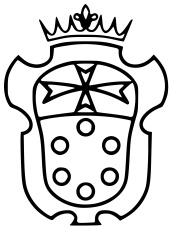


SCUOLA
NORMALE
SUPERIORE

Scuola Normale Superiore

Classe di Scienze

Tesi di Perfezionamento in Methods and Models for Molecular Sciences

XXXIV CICLO

Accurate Multiscale Approaches for Embedded Systems

Supervisors:

Prof. Chiara Cappelli

Dr. Tommaso Giovannini

Gioia Marrazzini

Academic Year: 2021/2022

To my husband Samuele

.

Acknowledgements

Vorrei innanzitutto ringraziare la Prof.ssa Chiara Cappelli, che in questi anni ha supervisionato il mio operato, prendendosi cura non solo dell'aspetto lavorativo, ma soprattutto di quello umano. A Lei è rivolta la mia stima, il mio affetto e la mia gratitudine.
Grazie a Tommaso, per la presenza continua, l'aiuto indispensabile, le risate, i pettegolezzi e per essere stato un relatore TOP.
Grazie ai colleghi dell'EmbedLab, in particolare a Luca, prezioso compagno d'avventura.
Grazie agli Amici, per il sostegno e l'affetto costante.
Grazie a mamma Giovanna, babbo Carlo e Leo, senza di loro non sarei "Gioia di nome e di fatto".
Infine, grazie a Samuele, la mia famiglia, colui che ogni giorno supporta ogni mia scelta e sopporta ogni mio cambiamento di umore... "per almeno sessant'anni".

Contents

Papers Attached

1. Calculation of Linear and Non-linear Electric Response Properties of Systems in Aqueous Solution: A Polarizable Quantum/Classical Approach with Quantum Repulsion Effects
   G. Marrazzini, T. Giovannini, F- Egidi, and C. Cappelli, *J. Chem. Theory Comput.*, **2020**, 6993-7004

2. Multilevel Density Functional Theory
   G. Marrazzini, T. Giovannini, M. Scavino, C. Cappelli and H. Koch, *J. Chem. Theory Comput.*, **2021**, 791–803

3. Integrated Multiscale Multilevel Approach to Open Shell Molecular Systems
   T. Giovannini, **G. Marrazzini**, M. Scavino, H. Koch and C. Cappelli, *submitted*

CHAPTER 1

Introduction

The aim of this Thesis is the development and application of theoretical models to describe molecular systems interacting with a solvating environment.[1–60] From the theoretical point of view, a full Quantum Mechanical (QM) description of this kind of systems would be particularly challenging, because a large number of degrees of freedom would need to be taken into account. Luckily, in most cases, a QM description of the whole system is not necessary, because the environment modifies but not determines the investigated property, which is instead connected to the molecular target. This is the fundamental idea behind the so-called focused models.[61–63] Within this formalism, the total system is partitioned into two portions: the target fragment, which is described at high level of theory (usually QM), and the environment (solvent), which is treated by means of a less sophisticated description.

The idea is thus very general, and encompasses a wide variety of methods.

What concretely distinguishes one focused model from another, is the particular way in which the environment interacts with the target fragment, and vice versa.[62] Such interaction can be limited to electrostatics (including polarization) or may retain the most quantum forces.[8]

The most widespread focused models are the QM/classical approaches,[1,2,61,64–66] in which the molecular fragment is described at the QM level, whereas the environment is treated by means of classical mechanics, either at the continuum or at the atomistic level. In particular, when an

atomistic description is exploited, strong specific interactions, such as Hydrogen Bonding, can be accurately modeled, thus overcoming the well-known limits of the continuum description. Atomistic Molecular Mechanics (MM) force fields are used for the classical part, resulting in the so-called QM/MM methods.[1,8] In most QM/MM approaches, the QM and MM parts interact at an electrostatic level, by including mutual polarization between the two regions in polarizable embedding approaches. Therefore, purely quantum forces, such as Pauli repulsion and dispersion, are neglected, although they may play a significant role in many systems/processes. Quantum interactions may be described by exploiting quantum-embedding approaches,[67–81] which are usually based on the partitioning of the density of the system into two parts: an active part, which leads the property, and an inactive one, which perturbs the active moiety, in a focused model fashion. Thus, target-environment interactions are retained at the QM level.
In this Thesis, different embedding methods, ranging from QM/MM to quantum embedding are developed and applied to different properties: polarizability, first hyperpolarizability and hyperfine coupling constants. In this chapter, the theoretical foundations of the approaches are presented and discussed, and a theoretical overview of the studied properties is reported.

1.1. Classical Embedding Methods

QM/classical approaches are characterized by a reduced computational scaling with respect to system size as compared to a full QM treatment of the whole system, which quickly becomes unfeasible from the computational point of view.[82–84] In fact, the computational cost associated to these approaches is almost the same as the corresponding QM calculation of the isolated molecule. Thus, QM/classical approaches enable efficient calculations of all kinds of chemical systems where a division into a target region (treated using quantum mechanics) and its environment (treated classically) is possible. One example of such a system is the case of a solute in a solvent, which is usually partitioned into two layers.
QM/classical models can be divided into main classes, depending on the way the classical portion is modelled. The classical part can be described in terms of a continuum, polarizable dielectric (the most popular approach being the

Polarizable Continuum Model – PCM[61,63–66,85–103]). Alternatively, the atomistic nature of the environment can be retained, overcoming the common limitations of continuum approaches in which specific solute-solvent interactions cannot be described.

1.1.1. QM/Molecular Mechanics Models

QM/MM methods[1,2,8,9] were first developed by Warshel & Levitt, who received the Nobel Prize in 2013, together with Karplus.[1] One of the crucial differences with respect to continuum approaches is that within the classical MM portion a fully atomistic description is retained. This automatically implies that specific QM-MM interactions, such as Hydrogen Bonds, can be adequately modelled. However, such a feature is not gained for free: in fact, whereas QM/PCM[61,64–66] implicitly includes the statistical average of the possible configurations of the environment, QM/MM approaches need an explicit sampling of the phase space. Such a sampling is usually carried out by firstly performing a Molecular Dynamics (MD) simulation, and then extracting some uncorrelated snapshots, whose number can vary depending on the property under consideration, until convergence is reached.[50,54,104–113] Therefore, the computational cost associated with a QM/MM simulation intrinsically increases. However, in this way, the dynamical aspects of embedding phenomena can be taken into account, thus providing a physically consistent picture.

The interaction between the QM and the MM portions is usually limited to the electrostatics. Three possible formulations have been proposed in the literature:

Mechanical embedding treats the interactions between the QM and the MM part at the MM level. The method requires an accurate set of parameters such as atom-centered point charges.

Electrostatic embedding, the QM Hamiltonian is modified by including the interaction between the potential generated by the MM fixed charges and the QM electron density:

$$H^{\text{int}}_{\text{QM/MM}} = \sum_{i}^{\text{EE charges}} q_i V_i(\mathbf{r}_i), \tag{1.1}$$

where the sum runs over the MM charges q_i and $V_i(\mathbf{r}_i)$ is the QM potential calculated at charge positions $\mathbf{r}_i$. The quality of electrostatic embedding results crucially depends on a proper choice of for MM fixed charges representing the MM electron density. The polarization of the QM fragment due to the MM charge distribution is automatically taken into account.

Polarizable Embedding differently from QM/EE, mutual polarization between QM and MM portions is taken into account. In particular, the density of the QM region is modified by the presence of the MM portion, which in turn is polarized by the QM density, in a self-consistent manner. This introduces and recovers one of the most relevant aspects of PCM, namely the mutual polarization between the QM and the classical part, by introducing a non-linear contribution in the QM Hamiltonian. Clearly, polarizable embeddings are the most sophisticated QM/MM approaches, and in recent years have become the golden standard for embedded systems.

1.2. Focus on MM Polarizable Embedding

Different polarizable embeddings have been proposed previously in the literature based on different polarizable FFs, namely Drude model,[55,114–120] Induced dipoles,[6,121] AMOEBA,[122,123] Fluctuating Charge (FQ)[50–52,60,104–111,124–128] and FQ Fluctuating dipoles (FQFμ).[128–133]

In Drude model,[55,114–120] the MM polarization is taken into account by considering an induced electric dipole at each MM site, which is defined in terms of a couple of charges, of the same magnitude but opposite sign, linked by a harmonic spring. The first charge is located at the nucleus of the MM atom, whereas the second is mobile, so that polarization arises from the competition between the forces acting on the charges, i.e. the harmonic potential and the electrostatic interactions with the remaining environment.[134]

In the Induced dipole formalism, polarization effects are treated by including induced dipoles on each MM atom. Therefore, each atom is endowed with an isotropic atomic polarizability, from which the induced dipoles are originating

as a response to the external environment. The induced dipole formalism is the basis of different polarizable QM/MM approaches, which mainly differs for the treatment of the purely electrostatic part of the force field. In QM/ID,[33,35,113,135–140] QM/MMPol[18,31,32] and QM/Discrete Reaction Field (DRF),[141–146] fixed charges are also considered, whereas in QM/Polarizable Embedding (PE)[56–58] and QM/AMOEBA[54,112,147–151] also other terms in the multipolar expansion are included.

1.2.1. Fluctuating Charges

In the Fluctuating Charges (FQ) force field[152–154] each atom of the MM portion is endowed with an electric charge (q). Such charges are not fixed but can vary (fluctuate) in accordance with the Electronegativity Equalization Principle (EEP),[155,156] which provides a minimum principle based on the consideration that at equilibrium the electronegativity of each atom have the same value.[157] The FQ force field[50–52,104–111,124–126] can be rigorously defined in conceptual density functional theory (CDFT),[158–163] because it is based on atomic electronegativity (χ) and chemical hardness (η), which are the first and the second energy derivatives with respect to the charge, respectively. By exploiting EEP, the FQs can be defined as the constrained minimum of the following functional expression (F):

$$F(\mathbf{q},\lambda) = E_0 + \sum_i \Big[\chi_i q_i + \eta_i q_i^2 + \sum_{j>i} J_{ij} q_i q_j\Big] + \lambda\Big(\sum_i q_i - Q_{\text{tot}}\Big), \qquad (1.2)$$

where i runs over the number of atoms, $\mathbf{J}$ is the interaction kernel between the FQs and Q_{tot} is the total charge. The constraint, imposed by Lagrangian multiplier λ, preserves the total charge.

Generally, the formalism adopted is based on the assumption that if it is present more than one molecule, each molecule is constrained to assume a fixed total charge (Q_α) that have to sum to the total charge of the whole system (Q_{tot}). In this way, the electronegativity of each atom in the same molecule will be the same but will have in general different values among

different molecules. The functional that has to be minimized became:[164]

$$\begin{aligned} F(\mathbf{q},\lambda) &= \sum_{\alpha,i} q_{\alpha,i}\chi_{\alpha,i} + \frac{1}{2}\sum_{\alpha,i}\sum_{\beta,j} q_{\alpha,i}J_{\alpha i,\beta j}q_{\beta,j} + \sum_{\alpha}\lambda_{\alpha}\sum_{i}\Big(q_{\alpha,i} - Q_{\alpha}\Big) \\ &= \mathbf{q}^{\dagger}\chi + \frac{1}{2}\mathbf{q}^{\dagger}\mathbf{J}\mathbf{q} + \lambda^{\dagger}\mathbf{q} \end{aligned} \tag{1.3}$$

where α and β run over the different molecules. Differentiating eq.1.3 with respect to charges and Lagrange multipliers, the constrained minimum conditions are obtained:

$$\begin{cases} \sum_{\beta,j} J_{\alpha i,\beta j}q_{\beta j} + \lambda_{\alpha} = -\chi_{\alpha i} \\ \sum_{i} q_{\alpha_i} = Q_{\alpha} \end{cases} \tag{1.4}$$

By switching to a matrix-vector formalism:

$$\begin{pmatrix} \mathbf{J} & \mathbf{1}_{\lambda} \\ \mathbf{1}_{\lambda}^{\dagger} & \mathbf{0} \end{pmatrix}\begin{pmatrix} \mathbf{q} \\ \boldsymbol{\lambda} \end{pmatrix} = \begin{pmatrix} -\boldsymbol{\chi} \\ \mathbf{Q} \end{pmatrix}$$

The linear system then reads:

$$\mathbf{M}\mathbf{q}_{\lambda} = -\mathbf{C}_{Q}, \tag{1.5}$$

where $\mathbf{q}_{\lambda}$ is a vector containing charges and Lagrange multipliers, whereas $\mathbf{C}_{Q}$ is a vector containing total charge constraints and electronegativity of each atom.

The FQ force field can be coupled to a QM description in a QM/MM framework.[50] At the Hartree-Fock level, the global energy functional reads as:

$$\begin{aligned} F(\mathbf{D},\mathbf{q},\lambda) = \mathrm{Tr}\mathbf{h}\mathbf{D} &+ \frac{1}{2}\mathrm{Tr}\mathbf{D}\mathbf{G}(\mathbf{D}) \\ &+ \mathbf{q}^{\dagger}\chi + \frac{1}{2}\mathbf{q}^{\dagger}\mathbf{J}\mathbf{q} + \lambda^{\dagger}\mathbf{q} + \mathbf{q}^{\dagger}\mathbf{V}(\mathbf{D}), \end{aligned} \tag{1.6}$$

where $\mathbf{h}$ and $\mathbf{G}$ are the common one- and two-electron matrices and $\mathbf{D}$ is the QM density matrix. $\mathbf{q}^{\dagger}\mathbf{V}(\mathbf{D})$ represents the electrostatic interaction between the charges and the electric potential due to the QM density. To obtain the FQs, we differentiate the functional in Eq. 1.6 with respect to the charges and the langrangian multipliers, yielding to the following set of linear equations:

$$\mathbf{M}\mathbf{q}_{\lambda} = -\mathbf{C}_{Q} - \mathbf{V}(\mathbf{D}), \tag{1.7}$$

Differently from Eq. 1.4, an additional term, modeling the interaction with the QM density ($\mathbf{V}(\mathbf{D})$) is present. The QM Fock matrix ($F_{\mu\nu}$ in the AO basis $\{\chi_\mu\}$), which is obtained by differentiating the energy functional in Eq. 1.6 with respect to the density matrix, is also modified by the electrostatic interaction with the FQ charges:

$$\mathbf{F}_{\mu\nu} = \frac{\partial F}{\partial D_{\mu\nu}} = h_{\mu\nu} + G_{\mu\nu}(\mathbf{D}) + \sum_i q_i V_{\mu\nu,i} \tag{1.8}$$

Differently from non-polarizable QM/MM, in which the QM/MM contribution to the Fock matrix is fixed because the charges are constant, and it can be computed once at the beginning of the Self Consistent Field (SCF) procedure, in QM/FQ, the charges depend on the QM density, and thus the QM/MM contribution to the Fock matrix has to be computed at each SCF step. In this way, the mutual polarization between the QM and FQ parts is taken into account.

1.2.2. Fluctuating Charges and Dipoles

A pragmatical extension of FQ model has been proposed in the last few years: the Fluctuating Charges and Dipoles (FQFμ) approach.[128–133] In this force field, atomic fluctuating dipoles (μ) are included in the MM portion in addition to the fluctuating charges (q). In this way, out of plane polarization, which contributes to model the anisotropy of the electrostatic interaction, is automatically taken into account. Charge values are defined by the same charge equilibration as FQ, but their values depend also on the interaction with dipoles. The peculiarity of FQFμ with respect to all the aforementioned polarizable MM methods, stands in the fact that both the FQs and Fμs vary as a response to both the external electric potential and electric field.
In FQFμ, charges (q) and dipoles (μ) are described by Gaussian distribution functions ($\rho_{q_i}(\mathbf{r}), \rho_{\mu_i}(\mathbf{r})$):

$$\begin{aligned}
\rho_{q_i}(\mathbf{r}) &= \frac{q_i}{\pi^{3/2} R_{q_i}^3} \exp\left(- \frac{|\mathbf{r} - \mathbf{r}_i|^2}{R_{q_i}^2} \right), \\
\rho_{\mu_i}(\mathbf{r}) &= \frac{|\mu_i|}{\pi^{3/2} R_{\mu_i}^3} \hat{\mathbf{n}}_i \cdot \nabla \left[\exp\left(- \frac{|\mathbf{r} - \mathbf{r}_i|^2}{R_{\mu_i}^2} \right) \right],
\end{aligned} \tag{1.9}$$

where R_{q_i} and R_{μ_i} are the widths of the corresponded Gaussian distribution functions centered at $\mathbf{r}_i$, and $\hat{\mathbf{n}}_i$ is the versor that indicates the dipole direction. The total energy E associated with a distribution of charges and dipoles is equal to:

$$\begin{aligned} E(\mathbf{q},\mu) = &\sum_i q_i\chi_i + \frac{1}{2}\sum_i q_i\eta_i q_i + \frac{1}{2}\sum_i\sum_{j\neq i} q_i T_{ij}^{qq} q_j \\ &+ \sum_i\sum_{j\neq i} q_i \mathbf{T}_{ij}^{q\mu}\mu_j + \frac{1}{2}\sum_i\sum_{j\neq i}\mu_i^\dagger \mathbf{T}_{ij}^{\mu\mu}\mu_j - \frac{1}{2}\sum_i \mu_i^\dagger \alpha_i^{-1}\mu_i, \end{aligned} \tag{1.10}$$

where χ is the atomic electronegativity, η the chemical hardness, and α is the atomic polarizability. $\mathbf{T}$ terms represent the charge-charge (T_{ij}^{qq}), charge-dipole ($\mathbf{T}_{ij}^{q\mu}$) and dipole-dipole ($\mathbf{T}_{ij}^{\mu\mu}$) interaction kernels. For the gaussian distributions defined in Eqs. 1.9, they can be written as:

$$\begin{aligned} \mathrm{T}_{ii}^{qq} &= \frac{1}{|\mathbf{r}_{ij|}}\,\mathrm{erf}\left(\frac{|\mathbf{r}_{ij}|}{R_{q_i-q_j}}\right), \\ \mathbf{T}_{ij}^{q\mu} &= \nabla_{r_j}\mathrm{T}_{ij}^{qq} = -\frac{\mathbf{r}_{ij}}{|\mathbf{r}_{ij}|^3}\left[\mathrm{erf}\left(\frac{|\mathbf{r}_{ij}|}{R_{q_i-\mu_j}}\right) - \frac{2|\mathbf{r}_{ij}|}{\sqrt{\pi}R_{q_i-\mu_j}}\exp\left(\frac{-|\mathbf{r}_{ij}|}{R_{q_i-\mu_j}}\right)^2\right], \\ \mathbf{T}_{ij}^{\mu\mu} &= -\nabla_{r_j}\mathbf{T}_{ij}^{q\mu} \\ &= -\frac{3\mathbf{r}_{ij}\times\mathbf{r}_{ij} - |\mathbf{r}_{ij}|^2\mathbf{1}}{|\mathbf{r}_{ij}|^5}\left[\mathrm{erf}\left(\frac{|\mathbf{r}_{ij}|}{R_{\mu_i-\mu_j}}\right) - \frac{2|\mathbf{r}_{ij}|}{\sqrt{\pi}R_{\mu_i-\mu_j}}\exp\left(-\frac{\mathbf{r}_{ij}}{R_{\mu_i-\mu_j}}\right)^2\right] \\ &+ \frac{4}{\sqrt{\pi}R_{\mu_i-\mu_j}^3}\frac{\mathbf{r}_{ij}\times\mathbf{r}_{ij}}{|\mathbf{r}_{ij}|^2}\exp\left(-\frac{|\mathbf{r}_{ij}}{R_{\mu_i-\mu_j}}\right)^2, \end{aligned} \tag{1.11}$$

where $R_{(q/\mu)_i-(q/\mu)_j} = \sqrt{R_{(q/\mu)_i}^2 + R_{(q/\mu)_j}^2}$. The charge-dipole and the dipole-dipole interaction kernels are therefore defined as the first and the second derivatives oh the charge-charge kernel. When $\mathbf{r}_i$ tends to $\mathbf{r}_j$ the three interaction kernels are:

$$\begin{aligned} \lim_{\mathbf{r}_{ij}\to 0}\mathrm{T}_{ij}^{qq} &= \mathrm{T}_{ii}^{qq} = \eta_i = \frac{2}{\sqrt{\pi}R_{q_i-q_i}}, \\ \lim_{\mathbf{r}_{ij}\to 0}\mathbf{T}_{ij}^{q\mu} &= \mathbf{T}_{ii}^{q\mu} = 0, \\ \lim_{\mathbf{r}_{ij}\to 0}\mathbf{T}_{ij}^{\mu\mu} &= \mathbf{T}_{ii}^{\mu\mu} = \frac{1}{\alpha_i} = \sqrt{\frac{2}{\pi}}\frac{\mathbf{1}}{3R_{\mu_i}^3}. \end{aligned} \tag{1.12}$$

Therefore, charge and dipole distributions (R_{q_i} and R_{μ_i}, respectively) can directly be computed in terms of η_i and α_i:

$$\begin{aligned} R_{q_i} &= \sqrt{\frac{2}{\pi}}\frac{1}{\eta_i} \\ R_{\mu_i} &= \left(\sqrt{\frac{2}{\pi}}\frac{1}{3}\alpha_i\right)^{\frac{1}{3}} \end{aligned} \tag{1.13}$$

As a consequence, FQFμ equations are defined in terms of three parameters only: electronegativity, chemical hardness and polarizability for each atom type.

The energy expression presented in eq.1.10 then reads:

$$\begin{aligned} E(\mathbf{q},\boldsymbol{\mu}) &= \sum_i q_i\chi_i + \frac{1}{2}\sum_{i,j} q_i \mathrm{T}^{qq}_{ij} q_j + \sum_{i,j} q_i \mathbf{T}^{q\mu}_{ij}\boldsymbol{\mu}^\dagger_j + \frac{1}{2}\sum_{i,j}\boldsymbol{\mu}^\dagger_i \mathbf{T}^{\mu\mu}_{ij}\boldsymbol{\mu}_j \\ &= \chi^\dagger\mathbf{q} + \frac{1}{2}\mathbf{q}^\dagger\mathbf{T}^{qq}\mathbf{q} + \mathbf{q}^\dagger\mathbf{T}^{q\mu}\boldsymbol{\mu} + \frac{1}{2}\boldsymbol{\mu}^\dagger\mathbf{T}^{\mu\mu}\boldsymbol{\mu} \end{aligned} \tag{1.14}$$

Similarly to FQ, to satisfy the Electronegativity Equalization Principle (EEP), the total molecular charge is constrained to Q_{tot}.

Similarly to FQ force field, Eq. 1.14 can be translated to the molecular case, by properly avoiding the charge transfer between different molecules. In this case, the energy functional (F) can be rewritten as:

$$\begin{aligned} F(\mathbf{q},\boldsymbol{\mu},\lambda) &= E(\mathbf{r},\mathbf{q},\boldsymbol{\mu}) + \sum_\alpha \lambda_\alpha \sum_i q_{\alpha,i} - Q_\alpha \\ &= \sum_{\alpha,i} q_{\alpha,i}\chi_{\alpha,i} + \frac{1}{2}\sum_{\alpha,i}\sum_{\beta,j} q_{\alpha,i}\mathrm{T}^{qq}_{\alpha i,\beta j} q_{\beta,j} + \sum_{\alpha,i}\sum_{\beta,j} q_{\alpha,i}\mathbf{T}^{q\mu}_{\alpha i,\beta j}\boldsymbol{\mu}^\dagger_{\beta,j} \\ &+ \frac{1}{2}\sum_{\alpha,i}\sum_{\beta,j}\boldsymbol{\mu}^\dagger_{\alpha,i}\mathbf{T}^{\mu\mu}_{\alpha i,\beta j}\boldsymbol{\mu}_{\beta,j} + \sum_\alpha \lambda_\alpha \sum_i q_{\alpha,i} - Q_\alpha \\ &= \chi^\dagger\mathbf{q} + \frac{1}{2}\mathbf{q}^\dagger\mathbf{T}^{qq}\mathbf{q} + \mathbf{q}^\dagger\mathbf{T}^{q\mu}\boldsymbol{\mu} + \frac{1}{2}\boldsymbol{\mu}^\dagger\mathbf{T}^{\mu\mu}\boldsymbol{\mu} + \lambda^\dagger\mathbf{q}, \end{aligned} \tag{1.15}$$

By differentiating eq.1.15 with respect to FQFμ variables (charges, dipoles and Lagrange multipliers) the constrained minimum condition is obtained:

$$\begin{cases} \sum_{\beta,j}\mathbf{T}^{qq}_{\alpha i,\beta j} q_{\beta,j} + \lambda_\alpha + \sum_{\beta,j}\mathbf{T}^{q\mu}_{\alpha i,\beta j}\boldsymbol{\mu}_{\beta,j} = -\chi_{\alpha,i} \\ \sum_{\beta,j}\mathbf{T}^{\mu\mu}_{\alpha i,\beta j}\mu_{\beta,j} + \sum_{\beta,j}\mathbf{T}^{q\mu}_{\alpha i,\beta j} q_{\beta,j} = 0 \\ \sum_{\alpha,i} q_{\alpha,i} = Q_\alpha \end{cases} \tag{1.16}$$

which can be recast as:

$$\begin{pmatrix} \mathbf{T}^{qq} & \mathbf{1}_\lambda & \mathbf{T}^{q\mu} \\ \mathbf{1}_\lambda^\dagger & \mathbf{0} & \mathbf{0} \\ -\mathbf{T}^{q\mu\dagger} & \mathbf{0} & \mathbf{T}^{\mu\mu}, \end{pmatrix} \begin{pmatrix} \mathbf{q} \\ \boldsymbol{\lambda} \\ \boldsymbol{\mu} \end{pmatrix} = \begin{pmatrix} -\boldsymbol{\chi} \\ \mathbf{Q} \\ \mathbf{0} \end{pmatrix} \tag{1.17}$$

$$\mathbf{M}\mathbf{Q}_\lambda = -\mathbf{C}_Q, \tag{1.18}$$

where Q_λ vector contains charges, dipoles and Lagrangian multipliers, whereas $\mathbf{C}_Q$ vector contains total charge constraint and atomic electronegativity.
The QM/FQFμ approach can be defined by exploiting the same strategy proposed for QM/FQ. Thus, for an HF description of the QM portion, the global energy functional is:

$$\begin{aligned} F(\mathbf{D}, \mathbf{q}, \mu, \lambda) = \mathrm{Tr}\mathbf{h}\mathbf{D} + \frac{1}{2}\mathrm{Tr}\mathbf{D}\mathbf{G}(\mathbf{D}) \\ + \boldsymbol{\chi}^\dagger\mathbf{q} + \frac{1}{2}\mathbf{q}^\dagger\mathbf{T}^{qq}\mathbf{q} + \mathbf{q}^\dagger\mathbf{T}^{q\mu}\boldsymbol{\mu} + \frac{1}{2}\boldsymbol{\mu}^\dagger\mathbf{T}^{\mu\mu}\boldsymbol{\mu} + \boldsymbol{\lambda}^\dagger\mathbf{q} \\ + \mathbf{q}^\dagger\mathbf{V}(\mathbf{D}) - \boldsymbol{\mu}^\dagger\mathbf{E}(\mathbf{D}), \end{aligned} \tag{1.19}$$

where $\mathbf{V}(\mathbf{D})$ is the electric potential and $\mathbf{E}(\mathbf{D})$ is the electric field. Thus, similarly to QM/FQ, the QM/MM interaction energy is limited to electrostatics.
To solve for FQFμ variables, we differentiate Eq. 1.19 with respect to charges, dipoles and lagrangian multipliers. Thus, the linear system defined in eq.1.18, now reads:

$$\mathbf{M}\mathbf{Q}_\lambda = -\mathbf{C}_Q - \mathbf{S}(\mathbf{D}). \tag{1.20}$$

where $\mathbf{S(D)}$ is a vector which contains the QM potential and electric field.
By finally differentiating the functional defined in eq. 1.19 with respect to the density matrix elements in the AO basis set, the QM Fock is obtained:

$$\mathbf{F}_{\mu\nu} = \frac{\partial F}{\partial D_{\mu\nu}} = h_{\mu\nu} + G_{\mu\nu}(\mathbf{D}) + \sum_i q_i V_{\mu\nu,i} - \sum_j \mu_j E_{\mu\nu,j} \tag{1.21}$$

where both the charges and the dipoles have to be updated at each SCF cycles, due to their explicit dependence on the QM density.

