the wave function has unexpected consequences for the dynamics of a system. Although costly, a strict convergence of the interatomic forces during the wave function optimization is necessary to ensure energy conservation. We propose
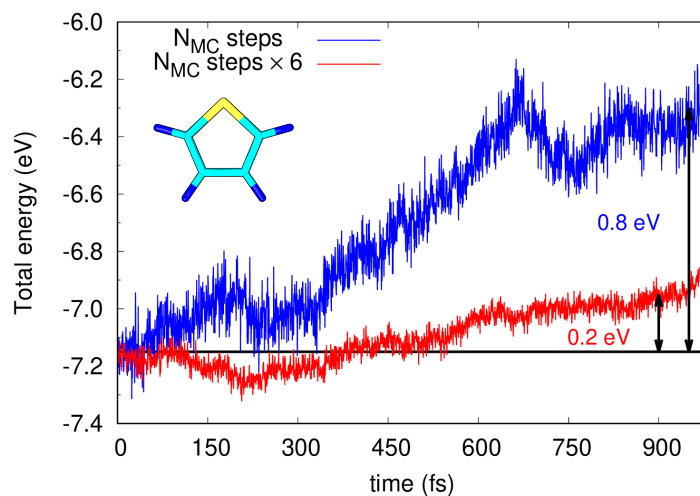
Figure 1.3: Total (kinetic and potential) energy of a thiophene molecule in which we drive the molecular dynamics simulation with QMC forces. The energy increases in time faster for the simulation where fewer Monte Carlo steps are used to sample the forces, resulting in larger noise.

a method to speed up the variational optimization during a molecular dynamics run, which we tested so far on a simple C_2 molecule. Finally, we combine all our findings and perform molecular dynamics simulations with QMC forces on C_2 and thiophene. We conclude by showing some promising preliminary relaxation runs of thiophene in the excited state.

Bibliography

[1] C. E. Crespo-Hernández, B. Cohen, P. M. Hare, and B. Kohler, Chem Rev **104**, 1977 (2004).

[2] M. Barbatti, A. C. Borin, and S. Ullrich, Top Curr Chem **355**, 1 (2015).

[3] T. van Leeuwen, A. S. Lubbe, P. Štacko, S. J. Wezenberg, and B. L. Feringa, Nature Reviews Chemistry **1**, 0096 (2017).

[4] J. Wang, Z. Xiong, J. Zheng, X. Zhan, and J. Tang, Acc. Chem. Res. **51**, 1957 (2018).

[5] R. Bongiovanni, S. D. Vacche, and A. Vitale, Polymers **13**, (2021).

[6] J. Herbst, K. Heyne, and R. Diller, Science **297**, 822 (2002).

[7] E. Tapavicza, I. Tavernelli, and U. Rothlisberger, Phys. Rev. Lett. **98**, 023001 (2007).

[8] S. Sen, R. K. Kar, V. A. Borin, and I. Schapiro, Wiley Interdiscip. Rev. Comput. Mol. Sci. **12**, e1562 (2022).

[9] G. Cerullo, D. Polli, G. Lanzani, S. De Silvestri, H. Hashimoto, and R. J. Cogdell, Science **298**, 2395 (2002).

[10] Y.-C. Cheng and G. R. Fleming, Annu. Rev. Phys. Chem. **60**, 241 (2009).

[11] G. J. Hedley, A. Ruseckas, and I. D. W. Samuel, Chem. Rev. **117**, 796 (2016).

[12] P. D. Dahlberg, P.-C. Ting, S. C. Massey, M. A. Allodi, E. C. Martin, C. N. Hunter, and G. S. Engel, Nat. Commun. **8**, 988 (2017).

[13] S. J. Jang and B. Mennucci, Rev. Mod. Phys. **90**, 035003 (2018).

[14] N. M. O'Boyle, C. M. Campbell, and G. R. Hutchison, J. Phys. Chem. C **115**, 16200 (2011).

[15] S. Mathew, A. Yella, P. Gao, R. Humphry-Baker, B. F. E. Curchod, N. Ashari-Astani, I. Tavernelli, U. Rothlisberger, M. K. Nazeeruddin, and M. Grätzel, Nat. Chem. **6**, 242 (2014).

[16] S. Matsika and A. I. Krylov, Chem. Rev. **118**, 6925 (2018).

[17] E. Brunk and U. Rothlisberger, Chem. Rev. **115**, 6217 (2015).

[18] B. Mennucci and S. Corni, Nat. Rev. Chem. **3**, 315 (2019).

[19] F. Mouvet, J. Villard, V. Bolnykh, and U. Rothlisberger, Acc. Chem. Res. **55**, 221 (2022).

[20] H. M. Senn and W. Thiel, Angew. Chem. Int. Ed. **48**, 1198 (2009).

[21] A. V. Akimov and O. V. Prezhdo, Chem. Rev. **115**, 5797 (2015).

[22] D. R. Yarkony, Chem. Rev. **112**, 481 (2012).

[23] R. Crespo-Otero and M. Barbatti, Chem. Rev. **118**, 7026 (2018).

[24] N. J. Hestand and F. C. Spano, Chem. Rev. **118**, 7069 (2018).

[25] H. Lischka, D. Nachtigallová, A. J. Aquino, P. G. Szalay, F. Plasser, F. B. C. Machado, and M. Barbatti, Chem. Rev. **118**, 7293 (2018).

[26] M. Nightingale and C. Umrigar, *Quantum Monte Carlo Methods in Physics and Chemistry*, *Nato Science Series C:* (Springer Netherlands, 1998).

[27] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001).

[28] B. M. Austin, D. Y. Zubarev, and W. A. Lester, Chem. Rev. **112**, 263 (2012).

[29] A. Lüchow, Wiley Interdiscip. Rev. Comput. Mol. Sci. **1**, 388 (2011).

[30] J. Feldt and C. Filippi, in *Quantum Chemistry and Dynamics of Excited States: Methods and Applications*, edited by R. Lindh and L. Gonzalez (John Wiley and Sons Ltd, 2019).

[31] M. Dash, J. Feldt, S. Moroni, A. Scemama, and C. Filippi, J. Chem. Theory Comput. **15**, 4896 (2019).

[32] M. Dash, S. Moroni, C. Filippi, and A. Scemama, J. Chem. Theory Comput. **17**, 3426 (2021).

[33] A. Cuzzocrea, S. Moroni, A. Scemama, and C. Filippi, J. Chem. Theory Comput. **18**, 1089 (2022).

[34] G. Carleo and M. Troyer, Science **355**, 602 (2017).

[35] D. Pfau, J. S. Spencer, A. G. D. G. Matthews, and W. M. C. Foulkes, Phys. Rev. Research **2**, 033429 (2020).

[36] J. Hermann, Z. Schätzle, and F. Noé, Nature Chemistry **12**, 891 (2020).

[37] M. Entwistle, Z. Schätzle, P. A. Erdman, J. Hermann, and F. Noé, Electronic excited states in deep variational Monte Carlo, 2022.

[38] J. Hermann, J. Spencer, K. Choo, A. Mezzacapo, W. M. C. Foulkes, D. Pfau, G. Carleo, and F. Noé, Ab-initio quantum chemistry with neural-network wave-functions, 2022.

[39] C. Filippi, M. Zaccheddu, and F. Buda, J. Chem. Theory Comput. **5**, 2074 (2009).

[40] P. M. Zimmerman, J. Toulouse, Z. Zhang, C. B. Musgrave, and C. Umrigar, J. Chem. Phys. **131**, 124103 (2009).

[41] T. Bouabça, N. Ben Amor, D. Maynau, and M. Caffarel, J. Chem. Phys. **130**, 114107 (2009).

[42] M. Dubecký, R. Derian, L. Mitas, and I. Štich, J. Chem. Phys. **133**, 244301 (2010).

[43] O. Valsson and C. Filippi, J. Chem. Theory Comput. **6**, 1275 (2010).

[44] R. Send, O. Valsson, and C. Filippi, J. Chem. Theory Comput. **7**, 444 (2011).

[45] C. Daday, C. Curutchet, A. Sinicropi, B. Mennucci, and C. Filippi, J. Chem. Theory Comput. **11**, 4825 (2015).

[46] M. Dash, S. Moroni, A. Scemama, and C. Filippi, J. Chem. Theory Comput. **14**, 4176 (2018).

[47] A. Scemama, A. Benali, D. Jacquemin, M. Caffarel, and P.-F. Loos, J. Chem. Phys. **149**, 034108 (2018).

[48] A. Scemama, M. Caffarel, A. Benali, D. Jacquemin, and P.-F. Loos, Results Chem. **1**, 100002 (2019).

[49] S. D. Pineda Flores and E. Neuscamman, J. Phys. Chem. A **123**, 1487 (2019).

[50] P. R. C. Kent, A. Annaberdiyev, A. Benali, M. C. Bennett, E. J. Landinez Borda, P. Doak, H. Hao, K. D. Jordan, J. T. Krogel, I. Kylänpää, J. Lee, Y. Luo, F. D. Malone, C. A. Melton, L. Mitas, M. A. Morales, E. Neuscamman, F. A. Reboredo, B. Rubenstein, K. Saritas, S. Upadhyay, G. Wang, S. Zhang, and L. Zhao, J. Chem. Phys. **152**, 174105 (2020).

[51] C. J. Umrigar and C. Filippi, Phys. Rev. Lett. **94**, 150201 (2005).

[52] C. J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R. G. Hennig, Phys. Rev. Lett. **98**, 110201 (2007).

[53] C. J. U. Julien Toulouse, J. Chem. Phys. **126**, 084102 (2017).

[54] S. Sorella and L. Capriotti, Phys. Rev. B **61**, 2599 (2000).

[55] S. Sorella, M. Casula, and D. Rocca, J. Chem. Phys. **127**, 014105 (2007).

[56] L. Otis and E. Neuscamman, Phys. Chem. Chem. Phys. **21**, 14491 (2019).

[57] A. Ammar, E. Giner, and A. Scemama, J. Chem. Theory Comput. **18**, 5325 (2022).

[58] E. Neuscamman, C. J. Umrigar, and G. K.-L. Chan, Phys. Rev. B **85**, 045103 (2012).

[59] L. Zhao and E. Neuscamman, J. Chem. Theory Comput. **13**, 2604 (2017).

[60] I. Sabzevari, A. Mahajan, and S. Sharma, J. Chem. Phys. **152**, 024111 (2020).

[61] A. Cuzzocrea, A. Scemama, W. J. Briels, S. Moroni, and C. Filippi, J. Chem. Theory Comput. **16**, 4203 (2020).

[62] S. Jacqueline A. R. and N. Eric, J. Chem. Theory Comput. **13**, 6078 (2017).

[63] S. Pathak, B. Busemeyer, J. N. B. Rodrigues, and L. K. Wagner, J. Chem. Phys. **154**, 034101 (2021).

[64] N. S. Blunt, S. D. Smart, G. H. Booth, and A. Alavi, J. Chem. Phys. **143**, 134117 (2015).

[65] K. Choo, G. Carleo, N. Regnault, and T. Neupert, Phys. Rev. Lett. **121**, 167204 (2018).

[66] S. Shepard, R. L. Panadés-Barrueta, S. Moroni, A. Scemama, and C. Filippi, Double excitation energies from quantum Monte Carlo using state-specific energy optimization, 2022.

[67] C. Filippi, R. Assaraf, and S. Moroni, J. Chem. Phys. **144**, 194105 (2016).

[68] R. Assaraf, S. Moroni, and C. Filippi, J. Chem. Theory Comput. **13**, 5273 (2017).

[69] C. Attaccalite and S. Sorella, Phys. Rev. Lett. **100**, 114501 (2008).

[70] G. Mazzola, S. Yunoki, and S. Sorella, Nat. Commun. **5**, 3487 (2014).

[71] F. Mouhat, S. Sorella, R. Vuilleumier, A. M. Saitta, and M. Casula, J. Chem. Theory Comput. **13**, 2400 (2017).

[72] C. Filippi and C. J. Umrigar, Phys. Rev. B **61**, R16291 (2000).

[73] A. Badinski, P. D. Haynes, J. R. Trail, and R. J. Needs, J. Phys. Condens. Matter **22**, 074202 (2010).

[74] R. Assaraf, M. Caffarel, and A. C. Kollias, Phys. Rev. Lett. **106**, 150601 (2011).

[75] S. Moroni, S. Saccani, and C. Filippi, J. Chem. Theory Comput. **10**, 4823 (2014).

[76] J. Tiihonen, R. C. Clay, and J. T. Krogel, J. Chem. Phys. **154**, 204111 (2021).

[77] J. van Rhijn, C. Filippi, S. De Palo, and S. Moroni, J. Chem. Theory Comput. **18**, 118 (2022).

[78] L. Goerigk and S. Grimme, J. Chem. Phys. **132**, 184103 (2010).

[79] H. S. Yu, S. L. Li, and D. G. Truhlar, J. Chem. Phys. **145**, 130901 (2016).

[80] N. T. Maitra, J. Chem. Phys. **144**, 220901 (2016).

[81] A. Charaf-Eddin, A. Planchat, B. Mennucci, C. Adamo, and D. Jacquemin, J. Chem. Theory Comput. **9**, 2749 (2013).

[82] D. Jacquemin, B. Moore, A. Planchat, C. Adamo, and J. Autschbach, J. Chem. Theory Comput. **10**, 1677 (2014).

[83] C. L. Benavides-Riveros, N. N. Lathiotakis, and M. A. L. Marques, Phys. Chem. Chem. Phys. **19**, 12655 (2017).

[84] V. Veryazov, P. r. Malmqvist, and B. O. Roos, Int. J. Quantum Chem. **111**, 3329 (2011).

[85] A. I. Krylov, Annual Review of Physical Chemistry **59**, 433 (2008).

[86] P.-F. Loos, M. Boggio-Pasqua, A. Scemama, M. Caffarel, and D. Jacquemin, J. Chem. Theory Comput. **15**, 1939 (2019).

[87] G. H. Booth, A. J. W. Thom, and A. Alavi, The Journal of Chemical Physics **131**, 054106 (2009).

[88] J. Westermayr, M. Gastegger, M. F. S. J. Menger, S. Mai, L. González, and P. Marquetand, Chem. Sci. **10**, 8100 (2019).

[89] J. Westermayr and P. Marquetand, Chem. Rev. **121**, 9873 (2021).

Chapter 2

Theoretical methods

In this Chapter, we describe the electronic structure methods that are at the core of this dissertation and briefly introduce the basics of molecular dynamics simulations.

2.1   Separating the electronic and nuclear motion

The time-dependent Schrödinger equation in atomic units ($\hbar = m_e = e = 1$) has the following form

$$i\frac{\partial}{\partial t}\Psi(\mathbf{r},\mathbf{R},t) = \hat{\mathcal{H}}(\mathbf{r},\mathbf{R})\Psi(\mathbf{r},\mathbf{R},t)\,, \tag{2.1}$$

where Ψ is the total wave function, and $\mathbf{r}$ and $\mathbf{R}$ represent the electronic and nuclear positions, respectively.

The non-relativistic Hamiltonian $\hat{\mathcal{H}}$ is given by

$$
\begin{aligned}
\hat{\mathcal{H}}(\mathbf{r},\mathbf{R}) &= -\sum_I \frac{\nabla_I^2}{2M_I} - \sum_i \frac{\nabla_i^2}{2} + \sum_{i<j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \\
&\quad - \sum_{I,i}\frac{Z_I}{|\mathbf{R}_I - \mathbf{r}_i|} + \sum_{I<J}\frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \\
&= \hat{T}_N(\mathbf{R}) + \hat{\mathcal{H}}_{\mathrm{el}}(\mathbf{r},\mathbf{R})\,,
\end{aligned}
\tag{2.2}
$$

where we separate the nuclear kinetic energy operator from the electronic Hamiltonian, which includes the kinetic operator for the electrons and the Coulomb interactions of the electrons and nuclei with charge Z. The sums over capitalized indices are relative to the nuclei, while lower-case indices label the electrons.

For fixed nuclear positions $\mathbf{R}$, the solution of the time-independent electronic Schrödinger equation is

$$\hat{\mathcal{H}}_{\mathrm{el}}(\mathbf{r},\mathbf{R})\psi_j(\mathbf{r},\mathbf{R}) = E_j^{\mathrm{el}}\psi_j(\mathbf{r},\mathbf{R})\,, \tag{2.3}$$

with $\{\psi_j\}$ the electronic eigenfunctions with eigenvalues E_j^{el}.

In the Born-Huang (BH) representation [1, 2], we write the full molecular wave function in terms of the electronic eigenstates

$$\Psi(\mathbf{r}, \mathbf{R}, t) = \sum_{j}^{\infty} \psi_j(\mathbf{r}, \mathbf{R}) \chi_j(\mathbf{R}, t) \,, \tag{2.4}$$

and therefore include a dependency on time only in the nuclear solutions $\{\chi_j\}$. Inserting this wave function in Eq. 2.1 and using the orthogonality of the electronic eigenstates, we obtain

$$
\begin{aligned}
i \frac{\partial \chi_k}{\partial t} &= [\hat{T}_N(\mathbf{R}) + E_k^{el}(\mathbf{R})] \chi_k(\mathbf{R}, t) \\
&\quad - \sum_{j}^{\infty} \sum_{I}^{N_n} \frac{1}{2M_I} [\langle \psi_k | \nabla_\mathbf{R}^2 | \psi_j \rangle + 2 \langle \psi_k | \nabla_\mathbf{R} | \psi_j \rangle \nabla_\mathbf{R}] \chi_j(\mathbf{R}, t) \,.
\end{aligned}
\tag{2.5}
$$

The equation in the first line evolves the nuclear wave function on a specific potential energy surface (PES) given by the electronic energy of an eigenstate, while the second part contains non-adiabatic couplings to the other eigenstates.

In this thesis, we follow the Born-Oppenheimer approximation (BOA) [3] and assume that the coupling terms are negligible so that we can approximate the wave function with a single term as

$$\Psi(\mathbf{r}, \mathbf{R}, t) \approx \psi_j(\mathbf{r}, \mathbf{R}) \chi_j(\mathbf{R}, t) \,. \tag{2.6}$$

This means that, during the time evolution, we assume that the system always stays on one PES, $E_j^{el}(\mathbf{R})$, and that the electrons relax instantaneously when the nuclei move [4, 5]. The BOA is usually a safe choice when studying ground-state properties but, if one is interested in dynamical processes in the excited state, one needs to take in account that, when two PESs become close to each other, the coupling terms become relevant, and the BOA does no longer hold. In this case, one may adopt non-adiabatic molecular dynamics techniques such as surface hopping [6,7] to enable the system to "hop" between different PESs. In this thesis, we will explore the use of quantum Monte Carlo forces in dynamics and simply stop the evolution when the chosen excited-state PES is energetically close to another one.

2.2 Solving the electronic problem

We focus now on the computation of the electronic energy for given nuclear coordinates $\mathbf{R}$, and write the electronic Hamiltonian as

$$\hat{\mathcal{H}}_{el}(\mathbf{r}, \mathbf{R}) = -\sum_{i} \frac{\nabla_i^2}{2} + \sum_{i} v_{ext}(\mathbf{r}_i) + \sum_{i<j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \,, \tag{2.7}$$

where we omit the nuclear repulsion term, which is a constant, and denote the electron-nucleus Coulomb interaction as v_{ext}. While we can rewrite the first two

terms as a sum of one-body operators, the electron-electron interaction is of course non separable and an approximate solution to the problem must be sought.

Electronic structure theory focuses on the solution of this equation using different approximate schemes, which can be mainly divided into two large categories, namely, density functional theory (DFT) and wave function methods. In the first case, the interacting problem is exactly mapped to a non-interacting one in an effective potential which accounts for exchange and correlation among the electrons and, in practice, must be approximated. In the second case, one attempts to directly solve the fully interacting Schrödinger equation, always introducing some approximations. In this thesis, we are mainly concerned with the second category of approaches, which we briefly describe below.

2.2.1 Hartree-Fock approximation

The starting point of all wave function methods is the Hartree-Fock (HF) approximation, which uses a non-interacting ansatz for the wave function to describe a system of interacting electrons [8–13]. The wave function is a Slater determinant (SD) [14] of single-particle spin-orbitals, which accounts for the anti-symmetry of the fermionic solution.

If we denote by $\mathbf{x} = (\mathbf{r}, \sigma)$ the electron spatial and spin coordinates ($\sigma = \pm 1$), a SD is given by

$$\psi_{\mathrm{SD}}(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \Phi_1(\mathbf{x}_1) & \Phi_1(\mathbf{x}_2) & \cdots & \Phi_1(\mathbf{x}_N) \\ \Phi_2(\mathbf{x}_1) & \Phi_2(\mathbf{x}_2) & \cdots & \Phi_2(\mathbf{x}_N) \\ \vdots & \vdots & \vdots & \vdots \\ \Phi_N(\mathbf{x}_1) & \Phi_N(\mathbf{x}_2) & \cdots & \Phi_N(\mathbf{x}_N) \end{vmatrix}.$$

where N is the total number of electrons and $\{\Phi_i\}$ is a set of orthonormal spin-orbitals usually written as a product of spatial and spin orbitals, $\Phi_i = \phi_i(\mathbf{r})\chi_i(\sigma)$. Being a determinant, the wave function is zero if two electrons occupy the same space-spin position ($\mathbf{x}$) or if two orbitals are the same.

To find the optimal non-interacting solution for the interacting system, one minimizes the expectation value of $\hat{\mathcal{H}}_{\mathrm{el}}$ on the wave function ψ_{SD} with respect to the orbitals, which results in the HF equation

$$\hat{f} \left|\phi_i\right\rangle = \sum_{j=1}^{N} \epsilon_{ij} \left|\phi_j\right\rangle, \tag{2.8}$$

where $\hat{f}$ is the Fock operator and the sum runs over all the occupied orbitals. The real-space representation of the left-hand-side of the HF equation is given by

$$\langle\mathbf{r}| \hat{f} \left|\phi_i\right\rangle = -\frac{1}{2}\nabla^2\phi_i(\mathbf{r}) + v_{\mathrm{ext}}(\mathbf{r})\phi_i(\mathrm{r}) + \left(v^{\mathrm{HF}}\phi_i\right)(\mathbf{r}), \tag{2.9}$$

17

which, in addition to the external potential, contains an effective non-local potential, v_{HF}. This consists of two terms, a Coulomb and an exchange term as

$$
\begin{aligned}
\left(v^{\text{HF}}\phi_i\right)(\mathbf{r}) &= \left(v_{\text{Coulomb}}\phi_i\right)(\mathbf{r}) + \left(v_{\text{x}}^{\text{HF}}\phi_i\right)(\mathbf{r}) \\
\left(v_{\text{Coulomb}}\phi_i\right)(\mathbf{r}) &= \left[\sum_{j=1}^{N}\int \frac{|\phi_j(\mathbf{r}')|^2}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}'\right]\phi_i(\mathbf{r}) \\
\left(v_{\text{x}}^{\text{HF}}\phi_i\right)(\mathbf{r}) &= -\sum_{j=1}^{N}\delta_{s_i s_j}\left[\int \frac{\phi_j^*(\mathbf{r}')\phi_i(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}'\right]\phi_j(\mathbf{r})\,. \qquad (2.10)
\end{aligned}
$$

where v_{Coulomb} describes the electrostatic interaction of an electron with the charge distribution of the system. The $j = i$ component in the exchange potential, v_{x}^{HF}, cancels the self-interaction term in the Coulomb potential, whereas the $j \neq i$ terms account for the exchange interaction with electrons in orbitals with the same spin quantum number. This last term is a direct consequence of the wave function anti-symmetry, so it has no classical counterpart.

The Hartree-Fock equation is solved iteratively to find the optimal $\{\phi_i\}$. Starting from a guess for the initial spin orbitals, one constructs the HF potential, solves the HF equations to obtain a new set of spin orbitals, and iterates until convergence. For molecular systems, a common choice is to represent the spin-orbitals as a linear combination of atomic orbitals (LCAO) as

$$
\phi_i(\mathbf{r}) = \sum_{I=1}^{N_{\text{n}}}\sum_{j=1}^{M_{\text{b}}^I} a_{ji}^I\,\eta_{jI}(\mathbf{r}-\mathbf{R}_I)\,, \qquad (2.11)
$$

where M_{b}^I is the number of atomic basis functions centered on nucleus I. The functions $\{\eta\}$ are usually chosen to be Gaussians,

$$
\eta(\mathbf{r}) = x^a y^b z^c \exp(-\alpha r^2)\,. \qquad (2.12)
$$

since this choice allows the analytical computation of the one- and two-body integrals obtained from expressing the Fock operator in the atomic basis. Finding the HF solution becomes equivalent to the problem of determining the optimal LCAO coefficients, a_{ji}^I.

2.2.2   Beyond the Hartree-Fock approximation

The HF method gives the optimal non-interacting solution to an interacting problem and the difference between the resulting energy, E_{HF}, and the exact one, E, is called correlation energy,

$$
E_{\text{corr}} = E - E_{\text{HF}}\,. \qquad (2.13)
$$

Post-HF quantum chemical methods attempt to recover this missing part of the energy with the use of a better ansatz for the wave function.

The most straightforward way to improve on HF is to expand the wave function in terms of multiple Slater determinants and build a so-called configuration interaction (CI) wave function. Such an expansion is constructed by exciting one or more electrons from the orbitals occupied in the HF determinant to the virtual orbitals as

$$\psi_{\text{CI}} = c_0 D_{\text{HF}} + \sum_{ac} c_{a \to c} D^{a \to c} + \sum_{abcd} c_{ab \to cd} D^{ab \to cd} + \dots , \qquad (2.14)$$

where, in addition to the HF determinant, one has determinants consisting of single excitation ($a \to c$), double excitation $ab \to cd$, and so on. Depending on the problem of interest, one can decide how many and which contributions to include in the expansion.

For a given expansion in Slater determinants, it is then possible to minimize the energy and find the optimal set of CI coefficients, $\{c_i\}$, by solving the secular equation,

$$\sum_{j=1}^{K} \langle D_i | \hat{\mathcal{H}}_{el} | D_j \rangle c_j^{(k)} = E_{\text{CI}}^{(k)} \sum_{j=1}^{K} \langle D_i | D_j \rangle c_j^{(k)} , \qquad (2.15)$$

where $\langle D_i | D_j \rangle = \delta_{ij}$ since the orbitals are orthonormal. It is important to note that going beyond the single determinant ansatz also enables the description of excited states. In fact, diagonalizing the secular problem, we obtain a set of energies, $\{E_{\text{CI}}^{(k)}\}$, each being variational with respect to the corresponding eigenstate [15, 16].

In principle, for a given basis set, one can consider expansions including all possible excitations (up to N electrons and to behind all orbitals) and find the best possible solution in the given basis. This procedure, called full CI (FCI), leads to very accurate results but is computationally too expensive and can only be applied to very small systems. For this reason, one usually adopts smaller expansions such as a CI singles (CIS), where one only considers the determinants relatives to single excitations, or CI singles and doubles (CISD), which also includes double excitations.

A more involved treatment is the multi-configurational self-consistent-field (MC-SCF) approach where one optimizes the molecular orbitals in the expansion in addition to the CI coefficients. This approach considerably improves the description of a given state since the optimal LCAO coefficients in a multi-determinantal expansion can be quite different from the starting HF ones. In this thesis, we often use the complete-active-space SCF (CASSCF) method, one of the most popular among the MCSCF approaches. In a CASSCF calculation, one divides the orbital space into three groups, namely, doubly occupied, active, and virtual orbitals, and generates a FCI expansion in the restricted space of the active orbitals, optimizing both the CI and the LCAO coefficients. The active space is usually labeled as CAS(n, m) where one considers all possible excitation of n electrons in m active orbitals [17]. The choice of orbitals included in the active space is often done based on a chemical intuition and represents the most delicate part of the procedure as we discuss later in this thesis.

Finally, if one is interested in describing an excited state which has the same symmetry as lower-lying state(s), one typically performs a state-average (SA-CASSCF)

calculation over the states of interest to obtain a set of orbitals, which offers a balanced description of the multiple states by minimizing a weighted average of their energies,

$$E_{\text{SA}} = \sum_J w_J \frac{\langle \psi^J | \hat{\mathcal{H}}_{\text{el}} | \psi^J \rangle}{\langle \psi^J | \psi^J \rangle}\,, \tag{2.16}$$

with $\sum_J w_J = 1$, where the index J labels the states. With this procedure, the wave functions share the same orbitals and the CI coefficients are used to ensure orthogonality. It is important to note that, since the orbitals are common, they are not optimal for any of the states involved but quantities such as the excitations energy are reasonably described.

2.2.3 Perturbation methods

One route to improve upon methods such as HF, CI, or CASSCF is to use perturbation theory. In the Rayleigh-Schrödinger formulation, the total electron Hamiltonian is partitioned into a zeroth-order Hamiltonian, $\hat{\mathcal{H}}_{\text{el}}^{(0)}$, and a perturbation operator, $\hat{\mathcal{V}}$,

$$\hat{\mathcal{H}}_{\text{el}} = \hat{\mathcal{H}}_{\text{el}}^{(0)} + \hat{\mathcal{V}}\,. \tag{2.17}$$

The Møller-Plesset perturbation theory at second order (MP2) or the complete-active-space perturbation theory at second order (CASPT2) are examples of perturbation methods to the same order, which differ in the choice of the zeroth-order Hamiltonian. In MP2, one uses the Hartree-Fock operator as the zeroth-order Hamiltonian [18], while in CASPT2 one constructs a one-body Hamiltonian with the CASSCF wave function as eigenstate [19, 20].

In this thesis, we extensively use the Epstein-Nesbet partitioning of the Hamiltonian [21, 22], where the FCI space is divided in a space $\mathcal{S}$, spanned by the determinants $\{|D_i\rangle\}$, and the external space of all other determinants $\{|\alpha\rangle\}$. The zeroth-order Hamiltonian is then constructed to have the same matrix elements of the original Hamiltonian within the space $\mathcal{S}$,

$$\langle D_i | \hat{\mathcal{H}}_{\text{el}}^{(0)} | D_j \rangle = \langle D_i | \hat{\mathcal{H}}_{\text{el}} | D_j \rangle\,, \tag{2.18}$$

and to be diagonal in the space external space as

$$\langle D_i | \hat{\mathcal{H}}_{\text{el}}^{(0)} | \alpha \rangle = 0 \tag{2.19}$$

$$\langle \alpha | \hat{\mathcal{H}}_{\text{el}}^{(0)} | \alpha \rangle = \langle \alpha | \hat{\mathcal{H}}_{\text{el}} | \alpha \rangle\,, \tag{2.20}$$

Consequently, $\hat{\mathcal{H}}_{\text{el}}^{(0)}$ is defined as

$$\hat{\mathcal{H}}_{\text{el}}^{(0)} = \sum_{i,j \in \mathcal{S}} \langle D_i | \hat{\mathcal{H}}_{\text{el}} | D_j \rangle | D_i \rangle \langle D_j | + \sum_{\alpha \notin \mathcal{S}} \langle \alpha | \hat{\mathcal{H}}_{\text{el}} | \alpha \rangle | \alpha \rangle \langle \alpha |\,, \tag{2.21}$$

and the perturbative potential has then the following properties

$$\langle D_i | \hat{\mathcal{V}} | \alpha \rangle = \langle D_i | \hat{\mathcal{H}}_{\text{el}} | \alpha \rangle$$
$$\langle D_i | \hat{\mathcal{V}} | D_j \rangle = 0\,. \tag{2.22}$$

Within this formulation, a CI wave function which diagonalizes $\hat{H}_{el}$ in the current space $\mathcal{S}$ with variational energy $E^{(0)}$ is an eigenstate of $\hat{\mathcal{H}}_{el}^{(0)}$ with the same energy. Furthermore, the first-order perturbation correction to the energy is zero and the second-order one equals

$$E_{\mathrm{PT2}} = \sum_\alpha \frac{|\langle \alpha | \hat{\mathcal{H}}_{el} | \Psi^{(0)} \rangle|^2}{E^{(0)} - \langle \alpha | \hat{\mathcal{H}}_{el} | \alpha \rangle} = \sum_\alpha \delta E_\alpha^{(2)}, \qquad (2.23)$$

which we have rewritten as the sum of the contributions from the determinants, $\{|\alpha\rangle\}$ outside the space $\mathcal{S}$. The Epstein-Nesbet perturbation theory is used to estimate the second-order corrections in the CI perturbatively selected iteratively (CIPSI) method.

2.2.4 CIPSI

CIPSI is a selective CI (sCI) approach, which relies on perturbation theory for the selection of the important determinants in the expansion. Given a starting set of determinants in the space $\mathcal{S}$, one iteratively selects other relevant contributions, $|\alpha\rangle$, outside the space according to their second-order perturbation energy contribution [23–25].

After constructing a starting wave function which defines the initial space $\mathcal{S}$, the CIPSI procedure can be summarized into the following main steps:

1. For all determinants $\{|D_i\rangle\}$ in $\mathcal{S}$, we construct the determinants of the external space $\{|\alpha\rangle\}$ which are given by all possible single and double excitations from $\{|D_i\rangle\}$.

2. For each $|\alpha\rangle$, we compute the second-order perturbative correction, $\delta E_\alpha^{(2)}$.

3. A subset $\{|\alpha*\rangle\}$ of $\{|\alpha\rangle\}$ is selected based on a given threshold on $\delta E_\alpha^{(2)}$. This subset is included in the reference space $\mathcal{S}$ as

$$\mathcal{S}' \leftarrow \mathcal{S} \cup \{|\alpha*\rangle\}, \qquad (2.24)$$

where we denote the enlarged reference space as S'.

4. By diagonalizing the Hamiltonian in the space S', we find the new reference wave function with its optimal linear coefficients and variational energy. A new expression for the zeroth-order Hamiltonian $\hat{\mathcal{H}}_{el}^{(0)}$ (Eq. 2.21) can then be constructed.

5. We iterate from Step 1 until a given criterion is satisfied, for instance, if a target number of determinants is reached or if E_{PT2} is such that the variational CI energy can be reliably extrapolated to the full CI energy. In this thesis, we will typically stop the CIPSI procedure when we obtain a target value of PT2 energy or CI variance.

The choice of the starting space $\mathcal{S}$ depends on the nature of the states of interest and it is usually a HF determinant when treating the ground state. In this thesis, we consider excited states of predominantly single-excitation nature and different symmetry than the ground state, and we therefore perform a CIS calculation to construct a starting expansion from the subset of single excitations in the same symmetry class of the target state. From this subset, we retain the CIS determinants with higher CI coefficient.

The CIPSI procedure outlined above is rather schematic and the actual scheme is in fact more complicated to render the algorithm more memory efficient [23, 26]. In particular, the number of single and double excitations from the reference determinants scales like $\mathcal{O}(N_{\mathrm{det}} \times N^2 \times N_{\mathrm{bas}}^2)$ where N_{det} is the number of determinants in $\mathcal{S}$, N the number of electrons, and N_{bas} the total number of basis functions. Storing all their second-order contributions $\delta E_\alpha^{(2)}$ for the subsequent selection would be too memory demanding. To address this issue, the present algorithm uses a recently introduced stochastic PT2 sampling method [26].

In this procedure, one only generates single and double excitations out of a subset of determinants in the reference space $\mathcal{S}$, which are drawn stochastically using Monte Carlo (MC) sampling from the $\mathcal{S}$ with probabilities proportional to their weights. For each drawn determinant, all possible single and double excitations are built, and their net second-order energy contribution is computed. In order to describe this algorithm, we first need to define a partitioning of the external determinantal space $\{|\alpha\rangle\}$ into N_{det} subsets $\{|\alpha_i\rangle\}$ so that

$$\{|\alpha\rangle\} = \bigcup_{i=1}^{N_{\mathrm{det}}} \{|\alpha_i\rangle\}, \quad \text{where} \quad \{|\alpha_i\rangle\} \cap \{|\alpha_j\rangle\} = \emptyset \quad \forall \quad i \neq j, \tag{2.25}$$

where each $\{|\alpha_i\rangle\}$ contains all *unique* single and double excitations that can be generated out of an internal determinant $|D_i\rangle$, which do not already exist in $\mathcal{S}$ or in $\{|\alpha_{j<i}\rangle\}$. Note that the internal determinants are ordered in decreasing order of their weights, i.e. $c_i^2 \geq c_{i+1}^2$. As it is done for the external determinantal space $\{|\alpha\rangle\}$, one can also decompose the total PT2 contribution of $\{|\alpha\rangle\}$ into the sum of PT2 contributions corresponding to the different subspaces $\{|\alpha_i\rangle\}$

$$E_{\mathrm{PT2}} = \sum_{i=1}^{N_{\mathrm{det}}} e_i, \tag{2.26}$$

where

$$e_i = \sum_{|\alpha\rangle \in \{|\alpha_i\rangle\}} \frac{\left| \langle \alpha | \hat{\mathcal{H}}_{\mathrm{el}} | \Psi^{(n)} \rangle \right|^2}{E^{(n)} - \langle \alpha | \hat{\mathcal{H}}_{\mathrm{el}} | \alpha \rangle}, \tag{2.27}$$

and $\Psi^{(n)}$ is the wave function at the nth iteration of the CIPSI selection. Because of the ordering mentioned above, we have that the $\{|\alpha\rangle\} \in \{|\alpha_i\rangle\}$ are not connected to any $|D_{j<i}\rangle$ with a non-zero coefficient. It follows that, by introducing a subspace wave function

$$\Psi_i^{(n)} = \sum_{m=i}^{N_{\mathrm{det}}} c_m |D_m\rangle, \tag{2.28}$$

we can simplify Eq. 2.27 as

$$e_i = \sum_{|\alpha\rangle \in \{|\alpha_i\rangle\}} \frac{|\langle\alpha|\hat{\mathcal{H}}_{\text{el}}|\Psi_i^{(n)}\rangle|^2}{E_{\text{var}}^{(n)} - \langle\alpha|\hat{\mathcal{H}}_{\text{el}}|\alpha\rangle}. \tag{2.29}$$

In the same way, we can now rewrite the total PT2 contribution (Eq. 2.26) as

$$E_{\text{PT2}} = \sum_{i=1}^{N_{\text{det}}} \sum_{|\alpha\rangle \in \{|\alpha_i\rangle\}} \frac{|\langle\alpha|\hat{\mathcal{H}}_{\text{el}}|\Psi_i^{(n)}\rangle|^2}{E_{\text{var}}^{(n)} - \langle\alpha|\hat{\mathcal{H}}_{\text{el}}|\alpha\rangle}. \tag{2.30}$$

It is good to note that, given the way in which the $\{|\alpha_i\rangle\}$ are constructed, the absolute value of the e_i rapidly decays for increasing values of i [26]. Moreover, with increasing values of i, the number of determinants involved in the sum of Eq. 2.29 decreases, making the computation of e_i faster. Furthermore, the magnitude of the denominator of Eq. 2.29 increases (with increasing i) and the norm of the wave function $\Psi_i^{(n)}$ decreases.

To get the unbiased estimates of E_{PT2}, we want to compute the e_i contributions through a MC stochastic sampling. With this aim, we can rewrite the expression for the PT2 energy as

$$E_{\text{PT2}} = \sum_{i=1}^{N_{\text{det}}} e_i = \sum_{i=1}^{N_{\text{det}}} p_i \left(\frac{e_i}{p_i}\right) = \left\langle \frac{e_i}{p_i} \right\rangle_{p_i}, \tag{2.31}$$

where p_i is an arbitrary probability distribution. The optimal choice of p_i, given by a zero-variance condition, is

$$p_i^{\text{opt}} = \frac{e_i}{E_{\text{PT2}}}. \tag{2.32}$$

Starting from this formula, Garniron *et al.* note that the magnitude of e_i is related to the norm of the truncated wave function $\Psi_i^{(n)}$ [23, 26], so they choose the probability distribution p_i as

$$p_i = \frac{\langle\Psi_i^{(n)}|\Psi_i^{(n)}\rangle}{\sum_{j=1}^{N_{\text{det}}} \langle\Psi_j^{(n)}|\Psi_j^{(n)}\rangle} = \frac{\sum_{m=i}^{N_{\text{det}}} c_m^2}{\sum_{j=1}^{N_{\text{det}}} \sum_{k=j}^{N_{\text{det}}} c_k^2}. \tag{2.33}$$

Finally, in the code implementation, they restrict the numerator to the leading coefficient and reduce the denominator into a simpler form using as p_i the expression

$$p_i = \frac{c_i^2}{\sum_{j=1}^{N_{\text{det}}} c_j^2}, \tag{2.34}$$

which, according to the simulations of Ref. [26], does not significantly change the results compared to using Eq. 2.33.

To summarize, one proceeds by sampling the probability distribution p_i, drawing the corresponding internal determinant $|D_i\rangle$, and generating all the *unique* single and double excitations. At this point, one computes the value e_i/p_i and adds it to the

average in order to estimate E_{PT2}. In practice, the value of e_i is calculated only the first time a $|D_i\rangle$ is drawn; it is then stored in memory for the following times that the same $|D_i\rangle$ is drawn. Indeed, since the computation of e_i is the most expensive part, computing them only once makes the algorithm extremely efficient. Furthermore, the first few determinants $|D_i\rangle$ (smallest i) are removed from the stochastic sampling process since they are often responsible for the most significant contributions to E_{PT2}. The contributions computed from external determinants that are connected to the dominant ones (the one removed from the stochastic sampling) are stored as a deterministic component of E_{PT2}. The remaining contributions are stochastically sampled as previously discussed. In practice, we follow a hybrid stochastic-deterministic scheme, where, starting from the deterministic component, at every MC step, the average is updated by adding the e_i contribution of all the newly drawn generators while keeping track of all previous draws.

2.2.5 Quantum Monte Carlo

Quantum Monte Carlo (QMC) methods are a large class of electronic structure techniques which solve the Schrödinger equation stochastically. In this dissertation, we focus on two of the most commonly used variants of real-space QMC, namely, variational (VMC) and diffusion Monte Carlo (DMC). In the following, we will briefly describe these two methods and refer to Refs. [27–30] for a more in-depth description.

Variational Monte Carlo

Let us assume that we want to compute a quantum mechanical observable, namely, the expectation value of an operator $\hat{O}$ on a given wave function. In the spatial representation, we have

$$\langle \hat{O} \rangle = \frac{\langle \psi | \hat{O} | \psi \rangle}{\langle \psi | \psi \rangle} = \frac{\int d\mathbf{r}\, \psi^*(\mathbf{r}) \hat{O} \psi(\mathbf{r})}{\int d\mathbf{r}\, \psi^*(\mathbf{r}) \psi(\mathbf{r})} \,. \tag{2.35}$$

where $\mathbf{r}$ labels here the $3N$ electronic spatial coordinates. Monte Carlo integration can now be used as an effective method to estimate this high-dimensional integral.

To understand how to proceed, we need to further manipulate the integral and we do so for the computation of the energy but the approach holds also for other operators. In particular, we rewrite the expectation value of the Hamiltonian as

$$
\begin{aligned}
E_V = \langle \hat{\mathcal{H}}_{\mathrm{el}} \rangle &= \frac{\int d\mathbf{r}\, \psi^*(\mathbf{r}) \hat{\mathcal{H}}_{\mathrm{el}} \psi(\mathbf{r})}{\int d\mathbf{r}\, \psi^*(\mathbf{r}) \psi(\mathbf{r})} \\
&= \int d\mathbf{r}\, \frac{|\psi(\mathbf{r})|^2}{\int d\mathbf{r}\, |\psi(\mathbf{r})|^2} \frac{\hat{\mathcal{H}}_{\mathrm{el}} \psi(\mathbf{r})}{\psi(\mathbf{r})} \\
&= \int d\mathbf{r}\, \rho(\mathbf{r}) E_L(\mathbf{r}) \,,
\end{aligned}
\tag{2.36}
$$

where ρ is the normalized square of the wave function and the local energy, E_L, is the result of the action of the Hamiltonian on the wave function divided by the wave function.

Since ρ is positive and integrates to one, we can interpreted it as a probability distribution and use the Metropolis algorithm to sample M electronic coordinates, $\{\mathbf{r}^m\}$, that are asymptotically distributed as ρ. The energy can then be estimated as the average of the local energies computed on the sampled configurations,

$$E_V = \langle E_L \rangle_\rho \approx \frac{1}{M} \sum_{m=1}^{M} E_L(\mathbf{r}^m)\,. \tag{2.37}$$

In VMC, the energy like any other observable is therefore estimated with a statistical error. For the statistical error to be meaningful, the wave function ψ should yield a finite variance σ^2[1]. In the case of the variational energy, the variance has the form

$$\sigma_V^2 = \frac{\langle \psi | (\hat{\mathcal{H}}_{\text{el}} - E_V)^2 | \psi \rangle}{\langle \psi | \psi \rangle} = \langle (E_L(\mathbf{r}) - E_V)^2 \rangle_\rho\,, \tag{2.38}$$

and the error as a function of the number of MC steps behaves as

$$\text{Err}(E_V) \sim \frac{\sigma_V}{\sqrt{M}}\,. \tag{2.39}$$

In practice, if we want to approximate the energy of an eigenstate of the Hamiltonian, there are two sources of errors: i) the systematic error coming from the use of an approximate wave function and ii) the statistical uncertainty $\text{Err}(E_V)$.

In VMC, when ψ approaches an eigenstate of the Hamiltonian, the local energy E_L becomes constant and equal to the correspondent eigenvalue of $\hat{\mathcal{H}}_{\text{el}}$. Consequently, the energy becomes the exact eigenvalue energy and the variance goes to zero. This property, called the *zero-variance principle*, implies that, as the wave function improves, the MC estimate of the energy converges more rapidly with the number (M) of configuration sampled. For this reason, a great effort in VMC is put into improving the input wave function used.

Chapters 3 and 4 discuss the choices we make to construct accurate VMC wave functions for ground and excited states, and how to optimize them. Here, we briefly introduce the Jastrow-Slater functional form typically adopted for the wave function,

$$\psi(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \mathcal{J}(\mathbf{r}_1, \ldots, \mathbf{r}_N) \sum_k d_k D_k^\uparrow(\mathbf{r}_1, \ldots, \mathbf{r}_{N_\uparrow}) D_k^\downarrow(\mathbf{r}_{N_\uparrow+1}, \ldots, \mathbf{r}_N)\,, \tag{2.40}$$

where $(\mathbf{r}_1, \ldots, \mathbf{r}_N)$ explicitly denote the $3N$ coordinates of the electrons. The determinants $D_k^\uparrow$ and $D_k^\downarrow$ are spin-assigned (up and down, respectively) Slater determinants of single-particle orbitals [31, 32].

---

[1] If the variance is infinite but the expectation value of the quantity of interest is finite, the estimate of the integral will still converge for the law of big numbers but the statistical uncertainty will decrease slower than $1/\sqrt{M}$.

The Jastrow factor, $\mathcal{J}$, is here an exponential function of polynomials of the inter-particle distances and accounts for so-called dynamical correlation. It is also used to impose Kato's cusp conditions at the coalescence points of the particles, where the inter-particle distance goes to zero and the potential diverges. In particular, the functional form employed in this work is the following product,

$$\mathcal{J} = \prod_{I,i} e^{A(r_{iI})} \times \prod_{i<j} e^{B(r_{ij})} \times \prod_{I,i<j} e^{C(r_{iI}, r_{jI}, r_{ij})} , \qquad (2.41)$$

where lower-case indices label electrons and the upper-case ones the nuclei. The A, B, and C terms account for electron-nucleus, electron-electron, and electron-electron-nucleus correlations, respectively. In this thesis, we mainly restrict the description to a 2-body interaction, namely, A and B terms, keeping the 3-body C part equal to zero and only using it for testing purposes.

While we keep the form of the Jastrow factor fixed, we explore here how to construct the determinantal part of the Jastrow-Slater wave function, maintaining a compromise between accuracy and computational effort. When studying excited states, it is also important to maintain a balanced description between the states of interest to ensure a good cancellation of errors and an accurate computation of excitation energies. Here, we typically employ wave functions generated in a CASSCF calculation or via a CIPSI selection as discussed in Chapter 4

In summary, in VMC, given a wave function ψ, we estimate the expectation value of an operator on that wave function by MC integration. Since we are interested in the eigenstates/eigenvalues of the Hamiltonian, we try to construct a wave function as close as possible to the desired state, and improve it by optimizing its parameters. In Chapter 3, we further discuss the use of different variational principles to optimize ground and excited states, namely, variance versus energy minimization.

Diffusion Monte Carlo

The other variant of real-space QMC we employ here is diffusion Monte Carlo. DMC is a so-called projection method where, starting from a given wave function, one projects a better approximation by stochastically applying an operator which inverts the spectrum of the Hamiltonian. In DMC, one uses the following projection operator

$$e^{-t(\hat{\mathcal{H}}_{el} - E_T)} , \qquad (2.42)$$

and, starting for instance from the best wave function obtained in VMC, one iteratively project the ground-state solution ψ_0 of the Hamiltonian.

