Ab initio study of the optical properties of green fluorescent protein

Zaccheddu, M.

Chapter 2

Computational methods

2.1 Introduction

To investigate the photophysics of Green Fluorescent Protein, we will employ a variety of computational methods as there is not a single theoretical approach to date, which is capable to cover the different spatial and temporal scales which characterize this complex problem. Starting from the x-ray structure of the protein, we will build a realistic model of the protein environment surrounding the chromophore, that is, the optically active component of the protein. As the protein exhibits multiple forms corresponding to different protonations of the chromophore, the system must be properly relaxed to describe the different conformations. The electronic properties of the chromophore within its protein environment are then investigated quantum mechanically in the ground and excited states. These steps translate in a series of computations involving a hierarchy of theoretical approaches, ranging from classical molecular dynamics to correlated many-body techniques. In particular, we will present the results of the following calculations:

- Classical molecular mechanics (MM) calculations to equilibrate the protein in water solution at room temperature.
- Hybrid quantum mechanics in molecular mechanics (QM/MM) calculations based on density functional theory (DFT) to obtain an accurate description of the ground state geometry of the optically active chromophore.
- Time-dependent density functional theory (TDDFT) calculations to compute the excitation spectrum and access how the protein environment modulates the response of the chromophore to light.
- Correlated post-Hartree-Fock quantum chemical approaches to investi-

gate the role of correlation and possible shortcomings in the description of excited states within density functional theory. These calculations are also a prerequisite for the construction of the many-body wave function needed in the next step.

- Quantum Monte Carlo (QMC) calculations of the excitation spectrum. As the application of quantum Monte Carlo to excited states is rather new, further methodological developments were needed.

In this chapter, we give a short description of the theoretical methods we employ and of the relevant technical details. We begin with the quantum mechanical methods, in particular the multi-configuration self-consistent (MCSCF) approach and its state average (SA) version for the computation of excited states, the variational (VMC) and diffusion (DMC) Monte Carlo methods, and time-dependent density functional theory (TDDFT). We then describe how a quantum mechanical approach can be combined with classical molecular mechanics (QM/MM) for the hybrid treatment of a quantum site embedded in a larger classical system. Novel methodological developments will be presented in the following chapters to more clearly illustrate the context in which they are needed.

2.2 Quantum mechanical calculations

The many-electron Schrödinger equation gives an accurate description of materials at the quantum mechanical level but is an intractable $3N + 1$ dimensional partial differential equation, where the number of electrons N may be very large. In this thesis, we will consider molecular systems with typically 100-500 electrons for which we want to investigate the electronic properties of the ground and lowest excited states.

Most computational quantum mechanical studies of such large electronic systems circumvent the problem of the high dimensionality by employing simpler one-electron theories such as Kohn-Sham density-functional theory (DFT), which replaces the electron-electron interactions by an effective potential, thereby reducing the problem to a set of one-electron equations. Despite the successes of DFT in describing the electronic structure of complex molecular systems, the treatment of electronic correlation within DFT is only approximate, sometimes leading to incorrect results as we will see in the case of the excitation spectrum of the Green Fluorescent Protein. Therefore, one needs to resort to alternative approaches as the more costly wave function based methods. Here, we will not employ traditional quantum chemistry wave function methods as for instance the complete active space second or-

der perturbation theory (CASPT2) technique which is often used to treat excitations of organic molecules, but we will focus on quantum Monte Carlo (QMC) techniques, which, for ground state problems, have yielded in the past very accurate description of correlated properties also of large systems, where conventional quantum chemistry methods are extremely difficult to apply. We review here only those aspects of traditional quantum chemistry approaches which are need to understand how the wave function is constructed for quantum Monte Carlo calculations.

Let us define the notation we adopt in this thesis. As we neglect relativistic effects and work in the Born-Oppenheimer approximation, we will assume that we have a non-relativistic system of N interacting electrons described by the Hamiltonian:

$$\mathcal{H} = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 + \sum_{i=1}^{N} v_{\text{ext}}(\mathbf{r_i}) + \sum_{i<j}^{N}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}\,, \tag{2.1}$$

where we used atomic units ($\hbar = m = e = 1$). The external potential $v_{\text{ext}}(\mathbf{r})$ is given either by the bare electron-ion Coulomb potential $-Z/r$ where Z is the charge of the ion, or by a pseudopotential describing the ion plus the core electrons which have been eliminated from the calculation. We denote with $\mathbf{R}$ the $3N$ particle coordinates, and with $\mathbf{x} = (\mathbf{r}, \sigma)$ the 3 spatial and 1 spin coordinates of one electron where $\sigma = \pm 1$.

2.2.1 Traditional quantum chemistry methods

The simplest approach for the description of a system of N interacting electrons is the the Hartree-Fock (HF) method, where the ground state many-body wave function is approximated as the optimal non-interacting solution, that is a Slater determinant of single-particle spin-orbitals $\{\Phi_i\}$:

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

The optimal single-particle orbitals are determined by minimizing the expectation value E_{HF} of the interacting Hamiltonian $\mathcal{H}$ on the wave function Ψ_{HF}. If the spin-orbitals are written as the product of a spatial and a spin components, $\Phi_i(\mathbf{x}) = \phi_i(\mathbf{r})\chi_{s_i}(\sigma)$, one obtains that the spatial orbitals must

satisfy the self-consistent HF equations:

$$\left[-\frac{1}{2}\nabla^2 + v_{\text{ext}}(\mathbf{r}) + \sum_{j=1}^{N}\int \mathrm{d}\mathbf{r}'\,\frac{|\phi_j(\mathbf{r}')|^2}{|\mathbf{r}-\mathbf{r}'|}\right]\phi_i(\mathbf{r}) - \sum_{j=1}^{N}\delta_{s_i,s_j}\int \mathrm{d}\mathbf{r}'\,\frac{\phi_j^*(\mathbf{r}')\phi_i(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\phi_j(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r})\,, \tag{2.2}$$

where the Lagrange multipliers ϵ_i arise from the orthonormality constraints between the orbitals. Each orbital sees the external potential, the Hartree electrostatic component, and the non-local Hartree-Fock exchange potential. The HF potential cancels the interaction of the electron with itself, that is the self-interaction contribution coming from the the Hartree potential, and keeps the electrons of the same spin apart so that each electron has a hole around it, known as the exchange hole, containing unit positive charge.

For atoms, the HF equations can be solved directly on a grid but, for molecular systems, the orbitals are expanded as a linear combination of atomic orbitals (LCAO) centered on the nuclear positions:

$$\phi_i(\mathbf{r}) = \sum_{\mu}^{\text{nuclei}}\sum_{j} a_{ji}^{\mu}\,\eta_{j\mu}(\mathbf{r}-\mathbf{r}_\mu)\,, \tag{2.3}$$

where $\mathbf{r}_\mu$ denotes the position of a nucleus. The LCAO coefficients, a_{ji}^{μ}, are optimized to yield the lowest variational energy. In general, most quantum chemistry codes work with a Gaussian atomic basis:

$$\eta(\mathbf{r}) = x^m y^n z^k \exp\left(-\alpha r^2\right), \tag{2.4}$$

as this choice allows all integrals to be computed analytically.

The difference between the exact energy E and the HF energy is called the correlation energy, $E_{\text{corr}} = E - E_{\text{HF}}$.

Post Hartree-Fock methods

Quantum chemical post-HF approaches express the many-body wave function $\Psi(\mathbf{x}_1,\ldots,\mathbf{x}_N)$ in terms of a non-interacting basis as they rely implicitly or explicitly in writing the wave function as an expansion in determinants of single-particle orbitals. With such an expansion, the matrix elements of the Hamiltonian on the basis and the overlap of the basis functions can be readily expressed and even computed analytically if a Gaussian basis set is employed to express the single-particle orbitals.

Conceptually, we can imagine to start from the solutions of the HF equations which give us a complete set of orthonormal orbitals, comprising the N occupied orbitals and $M - N$ virtual orbitals, where M is the size of the atomic basis set. We can then proceed as in the configuration interaction (CI) approach and construct a correlated wave function as

$$\Psi_{\mathrm{CI}} = c_0 D_{\mathrm{HF}} + \sum_{ab} c_{a\to b} D^{a\to b} + \sum_{abcd} c_{ab\to cd} D^{ab\to cd} + \dots \,, \tag{2.5}$$

where $D^{a\to b}$ denotes a single excitation from the HF determinant where the occupied orbital a is substituted with the virtual orbital b. Similarly, $D^{ab\to cd}$ indicates a double excitation with the orbitals a and b substituted with the virtual orbitals c and d. A full CI expansion is obtained if one includes up to N-body excitations to all virtual orbitals, and the result should then be extrapolated to the infinite basis limit by considering larger basis sets. We can rewrite a CI expansion in more compact form as

$$\Psi_{\mathrm{CI}} = \sum_{i=1}^{K} c_i C_i \,, \tag{2.6}$$

where C_i are spin and space-adapted configuration state functions (CSF), that is, fixed linear combination of determinants with proper spin and space symmetry. By applying the variational principle, one obtains the secular equations for the coefficients c_i:

$$\sum_{j=1}^{K} \langle C_i|\mathcal{H}|C_j\rangle c_j^{(k)} = E_{\mathrm{CI}}^{(k)} \sum_{j=1}^{K} \langle C_i|C_j\rangle c_j^{(k)} \,, \tag{2.7}$$

where $\langle C_i|C_j\rangle = \delta_{ij}$ as the orbitals are orthonormal.