1.3. Non-electrostatic interactions

In all the aforementioned QM/MM approaches, the QM/MM interaction is treated at the purely electrostatic level, with the inclusion of mutual polarization effects in polarizable QM/MM methods. Non-electrostatic interactions are usually included in the QM/MM modeling by means of classical, parameterized functions. The most used approach is the Lennard-Jones potential (V_{LJ}),[165] which reads:

$$V_{LJ}(\mathbf{r}) = 4\epsilon\left[\left(\frac{\sigma}{\mathbf{r}}\right)^{12} - \left(\frac{\sigma}{\mathbf{r}}\right)^{6}\right], \tag{1.22}$$

where $\mathbf{r}$ is the distance between two interacting atoms, ε is the depth of the potential well, and σ is the distance at which the atom-atom potential energy V is zero.

Clearly, a classical, parameterized function, as the one in Eq. 1.22, only affects the total energy of the QM/MM system. In fact, contrary to the electrostatic term which acts as an external potential in the Hamiltonian in both electrostatic and polarizable embeddings, the Lennard-Jones energetic term is only added to the total energy of the system, because it does not depend on the QM density but only on atom positions. As a consequence, only indirect contributions to molecular properties are considered, because a parametrized function does not directly affect the Hamiltonian, its derivatives and response equations. Such an approximation might be justified in case of polar embeddings. However, non-electrostatics dominates the interaction between many biological systems and it is indeed the leading force in many chemical reactions.[166]

For this reason, in the last years, different approaches have been developed to describe both Pauli repulsion and quantum dispersion in QM/MM approaches.[56–58,167–175] Among them, in this thesis, we focus on a novel approach to describe Pauli repulsion,[133] which is general enough to be coupled to any description of the electrostatic interaction. In particular, it has been coupled to non-polarizable, FQ and FQFμ force fields.[176,177]

In such a model, each MM molecule is endowed with a set of s-type Gaussian functions, which simulate the density of the MM portion (ρ^{MM}). The repulsion energy term is then written as the opposite of an exchange integral between

the QM ρ^{QM} and MM density. This approximation is justified by the fact that the repulsive penetration energy terms in the Van der Waals is almost twice the attractive exchange interaction. Therefore, the QM/MM repulsion energy terms $E^{\mathrm{rep}}_{\mathrm{QM/MM}}$ reads:[178,179]

$$E^{\mathrm{rep}}_{\mathrm{QM/MM}} = \frac{1}{2}\int \frac{\mathrm{d}\mathbf{r_1}\,\mathrm{d}\mathbf{r_2}}{r_{12}}\rho_{\mathrm{QM}}(\mathbf{r_1},\mathbf{r_2})\rho_{\mathrm{MM}}(\mathbf{r_2},\mathbf{r_1}). \tag{1.23}$$

To define the ρ^{MM}, the aforementioned s-type Gaussian functions are located over bond and lone pair MM regions, as first proposed in Ref. 169. The ρ^{MM} is then defined as:

$$\rho_{\mathrm{MM}}(\mathbf{r_1},\mathbf{r_2}) = \sum_{\mathbf{R}} \xi^2_{\mathbf{R}} e^{-\beta_{\mathbf{R}}(\mathbf{r_1}-\mathbf{R})^2}\cdot e^{-\beta_{\mathbf{R}}(\mathbf{r_2}-\mathbf{R})^2}, \tag{1.24}$$

where $\mathbf{R}$ runs over the centroids of the s-type Gaussian functions, whereas β and ξ are two free parameters of the model, which can be determined by reproducing the exchange-repulsion energy contribution calculated by some energy decomposition analysis (EDA).[180–182]

By substituting eq. 1.24 in eq.1.23, the working expression for QM/MM Pauli-repulsion is finally obtained:

$$E^{rep}_{\mathrm{QM/MM}} = \frac{1}{2}\sum_{\mathbf{R}}\int \frac{\mathrm{d}\mathbf{r_1}\,\mathrm{d}\mathbf{r_2}}{r_{12}}\rho_{\mathrm{QM}}(\mathbf{r_1},\mathbf{r_2})\cdot\left[\xi^2_{\mathbf{R}} e^{-\beta_{\mathbf{R}}(\mathbf{r_1}-\mathbf{R})^2}\cdot e^{-\beta_{\mathbf{R}}(\mathbf{r_2}-\mathbf{R})^2}\right] \tag{1.25}$$

The energy expression in eq.1.25 can be rewritten as:

$$E^{rep}_{\mathrm{QM/MM}} = \frac{1}{2}\mathrm{Tr}\mathbf{D_{QM}}\mathbf{K}(\mathbf{D_{MM}}) \tag{1.26}$$

where $\mathbf{K}$ is the exchange integral and $\mathbf{D}_{QM/MM}$ are the QM and MM densities. Due to the fact that $\mathbf{D_{MM}}$ is a fixed parameterized quantity, the QM Hamiltonian is modified by an additional term to the one-electron integrals:

$$\tilde{h}_{\mu\nu} = h_{\mu\nu} + \frac{1}{2}K_{\mu\nu}(\mathbf{D_{MM}}) \tag{1.27}$$

This explicit contribution to the QM Hamiltonian propagates to the calculation of molecular properties and spectra, through the definition of suitable analytical procedures. This method, is first applied to the calculation of the quantum Pauli repulsion interaction energy of molecular

systems in aqueous solution.[176] It has been extended to spectroscopy starting from a formulation of QM/MM terms for Electron Paramagnetic Resonance (EPR) Hyperfine Coupling Constant.[177] Also this Pauli repulsion energy term formulation has been considered to show that density confinement needs to be included to get a physically consistent picture of solvatochromism.[128] In **Chapter 1** of this thesis the model has been applied to non-linear optical response properties for which repulsion forces can indeed be important to the determination of a solute's (Hyper)-Polarizability as electrostatic interactions, even for a solvent as polar as water.[183]

1.4. Quantum Embedding Methods

Highly accurate results can be obtained by exploiting polarizable QM/MM approaches, eventually coupled to a refined treatment of non-electrostatic interactions. However, the quality of a QM/MM embedding strongly depends on the values assigned to the parameters defining the model (e.g. χ and η in QM/FQ approach). Also, as we have shown above, the inclusion of non-electrostatic interactions at a mixed quantum-classical level is far from trivial and requires ad-hoc models to mimic the quantum nature of such energy terms.

To overcome such limitations of atomistic QM/classical approaches, quantum embedding methods can be exploited.[67–70,70–81]

In such approaches, the entire system is treated by resorting to a QM description, thus, all the interactions are retained at the purely QM level. The reduction in computational cost with respect to a QM calculation on the whole system is obtained by partitioning the system in at least one active and one inactive part. In the fashion of focused models, only the active wavefunction/density is usually optimized, whereas the inactive wavefunction/density remains frozen.

Different Quantum Embedding Models have been presented in literature. Most embedding theories are developed within the framework of density functional theory (DFT) due to the favourable compromise between computational cost and accuracy.[75,184,185] Among the many developed methods, the most exploited is probably the Frozen Density Embedding model.[186–191] FDE describes a QM complex system based on the partitioning of the electronic density into at

least two subsystems, each described by a separate electronic density. The two subsystems may be just two molecules, a solute and the surrounding solvent, or any other partition of a chemical system. The FDE method is based on the assumption that one of the two densities may be kept frozen, while the total energy is variationally optimized with respect to the other density. By construction, in the embedding potential, defining the interaction between the two portions, a nonadditive kinetic energy potential is introduced.[78,80,81]
In this thesis, a novel quantum embedding method based on DFT theory is proposed, named Multilevel DFT (MLDFT).[192] Such an approach has its theoretical foundations in multilevel Hartree-Fock (MLHF),[193] in which a complex system is partitioned into an active and an inactive portions. Such a partitioning is performed by decomposing the density matrix into its two components by means of a Cholesky decomposition.[194–196] While the obtained active density is optimized in the field of the inactive density, the latter kept frozen during the SCF procedure. In the following MLHF and MLDFT models are detailed.

1.4.1. Multilevel Hartree-Fock

The starting point of MLHF is the HF energy expression for a closed-shell molecule in the molecular orbitals (MO) basis:

$$E = \mathrm{Tr}\mathbf{h}\mathbf{D} + \frac{1}{2}\mathrm{Tr}\mathbf{D}\mathbf{G}(\mathbf{D}) + h_{nuc}, \tag{1.28}$$

where $\mathbf{D}$ is the MO density matrix, $\mathbf{h}$ is the one-electron operator, $\mathbf{G}$ the two-electron repulsion integrals matrix and h_{nuc} contains the nuclear repulsion. Particularly $\mathbf{D}$ for a set of orthogonal MOs is given by:

$$\begin{pmatrix} 1_{oo} & 0_{ov} \\ 0_{vo} & 0_{vv} \end{pmatrix}$$

in which o indicates the occupied MOs, whereas v the virtual ones. The density matrix that represents a closed-shell system with N_e electrons must satisfy the conditions of idempotency, symmetry and trace[197]

$$\begin{aligned} \mathbf{D}^{\mathrm{T}} &= \mathbf{D} \\ \mathrm{Tr}\mathbf{D} &= \frac{1}{2}N_e \\ \mathbf{D}^2 &= \mathbf{D} \end{aligned} \tag{1.29}$$

A density matrix $\mathbf{D}$ for N_e-electron closed-shell system, that fulfills the conditions presented in eq.1.29, can be partitioned into an active part, $\mathbf{D}^A$, and an inactive (environment) one, $\mathbf{D}^B$. This is the starting point of the MLHF approach.[193] In particular, in a standard HF calculation, an initial density is usually constructed by means of the Superposition of Atomic Densities (SAD) approach.[198] However, the resulting SAD density is not an idempotent density, thus it cannot be decomposed. Two potential solutions to this problem can be exploited: (i) the McWeeny purification scheme to make idempotent a general density matrix;[199–201] (ii) construct the Fock matrix from the SAD density, diagonalize the HF equations, use the resulting MOs to construct a density matrix which is idempotent by definition.
Independently from the selected method to obtain the starting idempotent density matrix, in MLHF, the active $\mathbf{D}^A$ and the inactive densities $\mathbf{D}^B$ are generated by a partial Cholesky decomposition:[194,195]

$$\mathbf{D} = \mathbf{D}_A + \mathbf{D}_B \tag{1.30}$$

In particular, the exploited Cholesky decomposition provides a set of orthonormalized occupied active orbitals, which can be used to construct the active density matrix. The inactive density matrix is instead obtained by reverting Eq. 1.30. Both $\mathbf{D}^A$ and $\mathbf{D}^B$ satisfy the conditions of symmetry, trace and idempotency, in particular, the requirement of idempotency for the sum of the two densities is equivalent to requiring orthogonality between the active and inactive orbital fragments. Therefore, by construction, the active and inactive MOs are orthogonal by definition, preserving the Pauli principle. Finally, the active virtual MOs are instead constructed in terms of orthonormalized projected atomic orbitals, which have been shown to be among the most localized virtual MOs.[202]
By exploiting the partitioning introduced in Eq. 1.30, Eq. 1.28 can be rewritten as:

$$\begin{aligned} E(\mathbf{D^A}, \mathbf{D^B}) = &\mathrm{Tr}\mathbf{h}\mathbf{D^A} + \frac{1}{2}\mathrm{Tr}\mathbf{D^A}\mathbf{G}(\mathbf{D^A}) \\ &+ \mathrm{Tr}\mathbf{h}\mathbf{D^B} + \frac{1}{2}\mathrm{Tr}\mathbf{D^B}\mathbf{G}(\mathbf{D^B}) \\ &+ \mathrm{Tr}\mathbf{D^A}\mathbf{G}(\mathbf{D^B}) + h_{nuc} \end{aligned} \tag{1.31}$$

The structure of the active and inactive density matrices in the MO basis can be expressed as:

$$\mathbf{D^A} = \begin{array}{c} \begin{matrix} o_A & o_B & v_A & v_B \end{matrix} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array} \quad \mathbf{D^B} = \begin{array}{c} \begin{matrix} o_A & o_B & v_A & v_B \end{matrix} \\ \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

where o_A and o_B represent the active and inactive occupied space respectively, and v_A and v_B the active and inactive virtual space, respectively. Due to the structure of $\mathbf{D^A}$, itS dimension can be reduced to N_A, where $N_A = o_A + v_A$ is the number of the total active MOs, as:

$$\mathbf{D_r^A} = \begin{array}{c} \begin{matrix} o_A & v_A \end{matrix} \\ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \end{array}$$

Eq.1.31 can finally be rewritten as:

$$\begin{aligned} E(\mathbf{D^A}, \mathbf{D^B}) &= E(\mathbf{D_r^A}, \mathbf{D^B}) \\ &= \mathrm{Tr}\mathbf{h_r}\mathbf{D_r^A} + \frac{1}{2}\mathrm{Tr}\mathbf{D_r^A}\mathbf{G_r}(\mathbf{D_r^A}) \\ &+ \mathrm{Tr}\mathbf{h}\mathbf{D^B} + \frac{1}{2}\mathrm{Tr}\mathbf{D^B}\mathbf{G}(\mathbf{D^B}) \\ &+ \mathrm{Tr}\mathbf{D_r^A}\mathbf{G_r}(\mathbf{D^B}) + h_{nuc}, \end{aligned} \tag{1.32}$$

where the quantities highlighted with r are expressed in the active MO basis. The effective Fock matrix for the optimization of the active density matrix reads:

$$\mathbf{F_{eff,r}} = \mathbf{F_r}(\mathbf{D_r^A}) + \mathbf{G_r}(\mathbf{D^B}), \tag{1.33}$$

where $\mathbf{F}_r(\mathbf{D^A})$ is the MO Fock matrix of the active part only. Being $\mathbf{D}^B$ fixed during the SCF cycles, the $\mathbf{G_r}(\mathbf{D^B})$ term entering the Fock is thus a one-electron term, which can be computed once at the beginning of the SCF procedure in the AO basis, and then converted in the updated MO basis during the SCF. Differently from most QM/MM approaches, in MLHF electrostatics and Pauli repulsion interactions are automatically taken into account at the HF level. However, since the inactive density is kept frozen during the optimization of the active density, the active-inactive mutual

polarization is neglected. However, a large part of such interaction is recovered by the initial diagonalization of the Fock matrix constructed on the basis of the SAD density matrix.

1.4.2. Multilevel Density Functional Theory

In this section, we present a novel multilevel approach entirely based on Density Functional Theory (DFT), namely Multilevel DFT (MLDFT).[133] The MLDFT conceptually differs from the other quantum embedding methods in DFT framework because the partitioning in different subsystems is performed on the density matrix instead of the density function as, for instance, in FDE.[186–191] Similarly to MLHF,[193] the density of the system under investigation (active) is optimized, whereas the environment (inactive) density is kept frozen during the active fragment optimization procedure. The choice of the partitioning is arbitrarily and depends on the specific features of the system and in particular from the properties one wishes to simulate. The starting point to derive the MLDFT energy expression is the DFT expression for the electronic energy of the system:

$$\begin{aligned} E &= \mathrm{Tr}\mathbf{hD} + \frac{1}{2}\mathrm{Tr}\mathbf{DJ}(\mathbf{D}) - \frac{1}{2}c_x\mathrm{Tr}\mathbf{DK}(\mathbf{D}) + (1+c_x)E_x + Ec \\ &= \mathrm{Tr}\mathbf{hD} + \frac{1}{2}\mathrm{Tr}\mathbf{DJ}(\mathbf{D}) - \frac{1}{2}c_x\mathrm{Tr}\mathbf{DK}(\mathbf{D}) \\ &\quad + (1-c_x)\int \rho(\mathbf{r})\varepsilon_x(\rho(\mathbf{r}))\,\mathrm{d}\mathbf{r} + \int \rho(\mathbf{r})\varepsilon_c(\rho(\mathbf{r}))\,\mathrm{d}\mathbf{r} \end{aligned}$$

where $\mathbf{D}$ is the density matrix, $\mathbf{h}$ is the one-electron operator, whereas $\mathbf{J}$ and $\mathbf{K}$ are the Coulomb and the Exchange matrices, respectively. E_x and E_c terms are DFT exchange and correlation energies; $\rho(\mathbf{r})$ is the DFT density function and $\varepsilon_x, \varepsilon_c$ are the exchange and correlation energy densities, respectively. The coefficient c_x defines if the DFT functional used is pure ($c_x = 0$), or hybrid ($c_x \neq 0$).

The DFT density $\rho(\mathbf{r})$ can be expressed in terms of the density matrix: ($\mathbf{D}$) as:

$$\rho(\mathbf{r}) = \sum_{\mu\nu} D_{\mu\nu}\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r}) \tag{1.34}$$

where $\{\chi_\mu\}$ are the AO basis functions.

Generally the total DFT energy is minimized in the AO basis. In order to reformulate the minimization in the MO basis, the same strategy present for

Multilevel Hartree-Fock in the previous section can be used. Within the MLDFT framework, the subdivision of the entire system into two fragments is based on the following decomposition of the density $\mathbf{D}$ and $\rho(\mathbf{r})$:

$$\mathbf{D} = \mathbf{D}^A + \mathbf{D}^B \ , \qquad \rho(\mathbf{r}) = \rho^A(\mathbf{r}) + \rho^B(\mathbf{r}) \ , \tag{1.35}$$

where, A correspond to the active layer, whereas B indicates the inactive one. As stated above, the active and inactive densities are usually defined on a physico-chemical basis. In case of a molecule embedded in an external environment, it is natural to define the target molecule as the active fragment, whereas the environment system is included the inactive part. Notice, however, that the partitioning in Eq. 1.35 is arbitrary and strongly depends on the method which is selected to mathematically decompose the total density matrix $\mathbf{D}$. Similarly to MLHF, a Cholesky decomposition of the total density is performed for the active occupied MOs, from which the active density $\mathbf{D}^A$ is calculated.[193–196] The procedure ensures the all active and inactive orbitals are orthogonal.
Now using Eq. 1.35, the total electronic energy in Eq. 1.34 can be written as:

$$\begin{aligned} E = & \mathrm{Tr}\mathbf{h}\mathbf{D}^A + \mathrm{Tr}\mathbf{h}\mathbf{D}^B \\ & + \frac{1}{2}\mathrm{Tr}\mathbf{D}^A\mathbf{J}(\mathbf{D}^A) + \frac{1}{2}\mathrm{Tr}\mathbf{D}^B\mathbf{J}(\mathbf{D}^B) + \mathrm{Tr}\mathbf{D}^A\mathbf{J}(\mathbf{D}^B) \\ & - c_x\left(\frac{1}{2}\mathrm{Tr}\mathbf{D}^A\mathbf{K}(\mathbf{D}^A) + \frac{1}{2}\mathrm{Tr}\mathbf{D}^B\mathbf{K}(\mathbf{D}^B) + \mathrm{Tr}\mathbf{D}^A\mathbf{K}(\mathbf{D}^B)\right) \\ & + (1-c_x)\int \rho(\mathbf{r})\varepsilon_x(\rho(\mathbf{r}))\,\mathrm{d}\mathbf{r} + \int \rho(\mathbf{r})\varepsilon_c(\rho(\mathbf{r}))\,\mathrm{d}\mathbf{r} \ , \end{aligned} \tag{1.36}$$

where the symmetry of $\mathbf{J}$ and $\mathbf{K}$ matrices have been used. Differently from MLHF, the latter term is not linear in the densities of the two subsystems. Therefore such a contribution can not be directly separated in two contributions that depend on ρ^A and ρ^B only. In order to get a deeper physico-chemical vision of Eq. 1.36, the last two terms can be written by using a trivial identity for the exchange-correlation energy density ($\varepsilon_{xc} = \varepsilon_x + \varepsilon_c$):

$$
\begin{aligned}
&\int [\rho^A(\mathbf{r}) + \rho^B(\mathbf{r})]\varepsilon_{xc}(\rho^A(\mathbf{r}) + \rho^B(\mathbf{r}))\,\mathrm{d}\mathbf{r} \\
&= \int [\rho^A(\mathbf{r}) + \rho^B(\mathbf{r})]\varepsilon_{xc}(\rho^A(\mathbf{r}) + \rho^B(\mathbf{r}))\,\mathrm{d}\mathbf{r} \\
&+ \int \rho^A(\mathbf{r})\varepsilon_{xc}(\rho^A(\mathbf{r}))\,\mathrm{d}\mathbf{r} - \int \rho^A(\mathbf{r})\varepsilon_{xc}(\rho^A(\mathbf{r}))\,\mathrm{d}\mathbf{r} \\
&+ \int \rho^B(\mathbf{r})\varepsilon_{xc}(\rho^B(\mathbf{r}))\,\mathrm{d}\mathbf{r} - \int \rho^B(\mathbf{r})\varepsilon_{xc}(\rho^B(\mathbf{r}))\,\mathrm{d}\mathbf{r} \\
&+ \int \rho^A(\mathbf{r})\varepsilon_{xc}(\rho^B(\mathbf{r}))\,\mathrm{d}\mathbf{r} - \int \rho^A(\mathbf{r})\varepsilon_{xc}(\rho^B(\mathbf{r}))\,\mathrm{d}\mathbf{r} \\
&+ \int \rho^B(\mathbf{r})\varepsilon_{xc}(\rho^A(\mathbf{r}))\,\mathrm{d}\mathbf{r} - \int \rho^B(\mathbf{r})\varepsilon_{xc}(\rho^A(\mathbf{r}))\,\mathrm{d}\mathbf{r}. \qquad (1.37)
\end{aligned}
$$

Substituting Eq. 1.37 into Eq. 1.36, and reorganizing the different terms:

$$
\begin{aligned}
E[\mathbf{D}^A;\mathbf{D}^B] &= \mathrm{Tr}\mathbf{h}\mathbf{D}^A + \frac{1}{2}\mathrm{Tr}\mathbf{D}^A\mathbf{J}(\mathbf{D}^A) - \frac{1}{2}c_x\mathrm{Tr}\mathbf{D}^A\mathbf{K}(\mathbf{D}^A) \\
&+ (1-c_x)\int \rho^A(\mathbf{r})\varepsilon_x(\rho^A(\mathbf{r}))\,\mathrm{d}\mathbf{r} + \int \rho^A(\mathbf{r})\varepsilon_c(\rho^A(\mathbf{r}))\,\mathrm{d}\mathbf{r} \\
&+ \mathrm{Tr}\mathbf{h}\mathbf{D}^B + \frac{1}{2}\mathrm{Tr}\mathbf{D}^B\mathbf{J}(\mathbf{D}^B) - \frac{1}{2}c_x\mathrm{Tr}\mathbf{D}^B\mathbf{K}(\mathbf{D}^B) \\
&+ (1-c_x)\int \rho^B(\mathbf{r})\varepsilon_x(\rho^B(\mathbf{r}))\,\mathrm{d}\mathbf{r} + \int \rho^B(\mathbf{r})\varepsilon_c(\rho^B(\mathbf{r}))\,\mathrm{d}\mathbf{r} \\
&+ \mathrm{Tr}\mathbf{D}^A\mathbf{J}(\mathbf{D}^B) - c_x\mathrm{Tr}\mathbf{D}^A\mathbf{K}(\mathbf{D}^B) \\
&+ (1-c_x)\left(\int \rho^A(\mathbf{r})\varepsilon_x(\rho^B(\mathbf{r}))\,\mathrm{d}\mathbf{r} + \int \rho^B(\mathbf{r})\varepsilon_x(\rho^A(\mathbf{r}))\,\mathrm{d}\mathbf{r}\right) \\
&+ \left(\int \rho^A(\mathbf{r})\varepsilon_c(\rho^B(\mathbf{r}))\,\mathrm{d}\mathbf{r} + \int \rho^B(\mathbf{r})\varepsilon_c(\rho^A(\mathbf{r}))\,\mathrm{d}\mathbf{r}\right) + E^{AB}_{non-add}
\end{aligned} \qquad (1.38)
$$

where

$$
\begin{aligned}
&E^{AB}_{non-add} = \\
&(1-c_x)\left(\int \rho(\mathbf{r})\varepsilon_x(\rho(\mathbf{r}))\,\mathrm{d}\mathbf{r} - \int \rho(\mathbf{r})\varepsilon_x(\rho^A(\mathbf{r}))\,\mathrm{d}\mathbf{r} - \int \rho(\mathbf{r})\varepsilon_x(\rho^B(\mathbf{r}))\,\mathrm{d}\mathbf{r}\right) \\
&\quad + \int \rho(\mathbf{r})\varepsilon_c(\rho(\mathbf{r}))\,\mathrm{d}\mathbf{r} - \int \rho(\mathbf{r})\varepsilon_c(\rho^A(\mathbf{r}))\,\mathrm{d}\mathbf{r} - \int \rho(\mathbf{r})\varepsilon_c(\rho^B(\mathbf{r}))\,\mathrm{d}\mathbf{r}
\end{aligned} \qquad (1.39)
$$

In Eq. 1.38 the first four lines define the energy of the active and inactive fragments, whereas the last three lines define the active-inactive interaction, including non-additive $E_{non-add}$ term due to the non-linearity of DFT functionals. Similarly to MLHF, in MLDFT the density of the inactive part $\mathbf{D}^B$ (and $\rho^B(\mathbf{r})$) is kept frozen, acting as an external field on the active fragment density.
The total DFT Fock matrix is given by:

$$\begin{aligned} F_{\mu\nu} &= h_{\mu\nu} + J_{\mu\nu}(\mathbf{D}) - c_x K_{\mu\nu}(\mathbf{D}) \\ &\quad + (1 - c_x) \int v_x(\rho(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\,\mathrm{d}\mathbf{r} + \int v_c(\rho(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\,\mathrm{d}\mathbf{r}\ , \end{aligned} \tag{1.40}$$

where, $v_x(\rho(\mathbf{r}))$ and $v_c(\rho(\mathbf{r}))$ are the exchange and correlation potential densities, respectively. By introducing the partitioning in Eq. 1.35, Eq. 1.40:

$$\begin{aligned} F_{\mu\nu} &= h_{\mu\nu} + J_{\mu\nu}(\mathbf{D}^A) + J_{\mu\nu}(\mathbf{D}^B) - c_x \left(K_{\mu\nu}(\mathbf{D}^A) + K_{\mu\nu}(\mathbf{D}^B)\right) \\ &\quad + (1 - c_x) \int v_x(\rho^A(\mathbf{r}) + \rho^B(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\,\mathrm{d}\mathbf{r} \\ &\quad + \int v_c(\rho^A(\mathbf{r}) + \rho^B(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\,\mathrm{d}\mathbf{r}\ . \end{aligned} \tag{1.41}$$

By exploiting an identity similar to Eq. 1.37 for the exchange-correlation potential density ($v_{xc} = v_x + v_c$), the last two terms in Eq. 1.41 become:

$$\begin{aligned} &\int v_{xc}(\rho_A(\mathbf{r}) + \rho_B(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\,\mathrm{d}\mathbf{r} \\ &= \int v_{xc}(\rho_A(\mathbf{r}) + \rho_B(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\,\mathrm{d}\mathbf{r} + \\ &+ \int v_{xc}(\rho_A(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\,\mathrm{d}\mathbf{r} - \int v_{xc}(\rho_A(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\,\mathrm{d}\mathbf{r} + \\ &+ \int v_{xc}(\rho_B(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\,\mathrm{d}\mathbf{r} - \int v_{xc}(\rho_B(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\,\mathrm{d}\mathbf{r}\ . \end{aligned} \tag{1.42}$$

Reorganizing the terms in Eq. 1.41, the working expression for the MLDFT

Fock matrix can be obtained:

$$
\begin{aligned}
&F_{\mu\nu} = h_{\mu\nu} \\
&2e_A \begin{cases} +J_{\mu\nu}(\mathbf{D}^A) - c_x K_{\mu\nu}(\mathbf{D}^A) \\ + \int v_{xc}(\rho^A(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\,\mathrm{d}\mathbf{r} \end{cases} \\
&2e_B \begin{cases} +J_{\mu\nu}(\mathbf{D}^B) - c_x K_{\mu\nu}(\mathbf{D}^B) \\ + \int v_{xc}(\rho^B(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\,\mathrm{d}\mathbf{r} \end{cases} \\
&2e_{non-add} \Big\{ + \int [v_{xc}(\rho(\mathbf{r})) - v_{xc}(\rho^A(\mathbf{r})) - v_{xc}(\rho^B(\mathbf{r}))]\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\,\mathrm{d}\mathbf{r}
\end{aligned}
\tag{1.43}
$$

where, the two-electron contributions of A and B fragments and the interaction term AB are highlighted as $2e_X$, $\{X = A, B, AB\}$. Note that $2e_B$ is a one-electron contribution being $\mathbf{D}^B$ (and consequently $\rho^B(\mathbf{r})$) constant during the active density optimization.
There are two main advantages of using MLDFT compared to full DFT:

1. in MLDFT, at each cycle, only the active exchange term is to be computed, because the exchange integral of the inactive fragment, constant during the optimization, is computed at the first SCF cycle only.