To demonstrate this, let us expand the initial wave function, ψ, on the basis of the eigenstates of the Hamiltonian $\{\psi_i\}$ with eigenvalues $\{E_i\}$ and apply the operator (Eq. 2.42). In the limit of large t, we have

$$\lim_{t \to \infty} e^{-t(\hat{\mathcal{H}}_{el} - E_T)} |\psi\rangle = \lim_{n \to \infty} \sum_i e^{-t(E_i - E_T)} |\psi_i\rangle \langle \psi_i | \psi \rangle$$

$$\approx e^{-t(E_0 - E_T)} |\psi_0\rangle \langle \psi_0 | \psi \rangle , \qquad (2.43)$$

since the excited states decay faster than the ground state, having energies $E_i > E_0$. If the trial energy, E_T, is adjusted so that $E_T \approx E_0$, we avoid the exponential decay of the signal and obtain the ground-state solution. We note that ψ_0 survives only if the starting wave function has non-zero overlap with the exact ground state.

To understand how to perform this projection stochastically, we express the action of the projection operator in the position representation as

$$\psi(\mathbf{r}', t + \tau) = \int d\mathbf{r}\, G(\mathbf{r}'|\mathbf{r}; \tau)\psi(\mathbf{r}, t)\,, \tag{2.44}$$

where

$$G(\mathbf{r}'|\mathbf{r}; \tau) = \langle \mathbf{r}'|e^{-\tau(\hat{\mathcal{H}}_{\mathrm{el}} - E_T)}|\mathbf{r}\rangle \tag{2.45}$$

is the Green function that describes the imaginary-time propagation from $|\mathbf{r}\rangle$ to $|\mathbf{r}\rangle'$. In practice, we employ the Trotter approximation to factorize the kinetic and potential operators, and write the projection operator for small time-step τ as

$$e^{-\tau\hat{\mathcal{H}}_{\mathrm{el}}} = e^{-\tau(\hat{T} + \hat{V})} = e^{-\tau\hat{V}/2}e^{-\tau\hat{T}}e^{-\tau\hat{V}/2} + O(\tau^3)\,. \tag{2.46}$$

The imaginary-time Green function in the *short-time approximation* is then given by

$$\begin{aligned} G(\mathbf{r}'|\mathbf{r}; \tau) &\approx& \frac{1}{(2\pi\tau)^{3N/2}} \exp\left[-\frac{(\mathbf{r} - \mathbf{r}')^2}{2\tau} \right] \\ &\times& \exp\left[-\left(\frac{V(\mathbf{r}') + V(\mathbf{r})}{2} - E_T \right)\tau \right]. \end{aligned} \tag{2.47}$$

We stress that the repeated application of the short-time Green function yields a solution which is affected by a time-step error. Consequently, in DMC, one needs to perform calculations for different time-steps and extrapolate the results for $\tau \to 0$.

In a DMC simulation, one starts by sampling a set of configurations in $3N$-dimensions, the so-called walkers, distributed according to the initial trial wave function and performs the DMC projection by evolving the walkers through a diffusion and a branching process, namely, by sampling the Gaussian distribution and using the exponential factor in the approximate Green function (Eq. 2.47) as a re-weight. The basic DMC algorithm can be summarized as follows

1. Generate M_0 walkers by sampling the trial wave function ψ with the Metropolis algorithm.

2. Sample a $3N$-dimensional Gaussian distribution, $G(\xi)$, with standard deviation $\sqrt{\tau}$ and diffuse each walker as $\mathbf{r}' = \mathbf{r} + \xi$.

3. For each walker, compute the factor

$$p = \exp\left[-\left(\frac{V(\mathbf{r}) + V(\mathbf{r}')}{2} - E_T \right)\tau \right], \tag{2.48}$$

which we interpret as its probability of survival to the next iteration: if $p < 1$, the walker survives with probability p; if $p > 1$, the walker continues and new walkers are created at the same configuration with probability $p - 1$.

The branching process will increase the walker population in regions where the potential is higher than the trial energy and suppress it where the potential is lower.

4. Adjust E_T so that the number of walkers remains roughly equal to the initial value M_0.

In practice, this simple algorithm is highly inefficient since the potential can vary significantly and also diverge at the particle coalescence points, leading to large fluctuations in the walker population.

To overcome this problem, we introduce importance sampling and rewrite the integral in Eq. 2.44 in terms of the probability distribution $f(\mathbf{r}, t) = \psi(\mathbf{r})\psi(\mathbf{r}, t)$ which is a product of the trial wave function and the projected solution. It is easy to show that f satisfies

$$f(\mathbf{r}', t + \tau) = \int d\mathbf{r}\, \tilde{G}(\mathbf{r}'|\mathbf{r}; \tau) f(\mathbf{r}, t)\,, \tag{2.49}$$

where the so-called importance-sampled Green function is given by

$$\tilde{G}(\mathbf{r}'|\mathbf{r}; \tau) = \psi(\mathbf{r}')\langle\mathbf{r}'|e^{-\tau(\hat{\mathcal{H}}_{\mathrm{el}} - E_{\mathrm{T}})}|\mathbf{r}\rangle/\psi(\mathbf{r})\,. \tag{2.50}$$

In the short-time approximation, this Green function becomes

$$\begin{aligned}\tilde{G}(\mathbf{r}'|\mathbf{r}; \tau) &\approx \frac{1}{(2\pi\tau)^{3N/2}} \exp\left[-\frac{(\mathbf{r}' - \mathbf{r} - \mathbf{v}(\mathbf{r})\tau)^2}{2\tau}\right] \\ &\times \exp\left[-\left(\frac{E_L(\mathbf{r}') + E_L(\mathbf{r})}{2} - E_T\right)\tau\right]\,.\end{aligned} \tag{2.51}$$

where $\mathbf{v}(\mathbf{r}) = \nabla\psi(\mathbf{r})/\psi(\mathbf{r})$ is the drift velocity and the local energy $E_L(\mathbf{r})$ now enters the re-weighting factor. This imaginary-time propagator describes a drift-diffusion-branching process with the drift velocity $\mathbf{v}(\mathbf{r})$ pushing the walkers to regions of the space where the wave function is significant. Furthermore, the local energy substitutes the potential V in the re-weighting factor and, for a good trial wave function ψ, this guarantees smaller fluctuations (e.g. imposing the cusp conditions ensures that the local energy does not diverge at the coalescence points). In fact, in the limit of ψ being an eigenstate of the Hamiltonian, E_L becomes a constant and the walkers simply drift and diffuse, being distributed in the small time-step limit as ψ^2.

With this modified Green function, the DMC algorithm is stable and can be used to treat bosonic systems. However, for electrons, we need to introduce additional approximations to address what is known as the *fermionic sign problem*: since the bosonic ground state has a lower energy than the fermionic one, projecting the wave function without any constraint will ultimately yield the bosonic state. To prevent

the solution from collapsing on the bosonic ground state, we employ the so-called *fixed-node approximation* in which one imposes that the solution has the same nodes as the trial wave function. In the simple DMC algorithm, one would need to kill the walkers crossing the nodes but, when the drift-diffusion-branching Green function is employed, the fixed-node approximation is automatically satisfied in the limit of $\tau \to 0$ since the drift velocity pushes the walkers away from the nodes. The fixed-node solution is exact only if the nodes of the trial wave function are exact. In general, the DMC algorithm gives an upper bound to the exact energy.

In conclusion, even though DMC projects a better approximation starting from a given trial wave function, a bias remains in the results because of the fixed-node approximation. For this reason, we typically invest significant effort to construct flexible wave functions and fully optimize them in VMC before carrying out a DMC calculation. In this thesis, we also perform DMC calculations to compute excited state energies. In most of the cases presented here, the excited state is a ground state in its own symmetry class, so we are ensured that the resulting energy is variational. In Chapter 3, we also use DMC to compute the energy of excited states with the same symmetry as the ground state. In this case, all what we know is that, if the nodes are exact, we obtain the exact solution and the variationality of the energy is not guaranteed. In our studies, we find that, unless the trial wave function is really poor, DMC approaches the exact excited-state solution from above.

2.3   Solution for the nuclei

Part of this thesis aims at extending the use of QMC forces to the study of dynamical processes. Employing stochastic forces to drive the dynamics generates a series of challenges which we discuss in Chapters 5 and 6. Here, we shortly introduce the equations used to solve Newton equations and discuss Langevin dynamics. We note that we always treat the nuclei as classical and work in the Born-Oppenheimer approximation.

2.3.1   Molecular dynamics

The objective of a molecular dynamics (MD) calculation is to find the solution $x(t)$ to the Newton equation of motion

$$m\frac{d^2x}{dt^2} = F^C \,, \tag{2.52}$$

where x is the position of a classical particle (a nucleus), m its mass, and F^C the conservative force applied to the system. In our case, we would like to use the force computed from the solution of the electronic problem.

In this work, we use the so-called velocity-Verlet algorithm to evolve the system [33–35]. We start by expanding the velocities to second order as

$$v(t + \Delta t) = v(t) + \frac{F^C(t)}{m}\Delta t + \frac{\ddot{v}}{2}\Delta t^2 + O(\Delta t^3) \,. \tag{2.53}$$

If we also expand $\dot{v}(t + \Delta t)$ to first order and solve for $\ddot{v}$, we obtain

$$\frac{\ddot{v}}{2}\Delta t^2 = \frac{\Delta t}{2m}(F^C(t + \Delta t) - F^C(t)) + O(\Delta t^3),$$

where we used that $\dot{v} = F^C(t)/m$. Inserting this equation in Eq 2.53, we obtain

$$v(t + \Delta t) = v(t) + \frac{\Delta t}{2m}(F^C(t + \Delta t) + F^C(t)) + O(\Delta t^3). \tag{2.54}$$

If we expand also the position to second order, the velocity-Verlet equations become

$$
\begin{aligned}
x(t + \Delta t) &= x(t) + v(t)\Delta t + \frac{F^C(t)}{2m}\Delta t^2 + O(\Delta t^3) \\
v(t + \Delta t) &= v(t) + \frac{F^C(t) + F^C(t + \Delta t)}{2m}\Delta t + O(\Delta t^3).
\end{aligned} \tag{2.55}
$$

The standard implementation of the algorithm follows four steps:

1. Compute the half-step velocities $v(t + \Delta t/2) = v(t) + \frac{\Delta t}{2m}F^C(t)$.

2. Update the positions as $x(t + \Delta t) = x(t) + v(t + \Delta t/2)\Delta t$.

3. Compute the force $F^C(t + \Delta t)$ at the new configuration.

4. Compute the full-step velocities $v(t + \Delta) = v(t + \Delta t/2) + \frac{\Delta t}{2m}F^C(t + \Delta t)$.

To prepare the system, one only needs to input a starting configuration and initialize the velocities by sampling the Maxwell-Boltzmann at a specific temperature.

2.3.2 Langevin dynamics

Molecular dynamics simulations are widely employed to study both dynamical and thermodynamical properties of very diverse systems. However, if the system of interest is embedded in a complex environment, the treatment of all the particles involved might become computationally unfeasible. To address this problem, one can for instance rely on Langevin dynamics which is a method where the effect of the environment is included in the description of the system in an effective manner.

Langevin dynamics was first introduced through empirical considerations but it is possible to derive the equations starting from the Newtonian description of all the degrees of freedom of a system and averaging out the ones relative to the bath. We follow here such a derivation of Langevin equations, which we obtain in a generalized form (Eq. 2.73) less commonly encountered in the literature but which we will use in Chapter 5.

We start from a system which comprises two different sets of degrees of freedom, namely, $\{R_k, P_k\}$ for the N particles whose motion we want to describe, and $\{r_k, p_k\}$ for the n particle that we consider as a bath. For a classical Hamiltonian H that

describes the motion of the full system, we can write the time evolution of a quantity Q via the Louvillian operator $\hat{\mathbf{L}}$ as

$$\frac{dQ}{dt} = i\hat{\mathbf{L}}_{R,P}Q + i\hat{\mathbf{L}}_{r,p}Q = i\hat{\mathbf{L}}Q \,, \tag{2.56}$$

where $\hat{\mathbf{L}}_{r,p}$ is defined as

$$i\hat{\mathbf{L}}_{r,p} = \sum_{k=0}^{n} \frac{\partial H}{\partial p_k}\frac{\partial}{\partial r_k} - \frac{\partial H}{\partial r_k}\frac{\partial}{\partial p_k} \,, \tag{2.57}$$

and $\hat{\mathbf{L}}_{R,P}$ is similarly defined. By Taylor expanding $Q(t)$ at $t=0$, we obtain

$$Q(t) = \sum_k \frac{t^k}{k!}\frac{d^k}{dt^k}Q(t)\Big|_{t=0} = \sum_k \frac{t^k}{k!}(i\hat{\mathbf{L}})^k Q(0) = e^{i\hat{\mathbf{L}}t}Q(0) \,, \tag{2.58}$$

and, if we use this formula to describe the evolution of the momentum $P_k(\mathrm{t})$, we have for the force on particle k

$$\frac{dP_k}{dt} = i\hat{\mathbf{L}}P_k(t) = e^{i\hat{\mathbf{L}}t}i\hat{\mathbf{L}}P_k(0) \,. \tag{2.59}$$

The objective now is to average out the degrees of freedom of the bath and follow the evolution of the particle of interest. To do so, we rewrite the equation as

$$\frac{dP_k}{dt} = e^{i\hat{\mathbf{L}}t}\langle i\hat{\mathbf{L}}P_k \rangle_B + e^{i\hat{\mathbf{L}}t}\left(i\hat{\mathbf{L}}P_k - \langle i\hat{\mathbf{L}}P_k \rangle_B \right) \,, \tag{2.60}$$

where the subscript B indicates the average over all the possible initial conditions of the bath. The equation contains therefore two terms, an average force $\langle i\hat{\mathbf{L}}P_k \rangle_B$ and a fluctuating term, namely the force minus the average.

Let us analyze the first term, namely, the average force. Since we are averaging with respect to all initial configurations of the bath, the term relative to $\hat{\mathbf{L}}_{r,p}$ is zero. Moreover, R and P are independent, so the Louvillan term with the derivative $\partial P_k / \partial R_k$ is also zero. In conclusion, the average force is given by

$$\langle i\hat{\mathbf{L}}P_k \rangle_B = \left\langle -\frac{\partial H}{\partial R_k} \right\rangle_B = \left\langle -\frac{\partial \phi}{\partial R_k} \right\rangle_B \,, \tag{2.61}$$

where we assume that the potential, ϕ, is the only term of the Hamiltonian explicitly depending on R_k. In particular, ϕ depends both on R and r but the dependence on r drops out since we are averaging over the bath coordinates. As next step, we rewrite the average force in terms of the free energy of the system, A, as

$$\left\langle -\frac{\partial \phi}{\partial R_k} \right\rangle_B = -\frac{\partial A(R)}{\partial R_k} \,, \tag{2.62}$$

where

$$A(R) = -k_B T \int dr \int dp \, e^{-\beta H(r,q,R,P)} \,, \tag{2.63}$$

with $\beta = 1/k_B T$ and k_B the Boltzmann constant. The integral in the definition of A runs over all the positions and momenta of the bath.

Finally, we obtain the average force given by the first term in Eq. 2.60 as

$$e^{i\hat{\mathbf{L}}t}\langle i\hat{\mathbf{L}}P_k\rangle_B = -e^{i\hat{\mathbf{L}}t}\frac{\partial A(R)}{\partial R_k} = -\frac{\partial A(R;t)}{\partial R_k}\,, \qquad (2.64)$$

where we use that also the free energy evolves in time under the action of the Louvillian operator as given in Eq. 2.58.

We now focus on the fluctuating contribution to the force in Eq. 2.60. To this aim, we first introduce the projector notation for the average of a quantity Q over the bath B

$$\langle Q\rangle_B = \wp Q\,, \qquad (2.65)$$

where $\wp^2 = \wp$ and, consequently, $(1 - \wp)^2 = (1 - \wp)$ and $\wp(1 - \wp) = (1 - \wp)\wp = 0$. With this notation, the fluctuating term of Eq. 2.60 can be written as

$$e^{i\hat{\mathbf{L}}t}\left(i\hat{\mathbf{L}}P_k - \langle i\hat{\mathbf{L}}P_k\rangle_B\right) = e^{i\hat{\mathbf{L}}t}(1 - \wp)i\hat{\mathbf{L}}P_k\,. \qquad (2.66)$$

We first note that, since $\wp(1 - \wp) = 0$, the average on the bath of the fluctuating force is zero,

$$\wp(1 - \wp)i\hat{\mathbf{L}}P_k = 0\,, \qquad (2.67)$$

but this does not hold for the average of the time evolution of the fluctuating force, so

$$\wp(e^{i\hat{\mathbf{L}}t}(1 - \wp)i\hat{\mathbf{L}}P_k) \neq 0\,. \qquad (2.68)$$

In order to gain some understanding on the fluctuating term, we use the identity

$$\begin{aligned} e^{i\hat{\mathbf{L}}t}(1 - \wp)i\hat{\mathbf{L}}P_k &= e^{(1-\wp)i\hat{\mathbf{L}}t}(1 - \wp)i\hat{\mathbf{L}}P_k \\ &+ \int_0^t d\tau e^{i\hat{\mathbf{L}}(t-\tau)}\wp i\hat{\mathbf{L}}e^{(1-\wp)i\hat{\mathbf{L}}\tau}(1 - \wp)i\hat{\mathbf{L}}P_k\,, \end{aligned} \qquad (2.69)$$

which is proven in Refs. [36, 37].

We now define

$$F_k^R \equiv (1 - \wp)i\hat{\mathbf{L}}P_k \quad \text{and} \quad F_{k,t}^R = e^{(1-\wp)i\hat{\mathbf{L}}t}(1 - \wp)i\hat{\mathbf{L}}P_k\,, \qquad (2.70)$$

where we use the superscript R to indicate that, as a random force, its average is zero. Then, Eq. 2.69 becomes

$$e^{i\hat{\mathbf{L}}t}(1 - \wp)i\hat{\mathbf{L}}P_k = F_{k,t}^R + \int_0^t d\tau e^{i\hat{\mathbf{L}}(t-\tau)}\wp i\hat{\mathbf{L}}F_{k,\tau}^R\,. \qquad (2.71)$$

With further manipulations [38] , one can approximate the integral as

$$\int_0^t d\tau e^{i\hat{\mathbf{L}}(t-\tau)} \wp i\hat{\mathbf{L}} F_{k,\tau}^R \approx -\beta \int_0^t d\tau \sum_j \frac{P_j(t-\tau)}{M} \langle F_j^R F_{k,\tau}^R \rangle_B (t-\tau) . \quad (2.72)$$

where we assume for simplicity that all particles have mass M.

In conclusion, inserting our findings for the average and random components in Eq. 2.60, we have

$$\frac{dP_k}{dt} = -\frac{\partial A(R;t)}{\partial R_k} + F_{k,t}^R - \beta \int_0^t d\tau \sum_j \frac{P_j(\tau)}{M} \langle F_j^R F_{k,t-\tau}^R \rangle_B (\tau) , \quad (2.73)$$

where the first term is a conservative force and the second a random force with $\langle F_{k,t}^R \rangle_B = 0$. The last term contains the contribution of the velocities $v_j = P_j/M$ at previous times through the correlation of the random forces at different times.

Typically, one assumes that the momenta barely change in the time interval in which the random forces decorrelate. If this approximation holds, we can bring the momenta out of the integral and simplify the expression as

$$\frac{dP_k}{dt} = -\frac{\partial A(R;t)}{\partial R_k} + F_{k,t}^R - \sum_j \frac{P_j(t)}{M} \gamma_{j,k}(t) . \quad (2.74)$$

Finally, comparing the last two equations, we see that the parameter $\gamma_{j,k}(t)$ satisfies

$$\langle F_j^R F_{k,t-\tau}^R \rangle_B (\tau) = 2k_B T \gamma_{j,k}(t) \delta(t-\tau) , \quad (2.75)$$

which is the fluctuation-dissipation theorem and allows us to physically interpret the friction as a time correlation at the microscopic level.

While this is the general Langevin formulation, the random contributions of the various degrees of freedom are considered to be uncorrelated so that $\gamma_{j,k}(t) = \gamma(t)\delta_{j,k}$. In this way, we recover the more familiar expression of the Langevin equation,

$$\frac{dP_k}{dt} = -\frac{\partial A(R;t)}{\partial R_k} + F_{k,t}^R - \frac{P_k}{M} \gamma(t) . \quad (2.76)$$

In this dissertation, we mainly consider the last expression but, in Chapter 5, we deal with random forces with memory so that we strat from an equation more similar to Eq. 2.73.

2.4   Codes used in this thesis

We list here the codes employed to perform the simulations in this thesis and give a more in-depth description of the computational details in each Chapter.

- `CHAMP` [39] is a quantum Monte Carlo code, which we use to perform total energy calculations and obtain the forces for molecular dynamics. The code can perform both VMC and DMC calculations with wave functions of the Jastrow-Slater form.

- `Tinker` [40] is a molecular dynamics package. In this thesis, we modify this code to mimic a molecular dynamics simulation in which the forces have noise and to quickly explore possible solutions.

- `GAMESS` [41, 42] is a quantum chemistry code which we use to generate the starting wave function for our QMC calculations in either a Hartree-Fock or a CASSCF calculation.

- `Gaussian09` [43] is an electronic structure program that we eploy to compute optimized geometries with various DFT functionals as well as time-dependent (TD)DFT excitation energies.

- `QP2` [24] is a code which can perform CIPSI calculations. We use the code both to determine extrapolate FCI (exFCI) estimates of excitation energies and to obtain determinantal expansions for QMC.

- `Molcas` [44] is another quantum chemistry code specialized in multiconfigurational methods. We employ it to perform CASPT2 calculations.

- `PSi4` [45] is a coupled cluster code that we use for iterative approximate coupled cluster singles, doubles, and triples (CC3) calculations.

- `CFOUR` [46] is also a coupled cluster code that we employ for CC3 calculations for big molecules since the code is parallelized.

- `VMD` [47] is a program which we use to visualize trajectories from molecular dynamics and extract bond information.

Bibliography

[1] M. Born, K. Huang, and M. Lax, Am. J. Phys. **23**, 474 (1955).

[2] W. Domcke, D. Yarkony, and H. Köppel, *Conical intersections: electronic structure, dynamics & spectroscopy* (World Scientific, 2004), Vol. 15.

[3] M. Born and R. Oppenheimer, Ann. Phys. **389**, 457 (1927).

[4] B. F. E. Curchod and T. J. Martínez, Chem. Rev. **118**, 3305 (2018).

[5] L. M. Ibele, A. Nicolson, and B. F. E. Curchod, Mol. Phys. **118**, e1665199 (2020).

[6] J. C. Tully, J. Chem. Phys. **93**, 1061 (1990).

[7] M. Barbatti, WIREs Comput. Mol. Sci. **1**, 620 (2011).

[8] D. R. Hartree, Math. Proceed. Camb. Phil. Society **24**, 89 (1928).

[9] D. R. Hartree, Math. Proceed. Camb. Phil. Society **24**, 111 (1928).

[10] V. Fock, Z. Phys. **61**, 126 (1930).

[11] V. Fock, Z. Phys. **62**, 795 (1930).

[12] D. R. Hartree and W. Hartree, Proc. R. Soc. Lond. A **150**, 9 (1935).

[13] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, 1st ed. (Dover Publications, Inc., 1996).

[14] J. C. Slater, Phys. Rev. **34**, 1293 (1929).

[15] E. A. Hylleraas and B. Undheim, Z. Phys. **65**, 759 (1930).

[16] J. K. L. MacDonald, Phys. Rev. **43**, 830 (1933).

[17] B. O. Roos, P. R. Taylor, and P. E. Sigbahn, Chem. Phys. **48**, 157 (1980).

[18] C. Møller and M. S. Plesset, Phys. Rev. **46**, 618 (1934).

[19] K. Andersson, P. A. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski, J. Phys. Chem. **94**, 5483 (1990).

[20] K. Andersson, P. Malmqvist, and B. O. Roos, J. Chem. Phys. **96**, 1218 (1992).

[21] P. S. Epstein, Phys. Rev. **28**, 695 (1926).

[22] R. K. Nesbet and D. R. Hartree, Proc. R. Soc. Lond. A **230**, 312 (1955).

[23] Y. Garniron, *Development and parallel implementation of selected configuration interaction methods* (Université de Toulouse, 2019).

[24] Y. Garniron, T. Applencourt, K. Gasperich, A. Benali, A. Ferté, J. Paquier, B. Pradines, R. Assaraf, P. Reinhardt, J. Toulouse, P. Barbaresco, N. Renon, G. David, J.-P. Malrieu, M. Véril, M. Caffarel, P.-F. Loos, E. Giner, and A. Scemama, J. Chem. Theory Comput. **15**, 3591 (2019).

[25] M. Dash, *Accurate excited states with quantum Monte Carlo: looking beyond conventions* (University of Twente, 2020).

[26] Y. Garniron, A. Scemama, P.-F. Loos, and M. Caffarel, J. Chem. Phys. **147**, 034101 (2017).

[27] M. Nightingale and C. Umrigar, *Quantum Monte Carlo Methods in Physics and Chemistry*, *Nato Science Series C:* (Springer Netherlands, 1998).

[28] A. Lüchow, WIREs Comput. Mol. Sc. **1**, 388 (2011).

[29] J. Toulouse, R. Assaraf, and C. J. Umrigar, in *Advances in Quantum Chemistry*, Vol. 73 of *Electron Correlation in Molecules – ab initio Beyond Gaussian Quantum Chemistry*, edited by P. E. Hoggan and T. Ozdogan (Academic Press, 2016), pp. 285–314.

[30] J. Feldt and C. Filippi, in *Quantum Chemistry and Dynamics of Excited States: Methods and Applications*, edited by R. Lindh and L. Gonzalez (John Wiley and Sons Ltd, 2019).

[31] C.-J. Huang, C. Filippi, and C. J. Umrigar, J. Chem. Phys. **108**, 8838 (1998).

[32] M. Zaccheddu, *Ab initio study of the optical properties of green fluorescent protein* (Leiden University, 2008).

[33] L. Verlet, Phys. Rev. **159**, 98 (1967).

[34] M. Allen, M. Allen, D. Tildesley, and D. Tildesley, *Computer Simulation of Liquids*, *Oxford Science Publications* (Clarendon Press, 1989).

[35] D. Frenkel and B. Smit, in *Understanding Molecular Simulation (Second Edition)*, second edition ed., edited by D. Frenkel and B. Smit (Academic Press, 2002), pp. 63–107.

[36] S. Nordholm and R. Zwanzig, J. Stat. Phys. **13**, 347 (1975).

[37] R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford University Press, 2001).

[38] R. L. C. Akkermans and W. J. Briels, J. Chem. Phys. **113**, 6409 (2000).

[39] CHAMP is a quantum Monte Carlo program package written by C. J. Umrigar, C. Filippi, S. Moroni and collaborators.

[40] J. A. Rackers, Z. Wang, C. Lu, M. L. Laury, L. Lagardère, M. J. Schnieders, J.-P. Piquemal, P. Ren, and J. W. Ponder, J. Chem. Theory Comput. **14**, 5273 (2018).

[41] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery Jr, J. Comput. Chem. **14**, 1347 (1993).

[42] M. S. Gordon and M. W. Schmidt, in *Theory and Applications of Computational Chemistry: the first forty years*, edited by C. Dykstra, G. Frenking, K. Kim, and G. Scuseria (Elsevier, 2011), Chap. 41, pp. 1167–1190.

[43] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox. Gaussian09, Revision A.02; Gaussian, Inc., Wallingford CT, 2016.

[44] F. Aquilante, J. Autschbach, R. K. Carlson, L. F. Chibotaru, M. G. Delcey, L. De Vico, I. Fdez. Galván, N. Ferré, L. M. Frutos, L. Gagliardi, M. Garavelli, A. Giussani, C. E. Hoyer, G. Li Manni, H. Lischka, D. Ma, P. Å. Malmqvist, T. Müller, A. Nenov, M. Olivucci, T. B. Pedersen, D. Peng, F. Plasser, B. Pritchard, M. Reiher, I. Rivalta, I. Schapiro, J. Segarra-Martí, M. Stenrup, D. G. Truhlar, L. Ungur, A. Valentini, S. Vancoillie, V. Veryazov, V. P. Vysotskiy, O. Weingart, F. Zapata, and R. Lindh, J. Comput. Chem. **37**, 506 (2016).

[45] R. M. Parrish, L. A. Burns, D. G. A. Smith, A. C. Simmonett, A. E. DePrince, E. G. Hohenstein, U. Bozkaya, A. Y. Sokolov, R. Di Remigio, R. M. Richard, J. F. Gonthier, A. M. James, H. R. McAlexander, A. Kumar, M. Saitow, X. Wang, B. P. Pritchard, P. Verma, H. F. Schaefer, K. Patkowski, R. A. King, E. F. Valeev, F. A. Evangelista, J. M. Turney, T. D. Crawford, and C. D. Sherrill, J. Chem. Theory Comput. **13**, 3185 (2017).

[46] J. F. Stanton, J. Gauss, L. Cheng, M. E. Harding, D. A. Matthews, and P. G. Szalay, CFOUR, Coupled-Cluster techniques for Computational Chemistry, a quantum-chemical program package, With contributions from A.A. Auer, R.J. Bartlett, U. Benedikt, C. Berger, D.E. Bernholdt, S. Blaschke, Y. J. Bomble, S. Burger, O. Christiansen, D. Datta, F. Engel, R. Faber, J. Greiner, M. Heckert, O.

Heun, M. Hilgenberg, C. Huber, T.-C. Jagau, D. Jonsson, J. Jusélius, T. Kirsch, K. Klein, G.M. KopperW.J. Lauderdale, F. Lipparini, T. Metzroth, L.A. Mück, D.P. O'Neill, T. Nottoli, D.R. Price, E. Prochnow, C. Puzzarini, K. Ruud, F. Schiffmann, W. Schwalbach, C. Simmons, S. Stopkowicz, A. Tajti, J. Vázquez, F. Wang, J.D. Watts and the integral packages MOLECULE (J. Almlöf and P.R. Taylor), PROPS (P.R. Taylor), ABACUS (T. Helgaker, H.J. Aa. Jensen, P. Jørgensen, and J. Olsen), and ECP routines by A. V. Mitin and C. van Wüllen. For the current version, see http://www.cfour.de.

[47] W. Humphrey, A. Dalke, and K. Schulten, J. Mol. Graph. **14**, 33 (1996).

Chapter 3

Variational principles in QMC: the troubled story of variance minimization

†

We investigate the use of different variational principles in quantum Monte Carlo, namely energy and variance minimization, prompted by the interest in the robust and accurate estimate of electronic excited states. For two prototypical, challenging molecules, we readily reach the accuracy of the best available reference excitation energies using energy minimization in a state-specific or state-average fashion for states of different or equal symmetry, respectively. On the other hand, in variance minimization, where the use of suitable functionals is expected to target specific states regardless of the symmetry, we encounter severe problems for a variety of wave functions: as the variance converges, the energy drifts away from that of the selected state. This unexpected behavior is sometimes observed even when the target is the ground state, and generally prevents the robust estimate of total and excitation energies. We analyze this problem using a very simple wave function and infer that the optimization finds little or no barrier to escape from a local minimum or local plateau, eventually converging to a lower-variance state instead of the target state. For the increasingly complex systems becoming in reach of quantum Monte Carlo simulations, variance minimization with current functionals appears to be an impractical route.

3.1   Introduction

Light-induced processes are at the heart of a variety of phenomena and applications which range from harnessing the response to light of biological systems to improving the technologies for renewable energies. The contribution of electronic structure

theory in this field hinges on its ability to efficiently and accurately compute excited-state properties. In this context, the use of quantum Monte Carlo (QMC) methods is relatively recent and quite promising [1–9]: QMC approaches provide an accurate (stochastic) solution of the Schrödinger equation and benefit from a favorable scaling with system size and great ease of parallelization [10–12]. Importantly, recent methodological advancements [13–16] enable the fast calculation of energy derivatives and the optimization of many thousands of parameters for the internally consistent computation of QMC wave functions and geometries in the ground and excited states [9, 17].

Here, we investigate the use of two different variational principles for ground and excited states in QMC, namely, variance and energy minimization, to assess whether they allow us to fully capitalize on the increased power of minimization algorithms and availability of accurate wave functions. Variance minimization techniques [18–22] have been extensively employed in QMC for the last 30 years but their potential for the computation of excited states has only recently been revisited and exploited to compute vertical excitation energies of various small molecules [23,24]. Different functionals for the optimization of the variance [19,22,23] have also been put forward with the common attractive feature of the built-in possibility to target a specific state and avoid in principle the complications encountered in energy minimization where, without constraints, one would generally collapse to lower-energy states.

For our study, we select two molecules, a small cyanine dye and a retinal model, because of the difficulties they pose in the computation of the lowest vertical excitation energy [4, 25–28], and the different requirements in the procedure adopted in energy minimization: while the ground and excited states of the cyanine belong to different symmetries and can therefore be treated in a state-specific manner, this is not the case for the retinal model, where energy minimization must be performed in a state-average fashion. For both molecules and therefore regardless of the nature of the optimization, we find that energy minimization leads to the stable and fast convergence of the total energies of the states of interest. Furthermore, with the use of compact and balanced energy-minimized wave functions constructed through a selected configuration interaction (CI) approach, we recover vertical excitation energies which are already at the variational Monte Carlo (VMC) level within chemical accuracy (about 0.04 eV) of the reference coupled cluster or extrapolated CI values. On the other hand, for both molecules and for nearly all wave functions investigated, the optimization of all parameters in variance minimization is problematic since it results in the apparent loss of the state of interest over sufficiently long optimization runs, precluding the estimate of the excitation energy. This occurs for the different functionals originally proposed to stabilize the optimization and, surprisingly, in some cases also when targeting the ground state. This finding is unexpected, especially considering that variance minimization has been the method of choice in QMC for decades and is still routinely used, albeit for simpler systems and/or for wave functions with a small number of parameters, often limited to the Jastrow factor.

To understand these newly-found issues, we examine how variance minimization behaves when optimizing the linear coefficients of a very small CI expansion: working in the linear sub-space spanned by a few approximate eigenvectors, we discover

40

that the optimization of the CI parameters in variance minimization does not converge to the target eigenstate but to a different one. In this specific example, during the minimization, the system slowly reaches the eigenstate corresponding to the absolute minimum of the variance, no matter what the starting state is. We verify that a similar pattern explains the unexpected behavior observed for more complicated wave functions.

It is well known that the variance reaches its minimum value of zero for every exact eigenstate [19]. This is the very basis of variance minimization. Whether the variance maintains a minimum when any particular eigenstate is described by a given approximate wave function is a question that can only be assessed empirically on case by case basis. Our calculations identify missing minima in several instances of current interest for QMC simulations. Systematic improvement of the wave function to recover the zero-variance property of the exact eigenstates would be possible in principle, but impractically demanding.

Our findings pose severe limitations on the application of variance minimization for the increasingly complex systems that are becoming accessible to QMC simulations.

In Section 3.2, we recap the equations used for energy and variance optimization, discuss the procedure employed for the state-average case, and introduce the ingredients for a stable version of the Newton method in variance minimization. In Section 3.3, we summarize the computational details and, in Section 3.4, present the accurate vertical excitation energies obtained in energy minimization and the difficulties encountered in variance minimization for both molecules. We elucidate these findings and conclude in Section 3.5.

3.2 Methods

We briefly introduce below the variance and energy minimization approaches used to optimize the wave functions in variational Monte Carlo. While we employ variance minimization as a state-specific approach to target a given state, we must distinguish between a state-specific and a state-average route for energy optimization when the excited state of interest is of different or equal symmetry, respectively, than other lower-lying states.

3.2.1 Wave function form

The wave functions employed in this work are of the Jastrow-Slater type, namely, the product of a determinantal expansion and a Jastrow correlation function, $\mathcal{J}$, as

$$\Psi = \mathcal{J} \sum_{i=1}^{N_{\text{det}}} c_i D_i \,, \tag{3.1}$$

where the determinants are expressed on single-particle orbitals and the Jastrow factor includes an explicit dependence on the electron-electron distances. Here, the

41

Jastrow factor is chosen to include electron-electron and electron-nucleus correlation terms [29]. For the determinantal component, we select the relevant determinants according to different recipes: i) very simple ansatzes such as Hartree-Fock (HF) or a CI singles (CIS) expansion recently put forward as a computationally cheap and sufficiently accurate wave function for excited states in QMC [8, 30]; ii) complete-active-space (CAS) expansions where small sets of important active orbitals are manually identified; iii) CI perturbatively selected iteratively (CIPSI) expansions generated to yield automatically balanced multiple states for a fast convergence of the QMC excitation energy with the number of determinants. All expansions are expressed in terms of spin-adapted configuration state functions (CSF) to reduce the number of variational parameters.

3.2.2 Energy minimization

For state-specific optimization in energy minimization, we employ the stochastic reconfiguration (SR) method [14, 31] in a low-memory conjugate-gradient implementation [14]. Given a starting wave function Ψ depending on a set of parameters $\mathbf{p}$, we denote the derivatives of Ψ with respect to a parameter p_i as $\Psi_i = \partial_i \Psi$. At every step of the SR optimization, the parameter variations, Δp_i, are computed according to the equation:

$$\bar{\mathbf{S}}\Delta\mathbf{p} = -\tau\mathbf{g}\,,\tag{3.2}$$

where τ is a positive quantity chosen small enough to guarantee the convergence. The vector $\mathbf{g}$ is the gradient of the energy with components:

$$
\begin{aligned}
g_i = \frac{\partial E}{\partial p_i} &= 2\left[\frac{\langle\Psi_i|\hat{\mathcal{H}}|\Psi\rangle}{\langle\Psi|\Psi\rangle} - E\frac{\langle\Psi|\Psi_i\rangle}{\langle\Psi|\Psi\rangle}\right] \\
&= 2\left[\left\langle\frac{\Psi_i}{\Psi}E_{\mathrm{L}}\right\rangle - \langle E_{\mathrm{L}}\rangle\left\langle\frac{\Psi_i}{\Psi}\right\rangle\right]\,,
\end{aligned}\tag{3.3}
$$

where $E_{\mathrm{L}} = \hat{H}\Psi/\Psi$ is the so-called local energy and $\langle.\rangle$ denotes the Monte Carlo average of the quantity in brackets over the electron configurations sampled from $\Psi^2/\langle\Psi|\Psi\rangle$. The matrix $\bar{\mathbf{S}}$ has components:

$$
\begin{aligned}
\bar{S}_{ij} &= \frac{\langle\Psi_i|\Psi_j\rangle}{\langle\Psi|\Psi\rangle} - \frac{\langle\Psi|\Psi_i\rangle}{\langle\Psi|\Psi\rangle}\frac{\langle\Psi|\Psi_j\rangle}{\langle\Psi|\Psi\rangle} \\
&= \left\langle\frac{\Psi_i}{\Psi}\frac{\Psi_j}{\Psi}\right\rangle - \left\langle\frac{\Psi_i}{\Psi}\right\rangle\left\langle\frac{\Psi_j}{\Psi}\right\rangle \equiv \left\langle\frac{\bar{\Psi}_i}{\Psi}\frac{\bar{\Psi}_j}{\Psi}\right\rangle\,,
\end{aligned}\tag{3.4}
$$

which is expressed in the last equality as the overlap matrix in the semi-orthogonal basis, $\bar{\Psi}_i = \Psi_i - [\langle\Psi|\Psi_i\rangle/\langle\Psi|\Psi\rangle]\Psi$.

When the state of interest is energetically not the lowest in its symmetry class, we start from a set of wave functions for the multiple states which share the same Jastrow factor and orbitals but are characterized by different linear CI coefficients as

$$\Psi^I = \mathcal{J}\sum_{i=1}^{N_{\mathrm{det}}} c_i^I D_i\,,\tag{3.5}$$

where the superscript I indicates a particular state. To obtain a balanced description of the states of interest, we optimize the non-linear parameters of the orbitals and the Jastrow factor by minimizing the state-average energy [1]:

$$E^{\mathrm{SA}} = \sum_I w_I \frac{\langle \Psi^I | \hat{\mathcal{H}} | \Psi^I \rangle}{\langle \Psi^I | \Psi^I \rangle} \,, \tag{3.6}$$

where the weights w_I are kept fixed and $\sum_I w_I = 1$. To this aim, we follow the SR scheme (Eq. 3.2) and use the gradient of the state-average energy

$$g_i^{\mathrm{SA}} = \sum_I w_I g_i^I \,, \tag{3.7}$$

where g_i^I is the gradient with respect to a parameter p_i of the energy of state I, which is computed from the wave function Ψ^I and its derivatives as in Eq. 3.3. Moreover, in analogy to the single-state optimization, we introduce a weighted-average overlap matrix defined as

$$\bar{S}_{ij}^{\mathrm{SA}} = \sum_I w_I \bar{S}_{ij}^I \,, \tag{3.8}$$

where the overlap matrix for each state is computed from the corresponding wave function as in Eq. 3.4. We stress that, while the state-average SR procedure is defined simply by analogy with the single-state case, it employs the correct gradients of the SA energy ($\mathbf{g}^{\mathrm{SA}}$) and, therefore, at convergence, leads to the minimization of the state-average energy.

We alternate a number of optimization steps of the non-linear parameters with the optimization of the linear coefficients c_i^I, whose optimal values are the solution of the generalized eigenvalue equations

$$\mathbf{H}^{\mathrm{CI}} \mathbf{c}^I = E_I \mathbf{S}^{\mathrm{CI}} \mathbf{c}^I \,, \tag{3.9}$$

where the Hamiltonian and overlap matrix elements are defined in the basis of the functions $\{\mathcal{J} D_i\}$ and estimated through Monte Carlo sampling. After diagonalization of Eq. 3.9, orthogonality between the individual states is automatically enforced. To solve the eigenvalue equation with a memory efficient algorithm, we use the Davidson diagonalization scheme in which the lowest energy eigenvalues are computed without the explicit construction of the entire Hamiltonian and overlap matrices [14]. A similar procedure was recently followed in Ref. [32].

3.2.3 Variance minimization

To perform variance minimization, we can directly minimize the variance of the state of interest,

$$\sigma^2 = \frac{\langle \Psi | (\hat{\mathcal{H}} - E)^2 | \Psi \rangle}{\langle \Psi | \Psi \rangle} \,, \tag{3.10}$$

or follow a somewhat more stable optimization procedure by minimizing the expression

$$\sigma_\omega^2 = \frac{\langle \Psi | (\hat{\mathcal{H}} - \omega)^2 | \Psi \rangle}{\langle \Psi | \Psi \rangle} \, , \tag{3.11}$$

where the energy ω is fixed during the optimization step and then appropriately modified to follow the current value of the energy as originally proposed in Ref. [19]. Recently, a functional Ω has been put forward,

$$\Omega = \frac{\langle \Psi | (\omega - \hat{\mathcal{H}}) | \Psi \rangle}{\langle \Psi | (\omega - \hat{\mathcal{H}})^2 | \Psi \rangle} \, , \tag{3.12}$$

whose minimization is equivalent to variance minimization if ω is eventually updated to the running value of $E - \sigma$ [23].

Because of its simplicity, we choose here the functional σ_ω^2 but also compare the convergence behavior obtained with the functional Ω. To this aim, we use the Newton optimization method as in Ref. [22] and update the parameters as

$$\Delta \mathbf{p} = -\tau \mathbf{h}^{-1} \mathbf{g} \, , \tag{3.13}$$

where $\mathbf{g}$ is here the gradient of σ_ω^2 and $\mathbf{h}$ its Hessian matrix, and the parameter τ is introduced to damp the size of the variations.

The components of the gradient are given by

$$\begin{aligned} g_i = & \; 2 \left[\left\langle \frac{\hat{\mathcal{H}} \Psi_i}{\Psi} E_L \right\rangle - \left\langle \frac{\Psi_i}{\Psi} \right\rangle \langle E_L^2 \rangle \right. \\ & \left. - \; \omega \left(\left\langle \frac{\Psi_i}{\Psi} E_\mathrm{L} \right\rangle + \left\langle \frac{\hat{H} \Psi_i}{\Psi} \right\rangle - 2 \left\langle \frac{\Psi_i}{\Psi} \right\rangle \langle E_L \rangle \right) \right] \, , \end{aligned} \tag{3.14}$$

and we discuss other possible equivalent expressions and their relative fluctuations in Appendix 3.A. The Hessian matrix elements require the second derivatives of the wave function and, to avoid their computation, we follow the same approximation strategy of the Levenberg-Marquardt algorithm [33] and manipulate the expression of the variance in a somewhat different way than proposed in Refs. [20, 22, 34], to obtain the approximate expression of the Hessian matrix

$$\begin{aligned} h_{ij} = & \; \left\langle \left[\partial_i E_\mathrm{L} + (E_\mathrm{L} - \omega) \left(\frac{\Psi_i}{\Psi} - \left\langle \frac{\Psi_i}{\Psi} \right\rangle \right) \right] \right. \\ & \; \times \quad \left. \left[\partial_j E_\mathrm{L} + (E_\mathrm{L} - \omega) \left(\frac{\Psi_j}{\Psi} - \left\langle \frac{\Psi_j}{\Psi} \right\rangle \right) \right] \right\rangle \, , \end{aligned} \tag{3.15}$$

Details of the derivation and alternative expressions for the Hessian are given in Appendix 3.A.

We use the Newton method and the Hessian $\mathbf{h}$ (Eq. 3.15) when optimizing both σ_ω^2 and the Ω functional in combination with the corresponding gradient. Furthermore, we follow Ref. [23] in keeping ω fixed to an appropriate guess energy for an initial number of minimization steps, upgrading it linearly to the running energy (or $E - \sigma$ in the case of Ω) over some intermediate iteration steps, and then setting it equal to the current energy estimate for the rest of the run.

3.3 Computational details

All QMC calculations are carried out with the program package CHAMP [35]. We employ scalar-relativistic energy-consistent HF pseudopotentials and the correlation-consistent Gaussian basis sets specifically constructed for these pseudopotentials [36, 37]. Unless otherwise specified, we use a double-ζ basis set minimally augmented with s and p diffuse functions on the heavy atoms and denoted here as maug-cc-pVDZ. Basis-set convergence tests are performed with the fully augmented double (aug-cc-pvDZ) and triple (aug-cc-pvTZ) basis sets. In all cases, the exponents of the diffuse functions are taken from the corresponding all-electron Dunning's correlation-consistent basis sets [38].

In the state-specific (energy and variance) optimization runs, we sample a guiding wave function that differs from the current wave function close to the nodes [39] to guarantee finite variances of the estimators of the gradient, overlap, and Hessian matrix elements. In the state-average energy minimizations, we employ equal weights for the multiple states and sample a guiding wave function constructed as $\Psi_g^2 = \sum_I |\Psi^I|^2$, to ensure that the distribution sampled has a large overlap with all states of interest [1]. All wave function parameters (Jastrow, orbital, and CI coefficients) are optimized and the damping factor, τ, in the SR and the Newton method is set to 0.05 and 0.1, respectively, unless otherwise specified. In the DMC calculations, we treat the pseudopotentials beyond the locality approximation using the T-move algorithm [40] and employ an imaginary time-step of 0.05 a.u. which yields excitation energies converged to better than 0.01 eV.

The HF, CIS, and complete-active-space self-consistent-field (CASSCF) calculations are carried out with the program GAMESS(US) [41, 42]. For the cyanine dye, we consider different CAS expansions: a CAS(6,5) and CAS(6,10) correlating 6 π electrons in the orbitals constructed from the $2p_z$ and $3p_z$ atomic orbitals; a truncated CAS(14,13) consisting of 6 π and 8 σ electrons in 13 bonding and antibonding orbitals. For the retinal model, we employ a minimal CAS(6,6) active space of 6 π electrons in the orbitals constructed from the $2p_z$ atomic orbitals.

The CIPSI calculations are performed with Quantum Package [43] and the determinantal expansions are constructed to be eigenstates of $\hat{S}^2$. For the cyanine dye where ground and excited states have different symmetry, we follow two paths to construct the CIPSI expansions: i) we perform separate expansions for the two states starting from the corresponding CASSCF(6,10) orbitals, and match the variances of the CI wave functions to obtain a balanced description of the states. We find that this procedure leads to an automatic match of the second-order perturbation theory (PT2) energy contributions, which are an estimate of the errors of the wave functions with respect to the corresponding full CI (FCI) limit. Using expansions with matched PT2 corrections has recently been shown to lead to accurate QMC excitation energies also for a relatively small number of determinants [9]. ii) We perform the expansion of the two states simultaneously, using a common set of orbitals (the excited-state CASSCF(6,10) orbitals), and obtain automatically matched PT2 energy corrections during the expansion [9]. For the retinal model where the ground and excited states have the same symmetry, we have only one set of orbitals for the
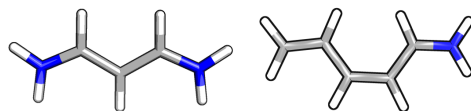
Figure 3.1: Schematic representations of the CN5 (left) and PSB3 (right) molecules. White, gray, and blue denote hydrogen, carbon, and nitrogen, respectively.