An advantage of the CI approach is that one obtains not only an approximation to the ground state wave function but also to the higher excited states via the coefficients $c_i^{(k)}$. In fact, a generalized variational principle applies, known as the the McDonald's theorem, which states that the approximate solutions with energies $E_{\mathrm{CI}}^{(0)} \leq E_{\mathrm{CI}}^{(1)} \leq \ldots \leq E_{\mathrm{CI}}^{(K)}$ satisfy

$$E_i \leq E_{\mathrm{CI}}^{(i)} \,, \tag{2.8}$$

where E_i are the exact energies of the eigenstates of the Hamiltonian $\mathcal{H}$. A disadvantage of a CI expansion is that a great number of determinants must be included due to the lack of explicit dependence of the wave function from inter-electron coordinates which makes difficult the description of the cusp

occurring at the electron coalescence points. Moreover, the number of determinants increases very fast with the system size, in particular exponentially with the number of electrons N. A way to limit the number of determinants is to include the most important excitations, for instance single and double (CISD), which yields a computational cost of N^6 with consequent loss of size consistency.

In the multi-configuration self consistent field (MCSF) approach, one optimizes not only the linear coefficients c_i but also the LCAO coefficients a_{ji} to minimize the total energy. A particular type of MCSCF calculation is the complete active space self-consistent (CASCF) approach, where a set of active orbitals is selected, whose occupancy is allowed to vary, while all other orbitals are fixed as either doubly occupied or unoccupied. In a CASSCF(n,m) calculation, n electrons are distributed among an active space of m orbitals and all possible resulting space- and spin-symmetry-adapted CSFs are constructed. The final CASSCF(n,m) wave function consists of a linear combination of these CSFs, like in a full CI calculation for n electrons in m orbitals, except that also the orbitals are now optimized to minimize the total energy.

When several states of the same symmetry are requested, there is a danger in optimizing the higher states that their energy is lowered enough to approach and mix with lower states, thus giving an unbalanced description of excitation energies. A well-established solution to this problem is the use of a state averaged (SA) CASSCF approach where the weighted average of the energies of the states under consideration is optimized

$$E_{\mathrm{SA}} = \sum_I w_I \frac{\langle \Psi_I | \mathcal{H} | \Psi_I \rangle}{\langle \Psi_I | \Psi_I \rangle}\,, \tag{2.9}$$

where $\sum_I w_I = 1$ and the states are kept orthogonal. The wave functions of the different states depend on their individual sets of CI coefficients using a common set of orbitals. Orthogonality is ensured via the CI coefficients and a generalized variational theorem applies. Obviously, the SA-CASSCF energy of the lowest state will be higher than the CASSCF energy obtained without SA. The most important step for a MCSCF/CASSCF calculation is the choice of the active space and, unfortunately, there is not a simple rule to select the proper orbitals. Usually, a great number of trial calculations are necessary to find out which orbitals must be included in the active space.

2.2.2 Density functional theory

When compared to conventional quantum chemistry methods, density functional theory (DFT) is particularly appealing since it does not rely on the

knowledge of the complete N-electron wave function but only of the electronic density. Density functional theory provides an expression for the ground state energy of a system of interacting electrons in an external potential as a functional of the ground state electronic density [14]. Let us assume for simplicity that the spin polarization of the system of interest is identically zero. In the Kohn-Sham formulation of density functional theory [15], the ground state density is written in terms of single-particle orbitals obeying the equations:

$$\left[-\frac{1}{2}\nabla^2 + v_{\text{eff}}\left([n]\,;\mathbf{r}\right)\right]\phi_i = \epsilon_i\phi_i, \tag{2.10}$$

where the electronic density is constructed by summing over the N lowest energy orbitals where N is the number of electrons:

$$n(\mathbf{r}) = \sum_{i=1}^{N} \left|\phi_i(\mathbf{r})\right|^2. \tag{2.11}$$

The effective Kohn-Sham potential is given by

$$v_{\text{eff}}\left([n]\,;\mathbf{r}\right) = v_{\text{ext}}(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\mathrm{d}\mathbf{r}' + v_{\text{xc}}\left([n]\,;\mathbf{r}\right) \tag{2.12}$$

$v_{\text{ext}}(\mathbf{r})$ is the external potential. The exchange-correlation potential $v_{\text{xc}}\left([n]\,;\mathbf{r}\right)$ is the functional derivative of the exchange-correlation energy $E_{\text{xc}}\left[n\right]$ that enters in the expression for the total energy of the system:

$$\begin{aligned} E &= -\frac{1}{2}\sum_{i=1}^{N}\int \phi_i\nabla^2\phi_i\,\mathrm{d}\mathbf{r} + \int n\left(\mathbf{r}\right)v_{\text{ext}}\left(\mathbf{r}\right)\,\mathrm{d}\mathbf{r} \\ &+ \frac{1}{2}\iint \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}' + E_{\text{xc}}\left[n\right]. \end{aligned} \tag{2.13}$$

Unfortunately, although the theory unlike HF is in principle exact, the energy functional contains an unknown quantity, called the exchange-correlation energy, $E_{\text{xc}}\left[n\right]$, that must be approximated in any practical implementation of the method. If the functional form of $E_{\text{xc}}\left[n\right]$, and consequently the exchange-correlation potential, were available, we could solve the N-electron problem by finding the self-consistent solution of a set of single-particle equations.

Approximate exchange-correlation functionals

Several approximate exchange-correlation functionals have been proposed in the literature, the most commonly used ones being the local density approximation (LDA), the generalized gradient approximation (GGA) and, more

recently, the hybrid functionals. The local density approximation [15] is the simplest functional:

$$E_{\mathrm{xc}}^{\mathrm{LDA}}[n] = \int \mathrm{d}\mathbf{r}\, \epsilon_{\mathrm{xc}}^{\mathrm{hom}}(n(\mathbf{r}))n(\mathbf{r}) \tag{2.14}$$

where $\epsilon_{\mathrm{xc}}^{\mathrm{hom}}(n)$ is the exchange correlation energy per electron of a uniform electron gas of density n. This functional is by construction exact for a homogeneous electron gas but has been shown to work surprisingly well also when the distribution of electrons is strongly inhomogeneous.

However, LDA does not always provide sufficiently accurate results. For example, it always overestimates the binding energy and the bond length of weak bonded molecules and solids. Therefore, a dependence of the exchange-correlation energy on the derivatives of the electronic density has been introduced in the so-called generalized gradient approximations (GGA), whose generic functional form (here restricted to second-order derivative) is

$$E_{\mathrm{xc}}^{\mathrm{GGA}}[n] = \int n(\mathbf{r})\, \epsilon_{\mathrm{xc}}^{\mathrm{GGA}}(n(\mathbf{r}), |\nabla n(\mathbf{r})|, \nabla^2 n(\mathbf{r}))\, \mathrm{d}\mathbf{r}. \tag{2.15}$$

Many different GGA's are available in the literature and, in this thesis, we will make use of the Becke-Lee-Yang-Parr (BLYP) [16] and the Perdew-Burke-Ehrennshof (PBE) [17] GGA functionals.

In recent years, hybrid functionals have become very popular in particular for chemical applications. These functionals introduce a dependence on the Kohn-Sham orbitals, and mix a portion of exact exchange from Hartree-Fock theory with the exchange and correlation GGA functional:

$$E_{\mathrm{xc}}^{\mathrm{hybrid}}[n] = E_{\mathrm{xc}}^{\mathrm{GGA}}[n] + c_{\mathrm{x}}(E_{\mathrm{x}}^{\mathrm{HF}}[n] - E_{\mathrm{x}}^{\mathrm{GGA}}[n])\,, \tag{2.16}$$

where $E_x^{\mathrm{GGA}}[n]$ and $E_{xc}^{\mathrm{GGA}}[n]$ are GGA exchange (x) and exchange-correlation (xc) energies, and $E_x^{\mathrm{HF}}[n]$ is the exact exchange which has the same form as the HF exchange energy:

$$E_{\mathrm{x}}[n] = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\delta_{s_i,s_j}\iint \frac{\phi_i^*(\mathbf{r})\phi_j^*(\mathbf{r}')\phi_j(\mathbf{r})\phi_i(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}'\,. \tag{2.17}$$

The coefficient c_x controls the amount of Hartree-Fock exchange: It is unity for Hartree-Fock, zero for pure DFT, and fractional (typically around 0.25 [18]) for hybrid functionals. This parameter is usually fitted to reproduce a set of properties as for instance atomization energies of first and second-row molecules. A widely used hybrid functional available in most DFT codes is

the three parameter B3LYP functional [19] which combines LDA and the BLYP GGA with exact exchange:

$$\begin{aligned} E_{\mathrm{xc}}^{\mathrm{B3LYP}} &= E_{\mathrm{xc}}^{\mathrm{LDA}} + a_0(E_x^{\mathrm{HF}} - E_x^{\mathrm{LDA}}) \\ &+ a_x(E_x^{\mathrm{GGA}} - E_x^{\mathrm{LDA}}) + a_c(E_c^{\mathrm{GGA}} - E_c^{\mathrm{LDA}}) \end{aligned} \tag{2.18}$$

where $a_0 = 0.20$, $a_x = 0.72$, and $a_c = 0.81$. In this thesis, we will use the hybrid functional B3LYP or PBE0. For more information about DFT, we refer the reader to Refs. [20–22].

Time-dependent density functional theory

Time-dependent density-functional theory (TDDFT) represents a rigorous formalism for the calculations of excitation energies. Similarly to ground state density functional theory, TDDFT is formally exact but relies in practice on the use of approximate exchange-correlation functionals.