2. MLDFT SCF procedure can be performed in the MO basis of the active part only, thus intrinsically reducing the computational time as previously observed for the MLHF method.[193]

1.4.3. Extension of MLHF and MLDFT to Open-Shell Systems

In this thesis, the described MLHF and MLDFT are further extended to the calculation of the electronic structure of open-shell systems. To this end, the unrestricted formalism is exploited, and the resulting methods are called unrestricted MLHF (UMLHF) and unrestricted MLDFT (UMLDFT).
The starting point to derive UMLHF/UMLDFT is the expression of the energy $E[\mathbf{D}^\alpha, \mathbf{D}^\beta]$ for open-shell systems in the unrestricted formalism:

$$
\begin{aligned}
E[\mathbf{D}^\alpha, \mathbf{D}^\beta] &= E[\mathbf{D}^\alpha] + E[\mathbf{D}^\beta] \\
&= \mathrm{Tr}\mathbf{h}\mathbf{D}^\alpha + \frac{1}{2}\mathrm{Tr}\mathbf{D}^\alpha\mathbf{J}(\mathbf{D}^\alpha + \mathbf{D}^\beta) \\
&\quad - c_x\left(\frac{1}{2}\mathrm{Tr}\mathbf{D}^\alpha\mathbf{K}(\mathbf{D}^\alpha)\right) \\
&\quad + (1 - c_x)\int \rho^\alpha(\mathbf{r})\varepsilon_x[\rho(\mathbf{r})]\mathrm{d}\mathbf{r} \\
&\quad + \int \rho^\alpha(\mathbf{r})\varepsilon_c[\rho(\mathbf{r})]\mathrm{d}\mathbf{r} \\
&\quad + \mathrm{Tr}\mathbf{h}\mathbf{D}^\beta + \frac{1}{2}\mathrm{Tr}\mathbf{D}^\beta\mathbf{J}(\mathbf{D}^\alpha + \mathbf{D}^\beta) \\
&\quad - c_x\left(\frac{1}{2}\mathrm{Tr}\mathbf{D}^\beta\mathbf{K}(\mathbf{D}^\beta)\right) \\
&\quad + (1 - c_x)\int \rho^\beta(\mathbf{r})\varepsilon_x[\rho(\mathbf{r})]\mathrm{d}\mathbf{r} \\
&\quad + \int \rho^\beta(\mathbf{r})\varepsilon_c[\rho(\mathbf{r})]\mathrm{d}\mathbf{r}
\end{aligned}
\tag{1.44}
$$

$E[\mathbf{D}^\alpha, \mathbf{D}^\beta]$ is split into two terms, E^α and E^β, which refer to α and β spin. The $\mathbf{D}^\sigma$ are spin density matrices (with $\sigma = \alpha, \beta$), $\mathbf{h}$ is the one-electron operator, and $\mathbf{J}$ and $\mathbf{K}$ are Coulomb and exchange matrices, respectively. Equation 5.1 is formulated for a generic DFT functional, where $\rho^\sigma(\mathbf{r})$ are spin DFT density functions and $\varepsilon_x, \varepsilon_c$ indicate exchange and correlation energy densities per unit particle. The UHF equations can easily be recovered by imposing $c_x = 1$ and $\varepsilon_c = 0$. The total density matrix $\mathbf{D}$ can be obtained from α and β spin density matrices as $\mathbf{D} = \mathbf{D}^\alpha + \mathbf{D}^\beta$.

Similarly to the closed shell case,[192,193] unrestricted MLHF/MLDFT are formulated by separating the total system into active (A) and inactive (B) parts. From a mathematical point of view, the separation is performed by decomposing spin-density matrices into active $\mathbf{D}^\sigma_A$ and inactive $\mathbf{D}^\sigma_B$ contributions:

$$
\begin{aligned}
\mathbf{D}^\alpha = \mathbf{D}^\alpha_A + \mathbf{D}^\alpha_B \qquad &\Rightarrow \qquad \rho^\alpha(\mathbf{r}) = \rho^\alpha_A(\mathbf{r}) + \rho^\alpha_B(\mathbf{r}) \\
\mathbf{D}^\beta = \mathbf{D}^\beta_A + \mathbf{D}^\beta_B \qquad &\Rightarrow \qquad \rho^\beta(\mathbf{r}) = \rho^\beta_A(\mathbf{r}) + \rho^\beta_B(\mathbf{r})
\end{aligned}
\tag{1.45}
$$

A similar partitioning applies to DFT spin density functions $(\rho^\alpha(\mathbf{r}), \rho^\beta(\mathbf{r}))$.

Note that in general $\mathbf{D}_X^\alpha \neq \mathbf{D}_X^\beta$, $\{X = A, B\}$. By substituting Eq. 5.2 into Eq. 5.1, we obtain:

$$E[\mathbf{D}_A^\alpha, \mathbf{D}_B^\alpha, \mathbf{D}_A^\beta, \mathbf{D}_B^\beta] = E[\mathbf{D}_A^\alpha, \mathbf{D}_B^\alpha] + E[\mathbf{D}_A^\beta, \mathbf{D}_B^\beta] \tag{1.46}$$

where $E[\mathbf{D}_A^\sigma, \mathbf{D}_B^\sigma]$ ($\{\sigma = \alpha, \beta\}$) is given by:

$$\begin{aligned}
&E[\mathbf{D}_A^\sigma, \mathbf{D}_B^\sigma] = \\
&\qquad E_A^\sigma \begin{cases} \mathrm{Tr}\mathbf{h}\mathbf{D}_A^\sigma + \frac{1}{2}\mathrm{Tr}\mathbf{D}_A^\sigma \mathbf{J}(\mathbf{D}_A) \\ -\frac{1}{2}c_x \mathrm{Tr}\mathbf{D}_A^\sigma \mathbf{K}(\mathbf{D}_A^\sigma) + \int \rho_A^\sigma(\mathbf{r})\varepsilon_{xc}(\rho_A(\mathbf{r}))\mathrm{d}\mathbf{r} \end{cases} \\
&\qquad E_B^\sigma \begin{cases} +\mathrm{Tr}\mathbf{h}\mathbf{D}_B^\sigma + \frac{1}{2}\mathrm{Tr}\mathbf{D}_B^\sigma \mathbf{J}(\mathbf{D}_B) \\ -\frac{1}{2}c_x \mathrm{Tr}\mathbf{D}_B^\sigma \mathbf{K}(\mathbf{D}_B^\sigma) + \int \rho_B^\sigma(\mathbf{r})\varepsilon_{xc}(\rho_B(\mathbf{r}))\mathrm{d}\mathbf{r} \end{cases} \\
&\qquad E_{int}^\sigma \begin{cases} +\mathrm{Tr}\mathbf{D}_A^\sigma \mathbf{J}(\mathbf{D}_B) - c_x \mathrm{Tr}\mathbf{D}_A^\sigma \mathbf{K}(\mathbf{D}_B^\sigma) \\ +\int \rho_A^\sigma(\mathbf{r})\varepsilon_{xc}(\rho_B(\mathbf{r}))\mathrm{d}\mathbf{r} + \int \rho_B^\sigma(\mathbf{r})\varepsilon_{xc}(\rho_A(\mathbf{r}))\mathrm{d}\mathbf{r} \end{cases} \\
&\qquad E_{non-add}^\sigma \begin{cases} +\int \rho^\sigma(\mathbf{r})\varepsilon_{xc}(\rho(\mathbf{r}))\mathrm{d}\mathbf{r} \\ -\int \rho^\sigma(\mathbf{r})\varepsilon_{xc}(\rho_A(\mathbf{r}))\mathrm{d}\mathbf{r} - \int \rho^\sigma(\mathbf{r})\varepsilon_{xc}(\rho_B(\mathbf{r}))\mathrm{d}\mathbf{r} \end{cases}
\end{aligned} \tag{1.47}$$

In Eq. 5.4, $\varepsilon_c + (1 - c_x)\varepsilon_x$ is substituted by ε_{xc} to make the notation compact, and energy terms are separated into active and inactive contributions ($E_{A,B}^\sigma$). Also, the coupling terms are divided into a purely interaction energy term, E_{int}^σ, and a non-additive contribution $E_{non-add}^\sigma$, which originates from the non linearity of ε_x and ε_c. It is worth noting that the non-additive energy terms vanish for UMLHF. As expected, the partitioning of the different terms in Eq. 5.4 is coherent with the MLDFT formulation for closed shell systems (see Ref.[192]).

As already discussed in the Introduction, the energy of the active fragment A is optimized while the inactive density B is kept fixed to the value resulting from the partitioning in Eq. 5.2. Therefore, the UMLHF/MLDFT spin-Fock matrix ($F_{\mu\nu}^\sigma$, in the AO basis $\{\chi_\mu\}$) can easily be recovered by differentiating the energy in Eq. 5.4 with respect to the active density ($D_A^\sigma, \rho_A^\sigma$), i.e.:

$$
\begin{aligned}
F^{\sigma}_{\mu\nu} &= h^{\sigma}_{\mu\nu} \\
2e^{\sigma}_{A} &\begin{cases} +J_{\mu\nu}(\mathbf{D}_A) - c_x K_{\mu\nu}(\mathbf{D}^{\sigma}_A) \\ +\int v^{\sigma}_{xc}(\rho^A(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\,\mathrm{d}\mathbf{r} \end{cases} \\
2e^{\sigma}_{B} &\begin{cases} +J_{\mu\nu}(\mathbf{D}_B) - c_x K_{\mu\nu}(\mathbf{D}^{\sigma}_B) \\ +\int v^{\sigma}_{xc}(\rho^B(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\,\mathrm{d}\mathbf{r} \end{cases} \\
2e^{\sigma}_{non-add} &\begin{cases} +\int [v^{\sigma}_{xc}(\rho(\mathbf{r})) - v^{\sigma}_{xc}(\rho^A(\mathbf{r})) - v^{\sigma}_{xc}(\rho^B(\mathbf{r}))] \\ \qquad \chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\,\mathrm{d}\mathbf{r} \end{cases}
\end{aligned}
\tag{1.48}
$$

where we have used the compact notation $v_{xc} = (1 - c_x)v_x + v_c$. The two-electron contributions to the Fock matrix can be are grouped into different $2e_X$ terms, with X = A, B. $2e^{\sigma}_{non-add}$ is due to the non-linearity in the DFT functional, and again vanishes for UMLHF. Finally, note that $2e^{\sigma}_{B}$ accounts for the frozen fragment, of which the density $(D_B, \rho^B(\mathbf{r}))$ does not change along SCF cycles. Therefore, $2e^{\sigma}_{B}$ is a constant one-electron contribution, which is computed only once, at the beginning of the SCF procedure, similarly to MLHF/MLDFT.[192,193]

1.4.4. Coupling with an outer layer described with MM force fields

The UMLHF(UMLDFT) can be coupled to an additional MM(FQ) layer. To this end, we introduce the total energy of the system:

$$
\mathcal{E} = E_{\text{UMLHF(DFT)}} + E_{\text{MM}} + E^{\text{int}}_{\text{UMLHF(DFT)/MM}} \,, \tag{1.49}
$$

where $E_{UMLHF(DFT)}$ is given in Eq. 5.4, whereas E_{MM} and $E^{\text{int}}_{\text{UMLHF(DFT)/MM}}$ are MM and UMLHF(DFT)/MM interaction energies, respectively. Electrostatic and polarizable QM/MM embedding differ from the way the interaction energy is specified (in our case $E^{\text{int}}_{\text{UMLHF(DFT)/MM}}$); electrostatic embedding approaches limit the description to electrostatic forces only, whereas mutual QM/MM polarization is modelled in polarizable embedding approaches.[8,62] In particular, non-polarizable embedding methods place fixed charges on MM atoms, which polarize the QM density. Different polarizable

QM/MM approaches exist;[56,58,127,129,131,203,204] in this work we exploit QM/Fluctuating Charges (FQ),[62,205] where each MM atom is assigned an atomic electronegativity (χ) and chemical hardness (η), which give rise to electric charges (q) as a response to the atomic chemical potential.[52,62,152–154] Therefore, for both non-polarizable UMLHF(DFT)/MM or polarizable UMLHF(DFT)/FQ, the UMLHF (DFT)/MM interaction energy can be written as follows:

$$E^{\text{int}}_{\text{UMLHF(DFT)/MM}} = \sum_i q_i V_i(\mathbf{D}^\alpha + \mathbf{D}^\beta) \ , \tag{1.50}$$

where $V_i(\mathbf{D}^\alpha + \mathbf{D}^\beta)$ is the electric potential generated by the total QM spin-density (i.e. both active and inactive contributions) on the i$-th$ charge (q_i). In the case of non-polarizable QM/MM, q_i values are fixed, whereas in QM/FQ they are obtained by minimizing the following energy expression:

$$\begin{aligned}\mathcal{E}[\mathbf{D}^\alpha_A, \mathbf{D}^\beta_A, \mathbf{D}^\alpha_B, \mathbf{D}^\beta_B, \mathbf{q}, \boldsymbol{\lambda}] = & E_{\text{UMLHF(DFT)}}[\mathbf{D}^\alpha_A, \mathbf{D}^\alpha_B] \\ & + E_{\text{UMLHF(DFT)}}[\mathbf{D}^\beta_A, \mathbf{D}^\beta_B] \\ & + \frac{1}{2}\mathbf{q}^\dagger_\lambda \mathbf{M}\mathbf{q}_\lambda + \mathbf{q}^\dagger_\lambda \mathbf{C}_Q \\ & + \mathbf{q}^\dagger_\lambda \mathbf{V}(\mathbf{D}^\alpha + \mathbf{D}^\beta),\end{aligned} \tag{1.51}$$

where $E_{\text{UMLHF(DFT)}}[\mathbf{D}^\alpha_A, \mathbf{D}^\alpha_B]$ and $E_{\text{UMLHF(DFT)}}[\mathbf{D}^\beta_A, \mathbf{D}^\beta_B]$ due to the total α- and β spin-densities, respectively. In Eq. 1.51, $\mathbf{q}_\lambda$ indicates a vector collecting FQ charges and a set of Lagrangian multipliers, which ensure charge conservation on each fragment composing the MM layer (e.g. on each solvent molecules for solvated systems). The $\mathbf{M}$ matrix is the interaction kernel between the FQ charges, which also contains the Lagrangian blocks,[110] and the vector $\mathbf{C}_Q$ accounts for the interaction between permanent moments, i.e. χ and charge constraints Q on each FQ moiety. The FQ charges equilibrated for the UMLHF(DFT)/FQ systems are obtained by minimizing the energy functional in Eq. 1.51. This procedure yields the following set of linear equations:

$$\mathbf{M}\mathbf{q}_\lambda = -\mathbf{C}_Q - \mathbf{V}(\mathbf{D}^\alpha + \mathbf{D}^\beta) \ . \tag{1.52}$$

In parallel, UMLHF(DFT)/MM spin Fock matrices are defined as follows:

$$F^\sigma_{\mu\nu} = F^{\sigma,\text{UMLHF(DFT)}}_{\mu\nu} + \sum_i q_i V_{i,\mu\nu} \ . \tag{1.53}$$

where $F_{\mu\nu}^{\sigma,\mathrm{UMLHF(DFT)}}$ is defined in Eq. 1.43. As for energy contributions, the additional QM/MM term is fixed and computed only at the first SCF cycle in case of non-polarizable QM/MM, whereas it changes in QM/FQ, because FQ charges depend on QM densities. Therefore, UMLHF(DFT)/FQ contribution to the Fock matrix needs to be updated at each SCF cycle, thus introducing mutual polarization effects between UMLHF(DFT) and FQ layers.

1.5. Properties Under Investigation

In this work, the multiscale models presented in the previous sections have been applied to different kind of molecular properties, such as: polarizability, first-hyperpolarizability and the Hyperfine Coupling Constant (hcc) of open shell systems.

1.5.1. Polarizability and First Hyperpolarizability

The discipline of nonlinear optics was born in 1875 when John Kerr first discovered that a material became birefringent when it was placed in an electric field: the refractive index was different in the directions parallel and perpendicular to the field.[206] Subsequently, such an effect (and many others) were understood in terms of the induced polarization which atoms and molecules sustain in the presence of a static or dynamic (oscillating) electric field.[207–221]

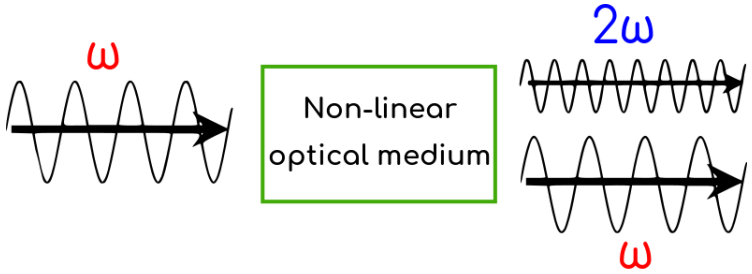


Figure 1.1. Scheme of a non-linear optical phenomenon occurring under the action of an external electric field.

Under the action of an external field **E**(t), the dipole moment can be expressed as a Taylor series:

$$\boldsymbol{\mu}(t) = \boldsymbol{\alpha}\mathbf{E}(t) + \frac{1}{2}\boldsymbol{\beta}\mathbf{E}(t)\mathbf{E}(t) + \frac{1}{6}\boldsymbol{\gamma}\mathbf{E}(t)\mathbf{E}(t)\mathbf{E}(t) + ... \tag{1.54}$$

$\boldsymbol{\alpha}$ is the dipole polarizability and governs the linear effect, and $\boldsymbol{\beta}, \boldsymbol{\gamma}$, ..., are the so-called dipole first, second, ..., hyperpolarizabilities, describing non-linear phenomena. Note that such quantities are generally n-elements tensors.
From the QM point of view, a general external electric field is usually seen as an external perturbation associated to a specific frequency:

$$\begin{aligned} H &= H_0 + V(t), \\ V(t) &= -\boldsymbol{\mu} \cdot \mathbf{E}(t)e^{\epsilon t}, \end{aligned} \tag{1.55}$$

where H_0 is the molecular Hamiltonian, $\boldsymbol{\mu}$ is the dipole moment and ϵ is an infinitesimal contribution that lead $V(t)$ to zero for $t \to -\infty$. To find a solution for the described problem, it is necessary consider $V(t)$ as a small perturbation and apply the time-dependent perturbation theory. In this framework the N-th order wave function is:

$$|\Psi^N(t)\rangle = \sum_n d_n^N e^{-iE_n t/\hbar} |n\rangle \tag{1.56}$$

The expectation value of dipole moment can be written as a sum of perturbed dipole moments

$$\langle\Psi(t)|\mu|\Psi(t)\rangle = \langle\mu\rangle^0 + \langle\mu\rangle^1 + \langle\mu\rangle^2 + \dots, \tag{1.57}$$

where the perturbed dipole moments are obtained as expectation values of the proper perturbed wavefunctions:

$$\begin{aligned} \langle\mu\rangle^0 &= \langle 0|\mu|0\rangle\,, \\ \langle\mu\rangle^1 &= \langle\Psi^0|\mu|\Psi^1\rangle + \langle\Psi^1|\mu|\Psi^0\rangle\,, \\ \langle\mu\rangle^2 &= \langle\Psi^0|\mu|\Psi^2\rangle + \langle\Psi^1|\mu|\Psi^1\rangle + \langle\Psi^2|\mu|\Psi^0\rangle\,, \end{aligned} \tag{1.58}$$

where $|0\rangle$ is the eigenvector of the time-independent molecular Schrodinger equation, whereas the zero-order, first-order and second-order wave function

expressions are:

$$
\begin{aligned}
|\Psi^0\rangle &= e^{-iE_0 t/\hbar}\,|0\rangle \\
d_m^1(t) &= \frac{1}{h}\sum_{\omega 1}\frac{\langle m|\mu_i|0\rangle\, E_i^{\omega_1}}{\omega_{m0}-\omega_1-i\epsilon}e^{i(\omega_{m0}-\omega_1)t}e^{\epsilon t} \\
d_m^2(t) &= \frac{1}{h^2}\sum_{\omega_1\omega_2}\sum_n \frac{\langle m|\mu_i|n\rangle\,\langle n|\mu_j|0\rangle\, E_i^{\omega_1}E_j^{\omega_2}}{(\omega_{m0}-\omega_1-\omega_2-i2\epsilon)(\omega_{n0}-\omega_2-i\epsilon)}e^{i(\omega_{m0}-\omega_1-\omega_2)t}e^{2\epsilon t}
\end{aligned}
\tag{1.59}
$$

where $E_i^{\omega 1}$ is the i-component of oscillating electric field at ω_1 frequency.

Polarizability

The first-order polarization is obtained from $\langle\mu\rangle^1$ expression by including the zero-order as well as the first-order correction to the wavefunction:

$$
\begin{aligned}
\langle\mu\rangle_i^1 &= \langle\psi^0|\mu|\psi^1\rangle + \langle\psi^1|\mu|\psi^0\rangle \\
&= \sum_{\omega_1}\frac{1}{\hbar}\sum_n\left[\frac{\langle 0|\mu_i|n\rangle\,\langle n|\mu_j|0\rangle}{\omega_{n0}-\omega_1-i\epsilon}+\frac{\langle 0|\mu_j|n\rangle\,\langle n|\mu_i|0\rangle}{\omega_{n0}-\omega_1+i\epsilon}\right]E_\beta^{\omega_1}e^{-i\omega_1 t}e^{\epsilon t}.
\end{aligned}
\tag{1.60}
$$

where it has been assumed that $[E^{\omega_1}]^* = E^{-\omega_1}$. By a direct comparison with Eq.1.54, the QM formula for the linear dipole polarizability can be identify as:

$$
\alpha_{\alpha,\beta}(-\omega;\omega) = \frac{1}{\hbar}\sum_n\left[\frac{\langle 0|\mu_i|n\rangle\,\langle n|\mu_j|0\rangle}{\omega_{n0}-\omega_1-i\epsilon}+\frac{\langle 0|\mu_j|n\rangle\,\langle n|\mu_i|0\rangle}{\omega_{n0}+\omega_1+i\epsilon}\right] \tag{1.61}
$$

It is clear that such a formula can be used directly as it stands for practical computations of the optical polarization once the excitation energies and transition moments of the system are known. For the two terms in the *sum-over-states* expression in Eq.1.61 that involve the ground state (n = 0), the transition frequency ω_{n0} is null. The two terms are of opposite sign and therefore cancel out. The positive infinitesimal ϵ has been added in the perturbation not only to avoid singularities or divergence in the time-integration step of the expansion coefficients d_n above but also to avoid divergences in the linear response function itself. The imaginary term in the denominator of Eq.1.61 can safely be neglected in the calculation of $\alpha(-\omega;\omega)$ for external frequencies far from the excitation energies of the molecular system. Since the set of eigenstates $|n\rangle$ can be chosen as real, without loss of

generality, the linear response function is real in the non-resonant region. The working formula will in this case take the form:

$$\alpha_{\alpha,\beta}(-\omega;\omega) = \frac{1}{\hbar}\sum_{n\neq 0}\left[\frac{\langle 0|\mu_i|n\rangle\langle n|\mu_j|0\rangle}{\omega_{n0}-\omega_1} + \frac{\langle 0|\mu_j|n\rangle\langle n|\mu_i|0\rangle}{\omega_{n0}+\omega_1}\right]. \tag{1.62}$$