CIPSI expansions. In this case, we perform a simultaneous expansion with a selection scheme that matches the CI variances and also attempts to balance the PT2 energy contributions of the two states [44].

All total energies are computed on the PBE0/cc-pVQZ ground-state geometry of the cyanine [45] and retinal molecules. The DFT geometry optimization of the retinal model is performed with the program Gaussian [46]. The coupled cluster results are obtained with Psi4 [47].

3.4 Results

We compute the lowest $\pi \rightarrow \pi^*$ vertical excitation energy of the cyanine dye $(C_3H_3(NH_2)_2^+)$ and the minimal model of the retinal protonated Schiff base $(C_5H_6NH_2^+)$ depicted in Fig. 3.1 and denoted as CN5 and PSB3, respectively. As already mentioned, while being generally challenging for electronic structure methods [4, 25–28], these examples are representative of the two cases of a ground (S0) and an excited (S1) state of different (CN5) and equal (PSB3) symmetry. Correspondingly, the energy minimization scheme is state-specific for CN5 and state-average for PSB3, while variance minimization affords a state-specific optimization for both molecules, at least in principle.

Ground and excited states of different symmetry

In Table 3.1, we list the ground- and excited-state energies, and corresponding excitation energies of CN5 computed in VMC and DMC with different wave functions optimized by (state-specific) energy minimization. The simplest case consists of a single determinant (HF) and a HOMO-LUMO (HL) two-determinant wave function for the ground and the excited state, respectively. We then consider configuration interaction singles (CIS) expansions, CAS expansions with increasing active spaces, and balanced CIPSI expansions with different choices of the starting orbitals, namely, independent sets for the two states (CIPSI-SS) or a common set of orbitals (CIPSI-B_1). The excitation energies are displayed in Fig. 3.2.

The general trend is a decrease of the excitation energy towards the extrapolated full CI (exFCI) and approximate coupled cluster singles, doubles and triples model (CC3) reference values for better wave functions. As an exception, when we move from the HF/HL to CIS wave functions, the VMC energies of both states decrease but the corresponding excitation energy becomes worse. With increasingly large CAS expansions, both the total and the excitation energies improve but the convergence is

very slow. For all these wave functions, the DMC excitation energy is lower than the VMC value and becomes within 0.1 eV of the reference results for the largest active spaces with about 50,000 and 70,000 determinants for the ground and the excited state, respectively. By comparison, the errors of TDDFT and CASPT2 can be as large as 0.4 and -0.2 eV, respectively [4, 45].
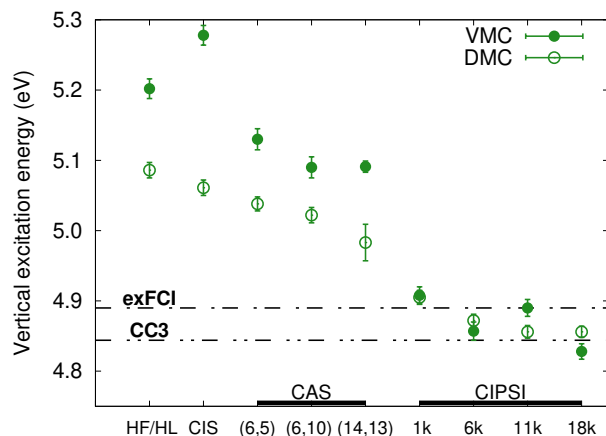

Figure 3.2: VMC and DMC excitation energies of CN5 calculated with different wave functions optimized in energy minimization. The exFCI/aug-cc-pVDZ [48] and CC3/aug-cc-pVTZ reference values are also shown. The approximate total number of determinants for the CIPSI-SS wave functions of the ground and excited states is indicated.

Table 3.1: VMC and DMC total energies (a.u.) and excitation energies (ΔE, eV) of CN5 obtained for different wave functions optimizing all parameters (Jastrow, orbital, and CI coefficients) in energy minimization.

WF	No. det		No. param		VMC			DMC		
	S0	S1	S0	S1	E(S0)	E(S1)	ΔE	E(S0)	E(S1)	ΔE
HF/HL	1	2	516	529	-40.8372(4)	-40.6460(3)	5.202(14)	-40.9378(3)	-40.7509(3)	5.086(11)
HF/CIS	1	980	516	4751	-40.8372(4)	-40.6505(3)	5.080(14)	-40.9378(3)	-40.7533(3)	5.020(11)
CIS	999	980	5260	4751	-40.8444(4)	-40.6505(3)	5.278(14)	-40.9393(3)	-40.7533(3)	5.061(11)
CAS(6,5)	52	48	567	561	-40.8468(4)	-40.6583(4)	5.130(15)	-40.9433(3)	-40.7582(2)	5.038(10)
CAS(6,10)	7232	7168	3134	3064	-40.8498(4)	-40.6628(4)	5.090(15)	-40.9439(3)	-40.7594(3)	5.022(11)
CAS(14,13)	48206	72732	9480	11727	-40.8583(3)	-40.6713(3)	5.091(10)	-40.9442(7)	-40.7611(7)	4.983(26)
CIPSI-SS	376	1094	1567	2609	-40.8646(3)	-40.6842(3)	4.908(12)	-40.9467(3)	-40.7665(3)	4.905(10)
	1344	4382	2478	4531	-40.8798(3)	-40.7013(3)	4.857(13)	-40.9502(2)	-40.7711(2)	4.872(09)
	2460	8782	3555	6561	-40.8896(3)	-40.7099(3)	4.890(12)	-40.9532(2)	-40.7748(2)	4.856(09)
	3913	14114	4842	8312	-40.8941(2)	-40.7167(3)	4.828(11)	-40.9559(2)	-40.7775(2)	4.856(08)
CIPSI-B$_1$	2456	6120	3971	5466	-40.8847(2)	-40.7053(2)	4.880(09)	-40.9521(2)	-40.7727(2)	4.881(09)
	4829	13130	5737	8021	-40.8945(3)	-40.7150(3)	4.889(13)	-40.9560(2)	-40.7766(2)	4.882(08)
exFCI/aug-cc-pVDZ [48]										4.89
CC3/aug-cc-pVDZ										4.851
CC3/aug-cc-pVTZ										4.844

The quality of the results exhibits a further, dramatic improvement with the use of CIPSI expansions. The VMC and DMC energies obtained with the smallest CIPSI wave function are lower than the corresponding values obtained with the largest CAS considered here. Furthermore, constructing ground- and excited-state CIPSI expansions with similar PT2 corrections leads to a balanced description of both states and

to VMC excitation energies which change very little with increasing expansion size, being irregularly scattered over a small energy range of 0.08 eV. Importantly, the DMC excitation energies are compatible with the VMC ones and in excellent agreement with the CC3 and exFCI values. Finally, employing two different sets of orbitals to generate the CIPSI expansions leads to marginal differences, namely, to DMC excitation energies of 4.856(8) and 4.882(8) eV, which are both bracketed by the reference values.

Having verified that state-specific energy optimization in combination with accurate wave functions allows the robust treatment of CN5, we now employ variance minimization with the σ_ω^2 functional to optimize the CAS(6,5) and CAS(6,10) wave functions of the ground and excited states. The convergence of the corresponding VMC variances and energies is shown in Fig. 3.3. For the smaller CAS(6,5), we observe that, while the variance converges rather quickly, the energy appears to do so more slowly and only after undershooting to a value which generally depends on the statistical error and initial conditions of the run. For an approximate wave function, the optimal parameters in variance minimization may differ from those obtained in energy minimization. Therefore, during the optimization of the variance, the energy can become lower than the final one.

As reported in Table 3.2, the optimal ground- and excited-state energies are higher by about 30 mHartree than the corresponding values obtained in energy minimization but the resulting excitation energy is compatible within statistical error. If
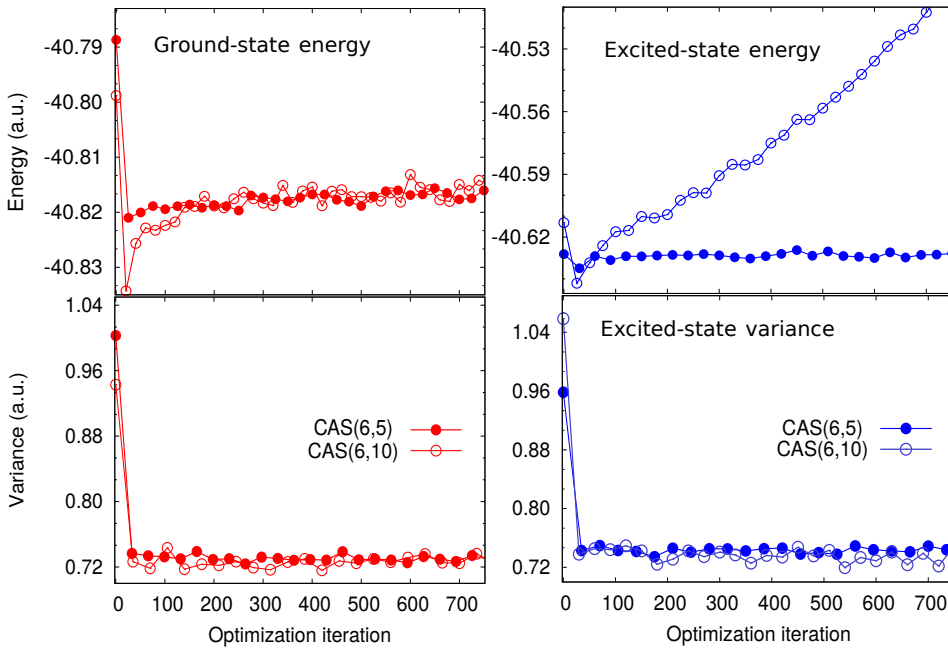

Figure 3.3: Convergence of the VMC energy (top) and variance (bottom) of the ground (left) and excited (right) states of CN5 in the optimization of the CAS(6,5) and CAS(6,10) wave functions in variance minimization.

we move to the larger CAS(6,10) determinantal expansion, we find however that,

while the variance reaches a stable value and the ground-state energy has a similar behavior to the CAS(6,5) case, the energy of the excited state grows steadily and it is therefore not possible to estimate the vertical excitation energy of the system. Surprisingly, even in the simplest case of the one-configuration (HF/HL) wave functions, the energy of the excited state keeps slowly rising even after 600 iterations as shown in Fig. 3.4, while the ground-state energy behaves similarly to the corresponding CAS cases.

Table 3.2: VMC energies and variances (a.u.) and vertical excitation energies (eV) of CN5 obtained with energy and variance minimization.

	Energy min.					Variance min.				
	E(S0)	E(S1)	ΔE	σ^2(S0)	σ^2(S1)	E(S0)	E(S1)	ΔE	σ^2(S0)	σ^2(S1)
CAS(6,5)	-40.8468(4)	-40.6583(4)	5.13(1)	0.862	0.885	-40.8170(5)	-40.6270(5)	5.17(2)	0.733	0.743
CAS(6,10)	-40.8498(4)	-40.6628(4)	5.09(1)	0.855	0.868	-40.8163(4)	–	–	0.731	–

Importantly, the apparently unstable behavior is independent of the initial value of ω and the number of steps over which we keep ω fixed (see Apendix 3.B). The use of a smaller or larger damping factors (i.e. $\tau = 0.04$ and 0.2) leads to the same pathological growth of the excited-state energy, characterized by the same slope as a function of time as shown in Fig. 3.13. Moreover, we recover the same behavior also when using a gradient-only-based optimizer (see Fig. 3.14). Finally, minimizing the Ω functional instead of σ_ω^2 yields an excited-state energy which ultimately rises with iterations as shown for the excited-state HL wave function in Fig. 3.4.
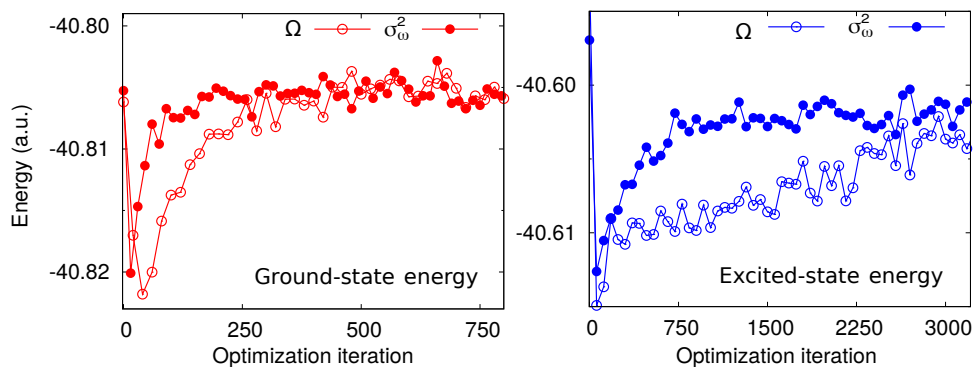

Figure 3.4: Convergence of VMC energy of the ground (left) and excited (right) states of CN5 in the optimization of the HF/HL wave functions within variance minimization with the σ_ω^2 (our default) and the Ω functional.

Ground and excited states of the same symmetry

For PSB3, we optimize the wave functions in energy minimization in a state-average fashion and report the resulting VMC and DMC total energies and vertical excitation energies in Table 3.3. As in the CN5 case, CIPSI wave functions are superior to CAS expansions of similar size and, with only about 400 determinants, the use of CIPSI

yields not only lower total energies but also a VMC vertical excitation energy in good agreement with the CC3 reference, largely correcting the error of 0.25 eV obtained with the CAS(6,6) wave function. For all CIPSI expansions, the DMC excitation energies are always quite close to the correspondent VMC results and, for the larger expansions, within 0.05 eV of the CC3 value.

When we perform state-specific variance minimization, we encounter great difficulties in the convergence of the energies as we show for the HF/HL and CAS(6,6) wave functions in Fig. 3.5. Differently from CN5, we find in general that not only the energy of the excited state but also that of the ground state grows steadily with iteration number.

Table 3.3: VMC and DMC total energies (a.u.) and excitation energies (ΔE, eV) of PSB3 obtained for different wave functions optimizing all parameters (Jastrow, orbital, and CI coefficients) in energy minimization.

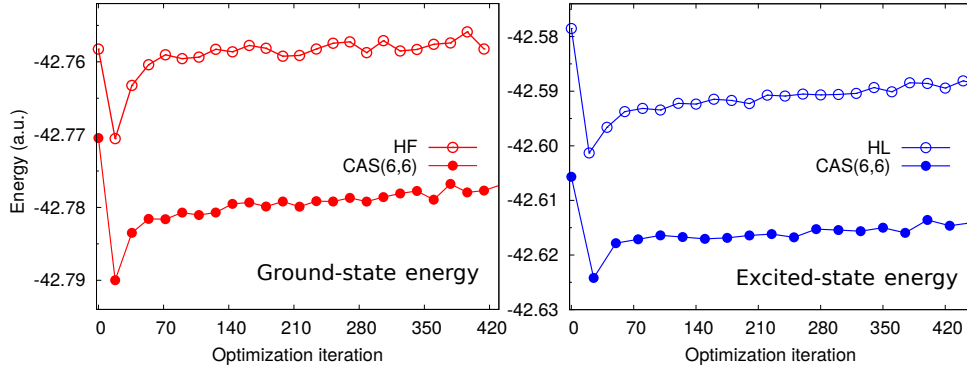| WF	No. det	No. param	VMC			DMC		
			E(S0)	E(S1)	ΔE	E(S0)	E(S1)	ΔE
CAS(6,6)	400	1645	-42.8091(2)	-42.6471(2)	4.409(9)	-42.9118(2)	-42.7541(2)	4.293(6)
CIPSI	422	4011	-42.8174(2)	-42.6623(2)	4.221(9)	-42.9133(2)	-42.7578(2)	4.233(6)
	1158	5968	-42.8297(2)	-42.6735(2)	4.252(9)	-42.9160(2)	-42.7609(2)	4.221(6)
	2579	8106	-42.8357(2)	-42.6796(2)	4.247(9)	-42.9169(2)	-42.7621(2)	4.214(6)
CC3/aug-cc-pVDZ								4.19
CC3/aug-cc-pVTZ								4.16

Figure 3.5: Convergence of the VMC energy of the ground (red) and excited (blue) states of PSB3 in the optimization of the RHF/HL and CAS(6,6) wave functions within variance minimization.

3.5 Discussion and conclusions

While our results confirm the high accuracy reachable in QMC with energy minimization, they evidence severe problems in variance minimization which, in most cases, preclude the estimation of the excitation energy. To gain a better understanding of the troublesome behavior of the energy during variance minimization, we further investigate the simple case of the HL wave function of CN5 (Fig. 3.4) and find that

the energy of the state drifts to higher values during variance minimization also when one optimizes only the LUMO orbital. Therefore, since optimization of an orbital can be achieved by mixing it with the unoccupied ones of the same symmetry, we can recast the LUMO optimization into the linear variation of the CI coefficients of the single excitations out of the LUMO orbital, which amount to only twelve additional CSFs in our basis set. With such a small expansion, we can then diagonalize the Hamiltonian in the basis of the CSFs times the Jastrow factor to estimate its thirteen eigenvalues and eigenvectors, and work directly in the basis of the eigenstates to assess the behavior of variance minimization when starting from the states which are optimal for energy minimization.
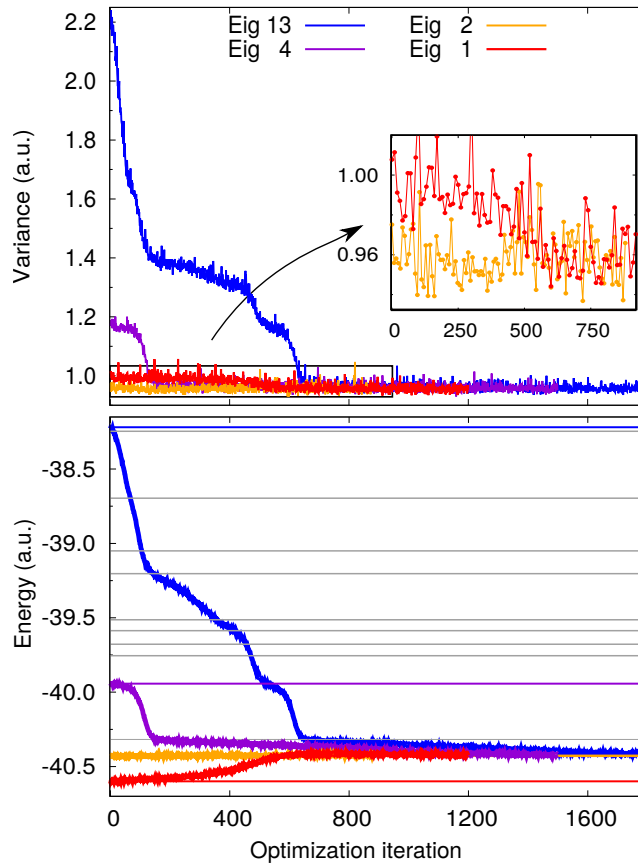

Figure 3.6: Convergence of the VMC variance (top) and energy (bottom) of CN5 in the CI optimization of a small expansion (see text) with variance minimization. The horizontal lines in the energy plot correspond to the eigenvalues in this reduced space, and the colored ones are the eigenstates used as starting point in four optimization runs. The damping factor used in the Newton method is $\tau = 0.2$.

In Fig. 3.6, we show the evolution of the VMC variance and energy for four variance minimization runs in which we start from different eigenvectors, taking the corresponding eigenvalues as initial target energies ω. In particular, we consider the lowest state in B_1 symmetry as well as the second, fourth, and thirteenth (corresponding to the highest energy) states. We note that, since our states are not exact eigen-

states of the full Hamiltonian, the corresponding variances of the local energy are non zero and are spread over about 0.5 a.u. with the lowest value in correspondence of the second state. In principle, one would expect to find a feature of the variance landscape –ideally a local minimum– near each of the approximate eigenstates since the functionals σ_ω^2 or Ω are designed to select a particular state through the initial value of ω, and minimize the variance of *that* state. Here, the selection of the state is further facilitated starting each run precisely from the chosen eigenstate, and variance minimization should perform minor adjustments of the initial parameters from their optimal values for the energy.

The behavior illustrated in Figs. 3.6 is totally different, with all optimization runs leaking down to successive lower-variance states and eventually converging to the absolute minimum corresponding to the second eigenstate. The staircase shape of the variance evolution points to the presence of flat regions of the variance landscape close to the eigenstates, from which the optimization can eventually escape. This is further corroborated if we follow the evolution of the CI coefficients as shown starting from the highest-energy state in Fig. 3.7: the initial coefficient quickly decreases to zero and other eigenstates become populated until convergence on the second state. In proximity of some eigenstates, the variance displays a more pronounced plateau, where the system spends enough time to acquire the full character of that particular state. It is also interesting to note that the states are populated sequentially with the order determined by decreasing energies. We stress that we observe a similar behavior of the variance also when using the Ω functional starting from the same set of approximate eigenstates (see Fig. 3.15).
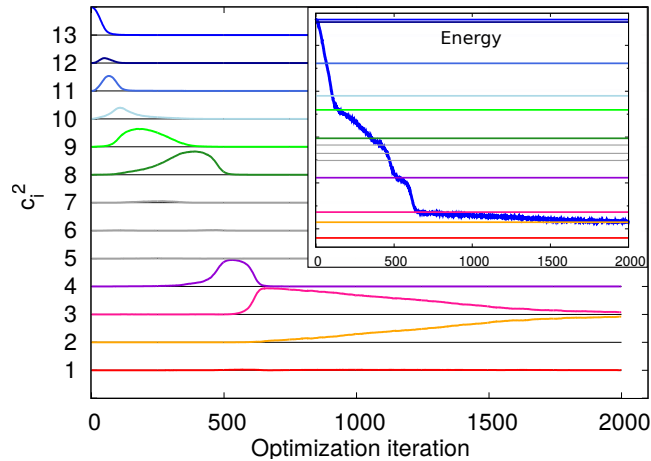
Figure 3.7: Evolution of the square of the CI coefficients c_i^2 (offset by i for clarity) of the small expansion of CN5 during variance minimization, for the run starting from the 13$^{\text{th}}$ eigenvector; in the inset, the evolution of the energy is replicated to emphasize flat regions in the energy landscape close to an eigenstate (i.e. when the corresponding $c_i \sim 1$).

In Fig. 3.8, we investigate the impact of the statistical error on the loss of the selected state. In particular, we focus on the evolution of the variance and the energy

starting from the 4th eigenvector for different lengths of the VMC runs used to compute the gradient and Hessian matrix. The shortest run (larger statistical error) looses the target state in a slightly smaller number of steps. However, the other runs give very similar results, suggesting that even longer VMC runs would not stabilize the target state.
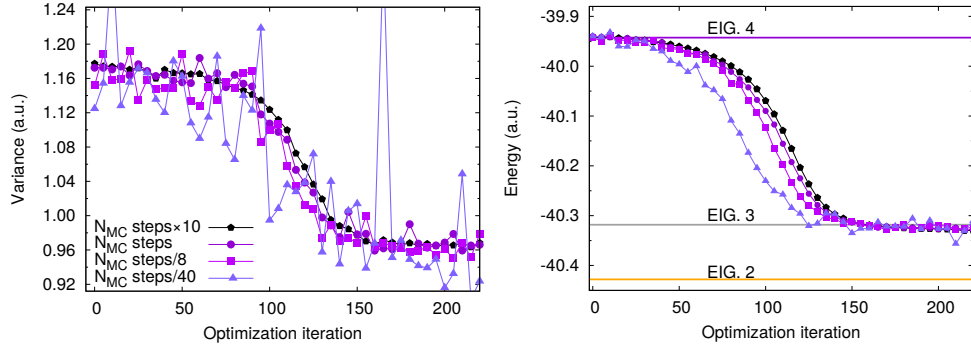

Figure 3.8: Convergence of the variance (left) and energy (right) for different lengths of the Monte Carlo runs used to compute the gradient and Hessian matrix during the optimization, starting from the 4th eigenvector. N_{MC} is the number of Monte Carlo steps used in Fig. 3.6.
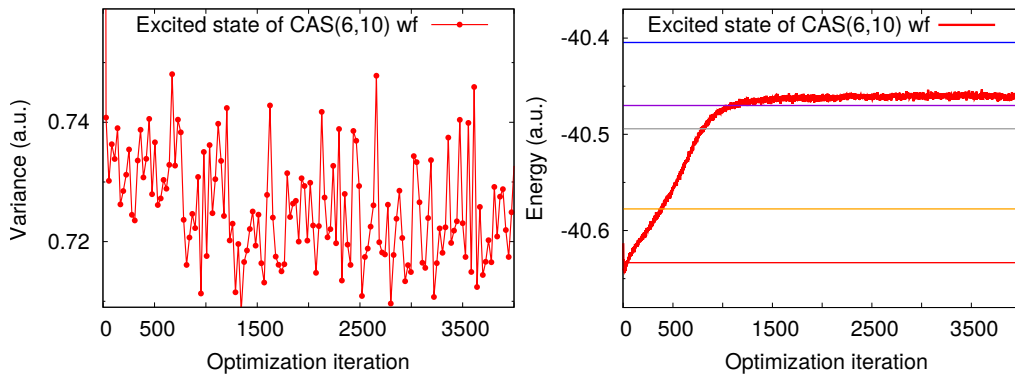

Figure 3.9: Variance (left) and energy (right) convergence for the optimization of the excited state of the CAS(6,10) wave function. The horizontal lines in the energy plot correspond to the firsts eigenvalue roots obtained with the Davidson optimization.

This simple wave function of CN5 is an explicit instance of missing one–to–one correspondence between minima of the variance and approximate eigenstates. Even if the actual number of minima and their correspondence to particular eigenstates remains unknown in general, the understanding gained here clearly applies to the behavior that we have observed for more complicated wave functions. As an explicit example, we revisit the very problematic optimization of the excited-state CAS(6,10) wave function (Fig. 3.3) and perform a much longer calculation, finding that the energy eventually converges as shown in Fig. 3.9. For the final set of Jastrow and orbital parameters, we determine the eigenvalues in the linear space of the determinants times the Jastrow factor and recover a similar behavior to what observed in

the simple example: the minimization of σ_ω^2 brings the system approximately to an eigenstate with a lower variance, which is in this case the 4$^\text{th}$ one.

By systematically improving the wave function, it is possible in principle to approach the exact eigenstate and its zero-variance property, thus recovering the corresponding minimum in the variance landscape. However, a CIPSI expansion which gives excellent results in energy minimization does not always prove sufficient to stabilize variance minimization(see Section 3.F). In general, going to extended determinantal expansions for the sake of a stable variance minimization, when energy minimization results are already satisfactory, appears unpractical, if feasible at all.

In summary, we have shown that the combination of energy minimization with an appropriate choice of the ground- and excited-state wave functions via a balanced CIPSI procedure leads to excitation energies that are in excellent agreement already at the VMC level with the reference values. In particular, we obtained a robust convergence of the total ground- and excited-state energies, and a very accurate excitation energy not only in the easier state-specific case of CN5 but also when employing energy minimization in a state-average fashion for PSB3. On the other hand, we encountered severe problems when employing variance minimization since, over sufficiently long optimization runs, one may loose the state of interest in favor of a state with lower variance, as we clearly demonstrated with a simple but realistic example. Even though, theoretically, the functionals σ_ω^2 and Ω have a built-in possibility to target the energy of a specific state, in practice, this is generally not sufficient to maintain the parameters close to the desired local minimum of the variance. Therefore, these considerations lead to the conclusion that, with the present functionals and no *a priori* knowledge of the parameter landscape of the variance for the system of interest, energy minimization is a safer and more stable procedure.

3.A   Appendix: Expressions for the gradient and Hessian

The functional σ_ω^2 is defined as

$$\sigma_\omega^2 = \frac{\langle \Psi | (\hat{\mathcal{H}} - \omega)^2 | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \langle E_{\mathrm{L}}^2 \rangle - 2\omega \langle E_{\mathrm{L}} \rangle + \omega^2 \,, \tag{3.16}$$

where $E_{\mathrm{L}} = \hat{H}\Psi/\Psi$ is the local energy and ω is a target energy kept fixed for some initial steps N_{F}, varied to reach the running energy average over the following N_{T} steps, and then simply updated to the current value of the energy for the remaining part of the optimization.

To minimize this functional with the Newton method, we need to compute the gradient and the Hessian matrix of σ_ω^2. To this aim, we can use the straightforward expressions, simplify them through some approximation, or further manipulate them to reduce the statistical noise as discussed below.

3.A.1   Computation of the gradient of σ_ω^2

As described in the main text (Eq. 14), we sample the following expression of the gradient of σ_ω^2:

$$\begin{aligned} g_i &= 2 \left[\left\langle \frac{\hat{\mathcal{H}}\Psi_i}{\Psi} E_{\mathrm{L}} \right\rangle - \left\langle \frac{\Psi_i}{\Psi} \right\rangle \langle E_{\mathrm{L}}^2 \rangle \right. \\ &\quad - \left. \omega \left(\left\langle \frac{\Psi_i}{\Psi} E_{\mathrm{L}} \right\rangle + \left\langle \frac{\hat{H}\Psi_i}{\Psi} \right\rangle - 2 \left\langle \frac{\Psi_i}{\Psi} \right\rangle \langle E_{\mathrm{L}} \rangle \right) \right] . \end{aligned} \tag{3.17}$$

Importantly, if we make use of the hermiticity of the Hamiltonian operator and modify the expression as:

$$g_{2,i} = 2 \left[\left\langle \frac{\hat{\mathcal{H}}\Psi_i}{\Psi} E_{\mathrm{L}} \right\rangle - \left\langle \frac{\Psi_i}{\Psi} \right\rangle \langle E_{\mathrm{L}}^2 \rangle - 2\,\omega \left(\left\langle \frac{\Psi_i}{\Psi} E_{\mathrm{L}} \right\rangle - \left\langle \frac{\Psi_i}{\Psi} \right\rangle \langle E_{\mathrm{L}} \rangle \right) \right] , \tag{3.18}$$

this leads to a significant increase of the noise and a very unstable minimization since the fourth term in Eq. 3.17 does no longer partially cancel the fluctuations of the first term (ω being equal to the running average energy). If we reintroduce a term to cancel the fluctuations of the first term, we obtain

$$
\begin{aligned}
g_{3,i} &= 2\left[\left\langle \frac{\hat{\mathcal{H}}\Psi_i}{\Psi} E_{\mathrm{L}} \right\rangle - \left\langle \frac{\hat{H}\Psi_i}{\Psi} \right\rangle \langle E_{\mathrm{L}} \rangle + \left\langle \frac{\hat{H}\Psi_i}{\Psi} \right\rangle \langle E_{\mathrm{L}} \rangle \right] \\
&\quad -2\left[\left\langle \frac{\Psi_i}{\Psi} \right\rangle \langle E_{\mathrm{L}}^2 \rangle - 2\omega\left(\left\langle \frac{\Psi_i}{\Psi} E_{\mathrm{L}} \right\rangle - \left\langle \frac{\Psi_i}{\Psi} \right\rangle \langle E_{\mathrm{L}} \rangle \right)\right] \\
&= 2\left[\left\langle \frac{\hat{\mathcal{H}}\Psi_i}{\Psi} E_{\mathrm{L}} \right\rangle - \left\langle \frac{\hat{H}\Psi_i}{\Psi} \right\rangle \langle E_{\mathrm{L}} \rangle + \left\langle \frac{\Psi_i}{\Psi} E_{\mathrm{L}} \right\rangle \langle E_{\mathrm{L}} \rangle \right] \\
&\quad -2\left[\left\langle \frac{\Psi_i}{\Psi} \right\rangle \langle E_{\mathrm{L}}^2 \rangle - 2\omega\left(\left\langle \frac{\Psi_i}{\Psi} E_{\mathrm{L}} \right\rangle - \left\langle \frac{\Psi_i}{\Psi} \right\rangle \langle E_{\mathrm{L}} \rangle \right)\right]. \quad (3.19)
\end{aligned}
$$

where we use Hermiticity to go from the first to the second line of the equation. This gradient expression has significantly lower fluctuations than $g_{2,i}$ since it is written in terms of covariances [22] and the optimization is stable. As expected, we find that the use of g_i (with ω updated to the running average energy) and $g_{3,i}$ yields equivalent speeds of convergence in the optimization.

3.A.2 Computation of the Hessian

The computation of the full Hessian matrix involves terms that depend on the second derivatives of the wave function. To avoid their computation, we follow the same approximation scheme of the Hessian one adopts when employing the Levenberg-Marquardt algorithm for the minimization of the mean square fluctuations of a fit function [33] as also done in the early work on variance minimization [19, 49]. To this aim, we rewrite the variance as the integral of the square of a function,

$$
\begin{aligned}
\sigma_\omega^2 &= \langle (E_{\mathrm{L}}(\mathbf{R}) - \omega)^2 \rangle = \frac{\int d\mathbf{R}\,(E_{\mathrm{L}}(\mathbf{R}) - \omega)^2 \Psi^2(\mathbf{R})}{\int d\mathbf{R}\,\Psi^2(\mathbf{R})} \\
&= \int d\mathbf{R}\,(E_{\mathrm{L}}(\mathbf{R}) - \omega)^2 w(\mathbf{R})^2 \equiv \int d\mathbf{R}\, Q^2(\mathbf{R}), \quad (3.20)
\end{aligned}
$$

where we defined

$$
Q(\mathbf{R}) = (E_{\mathrm{L}}(\mathbf{R}) - \omega) w(\mathbf{R}) = (E_{\mathrm{L}}(\mathbf{R}) - \omega)\frac{|\Psi(\mathbf{R})|}{\sqrt{\int d\mathbf{R}'\Psi^2(\mathbf{R}')^2}}. \quad (3.21)
$$

The resulting Hessian matrix with respect to the wave function parameters has components:

$$
h_{ij} = 2\int d\mathbf{R}\,[\partial_i Q(\mathbf{R})\partial_j Q(\mathbf{R}) + Q(\mathbf{R})\partial_{ij}Q(\mathbf{R})]. \quad (3.22)
$$

If $Q(\mathbf{R})$ and $\partial_{ij}Q(\mathbf{R})$ are weakly correlated, one can neglect the last term in this expression since we are minimizing the variance and therefore the value of $Q(\mathbf{R})$ over all space (in our case, one aims at $(E_{\mathrm{L}}(\mathbf{R}) - \omega) \approx 0$ with ω being equal to the running average energy). Within this approximation, the Hessian becomes

$$
h_{ij} \approx 2\int d\mathbf{R}\,\partial_i Q(\mathbf{R})\partial_j Q(\mathbf{R}). \quad (3.23)
$$

Since the gradient of the quantity Q is given by

$$
\begin{aligned}
\partial_i Q(\mathbf{R}) &= w(\mathbf{R})\left[\partial_i E_{\mathrm{L}}(\mathbf{R}) + (E_{\mathrm{L}}(\mathbf{R}) - \omega)\frac{\partial_i w(\mathbf{R})}{w(\mathbf{R})}\right] \\
&= w(\mathbf{R})\Big[\partial_i E_{\mathrm{L}}(\mathbf{R}) \\
&\quad + (E_{\mathrm{L}}(\mathbf{R}) - \omega)\left(\frac{\partial_i \Psi(\mathbf{R})}{\Psi(\mathbf{R})} - \int \mathrm{d}\mathbf{R}' w^2(\mathbf{R}')\frac{\partial_i \Psi(\mathbf{R}')}{\Psi(\mathbf{R}')}\right)\Big],
\end{aligned}
\tag{3.24}
$$

we obtain the following approximation for the Hessian

$$
\begin{aligned}
h_{ij} = 2 \ \Big\langle\ &\left[\partial_i E_{\mathrm{L}} + (E_{\mathrm{L}} - \omega)\left(\frac{\Psi_i}{\Psi} - \left\langle\frac{\Psi_i}{\Psi}\right\rangle\right)\right] \\
&\times \left[\partial_j E_{\mathrm{L}} + (E_{\mathrm{L}} - \omega)\left(\frac{\Psi_j}{\Psi} - \left\langle\frac{\Psi_j}{\Psi}\right\rangle\right)\right]\ \Big\rangle,
\end{aligned}
\tag{3.25}
$$

If we write out explicitly the derivative of the local energy, we have

$$
\begin{aligned}
h_{ij} = 2\Bigg\{ &\left\langle\left(\frac{\hat{\mathcal{H}}\Psi_i}{\Psi} - \omega\frac{\Psi_i}{\Psi}\right)\left(\frac{\hat{\mathcal{H}}\Psi_j}{\Psi} - \omega\frac{\Psi_j}{\Psi}\right)\right\rangle \\
&- \left\langle\frac{\Psi_i}{\Psi}\right\rangle\left\langle(E_{\mathrm{L}} - \omega)\left(\frac{\hat{\mathcal{H}}\Psi_j}{\Psi} - \omega\frac{\Psi_j}{\Psi}\right)\right\rangle \\
&- \left\langle\frac{\Psi_j}{\Psi}\right\rangle\left\langle(E_{\mathrm{L}} - \omega)\left(\frac{\hat{\mathcal{H}}\Psi_i}{\Psi} - \omega\frac{\Psi_i}{\Psi}\right)\right\rangle \\
&+ \left\langle(E_{\mathrm{L}} - \omega)^2\right\rangle\left\langle\frac{\Psi_i}{\Psi}\right\rangle\left\langle\frac{\Psi_j}{\Psi}\right\rangle\Bigg\}.
\end{aligned}
\tag{3.26}
$$

We note that, in refs. 22, 50, an approximate expression of the Hessian was put forward starting from the following expression of the variance on a finite sample of Monte Carlo configurations:

$$
\sigma_\omega^2 = \sum_{\mathbf{R}}(E_{\mathrm{L}}(\mathbf{R}) - \omega)^2,
\tag{3.27}
$$

where the weights $w(\mathbf{R})$ are omitted (Eq. 3.21), while they should be included when the wave function is changed with respect to the one used for the sampling. If we follow the derivation above to obtain an approximate Hessian, one formally finds

$$
h_{0,ij} = 2\sum_{\mathbf{R}}\partial_i E_{\mathrm{L}}(\mathbf{R})\partial_j E_{\mathrm{L}}(\mathbf{R}),
\tag{3.28}
$$

which can then be expressed as

$$
h_{0,ij} = 2\left\langle\left(\frac{\hat{\mathcal{H}}\Psi_i}{\Psi} - E_{\mathrm{L}}\frac{\Psi_i}{\Psi}\right)\left(\frac{\hat{\mathcal{H}}\Psi_j}{\Psi} - E_{\mathrm{L}}\frac{\Psi_j}{\Psi}\right)\right\rangle.
\tag{3.29}
$$

Alternatively, it is possible to obtain this approximate Hessian, starting from the appropriate expression of the variance and following Ref. 20.

While the use of h_0 leads to the efficient optimization of the Jastrow factor, it is unsuitable for the optimization of the determinant component since the terms,

$$\left\langle \frac{\hat{\mathcal{H}}\Psi_i}{\Psi} E_L \frac{\Psi_j}{\Psi} \right\rangle \quad \text{and} \quad \left\langle \frac{\hat{\mathcal{H}}\Psi_j}{\Psi} E_L \frac{\Psi_i}{\Psi} \right\rangle \tag{3.30}$$

are infinite. A possible, straightforward way to eliminate the divergences from the expression of the Hessian is to substitute E_L with the fixed value ω as:

$$h_{1,ij} = 2 \left\langle \left(\frac{\hat{\mathcal{H}}\Psi_i}{\Psi} - \omega \frac{\Psi_i}{\Psi} \right) \left(\frac{\hat{\mathcal{H}}\Psi_j}{\Psi} - \omega \frac{\Psi_j}{\Psi} \right) \right\rangle . \tag{3.31}$$

This corresponds to the first term of the approximate Hessian matrix (Eq. 3.25).

In Fig. 3.10, we show the convergence of the variance and energy during the variance optimization of the ground-state CAS(6,6) wave function of PSB3 using the approximate expressions of the Hessian, h and h_1. We see no significant difference in the convergence behavior of the variance and of the energy, so also the very simple expression in Eq. 3.31 results in the effective minimization of the variance.
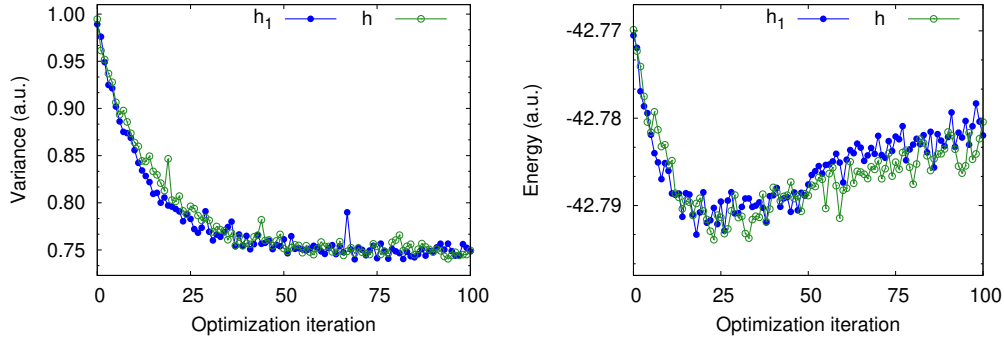

Figure 3.10: Convergence of the variance (left) and energy (right) in the variance optimization of the ground-state CAS(6,6) wave function of PSB3 with different approximate expressions of the Hessian.

3.B Appendix: Dependence of variance minimization on the choice of target energy ω

In Fig. 3.11, we show the convergence of the VMC energy and variance for the the excited state of CN5 with a CAS(6,10) wave function optimized in variance minimization. We compare the convergence when the target energy (ω) is kept fixed for the whole run and when it is updated following the scheme in Ref. 23.

For the runs with ω fixed, we use three different values of ω which are chosen to be lower than the energy of the starting wave function. The highest value of ω

(-40.65 a.u.) is close to the energy obtained in energy minimization. For the run
with varying ω, we keep ω fixed to -40.75 a.u. over $N_F =20$ steps and linearly
updated it to the running average energy over the subsequent N_T=30 steps.

As expected, we find that, when fixing the value of ω for the whole run, the energy
stabilizes to a certain value which is higher than the target ω and differs depending on
the choice of ω. Therefore, the procedure of keeping ω fixed is unsuitable for com-
puting excitation energies, not only because the method is not size consistent [23]
but also for the strong dependence of the result on the input value of ω.
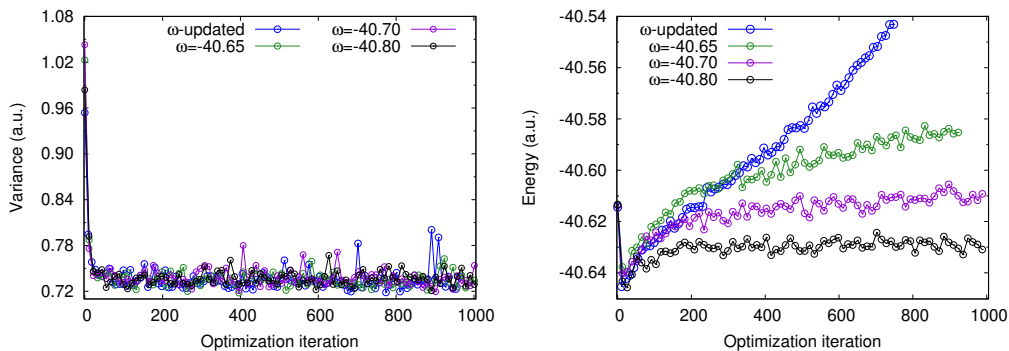

Figure 3.11: Convergence of the VMC variance (left) and energy (right) in the vari-
ance minimization of the excited-state CAS(6,10) wave function of CN5 when ω is
kept fixed over the whole run and when ω is updated to the value of the running
average energy [23].

In Fig. 3.12, we investigate the dependence of the results on the choice of N_T.
We start from a wave function optimized by maintaining ω fixed for the whole run,
and linearly update the value of ω to the current energy using different values of
steps N_T. We find that, no matter how slowly we tune this parameter, once we allow
ω to vary, the energy starts rising and, after a transient number of steps, all the runs
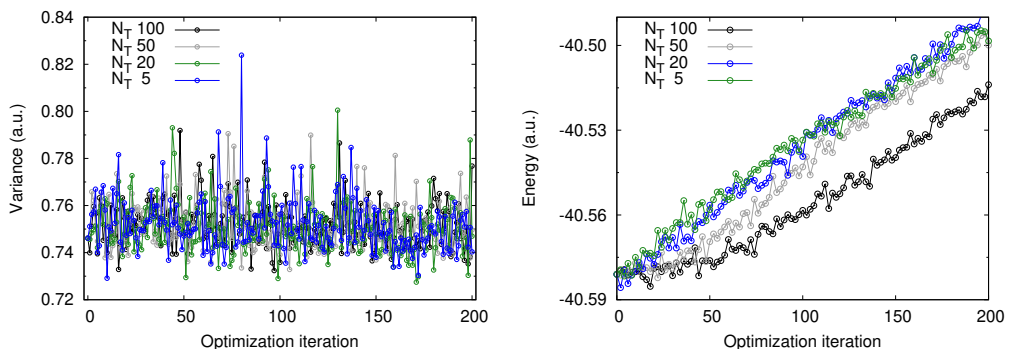display an equivalent behavior.

Figure 3.12: Convergence of the VMC variance (left) and energy (right) for the ex-
cited state of the CAS(6,10) of CN5 optimized in variance minimization with differ-
ent values of N_T.

3.C Appendix: Dependence of variance minimization on the damping factor τ

In Fig. 3.13, we report the convergence of the VMC excited-state energy and variance in the variance minimization of the CAS(6,10) wave function of CN5 for different values of the damping factor τ used in the Newton method. We find that, for all the tested values of τ (τ =0.04, 0.1, 0.2), the growing behavior of the energy is the same as a function of the product of the number of iterations and τ.
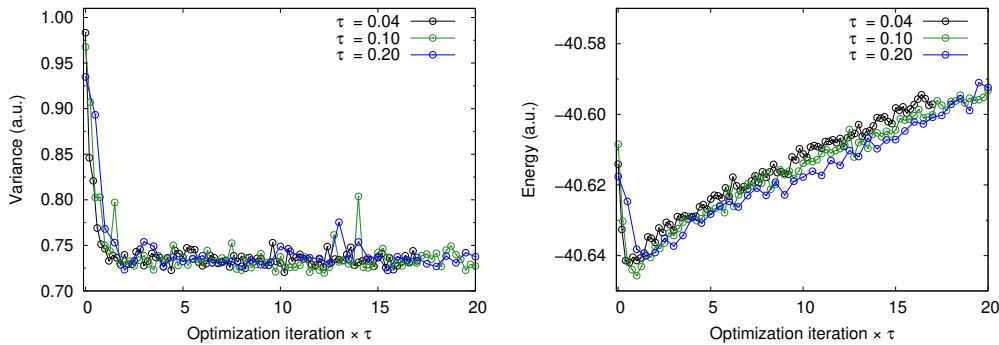

Figure 3.13: Convergence of the VMC variance (left) and energy (right) in the variance minimization of the excited-state CAS(6,10) wave function of CN5. The x-axis corresponds to the number of iteration times the damping factor τ in the Newton method.

3.D Appendix: Variance minimization with gradient-only optimizer

In Fig. 3.14, we show the convergence of the VMC variance and energy in variance optimization with the Newton method and with a gradient-only optimizer, namely, the ADAM approach, which has recently been successfully tested in the context of quantum Monte Carlo [51]. We use the HL wave function of the excited-state of CN5 and set the parameters in the ADAM method to $\beta_1 = 0.1$ and $\beta_2 = 0.001$.