The central theorem of TDDFT is the Runge-Gross theorem [23] which generalizes the Hohenberg-Kohn theorem to a time-dependent Hamiltonian, and proves the one-to-one correspondence between the external time-dependent potential $\mathrm{v_{ext}}(\mathbf{r},\mathrm{t})$ and the time-dependent electronic density, $\mathrm{n}(\mathbf{r},\mathrm{t})$. This theorem leads to construct a time-dependent Kohn-Sham scheme for a system of non-interacting electrons in an effective external time-dependent potential:

$$\left[-\frac{1}{2}\nabla^2 + v_{\mathrm{eff}}\left([n]\,;\mathbf{r},t\right)\right]\phi_i(\mathbf{r},t) = i\frac{\partial}{\partial t}\phi_i(\mathbf{r},t), \tag{2.19}$$

which yields the exact electronic density constructed from the Kohn-Sham orbitals as

$$n(\mathbf{r},t) = \sum_{i=1}^{N} \left|\phi_i(\mathbf{r},t)\right|^2 . \tag{2.20}$$

The Kohn-Sham effective potential is given by

$$v_{\mathrm{eff}}\left([n]\,;\mathbf{r},t\right) = v_{\mathrm{ext}}(\mathbf{r},t) + \int \frac{n(\mathbf{r}',t)}{|\mathbf{r}-\mathbf{r}'|}\mathrm{d}\mathbf{r}' + v_{\mathrm{xc}}\left([n]\,;\mathbf{r},t\right) , \tag{2.21}$$

where the first term in the external potential, the second term takes in account the electrostatic interaction between the electrons, and the last term is the exchange-correlation potential. It is important to stress that the time-dependent Kohn-Sham potential is not the same functional of the density

as the ground-state Kohn-Sham potential (Eq. 2.12) but equals the functional derivative of the exchange-correlation component of the action functional [23, 24].

Like in the ground-state DFT approach, the only fundamental approximation in TDDFT is the time-dependent exchange-correlation potential and the quality of the results crucially depends on the quality of this approximation. The simplest approximation is the so-called adiabatic approximation:

$$v_{\mathrm{xc}}^{\mathrm{adiab}}([n];\mathbf{r},t) = v_{\mathrm{xc}}^{\mathrm{gs}}([n];\mathbf{r})|_{n=n(\mathbf{r},t)} \tag{2.22}$$

where $v_{\mathrm{xc}}^{\mathrm{gs}}$ is some given ground-state exchange-correlation potential. The adiabatic approximation therefore assumes that the self-consistent potential is local in time and responds instantaneously and without memory to any temporal change in the charge density. As the $v_{\mathrm{xc}}^{\mathrm{gs}}$ is a ground-state property, we expect that this approximation works best for time-dependent systems whose density does not change too much from the ground-state one. By inserting the LDA or the BLYP potential (or whatever functional one prefers), we obtain what we denote as the approximate adiabatic TDDFT/LDA or TDDFT/BLYP approach.

The excitation energies can be readily obtained from a TDDFT calculation by knowing how the system responds to a small time-dependent perturbation. The key quantity is the linear density response function χ which measures the change in the density of the system due to a small perturbation in the external potential:

$$\delta n_\sigma(\mathbf{r},\omega) = \int \mathrm{d}\mathbf{r}'\chi(\mathbf{r},\mathbf{r}',\omega)\,\delta v_{\mathrm{ext}}(\mathbf{r}',\omega) \tag{2.23}$$

and which allows one to compute the dynamic polarizability and therefore access the photoabsorption cross section. Through the time-dependent Kohn-Sham scheme (Eqs. 2.19–2.21), we can rewrite the same change in the density as

$$\delta n_\sigma(\mathbf{r},\omega) = \int \mathrm{d}\mathbf{r}'\chi_{\mathrm{KS}}(\mathbf{r},\mathbf{r}',\omega)\,\delta v_{\mathrm{eff}}(\mathbf{r}',\omega)\,, \tag{2.24}$$

where χ_{KS} is the density response function of the non-interacting Kohn-Sham electrons which can be written in terms of the unperturbed time-independent Kohn-Sham orbitals. Then, using the definition of the exchange-correlation potential (Eq. 2.20), we can obtain the linear change in the potential as

$$\delta v_{\mathrm{eff}}(\mathbf{r},\omega) = \delta v_{\mathrm{ext}}(\mathbf{r},\omega) + \int \mathrm{d}\mathbf{r}' \left[\frac{1}{|\mathbf{r}-\mathbf{r}'|} + f_{\mathrm{xc}}(\mathbf{r},\mathbf{r}',\omega)\right] \delta n(\mathbf{r}',\omega)\,, \tag{2.25}$$

where $f_{\rm xc}([n];\mathbf{r},\mathbf{r}',\omega)$ is the Fourier transform of the exchange-correlation kernel:

$$f_{\rm xc}([n];\mathbf{r},\mathbf{r}',t-t') = \frac{\delta v_{\rm xc}([n];\mathbf{r},t)}{\delta n(\mathbf{r}',t')}\,. \tag{2.26}$$

Combining Eqs. 2.23–2.25, we derive a Dyson-like equation for the response function

$$\begin{aligned}\chi(\mathbf{r},\mathbf{r}',\omega) &= \chi_{\rm KS}(\mathbf{r},\mathbf{r}',\omega) \\ &+ \int\mathrm{d}\mathbf{x}\int\mathrm{d}\mathbf{x}'\,\chi(\mathbf{r},\mathbf{x},\omega)\left[\frac{1}{|\mathbf{x}-\mathbf{x}'|}+f_{\rm xc}(\mathbf{x},\mathbf{x}',\omega)\right]\chi_{\rm KS}(\mathbf{x}',\mathbf{r}',\omega)\,,\end{aligned} \tag{2.27}$$

which yields the response χ of the interacting system via a self-consistent solution if the exact exchange-correlation kernel is known. Since a full solution of this equation is numerically quite difficult, one obtains the excitation energies by knowing that the density response function, χ, has poles at the frequencies which correspond to the excitation energies of the interacting system. Similarly, the poles of the Kohn-Sham response function, $\chi_{\rm KS}$, correspond to the non-interacting excitation energies given by the difference of Kohn-Sham eigenvalues.

Through a series of algebraic manipulations, it is possible to reformulate linear-response TDDFT in terms of the so-called Casida's equations [25] where the poles of the response functions, $\Omega = E_m - E_0$, are determined as solutions of the non-Hermitian eigenvalue problem:

$$\begin{bmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{B} & \mathbf{A}\end{bmatrix}\begin{pmatrix}\vec{X}\\ \vec{Y}\end{pmatrix} = \Omega\begin{bmatrix}\mathbf{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{1}\end{bmatrix}\begin{pmatrix}\vec{X}\\ \vec{Y}\end{pmatrix}, \tag{2.28}$$

where the matrices $\mathbf{A}$ and $\mathbf{B}$ are defined as

$$\begin{aligned}A_{ia,i'a'} &= \delta_{ii'}\delta_{aa'}(\epsilon_a-\epsilon_i)+K_{ia,i'a'}\,,\\ B_{ia,i'a'} &= K_{ia,a'i'} = (ia|\frac{1}{|\mathbf{r}-\mathbf{r}'|}|a'i') + (ia|f_{\rm xc}|a'i')\,.\end{aligned} \tag{2.29}$$

The eigenvalues of these equations give the excitation energies and the eigenvectors can be used to compute the oscillator strengths.

It is important to note that TDDFT adiabatic approximation includes only dressed one-electron excitations [26]. This can be seen in the context of the Tamm-Dancoff approximation (TDA), which consists of neglecting the $\mathbf{B}$ matrices to obtain $\mathbf{A}\vec{X} = \Omega\vec{X}$. The number of possible solutions to this equation is the dimensionality of $\mathbf{A}$ which is the number of single excitations. In fact, the linear-response time-dependent Hartree-Fock TDA (exchange-only density functional theory) is simply the well-known configuration interaction

singles (CIS) method. The computational cost of TDDFT scales approximately like $O(N^3)$, so TDDFT represents a very appealing theoretical approach to compute the excitation energies of large molecular systems. While often reasonably accurate, the main difficulties encountered in conventional TDDFT include the underestimation of the ionization threshold [27], the underestimation of charge transfer excitations [28–30], and the lack of explicit two- and higher-electron excitations [26, 31]. These shortcomings may be fatal in describing the excitations of biomolecular systems as these systems are often characterized by charge transfer in the excited states and may display multi-configurational character.

2.2.3 Quantum Monte Carlo methods

The variational (VMC) and diffusion (DMC) Monte Carlo methods we present in this Section share with conventional quantum chemistry methods that they are wave function based approaches. However, differently from quantum chemistry approaches, they attempt to solve the Schrödinger equation stochastically, and have consequently significantly more freedom in the choice of the functional form of the many-body correlated wave function. Moreover, both approaches have a more favorable scaling with the number of electrons, that is, N^4 when compared to N^6 of CISD or N^7 of the coupled cluster single and double with perturbative triples approach. Therefore, even though they are more costly than DFT which only scales as N^3, they are significantly faster than conventional highly-correlated quantum chemistry methods, and can be applied to larger systems and to solids to provide accurate answers in situations when DFT is shown to be inadequate.