First Hyperpolarizability

The explicit formula for the linear polarizability is identified from the linear polarization. In the same way, the corresponding formula for the first hyperpolarizability can be derived from the second-order polarization equation. The starting point is the $\langle\mu\rangle^2$ expression presented above 1.58. By inserting the first-order (ψ^1) and the second-order (ψ^2) correction to the wavefunction, three terms can be recognized:

$$\begin{aligned}
\langle\psi^0|\mu|\psi^2\rangle &= \\
&= \sum_{\omega_1\omega_2}\frac{1}{\hbar^2}\sum_{np}\frac{\langle 0|\mu_i|n\rangle\langle n|\mu_j|p\rangle\langle p|\mu_k|0\rangle E_j^{\omega_1}E_k^{\omega_2}}{(\omega_{n0}-\omega_1-\omega_2-i2\epsilon)(\omega_{p0}-\omega_2-i\epsilon)}e^{-i(\omega_1+\omega_2)t}e^{2\epsilon t} \\
\langle\psi^1|\mu|\psi^1\rangle &= \\
&= \sum_{\omega_1\omega_2}\frac{1}{\hbar^2}\sum_{np}\frac{\langle 0|\mu_j|n\rangle\langle n|\mu_i|p\rangle\langle p|\mu_k|0\rangle E_j^{\omega_1}E_k^{\omega_2}}{(\omega_{n0}+\omega_1+i\epsilon)(\omega_{p0}-\omega_2-i\epsilon)}e^{-i(\omega_1+\omega_2)t}e^{2\epsilon t} \\
\langle\psi^2|\mu|\psi^0\rangle &= \\
&= \sum_{\omega_1\omega_2}\frac{1}{\hbar^2}\sum_{np}\frac{\langle 0|\mu_k|n\rangle\langle n|\mu_j|p\rangle\langle p|\mu_i|0\rangle E_j^{\omega_1}E_k^{\omega_2}}{(\omega_{n0}+\omega_1+\omega_2+i\epsilon)(\omega_{p0}+\omega_2+i\epsilon)}e^{-i(\omega_1+\omega_2)t}e^{2\epsilon t},
\end{aligned} \tag{1.63}$$

where, again, $[E^\omega]^* = E^{-\omega}$ equality has been adopted. Considering that j and k are merely summation indices, by the use of the Einstein summation convention, Eqs. 1.63 can be forced to be intrinsically symmetric without altering the physical polarization operating with $\frac{1}{2}\sum P_{1,2}$ which performs the summation of terms obtained by permuting the pairs (j,ω_1) and (k,ω_2). By introducing the operator $P_{1,2}$, which permutes the couples (j,ω_1) and (k,ω_2),

the second-order polarization becomes:

$$\begin{aligned}\langle\mu\rangle^2 &= \langle\psi^0|\mu|\psi^2\rangle + \langle\psi^1|\mu|\psi^1\rangle + \langle\psi^2|\mu|\psi^0\rangle \\ &= \frac{1}{2}\sum_{\omega_1\omega_2}\frac{1}{\hbar^2}\sum P_{1,2}\sum_{np}\left[\frac{\langle 0|\mu_i|n\rangle\langle n|\mu_j|p\rangle\langle p|\mu_k|0\rangle}{(\omega_{n0}-\omega_1-\omega_2-i2\epsilon)(\omega_{p0}-\omega_2-i\epsilon)}\right. \\ &+ \frac{\langle 0|\mu_j|n\rangle\langle n|\mu_i|p\rangle\langle p|\mu_k|0\rangle}{(\omega_{n0}+\omega_1+i\epsilon)(\omega_{p0}-\omega_2-i\epsilon)} \\ &+ \left.\frac{\langle 0|\mu_k|p\rangle\langle p|\mu_j|n\rangle\langle n|\mu_i|0\rangle}{(\omega_{n0}+\omega_1+\omega_2+i2\epsilon)(\omega_{p0}+\omega_2+i\epsilon)}\right]E_j^{\omega_1}E_k^{\omega_2}e^{-i(\omega_1+\omega_2)t}e^{2\epsilon t}\end{aligned} \tag{1.64}$$

and, by comparison with Eq.1.54, the expression for the first hyperpolarizability can be identified as:

$$\begin{aligned}\beta_{\alpha\beta\gamma}(-\omega_\sigma;\omega_1,\omega_2) = \frac{1}{\hbar^2}\sum P_{1,2}\sum_{np}&\left[\frac{\langle 0|\mu_i|n\rangle\langle n|\mu_j|p\rangle\langle p|\mu_k|0\rangle}{(\omega_{n0}-\omega_\sigma-i2\epsilon)(\omega_{p0}-\omega_2-i\epsilon)}\right. \\ &+ \frac{\langle 0|\mu_j|n\rangle\langle n|\mu_i|p\rangle\langle p|\mu_k|0\rangle}{(\omega_{n0}+\omega_1+i\epsilon)(\omega_{p0}-\omega_2-i\epsilon)} \\ &+ \left.\frac{\langle 0|\mu_k|p\rangle\langle p|\mu_j|n\rangle\langle n|\mu_i|0\rangle}{(\omega_{n0}+\omega_\sigma+i2\epsilon)(\omega_{p0}+\omega_2+i\epsilon)}\right].\end{aligned} \tag{1.65}$$

Analogously with the linear response, if all incidental frequencies (ω_1, ω_2) are non-resonant then the first hyperpolarizability will be real, which is equivalent to associate $\epsilon = 0$ in Eq.1.65 Therefore, the expression for the first hyperpolarizability can be written in a more compact form:

$$\beta_{\alpha\beta\gamma}(-\omega_\sigma;\omega_1,\omega_2) = \frac{1}{\hbar^2}\sum P_{-\sigma,1,2}\sum_{np}\frac{\langle 0|\mu_i|n\rangle\langle n|\mu_j|p\rangle\langle p|\mu_k|0\rangle}{(\omega_{n0}-\omega_\sigma)(\omega_{p0}-\omega_2)}. \tag{1.66}$$

where $P_{-\sigma,1,2}$ denote the sum of the six terms one gets by permuting pairs of $(i,-\omega_\sigma)$, (j,ω_1) and (k,ω_2). Similarly to polarizability, we can split Eq.1.66 as:

$$\begin{aligned}\beta_{\alpha\beta\gamma}(-\omega_\sigma;\omega_1,\omega_2) = \frac{1}{\hbar^2}\sum P_{-\sigma,1,2}&\left[\sum_{np\neq 0}\frac{\langle 0|\mu_i|n\rangle\langle n|\mu_j|p\rangle\langle p|\mu_k|0\rangle}{(\omega_{n0}-\omega_\sigma)(\omega_{p0}-\omega_2)}\right. \\ &+ \sum_{p\neq 0}\frac{\langle 0|\mu_i|0\rangle\langle 0|\mu_j|p\rangle\langle p|\mu_k|0\rangle}{-\omega_\sigma(\omega_{p0}-\omega_2)} \\ &+ \sum_{n\neq 0}\frac{\langle 0|\mu_i|n\rangle\langle n|\mu_j|p\rangle\langle 0|\mu_k|0\rangle}{-(\omega_{n0}-\omega_\sigma)\omega_2} \\ &+ \left.\frac{\langle 0|\mu_i|0\rangle\langle 0|\mu_j|0\rangle\langle 0|\mu_k|0\rangle}{\omega_\sigma\omega_2}\right],\end{aligned} \tag{1.67}$$

where

$$\frac{1}{\omega_\sigma \omega_2} = \frac{1}{\omega_1 \omega_2} - \frac{1}{\omega_\sigma \omega_1} \tag{1.68}$$

cancels all the permutations presented in Eq.1.67. Finally, by interchanging $i, -\omega_\sigma$ and j, ω_1 in the second term and k, ω_2 and j, ω_1 in the last contribution of Eq. 1.67, the expression of first hyperpolarizability can be rewritten in a compact way as:

$$\beta_{\alpha\beta\gamma}(-\omega_\sigma; \omega_1, \omega_2) = \frac{1}{\hbar^2} \sum P_{-\omega,1,2} \sum_{n,p\neq 0} \frac{\langle 0|\mu_i|n\rangle \langle n|\bar{\mu_j}|p\rangle \langle p|\mu_k|0\rangle}{(\omega_{n0} - \omega_\sigma)(\omega_{p0} - \omega_2)}. \tag{1.69}$$

where $\bar{\mu} = \mu - \langle 0|\mu|0\rangle$ is the fluctuation dipole moment.

Born Oppenheimer Approximation

In this work, all calculations of nonlinear optical vibrational phenomena have been carried out within the framework of the Born-Oppenheimer (BO) approximation.[222,223] Within the BO approximation, molecular polarizability and first hyperpolarizability (both static and frequency-dependent) can be split in three different terms: the purely electronic contribution of the molecule at equilibrium geometry, the pure vibrational correction and the zero-point vibrational correction.[224–227] The BO approximation is based on the assumption that the motions of nuclei and electrons can be treated separately, and so their wavefunctions. This is a physically-based approximation, considering that nuclei are heavier than electrons. By assuming that the complete set of vibronic product states can be considered as the exact states of the molecular systems, the polarizability ($\alpha(-\omega;\omega)$) can be written as[225–227]

$$\begin{aligned}\alpha_{\alpha\beta}(-\omega;\omega) = \frac{1}{\hbar} \sum_{K,k\neq 0} \Bigg[& \frac{\langle 0,0|\mu_i|K,k\rangle \langle k,K|\mu_j|0,0\rangle}{\omega_{K,k} - \omega} \\ & + \frac{\langle 0,0|\mu_j|K,k\rangle \langle k,K|\mu_i|0,0\rangle}{\omega_{K,k} + \omega} \Bigg],\end{aligned} \tag{1.70}$$

where K runs over electronic states, whereas k runs over vibrational one. $\hbar\omega_{K,k}$ indicates the energy difference between the K, k and $0, 0$ vibronic states. To simplify Eq.1.70, we consider that the energies involved in electronic transitions are much larger than those involved in vibrational excitations. Thus, $\omega_{K,k} \approx$

$\omega_{K,0}$. By then splitting the sum in Eq.1.70 in electronic and nuclear excitations, we obtain:

$$\begin{aligned}\alpha_{\alpha\beta}(-\omega;\omega) = \\ &= \frac{1}{\hbar}\sum_{K\neq 0}\langle 0|\left[\frac{\langle 0|\mu_i|K\rangle\langle K|\mu_j|0\rangle}{\omega_{K0}-\omega}+\frac{\langle 0|\mu_j|K\rangle\langle K|\mu_i|0\rangle}{\omega_{K0}+\omega}\right]|0\rangle \\ &+\frac{1}{\hbar}\sum_{k\neq 0}\left[\frac{\langle 0|\mu_i^{00}|k\rangle\langle k|\mu_j^{00}|0\rangle}{\omega_{k0}-\omega}+\frac{\langle 0|\mu_j^{00}|k\rangle\langle k|\mu_i^{00}|0\rangle}{\omega_{k0}+\omega}\right],\end{aligned} \tag{1.71}$$

where $\mu^{00} = \langle 0|\mu|0\rangle$. The last two terms represent the pure vibrational contribution.[209,211] The former term in Eq.1.70 is the corresponding electronic contribution to the polarizability averaged over the ground vibrational state. The latter contribution is usually divided into two terms, one corresponding to the purely electronic contribution as obtained at the equilibrium geometry, and one arising from the averaging of the polarizability over the vibrational ground state. Such a term is referred to as the zero-point vibrational averaging correction to the polarizability.

To conclude, under the BO approximation, both the total molecular polarizability and first hyperpolarizability are expressed as a sum of the three aforementioned contributions:

$$\begin{aligned}\alpha &= \alpha^{eq} + \Delta^{ZP}\alpha + \alpha^{v} \\ \beta &= \beta^{eq} + \Delta^{ZP}\beta + \beta^{v}\end{aligned} \tag{1.72}$$

where eq is the electronic contribution to the property calculated at the equilibrium geometry of the molecule, Δ^{ZP} the zero-point vibrational averaging contribution, and v the pure vibrational contribution.

The largest contribution to the total property is the electronic term (α^{eq} or β^{eq}), and generally, both the pure vibrational and the zero-point corrections are neglected, although they may also play an important role when comparing computed data with experiments.[224] Therefore, in the following, we focus on the computational techniques exploited to obtain the α^{eq} and β^{eq}. In particular, in this thesis, polarizabilities and first-Hyperpolarizabilities of selected organic dyes dissolved in aqueous solutions have been calculated. To this end, QM/FQ and QM/FQFμ approaches coupled to the quantum-Pauli Repulsion method described above have been exploited.[133]

By resorting to an SCF description of the QM portion, the two investigated properties have been obtained by exploiting the response theory formalism, because sum-over-state expressions are practically unfeasible from the computational point of view. In this framework, dynamic polarizability ($\alpha(-\omega;\omega)$) and first hyperpolarizability ($\beta(-2\omega;\omega,\omega)$) are calculated as:[228,229]

$$\begin{aligned}\alpha(-\omega;\omega) &= \mathrm{Tr}\boldsymbol{\mu}\mathbf{P}^{(1)} \\ \beta(-2\omega;\omega,\omega) &= 2\mathrm{Tr}\boldsymbol{\mu}\mathbf{P}^{(2)}\end{aligned} \tag{1.73}$$

where $\boldsymbol{\mu}$ is the electric dipole moment integral matrix and $\mathbf{P}^{(1)}$ and $\mathbf{P}^{(2)}$ are the first and second-order perturbed density matrices respectively. A generic second-order density matrix is obtained by solving perturbed equations up to the second order; however, when only one dynamic perturbation is involved, it is possible to avoid the solution of the second-order coupled perturbed equations by using an iterative procedure to reconstruct the density matrix.[228–230] Thus, only the first-order perturbed density matrix $\mathbf{P}^{(1)}$ is constructed, by resorting to linear response theory and by solving the first-order CPHF/CPKS equations.[178]

To calculate the first order perturbed density matrices, the first-order Coupled Perturbed Hartree-Fock or Kohn-Sham (CPHF-CPKS) equations need to be solved:

$$\left[\begin{pmatrix} \tilde{\mathbf{A}} & \tilde{\mathbf{B}} \\ \tilde{\mathbf{B}}^* & \tilde{\mathbf{A}}^* \end{pmatrix} - \omega \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix}\right] \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = - \begin{pmatrix} \mathbf{Q}_X \\ \mathbf{Q}_Y \end{pmatrix} \tag{1.74}$$

where, $\mathbf{1}$ is the identity matrix, $(\mathbf{X},\mathbf{Y})$ is the linear response vector (from which the perturbed density matrix is computed) and $(\mathbf{Q_X},\mathbf{Q_Y})$ is the perturbed gradient vector. $\tilde{A}$ and $\tilde{B}$ matrices are defined, in a general DFT/MM framework, as:[231]

$$\begin{aligned}\tilde{A}_{ai,bj} &= (\epsilon_a - \epsilon_i)\delta_{ab}\delta_{ij} + (ai|bj) + f^{xc}_{ai,bj} + C^{MM}_{ai,bj} \\ \tilde{B}_{ai,bj} &= (ai|bj) + C^{MM}_{ai,bj}\end{aligned} \tag{1.75}$$

where $(ai|bj)$ are two Coulomb integrals and ϵ are the molecular orbital energies, and f^{xc} is the DFT XC contribution. C^{MM} instead defined the QM/MM contribution, and its expression depends on the exploited embedding method, and it is different from zero for polarizable embedding

models only. For QM/FQ, it reads:[232]

$$C^{FQ}_{ai,bj} = \sum_{P}^{N_q} \left(\int_{R^3} \phi_a(\mathbf{r}) \frac{1}{|\mathbf{r} - \mathbf{r}_p|} \phi_i(\mathbf{r}) \mathrm{d}\mathbf{r} \right) \cdot q^T_P(\phi_b, \phi_i), \tag{1.76}$$

where q^T_P are the perturbed fluctuating charges adjusted to the transition density $\mathbf{P}^T_K = \mathbf{X}_K + \mathbf{Y}_K$. For QM/FQF$\mu$ an extra term due to the presence of fluctuating dipoles appears in C^{MM}:[232]

$$\begin{aligned} C^{FQ}_{ai,bj} = & \sum_{P}^{N_q} \left(\int_{R^3} \phi_a(\mathbf{r}) \frac{1}{|\mathbf{r} - \mathbf{r}_p|} \phi_i(\mathbf{r}) \mathrm{d}\mathbf{r} \right) \cdot q^T_P(\phi_b, \phi_i) \\ & - \sum_{P}^{N_q} \left(\int_{R^3} \phi_a(\mathbf{r}) \frac{(\mathbf{r} - \mathbf{r}_p)}{|\mathbf{r} - \mathbf{r}_p|^3} \phi_i(\mathbf{r}) \mathrm{d}\mathbf{r} \right) \cdot \boldsymbol{\mu}^T_P(\phi_b, \phi_i). \end{aligned} \tag{1.77}$$

Perturbed charges (q^T) and perturbed dipoles ($\boldsymbol{\mu}^T$) are obtained trough the system of equations:

$$\begin{pmatrix} \mathbf{T}^{qq} & \mathbf{1}_\lambda & \mathbf{T}^{q\mu} \\ \mathbf{1}^\dagger_\lambda & \mathbf{0} & \mathbf{0} \\ -\mathbf{T}^{q\mu^\dagger} & \mathbf{0} & \mathbf{T}^{\mu\mu} \end{pmatrix} \begin{pmatrix} \mathbf{q} \\ \boldsymbol{\lambda} \\ \boldsymbol{\mu} \end{pmatrix} = \begin{pmatrix} -\mathbf{V}(\mathbf{P}^T_K) \\ \mathbf{0} \\ \mathbf{E}(\mathbf{P}^T_K), \end{pmatrix} \tag{1.78}$$

where

$$\begin{aligned} V(\mathbf{P}^\mathbf{T}_\mathbf{K}) &= -\sum_{ai} \mathbf{P}^T_{K,ai} \int_{R^3} \phi_a(\mathbf{r}) \frac{1}{|\mathbf{r} - \mathbf{r}_p|} \phi_i(\mathbf{r}) \mathrm{d}\mathbf{r}, \\ E(\mathbf{P}^\mathbf{T}_\mathbf{K}) &= \sum_{ai} \mathbf{P}^T_{K,ai} \int_{R^3} \phi_a(\mathbf{r}) \frac{(\mathbf{r} - \mathbf{r}_p)}{|\mathbf{r} - \mathbf{r}_p|^3} \phi_i(\mathbf{r}) \mathrm{d}\mathbf{r}. \end{aligned} \tag{1.79}$$

The right hand side of Eq.1.78 contains both the electric potential and field due to the perturbed density matrix $\mathbf{P}^T_K$. Once CPHF/KS equations are solved, $\alpha(-\omega;\omega)$ and $\beta(-2\omega;\omega,\omega)$ can be calculated. The electrostatic terms in QM/FQ and QM/FQFμ, affect response properties both directly, through explicit terms in Eq.1.75 and by altering molecular orbitals. As a contrary, the QM/MM Pauli repulsion term instead gives no explicit, direct contribution, thus it is affecting the molecular properties only indirectly, i.e. by modifying the Molecular Orbitals (MOs) of the QM portion; however it is is worth remarking that the impact of QM/MM Pauli repulsion on the QM MOs can be crucial.[232]

1.5.2. Hyperfine Coupling Constant

EPR spectroscopy has emerged as the most effective technique to detect and characterize organic free radicals in different conditions and environments.[233,234] To effectively simulate such a property, the spin Hamiltonian, which describes the interaction between the electron spin (S) and an external magnetic field (B) is usually introduced:

$$H_S = \mu_B \vec{S} \cdot \mathbf{g} \cdot \vec{B} + \frac{1}{\hbar \gamma_I} \vec{S} \cdot \mathbf{A} \cdot \vec{\mu}_I \tag{1.80}$$

where $\mu_B \vec{S} \cdot \mathbf{g} \cdot \vec{B}$ is the Zeeman interaction between the electron spin and the external magnetic field. Such a contribution is written in terms of the Bohr magneton μ_B and $\mathbf{g} = g_e \mathbf{1}_3 + \Delta g_{corr}$ ($g_e = 2.0022319$). Δg_{corr} is a correction introduced to take into account the relativistic mass ($\Delta \mathbf{g}_{RM}$), the gauge first-order corrections ($\Delta \mathbf{g}_C$) and a term arising from the coupling of the orbital Zeeman (OZ) and the spin–orbit coupling (SOC) operator.[235,236] Finally, the hyperfine interaction between **S** and the nuclear spin **I** is defined in terms of the the hyperfine coupling tensor **A** (last term in Eq. 1.80) Such a tensor defined for each nucleus X can be written as:

$$\mathbf{A}(X) = \mathrm{A_X} \mathbf{1}_3 + \mathbf{A}_{\mathrm{dip}}(\mathrm{X}) \tag{1.81}$$

where $\mathbf{A}_{\mathrm{dip}}(\mathrm{X})$ is the dipolar term (a zero-trace tensor). Interestingly, $\mathbf{A}_{\mathrm{dip}}(\mathrm{X}) = 0$ in isotropic media (e.g. solutions). $\mathrm{A_X}$ is the Fermi-contact interaction, an isotropic contribution also known as hyperfine coupling constant (hcc). Its expression is given in terms of the spin density (ρ_X) at nucleus X:

$$\mathrm{A_X} = \frac{4\pi}{3} \mu_B \mu_X g_e g_X \left\langle S_Z \right\rangle^{-1} \rho_X^{\alpha-\beta} \tag{1.82}$$

where $\rho_X^{\alpha-\beta}$ can be obtained as:

$$\rho_X^{\alpha-\beta} = \sum_{\mu\nu} P_{\mu\nu}^{\alpha-\beta} \left\langle \chi_\mu(\mathbf{r}) | \delta(\mathbf{r} - \mathbf{r}_X) | \chi_\nu(\mathbf{r}) \right\rangle \tag{1.83}$$

$P^{\alpha-\beta}$ is the difference between α and β density matrices. In this framework, hccs have been evaluated by exploiting UMLHF and UMLDFT, as coupled with an additional layer treated at the MM level. Obviously, the

UMLHF(DFT) hccs are computed for the active part only. Thus, the difference between α and β density matrices in Eq. 1.83 only refers to active spin-density matrices.

CHAPTER 2

Overview of the Attached Papers

In **Paper 1**, a computational study of polarizabilities and first hyperpolarizabilities of different organic acids in aqueous solutions is reported. In particular, we focus on solute-water interactions, and how they affect solute's linear and non-linear electric response properties. To this end, the polarizable QM/FQ and QM/FQFμ approaches are used. In particular, as amply explained in the previous Chapter, both charges and dipoles dynamically respond to the solute's QM electron density, and viceversa, thus introducing mutual polarization effects. In addition, we also considered quantum confinement effects treated by endowing solvent molecules with a parametric electron density, which exerts Pauli repulsion forces upon the solute. Through the application and testing on a set of aromatic molecules in solution we show that, for both polarizabilities and first hyperpolarizabilities, the resulting values arise from a delicate balance between electrostatics, hydrogen-bonding, and non-electrostatic solute solvent interactions.

Paper 2 introduces a novel density matrix-based multilevel approach within the framework of density functional theory (DFT). In the resulting multilevel DFT, the system is partitioned into an active and an inactive fragment, and all interactions between the two parts are retained at the DFT level. In particular, electrostatics, polarization and Pauli repulsion interactions are quantum-mechanically described. Thus, differently from the method proposed in **Paper 1**, there is no need of specific parametrization (for χ, η (and α) for FQ(Fμ)). The partitioning of the total system into the two parts

is performed upon the density matrix, differently from alternative quantum embedding approaches based on DFT, such as FDE. The orthogonality between the two parts is maintained by solving the Kohn-Sham equations in the MO basis for the active part only, while keeping the inactive density matrix frozen. This results in intrinsic reduction of the computational cost. To show the quality of the proposed method, application to aqueous solutions of methyloxirane and glycidol is proposed.

In **Paper 3** a novel multiscale approach is presented, which can describe the electronic structure of open-shell molecular systems embedded in the external environment. This method is based on the coupling of multilevel Hartree-Fock (MLHF) and Density Functional Theory (MLDFT), suitably extended to the unrestricted formalism, to Molecular Mechanics (MM) force fields (FF). As for its restricted counterpart (MLDFT), the system is divided into active and inactive parts, thus describing the most relevant interactions (electrostatic, polarization and Pauli repulsion) at the quantum level. The partitioning isin this case performed upon the α and β density matrices, and orthogonality is maintained within each spin-density. The Kohn-Sham equations are solved in the MO basis for the active part only, while keeping the inactive α and β density matrices frozen. A surrounding MM part, which is formulated in terms of non-polarizable or polarizable FQ, is included. The polarizable FQ part in this case responds to the whole density of the UMLDFT portion, i.e. active + inactive densities. In this way, a physically consistent treatment of long-range electrostatics and polarization effects is introduced. Finally, the approach is extended to the calculation of hyperfine coupling constants of nitroxyl radicals in aqueous solution, which are used as spin-probes in many applications due to their highly sensitivity to the external environment.

CHAPTER 3

Calculation of Linear and Non-linear Electric Response Properties of Systems in Aqueous Solution: A Polarizable Quantum/Classical Approach with Quantum Repulsion Effects

JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

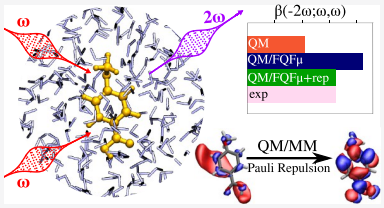
Calculation of Linear and Non-linear Electric Response Properties of Systems in Aqueous Solution: A Polarizable Quantum/Classical Approach with Quantum Repulsion Effects

Gioia Marrazzini, Tommaso Giovannini, Franco Egidi,* and Chiara Cappelli*

Cite This: J. Chem. Theory Comput. 2020, 16, 6993−7004 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: We present a computational study of polarizabilities and hyperpolarizabilities of organic molecules in aqueous solutions, focusing on solute−water interactions and the way they affect a molecule's linear and non-linear electric response properties. We employ a polarizable quantum mechanics/molecular mechanics (QM/MM) computational model that treats the solute at the QM level while the solvent is treated classically using a force field that includes polarizable charges and dipoles, which dynamically respond to the solute's quantum-mechanical electron density. Quantum confinement effects are also treated by means of a recently implemented method that endows solvent molecules with a parametric electron density, which exerts Pauli repulsion forces upon the solute. By applying the method to a set of aromatic molecules in solution we show that, for both polarizabilities and first hyperpolarizabilities, observed solution values are the result of a delicate balance between electrostatics, hydrogen-bonding, and non-electrostatic solute solvent interactions.



INTRODUCTION

The investigation of non-linear optical properties of molecular systems has for long been of particular interest owing to the peculiar optical behavior of materials that possess a high non-linear response, which have found applications in fields such as signal processing and telecommunications.[1] In parallel with experimental advances, a significant amount of effort has been devoted to the development of computational protocols to aid in both predicting and rationalizing the non-linear optical response a molecule or material in the condensed phase.

In fact, the problem of accurately simulating electric response properties of molecular systems in solution has been the object of many studies over the years, with research effort focusing on increasing the accuracy of the quantum mechanics (QM) methods employed for the simulation of the light−matter interaction, which is at the origin of the response, as well as investigating different strategies to incorporate environmental effects into the calculation, particularly in the case of molecules in liquid solutions.[2−9]

Ab initio calculations typically rely on a choice of a model to treat electron-correlation effects coupled to a suitable basis set, and different levels of theory have been explored in the literature.[10−24] The electronic component alone is sometimes not enough to properly reproduce both the linear and non-linear optical response of molecules, and vibrational effects can be quite relevant. Several studies have delved into this problem and offered computationally efficient solutions.[5,25−27] When it comes to the modeling of environmental properties, the literature has mostly focused on ways to model the purely electrostatic component of the solute−solvent interaction, both to produce general solvation models, and as it pertains to the calculation of linear and non-linear optical properties themselves.[28−35]

Because electrostatic interactions are long-range, an atomistic description of the solvent that properly accounts for the effect upon the solute has to include a large number of molecules. This fact, combined with the large configurational space of the solute−solvent system that should be sampled, makes a fully quantum-mechanical description computationally prohibitive. Mixed quantum-classical focused models that treat the solute quantum-mechanically while resorting to a classical description of the solvent, which can be treated as either a continuum or by preserving the atomistic detail and describing the latter using molecular mechanics (MM) models, are a suitable alternative.[36−39] In the most basic formulation, QM/MM models only account for the electrostatic solute−solvent interaction, modeling the solvent by means of fixed charges.[37] However, solvent polarization effects are crucial, especially if one is interested in linear and non-linear optical properties,[40−45] because otherwise the solvent remains insensitive to the polarization effects induced upon the molecule by the

Received: June 29, 2020
Published: October 15, 2020

https://dx.doi.org/10.1021/acs.jctc.0c00674
J. Chem. Theory Comput. 2020, 16, 6993−7004

probing electric field. Polarizable embedding methods establish a mutual polarization between the QM solute and its environment, and the solute−solvent interactions directly affect the former's response properties.[41−43,46−49]

In recent years, we have implemented a polarizable QM/MM method that endows solvent atoms with charges (FQ) and possibly dipoles (Fμ) that are allowed to fluctuate in response to the solute's electrostatic potential.[42,46,50,51] We have shown how the model can have tremendous success in describing a wide array of spectroscopic properties of molecules in water, a highly polar solvent that can form hydrogen bonds with the solute. The properties we have studied include Raman spectroscopy and Raman optical activity,[51,53] electronic and vibrational absorption and circular dichroism,[53−56] two-photon absorption,[57] optical rotation,[58,59] and electronic paramagnetic resonance.[60] The model describes electrostatic interactions through its fluctuating charges and dipoles that dynamically respond to changes in the solute's electronic density and has recently been extended to the treatment of non-electrostatic dispersion and repulsion effects.[60−62] These effects can be critical in determining linear and non-linear electronic properties of a system.[63] The quantum repulsion exerted by the solvent upon the solute's electron density, in particular, has the effect of confining it within the cavity occupied by the solute and is therefore expected to reduce the latter's polarizability and hyperpolarizability. Commonly employed solvation models, including the popular polarizable continuum model (PCM)[36] only account for solute−solvent electrostatics and therefore are missing any confinement effect due to repulsion forces. Note that alternative embedding methods that treat some solvent molecules quantum-mechanically can include repulsion effects naturally through the quantum treatment. These methods often include a classical solvent layer, resulting in a QM/QM/MM paradigm. The QM/FQ and related paradigms, however, find their strength in being "focused" models, where only the properties of the solute and solute−solvent interactions are accurately treated, while the properties of the solvent itself are not of interest, which helps limit the computational cost.