We find that both optimization schemes lead to the same qualitative results in terms of energy and variance: in both cases, the energy increases during the optimization and drifts to energies significantly higher than the value of $-40.6460(3)$ a.u. obtained in energy minimization. Therefore, even if convergence is later achieved, the converged state will not correspond to the desired lowest state.
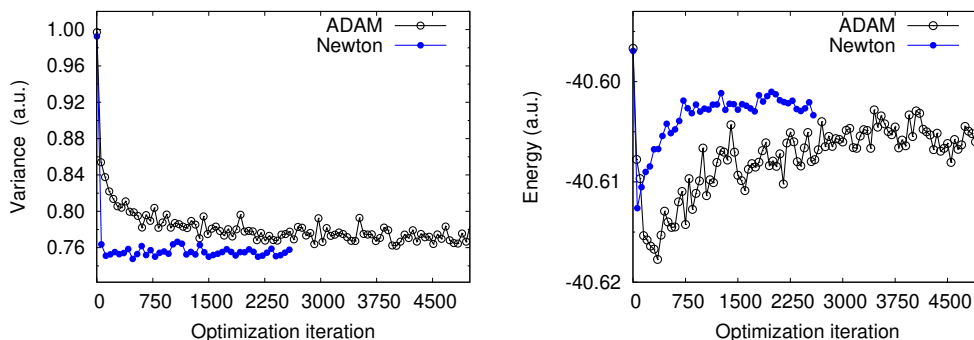
Figure 3.14: Convergence of the VMC variance (left) and energy (right) in the variance minimization of the excited-state HL wave function of CN5 with the ADAM optimizer and Newton method. Note that the energy obtained in energy minimization for the HL wave function is $-40.6460(3)$ a.u. which falls out of the scale.

3.E   Appendix: Variance minimization with Ω functional

In Fig. 3.15, we show the convergence of the VMC variance and energy in variance optimization for the optimization of the LUMO orbital of CN5 with the functional Ω. As explained in the main text, we recast the optimization of the LUMO as a CI optimization of thirteen CSFs. The curves shown correspond to variance minimization runs in which we start from different approximate eigenvectors and use the corresponding eigenvalue energies (E_i) to target the state ($\omega \simeq E_i - \sigma$). We observe a similar behavior to the one obtained with the functional σ_ω^2 (Fig. 6) with all four variance/energy curves converging to the second eigenstate.

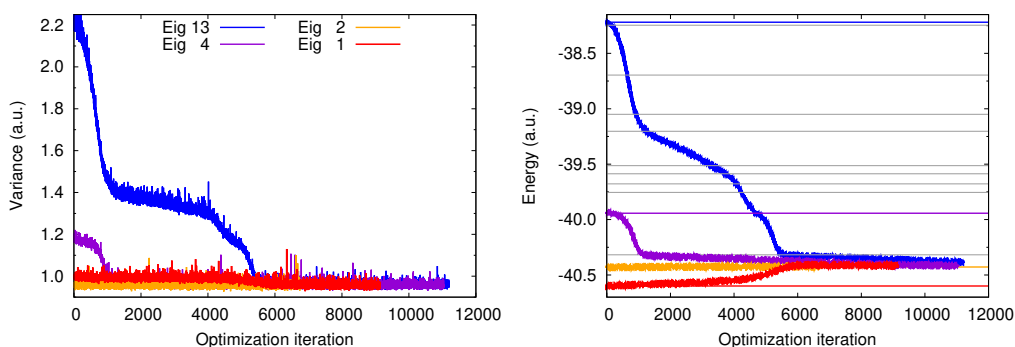Figure 3.15: Convergence of the VMC variance (left) and energy (right) of CN5 in the CI optimization of the LUMO orbital with the Ω functional in variance minimization. The horizontal lines in the energy plot correspond to the eigenvalues in this reduced space, and the colored ones are the eigenstates used as starting point for the optimization runs. The damping factor used in the Newton method is $\tau = 0.2$.

3.F Appendix: Variance minimization with CIPSI wave functions

In the respective ranges of determinantal expansions considered in this work, the CIPSI wave functions give significantly better energies and excitations than the CAS wave functions, as seen from Tables 1 and 3 of the main text. Their variances, listed in Table 3.4, are correspondingly lower, but not by a large amount: for instance, the variance of the best CIPSI wave function for the S1 state of CN5 is just 22% lower than that of the smallest CAS. Comparatively, we are still far from the limit of the zero-variance property of exact eigenstates where variance minimization is safe and stable. On this ground, our CIPSI wave functions may be expected to face problems in variance minimization to a similar extent than our CAS wave functions.

Table 3.4: Variance for the states S0 and S1 of CN5 and PSB3 with the wave functions of Tables 1 and 3 of the main text.

molecule	WF	N. of determinants		variance (a.u.)	
		S0	S1	S0	S1
CN5	HF/HL	1	2	0.885	0.893
	CIS	999	980	0.853	0.893
	CAS(6,5)	52	48	0.862	0.885
	CAS(6,10)	7232	7168	0.855	0.868
	CAS(14,13)	48206	72732	0.820	0.839
	CIPSI-SS	376	1094	0.798	0.799
		1344	4382	0.737	0.738
		2460	8782	0.699	0.712
		3913	14114	0.689	0.686
	CIPSI-B_1	2456	6120	0.719	0.728
		4829	13130	0.688	0.693
PSB3	CAS(6,6)	400	400	0.878	0.891
	CIPSI	422	422	0.848	0.851
		1158	1158	0.803	0.811
		2579	2579	0.786	0.793

This is indeed the case, as shown in Fig. 3.16 where we compare the performance of variance and energy minimization in the optimization of selected Jastrow-CIPSI wave functions. In particular, we show the convergence of the VMC energies during the optimization of the S1 excited-state wave function of CN5 and PSB3.

For CN5, we use the CIPSI expansion with 4382 determinants of Table 1 and observe that the energy converges with both energy and variance minimization, in analogy with the CAS(6,5) but not with the CAS(6,10) wave function. For PSB3, we use the CIPSI expansion with 1158 determinants of Table 3 and perform variance minimization using two different starting points, which yield the two curves labeled "variance min 1" and "variance min 2". We observe that the energy keeps increasing over a wide range of iterations ("variance min 1") as for the CAS(6,6) result. We also
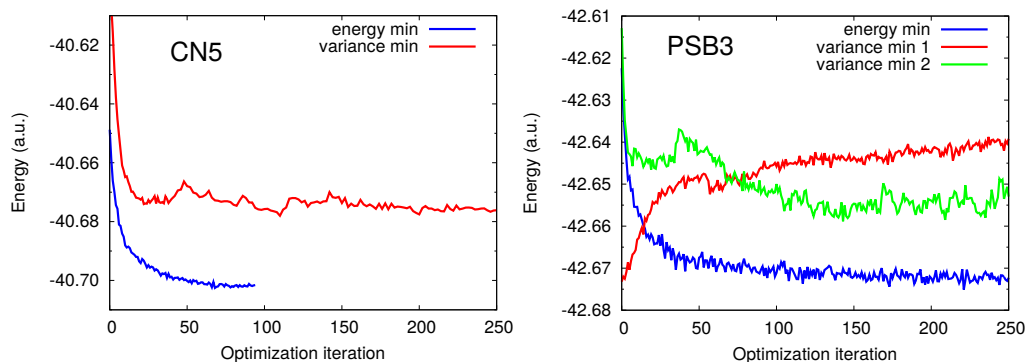
Figure 3.16: Convergence of the VMC energy in the energy and variance optimization of Jastrow, linear, and orbital parameters of an excited-state Jastrow-Slater wave function with a CIPSI expansion of 4382 determinants for CN5 (left) and 1158 determinants for PSB3 (right). The two variance minimization runs shown for PSB3 start from different initial parameters and are bound to converge to the same energy (barring different local minima).

note that the transient behavior of the energy depends significantly on the starting point, it can be quite slow (curve "variance min 2"), and is in general more noisy than in energy minimization.

Bibliography

[1] C. Filippi, M. Zaccheddu, and F. Buda, J. Chem. Theory Comput. **5**, 2074 (2009).

[2] P. M. Zimmerman, J. Toulouse, Z. Zhang, C. B. Musgrave, and C. Umrigar, J. Chem. Phys. **131**, 124103 (2009).

[3] O. Valsson and C. Filippi, J. Chem. Theory Comput. **6**, 1275 (2010).

[4] R. Send, O. Valsson, and C. Filippi, J. Chem. Theory Comput. **7**, 444 (2011).

[5] O. Valsson, P. Campomanes, I. Tavernelli, U. Rothlisberger, and C. Filippi, J. Chem. Theory Comput. **9**, 2441 (2013).

[6] R. Guareschi, H. Zulfikri, C. Daday, F. M. Floris, C. Amovilli, B. Mennucci, and C. Filippi, J. Chem. Theory Comput. **12**, 1674 (2016).

[7] R. J. Hunt, M. Szyniszewski, G. I. Prayogo, R. Maezono, and N. D. Drummond, Phys. Rev. B **98**, 075122 (2018).

[8] N. S. Blunt and E. Neuscamman, J. Chem. Theory Comput. **15**, 178 (2019).

[9] M. Dash, J. Feldt, S. Moroni, A. Scemama, and C. Filippi, J. Chem. Theory Comput. **15**, 4896 (2019).

[10] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001).

[11] A. Lüchow, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **1**, 388 (2011).

[12] B. M. Austin, D. Y. Zubarev, and W. A. Lester, Chem. Rev. **112**, 263 (2012).

[13] S. Sorella and L. Capriotti, J. Chem. Phys **133**, 234111 (2010).

[14] E. Neuscamman, C. J. Umrigar, and G. K.-L. Chan, Phys. Rev. B **85**, 045103 (2012).

[15] C. Filippi, R. Assaraf, and S. Moroni, J. Chem. Phys. **144**, 194105 (2016).

[16] R. Assaraf, S. Moroni, and C. Filippi, J. Chem. Theory Comput. **13**, 5273 (2017).

[17] M. Dash, S. Moroni, A. Scemama, and C. Filippi, J. Chem. Theory Comput. **14**, 4176 (2018).

[18] R. L. Coldwell, Int. J. Quantum Chem. **12**, 215 (1977).

[19] C. J. Umrigar, K. G. Wilson, and J. W. Wilkins, Phys. Rev. Lett. **60**, 1719 (1988).

[20] A. Malatesta, S. Fahy, and G. B. Bachelet, Phys. Rev. B **56**, 12201 (1997).

[21] P. R. C. Kent, R. J. Needs, and G. Rajagopal, Phys. Rev. B **59**, 12344 (1999).

[22] C. J. Umrigar and C. Filippi, Phys. Rev. Lett. **94**, 150201 (2005).

[23] J. A. R. Shea and E. Neuscamman, J. Chem. Theory Comput. **13**, 6078 (2017).

[24] S. D. Pineda Flores and E. Neuscamman, J. Phys. Chem. A **123**, 1487 (2019).

[25] O. Valsson, C. Angeli, and C. Filippi, Phys. Chem. Chem. Phys. **14**, 11015 (2012).

[26] M. Huix-Rotllant, M. Filatov, S. Gozem, I. Schapiro, M. Olivucci, and N. Ferré, J. Chem. Theory Comput. **9**, 3917 (2013).

[27] D. Tuna, D. Lefrancois, L. Wolański, S. Gozem, I. Schapiro, T. Andruniów, A. Dreuw, and M. Olivucci, J. Chem. Theory Comput. **11**, 5758 (2015).

[28] B. Le Guennic and D. Jacquemin, Acc. Chem. Res. **48**, 530 (2015).

[29] As Jastrow factor, we use the exponential of the sum of two fifth-order polynomials of the electron-nuclear and the electron-electron distances, respectively, and rescale the inter-particle distances as $R = (1 - \exp(-\kappa r))/\kappa$ with κ set to 0.6 a.u. We employ different electron-nucleus Jastrow factors to describe the correlation of an electron with H, C , and N. The total number of free parameters to be optimized in the Jastrow factor is 17 for the systems considered here.

[30] E. Neuscamman, J. Chem. Phys **145**, 081103 (2016).

[31] S. Sorella, M. Casula, and D. Rocca, J. Chem. Phys. **127**, 014105 (2007).

[32] I. Sabzevari, A. Mahajan, and S. Sharma, J. Chem. Phys **152**, 024111 (2020).

[33] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes 3rd Edition: The Art of Scientific Computing*, 3 ed. (Cambridge University Press, 2007), pp. 801–805.

[34] J. Toulouse and C. J. Umrigar, J. Chem. Phys **128**, 174101 (2008).

[35] CHAMP is a quantum Monte Carlo program package written by C. J. Umrigar, C. Filippi, S. Moroni and collaborators.

[36] M. Burkatzki, C. Filippi, and M. Dolg, J. Chem. Phys. **126**, 234105 (2007).

[37] For the hydrogen atom, we use a more accurate BFD pseudopotential and basis set. Dolg, M.; Filippi, C., private communication.

[38] R. A. Kendall, T. H. Dunning Jr, and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).

[39] C. Attaccalite and S. Sorella, Phys. Rev. Lett. **100**, 114501 (2008).

[40] M. Casula, Phys. Rev. B **74**, 161102 (2006).

[41] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery Jr, J. Comput. Chem. **14**, 1347 (1993).

[42] M. S. Gordon and M. W. Schmidt, in *Theory and Applications of Computational Chemistry*, edited by C. E. Dykstra, G. Frenking, K. S. Kim, and G. E. Scuseria (Elsevier, 2005), pp. 1167–1189.

[43] Y. Garniron, T. Applencourt, K. Gasperich, A. Benali, A. Ferté, J. Paquier, B. Pradines, R. Assaraf, P. Reinhardt, J. Toulouse, P. Barbaresco, N. Renon, G. David, J.-P. Malrieu, M. Véril, M. Caffarel, P.-F. Loos, E. Giner, and A. Scemama, J. Chem. Theory Comput. **15**, 3591 (2019).

[44] Dash, M.; Scemama, A., private communication.

[45] P. Boulanger, D. Jacquemin, I. Duchemin, and X. Blase, J. Chem. Theory Comput. **10**, 1212 (2014).

[46] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, J. J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox., Gaussian09, Revision A.02; Gaussian, Inc., Wallingford CT, 2016.

[47] R. M. Parrish, L. A. Burns, D. G. A. Smith, A. C. Simmonett, A. E. DePrince, E. G. Hohenstein, U. Bozkaya, A. Y. Sokolov, R. Di Remigio, R. M. Richard, J. F. Gonthier, A. M. James, H. R. McAlexander, A. Kumar, M. Saitow, X. Wang, B. P. Pritchard, P. Verma, H. F. Schaefer, K. Patkowski, R. A. King, E. F. Valeev, F. A. Evangelista, J. M. Turney, T. D. Crawford, and C. D. Sherrill, J. Chem. Theory Comput. **13**, 3185 (2017).

[48] Y. Garniron, A. Scemama, E. Giner, M. Caffarel, and P.-F. Loos, J. Chem. Phys **149**, 064103 (2018).

[49] C. Filippi and C. J. Umrigar, J. Chem. Phys. **105**, 213 (1996), as Jastrow correlation factor, we use the exponential of the sum of three fifth-order polynomials of the electron-nuclear (e-n), the electron-electron (e-e), and of pure 3-body mixed e-e and e-n distances, respectively. The e-n and e-e-n terms have different coefficients for different atom types. The Jastrow factor is further adapted to deal with pseudo-atoms, and the scaling factor κ is set to 0.60 a.u. The 3-body e-e-n terms are not included in the 2-body Jastrow factor.

[50] C. J. U. Julien Toulouse, J. Chem. Phys. **126**, 084102 (2017).

[51] L. Otis and E. Neuscamman, Phys. Chem. Chem. Phys. **21**, 14491 (2019).

Chapter 4

Reference excitation energies of increasingly large cyanine dyes: a QMC study

†

We revisit here the lowest vertical excitations of cyanine dyes using quantum Monte Carlo and leverage recent developments to systematically improve on previous results. In particular, we employ a protocol for the construction of compact and accurate multi-determinant Jastrow-Slater wave functions for multiple states, which we have recently validated on the excited-state properties of several small prototypical molecules. Here, we obtain quantum Monte Carlo excitation energies in excellent agreement with high-level coupled cluster for all the cyanines where the coupled cluster method is applicable. Furthermore, we push our protocol to longer chains, demonstrating that quantum Monte Carlo is a viable methodology to establish reference data at system sizes which are hard to reach with other high-end approaches of similar accuracy. Finally, we determine which ingredients are key to an accurate treatment of these challenging systems and rationalize why a description of the excitation based on only active π orbitals lacks the desired accuracy for the shorter chains.

4.1 Introduction

Cyanine dyes are a family of charged π-conjugated molecules which are employed in very diverse applications ranging from dye-synthesized solar cells to the labeling of bio-molecules [1–3]. Their characteristic structure consists of a chain of an odd number of carbons with two amine groups at the ends. While their photo-physical properties are strongly regulated by the length of the carbon chain, the lowest bright state of the cyanines always maintains a $\pi \to \pi^*$ character and can be predominantly

described as a HOMO to LUMO (HL) transition. Despite the apparent simplicity of this excitation, its accurate treatment is known to be challenging and, consequently, cyanine dyes have often been used as model systems to assess the quality of electronic structure methods for excited states [4–12].

Here, we employ quantum Monte Carlo (QMC) to revisit the vertical excitation energies of cyanine dyes of the simple form $C_nH_n(NH_2)_2^+$ with n an odd number ranging from 1 to 17, combining the use of sophisticated multi-determinant wave functions with recent developments for their efficient optimization in variational Monte Carlo (VMC) [13–16]. In particular, we build on our successful treatment at chemical accuracy of the excitation energies and optimal excited-state structures of small, prototypical molecules [17–19], where the determinantal components of the multiple states are generated in an automatic and balanced manner with the configuration interaction using a perturbative selection made iteratively (CIPSI) approach [20]. Studying the bright excitation of cyanine dyes enables us to demonstrate the accuracy of our protocol for the shorter chains, where high-level coupled cluster (CC) offers a good compromise in terms of accuracy versus computational cost. Importantly, it also establishes the applicability of QMC to larger sizes where the use of other high-level approaches is more challenging. Finally, we identify the key descriptors of orbital correlations for these systems and elucidate why earlier QMC studies with limited active space wave functions lacked the expected accuracy [4].

4.2 Methods

We employ QMC wave functions of the so-called Jastrow-Slater form, namely,

$$\Psi = \mathcal{J} \sum_{i=1}^{N_{\mathrm{det}}} c_i D_i \,, \tag{4.1}$$

where $\mathcal{J}$ is the Jastrow correlation factor and D_i are determinants of single-particle orbitals. The Jastrow factor explicitly depends on the inter-particle coordinates and includes here electron-electron and electron-nucleus correlation terms [21].

To generate the determinantal components for the two states, we employ the CIPSI approach which, starting from a given reference space, builds expansions by iteratively selecting determinants based on their second-order perturbation (PT2) energy contribution obtained via the Epstein-Nesbet partitioning of the Hamiltonian [22, 23],

$$\delta E_\alpha^{(2)} = \frac{|\langle \alpha | \mathcal{H} | \Psi^{\mathrm{CIPSI}} \rangle|^2}{\langle \Psi^{\mathrm{CIPSI}} | \mathcal{H} | \Psi^{\mathrm{CIPSI}} \rangle - \langle \alpha | \mathcal{H} | \alpha \rangle} \,, \tag{4.2}$$

where Ψ^{CIPSI} is the current CIPSI wave function for the state under consideration and $|\alpha\rangle$ denotes a determinant outside the current CI space. Since the ground and excited states of the cyanines have different symmetries, a state-specific approach can be used to perform the selection for the two states separately, using different orbitals.

We are here interested in computing excitation energies and, therefore, wish to achieve a balanced CIPSI description of the states of interest, which leads to converged excitation energies in QMC already for relatively small expansions.

A measure of the quality of a given CIPSI wave function is its PT2 energy contribution, which represents an approximate estimate of the error of the expansion with respect to the full CI (FCI) limit. Therefore, we can compute the excitation energies using expansions for the two (or more generally multiple) states with matched PT2 energy and, therefore, ensure comparable quality. We refer the reader to Ref. [19] on how to impose the "iso-PT2" criterion when treating multiple states of the same symmetry expanded on a common set of determinants.

Alternatively, one can match the CI variance of the relevant states, which is defined as the variance of the FCI Hamiltonian:

$$
\begin{aligned}
\sigma_{\mathrm{CI}}^2(\Psi^{\mathrm{CIPSI}}) &= \sum_{i \in \mathrm{FCI}} \langle \Psi^{\mathrm{CIPSI}} | \mathcal{H} | i \rangle \langle i | \mathcal{H} | \Psi^{\mathrm{CIPSI}} \rangle \\
&\quad - \langle \Psi^{\mathrm{CIPSI}} | \mathcal{H} | \Psi^{\mathrm{CIPSI}} \rangle^2 \\
&= \sum_{\alpha} | \langle \Psi^{\mathrm{CIPSI}} | \mathcal{H} | \alpha \rangle |^2 .
\end{aligned}
\tag{4.3}
$$

As the CIPSI wave function approaches the FCI limit, the CI variance goes to zero. For various small molecules [17–19], we have found that matching the PT2 energy contributions leads to expansions with also very similar variances. In general, this is not always the case and one of the two criteria might be more suitable than the other for the computation of the CI excitation energies of a particular system.

While we discuss in detail below the impact of this choice on the QMC excitation energies, we stress already here that the convergence of the QMC results is established not based on their agreement with available reference data but in an "internally consistent" manner based on the similarity of the VMC and DMC excitation energies [19] and their convergence with respect to the number of determinants.

Finally, as an alternative to the CIPSI expansions, we test complete active space (CAS) expansions for the determinantal components of our QMC wave functions. We start from separate CASSCF calculations for the two states and consider minimal active spaces by correlating the π electrons in the π orbitals constructed from the $2p_z$ orbitals. For the smaller cyanines with up to 7 heavy atoms, CN3–CN7 (we label a cyanine as CNm with m the total number of C and N atoms), we also explore the use of a larger active space with molecular orbitals constructed from the $2p_z$ and $3p_z$ atomic orbitals. Finally, in some cases, we also test the performance of a simple one-configuration ansatz, namely, the Hartree-Fock (HF) and HOMO-LUMO (HL) configurations for the ground and the excited state, respectively.

4.3 Computational details

Unless otherwise specified, we employ scalar-relativistic energy-consistent HF pseudopotentials and the correlation-consistent Gaussian basis sets specifically constructed

for these pseudopotentials [24, 25]. For most of the calculations, we use a double-ζ basis set minimally augmented with s and p diffuse functions on the heavy atoms and denoted here as maug-cc-pVDZ. Convergence tests are performed with the fully-augmented aug-cc-pVTZ basis set. The exponents of the diffuse functions are taken from the corresponding all-electron Dunning's correlation-consistent basis sets [26].

The HF and CASSCF computations are performed with the program GAMESS (US) [27, 28]. When using the CASSCF wave functions in QMC, we truncate the CAS expansion for CN11 and CN13, using a threshold on the CSF coefficients so that the configurations make up respectively about 0.9985 and 0.9765 of the weight of the total wave functions of the two states. The CIPSI expansions are generated with Quantum Package [29] and constructed to be eigenstates of $\hat{S}^2$ [17]. We perform the selection for the two states separately, starting from CASSCF orbitals obtained with the larger active spaces for the cyanine molecules up to CN7, and the minimal CAS from CN9 to CN15. We use the HF orbitals for CN17 and CN19. As shown in Fig. 4.4 and Table 4.5 for CN3 and Fig. 4.9 for CN15, the use of different orbitals to generate the CIPSI expansions has no appreciable impact on the CI or QMC excitation energies.

The QMC calculations are carried out with the CHAMP code [30]. The determinantal part of our QMC wave functions is expressed in terms of spin-adapted configuration state functions (CSF) to reduce the number of parameters during the VMC optimization. In the wave function optimization, we sample a guiding wave function that differs from the current wave function close to the nodes [31] to guarantee finite variances of the estimators of the gradients with respect to the wave function parameters. All wave function parameters (Jastrow, CI, and orbital coefficients) are optimized in state-specific energy minimization following the stochastic reconfiguration scheme [14, 32]. In the DMC calculations, we treat the pseudopotentials beyond the locality approximation using the T-move algorithm [33] and employ an imaginary time-step of 0.05 a.u. which we have already tested for one of the cyanine chains and shown to yield excitation energies converged to better than 0.01 eV [18]

We compute all energies on the ground-state geometries of CN3–CN11 determined with all-electron PBE0/cc-pVQZ in Ref. [8] and obtain the geometries for CN13 to CN19 at the same level of theory with the Gaussian 09 program [34]. We employ the programs CFour v2.1 [35] and Molcas [36] for the approximate coupled cluster singles and doubles (CC2) and singles, doubles, and triples model (CC3), and the CASPT2 calculations, respectively, using the all-electron aug-cc-pVDZ basis set and the frozen-core approximation, unless otherwise specified.

4.4 Results

We compute the lowest $\pi \to \pi^*$ vertical excitation energy of cyanine dyes of the form $C_n H_n (NH_2)_2^+$ with n ranging from 1 up to 17. The structures of the CN3 and CN9 molecules are shown in Figure 4.1. In all cases, the point group of the molecule is C_{2v} with the ground (GS) and excited (ES) states having A_1 and B_1 symmetry, respectively.

For CN3 up to CN15, we compare the QMC excitation energies with the all-electron CC3/aug-cc-pVDZ results. The use of the CC3 method as reference for the bright excitation of these systems is supported by the agreement of the CC3 excitation energies with the corresponding extrapolated FCI (exFCI) estimates in a small basis of the smaller CN3 and CN5 to better than 0.05 eV [12]. From a careful investigation, we know that employing the aug-cc-pVDZ basis set is sufficient given the agreement with the corresponding aug-cc-pVTZ values. Importantly, the all-electron CC3/aug-cc-pVDZ excitation energies are very close to the BFD CC3/aug-cc-pVDZ values, confirming that the use of pseudopotentials does not introduce appreciable errors. The reference CC3/aug-cc-pVDZ values also agree with the corresponding CC3 excitation energies computed with the BFD maug-cc-pVDZ basis set for all cyanines except the smallest CN3, where a fully-augmented double-ζ basis is needed also in the BFD calculations.

For dyes larger than CN15, we are however not able to run the CC3 calculations due to memory requirements [37] and the DMC excitation energy with our best CIPSI wave function becomes then the reference for other calculations.

4.4.1 Building the expansions

To compute accurate QMC excitation energies for the cyanine dyes, one needs balanced Jastrow-Slater wave functions to describe the ground and excited states. This is achieved in two stages, where the first is the construction of CIPSI expansions with the iso-PT2 and/or iso-variance scheme, and the second is a validation criterion that the resulting excitation energies in VMC and DMC are close to each other and converged with respect to the number of determinants. In particular, we generate the ground- and excited-state expansions at the CIPSI level to have either matched PT2 energy corrections or CI variances, which we use as measures of the "distance" of the wave functions from the FCI limit. Imposing that the determinantal components satisfy either the iso-PT2 or iso-variance criterion was previously found to lead to QMC excitation energies which were converged to the best reference values with a handful of determinants [17, 18], even when the error on the starting CI excitation energy was relatively large [19].

In Fig. 4.1, we illustrate the convergence of the CI excitation energies of CN3 and CN9 versus the total number of determinants for expansions characterized by similar PT2 corrections or CI variances. For CN3, the iso-PT2 construction leads to a somewhat faster convergence of the excitation energy for small expansions, but the two criteria become quickly equivalent beyond a few 1000 determinants. The situation is reversed for CN9, where matching the PT2 correction yields a much slower converging CI excitation energy, while the iso-variance criterion leads to a good agreement with the CC3 value in the same basis set for little more than 1000 determinants. In fact, we find that variance-matched expansions yield a faster converging CI excitation energy starting from CN7 and that, surprisingly, fewer determinants are needed to obtain a good estimate for the larger system sizes considered (see Fig. 4.7). Consequently, the CI treatment of the smallest cyanine, CN3, appears to be the most difficult as further elaborated in Sec. 4.4.3.
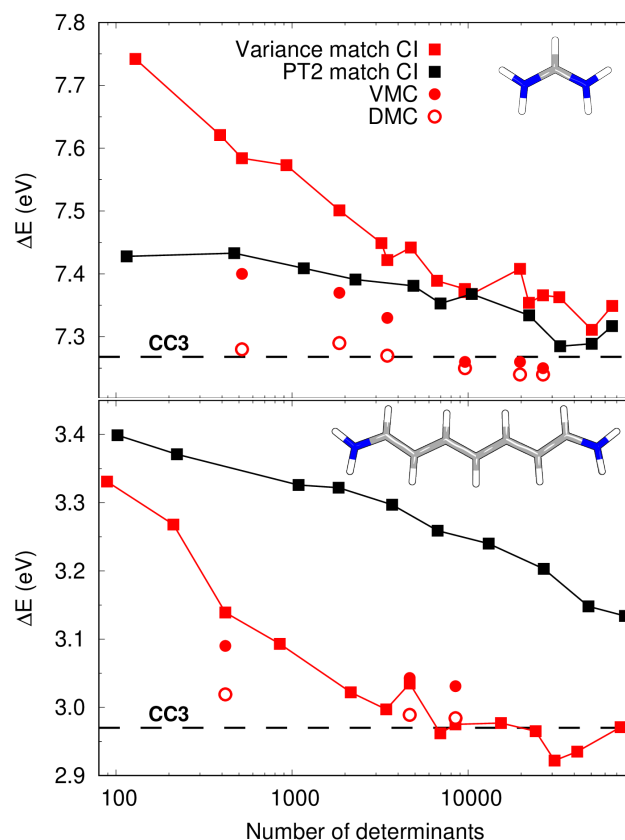
Figure 4.1: CI vertical excitation energies of CN3 (top) and CN9 (bottom) versus the total number of determinants, computed for ground- and excited-state CIPSI expansions having either matched PT2 energy contributions or CI variances. The VMC and DMC excitation energies obtained using the iso-variance expansions are also shown (the statistical error is smaller than the symbol size). The BFD pseudopotentials and the maug-cc-pVDZ basis are used here also for the CC3 calculations.

Importantly, in Fig. 4.1, we also show that QMC largely corrects for possible shortcomings of the starting CIPSI expansions, yielding excitation energies which display a rather small dependence on the number of determinants, especially at the DMC level. For CN3 and small expansions, where the iso-variance criterion significantly overestimates the CI excitation energy, VMC and DMC reduce the error at the CI level by about 0.2 and 0.3 eV, respectively. As the expansions become larger, the difference between the VMC and the DMC values diminishes, falling well below chemical accuracy (about 0.05 eV) for both CN3 and CN9. The robustness of the QMC results is further corroborated for CN7 in Table 4.6, where we show that, for comparable number of determinants, the use of PT2- and variance-matched wave functions yields excitation energies that are very close in VMC and completely equivalent in DMC (even though the two procedures give differences of about 0.2 eV at the CI level; see Fig. 4.7).

4.4.2 Best QMC vertical excitations

In Table 4.1, we summarize the VMC and DMC excitation energies of all cyanine dyes obtained with the largest CIPSI expansions of Table 4.6 and the iso-variance selection criterion. We also list the QMC and CASPT2 excitation energies computed with minimal CAS expansions, together with our CC2 and CC3 results and the exFCI estimates from the literature [12]. We refer the reader to Table 4.6 for additional QMC calculations with different numbers of determinants in the Jastrow-CIPSI wave functions.

For the reported CIPSI expansions, the VMC and DMC excitations energies are very close and also agree within chemical accuracy with the CC3 and exFCI values in all cases where these methods are applicable. This is in line with our previous findings that the agreement between VMC and DMC excitation energy is a strong indication of the balanced quality of the corresponding wave functions [19]. Furthermore, we find that the QMC values for the larger dyes are in very good agreement with the estimates given by the extrapolation of the CC3 results as a function of number of electrons. Since DMC can be employed in all cases, we plot all excitation energies in Fig. 4.2 in terms of their distance to the DMC-CIPSI results, which we use as reference values.
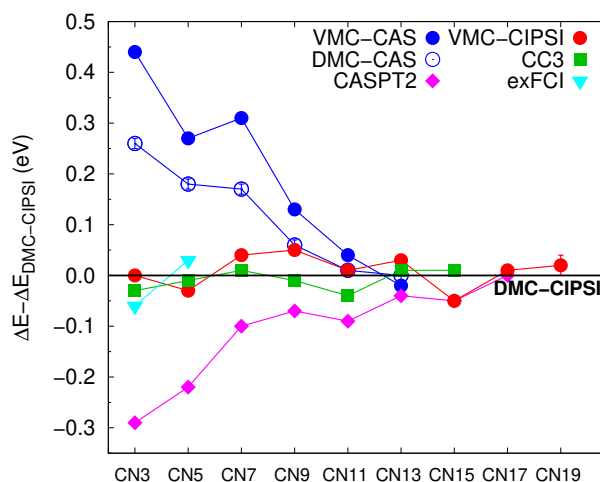
Figure 4.2: Excitation energies (eV) at different levels of theory with respect to the DMC values computed with the CIPSI wave functions (DMC-CIPSI line).

The CASPT2 and QMC-CAS energies computed with the minimal active spaces are instead very different from the DMC-CIPSI results: CASPT2 always underestimates the excitation energies, whereas QMC-CAS tends to overestimate them, similarly to what reported for CAS wave functions in Ref. [4]. For CN3–CN7, we test the effect of including more π orbitals in the active space, which somewhat ameliorates the VMC excitation energies but does not sufficiently affect the DMC values, which remain far from the DMC-CIPSI reference (see Table 4.7). Interestingly, we note that both the CASPT2 and QMC-CAS methods approach the best DMC results as the size of the molecule increases, suggesting an easier treatment of the longer

Table 4.1: Vertical excitation energies (eV) for the cyanine dyes computed with QMC and other highly-correlated methods. The BFD pseudopotentials are used in QMC, while all other calculations are all-electron. All energies are computed on PBE0/cc-pVQZ geometries.

Method	CN3[a]	CN5	CN7	CN9	CN11	CN13	CN15	CN17	CN19
VMC-CAS	7.67(1)	5.13(1)	3.97(1)	3.11(1)	2.58(1)	2.13(1)	–	–	–
DMC-CAS	7.49(1)	5.04(1)	3.83(1)	3.04(1)	2.55(1)	2.15(1)	–	–	–
VMC-CIPSI	7.23(1)	4.83(1)	3.65(1)	3.03(1)	2.55(1)	2.18(1)	1.85(1)	1.66(1)	1.59(2)
DMC-CIPSI	7.23(1)	4.86(1)	3.66(1)	2.98(1)	2.54(1)	2.15(1)	1.90(1)	1.65(1)	1.57(1)
CASPT2/aug-cc-pVDZ	6.94	4.64	3.56	2.91	2.45	2.11	1.85	1.65	–
CC2/aug-cc-oVDZ	7.29	4.97	3.80	3.10	2.64	2.30	2.04	1.84	–
CC3/aug-cc-pVDZ	7.20	4.85	3.67	2.97	2.50	2.16	1.91	–	–
exFCI/aug-cc-pVDZ [12]	7.17(2)	4.89(2)	–	–	–	–	–	–	–

[a] The QMC-CIPSI calculations for CN3 are performed with the aug-cc-pVTZ basis.

chains as already found at the CI level and as further discussed below.

4.4.3 Capturing orbital correlation

To understand the different performance of CAS and CIPSI expansions when used in QMC wave functions, we focus here on CN3 and analyze in Fig. 4.3 the VMC and DMC vertical excitation energies calculated using different determinantal components in the trial wave functions. As already mentioned, despite being the smallest cyanine dye, CN3 appears to be the most challenging one: the use of Jastrow-CAS wave functions leads to quite big errors and the number of CIPSI determinants needed to converge the excitation energy is larger than for the longer dyes. The
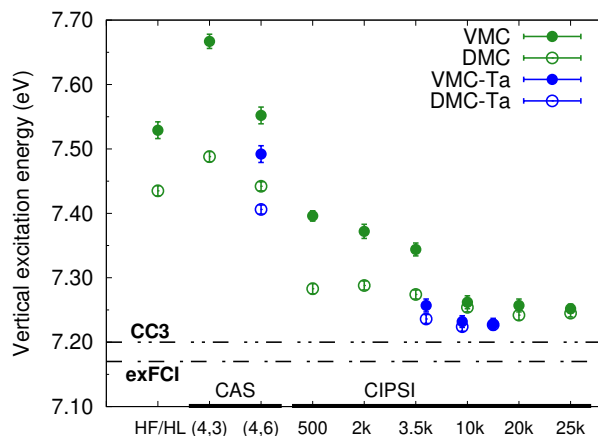

Figure 4.3: VMC (full circle) and DMC (empty circle) vertical excitation energies of CN3 for different wave functions. The maug-cc-pVDZ (green) and aug-cc-pVTZ (Ta, blue) are used.

simplest QMC calculations are performed with a one-configuration (HF/HL) wave function and the maug-cc-pVDZ basis set. We then proceed to CAS determinantal components and CIPSI expansions also employing the aug-cc-pVTZ basis set. The VMC excitation energy computed with the minimal CAS wave functions is worse than the HF/HL value since the active space comprises more determinants for the ground state but only the HL configuration for the excited state. DMC ameliorates the result but using a larger CAS space on the π orbitals only marginally helps (see Table 4.7). On the other hand, with the CIPSI selected determinants, we have a considerable improvement on the excitation energy and, with the use of just few hundred determinants, the DMC error reduces to less then 0.1 eV. Employing larger expansions with the maug-cc-pVDZ basis set, we finally converge to VMC values which are consistent with the DMC ones and approximately 0.04 eV higher than the reference. The use of the aug-cc-pVTZ basis set further reduces the excitation energy by about 0.02 eV. We note that, for the longer cyanine chains, the smaller maug-cc-pVDZ basis set is found to be sufficient for the computation of this excitation energy [18].

The superior performance of the use of a CIPSI with respect to the CAS expansions in QMC indicates that some key descriptor of correlation is missing from the active space and is not recovered through the addition of the Jastrow factor and the subsequent full optimization in VMC, nor through a DMC calculation with the optimal Jastrow-Slater wave function. In this work as in Ref. [4], the active space is chosen to correlate the π electrons in the π orbitals. From the QMC-CIPSI results, we can therefore infer that, while the excitation of interest is predominantly of $\pi \rightarrow \pi^*$ character, other orbital correlations are important and cannot be omitted in the QMC wave function of the shorter cyanines.

Table 4.2: CI total energies (a.u.) and vertical excitation energies (ΔE_{exc}, eV) of CN3 computed with the 6-31G basis set and different orbital sets. The last column reports the error with respect to the FCI excitation energy for CN3 and CN5, and with respect to the CC3 value for CN7.

	E(GS)	E(ES)	ΔE_{exc}	err
CN3				
1 CSF	-149.39966	-149.07223	8.91	1.39(2)
CAS-σ	-149.613(1)	-149.307(1)	8.32	0.80(2)
CAS-π	-149.44486	-149.14840	8.07	0.55(2)
CAS-π + SD-σ	-149.7151(5)	-149.4346(5)	7.65	0.13(2)
CC3	-149.74049	-149.46354	7.54	0.02(2)
FCI	-149.741(1)	-149.465(1)	7.52(2)	–
CN5				
1 CSF	-226.27705	-226.04468	6.32	1.48(1)
CAS-σ	-226.581(1)	-226.373(1)	5.66	0.82(1)
CAS-π	-226.34204	-226.15212	5.17	0.33(1)
CAS-π + SD-σ	-226.745(2)	-226.557(2)	5.03	0.19(1)
CC3	-226.80736	-226.62972	4.83	-0.01(1)
FCI	-226.809(1)	-226.631(1)	4.84(1)	–
CN7				
1 CSF	-303.14723	-302.95611	5.20	1.64
CAS-σ	-303.580(4)	-303.409(4)	4.73(4)	1.17(4)
CAS-π	-303.23260	-303.09606	3.72	0.16
CC3	-303.86766	-303.73676	3.56	–

To better understand this, we present a CI study for CN3–CN7 with the small 6-31G basis set in Table 4.2. We correlate only the valence electrons and use state-average natural orbitals obtained with a preliminary calculation at the CIPSI level. For each state, we compute the energy with only one CSF and, on top of this configuration, we perform a CAS-CI calculation restricted to the σ and the π orbitals in a CAS-σ and CAS-π, respectively. The reference FCI excitation energy in this basis and the associated confidence interval are computed following the scheme presented in Ref. [38] rather than extrapolating the variational energies of the individual states in the limit of the PT2 energy correction going to zero. Indeed, the uncertainties of the extrapolated FCI energies of both states is larger than the uncertainty on the

estimated excitation energy computed with this scheme. Since the CC3 estimate for CN3 and CN5 is in excellent agreement with the FCI value, we use the CC3 excitation energy as reference for CN7. We note that, because of the use of the simple 6-31G basis set, the FCI and CC3 excitation energies are much higher than the more accurate results presented above but this is not relevant for the present discussion.

For the CN3 molecule, the excitation energy obtained with a single CSF for each state is 8.91 eV, namely, higher by 1.4 eV than the FCI result. The CAS-π calculation corrects only 60% of the error, indicating that the σ orbitals also play an important role in the stabilization of the excited state. Similarly, the excitation energy obtained with the CAS-σ improves the excitation energy with respect to the single CSF by recovering about 31% of the error. These results indicate the importance of both σ and π orbitals in the calculation of the excitation energy of CN3.

Therefore, to partially account for both π and σ correlations, we perform a multi-reference CI calculation, applying all possible single and double excitations to the CAS-π determinants. Such a CAS-π+SD-σ calculation also enables the relaxation of the CAS-π CI coefficients in the presence of most of the σ correlation. The resulting excitation energy is now significantly improved but still 0.1 eV higher than the FCI reference, confirming that a similar computational effort needs to be made for the π and σ orbitals. This justifies the use of CIPSI where the most important Slater determinants will be chosen to describe σ, π, and $\sigma - \pi$ correlation in a "democratic" way based on their contribution to the second-order perturbation energy.

For CN5, the situation is somewhat different. While the single CSF still overestimates the excitation energy by 1.47 eV, the CAS-π wave function behaves better than for CN3, recovering 80% of the error. Consequently, omitting the σ orbitals in the active space results in an excitation energy closer to the reference than in the CN3 case. Once the σ orbitals are introduced, as for CN3, we improve the excitation energy but still observe an overestimation of the CAS-π+SD-σ result by almost 0.2 eV, pointing to the importance of describing the σ as well as the π correlation. The situation for CN7 is similar as for CN5, suggesting that the whole series behaves like CN5 and that CN3 is an exception because of the particularly small length of the chain.

4.5 Conclusion

We have presented a QMC benchmark study of the lowest vertical excitation energies of cyanine chains. We constructed the determinantal components of the Jastrow-Slater wave functions through an automatic selected-CI procedure and obtained a balanced description of the relevant states by ensuring similar quality of the corresponding expansions, for instance by matching their CI variances. With compact expansions of only a few thousand determinants, upon optimization of all parameters in our wave functions, we obtained QMC excitation energies which improve on the starting CI values and, for the shorter chain lengths where CC3 calculations are feasible, agree with the CC3 results to chemical accuracy. We also applied our protocol to longer cyanines and validated the accuracy of our estimates via the con-

sistent closeness of the determined VMC and DMC excitation energies. Finally, we showed that key to a successful description of this excitation over all chain lengths is to account for π, σ, and $\sigma - \pi$ correlations, therefore going beyond a CAS treatment based on π-orbitals only. In conclusion, we believe that the present study further establishes QMC methods as accurate and robust tools for the treatment of excited states of relatively large systems and parameter spaces.

4.A Appendix: CIPSI and QMC results for CN3

In Table 4.3, we report the CI energies, PT2 energies, and CI variances for the CIPSI expansions of CN3 used in the QMC-CIPSI calculations of Figure 4.3. The values are relative to single-state CIPSI wave functions of increasing size, that are separately expanded for the ground and excited states until when the corresponding CI variances match. We consider the CI variances matched if their difference is less than 0.002 a.u.

In Figure 4.4, we plot the CI excitation energies of CN3 versus the total (GS+ES) number of determinants. The excitation energies are obtained by matching the variance or the PT2 energy correction. Even though the PT2-matched excitation energy converges faster for small expansions, matching the CI variance automatically leads to expansions with also matched PT2 energy correction beyond few thousands total (GS+ES) determinants. We also show that the use of CAS(4,3) or CAS(4,6) optimized orbitals does not have an impact on the CIPSI convergence.

Table 4.3: CI energies, CI variances, and PT2 energy corrections (a.u.) of the ground- and excited-state CIPSI expansions of CN3. The expansions are matched according to the iso-variance criterion. The CI excitation energy ΔE_{CI} (eV) is also listed. We use the orbitals obtained from a CASSCF(6,10) calculation of symmetry A_1 and B_1 for the ground (GS) and excited (ES) states, respectively. The BFD pseudopotentials and corresponding basis sets are used.

No. det		E_{CI}		ΔE_{CI} (eV)	PT2		ΔPT2	σ^2_{CI}		$\Delta\sigma^2_{CI}$
GS	ES	GS	ES		GS	ES		GS	ES	
				maug-cc-pVDZ basis set						
220	300	-27.7557	-27.4770	7.584	-0.4625	-0.4705	-0.0080	1.5994	1.5984	0.0010
602	1254	-27.7898	-27.5141	7.501	-0.4167	-0.4217	-0.0050	1.5050	1.5039	0.0011
1028	2446	-27.8124	-27.5396	7.422	-0.3879	-0.3905	-0.0026	1.4391	1.4373	0.0018
2511	7092	-27.8618	-27.5907	7.376	-0.3296	-0.3300	-0.0004	1.2905	1.2885	0.0020
5005	14696	-27.9085	-27.6363	7.408	-0.2773	-0.2799	-0.0026	1.1420	1.1395	0.0025
6508	20144	-27.9285	-27.6578	7.366	-0.2557	-0.2575	-0.0018	1.0733	1.0744	-0.0011
				aug-cc-pVTZ basis set						
1483	3246	-27.8102	-27.5367	7.440	-0.5197	-0.5238	-0.0041	2.5206	2.5195	0.0011
2569	6432	-27.8291	-27.5574	7.395	-0.4945	-0.4976	-0.0031	2.4594	2.4587	0.0007
4451	12008	-27.8510	-27.5800	7.373	-0.4672	-0.4698	-0.0027	2.3868	2.3881	0.0013

Going to larger CIPSI expansions, we can also compute the extrapolated FCI excitation energy by fitting the ground- and excited-state CI energies versus the renormalized PT2 (rPT2) energy corrections [29], which become zero in the FCI limit. In Fig. 4.5, we show the convergence of the CI energies as well as their polynomial fits, which are performed over the interval [-0.1:0] of rPT2 values. We chose to extrapolate using the rPT2 energy corrections (instead of the PT2 ones) since the behavior is more linear as also shown in the Figure. The estimated FCI/maug-cc-pVDZ energy is 7.25 eV, close to the CC3/maug-cc-pVDZ (BFD) value of 7.27 eV.

In Table 4.4, we list the VMC and DMC excitation energies of CN3 computed with different determinantal components in the wave functions and with two basis sets. We employ CIPSI expansions with matched CI variances.
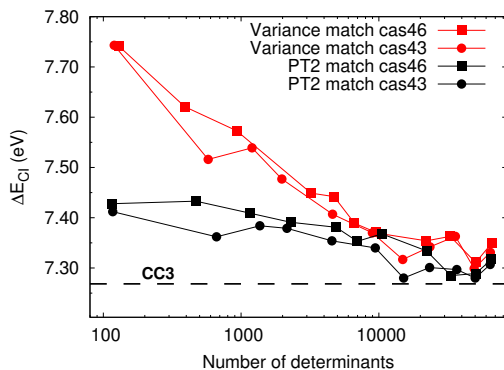
Figure 4.4: CI excitation energies of CN3 computed for expansions of increasing size with matched variance or PT2 energy correction and two different CASSCF orbital sets generated with a CAS(4,3) and a CAS(4,6) calculation. The BFD pseudopotentials and the corresponding maug-cc-pVDZ basis set is used for all calculations.
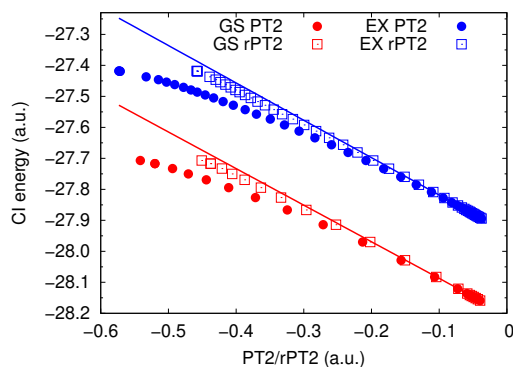

Figure 4.5: Ground- and excited-state CI energies of CN3 versus the rPT2 energy corrections and corresponding fits. The CI energies are also plotted against the PT2 energy corrections.