Variational Monte Carlo

The variational Monte Carlo method is the simplest QMC approach and uses Monte Carlo techniques to evaluate the expectation value of an operator for a given wave function. Let us assume we are given the many-body trial wave function Ψ_{T} and that we are interested in computing the expectation value of the Hamiltonian $\mathcal{H}$:

$$E_V = \frac{\int \Psi_{\mathrm{T}}^*(\mathbf{R})\mathcal{H}\Psi_{\mathrm{T}}(\mathbf{R})d\mathbf{R}}{\int \Psi_{\mathrm{T}}^*(\mathbf{R})\Psi_{\mathrm{T}}(\mathbf{R})\mathrm{d}\mathbf{R}} \tag{2.30}$$

This expectation value can be rewritten as

$$E_V = \frac{\int |\Psi_{\mathrm{T}}(\mathbf{R})|^2[\Psi_{\mathrm{T}}(\mathbf{R})^{-1}\mathcal{H}\Psi_{\mathrm{T}}(\mathbf{R})]d\mathbf{R}}{\int |\Psi_{\mathrm{T}}(\mathbf{R})|^2 d\mathbf{R}} = \int \rho(\mathbf{R})E_L(\mathbf{R})\mathrm{d}\mathbf{R}\,, \tag{2.31}$$

where

$$\rho(\mathbf{R}) = \frac{|\Psi_{\mathrm{T}}(\mathbf{R})|^2}{\int |\Psi_{\mathrm{T}}(\mathbf{R})|^2 \mathrm{d}\mathbf{R}}\,, \tag{2.32}$$

and the local energy is defined as

$$E_L(\mathbf{R}) = \Psi_{\mathrm{T}}(\mathbf{R})^{-1}\mathcal{H}\Psi_{\mathrm{T}}(\mathbf{R})\,, \tag{2.33}$$

Since $\rho(\mathbf{R})$ is a positive quantity and integrates to 1, we can interpret it as a probability distribution and use Monte Carlo techniques to sample a set of configurations $\{\mathbf{R}_m\}$ distributed according to $\rho(\mathbf{R})$. The expectation value can then be estimated as an average of the local energy $E_L(\mathbf{R})$ evaluated on these configurations:

$$E_V \approx \frac{1}{M}\sum_{m=1}^{M} E_L(\mathbf{R}_m) \tag{2.34}$$

Note that in this derivation, we can substitute the Hamiltonian $\mathcal{H}$ with any operator $\mathcal{O}$ diagonal in space representation.

For a realistic system of electrons, the square of the many-body wave function is a complicated probability distribution in a high-dimensional space, of which we do not usually know how to compute the normalization. Therefore, we cannot use direct sampling techniques but we employ the Metropolis algorithm to generate a sequence of configurations $\{\mathbf{R}_m\}$ distributed according to $\rho(\mathbf{R})$. The Metropolis algorithm is a general method to sample an arbitrary probability distribution without knowing its normalization, and is an application of a Markov chain. In a Markov chain, one changes the state of the system randomly from an initial state $\mathbf{R}_{\mathrm{i}}$ to a final state $\mathbf{R}_{\mathrm{f}}$ according to the stochastic transition matrix $M(\mathbf{R}_{\mathrm{f}}|\mathbf{R}_{\mathrm{i}})$ which satisfies

$$M(\mathbf{R}_{\mathrm{f}}|\mathbf{R}_{\mathrm{i}}) \geq 0 \quad \text{and} \quad \sum_{\mathrm{f}} M(\mathbf{R}_{\mathrm{f}}|\mathbf{R}_{\mathrm{i}}) = 1\,. \tag{2.35}$$

To sample the desired distribution $\rho(\mathbf{R})$, one evolves the the system by repeated application of a Markov matrix M which satisfies the *stationarity condition*

$$\sum_{i} M(\mathbf{R}_{\mathrm{f}}|\mathbf{R}_{\mathrm{i}})\ \rho(\mathbf{R}_i) = \rho(\mathbf{R}_f)\,,$$

for any state $\mathbf{R}_{\mathrm{f}}$. The stationarity condition condition tells us that if we start from the desired distribution ρ, we will continue to sample ρ. Moreover, if the stochastic matrix M is ergodic, this condition ensures that any initial distribution will evolve to ρ under repeated applications of M. Therefore, ρ

is the right eigenvector of M with eigenvalue 1 and it is also the dominant eigenvector.

In practice, one imposes the more stringent *detailed balance* condition

$$M(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})\ \rho(\mathbf{R}_i) = M(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})\ \rho(\mathbf{R}_f) \tag{2.36}$$

which is a sufficient but not necessary condition to satisfy the stationarity condition as can be easily seen by summing both sides of the equation over $\mathbf{R}_\mathrm{i}$ and using Eq. 2.35. The transition M is then rewritten as the product of a proposal matrix T and the acceptance A:

$$M(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i}) = A(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})\ T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})\,, \tag{2.37}$$

where M and T are stochastic matrices but A is not. The detailed balance condition finally becomes

$$\frac{A(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})}{A(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})} = \frac{T(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})\ \rho(\mathbf{R}_f)}{T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})\ \rho(\mathbf{R}_i)}. \tag{2.38}$$

For a given T, the choice originally made by Metropolis *et al.* [32] for the acceptance is

$$A(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i}) = \min\left\{1, \frac{T(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})\ \rho(\mathbf{R}_f)}{T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})\ \rho(\mathbf{R}_i)}\right\}\,, \tag{2.39}$$

and is the one which maximizes the acceptance. In choosing the proposal matrix T, we observe that the Metropolis algorithm generates points which are sequentially correlated so that the effective number of independent observations in a Monte Carlo run of M steps is M/T_corr, where T_corr is the autocorrelation time of the observable of interest. Therefore, to achieve a fast evolution and reduce T_corr, the optimal T should yield a high acceptance and at the same time allow large proposed moves. The choice of T will of course be limited by the fact that we need to be able to sample T directly. We use here the algorithm described in Ref. [33], which uses a non-symmetrical T and which we properly modified to deal with pseudopotentials.

In short, the generalized Metropolis algorithm will consist of the following steps:

1. Choose the distribution $\rho(\mathbf{R})$ and the transition probability $T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$.
2. Initialize the configuration $\mathbf{R}_\mathrm{i}$.
3. Advance the configuration from $\mathbf{R}_\mathrm{i}$ to $\mathbf{R}'$:
   a) Sample $\mathbf{R}'$ from $T(\mathbf{R}'|\mathbf{R}_\mathrm{i})$.

b) Calculate the ratio

$$q = \frac{T(\mathbf{R}_\mathrm{i}|\mathbf{R}')}{T(\mathbf{R}'|\mathbf{R}_\mathrm{i})} \frac{\rho(\mathbf{R}')}{\rho(\mathbf{R}_\mathrm{i})} \,. \tag{2.40}$$

c) Accept or reject: If $q > 1$ or $q > r$ where r is a uniformly distributed random number in (0,1), set the new configuration $\mathbf{R}_\mathrm{f} = \mathbf{R}'$. Otherwise, set $\mathbf{R}_\mathrm{f} = \mathbf{R}_\mathrm{i}$.

4. Throw away the first κ configurations corresponding to the equilibration time.
5. Collect the averages and block them to obtain the error bars.

Two final comments on the Metropolis algorithm. First, the distribution $\rho(\mathbf{R})$ does not have to be normalized since only ratios enter in the acceptance. Therefore, it is possible to sample the square of complex wave functions (Eq. 2.32) whose normalization we do not know. Second, if $M_1, M_2, \ldots, M_n$ are matrices which satisfy the stationarity condition, the matrix $M = \prod_{i=1}^{n} M_i$ also satisfies the stationarity condition. Consequently, particles can be moved one at the time, a necessary feature as the system size grows since the size of the move would need to be decreased to have a reasonable acceptance of a move of all particles.

Many-body wave functions used in quantum Monte Carlo

The use of VMC to compute the expectation values of quantum mechanical operators allows great freedom in the choice of the trial wave function which on the other hand determines the accuracy as well as the efficiency of the calculation. Therefore, the form of wave function should yield accurate results while being compact and easy to evaluate.

The ingredients entering in the wave function most commonly used in quantum Monte Carlo can be understood by inspecting the advantages and limitations of traditional quantum chemistry approaches. Methods such as configuration interaction (CI) expand the many body wave function in a linear combination of Slater determinants of single-particle spin-orbitals. This form allows the evaluation of the high-dimensional integrals in all expectation values but the convergence of the expansion is very slow, in part because of the difficulty in describing the cusps which occur as two electrons approach each other. Quantum Monte Carlo uses a much more compact representation of the wave function which is usually given by a sum of few determinants (tens and not millions like in a CI calculation) multiplied by a component which can exactly impose the cusps at the inter-particle coalescence points.

Slater-Jastrow wave function

The trial wave functions used in our quantum Monte Carlo calculations are of the Jastrow-Slater form, thus they are a product between a sum of determinants of single particle orbitals, and a Jastrow correlation factor:

$$\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N) = \mathcal{J}(\mathbf{r}_1,\ldots,\mathbf{r}_N)\sum_k d_k D_k^{\uparrow}(\mathbf{r}_1,\ldots,\mathbf{r}_{N_\uparrow}) D_k^{\downarrow}(\mathbf{r}_{N_\uparrow+1},\ldots,\mathbf{r}_N)\,, \tag{2.41}$$

where $D_k^{\uparrow}$ and $D_k^{\downarrow}$ are Slater determinants of single particle orbitals for the up and down spin electrons, respectively. The orbitals are a linear combination of Slater functions centered on the atoms for all-electron calculations while are expanded on a Gaussian basis when pseudopotentials are employed. The Jastrow correlation function is a positive function of the interparticle distances and explicitly depends on the electron-electron separations.