For these reasons, electrostatic, polarization, and quantum repulsion effects are all expected to be particularly relevant in the case of non-linear electric response properties, and it is therefore worth exploring the importance of these effects on model systems, both to confirm these intuitions and highlight the shortcomings in standard calculations based on environmental models, which often neglect one or more of these effects, as well as the magnitude of the errors that would be committed. To this end, we show how different solvation forces contribute to the overall linear and non-linear optical response on a set of six aromatic molecules in solution by employing different electrostatic models based on the QM/FQ(Fμ) paradigm, further enriched by the inclusion of repulsion forces. This is the first time this solvation model is applied to non-linear optical response properties. We show that repulsion forces can indeed be just as important, if not even more so, to the determination of a solute's (hyper)-polarizability as electrostatic interactions, even for a solvent as polar as water. In the next section, the theoretical model is briefly recalled in its various components followed by a description of the computational protocol and the analysis of the results. A summary of the work and future perspectives conclude the manuscript.

THEORETICAL BACKGROUND

Molecular polarizabilities and hyperpolarizabilities can be related to the microscopic response of a molecular system to an external electric field $\mathbf{E}(t)$, represented by an induced dipole moment $\mu(t)$:

$$\mathbf{E}(t) = \frac{1}{2}(\tilde{E}e^{-i\omega t} + \tilde{E}^{*}e^{i\omega t}) \tag{1}$$

$$\boldsymbol{\mu}(t) = \boldsymbol{\mu}^{0} + \boldsymbol{\mu}^{\omega}\cos(\omega t) + \boldsymbol{\mu}^{2\omega}\cos(2\omega t) + \Delta \tag{2}$$

where ω is the frequency of the monochromatic incident light, and $\tilde{E}$ is the complex constant amplitude of the field. The Fourier amplitude in eq 2 can be rewritten as a Taylor expansion with respect to the external electric field.[64] In particular, second harmonic generation (SHG), i.e., the generation of a photon at 2ω as a result of the interaction with an incident ω photon reads:[64]

$$\boldsymbol{\mu}^{2\omega} = \frac{1}{4}\boldsymbol{\beta}(-2\omega;\, \omega,\, \omega)\colon \mathbf{E}^{\omega}\mathbf{E}^{\omega} \tag{3}$$

The first hyperpolarizability β is a third-rank tensor that can be described by a $3 \times 3 \times 3$ matrix, whose 27 components are not independent and can be reduced assuming Kleinman's symmetry.[65]

By exploiting the response theory formalism, the first-order hyperpolarizability $\boldsymbol{\beta}(-2\omega;\, \omega,\, \omega)$ can be calculated as[66,67]

$$\boldsymbol{\beta}(-2\omega;\, \omega,\, \omega) = 2\ \mathrm{tr}\,\boldsymbol{\mu}\mathbf{P}^{(2)} \tag{4}$$

where μ is the electric dipole moment integral matrix and $\mathbf{P}^{(2)}$ is the second-order density matrix. A generic second-order density matrix is obtained by solving perturbed equations up to the second order; however, when only one dynamic perturbation is involved, it is possible to avoid the solution of the second-order coupled perturbed equations by using an iterative procedure to reconstruct the density matrix.[66−68]

Hyperpolarizabilities produced by QM calculations are three-indices tensor quantities. Any meaningful comparison between calculated and experimental data must refer to certain rotational invariants that can be obtained from the full tensor, depending on the specifics of the experimental setup one wishes to reproduce. In this work, we compare our results with those obtained from hyper-Rayleigh scattering (HRS)[69,70] experiments presented in ref 71. In that work, a comparison between computed and experimental results was done by referring to the following quantity:

$$|\beta| = \sqrt{\sum_{i}\left(\sum_{k}(\beta_{ikk} + \beta_{kik} + \beta_{kki})\right)^{2}} \tag{5}$$

Therefore, we refer to the same quantity for the sake of comparison between calculated and experimental data, as was also done in a previous work.[72] However, it is worth noticing that alternative definitions for HRS values have been proposed in the literature, giving computed results directly comparable with experimental data.[20,70,73]

In the following, within tables and figures, we use the notation $\beta(-2\omega;\, \omega,\, \omega)$ in order to emphasize the particular type of frequency dependence; however, note that the presented values always refer to eq 5.

Molecules in solution interact dynamically with the solvent through both electrostatic and non-electrostatic forces. The solute−solvent interaction energy depends on the solute's

https://dx.doi.org/10.1021/acs.jctc.0c00674
J. Chem. Theory Comput. 2020, 16, 6993−7004

electronic density, which is affected by the probing electromagnetic field. Therefore, an embedding model that seeks to capture solvation effects upon a measured linear and non-linear electric response property should take the dynamical aspects of the mutual solute−solvent interaction into account. In this work, we employ the fully atomistic QM/FQ and QM/FQFμ models to describe the electrostatic interactions between the solute and solvent, while resorting a recently implemented model to account for Pauli repulsion effects, the details of which are recalled in the following section.

Solvation Model. As explained above, in this work, we are adopting a multiscale QM/MM approach to describe solvent effects on a QM solute. In particular, the interaction energy $E^{\mathrm{int}}_{\mathrm{QM/MM}}$ between the QM and MM layers is formulated as

$$E^{\mathrm{int}}_{\mathrm{QM/MM}} = E^{\mathrm{ele}}_{\mathrm{QM/MM}} + E^{\mathrm{pol}}_{\mathrm{QM/MM}} + E^{\mathrm{rep}}_{\mathrm{QM/MM}} \tag{6}$$

where $E^{\mathrm{ele}}_{\mathrm{QM/MM}}$ and $E^{\mathrm{pol}}_{\mathrm{QM/MM}}$ are the electrostatic and polarization contributions, respectively, whereas the last term $E^{\mathrm{rep}}_{\mathrm{QM/MM}}$ is the Pauli repulsion, which acts a density confinement. It is worth remarking that we are not including any QM/MM dispersion interaction term. Because of the nature of QM/FQ being a focused model, by neglecting dispersion effects, the solute electronic density is not allowed to delocalize toward the solvent. It is however worth remarking that dispersion plays only a minor role in aqueous solutions, although eq 6 can be extended to account for such an interaction,[47,60,62,74] though of course it may be quite relevant for other solvents.

In order to treat the electrostatic QM/MM coupling, two different polarizable QM/MM approaches were considered, namely, QM/FQ[42,46,52,53,57,59] and QM/FQFμ.[50,51,75] In the former, each atom of the MM portion is endowed with a charge (q), which can vary in agreement with the electronegativity equalization principle (EEP), i.e., a charge flow occurs between two atoms at a different chemical potential. FQ force field is defined in terms of two atomic parameters, namely, electronegativity (χ) and chemical hardness (η). The latter (QM/FQFμ) is instead a pragmatical extension of FQ, in which fluctuating atomic dipoles (μ) and fluctuating atomic charges (q) are associated to each MM atom.[50] Charges values are defined by the same charge equilibration as FQ, but their values depend also on the interaction with dipoles. The peculiarity of FQFμ stands in the fact that both FQ's and Fμ's vary according to the electric potential and electric field.

In order to model Pauli repulsion, an approach recently proposed by some of the present authors is used.[60−62] There, each MM molecule is endowed with a set of s-type Gaussian functions, which mimic the presence of a QM density in the MM portion (Pauli repulsion interaction is a purely quantum effect due to Pauli principle). In our approach, the repulsion energy term is written as the opposite of an exchange integral:[63,76,77]

$$E^{\mathrm{rep}}_{\mathrm{QM/MM}} = \frac{1}{2}\int \frac{\mathrm{d}\mathbf{r}_1\,\mathrm{d}\mathbf{r}_2}{r_{12}}\rho_{\mathrm{QM}}(\mathbf{r}_1,\mathbf{r}_2)\rho_{\mathrm{MM}}(\mathbf{r}_2,\mathbf{r}_1) \tag{7}$$

In order to define the density ρ_{MM}, we localize fictitious valence electron pairs for MM molecules in bond and lone pair regions and represent them by s-Gaussian-type functions. The expression for ρ_{MM} becomes

$$\rho_{\mathrm{MM}}(\mathbf{r}_1,\mathbf{r}_2) = \sum_{\mathbf{R}} \xi^2_{\mathbf{R}} \mathrm{e}^{-\beta_{\mathbf{R}}(\mathbf{r}_1-\mathbf{R})^2}\cdot\mathrm{e}^{-\beta_{\mathbf{R}}(\mathbf{r}_2-\mathbf{R})^2} \tag{8}$$

where $\mathbf{R}$ runs over the centers of the Gaussian functions used to represent the fictitious MM electrons. The β and ξ parameters are generally different for lone pairs or bond pairs, their values being adjusted to the specific kind of environment (MM portion) to be modeled. See ref 61 for their definition in the case of the water molecule. By substituting eq 8 in eq 7, the QM/MM repulsion energy reads

$$E^{\mathrm{rep}}_{\mathrm{QM/MM}} = \frac{1}{2}\sum_{\mathbf{R}}\int \frac{\mathrm{d}\mathbf{r}_1\,\mathrm{d}\mathbf{r}_2}{r_{12}}\rho_{\mathrm{QM}}(\mathbf{r}_1,\mathbf{r}_2)\cdot [\xi^2_{\mathbf{R}}\mathrm{e}^{-\beta_{\mathbf{R}}(\mathbf{r}_1-\mathbf{R})^2}\cdot\mathrm{e}^{-\beta_{\mathbf{R}}(\mathbf{r}_2-\mathbf{R})^2}] \tag{9}$$

It is worth noticing that, in this formalism, QM/MM Pauli repulsion energy is calculated as a two-electron integral. Equation 9 is general enough to hold for any kind of MM environment (solvents, proteins, surfaces, etc.). The nature of the external environments is specified by defining the number of different electron-pair types and the corresponding β and ξ parameters in eq 8. Finally, the formalism is general so that it can be coupled to any kind of QM/MM approach.

All of the components of this solvation model require a specific parametrization.

COMPUTATIONAL DETAILS

For this work, we have selected six organic molecules (Figure 1) from ref 72, for which experimental measurements of their first hyperpolarizability values in aqueous solutions exist.[71] All QM and QM/MM calculations were performed using a locally modified version of Gaussian16 computational chemistry package[78] and employed the B3LYP,[79−81] CAM-B3LYP,[82] and M06-2X[83] density functionals in combination with the 6-311++G(d,p) basis set. Polarizable QM/MM calculations were performed with the fluctuating charge model (FQ)[42,46,84−86] with and without fluctuating dipoles (FQFμ).[50] QM/FQ

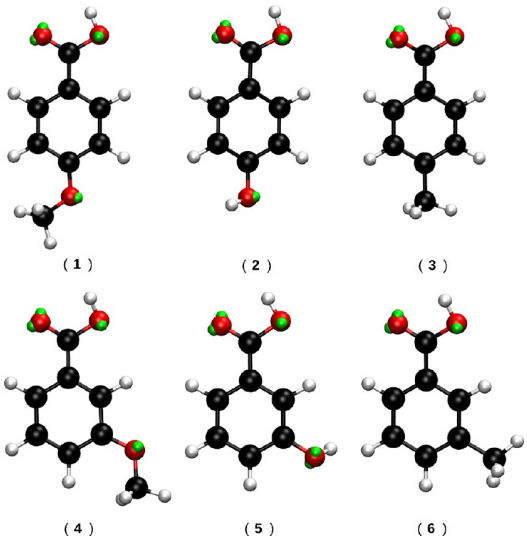


Figure 1. Structures of the molecules studied. The green spheres depicted close to the oxygen atoms represent the virtual sites (VS, *vide infra*).

https://dx.doi.org/10.1021/acs.jctc.0c00674
J. Chem. Theory Comput. 2020, 16, 6993−7004

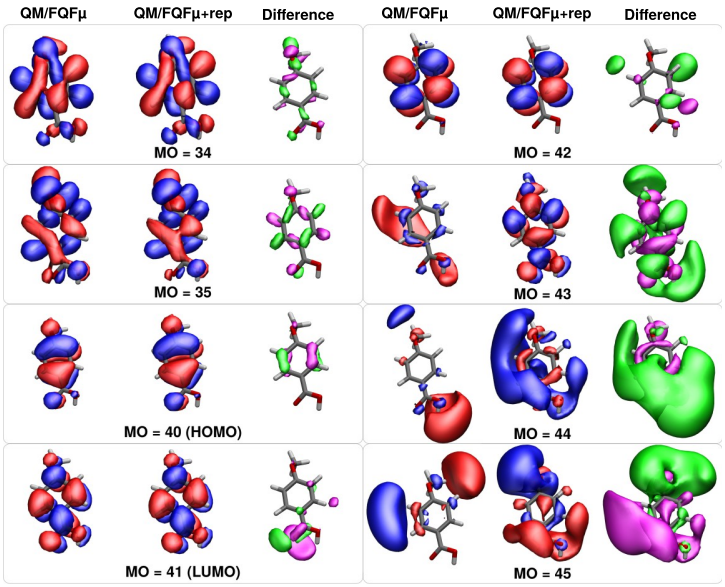


Figure 2. Selected molecule **1** molecular orbitals for a randomly chosen snapshot extracted from the MD simulation. QM/FQFμ and QM/FQFμ+rep orbitals and their difference are depicted.

calculations were performed using two distinct parametrizations, the one by Rick et al.,[84−86] which we here denote as FQ^a, and the one by Giovannini et al.,[60] denoted as FQ^b. Hyperpolarizabilities are reported in esu.[87] In order to compute hyperpolarizabilities with the QM/MM methodology described above, we followed a multistep procedure, which is here summarized:

1. Geometry optimization of the solute molecules. The structure of each system was optimized using the CAM-B3LYP density functional and by including solvent effects by means of the PCM.[88−90]
2. Calculation of atomic charges and definition of virtual sites. From the same CAM-B3LYP/PCM calculations on the optimized structures, we obtained the RESP atomic charges[91−93] and locations for the virtual sites (VS), which model the presence of non-bonding electron pairs. VS have a fixed position with respect to generating atoms and allow us to refine the description of hydrogen-bonding interactions. The positions were obtained by evaluating the centroids of Boys orbitals.[94,95]
3. Classical MD simulations in aqueous solutions. Each solute molecule was placed in a cubic box and then surrounded by water molecules under periodic boundary conditions (PBC). To sample the solute−solvent configuration space, a classical MD simulation on each system was run as detailed in ref 72.
4. Extraction of snapshots from the MD simulation. From each MD run, a total of 200 snapshots was extracted to be used in the QM/MM calculations for each system. For each snapshot, a solute-centered sphere with radius of 15 Å of explicit water molecules was cut.
5. Polarizable QM/MM calculations. The QM/MM calculations of static and dynamic polarizabilities and hyperpolarizabilities were performed on the full set of structures extracted from the MD. The results obtained for each spherical snapshot were extracted and averaged to produce the final value.

NUMERICAL RESULTS

Effect of Repulsion on the MOs. In this section, we wish to provide a more in-depth analysis of the effect of quantum repulsion and how it enters the computational results. As stated earlier, the addition of quantum repulsion affects the molecular orbitals (MO) of the system. This change then propagates to response equations and therefore computed electric response properties. Changes in the MOs caused by repulsion can be appreciated by plotting the matrix **J** that relates one set of MOs into the other:

$$\mathbf{J} = \mathbf{C}_{\mathrm{rep}}^{\dagger}\mathbf{S}\mathbf{C}_{\mathrm{norep}} \tag{10}$$

where $\mathbf{C}_{\mathrm{rep}}$ is the MO coefficient matrix calculated at the QM/FQFμ level with Pauli repulsion, **S** is the atomic orbital overlap matrix, and $\mathbf{C}_{\mathrm{norep}}$ is the MO matrix calculated at the same level without Pauli repulsion.

We performed this analysis for a randomly selected snapshot of the molecule **1**, and the result can be seen in Figure 3 where higher absolute values are represented by a darker square. As expected, occupied orbitals remain mostly unaffected, though this is not true in general (in particular for MO = 34 and MO = 35, which change somewhat, see Figure 2). Many virtual orbitals are instead mixed up, as is evident from Figure 3 and Figure 2. The latter figure shows isovalue plots of selected MOs with and without repulsion as well as the difference in the squared MOs to help visualize the regions of space where changes are most pronounced. In fact, the **J** matrix becomes so sparse in the block involving the first 100 virtual orbitals that it

https://dx.doi.org/10.1021/acs.jctc.0c00674
J. Chem. Theory Comput. 2020, 16, 6993−7004

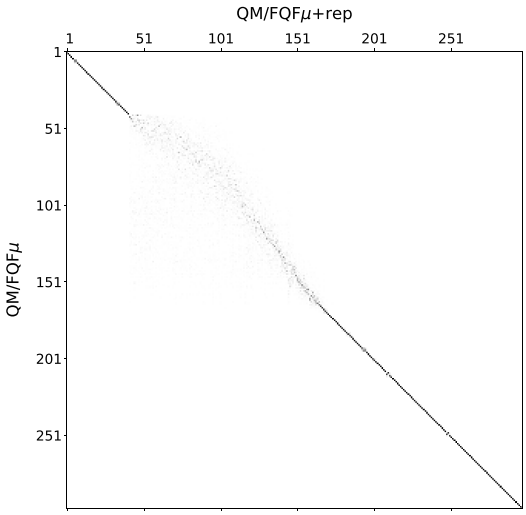


Figure 3. **J** matrix of a randomly snapshot extract from MD simulation (see eq 10)

is barely visible in the figure. This is true up to a point, with very high energy orbitals remaining unaltered.

It is worth investigating whether these changes how much these changes actually affect the density derivatives since they are what actually gives the hyperpolarizabilities according to eq 4. Given the large number of components, we only look at derivatives along the z component of the electric field. Derivatives with respect to the other components can be found in the Supporting Information. The first-order density derivative $\mathbf{P}^{(1)}$ (with respect to an electric perturbation along the z direction) is non-zero only in the occupied-virtual block. The difference between the two blocks (with and without repulsion) is shown on the left panel in Figure 4. Indeed, while differences are generally negligible, some deviations are observed, particularly in the blocks corresponding to the lowest-energy virtual orbitals that are most affected by repulsion. The same analysis can be carried out for the density second derivative $\mathbf{P}^{(2)}$, but this time only the occupied-occupied and virtual-virtual blocks are non-zero. Components belonging to higher energy occupied orbitals show a marked difference, while for virtual orbitals, we can draw a similar conclusion as for $\mathbf{P}^{(1)}$, whereupon only the block involving virtual orbitals that are actually affected by repulsion propagates to density derivatives.

Polarizability. We begin our investigation by studying the effect of water on static and dynamic polarizabilities.

Figure 5 reports the computed values for both the static $\alpha(0; 0)$ and dynamic $\alpha(-\omega; \omega)$ polarizability, evaluated with three different DFT functionals for the isolated and solvated molecules, with and without considering quantum repulsion effects. We start by looking at how a change in the underlying electronic structure model, i.e., the chosen density functional, affects the results, in order to verify that conclusions about solvation effects are consistent and do not depend too much on the functional. It can be immediately seen that the dynamic polarizabilities are substantially higher by about 1.7 units, compared with the static values (see the Supporting Information for tables reporting the numerical values). Solvation electrostatics leads to a significant and uniform increase in the polarizability values for all systems, and the magnitude is rather uniform among the three functionals. It should be noted that the inclusion of repulsion effects into the calculation brings about a significant decrease in the property, by about 8%, and this decrease is actually quite consistent and varies very little among the molecules. Nor are repulsion effects particularly affected by a change in DFT functional, even with the addition of a long-range correction as in CAM-B3LYP. This is not surprising since repulsion effects as modeled in this work directly influence the ground-state density of each

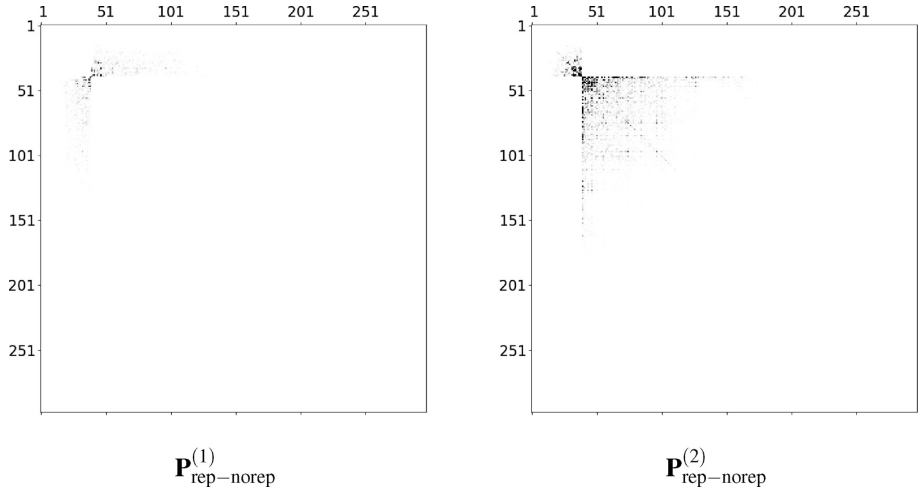


Figure 4. Difference between the density matrix derivatives with and without Pauli repulsion of a randomly snapshot of molecule **1** extracted from MD simulation. The first derivative $\mathbf{P}^{(1)}$ is on the left panel, and the second derivative $\mathbf{P}^{(2)}$ is on the right panel. Derivatives are taken with respect to the z component of the electric field.

https://dx.doi.org/10.1021/acs.jctc.0c00674
J. Chem. Theory Comput. 2020, 16, 6993−7004

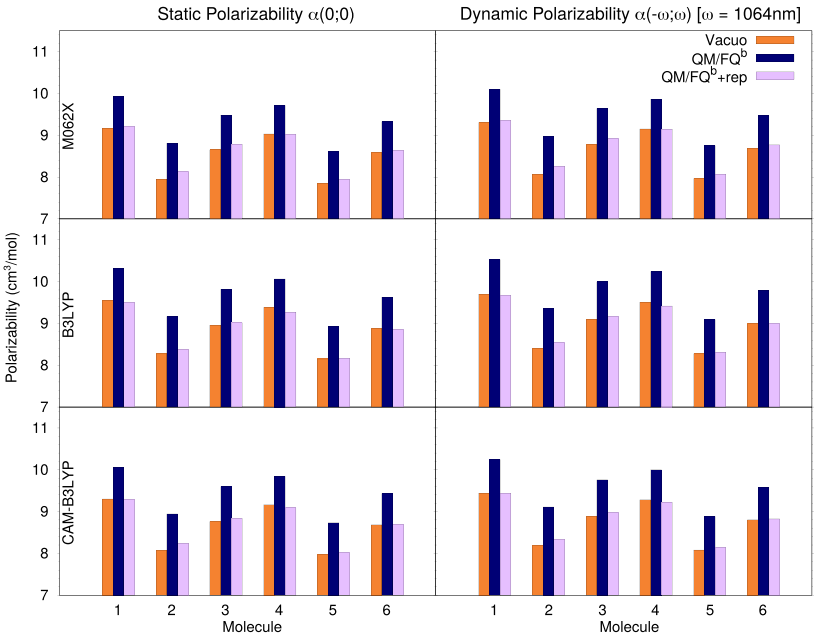


Figure 5. Static (left) and dynamic (right) polarizabilities of molecules **1**−**6** evaluated at 1064 nm in vacuo and in solution (with and without repulsion effects) with three different density functionals: M06-2X (top), B3LYP (middle), and CAM-B3LYP (bottom).

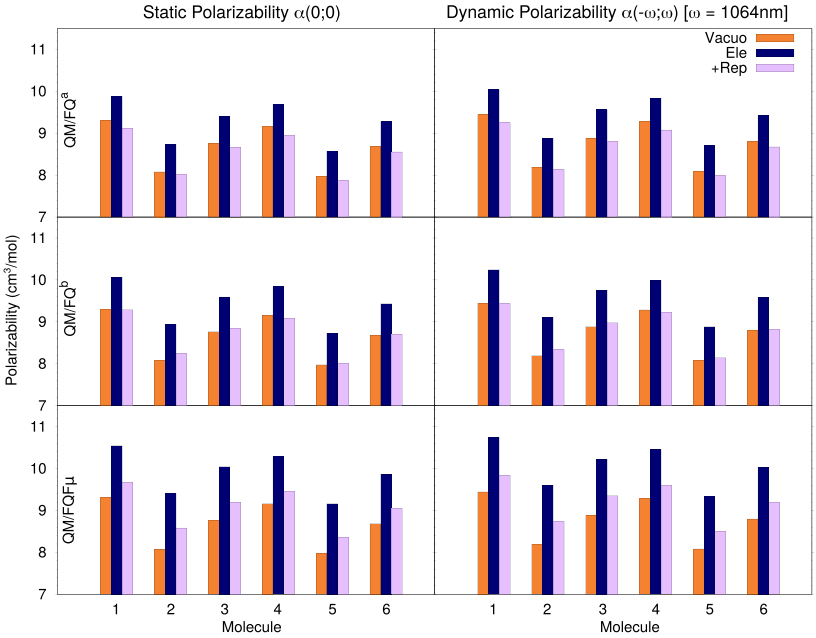


Figure 6. Static (left) and dynamic (right) polarizabilities of molecules **1**−**6** evaluated at 1064 nm in vacuo and in solution (with and without repulsion effects) with three different models for the electrostatic component: FQ^a (top), FQ^b (middle), and FQFμ (bottom).

system, though they do not directly affect the response functions, for which long-range corrections play their most important role.