In Fig. 4.6, we plot the VMC/maug-cc-pVDZ energies of the two states versus their VMC variances. As proposed in Refs. 39, 40, we fit the VMC energies for the ground and excited state against the corresponding VMC variances and estimate the excitation energy as the difference $\Delta E_{\text{VMC}}^{\text{fit}} = E_{\text{ES}}^{\text{fit}}(\sigma^2) - E_{\text{GS}}^{\text{fit}}(\sigma^2)$ of the fits of the energies. The resulting excitation energy is consistently overestimated and, as the variance becomes smaller, further departs from the reference CC3 value.

Table 4.4: VMC and DMC total energies (a.u.) and excitation energies (eV) of CN3 obtained for different wave functions. All the variational parameters (Jastrow, orbital, and CI coefficients) are optimized in energy minimization. We employ the BFD maug-cc-pVDZ basis set, unless we label the calculation with "Ta" which denotes the use of the BFD aug-cc-pVTZ basis set.

| WF	No. det		No. parm		E_{VMC}		ΔE_{VMC}	E_{DMC}		ΔE_{DMC}
	GS	ES	GS	ES	GS	ES		GS	ES	
HF/HL	1	2	220	230	-28.3005(3)	-28.0238(3)	7.529(13)	-28.3666(2)	-28.0934(2)	7.435(8)
CAS(4,3)	5	4	232	230	-28.3057(2)	-28.0238(3)	7.667(11)	-28.3693(2)	-28.0941(2)	7.488(8)
CAS(4,6)	113	112	302	293	-28.3077(4)	-28.0302(3)	7.552(13)	-28.3694(2)	-28.0959(2)	7.442(7)
CAS(4,6)-Ta	113	112	1008	999	-28.3165(3)	-28.0412(3)	7.492(13)	-28.3706(2)	-28.0985(2)	7.406(7)
CIPSI	220	300	690	880	-28.3201(2)	-28.0483(2)	7.396(08)	-28.3715(2)	-28.1039(2)	7.283(8)
	602	1254	1179	1893	-28.3334(3)	-28.0625(3)	7.372(11)	-28.3752(2)	-28.1074(2)	7.288(7)
	1028	2446	1653	2293	-28.3375(3)	-28.0682(3)	7.326(10)	-28.3768(2)	-28.1095(2)	7.274(7)
	2511	7092	2624	3872	-28.3444(3)	-28.0776(3)	7.262(10)	-28.3806(2)	-28.1140(2)	7.254(7)
	5005	14696	3688	6145	-28.3502(3)	-28.0835(3)	7.257(10)	-28.3830(2)	-28.1168(2)	7.242(8)
	6508	20144	4234	7648	-28.3526(2)	-28.0861(2)	7.252(07)	-28.3843(2)	-28.1180(2)	7.245(6)
CIPSI-Ta	1483	3246	8395	10445	-28.3427(2)	-28.0760(3)	7.257(09)	-28.3773(2)	-28.1114(2)	7.236(7)
	2569	6432	10988	14765	-28.3461(2)	-28.0803(2)	7.232(09)	-28.3783(2)	-28.1128(2)	7.224(7)
	4451	12008	14780	19237	-28.3515(2)	-28.0859(2)	7.228(09)	-28.3802(2)	-28.1146(2)	7.227(6)
CC3/aug-cc-pVDZ										7.20
CC3/aug-cc-pVTZ										7.19
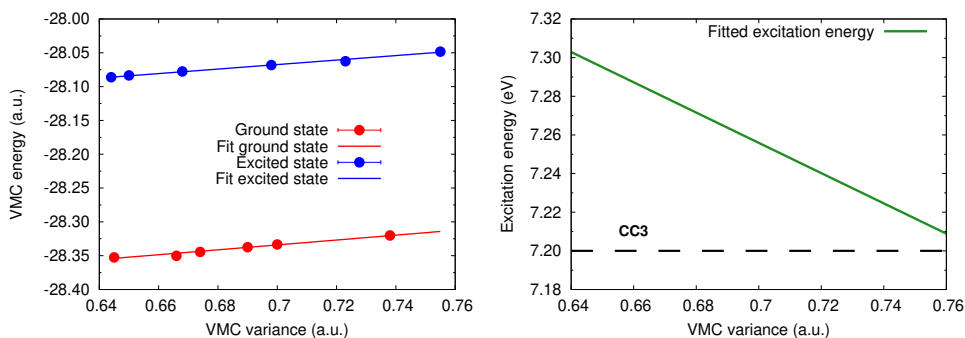
Figure 4.6: VMC ground- and excited-state energies versus the corresponding VMC variances (left) and resulting fitted excitation energy (right). The reference CC3/aug-cc-pVDZ value is shown.

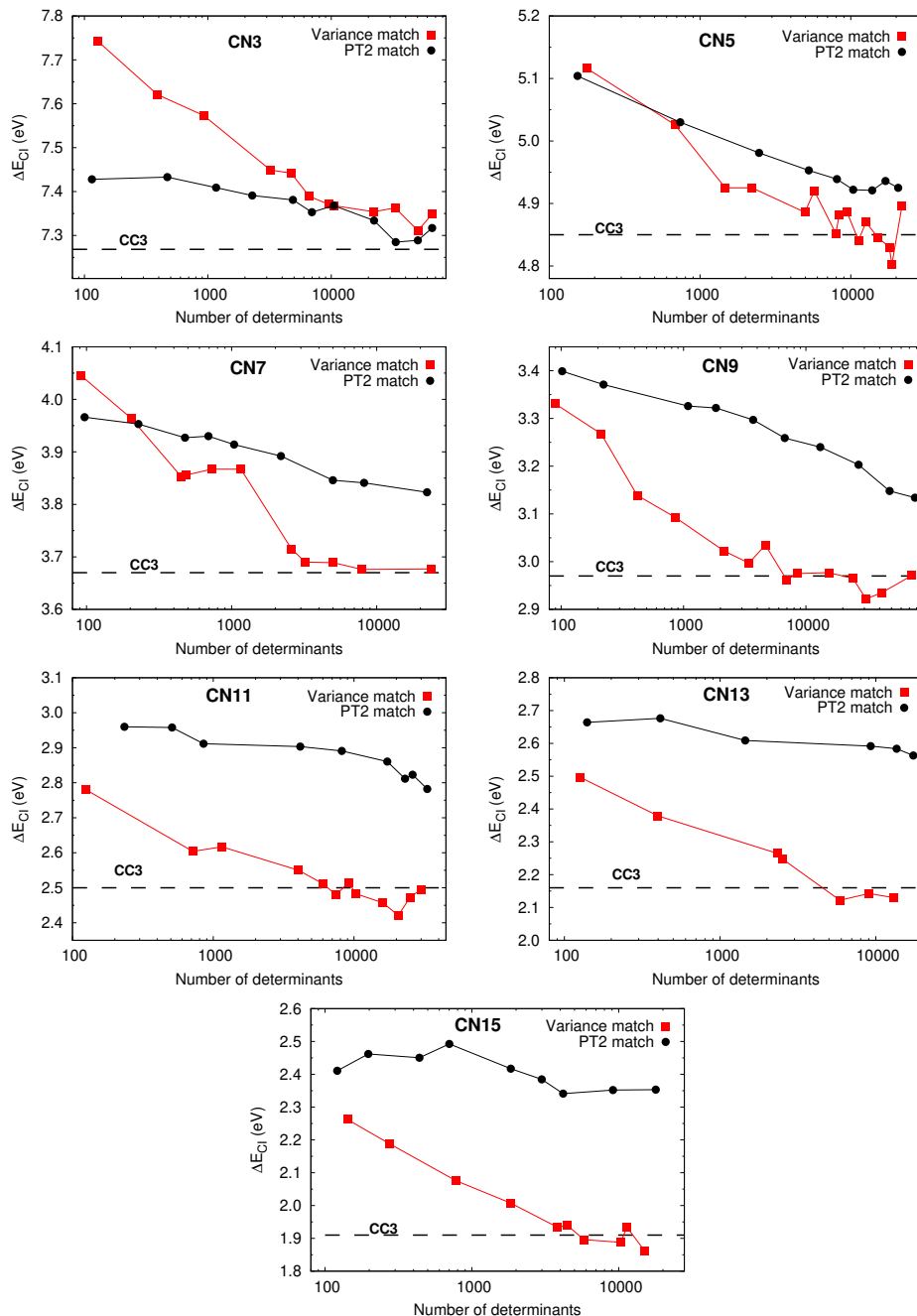# 4.B   Appendix: CIPSI results for all molecules

Figure 4.7: CI excitation energies of CN3–CN15 for iso-PT2 or iso-variance CIPSI expansions versus the total (GS+ES) number of determinants.

In Fig. 4.7, we show the convergence as a function of the total number of determinants of the CI excitation energies of CN3–CN15 obtained with ground- and excited-state CIPSI expansions with either matched PT2 energy corrections or CI variances. The iso-variance procedure leads to a faster convergence of the CI excitation energy

for the cyanine molecules larger than CN3 and is chosen to generate the CIPSI expansions for the calculations of the QMC excitation energies of all molecules. For all the cases in Fig. 4.7, the CIPSI expansions are performed on CASSCF optimized orbitals.

For CN17 and CN19, we perform instead the CIPSI expansions on HF orbitals and show the CI convergence of their excitation energies for the iso-variance procedure[1] in Fig. 4.8 and further discuss the use of HF orbitals in Section 4.B.1.
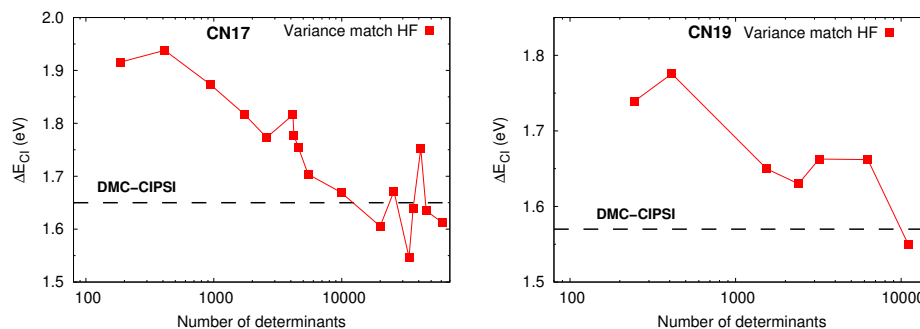

Figure 4.8: CI excitation energies of CN17 and CN19 for iso-variance CIPSI expansions performed on HF orbitals, versus the total (GS+ES) number of determinants.

4.B.1 Appendix: Use of HF orbitals in CIPSI expansions

Since, for large molecules, the CASSCF computation becomes quite expensive, we have used HF orbitals to perform the CIPSI calculation for CN17 and CN19. Here, we validate this choice with two different tests.
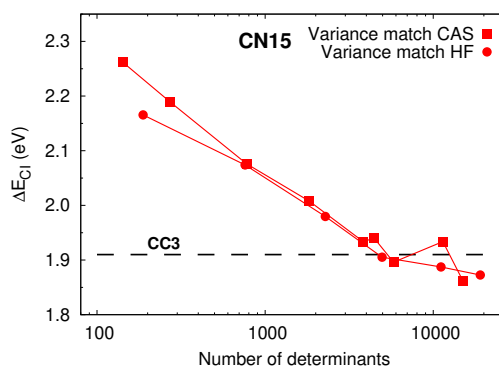

Figure 4.9: Convergence of variance-matched CI excitation energies of CN15 obtained with CAS(16,15) and HF orbitals, versus the total (GS+ES) number of determinants.

In Fig. 4.9, we compare the convergence of the iso-variance CI excitation energy of CN15 obtained with expansions on CASSCF and HF orbitals. We observe a very

---

[1]We do not show the iso-PT2 curves since, in the shown determinantal range, the iso-PT2 criterion leads to excitation energies very far from convergence.

similar convergence for the two different orbitals, a finding which makes us confident in the use of HF orbitals for the larger chains. As an additional test, in Table 4.5, we report the QMC excitation energies of CN3, obtained with CIPSI determinantal expansions on CASSCF or HF orbitals. We find that, both at the VMC and the DMC level, the results obtained with expansions on HF orbitals are compatible with the CASSCF-based results with a similar number of determinants.

Table 4.5: VMC and DMC total energies (a.u.) and excitation energies (eV) of CN3 obtained with CIPSI expansions on either CASSCF or HF orbitals. All the variational parameters (Jastrow, orbital, and CI coefficients) are optimized in energy minimization. The maug-cc-pVDZ basis set is employed.

Starting orb.	No. det		No. parm		E_{VMC}		ΔE_{VMC}	E_{DMC}		ΔE_{DMC}
CASSCF	2511	7092	2624	3872	-28.3444(3)	-28.0776(3)	7.262(10)	-28.3806(2)	-28.1140(2)	7.254(7)
HF	2525	7374	2587	4066	-28.3452(3)	-28.0775(3)	7.282(10)	-28.3808(2)	-28.1140(1)	7.259(6)

4.C   Appendix: QMC results

In Table 4.6, we present the VMC and DMC total and vertical excitation energies obtained with the CIPSI expansions. In addition to the "best" excitation energies presented in the main text, we report more estimates obtained with different wave functions. For CN7, we perform QMC calculations both for CIPSI expansions with matched PT2 corrections and variances, while we only employ the iso-variance wave functions for the other molecules.

For CN7, for a similar number of determinants, the iso-PT2 selection gives slightly less accurate VMC excitation energies than the iso-variance case. However, at the DMC level, the differences are cured and the excitation energies are compatible.

In Table 4.7, we report the VMC and DMC total and vertical excitation energies obtained with CAS expansions. We employ the minimal CAS over π orbitals and, for the smaller CN3–CN5, also a larger space always over π orbitals. For CN11 and CN13, we truncate the wave functions using the same percentage of the total weight for the ground and excited states. For larger dyes, we do not compute the QMC-CAS excitation energies since the CASSCF computation becomes computationally too demanding.

Table 4.6: VMC and DMC total energies (a.u.) and excitation energies (eV) of cyanine dyes obtained for CIPSI wave functions. All the variational parameters (Jastrow, orbital, and CI coefficients) are optimized in energy minimization. The maug-cc-pVDZ basis set is employed.

System	No. det		No. parm		E_{VMC}		ΔE_{VMC}	E_{DMC}		ΔE_{DMC}
	GS	ES	GS	ES	GS	ES		GS	ES	
					PT2 matched					
CN7	2147	6038	4349	6613	-53.4187(4)	-53.2810(3)	3.748(14)	-53.5172(2)	-53.3807(2)	3.714(8)
					Variance matched					
CN5[a]	376	1094	1567	2609	-40.8646(3)	-40.6842(3)	4.908(12)	-40.9467(3)	-40.7665(3)	4.905(10)
	1344	4382	2478	4531	-40.8798(3)	-40.7013(3)	4.857(13)	-40.9502(2)	-40.7711(2)	4.872(09)
	2460	8782	3555	6561	-40.8896(3)	-40.7099(3)	4.890(12)	-40.9532(2)	-40.7748(2)	4.856(09)
	3913	14114	4842	8312	-40.8941(2)	-40.7167(3)	4.828(11)	-40.9559(2)	-40.7775(2)	4.856(08)
CN7	1871	6038	4105	6613	-53.4162(4)	-53.2810(3)	3.681(15)	-53.5170(2)	-53.3807(2)	3.708(8)
	5578	18300	7783	12098	-53.4338(3)	-53.2996(3)	3.651(12)	-53.5227(2)	-53.3883(2)	3.656(8)
CN9	132	286	2074	2268	-65.9167(4)	-65.8031(4)	3.090(15)	-66.0727(3)	-65.9618(3)	3.019(11)
	1289	3372	4785	7272	-65.9374(3)	-65.8255(3)	3.043(12)	-66.0772(3)	-65.9674(3)	2.989(10)
	2120	6360	5741	9345	-65.9482(5)	-65.8373(3)	3.031(12)	-66.0800(3)	-65.9704(3)	2.984(10)
CN11	1071	2970	5320	7403	-78.4706(3)	-78.3726(4)	2.668(13)	-78.6406(3)	-78.5471(3)	2.544(11)
	1812	5634	6606	10453	-78.4749(4)	-78.3795(4)	2.598(13)	-78.6422(3)	-78.5479(2)	2.566(10)
	2284	6896	7412	11729	-78.4777(4)	-78.3841(4)	2.547(14)	-78.6441(3)	-78.5507(3)	2.542(11)
CN13	714	1810	5306	5985	-90.9883(4)	-90.9056(4)	2.250(14)	-91.1986(3)	-91.1206(3)	2.124(11)
	1406	4464	7351	10297	-90.9966(4)	-90.9165(4)	2.180(14)	-91.2013(3)	-91.1224(3)	2.147(10)
CN15	1356	4434	7549	11347	-103.5244(3)	-103.4538(3)	1.920(12)	-103.7647(3)	-103.6934(3)	1.938(11)
	2241	7996	9951	15479	-103.5349(3)	-103.4668(3)	1.854(12)	-103.7674(3)	-103.6975(3)	1.902(11)
CN17	559	2028	6995	9184	-116.0318(3)	-115.9735(4)	1.586(14)	-116.3160(3)	-116.2566(3)	1.616(13)
	943	3288	8404	11396	-116.0432(4)	-115.9814(4)	1.681(14)	-116.3220(3)	-116.2598(3)	1.692(12)
	1384	5068	10592	14978	-116.0502(4)	-115.9890(4)	1.663(14)	-116.3231(3)	-116.2624(3)	1.653(10)
CN19	515	1886	8496	10615	-128.5631(4)	-128.5069(4)	1.529(16)	-128.8792(3)	-128.8239(3)	1.505(13)
	1487	4806	12876	15512	-128.5828(4)	-128.5242(4)	1.595(16)	-128.8872(3)	-128.8294(3)	1.572(13)

[a] Data from our work in Ref. 18.

Table 4.7: VMC and DMC total energies (a.u.) and excitation energies (eV) of cyanine dyes obtained for CAS wave functions. All the variational parameters (Jastrow, orbital, and CI coefficients) are optimized in energy minimization. The maug-cc-pVDZ basis set is employed.

System	CAS	No. det		No. parm		E_{VMC}		ΔE_{VMC}	E_{DMC}		ΔE_{DMC}
	(n,m)	GS	ES	GS	ES	GS	ES		GS	ES	
CN3	4,3	5	4	232	230	-28.3057(2)	-28.0238(3)	7.667(11)	-28.3693(2)	-28.0941(2)	7.488(8)
	4,6	113	112	302	293	-28.3077(4)	-28.0302(3)	7.552(13)	-28.3694(2)	-28.0959(2)	7.442(7)
CN5[a]	6,5	52	48	567	561	-40.8468(4)	-40.6583(4)	5.130(15)	-40.9433(3)	-40.7582(2)	5.038(10)
	6,10	7232	7168	3134	3064	-40.8498(4)	-40.6628(4)	5.090(15)	-40.9439(3)	-40.7594(3)	5.022(11)
CN7	8,7	625	600	1250	1220	-53.3750(4)	-53.2291(4)	3.969(15)	-53.5056(2)	-53.3649(3)	3.828(10)
	8,14	6802	5158	2675	2302	-53.3781(4)	-53.2421(4)	3.700(16)	-53.5057(2)	-53.3702(2)	3.686(6)
CN9	10,9	7956	7920	4257	4197	-65.9151(4)	-65.8007(4)	3.114(15)	-66.0728(2)	-65.9610(2)	3.042(9)
CN11[b]	12,11	7642	13360	3937	4914	-78.4485(4)	-78.3535(4)	2.585(14)	-78.6376(3)	-78.5436(3)	2.555(12)
CN13[c]	14,13	529	3338	3445	3949	-90.9798(4)	-90.9017(4)	2.130(14)	-91.1993(3)	-91.1204(3)	2.146(11)

[a] Data from our work in Ref. 18.
[b] Truncated so that, for each state, $\sum c_i^2 = 0.9985$ of the total wave function.
[c] Truncated so that, for each state, $\sum c_i^2 = 0.9765$ of the total wave function.

Bibliography

[1] M. Levitus and S. Ranjit, Q. Rev. Biophys. **44**, 123 (2011).

[2] H. Shindy, Dyes Pigm. **145**, 505 (2017).

[3] C. I. Oprea, P. Panait, Z. M. Essam, R. M. Abd El-Aal, and M. A. Gîrțu, Nanomaterials **10**, 662 (2020).

[4] R. Send, O. Valsson, and C. Filippi, J. Chem. Theory Comput. **7**, 444 (2011).

[5] D. Jacquemin, Z. Zhao, R. Valero, C. Adamo, I. Ciofini, and D. Truhlar, J. Chem. Theory Comput. **8**, 1255 (2012).

[6] B. Moore and J. Autschbach, J. Chem. Theory Comput. **9**, 4991 (2013).

[7] H. Zhekova, M. Krykunov, J. Autschbach, and T. Ziegler, J. Chem. Theory Comput. **10**, 3299 (2014).

[8] P. Boulanger, D. Jacquemin, I. Duchemin, and X. Blase, J. Chem. Theory Comput. **10**, 1212 (2014).

[9] M. Filatov and M. Huix-Rotllant, J. Chem. Phys. **141**, 024112 (2014).

[10] L. G. Boris and J. Denis, Acc. Chem. Res. **48**, 530 (2015).

[11] N. Minezawa, Chem. Phys. Lett. **622**, 115 (2015).

[12] G. Yann, S. Anthony, G. Emmanuel, C. Michel, and P.-F. Loos, J. Chem. Phys. **149**, 064103 (2018).

[13] S. Sorella and L. Capriotti, J. Chem. Phys. **133**, 234111 (2010).

[14] E. Neuscamman, C. J. Umrigar, and G. K.-L. Chan, Phys. Rev. B **85**, 045103 (2012).

[15] C. Filippi, R. Assaraf, and S. Moroni, J. Chem. Phys. **144**, 194105 (2016).

[16] R. Assaraf, S. Moroni, and C. Filippi, J. Chem. Theory Comput. **13**, 5273 (2017).

[17] M. Dash, J. Feldt, S. Moroni, A. Scemama, and C. Filippi, J. Chem. Theory Comput. **15**, 4896 (2019).

[18] A. Cuzzocrea, A. Scemama, W. J. Briels, S. Moroni, and C. Filippi, J. Chem. Theory Comput. **16**, 4203 (2020).

[19] M. Dash, S. Moroni, C. Filippi, and A. Scemama, J. Chem. Theory Comput. **17**, 3426 (2021).

[20] B. Huron, J. Malrieu, and P. Rancurel, J. Chem. Phys. **58**, 5745 (1973).

[21] As Jastrow factor, we use the exponential of the sum of two fifth-order polynomials of the electron-nuclear and the electron-electron distances, respectively, and rescale the inter-particle distances as $R = (1 - \exp(-\kappa r))/\kappa$ with κ set to 0.6 a.u. We employ different electron-nucleus Jastrow factors to describe the correlation of an elecron with C, N, and H. The total number of free parameters to be optimized in the Jastrow factor is 17 for the systems considered here.

[22] P. S. Epstein, Phys. Rev. **28**, 695 (1926).

[23] R. K. Nesbet, in *Proc. R. Soc. Lond. A*, The Royal Society (PUBLISHER, 1955), No. 1182, pp. 312–321.

[24] M. Burkatzki, C. Filippi, and M. Dolg, J. Chem. Phys. **126**, 234105 (2007).

[25] For the hydrogen atom, we use a more accurate BFD pseudopotential and basis set. Dolg, M.; Filippi, C., private communication.

[26] R. A. Kendall, T. H. Dunning Jr, and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).

[27] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, and others, J. Comput. Chem. **14**, 1347 (1993).

[28] M. S. Gordon and M. W. Schmidt, *Theory and applications of computational chemistry* (Elsevier, 2005), pp. 1167–1189.

[29] Y. Garniron, T. Applencourt, K. Gasperich, A. Benali, A. Ferté, J. Paquier, B. Pradines, R. Assaraf, P. Reinhardt, J. Toulouse, P. Barbaresco, N. Renon, G. David, J.-P. Malrieu, M. Véril, M. Caffarel, P.-F. Loos, E. Giner, and A. Scemama, J. Chem. Theory Comput. **15**, 3591 (2019).

[30] CHAMP is a quantum Monte Carlo program package written by C. J. Umrigar, C. Filippi, S. Moroni and collaborators.

[31] C. Attaccalite and S. Sorella, Phys. Rev. Lett. **100**, 114501 (2008).

[32] S. Sorella, M. Casula, and D. Rocca, J. Chem. Phys. **127**, 014105 (2007).

[33] M. Casula, Phys. Rev. B **74**, 161102 (2006).

[34] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery,

J. J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox., Gaussian09, Revision A.02; Gaussian, Inc., Wallingford CT, 2016.

[35] J. F. Stanton, J. Gauss, L. Cheng, M. E. Harding, D. A. Matthews, and P. G. Szalay, CFOUR, Coupled-Cluster techniques for Computational Chemistry, a quantum-chemical program package, With contributions from A.A. Auer, R.J. Bartlett, U. Benedikt, C. Berger, D.E. Bernholdt, S. Blaschke, Y. J. Bomble, S. Burger, O. Christiansen, D. Datta, F. Engel, R. Faber, J. Greiner, M. Heckert, O. Heun, M. Hilgenberg, C. Huber, T.-C. Jagau, D. Jonsson, J. Jusélius, T. Kirsch, K. Klein, G.M. KopperW.J. Lauderdale, F. Lipparini, T. Metzroth, L.A. Mück, D.P. O'Neill, T. Nottoli, D.R. Price, E. Prochnow, C. Puzzarini, K. Ruud, F. Schiffmann, W. Schwalbach, C. Simmons, S. Stopkowicz, A. Tajti, J. Vázquez, F. Wang, J.D. Watts and the integral packages MOLECULE (J. Almlöf and P.R. Taylor), PROPS (P.R. Taylor), ABACUS (T. Helgaker, H.J. Aa. Jensen, P. Jørgensen, and J. Olsen), and ECP routines by A. V. Mitin and C. van Wüllen. For the current version, see http://www.cfour.de.

[36] F. Aquilante, J. Autschbach, R. K. Carlson, L. F. Chibotaru, M. l. G. Delcey, L. De Vico, I. Fdez. Galván, N. Ferré, L. M. Frutos, L. Gagliardi, M. Garavelli, A. Giussani, C. E. Hoyer, G. Li Manni, H. Lischka, D. Ma, P. r. Malmqvist, T. Müller, A. Nenov, M. Olivucci, T. B. Pedersen, D. Peng, F. Plasser, B. Pritchard, M. Reiher, I. Rivalta, I. Schapiro, J. Segarra-Martí, M. Stenrup, D. G. Truhlar, L. Ungur, A. Valentini, S. Vancoillie, V. Veryazov, V. P. Vysotskiy, O. Weingart, F. Zapata, and R. Lindh, J. Comput. Chem. **37**, 506 (2016).

[37] The CC calculations performed with CFour code are run on a dual socket AMD Epyc 7402 with 256GB of memory.

[38] M. Véril, A. Scemama, M. Caffarel, F. Lipparini, M. Boggio-Pasqua, D. Jacquemin, and P.-F. Loos, WIREs Comput. Mol. Sci. **n/a**, e1517 (2021).

[39] P. J. Robinson, S. D. Pineda Flores, and E. Neuscamman, J. Chem. Phys. **147**, 164114 (2017).

[40] S. D. Pineda Flores and E. Neuscamman, The Journal of Physical Chemistry A **123**, 1487 (2019).

Chapter 5

Dynamics with QMC forces: dealing with noisy forces

Thanks to recent methodological developments, it has become possible to employ accurate quantum Monte Carlo forces for the efficient structural optimization of molecular systems but their use for the study of dynamical processes is still unexplored. Here, we focus on one of the most commonly used quantum Monte Carlo variants, namely, variational Monte Carlo, and investigate its performance in a molecular dynamics setting. In particular, while employing quantum Monte Carlo forces to drive a molecular dynamics simulation, one encounters various issues which are not typical in deterministic quantum simulations. In this Chapter, we focus on the consequences on the dynamics of the presence of stochastic noise in the forces, and illustrate the strategies we have developed to minimize its effects.

5.1 Introduction

The field of atomistic simulations has undergone remarkable advances, and it is now possible to study the dynamics of large proteins or cell membranes for time scales as long as nanoseconds [1–3]. Concurrently, *ab initio* molecular dynamics has also been pushed to bigger systems sizes and, for example, with density functional theory (DFT), one can now treat thousands of atoms for several picoseconds [4–9]. Furthermore, the distance between quantum dynamical methods and realistic problems has been significantly reduced thanks to the enormous progress in multi-scale approaches combining quantum with classical models [10–12], most recently in the context of machine learning schemes [13, 14].

The description of photo-activated processes, however, remains complicated since it requires the balanced description of multiple electronic states whose nature might change significantly as the nuclear coordinates evolve. In the literature, such studies have typically been performed with a rather limited number of electronic structure methods, which are able to offer a reasonable compromise of accuracy and efficiency in the computation of the electronic forces in the excited state. Most commonly, the complete-active-space self-consistent (CASSCF) approach is employed in pho-

tochemical studies but, while the method can provide a satisfactory description of static correlation for small molecules, compromises must be made on the size of the active space as the system becomes larger.

Among available *ab initio* methods, quantum Monte Carlo (QMC) techniques have seen significant methodological improvements in recent years and have become an established electronic structure tool for studying molecular and extended systems. In particular, it is nowadays possible to compute energy derivatives at a similar computational cost to the one of computing the energy alone [15, 16]. Consequently, we can now routinely optimize the variational parameters in the many-body wave functions and compute interatomic forces within variational Monte Carlo (VMC). Moreover, thanks to recent work [17–20], we have stable procedures for the generation and simultaneous optimization of multiple wave functions for ground and excited states. Currently, these procedures have led to the computation of accurate excitation energies and geometries both in the ground and in the excited states [17, 18, 21, 22].

Encouraged by these positive results, we want to explore here for the first time the use of quantum Monte Carlo forces to perform molecular dynamics simulations in the excited state. To this aim, we must first address various issues which arise from the use of quantum Monte Carlo forces in molecular dynamics that are absent in deterministic electronic structure approaches. In this Chapter, we focus on the presence of noise in the quantum Monte Carlo forces, and propose different strategies to ameliorate its main impact, namely, the lack of conservation of energy during the molecular dynamics simulation.

5.2   Noisy forces in quantum Monte Carlo

Interatomic forces in QMC are estimated using Monte Carlo integration and are therefore characterized by a stochastic noise as any other observable in QMC. In particular, in variational Monte Carlo, interatomic forces are computed as the average of a local quantity over a set of N_{MC} electronic configurations, $\{\mathbf{R}\}$, sampled from a probability distribution given by the square of a given trial wave function, Ψ_T,

$$
\begin{aligned}
F^{\mathrm{QMC}} &= \langle f(\mathbf{R}) \rangle_{|\Psi_T|^2} \\
&\approx \frac{1}{N_{\mathrm{MC}}} \sum_{i=1}^{N_{MC}} f(\mathbf{R}_i) \,,
\end{aligned}
\tag{5.1}
$$

as further elaborated in the following Chapter. From the central limit theorem, for sufficiently large N_{MC}, each component of the QMC force will be characterized by a Gaussian distributed random noise and can therefore be rewritten as

$$
F^{\mathrm{QMC}} = F^{\mathrm{C}} + F^R \,,
\tag{5.2}
$$

where F^{C} is the conservative force one would obtain for infinite sampling, and F^R is the random contribution.
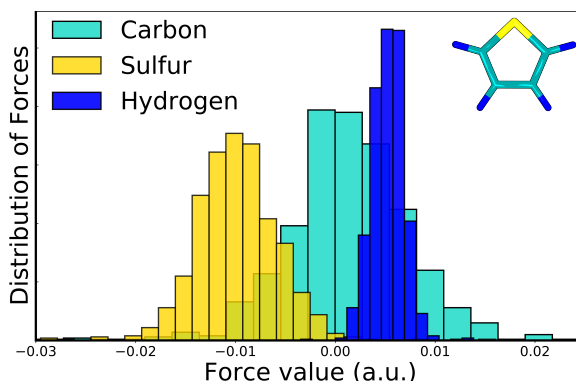
Figure 5.1: Distribution of the interatomic forces of thiophene at the CC3 ground-state geometry, computed in multiple QMC runs. We histogram the in-plane components of the QMC force along the principal axis of the molecule for sulfur, the adjacent carbon, and its hydrogen. The MC runs used to generate the forces were performed with $N_{MC} = 720000$ and yield a statistical error on the energy of about 1 mHa.

To illustrate this, we compute the interatomic forces for thiophene in 1000 variational Monte Carlo runs (characterized by different initial random seeds) at the same ground-state geometry. As shown in Fig. 5.1, the resulting 1000 forces are Gaussian distributed around their averages with a certain spread which we denote here as σ^{QMC}. The spread is proportional to the root mean square fluctuation of the corresponding local quantity (Eq. 5.1) divided by the square root of the number of MC steps, N_{MC} and, therefore, decreases with increasing length of the QMC run as

$$\sigma^{\mathrm{QMC}} \sim \frac{1}{\sqrt{N_{\mathrm{MC}}}} \sqrt{\langle f(\mathbf{R})^2 \rangle_{|\Psi_T|^2} - (\langle f(\mathbf{R}) \rangle_{|\Psi_T|^2})^2} \,. \tag{5.3}$$

We note that the spread is different for the different atoms and directions, and that the forces in Fig. 5.1 are rather small in absolute value since the structure is close to equilibrium.

Unfortunately, the use of noisy QMC forces in a Born-Oppenheimer molecular dynamics (BOMD) simulation affects the outcome of the dynamics in undesirable ways. As shown for thiophene over a time-scale of one picosecond in Fig. 5.2, an MD simulation with QMC forces in the micro-canonical (NVE) ensemble displays the following features:

- lack of conservation of total (kinetic plus potential) energy, which increases with time;

- appearance of nonphysical rotations and translations.

We stress that this MD simulation is started with a structurally distorted thiophene and no velocities. With these initial conditions, since the molecule is isolated, an MD simulation with deterministic conservative forces would lead to conservation of
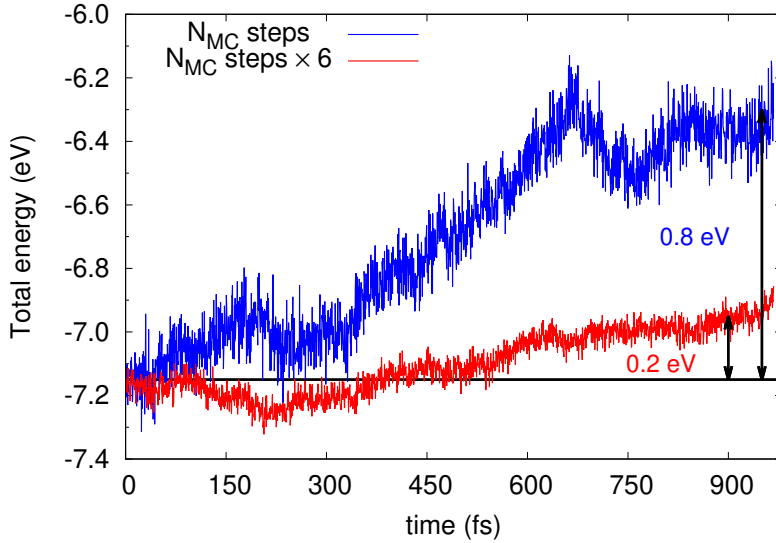
Figure 5.2: Total (kinetic plus potential) energy of thiophene during two MD simulations of 1 ps, where QMC forces are used at each step of the MD. The two simulations differ in the number of MC samples used to compute the forces in the variational Monte Carlo runs, namely, $N_{MC} = 24 \times 10^4$ and $6 \times N_{MC}$. The corresponding errors in the electronic energy are about 2.2 mHa (0.06 eV) and 0.8 mHa (0.02 eV), respectively.

energy without translations and rotations of the system. Since we are ultimately interested to study excited-state relaxation processes, we would generally aim at being able to follow a similarly stable evolution over a time-scale of at least a few picoseconds.

The behavior we observe in our MD with QMC forces stems from the fact that the noise spreads from the forces to the velocities during the simulation. The forces are indeed used to propagate the velocities and, since the QMC forces have a random contribution, this leads to the velocities having an additional component which "performs" a random walk and whose module increases as the square root of time. Since the kinetic energy depends on the square of the velocities, it increases linearly in time and the system heats up. Because of equipartition, the heat will be redistributed among internal as well as center of mass and rotational degrees of freedom, also resulting in translations and rotations.

As mentioned above, the magnitude of the noise in the QMC forces can be controlled by tuning the length of the QMC run used to estimate the forces at each MD iteration. If one increases the number of MC steps, the noise will be reduced and the corresponding increase in the total energy during the MD simulation will also diminish. This is shown in Fig. 5.2 where the growth of the total energy over 1 ps is smaller by about a factor of 4 when the length of the QMC run (i.e. N_{MC}) to compute the forces is appropriately increased. We note that, even though QMC runs starting from different random seeds result in a different sequence of random errors and yield a different energy increase during the MD simulation, the overall behavior remains in

average the same. Consequently, one can achieve conservation of energy as expected in an NVE simulation by performing longer QMC runs at each MD step. This path to a stable MD simulation is of course as much straightforward as it is expensive and, for this reason, we attempt here to reach the same goal through alternative strategies.

In the rest of the Chapter, to quickly test potential solutions to our problem, we use classical force fields in the MD simulations and add Gaussian-distributed errors to the components of the deterministic forces to mimic the effect of QMC fluctuations. The variances of the Gaussian distributions to sample the noise are obtained by fitting the histograms of QMC forces like those in Fig. 5.1.

5.3 Computational details

We perform the QMC (VMC) calculations with the program CHAMP [23]. We employ scalar relativistic energy-consistent Hartree-Fock (HF) pseudopotentials and the corresponding aug-cc-pVDZ basis set [24, 25]. We use here a simple Jastrow-Slater wave function with a single determinant and a 2-body Jastrow factor including only electron-electron and electron-nucleus terms. The starting orbitals are obtained in a HF calculation carried out with the program package GAMESS(US) [26]. For simplicity, in the QMC results shown in this Chapter, we keep the parameters of the wave function fixed and, at each MD step, we simply recenter the wave function at the new atomic positions. The nuclei are moved within CHAMP with a standard velocity-Verlet [27, 28] algorithm and a time-step Δt of 0.5 fs.

The classical force-field simulations are carried out with the Tinker code [29] and the OPLS-AA force field with the standard parameters supplied within Tinker. A time-step of 0.5 fs is used also in Tinker. The root-mean-square widths of the Gaussian distributions to sample the noise added to the classical forces of thiophene are listed in Table 5.1(A). The values are obtained from a Gaussian fit of the histogram of the forces similar to the one in Fig. 5.1. We prefer to perform a fit instead of actually computing the root mean square fluctuation since we have some outlier forces that need to be removed, as we explain in Chapter 6. The Table contains also other parameters which are relevant in later Sections of this Chapter. The library SLATEC [30] is linked to both CHAMP and Tinker to perform polynomial fits.

All QMC and force-field simulations for thiophene have the same starting configuration, namely, a distorted geometry and zero velocities, but differ either in the algorithm to perform the time evolution or in the sequence of random errors added to the deterministic forces during the MD. If not otherwise specified, we use a starting geometry whose potential energy is about 1 eV higher than the one corresponding to the ground-state optimal geometry. This choice allows us to sample configurations somewhat distant from the minimum as well as explore changes in the potential energy which are of the same order of magnitude as the excitation energy of thiophene, as we would like to encounter in future excited-state applications.

(A)

σ_C^{QMC}	σ_S^{QMC}	σ_H^{QMC}
0.0027	0.0020	0.0009

(B)

	FIT 1	FIT 2	FIT 3	FIT 4
$\tau_{C,S}$	1.5	1.2	1.3	1.2
τ_H	1.0	0.9	1.0	0.8

(C)

T (K)	0.01	0.05	0.1	1.0	10.0
γ_C	2.0960	0.4192	0.2096	0.0210	0.0021
γ_S	0.4340	0.0862	0.0434	0.0043	0.0004
γ_H	2.7751	0.5550	0.2775	0.0277	0.0028

(D)

		C	S	H
FIT 1	τ^E	1.22	0.89	1.30
	σ^E	0.00045	0.00073	0.00007
FIT 2	τ^E	1.22	0.93	1.17
	σ^E	0.00033	0.00059	0.00006

(E)

FIT 1	T (K)	0.003	0.010	0.100	0.500
	ω	0.655	0.910	0.991	0.995
FIT 2	T (K)	0.001	0.003	0.010	0.100
	ω	0.205	0.825	0.951	0.995

Table 5.1: (A) QMC errors (a.u.) on the forces in thiophene, averaged over all atoms of the same type and over all directions for a QMC run of N_{MC} =960000. (B) Values of dimensionless (in units of Δt) parameters, τ_{fit}, used for the exponential fit of the forces. (C) Correspondence between the input temperature (K) and the dimensionless parameter γ. (D) Values of dimensionless (in units of Δt) parameters, τ^E, and σ^E (a.u.) used in the modified Langevin approach with memory for two different sets of fit parameters as labeled in part (B) of this Table. (E) Correspondence between the input temperature (K) and the dimensionless parameter ω for two sets of fit parameters as labeled in part (B) of this Table.

5.4 Results

In this Section, we present various strategies we have explored to reduce the impact that noisy forces have on an MD simulation. We focus here on the thiophene molecule as prototypical example and, for simplicity, restrict ourselves to its ground-state description since our findings concerning noise will also be valid in the case of excited states. We use the conservation of the total energy as the primary indicator of the success of a given scheme since this is an easy way to judge our progress.

The Section is divided into three parts, which can be considered as additive or alternative steps one can employ to improve on a standard MD simulation in which the forces have a statistical noise.

5.4.1 Removing rotations and translations

As discussed above, the use of noisy forces results in a rise in temperature during an MD simulation and a redistribution of the resulting energy among all degrees of freedom, including the center of mass and the rotational degrees. To contrast the consequent appearance of nonphysical translations and rotations, for an isolated molecule, one can simply set the total angular momentum and the translation of the center of mass to zero at each time step of the MD simulation. To this aim, we rely here on the approach implemented in Tinker, which we summarize in Appendix 5.A.

Always starting from zero velocities and the same distorted geometry (see Computational Details Section), we perform MD simulations with noisy forces for thiophene both with all degrees of freedom left unconstrained as well as with rotations and translations removed from the system. In order to have statistically meaningful results, we compute 30 trajectories for both cases, which we obtain by starting with different random seeds for the Gaussian noise added to the classical forces. As shown in Fig. 5.3, constraining the translational and rotational motion of the molecule results in an improvement in the conservation of the total energy with a reduction in its growth of about 0.1 eV over the 2 picoseconds of MD.

This is however clearly not sufficient to solve the issue. Moreover, while removing the unwanted rotations and translations is possible for an isolated molecule, this solution cannot be applied when studying a molecule in a bath, for example, in a hybrid quantum-in-classical calculation where we use QMC for the quantum part.

We note that, in such an embedded calculation, one might hope that part of the statistical noise is dissipated from the molecule to the rest of the system, contributing to stabilizing the energy of the molecule treated with QMC. However, from preliminary QMC/molecular mechanics MD simulations of thiophene solvated in benzene, it appears that, over the time scales of few picoseconds, the combined system does not thermalize and the quantum part significantly heats up.

5.4.2 Fitting the forces to reduce the noise

A possible solution to the problem of statistical noise is to follow one of the most standard approaches to handle a set of data with random noise, namely, to perform a
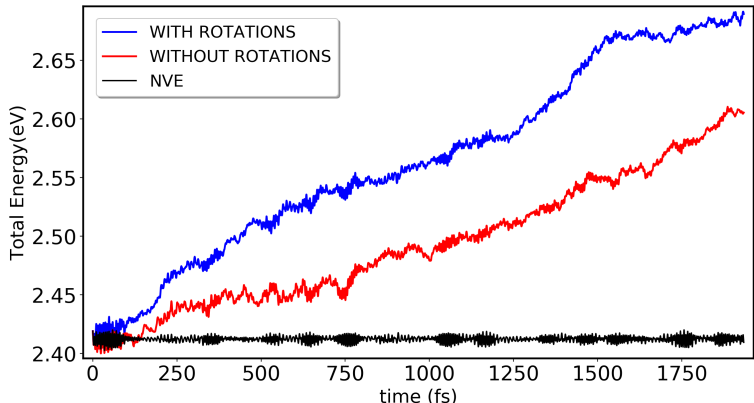
Figure 5.3: Total energy (eV) of thiophene averaged over 30 MD simulations using noisy forces (blue), and over 30 simulations where rotations and translations of the center of mass are removed at every time step (red). In both cases, the equations of motion are solved with the velocity-Verlet scheme.

fit. Fitting or filtering the noise is routinely applied to experimental data and is also at the core of many audio or video devices to remove white noise. When dealing with a potential energy surface, one is commonly interpolating between points of the configuration space (e.g. with machine learning force fields [14, 31]) rather than screening random noise. On the other hand, when using QMC, the problem is more similar to the latter with the interatomic forces being the noisy data.

Here, we would like to use the fit as a means to separate the conservative force, F^C, from its correspondent random component, F^R, (see Eq. 5.1) and do so on-the-fly during an MD simulation. If the simulation is on-the-fly, the fit can only be performed on past configurations to improve the present force prediction. Since the forces are the gradients of the potential energy with respect to the atomic coordinates, the most natural choice would be to fit the forces against the nuclear positions. However, in doing so, we encountered severe difficulties, especially close to the turning points where almost all of the points used to perform the fit have the same spacial configuration and, within the noise, the same values for the forces.

Therefore, to overcome this problem, we perform the fit of the noisy forces against time. For simplicity, we treat each force component (in each direction and for each atom) independently and use a third-order polynomial fit to be able to describe inflection points. We then use this fitted estimate of the force, F^E, instead of the noisy ones to drive the dynamics according to the velocity-Verlet integration

scheme,

Estimated force scheme:

(5.4)

$$v^E_{n+1/2} = v^E_n + \frac{1}{2m}F^E_n\Delta t$$
$$x^E_{n+1} = x^E_n + v^E_{n+1/2}\Delta t$$
$$\text{compute } F^{\text{QMC}}_{n+1} \rightarrow F^E_{n+1}$$
$$v^E_{n+1} = v^E_{n+1/2} + \frac{1}{2m}F^E_{n+1}\Delta t$$

where we add the superscript E to all the quantities which are dependent on the estimated forces (in this case, all but the QMC forces). Note that, when using a force-field force with added noise, we still denote the corresponding force with the label "QMC".

For all results presented in this Section, we constrain the motion of the molecule to enforce no rotations and no translations of the center of mass as previously explained.

Fitting procedure

We fit the noisy forces in time over a fixed number of points, N_{fit} and, for the first $(N_{\text{fit}} - 1)$ steps of the MD, collect the forces needed to perform the fit, using $F^E = F^{\text{QMC}}$ to evolve the system.

At the N_{fit}-th and subsequent steps, we use the noisy forces of the present and past points to get a better estimate, F^E, of the current force as illustrated in Fig. 5.4. We find it fundamental for the success of this procedure to include the force of the present time step, t_p, especially when the force changes sign and a prediction based only on the past values tends to dramatically overshoot or undershoot the real value.
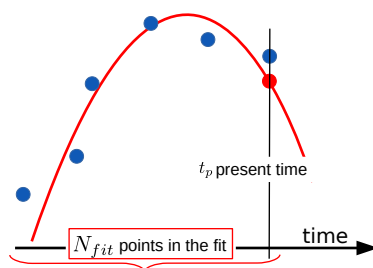
Figure 5.4: Illustrative scheme of the fitting procedure. We fit the force in time and, at a certain time t_p, we use the past $(N_{\text{fit}} - 1)$ noisy forces and the present one to determine a new estimate of the current force.

To perform the fit, we have to choose a value of N_{fit} and decide how to weigh the noisy forces in the fit. Intuitively, we would like that the forces closer to the present time have a higher impact on the final result. Therefore, in addition to the use of

equal weights for all points, we suggest the use of weighting functions decaying in time and investigate the impact on the conservation of the total energy.

In particular, we explore the following functional forms for the weights, $w(t)$:

a) $w(t) = 1$;

b) $w(t) = \dfrac{1}{(t_p - t)/\Delta t + 1}$;

c) $w(t) = \exp[-(t_p - t)/\Delta t]$;

where $t \leq t_p$. After experimenting with these weights and various values of N_{fit} we have found that the exponential weights perform the best. They have the additional benefit that N_{fit} can be chosen fixed while the range of actively used data points is tuned by a parameter τ_{fit} as in

$$w(t) = \exp[-(t_p - t)/(\tau_{\text{fit}} \Delta t)]. \tag{5.5}$$

This choice also enables us to differentiate between atom types via the values of τ_{fit}. Indeed, in the case of thiophene, the forces on the hydrogen oscillate faster in time than on the carbon and sulfur atoms, so hydrogen will require smaller decaying times than the other atoms: if a full oscillation of the hydrogen forces is completed in fewer time-steps, fewer points can be adequately fitted for hydrogen with a third-order polynomial relative to the other atoms. Consequently, including too many forces in the past will result in an inadequate fit, while, by including too few, we will not have enough information to perform an adequate fit. While it would be optimal to introduce three different values of τ_{fit} for the the different atom types, a good compromise is to use a common one for sulfur and carbon ($\tau_{\text{S,C}}$) and a different one for hydrogen (τ_{H}) in order to reduce the number of parameters to tune.