The wave function is here written in spin-assigned form where the dependence on the spin variables $\{\sigma_i\}$ has disappeared and the wave function appears to no longer be fully antisymmetric. To obtain such an expression, we start from a full wave function $\Psi(\mathbf{x}_1,\ldots,\mathbf{x}_N)$ depending on both spatial and spin coordinates, and expand it on its spin components. For a a system of N electrons with $N = N_\uparrow + N_\downarrow$ and $S_z = (N_\uparrow - N_\downarrow)/2$, we introduce a spin function ζ_1

$$\zeta_1(\sigma_1,\ldots,\sigma_N) = \chi_\uparrow(\sigma_1)\ldots\chi_\uparrow(\sigma_{N_\uparrow})\chi_\downarrow(\sigma_{N_\uparrow+1})\ldots\chi_\downarrow(\sigma_N)\,. \tag{2.42}$$

and construct a set of $K = N!/(N_\uparrow!N_\downarrow!)$ distinct spin functions ζ_i by permuting the indices in ζ_1. Since the spin functions ζ_i form a complete orthonormal set in spin space, we can decompose the wave function Ψ in terms of its spin components as

$$\Psi(\mathbf{x}_1,\ldots,\mathbf{x}_N) = \sum_{i=1}^{K} F_i(\mathbf{r}_1,\ldots,\mathbf{r}_N)\zeta_i(\sigma_1,\ldots,\sigma_N)\,. \tag{2.43}$$

As Ψ is antisymmetric under the interchange of particle indices, each function F_i is antisymmetric under the interchange of like-spin electrons and all F_i are the same except for a relabelling of the particle indices and a change in sign for odd permutations. Therefore, the wave function can be rewritten as:

$$\Psi(\mathbf{x}_1,\ldots,\mathbf{x}_N) = \mathcal{A}\left\{F_1(\mathbf{r}_1,\ldots,\mathbf{r}_N)\zeta_1(\sigma_1,\ldots,\sigma_N)\right\} \tag{2.44}$$

It is easy to show using orthonormality of the functions ζ_i that the expectation value of an operator $\mathcal{O}$ which is spin-independent is the same if we use the fully antisymmetric wave function Ψ or just one spatial function, say F_1:

$$\langle\Psi|\mathcal{O}|\Psi\rangle = \langle F_1|\mathcal{O}|F_1\rangle\,. \tag{2.45}$$

Since it is more convenient to use the function F_1 than the full wave function Ψ, in quantum Monte Carlo, we always work with spin-assigned wave functions. To obtain F_1, we simply assign the spin-variables of the particles as

$$\begin{array}{cccccccc} \text{Particle} & 1 & 2 & \dots & N_\uparrow & N_{\uparrow+1} & \dots & N \\ \sigma & 1 & 1 & \dots & 1 & -1 & \dots & -1 \end{array}$$

so that $F_1(\mathbf{r}_1, \dots, \mathbf{r}_N) = \Psi(\mathbf{r}_1, 1, \dots, \mathbf{r}_{N_\uparrow}, 1, \mathbf{r}_{N_\uparrow+1}, -1, \dots, \mathbf{r}_N, -1)$.

Finally, the Jastrow-Slater spin-assigned wave function obtained by imposing $\sigma = +1$ for first $N_\uparrow$ particles and $\sigma = -1$ for the others is given by

$$\begin{aligned} \Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) &= F_1(\mathbf{r}_1, \dots, \mathbf{r}_N) \\ &= \mathcal{J} \sum_k d_k \, D_k^\uparrow(\mathbf{r}_1, \dots, \mathbf{r}_{N_\uparrow}) D_k^\downarrow(\mathbf{r}_{N_\uparrow+1}, \dots, \mathbf{r}_N) \end{aligned} \quad (2.46)$$

where $\mathcal{J} = \mathcal{J}(\mathbf{r}_1, \dots, \mathbf{r}_N)$ is the Jastrow factor.

The Jastrow factor is generally chosen to be a positive function of the interparticle distances and therefore does not affect the sign of the wave function which is solely determined by the determinantal component. At a large interparticle distances, it plays no role since it becomes constant, which we achieve by using scaled variables as shown below. The Jastrow factor is of fundamental importance in describing correlation at short and intermediate distances. In particular, the electron-electron cusp conditions are imposed through the Jastrow factor: At the electron-electron coalescence points, the potential energy diverges at infinity, so the kinetic energy must have an opposite divergence to the potential to keep the local energy finite. It is possible to ensure this cancellation if the trial wave function satisfies a set of cusp conditions and displays a proper discontinuity of the derivatives at the coalescence points.

The form of the Jastrow factor which we use depends on electron-electron and the electron-nucleus distances and describes electron-electron, electron-nucleus and electron-electron-nucleus correlations:

$$\begin{aligned} \mathcal{J}(\mathbf{r}_1, \dots, \mathbf{r}_N) = \prod_{\alpha,i} \exp\left\{A(r_{i\alpha})\right\} \times \prod_{i<j} \exp\left\{B(r_{ij})\right\} \times \\ \times \prod_{\alpha, i<j} \exp\left\{C(r_{i\alpha}, r_{j\alpha}, r_{ij})\right\} . \end{aligned} \quad (2.47)$$

The electron-nucleus terms A should be included if the determinantal part is constructed using orbitals obtained from a DFT or a HF calculation and not reoptimized after the inclusion of the electron-electron Jastrow factor.

As the electron-electron term alters the single-particle density by reducing/increasing it in high/low density regions, the resulting density will in general be worse that the original DFT or HF density which can be repristinated by the inclusion of the electron-nucleus terms. The electron-electron term B is introduced to impose the electron-electron cusp conditions and to keep the electrons apart as the electron-electron interaction is repulsive. Finally, the electron-electron-nucleus terms C can in principle exactly describe a two-electron atom or ion in an S state. Higher body correlations are clearly less important as, due to the exclusion principle, it is rare for three or more electrons to be close since at least two electrons must necessarily have the same spin.

To keep the Jastrow factor finite at large distances, we use scaled variables $\bar{r} = (1 - e^{-\kappa r})/\kappa$ for the A and B terms, and $\bar{r} = e^{-\kappa r}$ for the C terms. The particular form we employ in this work is:

$$\begin{aligned} A(r_{i\alpha}) &= \frac{a_1 \bar{r}_{i\alpha}}{1 + a_2 \bar{r}_{i\alpha}} + \sum_{p=2}^{N^a_{\rm ord}} a_{p+1} \bar{r}^p_{i\alpha} \\ B(r_{ij}) &= \frac{b_1 \bar{r}_{ij}}{1 + b_2 \bar{r}_{ij}} + \sum_{p=2}^{N^b_{\rm ord}} b_{p+1} \bar{r}^p_{ij} \\ C(r_{i\alpha}, r_{j\alpha}, r_{ij}) &= \sum_{p=2}^{N^c_{\rm ord}} \sum_{0}^{k=p-1} \sum_{0}^{l=l_{\rm max}} c_{mkl} \bar{r}^k_{ij} (\bar{r}^l_{i\alpha} + \bar{r}^l_{j\alpha}) (\bar{r}_{i\alpha} \bar{r}_{j\alpha})^m, \end{aligned} \quad (2.48)$$

where $m = (p - k - l)/2$, and $l_{\rm max}$ is $p-k$ if $k \neq 0$ and $p-k-2$ if $k = 0$. Only terms for which $m = (p - k - l)/2$ is an integer are included. The a and c coefficients are different for different atom types. The only spin dependence is in b_1 which is is used to satisfy the electron-electron cusp conditions: $b_1 = 1/2$ for antiparallel spin, and $b_1 = 1/4$ for parallel electrons.

Diffusion Monte Carlo

VMC is a very useful tool which allows us to explore which type of correlation is relevant in the system under study with a relatively small amount of computer time. For instance, we can investigate how the complexity of the trial wave function influences the description of the state of interest. However, its major drawback is that the result will uniquely depend on the the quality of the trial wave function which cannot be constructed in an automatic way and whose functional form must be chosen for each particular problem. Therefore, we would like to have a way to remove (at least in part) the bias introduced by the wave function.

Projector Monte Carlo is a more powerful method than VMC and removes (at least in part) the bias of the trial wave function from the results. It is a stochastic implementation of the power method for finding the dominant eigenstate of a matrix or integral kernel. In a projector Monte Carlo method, one uses an operator that inverts the spectrum of $\mathcal{H}$ to project out the ground state of $\mathcal{H}$ from a given trial state. Different operators have been used as projectors but, here, for simplicity, we only discuss diffusion Monte Carlo (DMC) which we use in our calculations.

In DMC, we use as projection operator $\exp[-\tau(\mathcal{H} - E_{\mathrm{T}})]$, and given an initial trial wave function $\Psi^{(0)}$, we repeatedly apply the projection operator to obtain the sequence of wave functions:

$$\Psi^{(n)} = e^{-\tau(\mathcal{H}-E_{\mathrm{T}})}\Psi^{(n-1)} \,. \tag{2.49}$$

If we expand the initial wave function $\Psi^{(0)}$ on the eigenstates Ψ_i with energies E_i of $\mathcal{H}$, we obtain for $\Psi^{(n)}$:

$$\Psi^{(n)} = \sum_i \Psi_i \, \langle \Psi^{(0)} | \Psi_i \rangle e^{-n\tau(E_i - E_{\mathrm{T}})} \,, \tag{2.50}$$

where $\langle \Psi^{(0)} | \Psi_i \rangle$ is the overlap between $\Psi^{(0)}$ and the eigenstate Ψ_i. Since the coefficients of the excited states die off exponentially fast relative to the coefficient of the ground state, we obtain

$$\lim_{n\to\infty} \Psi^{(n)} = \Psi_0 \langle \Psi^{(0)} | \Psi_0 \rangle e^{-n\tau(E_0 - E_{\mathrm{T}})} \,. \tag{2.51}$$

Therefore, if we choose the trial energy $E_{\mathrm{T}} \approx E_0$ to keep the over all normalization of $\Psi^{(n)}$ fixed, the projection yields the ground state Ψ_0 of the Hamiltonian. Note that the starting wave function must have a non-zero overlap with the ground state one.