It is interesting to perform a more in-depth analysis of the different roles of electrostatics and non-electrostatics in determining the polarizability of the solvated systems. As

https://dx.doi.org/10.1021/acs.jctc.0c00674
J. Chem. Theory Comput. 2020, 16, 6993−7004

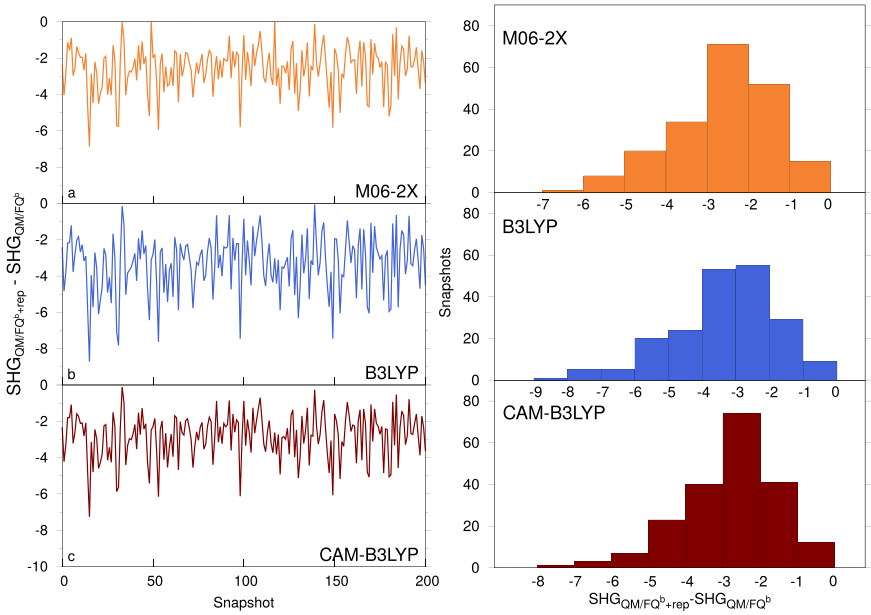


Figure 7. Difference between QM/FQ $\beta(-2\omega; \omega, \omega)$ (in esu) with and without repulsion for molecule **1** calculated for different snapshots extracted from the MD and for different functionals: CAM-B3LYP, B3LYP, and M906-2X. Values are shown both as they vary across the snapshots (left) and as interval distributions (right).

discussed in the Theoretical Background section, there are different sets of parameters to choose from when performing a QM/FQ calculation. Originally, parameters derived by Rick et al.[84] (hereby denoted as FQ^a) were the first to be developed, though they tend to underestimate the solvent polarization. New parameters specifically designed for QM/FQ calculations were recently adopted,[60] which allow for a higher solvent polarization. This may not necessarily result in better agreement with experimental values because a higher solvent polarization tends to have an opposite effect compared to the introduction of repulsion forces; therefore, underestimating solvent electrostatics may lead to a favorable error cancellation whenever repulsion effects are neglected. It is therefore interesting to compare values obtained with the different electrostatic models with and without repulsion effects. It is worth reiterating that QM/FQ results are always averages computed over a large set of snapshots obtained from a classical MD, and in order for the results to be reliable, they must be at convergence with respect to the number of snapshots. In the Supporting Information, we show that our results are indeed at convergence. In Figure 6, we present results obtained with the CAM-B3LYP functional only. Indeed, as is evident from the results, going from FQ^a to FQ^b, which leads to an increase in the electrostatics due to the parametrization, does have an opposite effect with respect to repulsion, though the magnitude is not comparable as the FQ^b parameters lead to computed polarizabilities which are about 2 units higher, whereas the reduction due to repulsion effects is significantly stronger. As mentioned in the Theoretical Background section, the basic FQ model can only account for in-plane polarization of solvent molecules; however, out-of-plane solvent polarization may not in principle be disregarded. The FQFμ model overcomes this limitation. Polarizabilities were therefore also evaluated using this electrostatic model with and without repulsion effects. The increase in polarizability that we observe, when going from the FQ^b to the FQFμ values, is of the same order of magnitude as the difference between the FQ^b values and the gas-phase results. Therefore, out-of-plane polarization effects, which can only be taken into account if solvent molecules are endowed with fluctuating dipoles, should not be neglected. It is interesting to note that, if we compare the FQ^b+rep results with the vacuum values (bottom panels in Figure 6), we see that they are very close to the gas-phase values. If the solution values that include all effects were simply compared to those for the isolated molecules, one might erroneously conclude that solvent effects are negligible. Our results show that the role of solvation in determining a system's polarizability rests on a delicate balance of different effects, none of which can be regarded as negligible; therefore, the use of a solvation model with the capability to include all such effects not only in the description of the system's ground state but also of its response properties is crucial. It should finally be remarked that solvation models that only treat one of these effects might lead to wrong computed values.

First Hyperpolarizabilities. We now move to first hyperpolarizabilities, which, being third-order properties, are expected to be much more sensitive to the polarizable environment of the molecule and thus a better probe for the different solvation effects.

As in the case of polarizabilities, gas-phase values are single-point calculations on the optimized structures while QM/FQ results are averages over the structure extracted from the classical MD.

The solvation effect observed for the average value is the result of changes on each of the extracted MD snapshots.

https://dx.doi.org/10.1021/acs.jctc.0c00674
J. Chem. Theory Comput. 2020, 16, 6993−7004

Before commenting on the averages, we therefore analyze the hyperpolarizability values for all snapshots with the different solvation models. Figure 7 reports the difference between the hyperpolarizability values of molecule **1** calculated with the QM/FQ^b model with and without repulsion for all snapshots. Data are also collected into distribution diagrams. The plots show the range of variability in MD time of the calculated property, which depends on the spatial arrangement of the solvent molecules around the solute as well as its instantaneous conformation. Our dynamical atomistic approach to the solvation phenomenon is able to give insight into such a variability, whereas mean-field approaches would instead focus on the average value only. Though the average effect of the hyperpolarizability is of course functional-dependent, it can be readily seen from the plots that they are highly correlated, i.e., given one snapshot if a high or low repulsion effect is obtained for one functional, a similar result will be observed when using the other two. One thing that stands out is that the effect of repulsion is very dishomogeneous across the snapshots, where some have almost no effect and others presenting a decrease in hyperpolarizability that is almost as high as the average value of the property itself.

Figure 8 reports the average values of the dynamical hyperpolarizabilities $\beta(-2\omega;\omega,\omega)$ computed with three

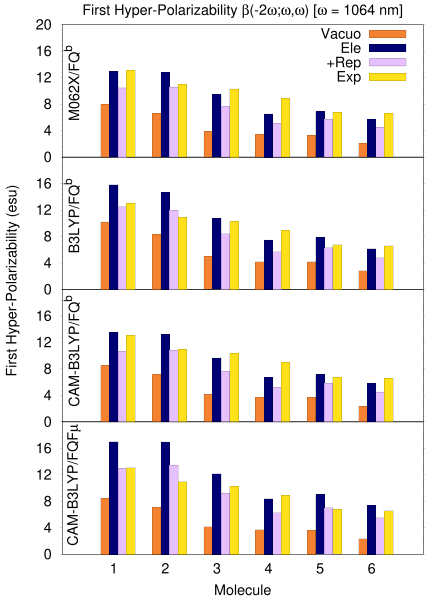


Figure 8. Dynamic hyperpolarizabilities of molecules **1**−**6** evaluated at 1064 nm in vacuo and in solution (with and without repulsion effects) evaluated with different functionals and solvation models. Experimental data from Ray et al.[71]

functionals, with and without quantum repulsion, as well as the experimental values obtained by means of hyper-Rayleigh scattering (HRS) measurements in ref 71. Numerical values for the solvated system are also reported in Table 1 for an easier reading.

Comparing gas-phase values with electrostatics-only solvated values (whether obtained with the FQ^b or the FQFμ model),

Table 1. CAM-B3LYP, B3LYP, and M06-2X with QM/FQ^b Parameters, with and without Repulsion $\beta(-2\omega;\omega,\omega)$ (± Standard Errors, Values in esu)

	$\beta(-2\omega;\omega,\omega)$															
	CAM-B3LYP			B3LYP			M06-2X			EXP[71]	CAM-B3LYP/FQFμ					
	w/o rep	rep	vacuum	w/o rep	rep	vacuum	w/o rep	rep	vacuum		w/o rep	rep	vacuum	w/o rep	rep	vacuum
1	13.45 ± 0.26	10.65 ± 0.19	8.47	15.81 ± 0.33	12.49 ± 0.25	10.19	12.96 ± 0.26	10.44 ± 0.21	7.94	13.06	16.99 ± 0.34	12.99 ± 0.24	10.19	15.81 ± 0.33	12.49 ± 0.25	10.19
2	13.19 ± 0.16	10.81 ± 0.12	7.10	14.70 ± 0.17	11.97 ± 0.13	8.33	12.72 ± 0.16	10.54 ± 0.12	6.58	10.93	16.96 ± 0.23	13.46 ± 0.16	8.33	14.70 ± 0.17	11.97 ± 0.13	8.33
3	9.65 ± 0.14	7.58 ± 0.10	4.15	10.72 ± 0.16	8.40 ± 0.11	4.98	9.44 ± 0.14	7.58 ± 0.11	3.80	10.28	12.17 ± 0.18	9.22 ± 0.12	4.98	10.72 ± 0.16	8.40 ± 0.11	4.98
4	6.76 ± 0.10	5.21 ± 0.07	3.71	7.47 ± 0.12	5.73 ± 0.08	4.18	6.51 ± 0.10	5.13 ± 0.08	3.42	8.91	8.39 ± 0.13	6.25 ± 0.09	4.18	7.47 ± 0.12	5.73 ± 0.08	4.18
5	7.19 ± 0.10	5.78 ± 0.07	3.63	7.91 ± 0.11	6.30 ± 0.07	4.16	6.87 ± 0.09	5.64 ± 0.08	3.28	6.78	9.02 ± 0.12	7.01 ± 0.08	4.16	7.91 ± 0.11	6.30 ± 0.07	4.16
6	5.80 ± 0.09	4.51 ± 0.06	2.35	6.18 ± 0.10	4.78 ±0.07	2.81	5.65 ± 0.09	4.43 ± 0.06	2.10	6.57	7.41 ± 0.11	5.52 ± 0.08	2.81	6.18 ± 0.10	4.78 ±0.07	2.81

https://dx.doi.org/10.1021/acs.jctc.0c00674
J. Chem. Theory Comput. 2020, 16, 6993−7004

we see that, in some cases, the computed property can even double in value. However, as was observed for polarizabilities, repulsion has a the opposite effect; however, in this case, the decrease is much more pronounced, being on average about 20%, compared to 8% of simple polarizabilities. This result emphasizes the important role played by repulsion effects in determining high-order electric properties of systems in the condensed phase and suggests that any quantitative calculation of such properties for systems in solution should not neglect them. The final result is the product of a delicate balance between these opposing effects, though all values in solution are larger than the corresponding gas-phase results. These results speak to a large extent about the fact that one must be careful when evaluating the performance of any solvation model that only accounts for electrostatics, such as plain QM/FQ^b or the popular polarizable continuum model (PCM). Results that are closer to the experiment might be achieved by lowering the solvent's polarization through a careful parametrization of the method, such as an increase in the dimension of the PCM cavity or tinkering with the FQ parameters, though this would only be so because of a fruitful and artificial error cancellation. The compensation between electrostatic and non-electrostatic forces, however, is not consistent across different molecular properties (as can be seen by simply comparing the data in Figure 6 for polarizabilities and Figure 8 for hyperpolarizabilities); therefore, error cancellation will not work for all properties leading to a systematic error in the results.

Finally, we can compare our calculations experimental data. We see that, in some cases, the QM/FQ^b+rep model apparently leads to a greater error compared to the simpler QM/FQ^b purely electrostatic model. This is observed for all systems except for molecule **2** when using the CAM-B3LYP and M06-2X functionals, though not in the case of the B3LYP functional where molecule **1** is also an exception. The inclusion of polarizable dipoles in the solvent's description leads to a further increase in the computed values, as observed in the case of static and dynamics polarizabilities and, with the exception of molecule **2**, produces values that are much closer to their experimental counterparts if repulsion is also included.

CONCLUSIONS AND PERSPECTIVES

In this paper, we have presented a computational study of polarizabilities and hyperpolarizabilities of molecules in aqueous solutions. We dissected the solute−solvent interaction into its electrostatic and non-electrostatic components and then compared computed results with experimental findings to assess the role of each interaction. As a solvation model, we employed our recently developed polarizable QM/MM method based on fluctuating charges and dipoles (FQ and FQFμ) enriched by solute−solvent repulsion effects to the calculation of polarizabilities and hyperpolarizabilities of organic molecules in water. By dissecting the magnitude and role of each component of the solvation phenomenon as it applies to the set of studied systems, we showed that QM/FQ and QM/FQFμ models for solvation electrostatics can be combined with our recently implemented quantum repulsion model to successfully calculate linear and non-linear electric response properties of systems in solution in a "focused" solvation model paradigm. This is possible owing to the model's ability to be extended to high-order properties through the propagation of the solute−solvent interaction terms at all orders of the QM response functions. Our results show that all of the different effects we considered contribute to the computed value in similar measures, meaning that none of them can be safely neglected. In particular, the modeling of electrostatic effects with the FQ method leads to an expected increase in the computed polarizability values compared to the isolated molecule, which is further intensified by the addition of polarizable dipoles in the solute's description. Repulsion has an effect that is similar in magnitude but opposite in sign; therefore, the evaluation of such properties is the result of a delicate balance between all these contrasting forces, which in principle must all be included in the model and treated as accurately as possible. While numerically decent results might be obtained by neglecting repulsion altogether and tinkering with the magnitude of the solvent polarization (or removing it altogether as is done with standard non-polarizable QM/MM methods), this approach should not be regarded as "safe" or generally transferable to a wide array of systems for which the one effect or the other may dominate. Our results therefore underline the complexity of the forces at play within a water solution, which, far from being simply a highly polar substance with the ability to form hydrogen bonds, can influence a solute's properties through effects such as quantum repulsion and electronic polarizability, which can be almost as important as the presence of hydrogen bonds themselves.

This work's results notwithstanding, much work remains to be done in this field. To fully appreciate the improvements offered by such refined models over more standard methodologies, a wide benchmark over a wider set of systems and solvents should be performed to estimate the expected error of the model for a given functional and basis set. In addition, a much wider array of response properties, particularly those involving a magnetic or mixed electric and magnetic response such as nuclear magnetic shields or optical rotatory dispersion should be investigated to fully appreciate the power of the method. Finally, one type of solvent effect that was neglected in this work is that, due to electron dispersion, while it has been shown to be negligible in the case of water,[47,60,62,74] it can be expected to be particularly relevant for solvents such as benzene, and models to include this effect in the evaluation of high-order response properties in an efficient manner should be investigated and will be the object of future work.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.0c00674.

Convergence of QM/MM polarizabilities and first hyperpolarizabilities as a function of the snapshots' number; data related to Figures 5, 6, and 8; and plotted density matrix derivatives (PDF)

AUTHOR INFORMATION

Corresponding Authors

Franco Egidi − *Scuola Normale Superiore, Pisa 56126, Italy;* orcid.org/0000-0003-3259-8863; Email: franco.egidi@sns.it

Chiara Cappelli − *Scuola Normale Superiore, Pisa 56126, Italy;* orcid.org/0000-0002-4872-4505; Email: chiara.cappelli@sns.it

https://dx.doi.org/10.1021/acs.jctc.0c00674
J. Chem. Theory Comput. 2020, 16, 6993−7004

Authors

Gioia Marrazzini − *Scuola Normale Superiore, Pisa 56126, Italy*

Tommaso Giovannini − *Department of Chemistry, Norwegian University of Science and Technology, Trondheim 7491, Norway;* orcid.org/0000-0002-5637-2853

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.0c00674

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are thankful for the computer resources provided by the high performance computer facilities of the SMART Laboratory (http://smart.sns.it/). T.G. acknowledges funding from the Research Council of Norway through its grant TheoLight (grant no. 275506).

■ REFERENCES

(1) Bhawalkar, J. D.; He, G. S.; Prasad, P. N. Nonlinear Multiphoton Processes in Organic and Polymeric Materials. *Rep. Prog. Phys.* **1996**, *59*, 1041−1070.

(2) Cammi, R.; Mennucci, B.; Tomasi, J. On the Calculation of Local Field Factors for Microscopic Static Hyperpolarizabilities of Molecules in Solution with the Aid of Quantum-Mechanical Methods. *J. Phys. Chem. A* **1998**, *102*, 870−875.

(3) Cammi, R.; Mennucci, B.; Tomasi, J. An Attempt To Bridge the Gap between Computation and Experiment for Nonlinear Optical Properties: Macroscopic Susceptibilities in Solution. *J. Phys. Chem. A* **2000**, *104*, 4690−4698.

(4) Rinkevicius, Z.; Li, X.; Sandberg, J. A. R.; Ågren, H. Non-Linear Optical Properties of Molecules in Heterogeneous Environments: a Quadratic Density Functional/Molecular Mechanics Response Theory. *Phys. Chem. Chem. Phys.* **2014**, *16*, 8981−8989.

(5) Egidi, F.; Giovannini, T.; Piccardo, M.; Bloino, J.; Cappelli, C.; Barone, V. Stereoelectronic, Vibrational, and Environmental Contributions to Polarizabilities of Large Molecular Systems: A Feasible Anharmonic Protocol. *J. Chem. Theory Comput.* **2014**, *10*, 2456−2464.

(6) Bartkowiak, W.; Zaleśny, R.; Niewodniczański, W.; Leszczynski, J. Quantum Chemical Calculations of the First- and Second-Order Hyperpolarizabilities of Molecules in Solutions. *J. Phys. Chem. A* **2001**, *105*, 10702−10710.

(7) Benassi, E.; Egidi, F.; Barone, V. General Strategy for Computing Non-Linear Optical Properties of Large Neutral and Cationic Organic Chromophores in Solution. *J. Phys. Chem. B* **2015**, *119*, 3155−3173.

(8) Steinmann, C.; Reinholdt, P.; Nørby, M. S.; Kongsted, J.; Olsen, J. M. H. Response Properties of Embedded Molecules through The Polarizable Embedding Model. *Int. J. Quant. Chem.* **2019**, *119*, No. e25717.

(9) Carlotti, B.; Cesaretti, A.; Cannelli, O.; Giovannini, T.; Cappelli, C.; Bonaccorso, C.; Fortuna, C. G.; Elisei, F.; Spalletti, A. Evaluation of Hyperpolarizability from the Solvatochromic Method: Thiophene Containing Push- Pull Cationic Dyes as a Case Study. *J. Phys. Chem. C* **2018**, *122*, 2285−2296.

(10) Champagne, B.; Botek, E.; Nakano, M.; Nitta, T.; Yamaguchi, K. Basis Set and Electron Correlation Effects on the Polarizability and Second Hyperpolarizability of Model Open-Shell π-Conjugated Systems. *J. Chem. Phys.* **2005**, *122*, 114315.

(11) Norman, P.; Bishop, D. M.; Jensen, H. J. A.; Oddershede, J. Nonlinear response theory with relaxation: The first-order hyperpolarizability. *J. Chem. Phys.* **2005**, *123*, 194103.

(12) Hättig, C.; Cacheiro, J. L.; Fernández, B.; Rizzo, A. *Ab initio* calculation of the refractivity and hyperpolarizability second virial coefficients of neon gas. *Mol. Phys.* **2003**, *101*, 1983−1995.

(13) Luo, Y.; Ågren, H.; Jørgensen, P.; Mikkelsen, K. V. Response theory and calculations of molecular hyperpolarizabilities. In *Advances in quantum chemistry*; Elsevier: 1995; *26*; 165−237.

(14) Limacher, P. A.; Mikkelsen, K. V.; Lüthi, H. P. On the accurate calculation of polarizabilities and second hyperpolarizabilities of polyacetylene oligomer chains using the CAM-B3LYP density functional. *J. Chem. Phys.* **2009**, *130*, 194114.

(15) Fernández, B.; Hättig, C.; Koch, H.; Rizzo, A. *Ab initio* calculation of the frequency-dependent interaction induced hyperpolarizability of Ar_2. *J. Chem. Phys.* **1999**, *110*, 2872−2882.

(16) Alparone, A.; Reis, H.; Papadopoulos, M. G. Theoretical Investigation of the (Hyper)polarizabilities of Pyrrole Homologues C_4H_4XH (X = N, P, As, Sb, Bi). A Coupled-Cluster and Density Functional Theory Study. *J. Phys. Chem. A* **2006**, *110*, 5909−5918.

(17) Isborn, C. M.; Leclercq, A.; Vila, F. D.; Dalton, L. R.; Brédas, J. L.; Eichinger, B. E.; Robinson, B. H. Comparison of Static First Hyperpolarizabilities Calculated with Various Quantum Mechanical Methods. *J. Phys. Chem. A* **2007**, *111*, 1319−1327.

(18) Silva, D. L.; Fonseca, R. D.; Vivas, M. G.; Ishow, E.; Canuto, S.; Mendonca, C. R.; De Boni, L. Experimental and theoretical investigation of the first-order hyperpolarizability of a class of triarylamine derivatives. *J. Chem. Phys.* **2015**, *142*, No. 064312.

(19) Vivas, M. G.; Silva, D. L.; Rodriguez, R. D. F.; Canuto, S.; Malinge, J.; Ishow, E.; Mendonca, C. R.; De Boni, L. Interpreting the First-Order Electronic Hyperpolarizability for a Series of Octupolar Push−Pull Triarylamine Molecules Containing Trifluoromethyl. *J. Phys. Chem. C* **2015**, *119*, 12589−12597.

(20) Plaquet, A.; Guillaume, M.; Champagne, B.; Castet, F.; Ducasse, L.; Pozzo, J.-L.; Rodriguez, V. *In silico* optimization of Merocyanine-Spiropyran Compounds as Second-Order Nonlinear Optical Molecular Switches. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6223−6232.

(21) Dutra, A. S.; Castro, M. A.; Fonseca, T. L.; Fileti, E. E.; Canuto, S. Hyperpolarizabilities of the methanol molecule: A CCSD calculation including vibrational corrections. *J. Chem. Phys.* **2010**, *132*, No. 034307.

(22) Sałek, P.; Helgaker, T.; Vahtras, O.; Ågren, H.; Jonsson, D.; Gauss, J. A comparison of density-functional-theory and coupled-cluster frequency-dependent polarizabilities and hyperpolarizabilities. *Mol. Phys.* **2005**, *103*, 439−450.

(23) Pawłowski, F.; Jørgensen, P.; Hättig, C. The hyperpolarizability of the Ne atom in the approximate coupled cluster triples model CC3. *Chem. Phys. Lett.* **2004**, *391*, 27−32.

(24) Johnson, L. E.; Dalton, L. R.; Robinson, B. H. Optimizing Calculations of Electronic Excitations and Relative Hyperpolarizabilities of Electrooptic Chromophores. *Acc. Chem. Res.* **2014**, *47*, 3258−3265.

(25) Naves, E. S.; Castro, M. A.; Fonseca, T. L. Dynamic (Hyper)Polarizabilities of the Sulphur Dioxide Molecule: Coupled Cluster Calculations Including Vibrational Corrections. *J. Chem. Phys.* **2012**, *136*, No. 014303.

(26) Zaleśny, R. Anharmonicity Contributions to the Vibrational First and Second Hyperpolarizability of Para-Disubstituted Benzenes. *Chem. Phys. Lett.* **2014**, *595-596*, 109−112.

(27) Campo, J.; Painelli, A.; Terenziani, F.; Van Regemorter, T.; Beljonne, D.; Goovaerts, E.; Wenseleers, W. First hyperpolarizability dispersion of the octupolar molecule crystal violet: multiple resonances and vibrational and solvation effects. *J. Am. Chem. Soc.* **2010**, *132*, 16467−16478.

(28) Loco, D.; Polack, É.; Caprasecca, S.; Lagardère, L.; Lipparini, F.; Piquemal, J.-P.; Mennucci, B. A QM/MM Approach Using the AMOEBA Polarizable Embedding: From Ground State Energies to Electronic Excitations. *J. Chem. Theory Comput.* **2016**, *12*, 3654−3661.

(29) Rizzo, A.; Coriani, S.; Fernández, B.; Christiansen, O. A coupled cluster response study of the electric dipole polarizability, first and second hyperpolarizabilities of HCl. *Phys. Chem. Chem. Phys.* **2002**, *4*, 2884−2890.