In Fig 5.5, we show the results obtained by evolving the system with forces fitted with exponential weights and four different sets of decaying times, τ_{fit} (FIT1–4), which are listed in Table 5.1(B). A value of $N_{\text{fit}} = 10$ is sufficient for all sets of fit parameters. For each case, we compute 100 trajectories and average their energy and the corresponding root-mean-square fluctuations as a function of time. Different choices of decay times in the weights can yield either an increase or a decrease of the total energy with respect to the conserved value. Importantly, by appropriately tuning the parameters τ_{fit}, we can stabilize the total energy over a time scale of the order of picoseconds. Moreover, for the chosen range of weighting parameters (0.8–1 for τ_{H} and 1.2–1.5 for $\tau_{\text{S,C}}$), we always improve over the MD result with noisy forces. Since the error on the average total energy is given by the root-mean-square fluctuations divided by the square root of the number of trajectories (and is about 0.02 eV after 2 ps), the average energy of FIT 3 is statistically compatible with the total NVE energy over the whole simulation.

A few remarks can be made concerning the values of τ_{fit}. First of all, our third order fit function can only be used over a time range that spans part of the period with which the atom is moving. Since the hydrogen atom moves much faster than the carbon and sulfur atoms, it is clear that its τ_{fit} value must be smaller than those
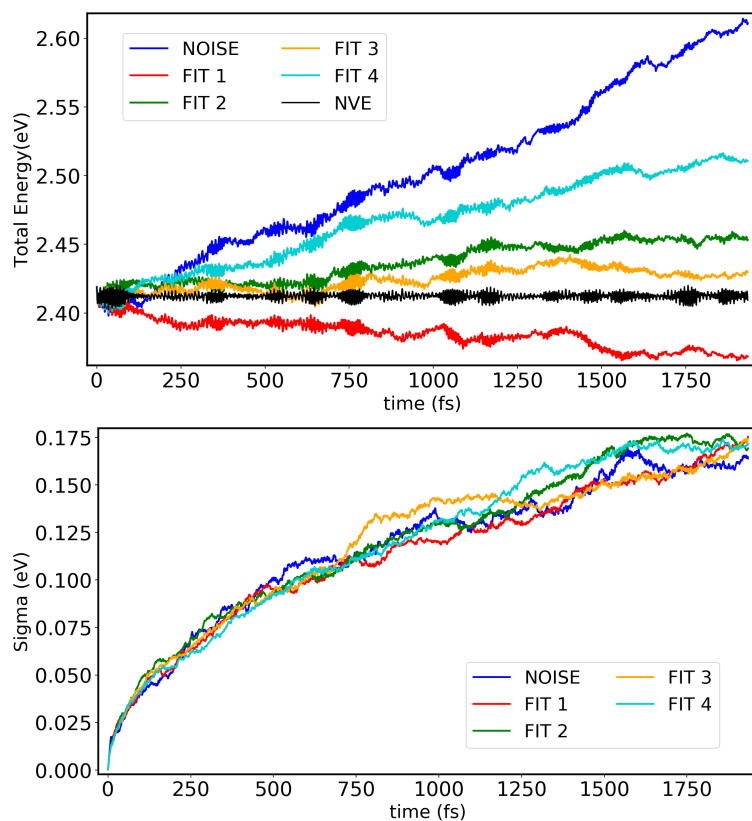
Figure 5.5: Total energies (eV) of thiophene (top) and relative root-mean-square fluctuations (bottom) averaged over 100 MD runs. We use the exponential fit (Eq. 5.5) with different sets of parameters, τ_{fit}, as detailed in Table 5.1(B).

for the other atoms. Second, similarly, in case a run is performed with a much lower initial energy, then, either, preferably, the time step must be adjusted or the τ_{fit} values must be adjusted such as to cover the same fraction of the period of the atom's motion. Thirdly, we find it possible to tune the parameters such that the average energy remains constant over a time span as long as three picoseconds (not shown). Since this is not possible in later applications, we do not pursue this further here. Finally, it is surprising that, with FIT 1, the total energy goes down. One explanation for this behavior can be that the fit systematically takes "shortcuts" at extrema of the potential energy. Of course, part of the randomness characteristic for the QMC forces remains in the fitted forces, which means that finally all curves must linearly increase like the one with the noisy "QMC" forces. We have checked that this is indeed the case, also when the forces are fitted according to FIT 1.

We report the spread of the trajectories also to remind us that these are average behaviors, which might not hold for a single trajectory. This is important because, when performing QMC calculations, one would like to extract the needed information with as few trajectories as possible. Surprisingly, performing a fit does not decrease the spread among the trajectories with respect to the use of noisy forces. Furthermore, the spread grows in time, becoming as large as 0.2 eV after 2 ps, so the longer the MD, the less representative a single trajectory becomes. Of course the spread is around a much better average, and therefore still less damaging than without the fit.

Finally, we note that, while we have removed here translations and rotations, the fit procedure improves the results of an MD with noisy forces also when we do not constraint the motion as shown in Fig. 5.10 in Appendix 5.B.

5.4.3 Thermalizing the noise

The fit procedure just presented combined with the removal of rotations and translations allows us to perform a nearly stable MD simulation with the use of noisy forces. The behavior of the average total energy differs however for different sets of τ_{fit} which are relatively close [see Fig 5.5 and Table 5.1(B)]. Furthermore, since tuning these parameters can become costly for more complex systems, we would like to make our approach more stable by decreasing the dependence of the results on the fit parameters. To this aim, we now investigate if it is possible to diminish the influence of the noisy component in F^{QMC} by adding a friction term to prevent the excess kinetic energy from increasing. In other words, we want to stabilize the excess kinetic energy as it is also done with Langevin dynamics.

Langevin dynamics

In many cases in physics, the system under investigation displays dynamics on a wide spectrum of time scales. In case one is only interested in the slow modes of the system, for example the slow motion of colloids in a system containing fluid molecules and colloidal particles, it is usually profitable to eliminate the fast modes altogether. This unavoidably leads to so-called Langevin dynamics for the slow modes. In this case, the acceleration of the slow degree of freedom is governed by a conservative

force to which are added a friction force, proportional to minus the velocity of the degree of freedom addressed, and a random force. In case the additional forces are related through the so called fluctuation-dissipation theorem, the system is guaranteed to behave as being thermalized to a preset temperature T. In particular, the average energy of the system is guaranteed to fluctuate around some average value, with the average value and the size of the fluctuations governed by the temperature. The question that poses itself is: Can we make use of Langevin dynamics to prevent the increase of energy that has plagued our runs up to now?

In order to point at some differences between the usual application of Langevin dynamics and the application that we are aiming for, consider the Langevin equation as it appears in most physical applications

$$m\frac{d^2x}{dt^2} = -\xi\frac{dx}{dt} + F^{\mathrm{C}} + F^{R}, \tag{5.6}$$

where the first and last terms in the right hand side represent the friction force and the random force respectively. The latter of these is a Gaussian random force with first and second moments

$$\langle F^{R} \rangle = 0 \tag{5.7}$$
$$\langle F^{R}(t)F^{R}(t') \rangle = 2k_{\mathrm{B}}T\xi\delta(t - t'). \tag{5.8}$$

The first of these is obvious, the second is the aforementioned fluctuation-dissipation theorem. Although we will meet more involved integration schemes later, let us here stick to the simplest one, where we approximate the Dirac delta as $\delta(t) = 1/\Delta t$. This approximation cannot be seen independent from a discretization of the time axis as in computer simulations, where $\int F(t)dt$ is replaced by $\sum_n F(n)\Delta t$. With this, we may calculate the random force as

$$F^{R}(n) = \sqrt{\frac{2k_B T\xi}{\Delta t}}G(n) =: \sigma G(n), \tag{5.9}$$

where G is a Gaussian random number with average zero and unit variance.[1] We conclude that the strength of the random force, denoted σ, is related to the friction coefficient and the temperature through $\sqrt{2k_B T\xi/\Delta t}$.

Now consider the differences between applying this set of equations in the usual physical application and the intended application with QMC forces. In the usual application, F^C is a known force derived from a known force-field model. Moreover, the friction coefficient ξ and the temperature T are physically determined quantities, from which the strength of the random force follows through the use of Eq. 5.9. So, in case the update of the velocities is done with the simplest possible integrator, the random forces effect a change of the velocities equal to $F^R\Delta t$, i.e. a change proportional to $\sqrt{\Delta t}$. Now, in the case of QMC forces, we do not know F^C nor do we know F^R, we only know the sum of the two, which is F^{QMC}. We do know, however, the strength of the random force since it can be calculated from the QMC run

---

[1]The fluctuation-dissipation theorem after discretization reads $\langle F^{R}(n)F^{R}(n') \rangle = 2k_B T\xi\langle G(n)G(n')\rangle/\Delta t = 2k_B T\xi\delta_{n,n'}/\Delta t.$

as discussed in Section 5.1. This means that of the two quantities ξ and T only one can be chosen independently, while the other follows from the fluctuation-dissipation theorem. An important consequence of all of this is that with the naive integration scheme, mentioned a few lines ago, the random forces lead to changes of the velocities equal to $F^R \Delta t$, which is now proportional to Δt. So the effect of the random force is similar to that of F^C, and the two forces can be combined again to become the measured F^{QMC}. As presented here, this seems to hold only for the naive integrator but we will see below that it also holds in a particular limit for the more involved, correct integrator. Finally, in the usual application to physical systems, one is interested in thermally averaged dynamics, i.e. in thermally averaged time correlation functions. This is what is guaranteed by the Langevin equation as written in Eq. 5.6. In our application to cure the QMC/MD runs from increasing energies, we are interested in following the real path as it follows from the exact forces and the initial conditions as closely as possible. For details, see below.

Let us mention at this point that Langevin dynamics in combination with QMC forces has been used before by Sorella and coworkers [32–34] with a rather different aim. These authors suggest a method to optimize the sampling efficiency of the algorithm when calculating thermal averages. To this end they add noise and corresponding random forces to the Langevin equation, beyond the noise already present in the QMC forces. This approach has been successfully applied for instance to study the phase diagram of liquid hydrogen [33,35] or to determine the molecular vibration properties of small molecules such as H_2S or SO_2 [32]. To compute such thermodynamic quantities, it is necessary to sample thermally relevant configurations on the potential energy surface (PES) but not to follow the evolution of the system on such a PES for any given amount of time. For example, in order to compute the molecular vibrations, the authors obtain configurations close to the minimum from a Langevin dynamics and estimate the average elements of the Hessian matrix.

Thermalizing the noise and not the signal

As stated above, in this work we are interested in time-related quantities and want to follow a realistic dynamics of the system. Since the Langevin equation as written in Eq. 5.6 quickly affects the trajectory as compared to a Newtonian evolution, we suggest to change it into

$$m \frac{d^2 x}{dt^2} = -\xi \left(\frac{dx}{dt} - v^E \right) + F^{\mathrm{QMC}}, \tag{5.10}$$

where v^E is the best estimate that we have for the actual path. The intuitive picture that we have is that the friction term combines with the random components in the QMC forces, such that together they stabilize the excess energy to be thermalized to some preset temperature. Surely, even then energy will gradually leak into the system but hopefully not very fast. We now must split the QMC force in a way that allows the calculation of v^E. Therefore we write $F^{\mathrm{QMC}} = F^E + F^Y$, and so

$$m \frac{d^2 x}{dt^2} = -\xi \left(\frac{dx}{dt} - v^E \right) + F^E + F^Y, \tag{5.11}$$

where F^E is the force that gives rise to the velocity v^E. This equation guarantees that if F^{QMC} is calculated increasingly more accurately and approaches F^E, then both F^Y and the friction force become zero, and the correct dynamics is obtained. Note that it is assumed that when F^{QMC} approaches F^C, also F^E approaches F^C, so the fit should do no harm in case it is applied in a purely classical MD run.

We now choose to let F^E be the force that we calculated in Section 5.4.2. When solving the equation, we treat F^E as a deterministic force, with a corresponding deterministic v^E. We treat F^Y as the random component of F^{QMC}, obeying the fluctuation-dissipation theorem

$$\langle F^Y(t)F^Y(t')\rangle = 2k_B T\xi\delta(t-t')\,. \tag{5.12}$$

It is not clear from the beginning if the procedure will improve on the case with fitted forces.

To solve the adjusted Langevin equation, we write the position and velocity as sums of estimated terms plus noisy contributions,

$$x = x^E + y \tag{5.13}$$

$$\frac{dx}{dt} = v^E + w\,, \tag{5.14}$$

where x^E and v^E are the position and velocity that result from the estimated force F^E. Consequently $w = \frac{dy}{dt}$, and the equations to solve become

$$m\frac{dv^E}{dt} = F^E \tag{5.15a}$$

$$m\frac{dw}{dt} = -\xi w + F^Y\,. \tag{5.15b}$$

At a first glance, these two equations appear to be independent of each other but they are linked via the computation of the force: the estimated force F^E is obtained from the fit of the QMC forces which are computed on the positions, $x = x^E + y$.

It is easy to show that the excess velocity, w, satisfies the following integral equation which we must evaluate,

$$w(t) = w(0)e^{-\xi t/m} + \frac{1}{m}\int_0^t d\tau e^{-\xi(t-\tau)/m}F^Y(\tau)\,. \tag{5.16}$$

Below we will suggest several ways to perform the integral.

Integrating the modified Langevin equations

As a first attempt to solve this equation, we follow Ref. [32] and consider F^Y to be constant during each time interval being considered. This allows us to easily perform the integral, obtaining

$$w(t+\Delta t) = w(t) + \frac{1}{\xi}\left(1 - e^{\xi\Delta t/m}\right)F^Y(t) \tag{5.17}$$

Again following Ref. [32], we calculate the change of y during the given time interval simply as $w\Delta t$. We update the estimated positions and velocities, x^E and v^E respectively, using the velocity-Verlet algorithm. In order to treat random and deterministic forces on a similar footing, after all they are both constant during the time step, we apply the velocity-Verlet algorithm to the random contributions as well. The resulting algorithm is

Langevin scheme 1

$$(5.18)$$

$$w_{n+1/2} = e^{-\xi\frac{\Delta t}{2m}}w_n + \frac{1}{\xi}(1 - e^{-\xi\frac{\Delta t}{2m}})F_n^Y$$

$$v_{n+1/2}^E = v_n^E + \frac{1}{2m}F_n^E\Delta t$$

$$x_{n+1} = x_n + v_{n+1/2}^E\Delta t + w_{n+1/2}\Delta t$$

$$\to F_{n+1}^{\mathrm{QMC}} \to F_{n+1}^E, F_{n+1}^Y$$

$$w_{n+1} = e^{-\xi\frac{\Delta t}{2m}}w_{n+1/2} + \frac{1}{\xi}(1 - e^{-\xi\frac{\Delta t}{2m}})F_{n+1}^Y$$

$$v_{n+1}^E = v_{n+1/2}^E + \frac{1}{2m}F_{n+1}^E\Delta t$$

We quickly investigate the two limits, i.e. when $\xi \to \infty$ or $\xi \to 0$; recall that σ^{QMC} remains constant while taking the limits. When $\xi \to \infty$, the random contribution to the velocity becomes zero and the scheme reduces to the simple velocity-Verlet integration of the dynamics governed by F^E. On the other hand, in the limit of $\xi \to 0$, we recover the evolution with the QMC force since

$$\lim_{\xi\to 0} e^{-\xi\frac{\Delta t}{2m}}w_n = w_n$$

$$\lim_{\xi\to 0} \frac{1}{\xi}(1 - e^{-\xi\frac{\Delta t}{2m}})F_{n+1}^Y = \frac{\Delta t}{2m}F_{n+1}^Y, \quad (5.19)$$

and, by substituting the equations for $v_{n+1/2}^E$ and w_{n+1} in x_{n+1}, we obtain

$$
\begin{aligned}
x_{n+1} &= x_n + v_{n+1/2}^E\Delta t + w_{n+1/2}\Delta t \\
&= x_n + (v_n^E + w_n)\Delta t + \frac{1}{2m}(F_n^E + F_n^Y)\Delta t^2 \\
&= x_n + v_n\Delta t + \frac{1}{2m}F_n^{\mathrm{QMC}}\Delta t^2 . \quad (5.20)
\end{aligned}
$$

Had we not applied the velocity-Verlet split to the random contributions as well, we would have found in the limit of $\xi \to 0$

$$
\begin{aligned}
v_{n+1}^E &= \frac{(F_n^E + F_{n+1}^E)}{2}\frac{\Delta t}{m} \\
w_{n+1} &= F_n^Y\frac{\Delta t}{m} . \quad (5.21)
\end{aligned}
$$

Consequently, the random force would enter differently from the deterministic force and the total would not exactly be equal to an update with F^{QMC}.

Results

We now investigate the performance of our modified Langevin dynamics with damped excess velocity and the integration Langevin scheme 1. In particular, to explore whether the approach ameliorates the conservation of total energy in thiophene, we perform runs with different values of the the damping parameter ξ (for fixed Δt and σ^{QMC}).

We summarize the results in Fig. 5.6, where we label the runs using the dimensionless parameter γ:

$$\gamma = \frac{\xi \Delta t}{m}\,, \tag{5.22}$$

since γ enters the exponentials in the formulas above and gives a direct indication of the size of the contribution coming from the noisy velocities. In Table 5.1(C), we report the values of γ for the different atoms and the corresponding temperature, T, since T is the actual parameter that we vary in the simulations.
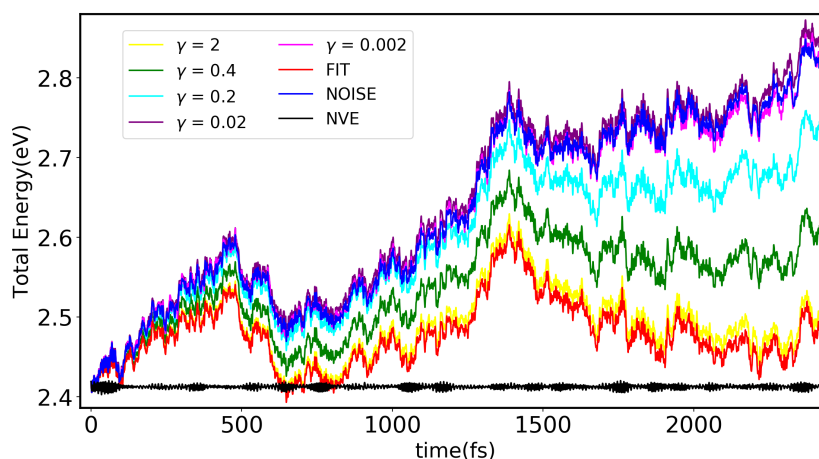
Figure 5.6: Total energies (eV) for different values of the γ parameters ($\gamma = \gamma_{\mathrm{C}}$ is given), obtained with integration Langevin scheme 1, compared with the curves obtained with the noisy (NOISE) and fitted (FIT) forces. All runs are performed with the same starting seed and the same weights in the fit, labeled FIT 1 in Table 5.1(B).

To easily compare the behavior as a function of γ, we use the same sequence of random numbers to introduce the Gaussian noise in the forces. When nothing is said to the contrary, we use the shorthand $\gamma = \gamma_C$. Clearly, already for $\gamma = 2$ ($T = 0.01$ K), we are approaching the limit of $\gamma \to \infty$ and closely following the curve obtained with fitted forces [FIT 1 in Table 5.1(B)]. On the other hand, for values of γ smaller than 0.02 ($T \geq 1$ K), we recover the curve obtained with noisy forces. For the values in between, the total energy lies between the two cases so that, by varying γ, we gradually move between the two limits.

We however do not find a range of γ for which the total energy is not between the two limits, namely, neither being conserved better than the fitted case or worse

than the noisy one. We conclude that, as an alternative to the fitted case, the modified Langevin scheme presented so far is not a viable candidate. We therefore suggest to abandon some of the assumptions we have adopted to solve our equations.

5.4.4 Treating the random forces as properly random

In an attempt to improve our integration scheme, we notice that treating F^Y as a constant during the integration interval in Eq. 5.16 may not be correct. In a more refined integration scheme, F^Y would occur many times, each time with its own random number. We therefore should treat it as a real random function, also when integrating the equation of motion over an interval of one time step Δt.

In this Section, we will make use of the fluctuation-dissipation theorem (Eq. 5.12) at several occasions. As a result, both ξ and T will enter the equations. As we mentioned before, since the strength σ^{QMC} of the random force F^Y is fixed, only one of ξ and T can be chosen freely. Approximating $\delta(t) = 1/\Delta t$ in the fluctuation-dissipation theorem, we obtain the relation between the three of them

$$\xi = \frac{(\sigma^{\mathrm{QMC}})^2 \Delta t}{2k_B T}. \tag{5.23}$$

It may seem inconsistent to approximate the Dirac delta as $1/\Delta t$, while still integrating stochastic forces over the entire time interval. We will show however in Appendix 5.C that the results that we obtain are correct.

With this in mind, we notice that the integral in Eq. 5.16 represents a sum of many random numbers, and therefore must be a Gaussian distributed random number, which we write as $\sigma_w G^w(t)$ with $G^w(t)$ being drawn from a standard Gaussian with unit variance; the actual variance of the stochastic integral is σ_w^2. The latter we compute according to

$$\begin{aligned}
\sigma_w^2 &= \int_0^{\Delta t} d\tau \int_0^{\Delta t} d\tau' e^{-\xi(\Delta t - \tau)/m} e^{-\xi(\Delta t - \tau')/m} \langle F^Y(\tau) F^Y(\tau') \rangle \\
&= 2k_B T \xi \int_0^{\Delta t} d\tau e^{-2\xi(\Delta t - \tau)/m} \\
&= \frac{k_B T}{m}(1 - e^{-2\xi \Delta t/m}),
\end{aligned}$$

where we have made use of the fluctuation-dissipation theorem Eq. 5.12. Inserting this into Eq. 5.16, we get

$$w(t + \Delta t) = w(t)e^{-\xi \Delta t/m} + \sqrt{\frac{k_B T}{m}} \sqrt{1 - e^{-2\xi \Delta t/m}} G^w. \tag{5.24}$$

To obtain the position y, we integrate $\frac{dy}{dt} = w(t)$ with $w(t)$ given in Eq. 5.16, we swap the integrations in the resulting double integral according to $\int_0^t d\tau' \int_0^{\tau'} d\tau = \int_0^t d\tau \int_\tau^t d\tau'$ and perform the integral over τ', obtaining

$$y(t) = y(0) + \frac{m}{\xi} w(0)(1 - e^{-\xi t/m}) + \frac{1}{m} \int_0^t d\tau (1 - e^{-\xi(t' - \tau)/m}) F^Y(\tau).$$

We next treat the stochastic integral as we did for the velocities and write the result as $\sigma_y G^y$, obtaining

$$
\begin{aligned}
y(t + \Delta t) \;\; &= \;\; y(t) + \frac{m}{\xi} w(t)(1 - e^{-\xi \Delta t/m}) \\
&+ \;\; \sqrt{\frac{2k_B T}{\xi}} \sqrt{\Delta t - \frac{2m}{\xi}(1 - e^{-\frac{\xi \Delta t}{m}}) + \frac{m}{2\xi}(1 - e^{-\frac{2\xi \Delta t}{m}})} G^y
\end{aligned}
$$

$$(5.25)$$

Here, G^w and G^y are two independent random numbers. With these new equations for the noisy evolution, we revisit our algorithm. For the systematic part, we still choose a velocity-Verlet integration, while the noisy part will be included according to the equations above. Now that we have an analytical expression for the stochastic displacement, making a half-step will not improve the algorithm. Therefore we suggest to apply Langevin scheme 2.

Langevin scheme 2

$$(5.26)$$

$$
\begin{aligned}
v_{n+1/2}^E \;\; &= \;\; v_n^E + \frac{1}{2m} F_n^E \Delta t \\
y_{n+1} \;\; &= \;\; w_n \frac{m}{\xi}(1 - e^{-\frac{\xi \Delta t}{m}}) \\
&+ \;\; \sqrt{\frac{2k_B T}{\xi}} \sqrt{\Delta t - \frac{2m}{\xi}(1 - e^{-\frac{\xi \Delta t}{m}}) + \frac{m}{2\xi}(1 - e^{-\frac{2\xi \Delta t}{m}})} G_n^y \\
x_{n+1} \;\; &= \;\; x_n + v_{n+1/2} \Delta t + y_{n+1} \\
&\quad\quad\quad \rightarrow F_{n+1}^{\mathrm{QMC}} \rightarrow F_{n+1}^E \\
w_{n+1} \;\; &= \;\; e^{-\xi \frac{\Delta t}{m}} w_n + \sqrt{\frac{k_B T}{m}} \sqrt{1 - e^{-\frac{2\xi \Delta t}{m}}} G_n^w \\
v_{n+1}^E \;\; &= \;\; v_{n+1/2}^E + \frac{1}{2m} F_{n+1}^E \Delta t
\end{aligned}
$$

Note that, in Langevin scheme 1, we skipped an independent computation of y_n since that could implicitly be done by adding the velocities v_n^E and w. This can not be done in the present case. It is easy to see that, in the limit of $\xi \rightarrow \infty$, the stochastic displacement y goes to zero, and Langevin scheme 2 becomes equal to the Estimated force scheme. At first sight, it may seem that the stochastic velocity w remains finite and equal to $\sqrt{k_B T/m}$, but since $\xi = (\sigma^{\mathrm{QMC}})^2 \Delta t/(2k_B T)$, we find that T goes to zero when ξ goes to infinity. The other limit, when ξ goes to zero, yields

$$
w_{n+1} \;\; = \;\; w_n + \frac{\sigma^{\mathrm{QMC}}}{m} G_n^w \Delta t \tag{5.27a}
$$

$$
y_{n+1} \;\; = \;\; w_n \Delta t + \sqrt{\frac{4}{3}} \frac{\sigma^{\mathrm{QMC}}}{m} G_n^y \frac{(\Delta t)^2}{2}. \tag{5.27b}
$$

As argued above, choosing independent random contributions to velocity changes, unrelated to the ones already induced by the random component included in F^{QMC} may introduce additional randomness into the procedure. One may argue that in the limit of very small values of γ the algorithm must approach that of a normal run with QMC forces. This can be achieved by applying in Langevin scheme 2 the following substitutions

$$G^w \quad \longmapsto \quad \frac{F^Y(t)}{\sigma^{\mathrm{QMC}}} \tag{5.28a}$$

$$G^y \quad \longmapsto \quad \sqrt{\frac{3}{4}} \frac{F^Y(t)}{\sigma^{\mathrm{QMC}}}, \tag{5.28b}$$

as is is clear from the limits for small γ shown above. Results obtained with Langevin scheme 2 modified like this are shown in Fig. 5.7.
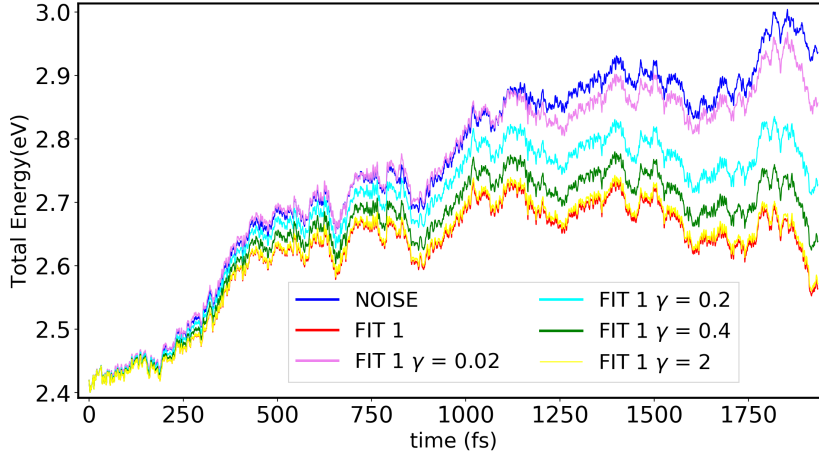
Figure 5.7: Total energies (eV) obtained for different values of $\gamma = \gamma_{\mathrm{C}}$ and Langevin scheme 2 with substitutions according to Eq. 5.28. All runs start with the same random seed. The weight parameters in the fit are labeled FIT 1 in Table 5.1(B).

As is clearly seen, all curves obtained with varying values of γ change monotonously from the fitted curve to the QMC curve. In order to have more flexibility, we suggest to keep Langevin scheme 2 as it is presented, i.e. with independent random numbers G^w and G^y.

Using this new formulation, it is not possible to compare the results of the MD simulation with fitted (FIT) and noisy (NOISE) forces with a fixed sequence of random number, since we are introducing two extra series of random numbers (G^w_n and G^y_n). For this reason, we directly plot the results as averages on 100 trajectories. In Fig. 5.8, we show how the total energy varies in dependence of the value of $\gamma = \gamma_{\mathrm{C}}$; the values of τ_{fit} are taken according to FIT 1 [Table 5.1(B)].

As expected, for small values of γ and correspondingly high temperatures [see Table 5.1(C)], we introduce noise in the system and the total energy grows in time similarly to the behavior obtained with noisy forces. In the limit of large values of γ,

we converge instead to the estimated curve (FIT 1). Differently from the findings so far, the use of Langevin scheme 2 yields a range of values of γ (around 0.2–0.3) for which the energy is conserved over the entire 2 ps of the calculation.

With this range being appropriate in combination with FIT 1, we are interested if it is possible to find a range of γ values that can be applied to all sets of FIT parameters studied in this work. In Fig. 5.9, we present the results for different FIT parameters and values of γ equal to 0.4 and 2 respectively. In all cases, for averages of 100 trajectories, the new algorithm contributes to a stabilization of the total energy. In some cases the improvement is minor but this is to be expected since we are fixing γ instead of using the optimal value for each set of τ_{fit} independently. Importantly, we note that the results for the curves FIT 2–4 clearly indicate that, with the present integration scheme, we are not bounded between the two limits of small and large ξ: we can now find a range of parameters for which we thermalize the noise and obtain nearly constant energy even when both the FIT and NOISE curves result in increasing total energy.

Therefore, our procedure is effective in thermalizing the noise for different sets of τ_{fit} over a range of γ values centered around $\gamma = 1$. This is important since, when finally using forces from an actual QMC simulation to drive the dynamics (as opposed to the forces from an OPLS-AA force field with added noise), the optimal values of τ_{fit} may change. Depending on the size of the system, it could then be too expensive to perform enough simulations to determine the new optimal parameters.

In Fig. 5.11 in Appendix 5.B, we show the results for Langevin scheme 2 when we do not remove rotations and translations. It turns out that, in this case, we are not able to improve on the results with fitted forces when using a value of γ equal to 2. We did not investigate further if an optimal γ exists in this case.
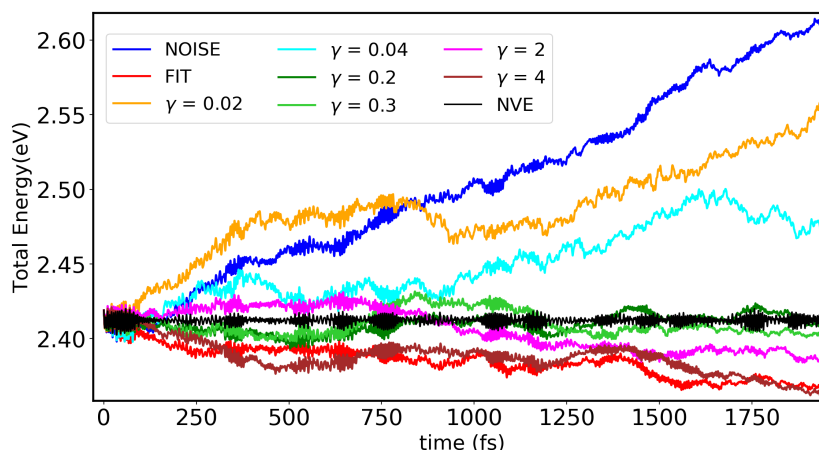

Figure 5.8: Total energies (eV) obtained for different values of $\gamma = \gamma_C$ and the Langevin scheme 2 (Eq. 5.26). The weights parameters in the fit are labeled FIT 1 in Table 5.1(B). The results are averages of 100 trajectories.
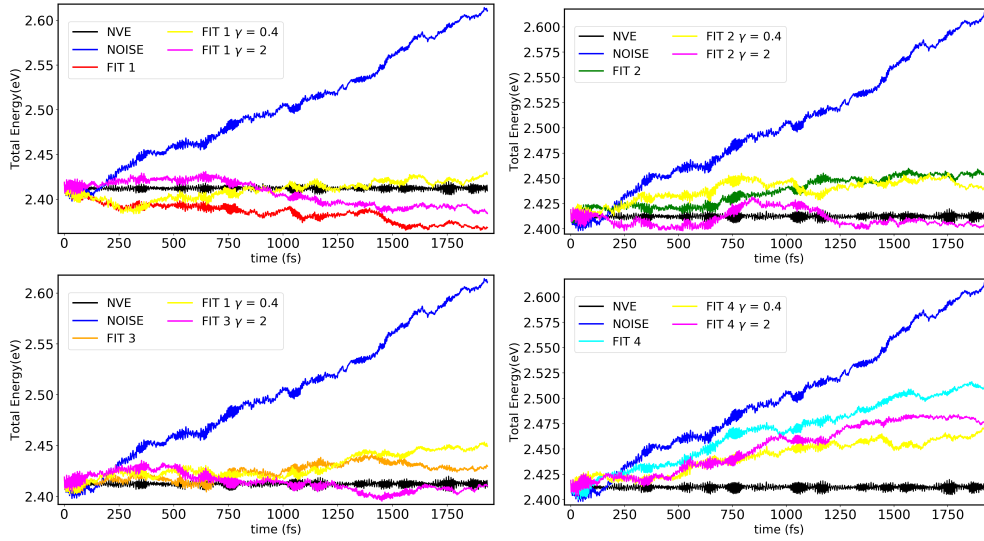
Figure 5.9: Total energies resulting from Langevin scheme 2 (Eq. 5.26) with $\gamma = 0.4$ and $\gamma = 2$ respectively, and for different sets of τ_{fit}. We compare each curve with the noisy one and the respective FIT. The parameters of the fits are given in Table 5.1(B). The results are averages of 100 trajectories.

5.5 Conclusions

In this Chapter, we investigated the impact of driving a molecular dynamics simulation with noisy forces and explored some strategies to minimize undesirable consequences such as the lack of conservation of total energy. Our interest stems from the desire to use QMC forces, which are inherently affected by a stochastic error, to follow the evolution of a system on an excited-state potential energy surface for time scales of at least a few picoseconds. We have shown that we obtain significant amelioration of the problem by removing nonphysical rotations and translations, and by employing an appropriate fit procedure of the forces. We also extensively investigated the possibility to thermalize and, therefore, control the noise with a modified Langevin scheme which we have found to add stability to the molecular dynamics simulations in some cases. That said, further studies are needed since we are not fully understanding yet how these positive occurrences are connected to specific features of the algorithm. In the next Chapter, we will apply the different strategies proposed here to dynamical simulations with actual QMC forces.

5.A   Appendix: Removing rotations and translations

We present here the equations used to impose that the center of mass velocity and the total angular momentum are zero at each time step, as implemented in the code Tinker [29].

We denote the center of mass position, velocity, and angular momentum as

$$\mathbf{R}^{\mathrm{CM}} = \frac{1}{M} \sum_{i=1}^{N} m_i \mathbf{r}_i \tag{5.29}$$

$$\mathbf{V}^{\mathrm{CM}} = \frac{1}{M} \sum_{i=1}^{N} m_i \mathbf{v}_i \tag{5.30}$$

$$\mathbf{L}^{\mathrm{CM}} = \mathbf{R}^{\mathrm{CM}} \times M \mathbf{V}^{\mathrm{CM}} \tag{5.31}$$

where N is the total number of particle and $M = \sum_{i=1}^{N} m_i$ is the total mass of the system.

To impose a zero velocity of the center of mass, we simply shift the velocities of all particles at each time step as

$$\mathbf{v}_i \to \mathbf{v}_i - \mathbf{V}^{\mathrm{CM}} . \tag{5.32}$$

Then, we want to impose that the total angular momentum of the system is zero. Since the center-of-mass velocity is zero, we define the angular momentum with respect to the center of mass as

$$\mathbf{L} = \sum_{i=1}^{N} \mathbf{x}_i \times m_i \mathbf{v}_i , \tag{5.33}$$

where $\mathbf{x}_i = \mathbf{r}_i - \mathbf{R}^{\mathrm{CM}}$.

We now subtract from each velocity $\mathbf{v}_i$ a component that corresponds to a rigid body rotation

$$\mathbf{v}_i \to \mathbf{v}_i - \mathbf{\Omega} \times \mathbf{x}_i \tag{5.34}$$

where the angular velocity $\mathbf{\Omega}$ should be chosen such that the resulting system has zero angular momentum with respect to the center of mass. This can be achieved by writing the angular momentum from Eq. 5.33 as $\mathbf{L} = \mathbf{I}\mathbf{\Omega}$, where $\mathbf{I}$ is the inertia tensor

$$\mathbf{I} = \sum_{i=1}^{N} m_i \left[(\mathbf{x}_i \cdot \mathbf{x}_i)\ \mathbf{1} - \mathbf{x}_i \mathbf{x}_i \right], \tag{5.35}$$

and choosing $\mathbf{\Omega} = \mathbf{I}^{-1} \mathbf{L}$. Here, $\mathbf{1}$ is the three-dimensional unit tensor.

In order to prove that the resulting system has zero angular momentum, we must show that

$$\sum_{i=1}^{N} \mathbf{x}_i \times m_i \left(\mathbf{v}_i - \mathbf{\Omega} \times \mathbf{x}_i \right), \tag{5.36}$$

113

equals zero. Evaluating the cross products, we obtain

$$
\sum_{i=1}^{N} \mathbf{x}_i \times m_i \mathbf{v}_i - \sum_{i=1}^{N} m_i \mathbf{x}_i \times (\mathbf{\Omega} \times \mathbf{x}_i)
$$
$$
= \sum_{i=1}^{N} \mathbf{x}_i \times m_i \mathbf{v}_i - \sum_{i=1}^{N} m_i \left[(\mathbf{x}_i \cdot \mathbf{x}_i)\mathbf{\Omega} - \mathbf{x}_i(\mathbf{x}_i\mathbf{\Omega}) \right]
$$
$$
= \sum_{i=1}^{N} \mathbf{x}_i \times m_i \mathbf{v}_i - \mathbf{I}\mathbf{\Omega} = 0, \tag{5.37}
$$

because $\mathbf{I}\,\mathbf{\Omega} = \mathbf{I}\,\mathbf{I}^{-1} \sum_{i=1}^{N} \mathbf{x}_i \times m_i \mathbf{v}_i = \sum_{i=1}^{N} \mathbf{x}_i \times m_i \mathbf{v}_i$.

5.B   Appendix: Not removing rotations and translations

Here, we quickly summarize the results of our main procedures but, now, when rotations and translations are not removed.

Fit without removing rotations and translations

We compute the total energy and the spread of the trajectories with fitted forces and integration Estimated force scheme (Eq. 5.4) and without imposing zero angular momentum and center-of-mass velocity. For the fit, we use the same parameters as in Fig. 5.5 and Table 5.1(B).

From Fig. 5.10, we conclude that, also in this case, the fit procedure increases the amount of time in which the total energy is constant. However, the spread of the noise in the kinetic energy is faster and, with this set of parameters, we do not manage to keep constant energy for more than 1 ps.
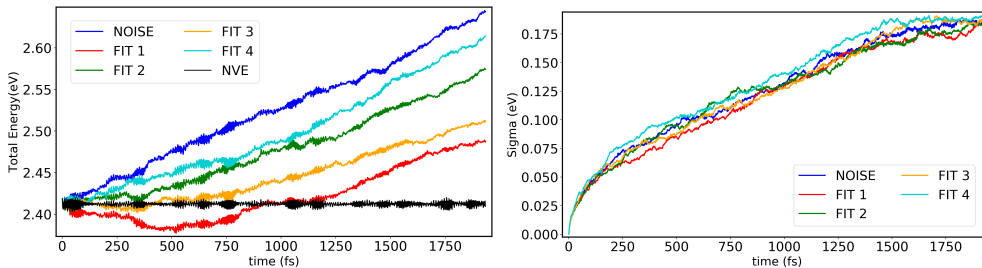
Figure 5.10: Total energies (eV) and relative spread (eV) averaged over 100 trajectories obtained with fitted forces (FIT1–4 in Tab. 5.1) without imposing the removal of translations and rotations.

Langevin scheme 2 without removing rotations and translations

In this Section, we report the results of the Langevin scheme 2 (Eq. 5.26) in the case in which we do not remove rotations and translations at each time step. We use the same value of $\gamma = 2$ and again average on 100 trajectories as in Fig 5.9 but, in this case, we see no clear improvement. It appears that, for this value of γ, we lose the improvement obtained for the case in which we do constrain the motion by removing rotations and translations. It is possible that, in the current situation, the noise is spreading in the system too fast and the algorithm loses its effectiveness. We did not investigate this further for different values of γ.
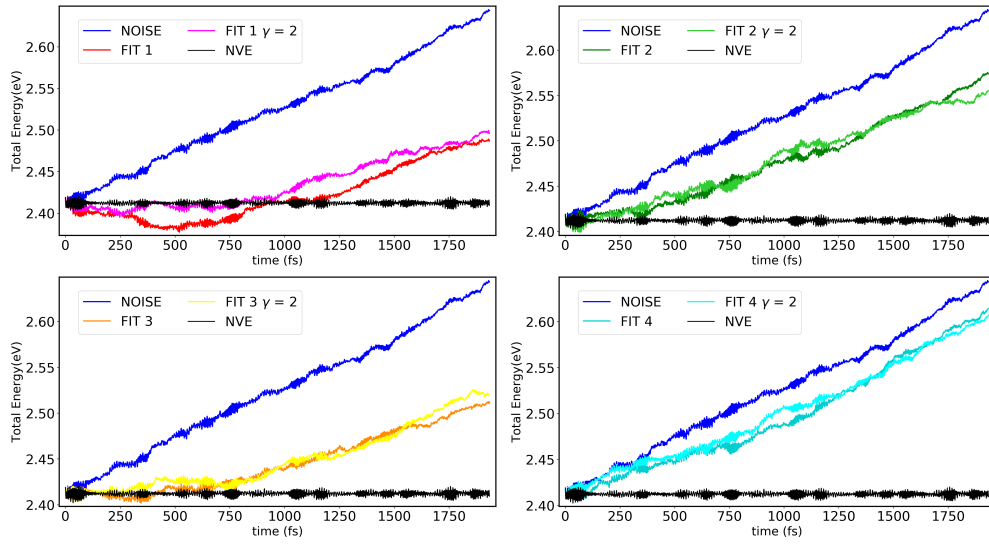
Figure 5.11: Total energies (eV) averaged over 100 trajectories, resulting from algorithm 5.26 with the same values of τ_{fit} as in Fig. 5.9. We again plot the simple fit curve (obtained with Eq. 5.4) and the noisy curve for comparison. We use, as previously, $\gamma = \xi \Delta t / m = 2$.

5.C   Appendix: Langevin with memory

When treating F^Y in Eq. 5.11 as a random contribution to the QMC forces, we assumed that they are delta correlated in time. This neglects the fact that, when calculating F^E, we make use of QMC forces from the recent past and that, therefore, F^Y acquires some memory. We therefore should adjust the equation of motion for the random component of the velocity accordingly

$$m\frac{dw}{dt} = -\frac{1}{k_B T}\int_0^t \langle F^Y(0)F^Y(t-t')\rangle w(t')dt' + F^Y(t), \qquad (5.38)$$

with $\langle F^Y(t)\rangle = 0$. The fluctuation-dissipation theorem, which guarantees that the velocities will be distributed according to the Maxwell-Boltzmann equation, has been explicitly incorporated by writing the friction kernel as $\langle F^Y(0)F^Y(t-t')\rangle/k_B T$.

Integration scheme with memory

We now choose the following analytical expression

$$\langle F^Y(t)F^Y(t')\rangle = (\sigma^E)^2 e^{-2|t-t'|/\tau^E}, \qquad (5.39)$$

to represent the memory of the random forces. There is no simple, unique way to relate σ^E to σ^{QMC} that we used in the main text. The reason is that σ^{QMC} represents the value of the random force time correlation function when it is approximated to be constant over an interval of length Δt and zero elsewhere. A relation between σ^{QMC} and σ^E can be obtained in different ways, for example by putting equal the time integral of the two representations of the random force time correlation function up to time Δt or up to infinity. It is possible however to make comparisons between the two approaches by defining

$$\xi^E = \frac{(\sigma^E)^2 \tau^E}{2k_B T}. \qquad (5.40)$$

The fluctuation-dissipation theorem then becomes

$$\lim_{\tau^E \to 0} \langle F^Y(t)F^Y(t')\rangle = 2k_B T \xi^E \frac{e^{-2|t-t'|/\tau^E}}{\tau^E}, \qquad (5.41)$$

which for small values of τ^E turns into $2k_B T \xi^E \delta(t - t')$.

The problem may now be solved by Laplace transformation, which yields

$$w(t) = w(0)w^H(t) + \int_0^t dt' \, w^H(t - t')\frac{F^Y(t')}{m} \qquad (5.42a)$$

$$w^H(t) = e^{-t/\tau^E}\cosh(\omega t/\tau^E) + \frac{e^{-t/\tau^E}}{\omega}\sinh(\omega t/\tau^E). \qquad (5.42b)$$

The dimensionless parameter ω is given by

$$\omega = \sqrt{1 - \frac{(\sigma^E \tau^E)^2}{m k_B T}}. \qquad (5.43)$$

We calculate the position $y(t)$ by integrating the velocity $w(t)$. After swapping the time integrals according to $\int_0^t d\tau' \int_0^{\tau'} d\tau = \int_0^t d\tau \int_\tau^t d\tau'$ and simplifying the equations, we get

$$y(t) = y(0) + w(0)y^H(t) + \int_0^t dt' \, y^H(t - t')\frac{F^R(t')}{m} \qquad (5.44a)$$

$$y^H(t) = \frac{2\tau^E}{1 - \omega^2} - \frac{e^{-t/\tau^E}}{1 - \omega^2}\tau^E\left[2\cosh(\omega t/\tau^E) + \frac{(1 + \omega^2)}{\omega}\sinh(\omega t/\tau^E)\right] \qquad (5.44b)$$

This concludes the formal solution of the Eq. 5.38 with kernel given in Eq. 5.39.

As before, for numerical applications, the stochastic contribution to the change of the velocity during one time step becomes $\sigma_w G^w$ with the variance given by

$$\sigma_w^2 = \frac{(\sigma^E \tau^E)^2}{(m\omega)^2} \frac{e^{-2\Delta t/\tau^E}}{2(\omega^2 - 1)} \times \tag{5.45}$$
$$\left[-1 + \omega^2 - 2e^{2\Delta t/\tau^E}\omega^2 + (1 + \omega^2)\cosh(2\omega\Delta t/\tau^E) + 2\omega\sinh(2\omega\Delta t/\tau^E) \right].$$

Similarly we write the stochastic contribution as $\sigma_y G^y$ with σ_y for one time step Δt being given by

$$\sigma_y^2 = \Omega \left[4\Delta t(1 - \omega^2) - 2\tau^E(5 + \omega^2) \right] \tag{5.46}$$
$$- \Omega e^{-2\Delta t/\tau^E}\tau^E \left[2\cosh(\omega\Delta t/\tau^E) + \frac{1 + \omega^2}{\omega}\sinh(\omega\Delta t/\tau^E) \right]^2$$
$$+ 2\Omega e^{-\Delta t/\tau^E}\tau^E \left[(7 + \omega^2)\cosh(\omega\Delta t/\tau^E) + \frac{3 + 5\omega^2}{\omega}\sinh(\omega\Delta t/\tau^E) \right],$$

where Ω is given by

$$\Omega = \frac{(\sigma^E)^2}{m^2}\frac{(\tau^E)^3}{(1 - \omega^2)^3}. \tag{5.47}$$

This concludes the solution of the model.