To see how to perform the projection, let us first rewrite Eq. 2.49 in integral from and obtain

$$\Psi^{(n)}(\mathbf{R}', t+\tau) = \int \mathrm{d}\mathbf{R} \, G(\mathbf{R}', \mathbf{R}, \tau) \Psi^{(n-1)}(\mathbf{R}, t) \,, \tag{2.52}$$

where the coordinate Green's function is defined as

$$G(\mathbf{R}', \mathbf{R}, \tau) = \langle \mathbf{R}' | e^{-\tau(\mathcal{H}-E_{\mathrm{T}})} | \mathbf{R} \rangle \,. \tag{2.53}$$

If we can sample the trial wave function and the Green's function in Eq. 2.52, we can perform this high-dimensional integral by Monte Carlo integration. For fermions, since the wave function must be antisymmetric, it cannot be

interpreted as a probability distribution. Therefore, for the moment, we will assume that we are dealing with bosons which are characterized by a positive ground state wave function.

Using the Trotter-Suzuki formula it is possible to show that the approximate Green's function for small τ is given by:

$$\langle \mathbf{R}'|e^{-\mathcal{H}\tau}|\mathbf{R}\rangle \approx \frac{1}{(2\pi\tau)^{3N/2}} \exp\left[-\frac{(\mathbf{R}'-\mathbf{R})^2}{2\tau}\right] \exp\left[-\tau\,\mathcal{V}(\mathbf{R})\right] . \quad (2.54)$$

Therefore, the iteration in Eq. 2.52 can be interpreted as a Markov process with the difference that the Green's function is not normalized and we obtain a branching random walk: the first factor in the short-time Green's function is the Green's function for diffusion while the second term multiplies the distribution by a positive scalar. Since the short-time expression of the Green's function is only valid in the limit of τ approaching zero, in practice, DMC calculations must be performed for different values of τ and the result extrapolated for τ which goes to zero.

The use of this Green's function would however yield a highly inefficient and unstable algorithm since the potential can vary significantly from configuration to configuration or also be unbounded like the Coulomb potential. For example, the electron-nucleus potential diverges to minus infinity as the two particles approach each other, and the branching factor will give raise to an unlimited number of walkers. Even if the potential is bounded, the approach becomes inefficient with increasing size of the system since the branching factor also grows with the number of particles.

These difficulties can be overcame by using *importance sampling* which was originally proposed by Kalos [34] for Green's function Monte Carlo and extended by Ceperley and Alder [35] to DMC. We start from Eq. 2.52, multiply each side by a trial wave function Ψ and define the probability distribution $f^{(n)}(\mathbf{R}) = \Psi(\mathbf{R})\Psi^{(n)}(\mathbf{R})$ which satisfies

$$f^{(n)}(\mathbf{R}', t+\tau) = \int \mathrm{d}\mathbf{R}\, \tilde{G}(\mathbf{R}', \mathbf{R}, \tau) f^{(n-1)}(\mathbf{R}, t)\,, \quad (2.55)$$

where the importance sampled Green's function is given by

$$\tilde{G}(\mathbf{R}', \mathbf{R}, \tau) = \Psi(\mathbf{R}')\langle \mathbf{R}'|e^{-\tau(\mathcal{H}-E_\mathrm{T})}|\mathbf{R}\rangle/\Psi(\mathbf{R})\,. \quad (2.56)$$

It is possible to show that resulting drift-diffusion-branching short-time Green's function is given by

$$\tilde{G}(\mathbf{R}', \mathbf{R}, \tau) = (2\pi\tau)^{3N/2} \exp\left[-\frac{(\mathbf{R}'-\mathbf{R}-\tau\mathbf{V}(\mathbf{R}))^2}{2\tau}\right] \times \\ \times \exp\left\{-\tau\left[(E_\mathrm{L}(\mathbf{R}) + E_\mathrm{L}(\mathbf{R}'))/2 - E_\mathrm{T}\right]\right\} + O(\tau^2)\,. \quad (2.57)$$

where the quantum velocity is defined as

$$\mathbf{V}(\mathbf{R}) = \frac{\nabla\Psi(\mathbf{R})}{\Psi(\mathbf{R})}\,. \tag{2.58}$$

There are two important new features of $\tilde{G}(\mathbf{R}', \mathbf{R}, \tau)$. First, the quantum velocity $\mathbf{V}(\mathbf{R})$ pushes the walkers to regions where $\Psi(\mathbf{R})$ is large. In addition, the local energy $E_{\mathrm{L}}(\mathbf{R})$ instead of the potential $\mathcal{V}(\mathbf{R})$ appears in the branching factor. Since the local energy becomes constant and equal to the eigenvalue as the trial wave function approaches the exact eigenstate, we expect that, for a good trial wave function, the fluctuations in the branching factor will be significantly smaller. In particular, imposing the cusp conditions on the wave function will remove the instabilities coming from the singular Coulomb potential.

The DMC algorithm will now be:

1. A set of M_0 configurations is sampled from $|\Psi(\mathbf{R})|^2$ using the Metropolis algorithm. This is the zero-th *generation* and the number of configurations is the *population* of the zero-th generation.
2. The walkers are advanced as $\mathbf{R}' = \mathbf{R} + \xi + \tau\mathbf{V}(\mathbf{R})$ where ξ is a normally distributed $3N$ dimensional random vector, and the last term is the drift.
3. For each walker, compute the factor

$$p = \exp\left\{-\tau[(E_{\mathrm{L}}(\mathbf{R}) + E_{\mathrm{L}}(\mathbf{R}'))/2 - E_{\mathrm{T}}]\right\}\,. \tag{2.59}$$

   Branch the walker by treating p as the probability to survive at the next step: if $p < 1$, the walker survives with probability p while, if $p > 1$, the walker continues and new walkers with the same coordinates are created with probability $p - 1$. This is achieved by creating a number of copies of the current walker equal to the integer part of $p + \eta$ where η is a random number between (0,1).
4. The trial energy E_{T} is adjusted to keep the population stable around the target population M_0.

In Ref. [36], the reader can find a thourough description of several improvements one can bring to the simple algorithm outlined above.

So far, we have assumed that the wave function is positive everywhere and we have not yet addressed the problem posed by the fact that electrons are fermions and that the trial wave function must be antisymmetric. Unfortunately, straightforward generalizations of the DMC algorithm to handle

both signs of the wave functions, even if formally correct, lead to the fermion sign problem: The bosonic component grows at the expenses of the fermionic one and the antisymmetric signal is lost in the noise. To avoid this problem, we can simply forbid moves in which the sign of the trial wave function changes and the walker crosses the nodes which are defined as the set of points where the trial wave function is zero. This procedure is known as the *fixed-node approximation*. Forbidding node crossing is equivalent to finding the solution of the evolution equation with the boundary condition that it has the same nodes as the trial wave function. The Schrödinger equation is therefore solved exactly inside the nodal regions but not at the nodes where the solution will have a discontinuity of the derivatives. The fixed-node solution will be exact only if the nodes of the trial wave function are exact. In general, the fixed-node energy will be an upper bound to the exact energy, in particular the best upper bound consistent with the boundary conditions given by the nodes of the trial wave function.

Wave function optimization

The quality of the trial wave function controls the statistical efficiency of the VMC and DMC algorithms and determines the final accuracy of the results. The ability of optimizing the parameters of the trial wave function is crucial for the success of quantum Monte Carlo methods. For the optimization of the parameters in the trial wave function of a system in its ground state, we use the optimization method within energy minimization recently proposed by Umrigar, Filippi, and Sorella [37], which we briefly describe below.

The Jastrow-Slater wave function depends on a set of parameters p:

$$\Psi(p, \mathbf{R}) = \mathcal{J}(\alpha, \mathbf{R}) \sum_{i=1}^{N_{CSF}} c_i C_i(\eta, \mathbf{R}) \tag{2.60}$$

where the parameters p are given by parameters α in the Jastrow factors, the linear parameters c_i in from of the CSFs, and the LCAO coefficients η which enter in the single-particle orbitals. An optimal set of linear coefficients is readily obtained by solving the generalized eigenvalue problem

$$\sum_{j=1}^{N_{\mathrm{CSF}}} H_{ij} c_j = E_I \sum_{j=1}^{N_{\mathrm{CSF}}} S_{ij} c_j \,. \tag{2.61}$$

The Hamiltonian and overlap matrix elements are estimated by a finite-sample average in variational Monte Carlo as

$$H_{ij} = \left\langle \frac{\mathcal{J}C_i}{\Psi} \frac{\mathcal{H}\mathcal{J}C_j}{\Psi} \right\rangle_{\Psi^2} , \quad S_{ij} = \left\langle \frac{\mathcal{J}C_i}{\Psi} \frac{\mathcal{J}C_j}{\Psi} \right\rangle_{\Psi^2} \tag{2.62}$$

where the statistical average is over the Monte Carlo configurations sampled from Ψ^2. Importantly, the use of the non-symmetric estimator of the Hamiltonian matrix of Eq. (2.62) yields a strong zero-variance principle [38] and results in a particularly efficient approach.