(30) Kongsted, J.; Pedersen, T. B.; Strange, M.; Osted, A.; Hansen, A. E.; Mikkelsen, K. V.; Pawlowski, F.; Jørgensen, P.; Hättig, C. Coupled cluster calculations of the optical rotation of S-propylene oxide in gas phase and solution. *Chem. Phys. Lett.* **2005**, *401*, 385−392.
(31) Jacquemin, D.; Beljonne, D.; Champagne, B.; Geskin, V.; Brédas, J.-L.; André, J.-M. Analysis of the sign reversal of the second-order molecular polarizability in polymethineimine chains. *J. Chem. Phys.* **2001**, *115*, 6766−6774.
(32) Caricato, M. Implementation of the CCSD-PCM linear response function for frequency dependent properties in solution: Application to polarizability and specific rotation. *J. Chem. Phys.* **2013**, *139*, 114103.
(33) Nielsen, C. B.; Christiansen, O.; Mikkelsen, K. V.; Kongsted, J. Density functional self-consistent quantum mechanics/molecular mechanics theory for linear and nonlinear molecular properties: Applications to solvated water and formaldehyde. *J. Chem. Phys.* **2007**, *126*, 154112.
(34) Cardenuto, M. H.; Champagne, B. The first hyperpolarizability of nitrobenzene in benzene solutions: investigation of the effects of electron correlation within the sequential QM/MM approach. *Phys. Chem. Chem. Phys.* **2015**, *17*, 23634−23642.
(35) Ferrighi, L.; Frediani, L.; Cappelli, C.; Sałek, P.; Ågren, H.; Helgaker, T.; Ruud, K. Density-functional-theory study of the electric-field-induced second harmonic generation (EFISHG) of push−pull phenylpolyenes in solution. *Chem. Phys. Lett.* **2006**, *425*, 267−272.
(36) Corni, S.; Cammi, R.; Mennucci, B.; Tomasi, J. Electronic Excitation Energies of Molecules in Solution within Continuum Solvation Models: Investigating the Discrepancy Between State-Specific and Linear-Response Methods. *J. Chem. Phys.* **2005**, *123*, 134512.
(37) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198−1229.
(38) Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103*, 227−249.
(39) Warshel, A.; Karplus, M. Calculation of ground and excited state potential surfaces of conjugated molecules. I. Formulation and parametrization. *J. Am. Chem. Soc.* **1972**, *94*, 5612−5625.
(40) Olsen, J. M.; Aidas, K.; Kongsted, J. Excited states in solution through polarizable embedding. *J. Chem. Theory Comput.* **2010**, *6*, 3721−3734.
(41) Olsen, J. M. H.; Kongsted, J. Molecular properties through polarizable embedding. In *Advances in Quantum Chemistry*; Elsevier: 2011; *61*; 107−143.
(42) Lipparini, F.; Cappelli, C.; Barone, V. Linear Response Theory and Electronic Transition Energies for a Fully Polarizable QM/Classical Hamiltonian. *J. Chem. Theory Comput.* **2012**, *8*, 4153−4165.
(43) Curutchet, C.; Muñoz-Losa, A.; Monti, S.; Kongsted, J.; Scholes, G. D.; Mennucci, B. Electronic energy transfer in condensed phase studied by a polarizable QM/MM model. *J. Chem. Theory Comput.* **2009**, *5*, 1838−1848.
(44) Jurinovich, S.; Curutchet, C.; Mennucci, B. The Fenna−Matthews−Olson Protein Revisited: A Fully Polarizable (TD) DFT/MM Description. *ChemPhysChem* **2014**, *15*, 3194−3204.
(45) Jensen, L.; Van Duijnen, P. T.; Snijders, J. G. A discrete solvent reaction field model for calculating molecular linear response properties in solution. *J. Chem. Phys.* **2003**, *119*, 3800−3809.
(46) Cappelli, C. Integrated QM/Polarizable MM/Continuum Approaches to Model Chiroptical Properties of Strongly Interacting Solute-Solvent Systems. *Int. J. Quant. Chem.* **2016**, *116*, 1532−1542.
(47) Reinholdt, P.; Kongsted, J.; Olsen, J. M. H. Polarizable Density Embedding: A Solution to the Electron Spill-Out Problem in Multiscale Modeling. *J. Phys. Chem. Lett.* **2017**, *8*, 5949−5958.
(48) Olsen, J. M. H.; Steinmann, C.; Ruud, K.; Kongsted, J. Polarizable Density Embedding: A New QM/QM/MM-Based Computational Strategy. *J. Phys. Chem. A* **2015**, *119*, 5344−5355.
(49) Fahleson, T.; Olsen, J. M. H.; Norman, P.; Rizzo, A. A QM/MM and QM/QM/MM study of Kerr, Cotton-Mouton and Jones linear birefringences in liquid acetonitrile. *Phys. Chem. Chem. Phys.* **2018**, *20*, 3831−3840.
(50) Giovannini, T.; Puglisi, A.; Ambrosetti, M.; Cappelli, C. Polarizable QM/MM Approach with Fluctuating Charges and Fluctuating Dipoles: The QM/FQFμ Model. *J. Chem. Theory Comput.* **2019**, *15*, 2233−2245.
(51) Giovannini, T.; Grazioli, L.; Ambrosetti, M.; Cappelli, C. Calculation of IR Spectra with a Fully Polarizable QM/MM Approach Based on Fluctuating Charges and Fluctuating Dipoles. *J. Chem. Theory Comput.* **2019**, *15*, 5495−5507.
(52) Giovannini, T.; Olszówka, M.; Egidi, F.; Cheeseman, J. R.; Scalmani, G.; Cappelli, C. Polarizable Embedding Approach for the Analytical Calculation of Raman and Raman Optical Activity Spectra of Solvated Systems. *J. Chem. Theory Comput.* **2017**, *13*, 4421−4435.
(53) Giovannini, T.; Del Frate, G.; Lafiosca, P.; Cappelli, C. Effective Computational Route Towards Vibrational Optical Activity Spectra of Chiral Molecules in Aqueous Solution. *Phys. Chem. Chem. Phys.* **2018**, *20*, 9181−9197.
(54) Giovannini, T.; Olszówka, M.; Cappelli, C. Effective Fully Polarizable QM/MM Approach To Model Vibrational Circular Dichroism Spectra of Systems in Aqueous Solution. *J. Chem. Theory Comput.* **2016**, *12*, 5483−5492.
(55) Egidi, F.; Lo Gerfo, G.; Macchiagodena, M.; Cappelli, C. On the Nature of Charge-Transfer Excitations for Molecules in Aqueous Solution: A Polarizable QM/MM Study. *Theor. Chem. Acc.* **2018**, *137*, 82.
(56) Egidi, F.; Russo, R.; Carnimeo, I.; D'Urso, A.; Mancini, G.; Cappelli, C. The Electronic Circular Dichroism of Nicotine in Aqueous Solution: A Test Case for Continuum and Mixed Explicit-Continuum Solvation Approaches. *J. Phys. Chem. A* **2015**, *119*, 5396−5404.
(57) Di Remigio, R.; Giovannini, T.; Ambrosetti, M.; Cappelli, C.; Frediani, L. Fully Polarizable QM/Fluctuating Charge Approach to Two-Photon Absorption of Aqueous Solutions. *J. Chem. Theory Comput.* **2019**, *15*, 4056−4068.
(58) Egidi, F.; Carnimeo, I.; Cappelli, C. The Optical Rotatory Dispersion of Methyloxirane in Aqueous Solution: Assessing the Performance of Density Functional Theory in Combination with a Fully Polarizable QM/MM/PCM Approach. *Opt. Mater. Express* **2015**, *5*, 196−209.
(59) Egidi, F.; Giovannini, T.; Del Frate, G.; Lemler, P. M.; Vaccaro, P. H.; Cappelli, C. A Combined Experimental and Theoretical Study of Optical Rotatory Dispersion for (*R*)-Glycidyl Methyl Ether in Aqueous Solution. *Phys. Chem. Chem. Phys.* **2019**, *21*, 3644−3655.
(60) Giovannini, T.; Lafiosca, P.; Chandramouli, B.; Barone, V.; Cappelli, C. Effective Yet Reliable Computation of Hyperfine Coupling Constants in Solution by a QM/MM approach: Interplay Between Electrostatics and Non-Electrostatic Effects. *J. Chem. Phys.* **2019**, *150*, 124102.
(61) Giovannini, T.; Lafiosca, P.; Cappelli, C. A General Route to Include Pauli Repulsion and Quantum Dispersion Effects in QM/MM Approaches. *J. Chem. Theory Comput.* **2017**, *13*, 4854−4870.
(62) Giovannini, T.; Ambrosetti, M.; Cappelli, C. Quantum Confinement Effects on Solvatochromic Shifts of Molecular Solutes. *J. Phys. Chem. Lett.* **2019**, *10*, 5823−5829.
(63) Mennucci, B.; Amovilli, C.; Tomasi, J. On the effect of Pauli repulsion and dispersion on static molecular polarizabilities and hyperpolarizabilities in solution. *Chem. Phys. Lett.* **1998**, *286*, 221−225.
(64) Boyd, R. W. *Nonlinear optics*; Academic press: 2003.
(65) Wortmann, R.; Krämer, P.; Glania, C.; Lebus, S.; Detzer, N. Deviations from Kleinman symmetry of the second-order polarizability tensor in molecules with low-lying perpendicular electronic bands. *Chem. Phys.* **1993**, *173*, 99−108.
(66) Rice, J. E.; Amos, R. D.; Colwell, S. M.; Handy, N. C.; Sanz, J. Frequency dependent hyperpolarizabilities with application to formaldehyde and methyl fluoride. *J. Chem. Phys.* **1990**, *93*, 8828−8839.

(67) Rice, J. E.; Handy, N. C. The calculation of frequency-dependent hyperpolarizabilities including electron correlation effects. *Int. J. Quantum Chem.* **1992**, *43*, 91−118.
(68) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *J. Phys. Chem. B* **2009**, *113*, 6378−6396.
(69) Clays, K.; Persoons, A. Hyper-Rayleigh scattering in solution. *Phys. Rev. Lett.* **1991**, *66*, 2980.
(70) Clays, K.; Persoons, A. Hyper-Rayleigh Scattering in Solution. *Rev. Sci. Instrum.* **1992**, *63*, 3285−3289.
(71) Ray, P. C.; Das, P. K.; Ramasesha, S. A comparative study of first hyperpolarizabilities of the acidic and basic forms of weak organic acids in water. *J. Chem. Phys.* **1996**, *105*, 9633−9639.
(72) Giovannini, T.; Ambrosetti, M.; Cappelli, C. A Polarizable Embedding Approach to Second Harmonic Generation (SHG) of Molecular Systems in Aqueous Solutions. *Theor. Chem. Acc.* **2018**, *137*, 74.
(73) Botek, E.; Spassova, M.; Champagne, B.; Asselberghs, I.; Persoons, A.; Clays, K. Hyper-Rayleigh Scattering of Neutral and Charged Helicenes. *Chem. Phys. Lett.* **2005**, *412*, 274−279.
(74) DeFusco, A.; Minezawa, N.; Slipchenko, L. V.; Zahariev, F.; Gordon, M. S. Modeling solvent effects on electronic excited states. *J. Phys. Chem. Lett.* **2011**, *2*, 2184−2192.
(75) Giovannini, T.; Riso, R. R.; Ambrosetti, M.; Puglisi, A.; Cappelli, C. Electronic Transitions for a Fully Polarizable QM/MM Approach Based on Fluctuating Charges and Fluctuating Dipoles: Linear and Corrected Linear Response Regimes. *J. Chem. Phys.* **2019**, *151*, 174104.
(76) Amovilli, C.; Mennucci, B. Self-Consistent-Field Calculation of Pauli Repulsion and Dispersion Contributions to the Solvation Free Energy in the Polarizable Continuum Model. *J. Phys. Chem. B* **1997**, *101*, 1051−1057.
(77) McWeeny, R. *Methods of Molecular Quantum Mechanics*; Academic press: London, U.K., 1992.
(78) Frisch, M. J. *Gaussian 16* Revision A.03. 2016; Gaussian Inc.: Wallingford CT.
(79) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098−3100.
(80) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(81) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37*, 785−789.
(82) Yanai, T.; Tew, D. P.; Handy, N. C. A New Hybrid Exchange-Correlation Functional Using the Coulomb-Attenuating Method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51−57.
(83) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215−241.
(84) Rick, S. W.; Stuart, S. J.; Berne, B. J. Dynamical Fluctuating Charge Force Fields: Application to Liquid Water. *J. Chem. Phys.* **1994**, *101*, 6141−6156.
(85) Rick, S. W.; Stuart, S. J.; Bader, J. S.; Berne, B. J. Fluctuating Charge Force Fields for Aqueous Solutions. *J. Mol. Liq.* **1995**, *65-66*, 31−40.
(86) Rick, S. W.; Berne, B. J. Dynamical Fluctuating Charge Force Fields: The Aqueous Solvation of Amides. *J. Am. Chem. Soc.* **1996**, *118*, 672−679.
(87) 1 esu = 115.75 atomic units.
(88) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005**, *105*, 2999−3094.
(89) Mennucci, B. Polarizable Continuum Model. *WIREs Comput. Mol. Sci.* **2012**, *2*, 386−404.
(90) Lipparini, F.; Mennucci, B. Perspective: Polarizable Continuum Models for Quantum-Mechanical Descriptions. *J. Chem. Phys.* **2016**, *144*, 160901.
(91) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: the RESP Model. *J. Phys. Chem.* **1993**, *97*, 10269−10280.
(92) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. Application of RESP Charges to Calculate Conformational Energies, Hydrogen Bond Energies, and Free Energies of Solvation. *J. Am. Chem. Soc.* **1993**, *115*, 9620−9631.
(93) Cieplak, P.; Cornell, W. D.; Bayly, C.; Kollman, P. A. Application of the Multimolecule and Multiconformational RESP Methodology to Biopolymers: Charge Derivation for DNA, RNA and Proteins. *J. Comput. Chem.* **1995**, *16*, 1357−1377.
(94) Macchiagodena, M.; Mancini, G.; Pagliai, M.; Barone, V. Accurate prediction of bulk properties in hydrogen bonded liquids: amides as case studies. *Phys. Chem. Chem. Phys.* **2016**, *18*, 25342−25354.
(95) Macchiagodena, M.; Mancini, G.; Pagliai, M.; Cardini, G.; Barone, V. New atomistic model of pyrrole with improved liquid state properties and structure. *Int. J. Quantum Chem.* **2018**, *118*, No. e25554.

https://dx.doi.org/10.1021/acs.jctc.0c00674
J. Chem. Theory Comput. 2020, 16, 6993−7004

CHAPTER 4

Multilevel Density Functional Theory

JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

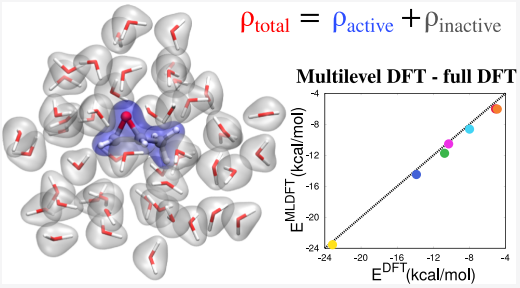
Multilevel Density Functional Theory

Gioia Marrazzini,[§] Tommaso Giovannini,*[,§] Marco Scavino, Franco Egidi, Chiara Cappelli, and Henrik Koch*

Cite This: *J. Chem. Theory Comput.* 2021, 17, 791−803 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Following recent developments in multilevel embedding methods, we introduce a novel density matrix-based multilevel approach within the framework of density functional theory (DFT). In this multilevel DFT, the system is partitioned in an active and an inactive fragment, and all interactions are retained between the two parts. The decomposition of the total system is performed upon the density matrix. The orthogonality between the two parts is maintained by solving the Kohn−Sham equations in the MO basis for the active part only, while keeping the inactive density matrix frozen. This results in the reduction of computational cost. We outline the theory and implementation and discuss the differences and similarities with state-of-the-art DFT embedding methods. We present applications to aqueous solutions of methyloxirane and glycidol.



1. INTRODUCTION

The study of the energetics and physico-chemical properties of large molecular systems is one of the most challenging problems in quantum chemistry.[1] Many processes of chemical interest take place in solution,[2−6] in biological matrices,[7,8] or at the interfaces between different materials.[9−11] The large size of such systems poses theoretical and computational challenges because high-level correlated electronic structure methods are usually unfeasible because of their high computational cost and unfavorable scaling.[12,13] A good compromise between accuracy and computational cost is provided by density functional theory (DFT),[14,15] which accounts for electron correlation in an approximate way. Because of the proven reliability of the results that can be obtained at the DFT level, it has become the most widely used approach for describing the electronic structure of large systems.

DFT permits the investigation of much larger systems than offered by highly correlated methods. However, it cannot be routinely applied to systems constituted by more than 500 atoms, unless implementations through graphical processing units are exploited.[16] The practical limit of 500 atoms makes applications of DFT to biological matrices, interfaces, and solutions particularly cumbersome and, in some cases, even impossible. For these reasons, several approximations have been developed in the past. The nature of any approximation aimed at reducing the computational cost of a calculation is often rooted in the information that one seeks to extract from a simulation. If one is interested in the interaction between two complex medium-sized subsystems and how they influence each other, then, one might seek a partition method which accurately models both moieties with a computational cost comparable to treating both systems separately. More often than not, however, one is more interested in modeling only a small part of a larger chemical system, while the rest is only relevant in as much as it can alter the physical and chemical properties of the smaller region. This is the case of systems embedded in an external environment, whose properties are modified by the latter, the archetypal example being a molecule dissolved in a solvent. For these problems, the system may be partitioned into two or more regions which are then treated at different levels of accuracy, and different approaches may be more or less suitable depending on the specificities of the system/environment. In the special case of systems in solution, particular success has been enjoyed by methods belonging to the family of the so-called focused models,[17−20] which are extremely useful when dealing with the property of a moiety or a chromophore embedded in an external environment.

In focused models, the target molecule is described at a higher level of theory with respect to the environment, which acts as a perturbation on the target system. Among the different focused models that have been developed in the past, the large majority belongs to the family of quantum mechanics (QM)/classical approaches, in which the target is treated at the

Received: September 10, 2020
Published: January 15, 2021

https://dx.doi.org/10.1021/acs.jctc.0c00940
J. Chem. Theory Comput. 2021, 17, 791−803

QM level. The environment is instead described classically, by means of either continuous descriptions, such as the polarizable continuum model,[17,18] or by retaining its atomistic nature in the so-called QM/molecular mechanics (QM/MM) approaches.[21−24] In all these methods, however, the interaction between the two parts of the whole system is usually described by classical electrostatics[20,25−27] and very rarely by including the interactions of the quantum nature, such as Pauli repulsion and dispersion.[28−30] Also, QM/classical methods allow for the treatment of very large systems; however, their accuracy crucially depends on the quality of the parametrization of the classical fragments. In order to avoid such a variability, quantum embedding methods can be exploited.[31−48] In these approaches, the whole system is treated by resorting to a QM description; thus, Pauli repulsion effects are introduced in the modeling. The reduction in the computational cost is then achieved by partitioning the system in at least one active and one inactive part. The former is described at a higher level of accuracy, whereas the latter may be kept frozen or described at a lower level of theory. Different approaches have been proposed in the past, ranging from projection-based methods, such as DFT-in-DFT or HF-in-DFT,[39,46,47,49,50] to frozen density embedding (FDE).[48,51−57] Quantum embeddings have become a popular choice to overcome the aforementioned shortcomings of QM/classical approaches; however, the partitioning of the electronic degrees of freedom for the whole system into the two layers is theoretically far from trivial. In fact, all such partitioning methods can be thought of as being based upon the same basic idea. However, the partition of the electronic degrees of freedom may be performed on molecular orbitals, the density function, or the density matrix. Which basis set must be used to expand the density and wavefunction of the different regions is also an important factor. All these choices lead to drastically different methods which have their unique advantages and complications (*vide infra*). In addition, all quantum-embedding approaches based on DFT suffer from the complication of the exchange−correlation functionals being nonlinear, which leads to nonadditive energy terms in the partitioned equations.

In this paper, we are proposing a novel quantum embedding approach defined in a DFT framework. We denote this method multilevel DFT (MLDFT), due to its similarity with multilevel Hartree−Fock (MLHF),[58,59] and multilevel coupled-cluster[12,60] that we have recently developed. Similar to FDE, in MLDFT, the system is partitioned into an active and an inactive region, and all the interactions are retained between the two parts. The MLDFT conceptually differs from the aforementioned quantum embedding methods because the partitioning is performed on the density matrix instead of the density function as in FDE.[61] As a consequence, the nonadditive term in MLDFT does not contain any kinetic energy term, thus avoiding the theoretical complications present in FDE (*vide infra*). Also, the MLDFT equations are defined in the MO basis of the active fragment only. This feature automatically allows for savings in the computational cost because the inactive MOs are not involved in the self-consistent field (SCF) procedure. In this paper, we derive MLDFT, and we apply the method to ground state (GS) energies of aqueous solutes. The results are compared, in all cases, to full DFT, in order to assess the quality of the multilevel partition.

The manuscript is organized as follows. In the next section, DFT theory is formulated in the MO basis, and MLDFT equations are presented with a particular focus on the computational savings that can be expected. Theoretical comparison with other quantum embedding approaches is also discussed. Then, after a brief section reporting on the computational details of the method, MLDFT is applied to selected aqueous systems, with emphasis on comparison with full DFT results. Conclusions and perspectives of the present work end the manuscript.

2. THEORY

Our starting point is the DFT expression for the electronic energy of the system

$$\begin{aligned} E[\mathbf{D}] &= \mathrm{Tr}\mathbf{h}\mathbf{D} + \frac{1}{2}\mathrm{Tr}\mathbf{D}\mathbf{J}(\mathbf{D}) - \frac{1}{2}c_x\mathrm{Tr}\mathbf{D}\mathbf{K}(\mathbf{D}) \\ &\quad + (1 - c_x)E_x[\mathbf{D}] + E_c[\mathbf{D}] \\ &= \mathrm{Tr}\mathbf{h}\mathbf{D} + \frac{1}{2}\mathrm{Tr}\mathbf{D}\mathbf{J}(\mathbf{D}) - \frac{1}{2}c_x\mathrm{Tr}\mathbf{D}\mathbf{K}(\mathbf{D}) \\ &\quad + (1 - c_x)\int \rho(\mathbf{r})\varepsilon_x(\rho(\mathbf{r}))\mathrm{d}\mathbf{r} \\ &\quad + \int \rho(\mathbf{r})\varepsilon_c(\rho(\mathbf{r}))\mathrm{d}\mathbf{r} \end{aligned} \tag{1}$$

$$\rho(\mathbf{r}) = \sum_{\mu\nu} D_{\mu\nu}\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r}) \tag{2}$$

Here, $\mathbf{D}$ is the one-particle density matrix expanded in the atomic orbital (AO) basis $\{\chi_\mu\}$, and $\mathbf{h}$ is the one-electron operator, whereas $\mathbf{J}$ and $\mathbf{K}$ are Coulomb and exchange matrices, respectively. The E_x and E_c terms are DFT exchange and correlation energy functionals; $\rho(\mathbf{r})$ is the DFT density function, and ε_x and ε_c are the exchange and correlation energy densities per unit particle, respectively. The coefficient c_x defines whether pure DFT ($c_x = 0$) or hybrid DFT functionals ($c_x \neq 0$) is used. Note that in the second equality in eq 1, it is implied that we are restricting the treatment to semilocal possibly hybrid functionals that are commonly used in quantum chemistry.

The energy defined in eq 1 is usually minimized in the AO basis. It is also possible to reformulate the minimization in the MO basis, which is the strategy employed in this work, similarly to what was presented for the Hartree−Fock case by Sæther et al.[58] This can be accomplished by parametrizing the density matrix $\mathbf{D}$ in terms of an antisymmetric rotation matrix, in which only the nonredundant occupied-virtual rotations are considered.[58]

2.1. Multilevel DFT. The MLDFT method belongs to the family of the so-called focused models. The part of the system which is under investigation (active) is described accurately, whereas the remaining (inactive) part remains frozen during the optimization of the active fragment. The choice of the partitioning intimately depends on the specificities of the system, its chemical nature, and the properties one wishes to simulate. Like other QM/QM methods, the details of the partitioning largely define the nature of the method, with its advantages as well as the limitations that arise from the simplifications that bring about the sought-after computational savings. Within the MLDFT formalism, the separation of the system into two parts is based on the following decomposition of the density matrix $\mathbf{D}$, which in turn defines the separation of the density function $\rho(\mathbf{r})$ as well

$$\mathbf{D} = \mathbf{D}^A + \mathbf{D}^B \Rightarrow \rho(\mathbf{r}) = \rho^A(\mathbf{r}) + \rho^B(\mathbf{r}) \tag{3}$$

https://dx.doi.org/10.1021/acs.jctc.0c00940
J. Chem. Theory Comput. 2021, 17, 791−803

where A and B indicate the active and inactive fragments, respectively. As stated above, the active and inactive densities are usually defined on a physico-chemical basis. In case of a molecular system in solution, it is natural to define the solute as the active fragment, whereas the solvent molecules are treated as the inactive part. It is crucial to notice that the partitioning in eq 3 is mathematical and therefore has some level of arbitrariness; that is, there is no unique way to perform the separation purely based on physico-chemical properties of the system. Therefore, a choice has to be made at this point. In this work, the partitioning is performed by means of Cholesky decomposition of the total density matrix for the active occupied MOs, from which the active density matrix $\mathbf{D}^{\mathrm{A}}$ is calculated, similarly to what was done in previous works presenting alternative multilevel methods.[12,58,60,62,63] One advantage of this procedure is that it ensures the all active and inactive orbitals are orthogonal and remain so during all the subsequent SCF procedure performed on the active subsystem.[64]

Now using eq 3, the total electronic energy in eq 1 can be written as

$$\begin{aligned} E[\mathbf{D}^{\mathrm{A}};\mathbf{D}^{\mathrm{B}}] &= \mathrm{Tr}\mathbf{h}\mathbf{D}^{\mathrm{A}} + \mathrm{Tr}\mathbf{h}\mathbf{D}^{\mathrm{B}} \\ &+ \frac{1}{2}\mathrm{Tr}\mathbf{D}^{\mathrm{A}}\mathbf{J}(\mathbf{D}^{\mathrm{A}}) + \frac{1}{2}\mathrm{Tr}\mathbf{D}^{\mathrm{B}}\mathbf{J}(\mathbf{D}^{\mathrm{B}}) + \mathrm{Tr}\mathbf{D}^{\mathrm{A}}\mathbf{J}(\mathbf{D}^{\mathrm{B}}) \\ &- c_{\mathrm{x}}\left(\frac{1}{2}\mathrm{Tr}\mathbf{D}^{\mathrm{A}}\mathbf{K}(\mathbf{D}^{\mathrm{A}}) + \frac{1}{2}\mathrm{Tr}\mathbf{D}^{\mathrm{B}}\mathbf{K}(\mathbf{D}^{\mathrm{B}}) + \mathrm{Tr}\mathbf{D}^{\mathrm{A}}\mathbf{K}(\mathbf{D}^{\mathrm{B}})\right) \\ &+ (1-c_{\mathrm{x}})\int \rho(\mathbf{r})\varepsilon_{\mathrm{x}}(\rho(\mathbf{r}))\mathrm{d}\mathbf{r} + \int \rho(\mathbf{r})\varepsilon_{\mathrm{c}}(\rho(\mathbf{r}))\mathrm{d}\mathbf{r} \end{aligned} \tag{4}$$

where the symmetry of $\mathbf{J}$ and $\mathbf{K}$ matrices has been used. So far, the method is analogous to the recently proposed MLHF approach.[58] However, in MLDFT, the last two terms are not linear in the densities of the two subsystems; therefore, we cannot directly split them into distinct contributions arising from $\rho^{\mathrm{A}}(\mathbf{r})$ and $\rho^{\mathrm{B}}(\mathbf{r})$. In order to acquire a physical understanding of eq 4, we rewrite the last two terms by using this trivial identity for the exchange−correlation energy density ($\varepsilon_{\mathrm{xc}} = \varepsilon_{\mathrm{x}} + \varepsilon_{\mathrm{c}}$)

$$\begin{aligned} &\int [\rho^{\mathrm{A}}(\mathbf{r}) + \rho^{\mathrm{B}}(\mathbf{r})]\varepsilon_{\mathrm{xc}}(\rho^{\mathrm{A}}(\mathbf{r}) + \rho^{\mathrm{B}}(\mathbf{r}))\mathrm{d}\mathbf{r} \\ &= \int [\rho^{\mathrm{A}}(\mathbf{r}) + \rho^{\mathrm{B}}(\mathbf{r})]\varepsilon_{\mathrm{xc}}(\rho^{\mathrm{A}}(\mathbf{r}) + \rho^{\mathrm{B}}(\mathbf{r}))\mathrm{d}\mathbf{r} \\ &+ \int \rho^{\mathrm{A}}(\mathbf{r})\varepsilon_{\mathrm{xc}}(\rho^{\mathrm{A}}(\mathbf{r}))\mathrm{d}\mathbf{r} - \int \rho^{\mathrm{A}}(\mathbf{r})\varepsilon_{\mathrm{xc}}(\rho^{\mathrm{A}}(\mathbf{r}))\mathrm{d}\mathbf{r} \\ &+ \int \rho^{\mathrm{B}}(\mathbf{r})\varepsilon_{\mathrm{xc}}(\rho^{\mathrm{B}}(\mathbf{r}))\mathrm{d}\mathbf{r} - \int \rho^{\mathrm{B}}(\mathbf{r})\varepsilon_{\mathrm{xc}}(\rho^{\mathrm{B}}(\mathbf{r}))\mathrm{d}\mathbf{r} \\ &+ \int \boldsymbol{\rho}^{\mathrm{A}}(\mathbf{r})\varepsilon_{\mathrm{xc}}(\rho^{\mathrm{B}}(\mathbf{r}))\mathrm{d}\mathbf{r} - \int \rho^{\mathrm{A}}(\mathbf{r})\varepsilon_{\mathrm{xc}}(\rho^{\mathrm{B}}(\mathbf{r}))\mathrm{d}\mathbf{r} \\ &+ \int \rho^{\mathrm{B}}(\mathbf{r})\varepsilon_{\mathrm{xc}}(\rho^{\mathrm{A}}(\mathbf{r}))\mathrm{d}\mathbf{r} - \int \rho^{\mathrm{B}}(\mathbf{r})\varepsilon_{\mathrm{xc}}(\rho^{\mathrm{A}}(\mathbf{r}))\mathrm{d}\mathbf{r} \end{aligned} \tag{5}$$