The solution remains valid even when ω becomes imaginary; just replace ω by $i\omega$ and work out the consequences. Since hyperbolic functions become trigonometric functions in this case, it does not seem to be useful for our application. In any case, this only occurs when T is nonphysically small.

The integration scheme is now the same as Langevin scheme 2, with σ_w and σ_y replaced by the new expressions.

Limits

It is interesting to consider the limit when τ^E is much less than Δt. Evidently, in this case the integral in Eq. 5.38 becomes small unless $(\sigma^E)^2\tau^E$ remains finite, i.e. unless ξ^E and T remain finite. Under these conditions, we have

$$\omega = \sqrt{1 - 2\xi^E\tau^E/m} \approx 1 - \xi^E\tau^E/m \tag{5.48a}$$
$$2\omega\Delta t/\tau^E \approx 2\Delta t/\tau^E - 2\xi^E\Delta t/m \tag{5.48b}$$

With this approximation, we find after some lengthy algebra

$$\sigma_w^2 = \frac{k_B T}{m}\left(1 - e^{2\xi^E\Delta t/m} \right) \tag{5.49a}$$
$$\sigma_y^2 = \frac{2k_B T}{\xi^E}\left(\Delta t - \frac{3}{2}\frac{m}{\xi^E} - 2\frac{m}{\xi^E}e^{-2\xi^E\Delta t/m} + \frac{2m}{\xi^E}e^{-\xi^E\Delta t/m} \right) \tag{5.49b}$$

These equations are equivalent to the ones that we used in Section 5.4.4, provided $\xi^E = \xi^{\text{QMC}}$, i.e. provided

$$(\sigma^E)^2\tau^E = (\sigma^{\text{QMC}})^2\Delta t \tag{5.50}$$

This gives us the required relation between σ^E and σ^{QMC}.

We have calculated τ^E and σ^E for the various atoms and for two fit curve (FIT 1 and FIT 2) [see Table 5.1(D)] and found that the above relation is not met exactly but well enough to justify the procedure used in the main text.

Results

We estimate the parameters τ^E and σ^E by performing a long NVE run (10 ps) where we evolve the system with the deterministic forces, F^C, and compute the fitted force, F^E, with weights parameters corresponding to FIT 1 at each time-step. We then calculate the auto-correlation function of the difference $F^Y = F^{\mathrm{QMC}} - F^E$ for each atom and fit its logarithm linearly in time in order to obtain τ^E ; σ^E follows immediately from the time zero value of the auto correlation function. All results are averaged over atoms of the same type and all directions of motion. The averaged results are collected in Table 5.1(D) and the resulting values of ω with the temperature correspondence are given in Tab.(E).

In Fig. 5.12, we show the total energies as averages over 100 runs, obtained with fit-parameters according to FIT 1 and with different values of ω corresponding to the temperatures listed in Table 5.1(E). In all cases, we have $\omega < 1$ so that the hyperbolic solution holds. For $\omega = 0.991$, the total energy is remarkably stable for times up to a bit more than one picosecond. Also for other cases with ω close to one, the results are encouraging; only with $\omega = 0.655$ does the energy decrease faster than when only the fit is applied.
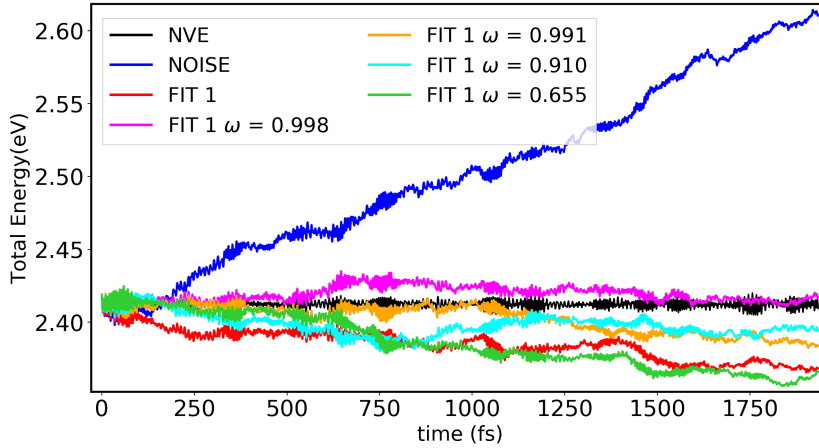
Figure 5.12: Average total energy over 100 trajectories obtained accounting for memory and using different values of ω. The fitted forces are estimated with the weights of FIT 1 and the parameters τ^E listed in Table 5.1.

In Fig. 5.13, we present similar calculations but, now, with the parameters that go with FIT 2. All curves are very close to the FIT curve, none of them improving on the latter. This is in contrast with previous calculations based on the simpler case of Langevin scheme 2 without memory.
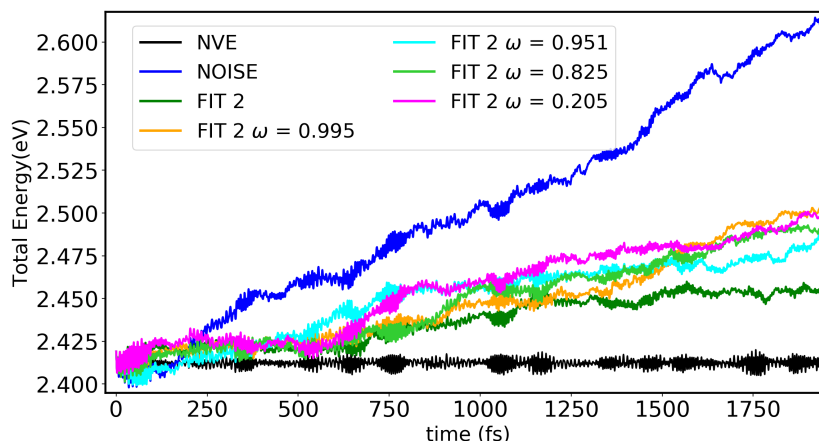
Figure 5.13: Average total energy over 100 trajectories obtained accounting for memory and using different values of ω. The fitted forces are estimated with the weights of FIT 2 and the parameters τ^E listed in Table 5.1.

In conclusion, we find that the scheme with memory yields similar results to the ones obtained without memory, with only small differences, sometimes to better sometimes to worse results. From a practical perspective, the memory case is more laborious since one has to determine the parameters σ^E and τ^E for each atom type. While this is easy and fast in the simulations with force fields and added noise, the task becomes more demanding in a QMC calculation. For this reason, we will prefer to adopt the simpler scheme with no memory in the following Chapter.

Bibliography

[1] G. C. Sosso, J. Chen, S. J. Cox, M. Fitzner, P. Pedevilla, A. Zen, and A. Michaelides, Chem. Rev. **116**, 7078 (2016).

[2] R. M. Venable, A. Krämer, and R. W. Pastor, Chem. Rev. **119**, 5954 (2019).

[3] S. J. Marrink, V. Corradi, P. C. Souza, H. I. Ingólfsson, D. P. Tieleman, and M. S. Sansom, Chem. Rev. **119**, 6184 (2019).

[4] L. Hung and E. A. Carter, Chem. Phys. Lett. **475**, 163 (2009).

[5] A. Erba, J. Baima, I. Bush, R. Orlando, and R. Dovesi, J. Chem. Theory Comput. **13**, 5019 (2017).

[6] M. Arita, D. R. Bowler, and T. Miyazaki, Journal of Chemical Theory and Computation **10**, 5419 (2014).

[7] X. Shao, Q. Xu, S. Wang, J. Lv, Y. Wang, and Y. Ma, Comput. Phys. Commun. **233**, 78 (2018).

[8] L. E. Ratcliff, S. Mohr, G. Huhs, T. Deutsch, M. Masella, and L. Genovese, WIREs Comput. Mol. Sci. **7**, e1290 (2017).

[9] R. Schade, T. Kenter, H. Elgabarty, M. Lass, O. Schütt, A. Lazzaro, H. Pabst, S. Mohr, J. Hutter, T. D. Kühne, and C. Plessl, Parallel Comput. **111**, 102920 (2022).

[10] H. M. Senn and W. Thiel, Angew. Chem., Int. Ed. **48**, 1198 (2009).

[11] E. Brunk and U. Rothlisberger, Chem. Rev. **115**, 6217 (2015).

[12] C. E. Tzeliou, M. A. Mermigki, and D. Tzeli, Molecules **27**, (2022).

[13] P. Gkeka, G. Stoltz, A. Barati Farimani, Z. Belkacemi, M. Ceriotti, J. D. Chodera, A. R. Dinner, A. L. Ferguson, J.-B. Maillet, H. Minoux, C. Peter, F. Pietrucci, A. Silveira, A. Tkatchenko, Z. Trstanova, R. Wiewiora, and T. Lelièvre, J. Chem. Theory Comput. **16**, 4757 (2020).

[14] O. T. Unke, S. Chmiela, H. E. Sauceda, M. Gastegger, I. Poltavsky, K. T. Schütt, A. Tkatchenko, and K.-R. Müller, Chem. Rev. **121**, 10142 (2021).

[15] C. Filippi, R. Assaraf, and S. Moroni, J. Chem. Phys. **144**, 194105 (2016).

[16] R. Assaraf, S. Moroni, and C. Filippi, J. Chem. Theory Comput. **13**, 5273 (2017).

[17] M. Dash, J. Feldt, S. Moroni, A. Scemama, and C. Filippi, J. Chem. Theory Comput. **15**, 4896 (2019).

[18] M. Dash, S. Moroni, C. Filippi, and A. Scemama, J. Chem. Theory Comput. **17**, 3426 (2021).

[19] S. Pathak, B. Busemeyer, J. N. B. Rodrigues, and L. K. Wagner, J. Chem. Phys. **154**, 034101 (2021).

[20] S. Shepard, R. L. Panadés-Barrueta, S. Moroni, A. Scemama, and C. Filippi, Double excitation energies from quantum Monte Carlo using state-specific energy optimization, 2022.

[21] M. Dash, S. Moroni, A. Scemama, and C. Filippi, J. Chem. Theory Comput. **14**, 4176 (2018).

[22] A. Cuzzocrea, S. Moroni, A. Scemama, and C. Filippi, J. Chem. Theory Comput. **18**, 1089 (2022).

[23] CHAMP is a quantum Monte Carlo program package written by C. J. Umrigar, C. Filippi, S. Moroni and collaborators.

[24] M. Burkatzki, C. Filippi, and M. Dolg, J. Chem. Phys. **126**, 234105 (2007).

[25] For the hydrogen atom, we use a more accurate BFD pseudopotential and basis set. Dolg, M.; Filippi, C., private communication.

[26] M. S. Gordon and M. W. Schmidt, *Theory and applications of computational chemistry* (Elsevier, 2005), pp. 1167–1189.

[27] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, J. Chem. Phys. **76**, 637 (1982).

[28] D. Frenkel and B. Smit, *Understanding Molecular Simulation*, 2nd ed. (Academic Press, Inc., 2001).

[29] J. A. Rackers, Z. Wang, C. Lu, M. L. Laury, L. Lagardère, M. J. Schnieders, J.-P. Piquemal, P. Ren, and J. W. Ponder, J. Chem. Theory Comput. **14**, 5273 (2018).

[30] SLATEC Common Mathematical Library, Version 4.1, July 1993 a comprehensive software library containing over 1400 general purpose mathematical and statistical routines written in Fortran 77.

[31] P. Gkeka, G. Stoltz, A. Barati Farimani, Z. Belkacemi, M. Ceriotti, J. D. Chodera, A. R. Dinner, A. L. Ferguson, J.-B. Maillet, H. Minoux, C. Peter, F. Pietrucci, A. Silveira, A. Tkatchenko, Z. Trstanova, R. Wiewiora, and T. Lelièvre, J. Chem. Theory Comput. **16**, 4757 (2020).

[32] Y. Luo, A. Zen, and S. Sorella, J. Chem. Phys. **141**, 194112 (2014).

[33] G. Mazzola and S. Sorella, Phys. Rev. Lett. **118**, 015703 (2017).

[34] E. Arnon, E. Rabani, D. Neuhauser, and R. Baer, J. Chem. Phys. **152**, 161103 (2020).

[35] C. Attaccalite and S. Sorella, Phys. Rev. Lett. **100**, 114501 (2008).

Chapter 6

Dynamics with QMC forces: exploring limits and possibilities

In the previous Chapter, we explored various approaches to reduce the impact of the stochastic noise in the forces on a molecular dynamics simulation. Here, we analyze other technical issues connected to the quantum Monte Carlo computation of the forces, namely, the reduction of the infinite variance of the force estimators and the need of well-optimized wave functions. Finally, we combine the findings of the two Chapters and perform Born-Oppenheimer molecular dynamics simulations driven by quantum Monte Carlo forces for two prototypical systems.

6.1   Introduction

We revisit here the various strategies we have developed to ameliorate the lack of energy conservation resulting from the use of noisy forces in molecular dynamics simulations. While, for most of the tests so far, we have utilized a classical force field to which we manually added Gaussian noise, we now test the approaches in actual QMC simulations where the noise comes directly from the use of Monte Carlo sampling to estimate the forces.

When moving from classical simulations with noise to quantum Monte Carlo runs, we encounter two additional problems. The first is the well-known issue of the force estimator having an infinite variance for approximate wave functions, which we solve following a standard procedure, namely, employing a guiding wave function finite at the nodes [22]. Despite the adoption of this scheme, we occasionally encounter forces that are more than three standard deviations from the average and must perform an additional check in the simulation to remove such forces during the molecular dynamics. The second problem concerns the optimization of the wave function. In particular, if the wave function is not fully optimized, we lose the correspondence between forces and potential energy, and the total energy is no longer conserved, decreasing in time during the molecular dynamics simulation. This problem is not present when performing a structural relaxation since, in that case, one is simply optimizing the potential energy with respect to the variational and structural

parameters. Having identified the need for a strict optimization of the wave function, we propose a strategy to speed up its convergence and, consequently, decrease the overall computational cost.

We test these methodological aspects in molecular dynamics simulations with quantum Monte Carlo forces for the carbon dimer and thiophene. We choose the C_2 molecule since the simplicity of its potential renders the analysis of the various issues more straightforward. Furthermore, thanks to its small size, C_2 represents a computationally affordable case for extensive testing. Thiophene is instead an interesting system since it constitutes a building block for a class of donor polymers in organic solar-cell devices [23–25]. For this reason, thiophene has been extensively investigated both with pump-probe experiments [26] and with theoretical studies trying to characterize the deactivation pathways following photo-excitation [27–31].

Because of its strong multireference character, the theoretical description of the lowest-energy bright state of thiophene is in fact quite challenging and, for this state, different highly-correlated methods predict vertical excitation energies which span a range as large as 0.5 eV [17]. Since QMC is able to accurately describe the vertical excitation energy as well as the excited-state optimal geometry of thiophene [17], this molecule is a good candidate for exploring the performance of QMC forces in molecular dynamics simulations in the excited state. In the last part of this Chapter, we present the preliminary results of such simulations, highlighting the remaining challenges posed by the use of QMC forces.

6.2   Computational details

We employ the program package CHAMP [33] to perform the variational Monte Carlo (VMC) calculations. For all calculations, we use scalar relativistic energy-consistent pseudopotentials and the corresponding aug-cc-pVDZ basis set [34, 35]. We employ Jastrow-Slater wave functions where the Jastrow factor contains 2-body terms accounting for electron-electron and electron-nucleus correlation. For the Slater part, we use either a Hartree-Fock (HF) determinant or complete active space (CAS) expansions, which are generated with the program GAMESS(US) [36]. At the start of the molecular dynamics, we fully optimize the wave function with the stochastic reconfiguration method [37, 38] and a damping factor $\tau_{\mathrm{SR}} = 0.05$. When optimizing the wave function in the excited state, we use a penalty method as introduced in Ref. [18] with the implementation described in Ref. [19] and $\lambda = 1$ a.u. Unless otherwise specified, the nuclei are moved within CHAMP with a standard velocity-Verlet algorithm [39, 40] and a time-step $\Delta t = 0.5$ fs. In order to perform a polynomial fit for the forces as introduced in Chapter 5, we link the library SLATEC [41] to CHAMP. The parameters used in the fit of the forces are listed in Table 6.1.

This work focuses on two prototypical molecules, namely, the simple C_2 dimer and the thiophene molecule. For C_2, we employ a CAS(4,4) wave function to correlate the two π_{1u} and the corresponding antibonding π_{1g} orbitals. When testing the divergences in the VMC estimator of the forces, we use a close-to-equilibrium geom-

etry with an inter-atomic distance of 1.26 Å. To test the effects of the wave function optimization and perform molecular dynamics simulations, we start from a stretched C_2 geometry with an inter-atomic distance of 1.58 Å. As in the previous Chapter, we always have zero velocities at the start of the MD simulations.

For thiophene in the ground state, we use a simple HF ansatz for the determinantal component of the Jastrow-Slater wave function to perform computationally affordable tests, and the same starting geometry as in the classical studies of the previous Chapter. When moving to the excited-state simulations of thiophene, we build a CAS(10,9) wave function comprising the two σ and corresponding σ^* orbitals on the CS bonds and the lowest three π and two π^* orbitals as suggested in Ref. [31]. As starting velocities and geometries, we use some configurations obtained from a Langevin simulation at 300 K performed with the classical force fields OPLS-AA and the package Tinker [42].

(A)	C_2	FIT 1	FIT 2	FIT 3
	τ_C	1.5	1.8	2.0

(B)	Thiophene	FIT 1	FIT 2	FIT 3
	$\tau_{S,C}$	1.8	1.9	2.0
	τ_H	1.0	1.3	1.4

Table 6.1: Values of the dimensionless (in units of Δt) parameters, τ_{fit}, used for the exponential fit of the forces (see Eq. 5.5) for (A) C_2 and (B) thiophene.

6.3 Forces with finite variance

As elaborated in the Methods Chapter, given a trial wave function Ψ_T, we write the energy in VMC as

$$E = \int dR\, E_L(R)\rho(R) \equiv \langle E_L(R) \rangle, \qquad (6.1)$$

where R denotes the positions of the N electrons, $E_L(R) = H\Psi_T(R)/\Psi_T(R)$ is the local energy, and $\rho(R) = |\Psi_T(R)|^2/\int dR\,|\Psi_T(R)|^2$ is the probability distribution that we sample during the Monte Carlo run.

Performing the gradient with respect to the nuclear positions, R_α, we obtain the nuclear force

$$F = -\nabla_\alpha E = \langle f(R) \rangle, \qquad (6.2)$$

with the local force given by [22, 43, 44]

$$f(R) = -(\nabla_\alpha E_L(R) + (E_L(R) - E)\nabla_\alpha \ln \rho(R)), \qquad (6.3)$$

where we assume that the wave function depends explicitly on the positions of the nuclei. This estimator of the forces obeys the so-called zero-variance principle: as

the wave function and its derivatives become exact, the variance goes to zero since E_L equals E for every electronic configuration, $\nabla_\alpha E_L$ becomes a constant, and, consequently, $f(R)$ is also constant.

Unfortunately, even though the estimator has zero variance in the limit of an exact eigenstate, it has an infinite variance for an approximate wave function. To understand this, let us consider for instance the term $E_L \nabla_\alpha \ln \rho$, which gives the following contribution close to the node

$$E_L(R) \nabla_\alpha |\Psi_T(R)|^2 = 2 \frac{H \Psi_T(R)}{\Psi_T(R)} \frac{\nabla_\alpha \Psi_T(R)}{\Psi_T(R)} \sim \frac{1}{\Psi_T(R)^2} \,, \tag{6.4}$$

which diverges as the wave function goes to zero. While the average of this term is finite since the distribution Ψ_T^2 goes to zero quadratically at the nodes canceling the divergence, the variance (see Eq. 5.3) is infinite since one needs to average the square of this term.

Various schemes have been proposed to regularize this estimator and overcome this problem [22,45,46]. Here, we employ a guiding wave function which is finite at the nodes as proposed in Ref. [22] and rewrite the estimator of the force as

$$F = \frac{\langle f(R) w(R) \rangle_{\Psi_g^2}}{\langle w(R) \rangle_{\Psi_g^2}} \quad \text{with} \quad w(R) = \frac{\Psi_T(R)^2}{\Psi_g(R)^2} \,, \tag{6.5}$$

where we are explicitly indicating that we are now sampling Ψ_g^2. The local quantity we now average does no longer diverge at the nodes and the variance of this new estimator is therefore finite. As guiding wave function, we use the same form as in Ref. [22],

$$\Psi_G(R) = \begin{cases} \Psi_T(R) & \text{for } d(R) \geq \epsilon \\ \epsilon (d(R)/\epsilon)^{d(R)/\epsilon} \Psi_T(R)/d(R) & \text{for } d(R) < \epsilon \,, \end{cases} \tag{6.6}$$

where $d(R)$ is a measure of the distance of the electronic configuration from the nodal surface, which we estimate here as

$$d(R) = \frac{1}{|\nabla \Phi_T(R)/\Phi_T(R)|} \,, \tag{6.7}$$

where Φ_T is the determinantal component of the trial wave function Ψ_T.

We test this regularization scheme on the simple C_2 at a fixed geometry close to equilibrium and perform multiple independent QMC runs of the same length for two different values of ϵ. As shown in Fig. 6.1, while the average of the force is the same for both values of ϵ, the spread of the force significantly decreases for the higher value of the node-distance parameter. That said, even when using this larger ϵ, we observe occasionally that the force is several standard deviations outside the average. We believe that this is related to numerical instabilities in the computation of the force and these issues are currently being investigated. However, for the purpose of performing MD simulations, such outliers cannot be used to update the atomic
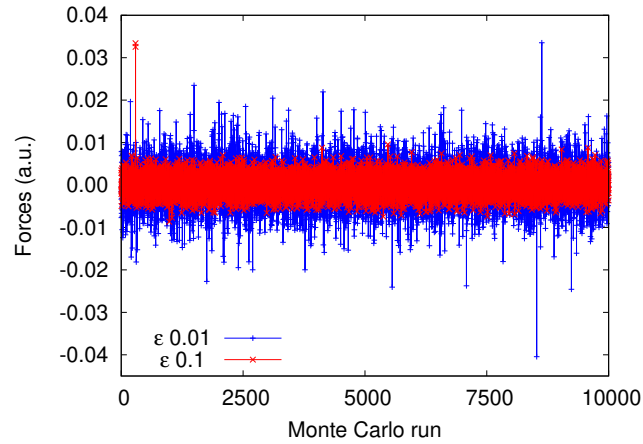
Figure 6.1: Force (a.u.) acting on one atom of a C_2 molecule close to equilibrium computed in 10.000 independent QMC runs of the same length with two different values of the node distance, ϵ.

configurations and we therefore include an additional check in the MD to manually remove them.

Rather than the force itself, we use its statistical error to select which forces must be discarded. This choice follows from the observation that, during a molecular dynamics run, the forces will naturally vary quickly but their errors should depend weakly on the configurations. Consequently, it is possible to use the error as an indicator of the quality of a force since a "kick" in the forces is typically mirrored by a similar variation in the error. This correlation is clear from Fig. 6.2 where we plot the absolute value of the QMC forces and the corresponding errors, and, in order to have more frequently large variations, we set the node-distance parameter to the lower value of $\epsilon = 0.01$.
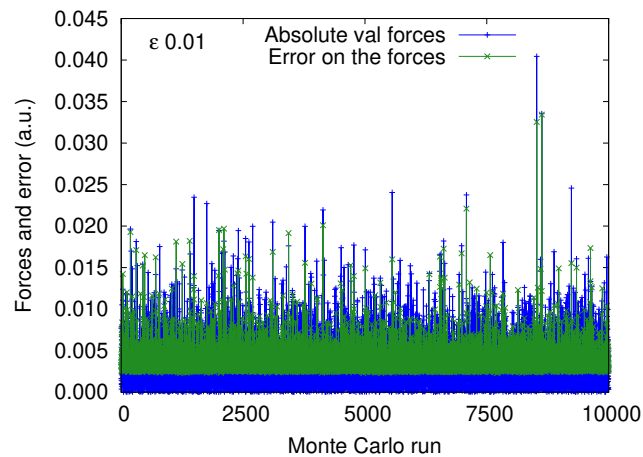
Figure 6.2: Absolute value of the force and its error (a.u.) in C_2, computed in multiple QMC runs with a parameter $\epsilon = 0.01$. The same geometry of C_2 is used as in Fig. 6.1.

127

If we denote the error as σ^{QMC} like in the previous Chapter (not to be confused with the root-mean-square fluctuation of the local force estimator), for each atom and each coordinate, we average this error over a short past history as

$$
\begin{aligned}
\overline{(\sigma_{i,j}^{\mathrm{QMC}})^2}(t) &= \frac{\int_{-\infty}^{t} dt'\, e^{-(t-t')/\tau} (\sigma_{i,j}^{\mathrm{QMC}})^2(t')}{\int_{-\infty}^{t} dt'\, e^{-(t-t')/\tau}} \\
&= \frac{1}{\tau} \int_{-\infty}^{t} dt'\, e^{-(t-t')/\tau} (\sigma_{i,j}^{\mathrm{QMC}})^2(t') \,,
\end{aligned}
\tag{6.8}
$$

where i and j label atoms and coordinates, respectively, and the past values enter in the average with an exponentially decaying weight as suggested in Ref. [47]. We introduce this time average instead of a fixed value of the error since also the error depends on the atomic configuration even though to a much lesser extent than the forces. By differentiating the last equation, we obtain

$$
\frac{d\overline{(\sigma_{i,j}^{\mathrm{QMC}})^2}(t)}{dt} = \frac{1}{\tau}[(\sigma_{i,j}^{\mathrm{QMC}})^2(t) - \overline{(\sigma_{i,j}^{\mathrm{QMC}})^2}(t)] \,,
\tag{6.9}
$$

and, to first order in the time step Δt,

$$
\overline{(\sigma_{i,j}^{\mathrm{QMC}})^2}(t + \Delta t) \simeq (1 - \frac{\Delta t}{\tau})\overline{(\sigma_{i,j}^{\mathrm{QMC}})^2}(t) + \frac{\Delta t}{\tau}(\sigma_{i,j}^{\mathrm{QMC}})^2(t) \,.
\tag{6.10}
$$

Therefore, at each time step of a molecular dynamics run, we update this time average and, if the current $(\sigma_{i,j}^{\mathrm{QMC}})^2$ for an atom/component of the force exceeds the corresponding $\overline{(\sigma_{i,j}^{\mathrm{QMC}})^2}$ by more than three times, we recompute the complete force (all atoms and coordinates) in the present configuration. Typically, we use $\tau = 15 \times \Delta t$ to average over a short history since the error varies slowly over time.

6.4 Optimization of the wave function

In the computation of the interatomic forces, we need to keep in account that the wave function can depend explicitly on the nuclear coordinates (e.g. via an atom-centered basis set) as well as implicitly via the wave function parameters. Therefore, the derivative with respect to the nuclear positions should be rewritten as

$$
\nabla_\alpha E = \frac{\partial E}{\partial R_\alpha} + \sum_i \frac{\partial E}{\partial c_i}\frac{dc_i}{dR_\alpha} \,,
\tag{6.11}
$$

where the first term was discussed in the previous Section and the second term contains the derivatives of the energy with respect to the parameters $\{c_i\}$ of the wave function. The second term can be neglected if one does not reoptimize the wave function or if the wave function is optimal and, consequently, the derivatives of the energy with respect to the parameters are zero.

Not surprisingly, non-optimizing the wave function is not possible when one considers large displacements with respect to the original set of nuclear coordinates

where the wave function was optimized. This is exemplified for a C_2 molecule in Fig. 6.3, where we can appreciate the dramatic difference between the potential energy curve obtained with a simple one-determinantal Jastrow-Slater wave function with optimal parameters at each MD step, and the one determined with fixed parameters (i.e. the wave function is just recentered at the new geometry).
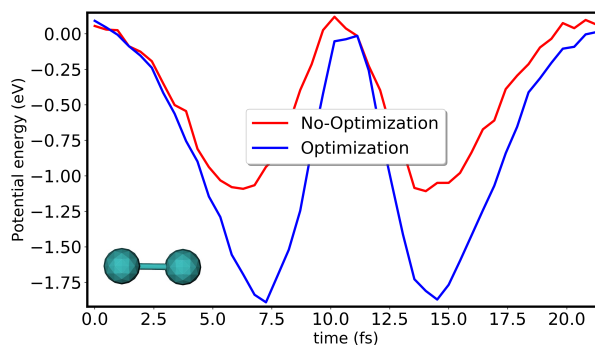

Figure 6.3: Potential energy curve (eV) of C_2 computed during an MD simulation with a wave function fully optimized at each time step (blue) and one in which the parameters are fixed and the wave function is recentered at the new positions (red).

In our simulations, we estimate the forces using only the first term in Eq. 6.11 and, since we aim at obtaining accurate results, we want to optimize the wave function at each MD step. A new question therefore arises, namely, how strict the variational optimization of the wave function in variational Monte Carlo should be to ensure that neglecting the second term yields forces which are compatible with the corresponding potential energy surface. To assess this, we perform several short MD runs for C_2 in which we employ a different number of optimization iterations of the wave function at each MD step, and show the results in Fig. 6.4. We find that a short variational optimization of 5 and 20 optimization iterations leads to a decrease of the total (kinetic plus potential) energy, even over the very short simulation times of one oscillation of C_2 (about 20 fs): with only five optimization iterations, we lose as much as 0.8 eV of energy. We recover conservation of energy by increasing the number of optimization iterations per MD step to fifty or one hundred. As expected, when we do not optimize the wave function so that the second term in Eq. 6.11 is identically zero, the energy is also conserved. As it will be clear below, the number of iteration steps needed depends on the total energy of the system, namely, on how fast and in which portion of the potential energy surface the system is moving. Here, we start indeed from a very stretched geometry of C_2 with a bond of 1.58 Å while the minimum is around 1.25 Å.

To understand this surprising decrease in total energy, we plot the single components of the energy, namely, kinetic and potential energies, as a function of the number of iterations in the optimization of the wave function in Fig. 6.14 in Appendix 6.A. For the case of five optimization iterations, the potential energy of the system starts to deviate from the 100 iteration case already after the first oscillation
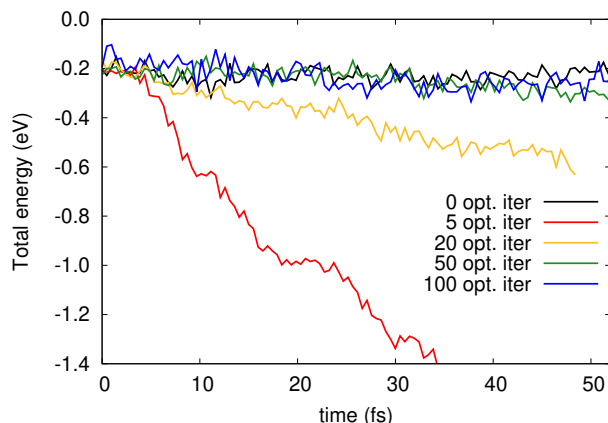
129

Figure 6.4: Total (kinetic plus potential) energy (eV) of a C_2 molecule during an MD run driven by QMC forces, where the variational parameters in the wave function are optimized with a different number of optimization iterations at each MD iteration.

and, as the simulation proceeds, the system loses kinetic energy, cooling down to relatively low energies. This behavior results from the fact that a few optimization steps are sufficient to give a seemingly converged potential energy during the first MD steps but are not enough to yield converged interatomic forces, and the system slowly starts to lag behind. In Fig. 6.5, we show the convergences of the forces during an MD simulation. In particular, we select the time interval over which the wave function changes the fastest (e.g. the steepest part of the potential in Fig. 6.3). We see that, in the part where the potential energy increases during the dynamics, stopping the variational optimization too early yields forces that are too large in absolute value. Oppositely, where the potential energy decreases, the forces are too small.

This means that effectively the system is moving on a different potential energy surface as schematically illustrated in Fig. 6.6. When the system is climbing the steepest part of the potential energy surface (PES), a not-fully optimized force is higher in absolute value than the actual force since, during the variational optimization, the energy is converging from a starting higher value and, therefore, also a higher force. For the same reason, the force will be too low when moving towards the minimum. Consequently, in the first case, the force will slow down the system too much, which makes the system lose too much kinetic energy. In the second case, the force will push the system down too little and, again, the system will lose kinetic energy. The decrease of kinetic energy means that the portion of potential energy visited gradually decreases, as shown in Appendix 6.A. The energy loss slows down in time and the total energy eventually stops decreasing when the chosen number of optimization iterations is enough to describe the slower change of the wave function characterizing the lower-energy portion of the PES. In fact, in a system with high kinetic energy (for example, when we start from a very stretched geometry), we will visit in less time more different configurations than if the kinetic energy is low. Consequently, the wave function from one-time step to the next will change more
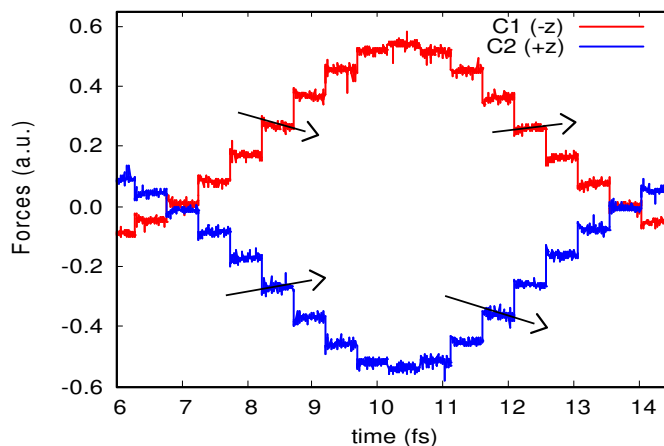
130

Figure 6.5: Forces (a.u.) on the two carbon atoms of the C_2 molecule during an MD run, where we also show the 100 steps of wave function optimization performed at each MD step. The time range corresponds to the central part of the potential energy surface shown in Fig. 6.4. The arrows are only intended as a guide for how the force changes during the variational optimization.

and need more extended optimization. On the other hand, if we have small kinetic energies, we visit very close configurations, so the wave function only needs minor adjustments at each MD step.
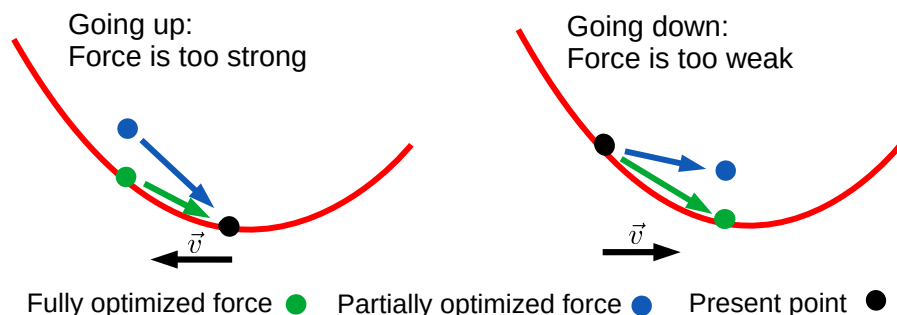
Figure 6.6: Schematic behavior of fully and partially-converged forces when moving upwards (left) and downwards (right) a PES. The unconverged forces correspond to an effective (not fully optimized) PES higher than the variational one. We add a black arrow to signal the direction of the velocity at the present point.

These findings underline the necessity of fully optimizing the wave function when performing an MD simulation driven by QMC forces. Since performing such optimization at each MD step can be quite costly, we exploit the robustness to noise of the stochastic reconfiguration optimization method to speed up the optimization. In particular, we begin the variational optimization using very short and therefore

cheap MC runs, and linearly increase the number of MC steps with the optimization iteration as to ensure well converged forces. In the Results Section, we also present a different method to speed up the wave function optimization by using information from previous MD steps, which we have tested so far only on the C_2 molecule, obtaining encouraging results.

6.5 Results

We assess here the performance of the various strategies introduced in the previous Chapter in actual QMC simulations as opposed to classical ones with force fields with added Gaussian noise. Furthermore, we investigate the additional effect potentially introduced by the optimization of the wave function. For all calculations, we use a node-cutoff value of $\epsilon = 0.05$ and control the forces as explained above in Section 6.3.

6.5.1 The carbon dimer

We begin by analyzing the molecular dynamics simulations driven by QMC forces for a simple C_2 molecule. We first investigate the case in which we do not optimize the wave function and perform averages over 20 trajectories of 5 ps. For C_2, we use longer times than for thiophene because the increase in total energy is slower given the smaller number of degrees of freedom. We always start the simulations from a stretched geometry with a bond of 1.587 Å and compute the forces in Monte Carlo runs of $N_{MC} = 250880$, which corresponds to an error on the energy of about 1 mHa.

In Fig. 6.7, we illustrate the cumulative effect of the different strategies to ameliorate the impact of noise presented in Chapter 5. We find that, also here, constraining the rotational and translational motion of the molecule slows down the increase in total energy but does not fully solve the problem (panel a). In addition to these constraints, we then introduce an exponential fit of the forces over previous steps (Eq. 5.5) and use the fitted forces to displace the nuclei. We employ three different sets of fit parameters τ_{fit} reported in Table 6.1(A) and obtain a significant improvement in the conservation of total energy with a nearly stable simulation over about 3 ps for one set of parameters (panel b, FIT 3). We note that, in this case, the optimal parameter is $\tau_C = 2$ while, in the classical case, for the carbon atom of thiophene, we have $1.2 < \tau_C < 1.5$. It is reasonable to expect that changing system and error on the forces (i.e. N_{MC}) leads to different optimal parameters. We will further comment on this aspect for thiophene, where the starting configurations of the classical and quantum simulations are identical.

Finally, we fix the fit parameter τ_{fit} corresponding to FIT 2 and use the Langevin scheme 2 (Eq. 5.26) to thermalize the noise. Differently from the previous Chapter, we let the damping parameter γ vary with the error in the QMC forces and use a

(a) Removing rotations

(b) Fitting the forces.
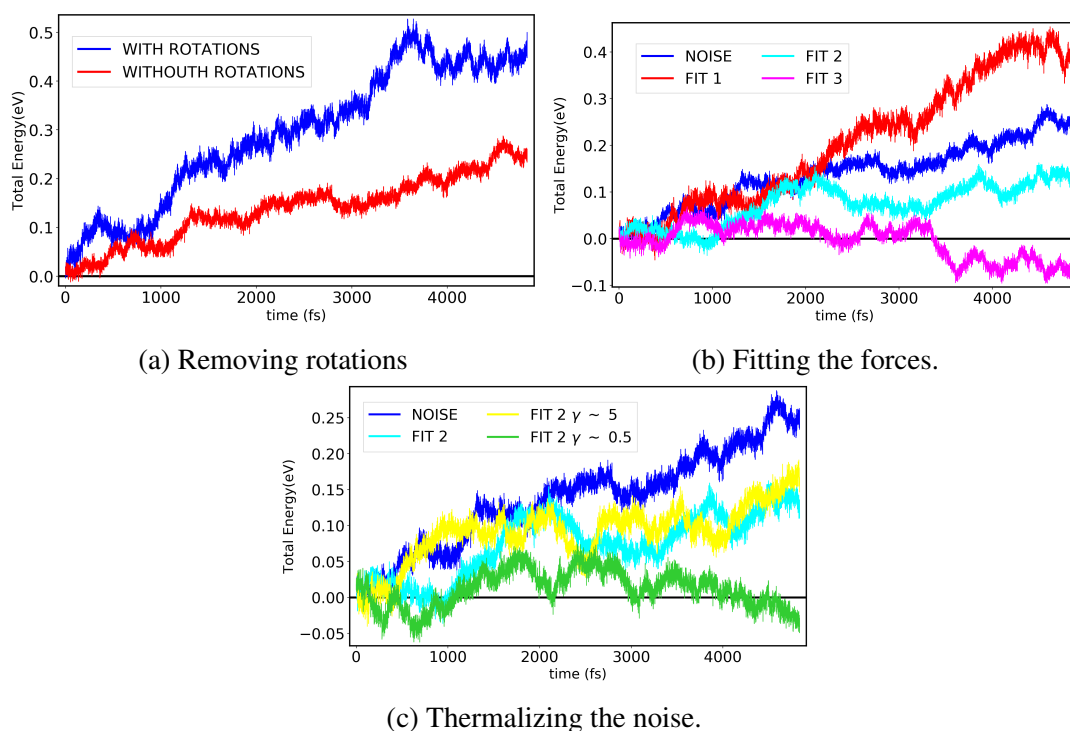

(c) Thermalizing the noise.

Figure 6.7: Total (kinetic plus potential) energies (eV) of C_2 averaged over 20 molecular dynamics runs, where the wave function is not optimized but simply recentered at each time step. In panel a (top left), we perform simulations constraining the motion of the center of mass and the rotation of the molecule. In panel b (top right), we use fitted forces with the fit parameter τ_{fit} reported in Table 6.1(A) and, in panel c (bottom), thermalize the noise where we use the fitted forces FIT 2 and report the value of γ averaged over all atoms and coordinates. In panels b and c, we constrain the rotational and translational motion.

$\gamma_{i,j}(t)$ which varies in time and is also atom and coordinate dependent as

$$\gamma_{i,j}(t) = \frac{\overline{(\sigma_{i,j}^{\text{QMC}})^2}(t)\Delta t^2}{2k_{\text{B}}Tm_i} \,, \tag{6.12}$$

where i and j label the different atoms and coordinates. As explained above, we prefer to employ the time-averaged value $\overline{(\sigma^{\text{QMC}})^2}(t)$ as defined in Eq. 6.10 to avoid unwanted sharp changes in γ that may come from the Monte Carlo computation of the forces. As in the previous Chapter, the temperature is an input parameter.

In Fig. 6.7c, we show the simulations for two values of the temperature, T = 0.01 and 0.1 K, which correspond to a value of γ at the first time step of about 5 and 0.5, respectively, where we report here a single value averaged over the two atoms and all coordinates as in the previous Chapter. The results are very encouraging: for the smaller γ, we improve on the fit and obtain a stable total energy for the all 5 ps of the calculation. For the higher value of γ, we have a similar behavior to the corresponding fit curve as expected from the limit of the integration scheme.

Having assessed the validity of the findings in Chapter 5 for the case where we do not optimize the wave function, we now add the additional complication of performing a variational optimization at each MD step, and present the results in Fig. 6.8. At each MD step, we perform 150 steps of wave function optimization with the stochastic reconfiguration method, where for safety, we increase the number of steps to 150 from the 100 used in the previous Section. Since having such a long variational optimization leads to a considerable increase in the computational cost, we perform here averages on five trajectories and stop the simulations after 2.5 ps, while keeping fixed the value of N_{MC} to compute the final forces.

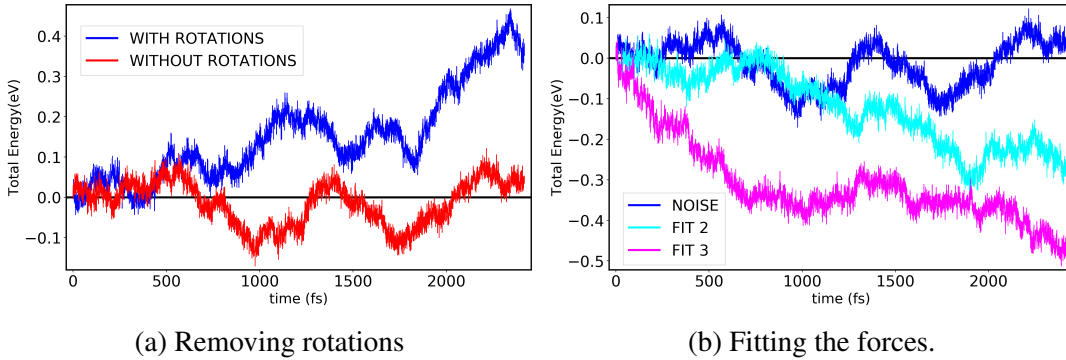| (a) Removing rotations | (b) Fitting the forces. |

Figure 6.8: Total (kinetic and potential) energies (eV) for C_2 averaged over five molecular dynamics simulations where the wave function is optimized with 150 steps of the stochastic reconfiguration method at each MD step. In panel a (left), we show the total energy when we remove rotations and translations and, in panel b (right), the results for different fit parameters as reported in Table 6.1(A), where we remove rotations and translations.

As regards constraining the translation and rotational motion, we observe the same trends as before, namely, a slowing down of the increase in total energy (panel

a). Moreover, even though one would expect additional noise with respect to the case with non-optimized wave functions, we do not observe significant differences in the spread of the trajectories. However, there are some surprising differences:

1) the total energy of the noisy curve (NOISE) does not increase in the 2.5 ps time of the simulation;

2) when fitting the forces (panel b), the fit parameters that were optimal before (FIT 2 and 3 in Fig. 6.7) are now clearly worsening the conservation of energy and, in particular, in both cases, the total energy decreases in time.

To understand these findings as compared to the case with no wave function optimization, there are three factors to keep in mind:

i) we have lower statistics since we average over 5 trajectories compared to the previous 20;

ii) the optimization of the wave function leads to a change in the potential so that the optimal fit parameter τ_{fit} may change;

iii) 150 iterations of wave function optimization with the stochastic method might not be enough to guarantee sufficiently converged forces and a stable total energy over long times.

While all factors can, in principle, contribute to the observed results, the constant total energy of the noisy curve (obtained without a fit and removing rotations and translations; Fig. 6.8a) is particularly unexpected. Indeed we kept the same total number of Monte Carlo steps (N_{MC}) in computing the final forces used to move the atoms as in the case without wave function optimization (Fig. 6.7a). Consequently, we would expect a similar increase in the total energy.

This consideration leads us to conclude that factor iii) is likely playing the most significant role. Therefore, we infer that, while setting the number of stochastic reconfiguration optimization steps to avoid energy loss (Fig. 6.4), we reached an incorrect conclusion. Indeed we were focusing on the conservation of energy over too short time scales, namely, 0.025 ps compared to the current 2.5 ps. We expect the optimal fit parameters for the case with and without wave function optimization to be in the same range of values. For this reason, we conclude that the energy conservation observed with the noisy curve without rotations and translations (Fig. 6.8a) is due to a fortuitous cancellation of the increase of energy due to noise and the decrease due to insufficient optimization steps. Similarly, the energy decrease we observe when fitting the forces (Fig. 6.8b) is again due to a lack of convergence of the forces in the wave function optimization. We further comment on this aspect while examining the thiophene molecule in an attempt to obtain a more complete picture.

6.5.2 Thiophene

For testing the case of thiophene, we use a simple HF starting wave function and, to limit the computational cost, a smaller number of Monte Carlo steps, N_{MC}, than

135

the one employed in the previous Chapter to set the standard deviation, σ^{QMC}, of the Gaussian noise added to the classical force-field forces. In particular, we use $N_{\mathrm{MC}} = 512000$, which corresponds to an error in the energy of about 1.5 mHa.

We present the results in a similar way as for C_2 starting with simulations where the wave function is not optimized, but restrict the averages to only 5 trajectories and the total time of the simulation to 1 ps. Having only 5 runs makes it more difficult to have quantitative estimates but still allows us to assess the trends in the simulations. As expected and shown in Fig. 6.9, constraining the translational and rotational motion of thiophene slows the increase in total energy during the molecular dynamics (panel a). Similarly, the fit procedure yields the intended improvement on the use of noisy forces (panel b). However, the optimal fitting parameters, τ_{fit}, listed in Table 6.1(B) are quite different from the values obtained with a classical force field in the previous Chapter (see Table 5.1(B)). In particular, both $\tau_{\mathrm{S,C}}$ and τ_{H} are larger than in the classical case. Given that we start from the same initial conditions as in the classical case, we attribute this discrepancy to the difference between the QMC and the classical potential energy surface. For example, while in the classical case, the initial conditions lead to an average kinetic energy of 1 eV, here, the kinetic energy is about 0.85 eV. While the higher error in the forces (i.e. the choice of N_{MC}) can ,in principle, also be responsible for the change in parameters, we rule this out since, for C_2 and different lengths of the Monte Carlo runs to compute the forces, we find the same range of optimal carbon parameters.

Finally, in Fig. 6.9c, we employ Langevin scheme 2 to thermalize the noise in simulations with fitted forces (FIT 2) and different values of the γ parameters, which are computed as in Eq. 6.12 for C_2. We use different temperatures of 0.05 and 0.1 K, which correspond to a value of γ at the first time step of 2 and 0.7, respectively (averaged over all carbon atoms and coordinates). The behavior of the total energy is similar to the case of C_2, where the larger γ does not improve on the fit. Given the poorer statistics than in the C_2 case, we cannot conclude that the results relative to the smaller γ ameliorate the fit ones: even though the increase in energy is slower at longer times, it also appears that the total energy is underestimated between 200 and 400 fs. Based on the current results, we can conclude that when we do not optimize the wave function, we find similar trends for thiophene in QMC as in the classical case, with the fitting procedure improving on the use of noisy forces. Due to the difference with the force field PES, the fit parameters must, however, be reset, and further investigation is needed on the effects of the thermalization of the excess noise with respect to the fit.