To find the optimal linear (c) as well as non-linear (α and η) parameters, we linearize the wave function around the current set of parameters p, and consider the changes in Ψ given by $\Psi_k = (\partial\Psi/\partial p_k)$, which can be made orthogonal to Ψ as

$$\bar{\Psi}_k = \Psi_k - \left\langle \frac{\Psi_k}{\Psi} \right\rangle_{\Psi^2} \Psi \,. \tag{2.63}$$

We now work in the semi-orthogonal basis of the functions $\{\bar{\Psi}_0, \bar{\Psi}_k\}$ where $\bar{\Psi}_0 = \Psi$, and find the variations Δp_i in the parameters as the lowest eigenvalue solution of the generalized eigenvalue problem

$$H_{ij}\Delta p_j = E\, S_{ij}\Delta p_j \tag{2.64}$$

where $\Delta p_0 = 1$. When the parameter values are far from optimal, the new parameters $p_i + \Delta p_i$ can be worse than the old ones. Therefore, to ensure convergence, a positive constant a_{diag} is added to the diagonal of the Hamiltonian matrix (apart the first element):

$$H_{ij} = H_{ij} + a_{diag}\delta_{ij}(1 - \delta_{i0}) \tag{2.65}$$

which is adjusted at each optimization step.

Wave function optimization and excited states

As we are not only interested in ground state problems, we want to be able to optimize the parameters of the multiple (ground and excited) states described by the wave functions:

$$\Psi_I = \sum_{i=1}^{N_{\mathrm{CSF}}} c_i^I \mathcal{J} C_i \,. \tag{2.66}$$

which share the same Jastrow factor and orbitals but different linear coefficients. To this end, we follow a state-average (SA) approach to determine a set of orbitals and a Jastrow factor which give a comparably good description of the states under considerations while preserving orthogonality among the states. We alternate the optimization of the linear coefficients as outlined above to a micro-iteration in which the optimized quantity with respect to

the orbital and Jastrow variations is the weighted average of the energies of the states under consideration:

$$E_{\mathrm{SA}} = \sum_{I \in \mathrm{A}} w_I \frac{\langle \Psi_I | \mathcal{H} | \Psi_I \rangle}{\langle \Psi_I | \Psi_I \rangle}\,, \tag{2.67}$$

where the weights w_I are fixed and $\sum_I w_I = 1$. Therefore, at convergence, the averaged energy E_{SA} is stationary with respect to all parameter variations subject to the orthogonality constraint while the individual state energies E_i are stationary with respect to variations of the linear coefficients but not with respect to variations of the orbital or Jastrow parameters. In this approach, the wave functions are kept orthogonal and a generalized variational theorem applies.

To improve the orbital and Jastrow parameters at each SA micro-iteration step, we extend the linear optimization approach of Ref. [37] we described above to the SA optimization of multiple states. Under a common variation in an orbital or Jastrow parameter p_i, the changes in the states Ψ_I are given by $\Psi_k^I = (\partial \Psi^I / \partial p_k)$ and can be made orthogonal to the corresponding state as

$$\bar{\Psi}_k^I = \Psi_k^I - \left\langle \frac{\Psi^I}{\Psi_g} \frac{\Psi_k^I}{\Psi_g} \right\rangle_{\Psi_g^2} \Psi^I\,. \tag{2.68}$$

To linearize the minimization with respect to the non-linear parameters, we work in the semi-orthogonal basis of the functions $\{\bar{\Psi}_0^I, \bar{\Psi}_k^I\}$ where $\bar{\Psi}_0^I = \Psi^I$, and find the variations Δp_i in the parameters as the lowest eigenvalue solution of the generalized eigenvalue problem

$$H_{ij}^{\mathrm{SA}} \Delta p_j = E\, S_{ij}^{\mathrm{SA}} \Delta p_j \tag{2.69}$$

where $\Delta p_0 = 1$. The SA Hamiltonian matrix is computed as

$$H_{ij}^{\mathrm{SA}} = \sum_{I \in \mathrm{A}} w_I \left\langle \frac{\bar{\Psi}_i^I}{\Psi_g} \frac{H \bar{\Psi}_j^I}{\Psi_g} \right\rangle_{\Psi_g^2}\,, \tag{2.70}$$

and an analogous definition holds for the SA overlap matrix elements. The matrix elements for all states are computed in a single variational Monte Carlo run with guiding wave function Ψ_g. At convergence and for the optimal linear coefficients, the minimal energy E_{SA} (Eq. 2.67) is obtained: if the iterative scheme converges, the matrix elements H_{i0}^{SA} are zero and, consequently, the derivatives of E_{SA} with respect to the parameter p_i are zero as they equal H_{i0}^{SA}.

In summary, one iteration of excited state optimization consists of the following steps: *i*) Sample the quantities needed for the optimization of the linear coefficients with the appropriate guiding wave function; *ii*) diagonalize the matrix (Eq. 2.61) to obtain the optimal linear coefficients for the states under consideration; *iii*) sample for all states the quantities needed in the linear equations (Eq. 2.69) and obtain the parameters Δp_i; *iv*) construct a set of improved orbitals and Jastrow parameters as $p_i \rightarrow p_i + \Delta p_i$. As in the optimization of a ground state wave function, when the non-linear parameters are far from the optimal values, the optimization may need to be stabilized by shifting all diagonal elements except the first one as $H_{ij}^{\mathrm{SA}} \rightarrow H_{ij}^{\mathrm{SA}} + a_{\mathrm{adiag}}\delta_{ij}(1-\delta_{i0})$.

The use of pseudopotentials

While the QMC methods can be extended to large systems containing many electrons, the computational effort increases dramatically with the atomic number Z as the scaling is approximately Z^6, rendering all-electron calculations quickly intractable. The problem is caused by the core electrons which yield large energies and large fluctuations of the energy. The most common way to overcome this difficulty is to replace the core electrons by pseudopotentials, an approximation which is usually rather good as the core is chemically inert. An electron-nucleus pseudopotential is usually non-local and the most commonly used form is a potential which is local in the radial coordinate and non-local in the angular part as

$$\langle \mathbf{r}|v^{\mathrm{NL}}|\mathbf{r}'\rangle = \sum_{l=0}^{l_{max}} v^l(r)\,\delta(r-r') \sum_{m=-l}^{l} \mathrm{Y}_{lm}(\Omega)\mathrm{Y}_{lm}^*(\Omega')\,, \tag{2.71}$$

where l_{max} the maximum angular momentum considered, and the function v^l is radial and vanishes outside a core radius r_c. The non-local potential acting on the trial wave functions gives

$$\begin{aligned} & \langle \mathbf{R}|\mathcal{V}^{\mathrm{NL}}|\Psi\rangle \\ = & \sum_{i=1}^{N}\sum_{l=0}^{l_{max}} v^l(r_i) \sum_{m=-l}^{l} \mathrm{Y}_{lm}(\Omega_i) \int \mathrm{d}\Omega_i'\, \mathrm{Y}_{lm}^*(\Omega_i')\, \Psi(\mathbf{r}_1,\ldots,\mathbf{r}_i',\ldots,\mathbf{r}_N)\,, \end{aligned} \tag{2.72}$$

where the integral is over a sphere of radius $r_i' = r_i$ centered on the pseudoatom. This angular integration poses no particular problem in a VMC calculation and is done by a numerical quadrature on a regular polyhedron defined by a set of vertices whose number will depend on the value of l_{max} [39].

The use of nonlocal pseudopotential in DMC is more problematic since the Green's function (Eq. 2.53) is no longer positive. A possible way to circumvent this problem is to introduce the so-called *locality approximation* and define a new effective core potential by localizing the non-local potential on the trial wave function [40] as

$$\mathcal{V}_{\text{eff}}(\mathbf{R}) = \frac{1}{\Psi(\mathbf{R})}\langle \mathbf{R}|\mathcal{V}^{\text{NL}}|\Psi\rangle \,. \tag{2.73}$$

This new effective potential is explicitly many-body but is local and can be easily incorporated in a DMC algorithm. However, the potential depends now on the trial wave function, and the DMC energy computed with the mixed estimator is no longer necessarily variational and depends on the quality of the trial wave function. As the trial wave function approaches the fixed-node solution obtained without the locality approximation, the DMC energy converges to the correct fixed-node energy quadratically fast in the error on the trial wave function. A different approach to handle non-local pseudopotential which is variational and improves the accuracy upon the DMC approach with the locality approximation was recently proposed in Ref. [41].

Quantum Monte Carlo and excited states

Even though the DMC method was designed to project the ground state of a given Hamiltonian, it can also be used to study excited states. A straightforward way of obtain information on excited states is to construct a trial wave function which is a good approximation of the wave function of the state of interest. Naturally, this wave function can be used within VMC and, if the state is the lowest state of a one-dimensional representation of the point group of the molecules, DMC will yield a solution which is variational.

If the excited state is not the lowest state in its symmetry, we can still use DMC in the fixed-node approximation as the nodal constraints will prevent the collapse to the ground state solution. However, we may not be variational. It is only guaranteed that, if the nodal surface of the trial wave function is the same as the one of an exact state, fixed-node DMC yields the exact solution. Therefore, the role of the trial wave function is even more important when studying excited states within DMC since it not only imposes fermionic antisymmetry but also selects the state of interest.

2.3 Quantum mechanics in molecular mechanics techniques

Photochemical processes like absorption or fluorescence in biological systems are quite difficult to study because light-induced transitions require a quantum mechanical treatment of the system. However, as a biological system is very large and comprises thousands of atoms, the whole system cannot be treated at quantum level with the theoretical approaches available today. Fortunately, the primary process of photoabsorption often involves only few residues of the protein and is localized in a small spatial region, while the rest of the protein acts in this phase as spectator. Therefore, one can study the system by partitioning it into a quantum subsystem which is chemically active and small enough to be computationally treatable, and a larger environment which is assumed to be chemically inert and is simulated by less expensive classical molecular mechanics methods. These hybrid approaches are called quantum mechanics in molecular mechanics (QM/MM) approaches. In Fig. 2.1, we show how the QM/MM partitioning is applied in the simulation of the Green Fluorescent Protein.

The classical molecular mechanics calculations rely on empirical force fields to approximately describe the interactions of the classical particles. The most general form of these inter-particle potentials takes into account inter-molecular dispersion interactions between neutral atoms and electrostatic interactions between charged atoms, and intra-molecular interactions to describe the rotational, vibrational and torsional degrees of freedom of the molecule. To define a force field, one needs to specify the form of the potential and the values of the parameters whose number is very large as the same atom type may have different properties depending on which atom it is bonded to. As bonds cannot be created or broken in such a model, the bonding structure must be established at the beginning of the simulation.

Different QM/MM techniques are available in the literature, which differ for instance in how the interface between the MM and QM parts is treated and or in how the protein environment is simulated. In our work, we adopt the approach by Röthlisberger and coworkers [42] as implemented in the code CPMD [46]. The non-bonded interactions between the MM and the QM parts are modelled as

$$H_{\mathrm{NB}} = \sum_{I \in MM} q_I \int \mathrm{d}\mathbf{r}\, \frac{\rho(\mathbf{r})}{|\mathbf{r} - \mathbf{r}_I|} + \sum_{I \in MM, J \in QM} v_{\mathrm{vdW}}(r_{IJ}) \tag{2.74}$$

where $\rho(\mathbf{r})$ is the density of the electrons and the nuclei of the QM system, q_I are the MM partial charges at positions $\mathbf{r}_I$, and the van der Waals interactions

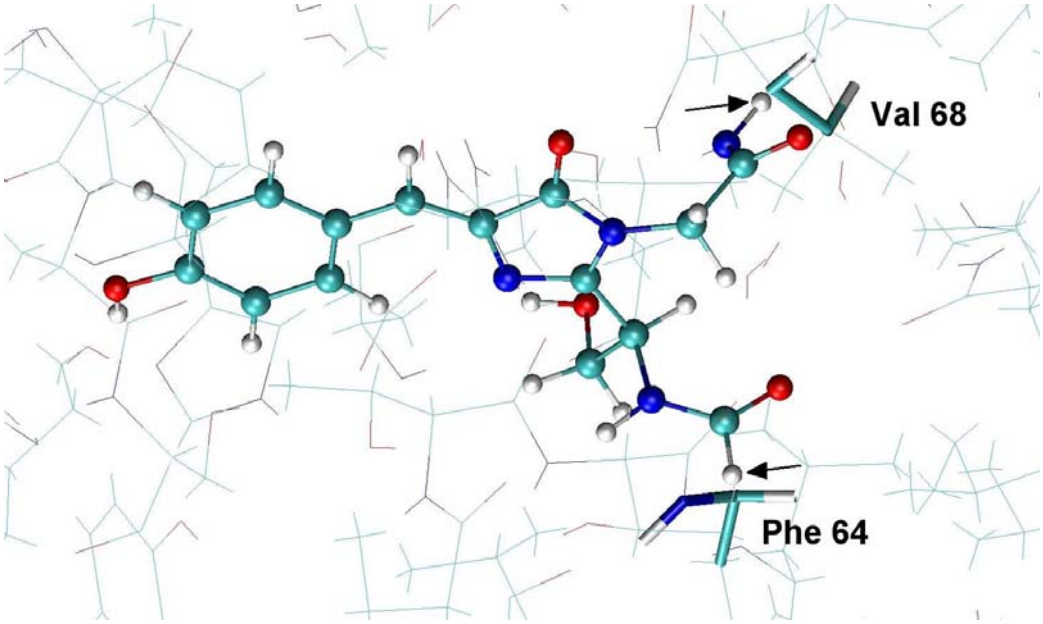

Figure 2.1: QM/MM partitioning in the simulation of Green Fluorescent Protein. The optically active part of the system is treated with a QM method while the rest of the protein is the MM part. The chromophore model is showed with a ball and sticks representation, while the MM part with solid lines. The two arrows indicate the two hydrogen-link atom between the QM and the MM. The MM part closest to the QM atoms is represented with cylinders.

between QM and MM atoms are described by the classical force field v_{vdW}.

The use of this expression is however problematic if the QM system is close to some positively charged MM atoms since the electrons will be attracted by the MM charges and the density overpolarized. This so-called spill-out effect is non physical and is particularly severe when a plane-wave basis is used as in the CPMD code. To avoid this problem in the CPMD implementation of QM/MM, a screening term is introduced for the point charges which are in proximity of the QM system. The electrostatic interaction of electrons with the close MM atoms (NN) in the non-bonded Hamiltonian is rewritten as

$$H_{\mathrm{NB}}^{\mathrm{el}} = \sum_{I \in \mathrm{NN}} q_I \int \mathrm{d}\mathbf{r}\, \rho^{\mathrm{el}}(\mathbf{r}) v_I(|\mathbf{r} - \mathbf{r}_I|) \tag{2.75}$$

where

$$v_I(r) = \frac{r_{cI}^n - r^n}{r_{cI}^{n+1} - r^{n+1}} \tag{2.76}$$

with r_{cI} the covalent radius of the atom type I and n=4. We refer the reader to Ref. [42] for a thorough discussion of the technical tricks used to reduce the large computational effort involved in computing the first term of Eq. 2.74 within plane-waves.

Finally, it is important to specify how to treat bonded interactions between the QM and MM parts, which arise when the QM/MM boundary cuts through a chemical bond of a molecule. In this case, we adopt the solution to "cap" the broken bond with a link atom, in particular a hydrogen atom. The QM system sees the hydrogen atom while the MM atoms do not.

2.4 Computational details

We employ different codes for the various steps of the calculations. For completeness, we conclude this chapter with a brief overview of the programs used, specifying in which step they were employed and briefly describing the technical details.

The **Amber Molecular Dynamics Package** [43] is a set of computational tools widely used to simulate bio-molecular systems. In particular, the module *Xleap* is used to add the missing hydrogens to the starting X-ray structure, to center the protein in a box of water solution (option *solvatebox*), to add the counterions to the water (option *addIons*) and finally to construct starting from the coordinates of the system the *Amber force field* which is then used in the Amber module *Sander* to perform classic molecular dynamics simulations. the Amber package to equilibrate the protein before starting the QM/MM calculations. The Amber force field 2003 is used to parameterize the MM part [44,45]. In all MM simulations, a cutoff of 8 Å is used for the non-bonded van der Walls interactions.

The **CPMD** [46] code is used to perform the QM/MM calculation within density functional theory to describe the QM system while the Gromos [47] force field is used for the MM part. The QM/MM interface was developed by Röthlisberger and coworkers [48,49], and, since it makes use of the MM Gromos libraries [47], it is necessary to convert the force field from the Amber to the Gromos format. The CPMD code is a plane-wave/pseudopotential code particularly designed for ab-initio molecular dynamics. The finite QM part is treated within a supercell approach using a sufficiently large periodic cell to avoid interactions between neighboring images. Moreover, we employ the isolated system module in CPMD which allows studying an isolated molecule or complex within periodic boundary conditions. We perform all calculations using a 70 Ry plane-wave cutoff and the Troullier-Martins pseudopotentials [50]. The Poisson equations are solved with the Tucker-

man's method [51]. All ground state calculations are carried out using the PBE [52] functional.

The **Gaussian03** [53] code is a quantum-chemistry code we use to perform ground state DFT and linear-response TDDFT calculations. It uses Gaussian basis sets, and a wide range of exchange-correlation functionals is available. We employ this code to perform geometry optimization of several molecular models of the chromophore in vacuum and to compute their TDDFT spectra. We also calculate the TDDFT spectra in the presence of the protein environment using the geometries obtained in the QM/MM approach with the CPMD code. The electrostatic effect of the protein environment is taken into account by using point charges with the coordinates obtained in the QM/MM calculation and the values of the charges as in the Amber force field. It is not possible to screen the charges with Gaussian03 but, as the basis is Gaussian, we do not expect significant spill-out problems as for a plane-wave basis.

The Amsterdam Density Functional **ADF** [54] code is a software package for first-principles electronic structure DFT calculations, using Slater functions for the construction of the orbitals. We employ this code to perform some TDDFT calculations with the SAOP [55] and LB94 [56] exchange-correlation potentials.

The **GAMESS** [57] code is a general ab-initio quantum chemistry package which we mostly use to generate the starting QMC wave functions either through a DFT or a SA-CASSCF calculation. This code employs Gaussian basis sets. The electrostatic effect of the protein environment can be introduced through the use of screened point charges described by the potential

$$v(r) = \frac{q}{r}\left(1 - A\, e^{-Br}\right) \tag{2.77}$$

where q is the value of the charge, and A and B are free parameters. We choose $A = 1$ to remove the divergence at the origin, and adjust the value of the B to reproduce the CPMD potential outside the maximum. We find that the choice $B = 1/(1.0828 \times r_c)^2$ where r_c is the CPMD core radius gives similar potentials in the valence region.

Finally, the code **CHAMP** is used for all the quantum Monte Carlo calculations. It can perform VMC and DMC calculations, and optimize the wave function parameters by energy minimization.