Also, for thiophene, we now move to the case where we optimize the wave function, using different numbers of steps in the variational optimization with the stochastic reconfiguration method at each MD step. As shown in Fig. 6.10 for averages over 5 trajectories of 0.3 ps, the use of 150 steps of variational optimization yields a nearly constant energy for the entire simulation time, while we recover the same increase with 200 steps as in the case with no optimization (panel a). This leads us to conclude that 150 steps are insufficient to converge the QMC forces and that we are experiencing the same issue encountered in Section 6.4: the decrease in energy due to lack of convergence in the forces balances by chance the increase due to noise. Therefore,

(a) Removing rotations.

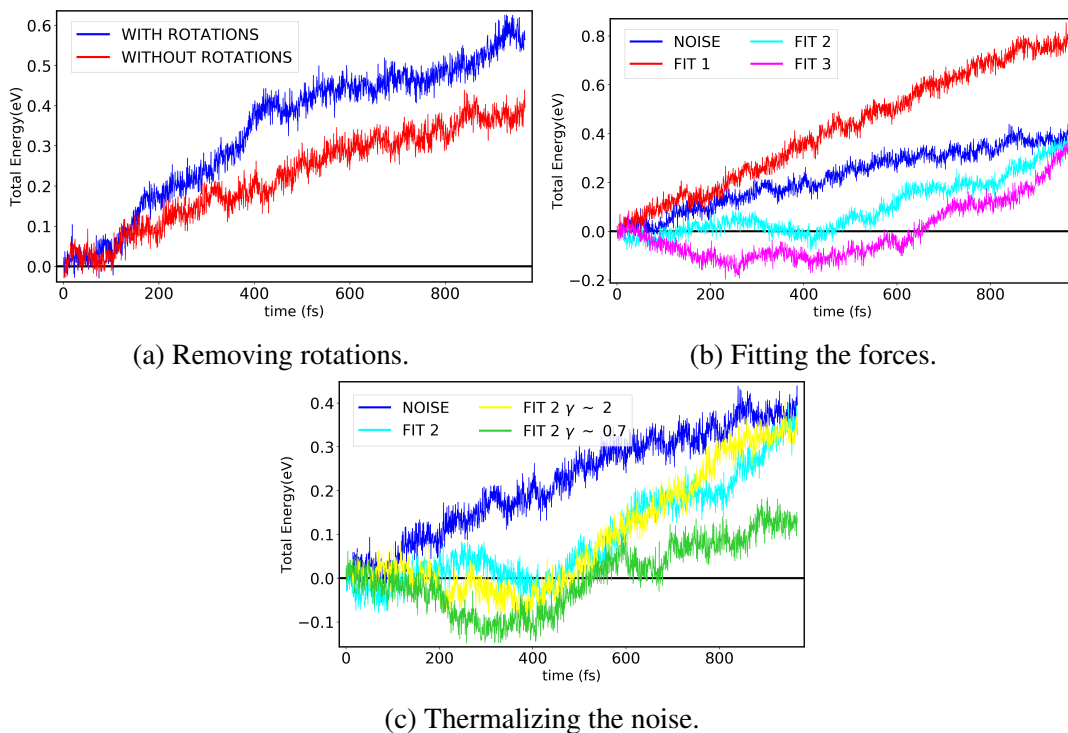(b) Fitting the forces.

(c) Thermalizing the noise.

Figure 6.9: Total (kinetic plus potential) energies (eV) of thiophene averaged over 5 molecular dynamics runs when we recenter the wave function without optimizing it. In panel a (top left), we show the total energy when we do and do not constrain the translational and rotational motion. In panel b (top right), we fit the forces with the different fit parameters, τ_{fit}, given in Table 6.1(B). In panel c (bottom), we thermalize the noise for the fitting procedure FIT 2 and different values of γ labeled with the average over all carbon atoms and coordinates. Both in panels b and c, we remove translations and rotations from the system.

(a) Runs with noisy forces.

(b) Runs with fitted forces.
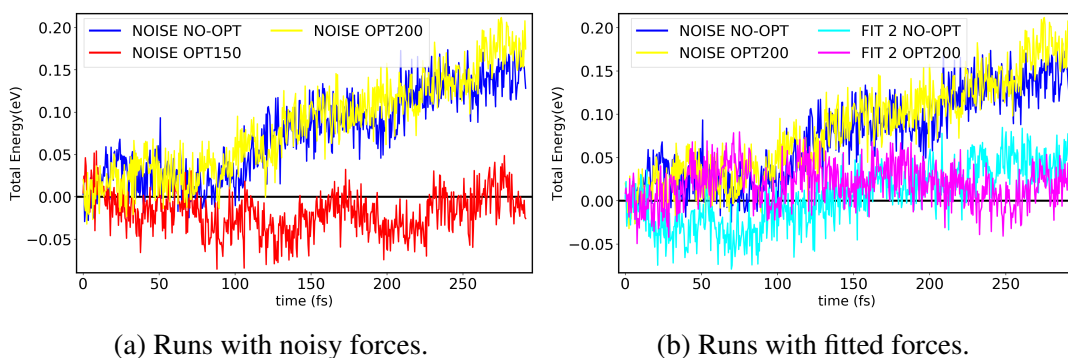
Figure 6.10: Total (kinetic plus potential) energies (eV) of thiophene averaged over 5 molecular dynamics runs. Panel a (left): energies obtained with 150 and 200 iterations of variational optimization at each MD step, compared with the no-optimization case. Panel b (right): results relative to the fit procedure with parameters FIT 2. In all cases, we constrain the translational and rotational motion.

comparing the behavior of the total energy with the equivalent run with no optimization appears to be a safe way to assess the number of needed steps in the variational optimization. We note that, unfortunately, the averages obtained with/without optimization and for different numbers of optimization steps behave in a very similar way in the first 0.1 ps, rendering difficult a quick determination of the optimal number of steps with short runs, as we had already inferred for C_2.

Finally, in Fig. 6.10b, we show the total energy obtained in simulations with fitted forces and the parameters labeled FIT 2, with and without variational optimization. In particular, we use 200 optimization iterations of optimization and, in the short-time investigated, find a similar behavior to the no-optimization case. Therefore, also when we turn on the optimization, we can stabilize the total energy following the fit procedure, provided a sufficiently high number of iterations of wave function optimization is performed. Additionally, the fit parameters appear to be transferable from the no-optimization case to the one when the wave function is optimized.

6.5.3 Excited-state relaxation of thiophene

All tests so far on the use of QMC forces in molecular dynamics clearly point to the slow convergence of the variational optimization as an even more severe problem than the presence of noise in the forces, at least over short times. In the examples of the previous Section, this issue was controlled by performing stricter optimizations. Here, as last test, we further raise the stakes and perform preliminary simulations of excited-state molecular dynamics with QMC forces. This study is at an early stage and should be seen as a proof of concept. In particular, we limit the analysis to two trajectories with different starting points, do not consider possible non-adiabatic effects, and employ a state-specific energy minimization approach based on a penalty, which was only recently implemented in our in-house QMC code [19].

For the time being, we do not constrain the translational and rotational motion, we do not fit the forces nor thermalize the noise. To ensure that the fluctuations in the forces are finite, we use a node-cutoff parameter $\epsilon = 0.05$ and further control the forces as explained in Section 6.3. In addition, we optimize the wave function at each time step, which is necessary if one wants to describe the system moving away from the ground-state minimum configuration towards the conical intersection region. To this aim, we perform 200 interactions of stochastic reconfiguration optimization as in the previous thiophene calculations.

In Fig. 6.11, we show the time evolution following the excited-state forces for two runs that start from different configurations. The runs employ the same number of Monte Carlo steps and are therefore affected by the same amount of stochastic noise. In one run (panel a), the system reaches a distorted configuration with an out-of-plane sulfur atom, which we know from previous work in the literature to be visited in the excited-state relaxation, being, in fact, the excited-state minimum. However, since the total energy is clearly not being conserved (panel b), we decided to stop the simulation. In the other run (panel c), the system reaches a conical intersection region in about 50 fs in line with the results of Ref. [29]. Moreover, the final configuration is the ring puckered one, which has a crucial role in the deactivation process and occurs

on the same time scales as the more common ring opening one [29].

(a) Potential energies of the two states (run 1).

(b) Total energy of the excited state (run 1).

(c) Potential energies of the two states (run 2).

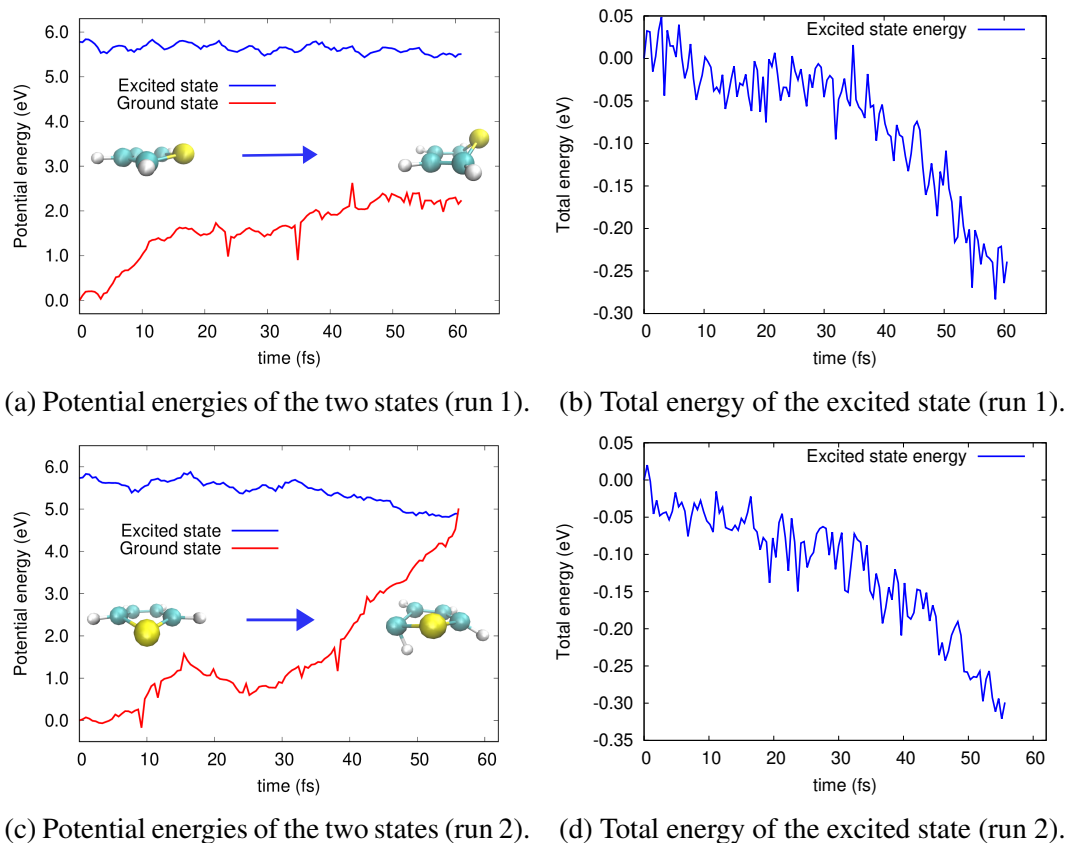(d) Total energy of the excited state (run 2).

Figure 6.11: Potential and total energies (eV) of two molecular dynamics simulations in which we follow the excited-state forces starting from two different initial conditions. In panels a and c (left), we plot the ground- and excited-state potential energies (with the starting ground-state energy as offset) during the two simulations and also depict the final structure of the molecule. In both runs, the statistical error on the potential energy is about 1 mHa (0.027 eV). In panels b and d (right), we plot the corresponding total energies (with the starting total energy as offset).

Unfortunately, both runs exhibit a large loss in total energy of about 0.3 eV in 60 fs. While the molecule does not come to rest as in the examples of the previous Section (see Appendix 6.B), this lack of energy conservation implies that we cannot trust the results, for example, regarding the time needed to reach the conical intersection. Based on our previous findings, the decrease in total energy indicates that 200 optimization iterations are insufficient to converge the forces. This is perhaps to be expected since we moved from a HF wave function in the previous tests to a more complex complete-active-space wave function with as many as 5292 linear coefficients to optimize. Since raising the number of optimization steps beyond 200 is computationally quite demanding, various routes for accelerating the variational optimization are currently being explored within the group. We also note that some jumps are present in the ground-state PES, which are possibly due to noise either in

the energy (we are following the forces in the excited state) or in the optimization of the multiple states, which must be further investigated.

In conclusion, more work is clearly needed to render the use of VMC forces combined with the multi-state optimization a viable tool to investigate excited-state relaxation processes. It is, of course, encouraging that the system reaches configurations that align with previous findings [28, 29], but the main obstacle over short time scales remains the optimization of the wave function. In the next Section, we explore a possible method aimed at speeding up the optimization, which looks promising for C_2. Further work is ongoing within the group to accelerate the stochastic reconfiguration method via the use of appropriate rescaling factors for the various parameters [48, 49] which is currently being tested on thiophene in the ground state.

6.5.4 Accelerating the optimization of the wave function

Here, we propose a strategy to accelerate the convergence of the wave function optimization by using the wave function of the previous MD configurations to predict better starting parameters at the present MD step. In Fig. 6.12, we show how the wave function parameters of C_2 vary in time during an MD run in which we optimize them at every step. Both the CI and orbital parameters follow a clear oscillatory pattern related to the motion of the molecule. Given their smooth variation in time, it is possible to save a short history of the parameters to predict the present ones through a fit.
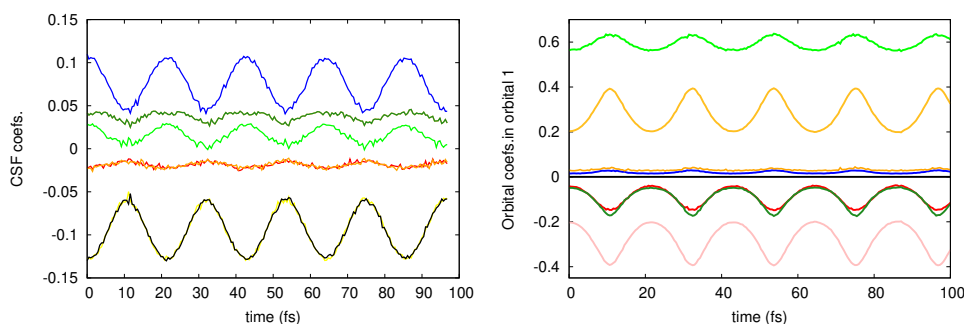
Figure 6.12: Time evolution of some linear and orbitals coefficients during an MD run of C_2 in which we optimize the parameters at each MD step. At every MD step, we plot the values of the parameters at the last of the 100 iterations of variational optimization.

The procedure is similar to the fit of the forces performed in Chapter 5: we employ a third-order polynomial fit and, since a prediction with no knowledge of the present is too unstable, we include some information from the present configuration. To this aim, we perform 50 optimization iterations at each MD step and save the optimal parameters of the previous 14 steps. We then perform a smaller number of optimization steps (20) at the present configuration and fit the previous and present parameter values (15 steps in total) with a third-order polynomial to estimate a better
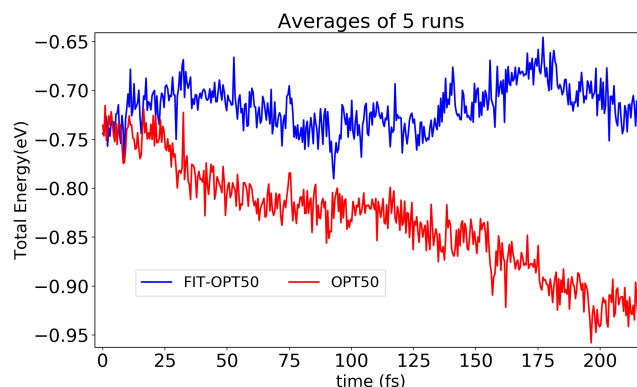
Figure 6.13: Total energies (eV) averaged over five MD runs where 50 iterations of variational optimization are performed at each MD step with a standard procedure (OPT50) and by fitting the orbital coefficients (FIT-OPT50) as described in the text.

starting point for the wave function at the present step. We then proceed with the optimization from the fitted wave function for 30 more iterations and finally compute the forces. In Fig. 6.13, we apply this procedure to the optimization of the orbital coefficients and obtain conservation of energy with only 50 optimization iterations, compared to the more than 100 iterations needed to have constant energy without the wave function fit. For completeness, in the same Figure, we also show that the energy decreases if we use 50 variational optimizations with no fit.

6.6 Conclusion

In this Chapter, we have investigated the use of VMC forces in molecular dynamics simulations. Our ultimate interest is to follow excited-state deactivation pathways, so we have focused on the time scale of picoseconds and encountered three main obstacles: i) the stochastic noise of the forces; ii) the infinite variance of the force estimators; iii) the need for a well-converged optimization of the wave function and, consequently, well-converged interatomic forces.

We have shown that it is possible to ameliorate the impact of all three issues, but more work remains to be done, especially on the first and the last one. Concerning the presence of noise in the forces, we have developed different strategies which help in reducing the increase of total energy over the time scales of interest. The main drawback is the need to tune various parameters (e.g. in the fit of the force), which will depend on the starting conditions and on the portions of PES being visited by the system. Ultimately, a functional form more flexible than the simple third-order polynomial used here will have to be employed. Second, we have shown that well-converged variational optimization is necessary for achieving energy conservation. This step is the most computationally costly, and more research is needed before we can perform affordable production runs. However, we believe that speeding up the convergence is possible either through modifications in the stochastic reconfiguration

method or through developments along the lines of what presented above.

In summary, this work takes a first, in-depth look at the possibilities and limitations of the use of QMC for molecular dynamics. Despite the remaining challenges, we believe that it shines a positive light on the potential use of QMC methods to study dynamical processes in the excited state on the time scale of picoseconds. We stress that such a study has only recently become possible thanks to the fast access to the computation of energy derivatives and to the continuous development of robust protocols to build excited-state wave functions and optimizers.

6.A   Appendix: Impact of the variational optimization on potential and kinetic energies

In Fig. 6.14, we show the kinetic and potential energies of C_2 computed in MD runs where we perform different numbers of optimization iterations of the wave function.

With 5 or 20 optimization iterations per MD step, we nearly recover the potential energy of a fully optimized wave function (100 iterations) during the first few MD steps. However, as time progresses, the system loses kinetic energy and the motion becomes confined to smaller parts (lower energy) of the PES.
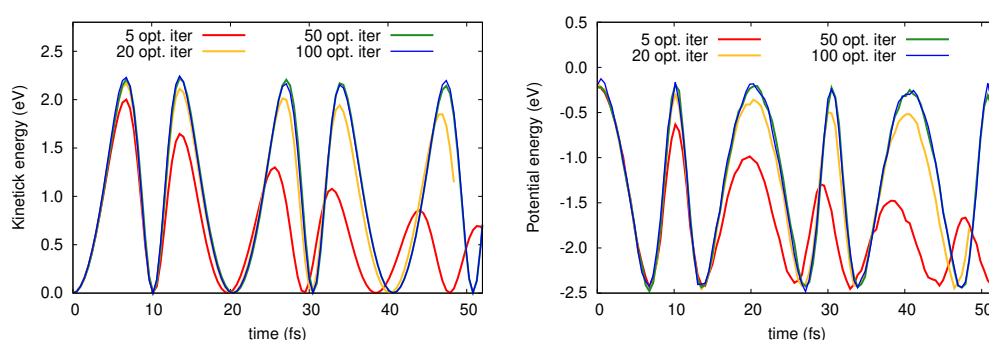
Figure 6.14: Kinetic and potential energy (eV) during an MD run in which we compute the QMC forces and energies after a different number of iterations of the variational optimization. The two plots correspond to the total energy presented in Fig. 6.4.

6.B   Appendix: Kinetic energy during the excited-state dynamics of thiophene

In Fig. 6.15, we plot the kinetic energy obtained during the two MD runs for thiophene in which we follow the excited-state forces. In these simulations, the total energy decreases by 0.3 eV over 60 fs and the results of Run 1 clearly indicate that the kinetic energy is slowly being damped. For Run 2, since the potential energy is quickly decreasing towards a conical intersection, the kinetic energy is somewhat increasing but not sufficiently to ensure energy conservation.

143
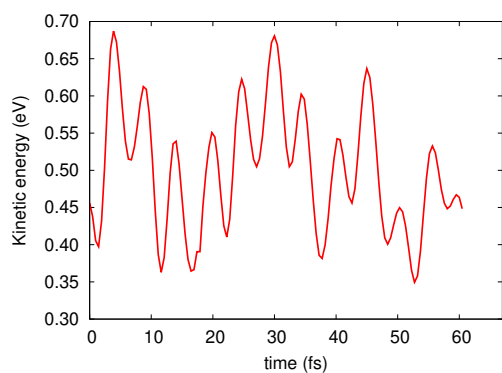
(a) Kinetic energy (run 1) (b) Kinetic energy (run 2)
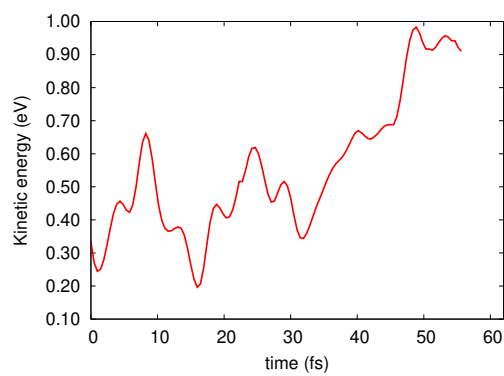
Figure 6.15: Kinetic energy (eV) of the two molecular dynamics runs of thiophene driven by the QMC forces in the excited state. The two plots correspond to the total energies of Fig 6.11.

Bibliography

[1] G. C. Sosso, J. Chen, S. J. Cox, M. Fitzner, P. Pedevilla, A. Zen, and A. Michaelides, Chem. Rev. **116**, 7078 (2016).

[2] R. M. Venable, A. Krämer, and R. W. Pastor, Chem. Rev. **119**, 5954 (2019).

[3] S. J. Marrink, V. Corradi, P. C. Souza, H. I. Ingólfsson, D. P. Tieleman, and M. S. Sansom, Chem. Rev. **119**, 6184 (2019).

[4] L. Hung and E. A. Carter, Chem. Phys. Lett. **475**, 163 (2009).

[5] A. Erba, J. Baima, I. Bush, R. Orlando, and R. Dovesi, J. Chem. Theory Comput. **13**, 5019 (2017).

[6] X. Shao, Q. Xu, S. Wang, J. Lv, Y. Wang, and Y. Ma, Comput. Phys. Commun. **233**, 78 (2018).

[7] L. E. Ratcliff, S. Mohr, G. Huhs, T. Deutsch, M. Masella, and L. Genovese, WIREs Comput. Mol. Sci. **7**, e1290 (2017).

[8] R. Schade, T. Kenter, H. Elgabarty, M. Lass, O. Schütt, A. Lazzaro, H. Pabst, S. Mohr, J. Hutter, T. D. Kühne, and C. Plessl, Parallel Comput. **111**, 102920 (2022).

[9] H. M. Senn and W. Thiel, Angew. Chem., Int. Ed. **48**, 1198 (2009).

[10] E. Brunk and U. Rothlisberger, Chem. Rev. **115**, 6217 (2015).

[11] C. E. Tzeliou, M. A. Mermigki, and D. Tzeli, Molecules **27**, (2022).

[12] P. Gkeka, G. Stoltz, A. Barati Farimani, Z. Belkacemi, M. Ceriotti, J. D. Chodera, A. R. Dinner, A. L. Ferguson, J.-B. Maillet, H. Minoux, C. Peter, F. Pietrucci, A. Silveira, A. Tkatchenko, Z. Trstanova, R. Wiewiora, and T. Lelièvre, J. Chem. Theory Comput. **16**, 4757 (2020).

[13] O. T. Unke, S. Chmiela, H. E. Sauceda, M. Gastegger, I. Poltavsky, K. T. Schütt, A. Tkatchenko, and K.-R. Müller, Chem. Rev. **121**, 10142 (2021).

[14] C. Filippi, R. Assaraf, and S. Moroni, J. Chem. Phys. **144**, 194105 (2016).

[15] R. Assaraf, S. Moroni, and C. Filippi, J. Chem. Theory Comput. **13**, 5273 (2017).

[16] M. Dash, J. Feldt, S. Moroni, A. Scemama, and C. Filippi, J. Chem. Theory Comput. **15**, 4896 (2019).

[17] M. Dash, S. Moroni, C. Filippi, and A. Scemama, J. Chem. Theory Comput. **17**, 3426 (2021).

[18] S. Pathak, B. Busemeyer, J. N. B. Rodrigues, and L. K. Wagner, J. Chem. Phys. **154**, 034101 (2021).

[19] S. Shepard, R. L. Panadés-Barrueta, S. Moroni, A. Scemama, and C. Filippi, Double excitation energies from quantum Monte Carlo using state-specific energy optimization, 2022.

[20] M. Dash, S. Moroni, A. Scemama, and C. Filippi, J. Chem. Theory Comput. **14**, 4176 (2018).

[21] A. Cuzzocrea, S. Moroni, A. Scemama, and C. Filippi, J. Chem. Theory Comput. **18**, 1089 (2022).

[22] C. Attaccalite and S. Sorella, Phys. Rev. Lett. **100**, 114501 (2008).

[23] F. Zhang, D. Wu, Y. Xu, and X. Feng, J. Mater. Chem. **21**, 17590 (2011).

[24] R. Wei, H. Chen, Y. Guo, H. Han, D. Zhang, Y. Zhu, F. He, and D. Zhao, Macromolecules **54**, 1499 (2021).

[25] C. Yang, S. Zhang, and J. Hou, Aggregate **3**, e111 (2022).

[26] R. Weinkauf, L. Lehr, E. W. Schlag, S. Salzmann, and C. M. Marian, Phys. Chem. Chem. Phys. **10**, 393 (2008).

[27] S. Salzmann, M. Kleinschmidt, J. Tatchen, R. Weinkauf, and C. M. Marian, Phys. Chem. Chem. Phys. **10**, 380 (2008).

[28] D. Fazzi, M. Barbatti, and W. Thiel, Phys. Chem. Chem. Phys. **17**, 7787 (2015).

[29] A. Prlj, B. F. E. Curchod, and C. Corminboeuf, Phys. Chem. Chem. Phys. **17**, 14719 (2015).

[30] P. Kölle, T. Schnappinger, and R. de Vivie-Riedle, Phys. Chem. Chem. Phys. **18**, 7903 (2016).

[31] T. Schnappinger, P. Kölle, M. Marazzi, A. Monari, L. González, and R. de Vivie-Riedle, Phys. Chem. Chem. Phys. **19**, 25662 (2017).

[32] A. Prlj, B. F. E. Curchod, A. Fabrizio, L. Floryan, and C. Corminboeuf, J. Phys. Chem. Lett. **6**, 13 (2015).

[33] CHAMP is a quantum Monte Carlo program package written by C. J. Umrigar, C. Filippi, S. Moroni and collaborators.

[34] M. Burkatzki, C. Filippi, and M. Dolg, J. Chem. Phys. **126**, 234105 (2007).

[35] For the hydrogen atom, we use a more accurate BFD pseudopotential and basis set. Dolg, M.; Filippi, C., private communication.

[36] M. S. Gordon and M. W. Schmidt, *Theory and applications of computational chemistry* (Elsevier, 2005), pp. 1167–1189.

[37] S. Sorella and L. Capriotti, Phys. Rev. B **61**, 2599 (2000).

[38] S. Sorella, M. Casula, and D. Rocca, J. Chem. Phys. **127**, 014105 (2007).

[39] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, J. Chem. Phys. **76**, 637 (1982).

[40] D. Frenkel and B. Smit, *Understanding Molecular Simulation*, 2nd ed. (Academic Press, Inc., 2001).

[41] SLATEC Common Mathematical Library, Version 4.1, July 1993 a comprehensive software library containing over 1400 general purpose mathematical and statistical routines written in Fortran 77.

[42] J. A. Rackers, Z. Wang, C. Lu, M. L. Laury, L. Lagardère, M. J. Schnieders, J.-P. Piquemal, P. Ren, and J. W. Ponder, J. Chem. Theory Comput. **14**, 5273 (2018).

[43] A. Badinski, P. D. Haynes, J. R. Trail, and R. J. Needs, J. Phys.: Condens. Matter **22**, 074202 (2010).

[44] S. Moroni, S. Saccani, and C. Filippi, J. Chem. Theory Comput. **10**, 4823 (2014).

[45] S. Pathak and L. K. Wagner, AIP Advances **10**, 085213 (2020).

[46] J. van Rhijn, C. Filippi, S. De Palo, and S. Moroni, Journal of Chemical Theory and Computation **18**, 118 (2022).

[47] A. van den Noort and W. J. Briels, Macromol. Theory Simul. **16**, 742 (2007).

[48] A. Scemama and C. Filippi, Phys. Rev. B **73**, 241101 (2006).

[49] J. Toulouse and C. J. Umrigar, J. Chem. Phys. **128**, 174101 (2008).

List of publications

- **A. Cuzzocrea**, A. Scemama, W. J. Briels, S. Moroni and C. Filippi
"Variational Principles in Quantum Monte Carlo: The Troubled Story of Variance Minimization"
J. Chem. Theory Comput. **16**, 4203–4212 (2020). (Chapter 3)

- **A. Cuzzocrea**, S. Moroni, A. Scemama, and C. Filippi
"Reference excitation energies of increasingly large molecules: a QMC study of cyanine dyes"
J. Chem. Theory Comput. **18**, 1089–1095 (2022). (Chapter 4)

- **A. Cuzzocrea**, C. Filippi and W. J. Briels
"Performing molecular dynamics with quantum Monte Carlo: exploring limits and possibilities"
In preparation. (Chapter 5/6)

Summary and Outlook

From vertical excitations towards excited-state relaxation: a journey with quantum Monte Carlo

With this dissertation, we contribute to the study of molecular excitations by assessing and expanding a specific electronic structure method: quantum Monte Carlo (QMC). The excited states of molecules play a central role in activating many natural processes such as human vision, and in the functioning of new technologies like solar panels. Their theoretical study is a very active field of research but, due to their complex nature, we still lack a standard procedure to analyze them. On one side, in many photo-induced processes, the effects of the environment on the molecule become decisive and have to be taken into account, for example, by using mixed quantum/classical approaches. On the other, outside the ground state equilibrium, the potential energy surfaces to be described is very complex, and especially when multiple states interact, the available quantum chemistry methods often fail.

In this thesis, we do not consider any effect of the environment and work on improving the quantum mechanical description of excited states by focusing on QMC methods, a class of techniques for solving the Schrödinger equation in a stochastic manner. Lately, they are attracting increasing interest in the electronic structure community thanks to their favorable scaling with the number of electrons and the natural ease in parallelization. In the context of excited states, they are affirming as a valid alternative to other methods, especially for those complex cases where the cheapest options (such as time-dependent density functional theory) fail in the description. Moreover, recent algorithmic advancements in the QMC community have made it possible to extend the description from small to medium and relatively large (about 100 non-hydrogen atoms) molecules and to compute accurate geometries in both ground and excited states. Encouraged by this work, in this dissertation, we further investigate the use of QMC for the study of molecular excited states. In particular, in the first half of the manuscript (Chapters 3 and 4), we try to build robust protocols to compute vertical excitations. In the second (Chapters 5 and 6), we explore and try to address the problems connected to using QMC methods to describe excited-state relaxation.

Building robust strategies to study vertical excitations. The use of QMC methods in the analysis of vertical excitation is a quite recent and fastly developing field. We here contribute to it by investigating two crucial aspects: the construction of the starting trial wave functions and their optimization. For the optimization of the

wave function, there are different quantities that one can minimize. In Chapter 3, we discuss which variational principle is more effective between variance and energy minimization in the context of excited states. In particular, we obtain accurate excitation energies for two prototypical molecules by minimizing the energy, but we encounter severe difficulties following the variance. By analyzing a simple model, we infer that the variance landscape has little or no barriers between its minima. For this reason, while minimizing the variance, the optimization leads to the, *a priori* unknown, global minimum of the variance, making it hard or impossible to target a specific state.

Using energy minimization, we then discuss in Chapter 3 and in more depth in Chapter 4 the possible ways of constructing the trial wave function when the aim is a balanced description of multiple states. We demonstrate how, by using a selected configuration interaction (sCI) scheme, we can build compact trial wave functions that, after being fully optimized in QMC, give vertical excitations energies in line with other accurate quantum chemistry methods. Differently than the scheme previously used (based on complete active space (CAS) calculations), with sCI, we have an automatic way of selecting important contributions to the wave function, removing any dangerous bias that could be introduced by the user's understanding of the problem. Moreover, with the sCI scheme, we obtain better accuracy with fewer parameters (less computationally demanding). Therefore, we can describe relatively large molecules (in Chapter 4 we go from 3 to 19 atoms, excluding hydrogens) that are not accessible by other accurate methods such us full CI (FCI) or approximate coupled cluster singles, doubles, and triples (CC3).

Exploring the use of QMC methods for describing excited-state relaxations. Establishing systematic protocols for accurately calculating vertical excitations opens the way for the study of excited-state relaxations. At the moment, few electronic structure methods can achieve the task; with this work, we try to asses if we can use QMC for such studies by investigating its strengths and limitations. In Chapter 5, we focus on the consequences of having a statistical error (typical of any quantity estimated with an MC integration) on the forces while performing a molecular dynamics (MD) calculation. The error in the forces creates a random walk in the velocities resulting in an increase in the total energy in time. Of course, the error would go to zero in the limit of an infinitely long MC run (infinite points for the integral), and the MD total energy would be stable. However, such long MC runs are currently too expensive, especially if we want to perform picoseconds long MD runs (as is needed to follow excited state relaxation processes). For this reason, we develop new strategies to obtain as stable as possible total energies without increasing the computational cost. In particular, we constrain the molecule's center of mass motion, removing unphysical rotations and translations, and we develop an on-the-fly fit procedure that uses the information of the past MD forces to improve the estimate of the present one. In this way, we obtain relatively stable total energies by tuning few fit parameters. To further improve on it, we also attempt to develop a Langevin-like scheme to thermalize the excess noise without corrupting the dynamical path. In Chapter 6, we examine two additional technical problems relative to QMC-driven MD calculations: the infinite estimator of the variance of the forces and the effects of the optimization

of the QMC wave function at every MD step. The first is a known problem that we solve in a standard way by introducing a guiding wave function designed to remove the infinity. Regarding the second problem, we find that a partial optimization of the wave function results in a decrease of the total energy in time, so at each MD, it is crucial to perform a strict optimization (although expensive). Combining all these findings and strategies, we show how we converge towards performing stable MD simulations; moreover, quite encouragingly, our preliminary results in the excited states align with previous findings.

To summarize, this dissertation contributes to establishing and enhancing QMC methods for the study of photo-excitations. Parallel to this work, new efficient procedures for performing wave function optimization have been proposed, and they have successfully been used for computationally challenging cases such as the excited states with a strong double excitation character. Moreover, with the rise of machine learning methods, an increasing effort is being devoted to creating wave function structures based on neural networks, and into borrowing and integrating efficient machine learning optimization algorithms into QMC. Additionally, big-scale research projects, such as the European center of excellence TREX project, are improving the efficiency and user-friendliness of the main QMC codes, enabling more users to experiment with these methods. In this very active research picture, this work has the advantage of not only contributing to the creation of robust backbones for QMC calculations in excited states, but also of exploring the new territory of molecular dynamics simulation with QMC forces. This direction is quite exciting since it allows us to move from the static description of vertical excitations toward following actual photo-excitation processes. Of course, more work is still necessary, for example in understanding how to build efficient wave functions that are suitable for molecular dynamic simulations. Furthermore, creating a more effective optimizer would reduce the cost of the optimization (currently this being the bottleneck of QMC-driven MD simulations). In summary, this dissertation helps setting up the pillars needed to extend the description to more complex scenarios. For example, together with the aforementioned advancements, a possibility is now to incorporate non-adiabatic effects, for example, by including hopping probabilities among energy surfaces.

Samenvatting en Vooruitblik

Van verticale excitaties naar relaxatie van de aangeslagen toestand: een reis met quantum Monte Carlo.

Met dit proefschrift leveren we een bijdrage aan de studie van moleculaire excitaties door een specifieke elektronische structuurmethode te beoordelen en uit te breiden: quantum Monte Carlo (QMC). Aangeslagen toestanden van moleculen spelen een centrale rol in het activeren van vele natuurlijke processen zoals het menselijk gezichtsvermogen, en in het functioneren van nieuwe technologieën zoals zonnepanelen. Theoretische studie van deze processen is een zeer actief onderzoeksgebied, maar door hun complexe aard ontbreekt het ons nog steeds aan een standaardprocedure om ze te analyseren. Enerzijds worden bij veel foto-geïnduceerde processen effecten van de omgeving op het molecuul doorslaggevend en moet er daarom rekening mee worden gehouden, bijvoorbeeld met een gemengde quantum/klassieke aanpak. Anderzijds zijn, buiten het evenwicht van de grondtoestand, de te beschrijven potentiële energieoppervlakken zeer complex. Vooral wanneer meerdere toestanden interactie met elkaar vertonen schieten de beschikbare quantumchemische methoden vaak tekort.

In dit proefschrifthouden we geen rekening met omgevingseffecten. We werken aan het verbeteren van de quantummechanische beschrijving van aangeslagen toestanden door ons te richten op QMC-methoden, een klasse van technieken voor het, op een stochastische manier, oplossen van de Schrödinger vergelijking. De laatste tijd krijgen deze technieken steeds meer belangstelling in de elektronische structuurgemeenschap dankzij de gunstige schaling met het aantal elektronen en door de eenvoud van parallellisatie. In de context van aangeslagen toestanden vestigen deze technieken zich als een goed alternatief voor andere methoden, vooral voor die complexe gevallen waarin de goedkoopste opties (zoals tijdsafhankelijke dichtheidsfunctionaaltheorie) falen in de beschrijving. Bovendien hebben recente ontwikkelingen van algoritmen in de QMC-gemeenschap het mogelijk gemaakt de beschrijving uit te breiden van kleine tot middelgrote en relatief grote (ongeveer 100 zware atomen) moleculen en nauwkeurige geometrieën te berekenen in zowel de grondtoestand als in de aangeslagen toestanden. Aangemoedigd door dit werk, onderzoeken we in dit proefschrift het gebruik van QMC verder voor de studie van moleculaire aangeslagen toestanden. In het bijzonder proberen we in de eerste helft van het manuscript (hoofdstukken 3 en 4) robuuste protocollen te ontwikkelen voor de berekening van verticale excitaties. In de tweede helft (hoofdstukken 5 en 6) verkennen we pro-

blemen die samenhangen met het gebruik van QMC-methoden om de relaxatie van aangeslagen toestanden te beschrijven. Het ontwikkelen van robuuste strategieën om verticale excitaties te bestuderen. Het gebruik van QMC methoden bij de analyse van verticale excitaties is een vrij actueel en een zich snel ontwikkelend onderzoeksgebied. In dit proefschrift leveren we hier een bijdrage aan door twee cruciale aspecten te onderzoeken: de constructie van zogenaamde *trial* golffuncties en hun optimalisatie. Voor de optimalisatie van de golffunctie zijn er verschillende grootheden die men kan minimaliseren. In de context van aangeslagen toestanden bespreken we in hoofdstuk 3 welk variatieprincipe effectiever is: variantie- of energie-minimalisatie. Nauwkeurige excitatie-energieën voor twee prototypische moleculen zijn door ons verkregen door de energie te minimaliseren, maar stuiten we op ernstige moeilijkheden bij het volgen van de variantie. Uit analyse van een eenvoudig model leiden we af dat het variantielandschap weinig of geen barrières kent tussen zijn minima. Daarom leidt de optimalisatie bij het minimaliseren van de variantie tot het, a priori onbekende, globale minimum van de variantie, waardoor het moeilijk of onmogelijk is om een specifieke toestand aan te pakken. Met behulp van energieminimalisatie bespreken we vervolgens in hoofdstuk 3 en meer diepgaand in hoofdstuk 4 de mogelijke manieren om de *trial* golffunctie te construeren wanneer het doel een evenwichtige beschrijving van meerdere toestanden is. We laten zien hoe we, met behulp van een selected configuration interaction (sCI) schema, compacte *trial*-golffuncties kunnen bouwen die, na volledig geoptimaliseerd te zijn in QMC, verticale excitaties energieën geven die in overeenstemming zijn met andere nauwkeurige quantumchemische methoden. Anders dan de eerder gebruikte procedure (die gebaseerd is op *complete active space* (CAS) berekeningen), hebben we met sCI een geautomatiseerde manier om de belangrijke bijdragen aan de golffunctie te selecteren, waardoor elke gevaarlijke vooringenomenheid van het begrip van het probleem van een gebruiker, wordt weggenomen. Bovendien verkrijgen we met het sCI-schema een grotere nauwkeurigheid met minder parameters (minder rekenintensief). Daarom kunnen we relatief grote moleculen beschrijven (in hoofdstuk 4 gaan we van 3 tot 19 zware atomen) die niet toegankelijk zijn met andere nauwkeurige methoden zoals volledige CI (FCI) of benaderde *coupled cluster singles, doubles en triples* (CC3). Onderzoek naar het gebruik van QMC-methoden voor de beschrijving van relaxaties van aangeslagen toestanden. Het opstellen van systematische protocollen voor het nauwkeurig berekenen van verticale excitaties opent de weg voor de studie van relaxaties van de aangeslagen toestand. Momenteel zijn er weinig elektronische structuurmethoden die deze taak kunnen uitvoeren; met dit werk proberen we na te gaan of we QMC voor dergelijke studies kunnen gebruiken door de sterke punten en beperkingen ervan te onderzoeken. In hoofdstuk 5 richten we ons op de gevolgen van een statistische fout (typisch voor elke grootheid die geschat wordt met een MC-integratie) op de krachten tijdens het uitvoeren van een moleculaire dynamica (MD) berekening. De fout in de krachten creëert een zgn. *random walk* in de snelheden die resulteert in een toename van de totale energie in de tijd. Uiteraard zou de fout naar nul gaan in de limiet van een oneindig lange MC-run (oneindig veel punten voor de integratie), en zou de totale MD-energie stabiel zijn. Dergelijke MC runs zijn momenteel echter te duur, vooral als we picoseconden lange MD runs

willen uitvoeren (zoals nodig is om relaxatieprocessen van de aangeslagen toestand te volgen). Daarom ontwikkelen we nieuwe strategieën om zo stabiel mogelijke totale energieën te verkrijgen zonder de rekenkosten te verhogen. Zo beperken we de beweging van het massamiddelpunt van het molecuul, waarbij we niet-fysieke rotaties en translaties verwijderen, en we ontwikkelen een *on-the-fly* fit-procedure die de informatie van de vroegere MD-krachten gebruikt om de schatting van de huidige te verbeteren. Op deze manier verkrijgen we relatief stabiele totale energieën door weinig fit-parameters in te stellen. Om dit verder te verbeteren, proberen we ook een Langevin-achtig schema te ontwikkelen om de overtollige ruis te thermiseren zonder het dynamische pad te corrumperen. In hoofdstuk 6 onderzoeken we twee bijkomende technische problemen in verband met QMC-gestuurde MD berekeningen: de zgn. *infinite estimator* van de variantie van de krachten en de effecten van de optimalisatie van de QMC golffunctie bij elke MD stap. Het eerste is een bekend probleem dat we op een standaard manier oplossen door een sturende golffunctie te introduceren die de oneindigheid wegneemt. Wat het tweede probleem betreft, vinden we dat een gedeeltelijke optimalisatie van de golffunctie resulteert in een afname van de totale energie in de tijd, dus bij elke MD is het van cruciaal belang om een strikte optimalisatie uit te voeren (hoewel duur). Door al deze bevindingen en strategieën te combineren, laten we zien hoe we convergeren naar het uitvoeren van stabiele MD simulaties; bovendien zijn onze voorlopige resultaten in de aangeslagen toestanden, heel bemoedigend, in overeenstemming met eerdere bevindingen. Samengevat draagt dit proefschrift bij aan het vaststellen en verbeteren van QMC methoden voor de studie van foto-excitaties. Parallel aan dit werk zijn nieuwe efficiënte procedures voor het uitvoeren van golffunctie-optimalisatie voorgesteld, en deze zijn met succes gebruikt voor computationeel uitdagende gevallen zoals de aangeslagen toestanden met een sterk dubbel excitatie karakter. Bovendien wordt, met de opkomst van *machine learning*, steeds meer aandacht besteed aan het creëren van golffunctiestructuren op basis van neurale netwerken, en aan het lenen en integreren van efficiënte *machine learning* optimalisatiealgoritmen in QMC. Bovendien verbeteren grootschalige onderzoeksprojecten, zoals het Europese *center of excellence TREX-project*, de efficiëntie en gebruiksvriendelijkheid van de belangrijkste QMC-codes, waardoor meer gebruikers met deze methoden kunnen experimenteren. In dit zeer actieve onderzoeksveld heeft dit werk het voordeel dat het niet alleen bijdraagt aan de totstandkoming van robuuste QMC berekeningen aan aangeslagen toestanden, maar ook dat het nieuwe terrein van moleculaire dynamica simulatie met QMC krachten wordt verkend. Deze richting is erg spannend, omdat het ons in staat stelt om van de statische beschrijving van verticale excitaties over te stappen op het volgen van werkelijke foto-excitatieprocessen. Natuurlijk is er nog meer werk nodig, bijvoorbeeld om te begrijpen hoe efficiënte golffuncties kunnen worden opgebouwd die geschikt zijn voor moleculair dynamische simulaties. Bovendien zou het maken van een effectievere optimalisator de kosten van de optimalisatie verminderen (momenteel is dit de *bottleneck* van QMC-gestuurde MD simulaties). Samengevat helpt dit proefschrift bij het opzetten van de pijlers die nodig zijn om de beschrijving uit te breiden naar complexere scenario's. Samen met de bovengenoemde verbeteringen is het nu bijvoorbeeld mogelijk om niet-adiabatische effecten op te nemen,

bijvoorbeeld door zgn. *hopping*-waarschijnlijkheden tussen energieoppervlakken op te nemen.

Acknowledgments

This manuscript is the result of four years of work in the computational chemical physics group of the University of Twente. This experience would not have been possible without the help and support of many people. As I approach the end of this journey, I want to express my gratitude to each of them.

I start by thanking my promotor Prof. Claudia Filippi for her help and constant presence in this journey. Her guidance and support throughout the whole time have been essential to me. She is a careful supervisor, a passionate researcher, and a model to follow. Her academic honesty and personal involvement with any of her students are admirable. I am grateful to her for this opportunity which has made me grow both personally and professionally. I also want to thank my unofficial co-supervisor Prof. Wim Briels, who has been patiently coming to Enschede to work with us. I appreciate having the opportunity to work closely with him and of going through the creative solutions he comes up with for any new problem. I thank him for his warm presence and for cheering my days by always being ready to discuss any kind of subject with both wisdom and curiosity.

Next, I would like to thank Dr. Anthony Scemama, who has collaborated with us for the first half of this work. Thanks for introducing us to the CIPSI code and for always being easy to reach for any doubt or difficulty with the code. Similarly, I would like to thank Dr. Saverio Moroni, whose sharp comments and considerations always give an exciting twist to the papers and become the motivation to dig deeper and gain a better understanding. I also want to thank Prof. Johannes Neugebauer, Prof. Geert-Jan Kroes, Dr. Bernd Ensing, Prof. Geert Brocks, and Prof. Wim Briels for agreeing to be part of my committee and for taking the time to read my thesis.

Of course, I could not forget to thank all the rest of the people who have created a positive and comfortable environment for me during these years. Firstly, Monika, who has made me feel welcome from the first day I arrived in the group and has guided me to my life in Enschede. She has been an invaluable help in my work and a great friend to share any activity with. I thank Emiel for sharing some parts of the dynamics project with me with his great enthusiasm. I also want to thank all the former and present group members: Vishal, Jonas, Ramon, Ravindra, Stu, Edgar, Jacopo, and Marco, for creating a supportive and relaxed atmosphere where to unload all complaints and stress that accompany the Ph.D. life. Moreover, seeing the group expand and get closer contact with the rest of the theory corridor has been great. I am grateful to Raisa, Linn, Menno, Jonathan, Geert, Paul, Max, and Kostas for making lunchtime more lively and for the nice outings. Finally, I thank Gerrit-Jan