Substituting eq 5 into 4 and reorganizing the different terms we obtain

$$\begin{aligned} E[\mathbf{D}^{\mathrm{A}};\mathbf{D}^{\mathrm{B}}] &= \mathrm{Tr}\mathbf{h}\mathbf{D}^{\mathrm{A}} + \frac{1}{2}\mathrm{Tr}\mathbf{D}^{\mathrm{A}}\mathbf{J}(\mathbf{D}^{\mathrm{A}}) - \frac{1}{2}c_{\mathrm{x}}\mathrm{Tr}\mathbf{D}^{\mathrm{A}} \\ &\quad \mathbf{K}(\mathbf{D}^{\mathrm{A}}) \\ &+ (1-c_{\mathrm{x}})\int \rho^{\mathrm{A}}(\mathbf{r})\varepsilon_{\mathrm{x}}(\rho^{\mathrm{A}}(\mathbf{r}))\mathrm{d}\mathbf{r} + \int \rho^{\mathrm{A}}(\mathbf{r})\varepsilon_{\mathrm{c}}(\rho^{\mathrm{A}}(\mathbf{r}))\mathrm{d}\mathbf{r} \\ &+ \mathrm{Tr}\mathbf{h}\mathbf{D}^{\mathrm{B}} + \frac{1}{2}\mathrm{Tr}\mathbf{D}^{\mathrm{B}}\mathbf{J}(\mathbf{D}^{\mathrm{B}}) - \frac{1}{2}c_{\mathrm{x}}\mathrm{Tr}\mathbf{D}^{\mathrm{B}}\mathbf{K}(\mathbf{D}^{\mathrm{B}}) \\ &+ (1-c_{\mathrm{x}})\int \rho^{\mathrm{B}}(\mathbf{r})\varepsilon_{\mathrm{x}}(\rho^{\mathrm{B}}(\mathbf{r}))\mathrm{d}\mathbf{r} + \int \rho^{\mathrm{B}}(\mathbf{r})\varepsilon_{\mathrm{c}}(\rho^{\mathrm{B}}(\mathbf{r}))\mathrm{d}\mathbf{r} \\ &+ \mathrm{Tr}\mathbf{D}^{\mathrm{A}}\mathbf{J}(\mathbf{D}^{\mathrm{B}}) - c_{\mathrm{x}}\mathrm{Tr}\mathbf{D}^{\mathrm{A}}\mathbf{K}(\mathbf{D}^{\mathrm{B}}) \\ &+ (1-c_{\mathrm{x}})\left(\int \rho^{\mathrm{A}}(\mathbf{r})\varepsilon_{\mathrm{x}}(\rho^{\mathrm{B}}(\mathbf{r}))\mathrm{d}\mathbf{r} + \int \rho^{\mathrm{B}}(\mathbf{r})\varepsilon_{\mathrm{x}}(\rho^{\mathrm{A}}(\mathbf{r}))\mathrm{d}\mathbf{r}\right) \\ &+ \int \rho^{\mathrm{A}}(\mathbf{r})\varepsilon_{\mathrm{c}}(\rho^{\mathrm{B}}(\mathbf{r}))\mathrm{d}\mathbf{r} + \int \rho^{\mathrm{B}}(\mathbf{r})\varepsilon_{\mathrm{c}}(\rho^{\mathrm{A}}(\mathbf{r}))\mathrm{d}\mathbf{r} + E^{\mathrm{AB}}_{\text{non-add}} \end{aligned} \tag{6}$$

where

$$\begin{aligned} E^{\mathrm{AB}}_{\mathrm{nonadd}}[\mathbf{D}^{\mathrm{A}};\mathbf{D}^{\mathrm{B}}] &= \int \rho(\mathbf{r})\varepsilon_{\mathrm{c}}(\rho(\mathbf{r}))\mathrm{d}\mathbf{r} \\ &\quad - \int \rho(\mathbf{r})\varepsilon_{\mathrm{c}}(\rho^{\mathrm{A}}(\mathbf{r}))\mathrm{d}\mathbf{r} \\ &\quad - \int \rho(\mathbf{r})\varepsilon_{\mathrm{c}}(\rho^{\mathrm{B}}(\mathbf{r}))\mathrm{d}\mathbf{r} \\ &\quad + (1-c_{\mathrm{x}})\left(\int \rho(\mathbf{r})\varepsilon_{\mathrm{x}}(\rho(\mathbf{r}))\mathrm{d}\mathbf{r}\right. \\ &\quad - \int \rho(\mathbf{r})\varepsilon_{\mathrm{x}}(\rho^{\mathrm{A}}(\mathbf{r}))\mathrm{d}\mathbf{r} \\ &\quad \left. - \int \rho(\mathbf{r})\varepsilon_{\mathrm{x}}(\rho^{\mathrm{B}}(\mathbf{r}))\mathrm{d}\mathbf{r}\right) \end{aligned} \tag{7}$$

In eq 6, the first four lines define the energy of the active and inactive fragments, whereas the last three lines define the active−inactive interaction. In MLDFT, the density matrix of the inactive part $\mathbf{D}^{\mathrm{B}}$ (and $\rho^{\mathrm{B}}(\mathbf{r})$) is frozen, and therefore, it acts as an external field on the active fragment. Therefore, only the energy terms containing the B labels are fixed during the SCF procedure. In this regard, the MLDFT is similar in construction to FDE approaches.[52,55,61] However, as it can be noticed from eq 7, one crucial difference is that the nonadditive energy terms do not involve the kinetic energy functional, thus allowing for a computational advantage with respect to common FDE approaches, which need to specifically employ nonadditive kinetic terms. This advantage stems from the fact that in MLDFT, the separation of the system into two subsystems is performed on the total density matrix, rather than the density function. However, one disadvantage with respect to FDE is that for MLDFT, we are forced to select a suitable decomposition algorithm [Cholesky/projected atomic orbitals (PAOs) are used in this work for active occupied and virtual orbitals, respectively], and the results may depend on it, while FDE does not need to make this choice. The total DFT Fock matrix is given by

https://dx.doi.org/10.1021/acs.jctc.0c00940
J. Chem. Theory Comput. 2021, 17, 791−803

$$F_{\mu\nu} = h_{\mu\nu} + J_{\mu\nu}(\mathbf{D}) - c_x K_{\mu\nu}(\mathbf{D}) + (1 - c_x)\int v_x(\rho(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\mathrm{d}\mathbf{r} + \int v_c(\rho(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\mathrm{d}\mathbf{r} \quad (8)$$

where $v_x(\rho(\mathbf{r}))$ and $v_c(\rho(\mathbf{r}))$ are the exchange and correlation potential densities, respectively. Using the partitioning in eqs 3 and 8 we get

$$F_{\mu\nu} = h_{\mu\nu} + J_{\mu\nu}(\mathbf{D}^A) + J_{\mu\nu}(\mathbf{D}^B) - c_x(K_{\mu\nu}(\mathbf{D}^A) + K_{\mu\nu}(\mathbf{D}^B)) + (1 - c_x)\int v_x(\rho^A(\mathbf{r}) + \rho^B(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\mathrm{d}\mathbf{r} + \int v_c(\rho^A(\mathbf{r}) + \rho^B(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\mathrm{d}\mathbf{r} \quad (9)$$

We exploit the same identity of eq 5 for the exchange−correlation potential density ($v_{xc} = v_x + v_c$). In this way, the last two terms in eq 9 become

$$\begin{aligned}&\int v_{xc}(\rho_A(\mathbf{r}) + \rho_B(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\mathrm{d}\mathbf{r}\\ &= \int v_{xc}(\rho_A(\mathbf{r}) + \rho_B(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\mathrm{d}\mathbf{r}\\ &+ \int v_{xc}(\rho_A(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\mathrm{d}\mathbf{r} - \int v_{xc}(\rho_A(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\mathrm{d}\mathbf{r}\\ &+ \int v_{xc}(\rho_B(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\mathrm{d}\mathbf{r} - \int v_{xc}(\rho_B(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\mathrm{d}\mathbf{r}\end{aligned} \quad (10)$$

Reorganizing the terms in eq 9, we can obtain the working expression for the MLDFT Fock matrix

$$\begin{aligned}F_{\mu\nu} = h_{\mu\nu}&\\ &+ J_{\mu\nu}(\mathbf{D}^A) - c_x K_{\mu\nu}(\mathbf{D}^A)\\ &\quad + (1 - c_x)\int v_x(\rho^A(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\mathrm{d}\mathbf{r}\\ &\quad + \int v_c(\rho^A(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\mathrm{d}\mathbf{r}\\ &\quad \underbrace{\qquad\qquad}_{2e_{A,\mu\nu}}\\ &+ J_{\mu\nu}(\mathbf{D}^B) - c_x K_{\mu\nu}(\mathbf{D}^B)\\ &\quad + (1 - c_x)\int v_x(\rho^B(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\mathrm{d}\mathbf{r}\\ &\quad + \int v_c(\rho^B(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\mathrm{d}\mathbf{r}\\ &\quad \underbrace{\qquad\qquad}_{2e_{B,\mu\nu}}\\ &+ (1 - c_x)\int \chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})[v_x(\rho^A(\mathbf{r}) + \rho^B(\mathbf{r}))\\ &\quad - v_x(\rho^A(\mathbf{r})) - v_x(\rho^B(\mathbf{r}))]\mathrm{d}\mathbf{r}\\ &\quad \underbrace{\qquad\qquad}_{2e^{\text{non-add}}_{AB,\mu\nu}}\\ &+ \int \chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})[v_c(\rho^A(\mathbf{r}) + \rho^B(\mathbf{r}))\\ &\quad - v_c(\rho^A(\mathbf{r})) - v_c(\rho^B(\mathbf{r}))]\mathrm{d}\mathbf{r}\\ &\quad \underbrace{\qquad\qquad}_{2e^{\text{non-add}}_{AB,\mu\nu}}\end{aligned} \quad (11)$$

where the two-electron contributions of A and B fragments and the interaction term AB are highlighted as $2e_{X,\mu\nu}$, {X = A, B, AB}.

There are two main advantages of using MLDFT compared to full DFT. First, the HF exchange contribution is usually the most expensive term in most hybrid functionals. In MLDFT, only the active exchange term is to be computed at each SCF cycle, whereas the exchange integral of the inactive fragment is computed at the first SCF cycle only, as it is constant during the optimization. Second, the MLDFT SCF procedure can be performed in the MO basis of the active part only, thus intrinsically reducing the computational time as previously observed for the MLHF method.[58]

2.2. Comparison between MLDFT and Other Quantum-Embedding Methods. As stated above, in MLDFT, the density matrix of the whole system is partitioned into active and inactive fragment densities (see eq 3), from which the active and inactive density functions are calculated. The inactive density matrix is kept frozen during the SCF cycles and enters the active Fock matrix as a one-electron term. The MLDFT shares many similarities with FDE methods;[42−45,52,54,61] however, some relevant differences are present, some of which are mentioned in the previous section. First, due to the partitioning performed on the density matrix instead of the density function as in FDE, we avoid the problems arising in the definition of the nonadditive kinetic potential terms that enforce Pauli exclusion between the electrons of the various subsystems in FDE. This is clearly evident in the definition of the nonadditive energy terms in eq 7, which are only due to the intrinsic nonlinearity of the DFT exchange−correlation functionals. However, a choice needs to be made on the decomposition algorithm to partition the density matrix.

One other difference that is worth mentioning is that if the decomposition is applied to the fully converged DFT density matrix, then, the MLDFT energy corresponds to the exact DFT energy of the full system, and it is in fact completely independent on the decomposition algorithm. This feature is shared with projection-based approaches, though the philosophy of MLDFT conceptually differs from DFT-in-DFT methods.[35,39,41,46,47,49] Of course, the point of MLDFT is to avoid a full DFT calculation on the entire system and instead find a way to approximate the total density matrix, which is then partitioned and the density of the active subsystem is then optimized through a subsequent SCF procedure while the density matrix of the inactive part is kept frozen. In this work, this initial full density matrix is obtained through superposition of molecular densities[51,65] followed by Fock matrix diagonalization, although we should emphasize that the method is general and a different choice may be made. Within this procedure, the solute−solvent and solvent−solvent polarization energy terms are only introduced in the starting full Fock diagonalization step. In projection-based approaches, a DFT calculation on the whole system is performed, and the occupied orbitals are then localized. The localized MOs are then assigned to the active or inactive fragments, based on a population threshold.[50] Neither step is required in MLDFT, in which a starting Fock matrix is diagonalized and the obtained density matrix is decomposed into fragment densities. This explains why the SCF procedure on the active subsystem has to be performed in the MO basis of the active fragment only, which is defined in terms of the full AO basis set. We do not perform any basis set truncation which is instead usually

https://dx.doi.org/10.1021/acs.jctc.0c00940
J. Chem. Theory Comput. 2021, 17, 791−803

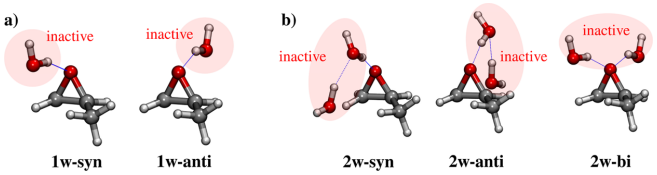


Figure 1. Structure of conformers methyloxirane/water clusters. MLDFT partition is constructed so that methyloxirane is the active part whereas the water molecules are the inactive fragments.

applied in projection-based approaches. This feature automatically allows for saving the computational cost because the inactive MOs are not involved in the SCF procedure.

3. COMPUTATIONAL DETAILS

The DFT and MLDFT are implemented in a development version of the electronic structure program e^T v.1.0.[66] In particular, the DFT grid is constructed using the widely employed Lebedev grid,[67] with the radial quadrature proposed in ref 68. All the calculations are performed using a quadrature of 25th order, and the radial threshold is set to 10^{-5}. The DFT functionals are implemented using the LibXC library.[69,70]

The MLDFT calculation follows this computational protocol:

1. Construction of the initial density matrix by means of superposition of molecular densities,[51,65] followed by the diagonalization of the initial Fock matrix.
2. Partitioning of the new density matrix into A and B densities, using Cholesky decomposition for the active occupied orbitals and PAOs for active virtual orbitals.[12,58,60,62−64,71,72] We refer the reader to ref 63 for the full details on this partitioning of the density matrix. The inactive density matrix is obtained by subtracting the active density matrix from the total one.
3. Calculation of the constant energy terms and the one-electron contributions due to the inactive density matrix B entering in eqs 6 and 11.
4. Minimization of the energy defined in eq 6 in the MO basis of the active part A only, until convergence is reached. All the equations reported in Section 2, which are expressed in the AO basis set, can be transformed in the active MO basis by using the active MO coefficients.

To show the robustness of MLDFT, three different functionals are used: LDA,[73] GGA (PBE[74]), and hybrid (B3LYP[75]). These are combined with three different basis sets: 6-31G, 6-31G*, and aug-cc-pVDZ.

4. NUMERICAL APPLICATIONS

In this section, the MLDFT is applied to some test cases to show the accuracy and the performance of the method. Solvation is one of the main physico-chemical phenomena in which such approaches can be exploited. We show the results of coupling MLDFT with two alternative, fully atomistic, strategies to model aqueous solutions. The first consists of a static modeling, which uses small clusters composed of the solute and a small number of surrounding water molecules. As an alternative, we apply MLDFT to snapshots extracted from a molecular dynamics (MD) simulation. In the latter framework, the dynamical aspects of the solvation phenomenon are retained, as those arising from the combination of conformational changes in the solute and the surrounding solvent. In addition, long range interactions are taken into account. This latter modeling of the solvation phenomenon has been amply and successfully exploited by some of us within the framework of QM/MM approaches.[20,30,76−78]

In the following sections, the combination of MLDFT to the two aforementioned solvation approaches is tested, with application to two relatively small molecules, that is, methyloxirane and glycidol in aqueous solution, which have been studied in the literature both theoretically and experimentally.[27,79−89] Such systems are chosen not only for their simplicity but also because methyloxirane is a rigid molecule, whereas glycidol is not. Therefore, in the latter case, the results depend on the selected QM level, and the approach used to solvation and conformational flexibility, which is instead discarded in the case of methyloxirane. In this way, we can dissect the various effects and highlight the quality of the MLDFT approach in details.

4.1. Cluster Models.

4.1.1. Methyloxirane/Water Clusters. The first studied solute is methyloxirane (MOXY), which is one of the smallest molecules that exhibits a chiral carbon. We have selected different clusters constituted by MOXY and one or two water molecules (see Figure 1) that have been previously studied by Su and Xu[86] to explain the unique characteristics of MOXY in aqueous solution.[86,89]

The two different conformers for the cluster composed of MOXY and one water molecule (MOXY + 1w) are depicted in Figure 1a. In the **1w-syn** structure, water interacts with MOXY through hydrogen bonding on the same side of the methyl group, whereas the opposite occurs for the **1w-anti** structure. In both cases, MOXY is the active fragment, and water is the inactive moiety in MLDFT calculations.

GS energy differences between DFT and MLDFT calculations are depicted in Figure 2, panel (a), left. Raw data are reported in Table S1 given in Supporting Information. We see that the error between MLDFT and full DFT is below 1 mHartree (<0.628 kcal/mol), irrespective of the combination of functional/basis set employed. The error due to the MLDFT partitioning is well below the chemical accuracy (i.e. 1 kcal/mol).

In the right panel of Figure 2a, DFT and MLDFT energy differences between **1w-anti** and **1w-syn** conformers are reported for all the considered combinations of the functional/basis set. The raw data are reported in Table S1 in the Supporting Information. We see that DFT and MLDFT values almost coincide. In particular, LDA and PBE functionals predict **1w-syn** to be the most stable conformer, both at DFT and MLDFT levels, independently of the selected basis set. Notice however that the energy difference between the two conformers decreases either as GGA functionals are employed or diffuse/polarization basis sets are used. The inclusion of HF exchange makes **1w-anti** the most stable conformer, if polarization/diffuse functions are considered. However, for

https://dx.doi.org/10.1021/acs.jctc.0c00940
J. Chem. Theory Comput. 2021, 17, 791−803

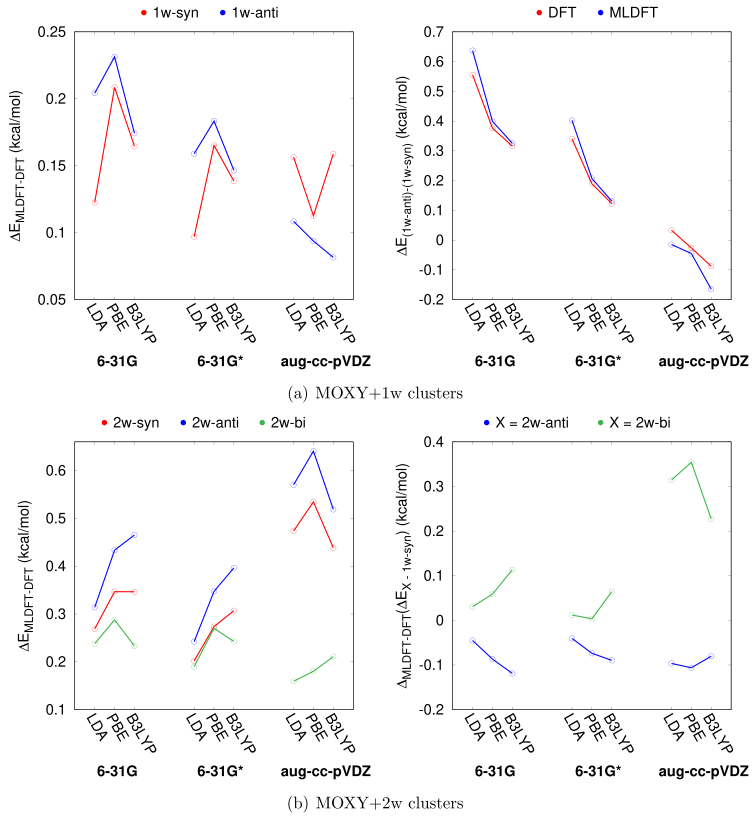


Figure 2. (a) (Left) **1w-syn** and **1w-anti** total energy differences between MLDFT and DFT. (Right) MLDFT and DFT energy difference between **1w-anti** and **1w-syn**. (b) (Left) **2w-syn**, **2w-anti,** and **2w-bi** total energy differences between MLDFT and DFT. (Right) MLDFT-DFT difference on relative energies of **2w-anti** and **2w-bi** with respect to **2w-syn**.

all the considered combinations of the functional/basis set, MLDFT and DFT values are almost perfectly in agreement, with the largest discrepancy being reported for B3LYP/aug-cc-pVDZ (0.08 kcal/mol). These findings clearly show that for this system, MLDFT is able to catch small energy differences, which are again well below the chemical accuracy.

We now turn to the clusters composed of MOXY and two water molecules (MOXY + 2w, Figure 1b). Three main conformers are considered, according to Su and Xu:[86] **2w-syn**, **2w-anti,** and **2w-bi**. The first two conformers differ from the position of water molecules, being both placed on the same side with respect to the methyl group in case of **2w-syn** or on the opposite side for **2w-anti**. In **2w-bi**, the two water molecules are instead placed on the opposite sides of the epoxyl oxygen atom. In all MLDFT calculations, MOXY is the active moiety, whereas the two water molecules are inactive.

In Figure 2b, left, GS energy differences between DFT and MLDFT for the three conformers are reported. The raw values associated with the data plotted in Figure 2b are given in Table S2 in the Supporting Information. The MLDFT and DFT results are, also in this case, in very good agreement with an absolute error below 1 kcal/mol for all combinations of functional/basis sets. However, the absolute deviation between DFT and MLDFT energies is larger than for the previous case (see Figure 2a). In particular, the MLDFT error is larger for **2w-syn** and **2w-anti** than for **2w-bi**, for which it is in line with what we have shown above for MOXY + 1w clusters (∼0.1−0.3 kcal/mol). The increase in the error may be justified by the fact that **2w-syn** and **2w-anti** feature one water molecule that is linked to another water molecule by means of intermolecular hydrogen bonding. The density matrix of the inactive fragments (the two water molecules) is kept frozen; therefore, the water molecule that is not directly bonded to the solute remains in its frozen electronic configuration, resulting in a larger error in the total energy. Such an hypothesis is confirmed by the fact that the error increases when the diffuse aug-cc-pVDZ basis set is used, and the same does not occur for **2w-bi**, where both water molecules are directly linked to methyloxirane through hydrogen bonding interactions.

The MLDFT-DFT deviations in energy differences between each conformer and **2w-syn** are shown in Figure 2b, right. We note small discrepancies between MLDFT and full DFT; however, also in this case, they are below the chemical accuracy, with the maximum error reported by PBE/6-31G* (∼0.35 kcal/mol). The error in the energy differences between the conformers is lower than for the total GS energies reported in Figure 2b, left.

https://dx.doi.org/10.1021/acs.jctc.0c00940
J. Chem. Theory Comput. 2021, 17, 791−803

4.1.2. Glycidol/Water Clusters. MOXY is a rigid molecule, so the different solvated conformers mainly differ by the position of the water molecules. In this section, we show how MLDFT can treat flexible solutes, and to this end, we have selected glycidol (GLY), which is a derivative of MOXY, where one hydrogen of the methyl group is replaced by the OH group (see Figures 3−5). In all MLDFT calculations, the GLY

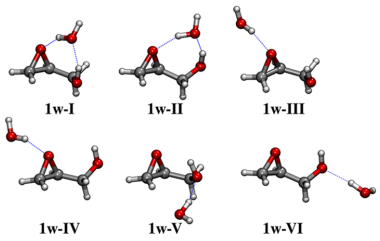


Figure 3. Structures of the six conformers of glycidol + 1 water clusters. In MLDFT calculations, the glycidol moiety is active whereas the water molecule is inactive.

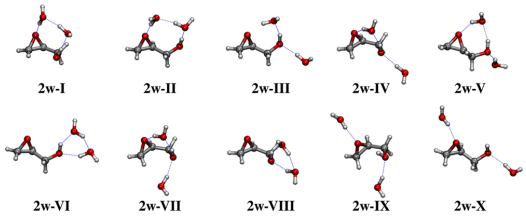


Figure 4. Structures of the ten conformers of glycidol + 2 waters clusters. In MLDFT calculations, glycidol is active whereas water molecules are inactive.

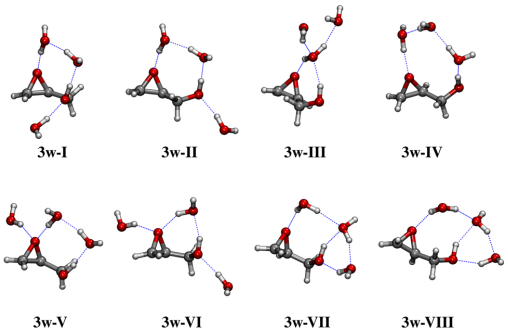


Figure 5. Structures of the eight conformers of glycidol + 3 waters clusters. In MLDFT calculations, glycidol is active whereas water molecules are inactive.

moiety is the active fragment and the water molecules are the inactive part. The presence of the hydroxyl group makes glycidol flexible up to the point that eight different conformers can be located in the gas phase potential energy surface (PES).[79,84]

To build up a glycidol/water clusters, different structures constituted by GLY and one, two, and three water molecules were constructed, by following the strategy reported in ref 84. Such structures are depicted in Figures 3−5. We note that the different structures not only differ by the position of the water molecules but also by the conformation of glycidol. In particular, the six conformers constituted by GLY and one water (GLY + 1w) are characterized by a different position of the water molecule. The latter interacts via hydrogen bonding with both the hydroxyl and epoxyl groups (**1w-I** and **1w-II**), with the epoxyl group only (**1w-III** and **1w-IV**) or with only the oxygen atom of the hydroxyl group (**1w-V** and **1w-VI**). The inclusion of additional water (GLY + 2w) results in ten different conformers, which are shown in Figure 4. These contain three or four center bridges (conformers **2w-I**, **2w-II**, **2w-IV**, **2w-V**, **2w-VI**, **2w-VII,** and **2w-VIII**) or are conformers where the two water molecules interact via hydrogen bonding with the epoxyl and hydroxyl groups (conformers **2w-III**, **2w-IX,** and **2w-X**). If three explicit water molecules are added to GLY (GLY + 3w), the conformational search provides eight main conformers, which are graphically depicted in Figure 5. Similarly to the previous case, some of them contain three or four center bridges (conformers **3w-I**, **3w-II**, **3w-V**, and **3w-VI**), whereas in conformers **3w-IV**, **3w-VII,** and **3w-VIII,** a five center bridge is present. In all cases, water molecules that are not involved in bridges interact with GLY through hydrogen bonding interaction. Conformer **3w-III** is instead characterized by a three center bridge and by the remaining water molecules hydrogen bonded to the bridge water.

We now move to discuss GS energy differences between DFT and MLDFT (see Figure 6a, raw data are given in Tables S3−S5 in the Supporting Information).

In Figure 6, panel (a), MLDFT - DFT GS energy differences for all the different conformers of GLY + 1w, GLY + 2w, and GLY + 3w water clusters are shown. The error reported by MLDFT is below 0.1 mH (<0.627 kcal/mol) when applied to GLY + 1w, at all the selected levels of theory. In particular, energy differences are perfectly in line with what is shown in Figure 2a, left panel, in case of MOXY + 1w clusters. Moving to GLY + 2w conformers, the agreement between DFT and MLDFT is almost perfect at all levels of theory, being the energy difference below 0.8 kcal/mol in all cases. We also see that at the B3LYP/aug-cc-pVDZ level, for **2w-I** and **2w-II,** the difference between MLDFT and full DFT is larger than for the other conformers (>0.1 mH, 0.627 kcal/mol). This is due to the specific spatial arrangement of water molecules, which create a four-center bridge connecting GLY hydroxyl and epoxyl groups (see Figure 4).

As stated above for MOXY + 2w clusters, MLDFT accounts for all the interactions between active and inactive parts, with the exception of dispersion; however, the inactive fragment(s) are described by a frozen density matrix. Therefore, polarization and charge transfer (and dispersion) effects are neglected in the inactive region. For **2w-I** and **2w-II,** we can speculate that such interactions may play a relevant role because the two inactive water molecules are hydrogen bonded. Also, their role is clearly increased when diffuse and polarization functions are included in the basis set (aug-cc-pVDZ) because such functions enhance the effects of these interactions. This does not occur in case of other conformers because of the different spatial arrangement of the solvent molecules.

We now focus on GLY + 3w conformers. The agreement between MLDFT and the reference full DFT values is generally worse than in the previous cases (see the right panel of Figure 6a). However, the average error is of about 0.67 kcal/mol (∼0.1 mH), that is, again well beyond the

https://dx.doi.org/10.1021/acs.jctc.0c00940
J. Chem. Theory Comput. 2021, 17, 791−803

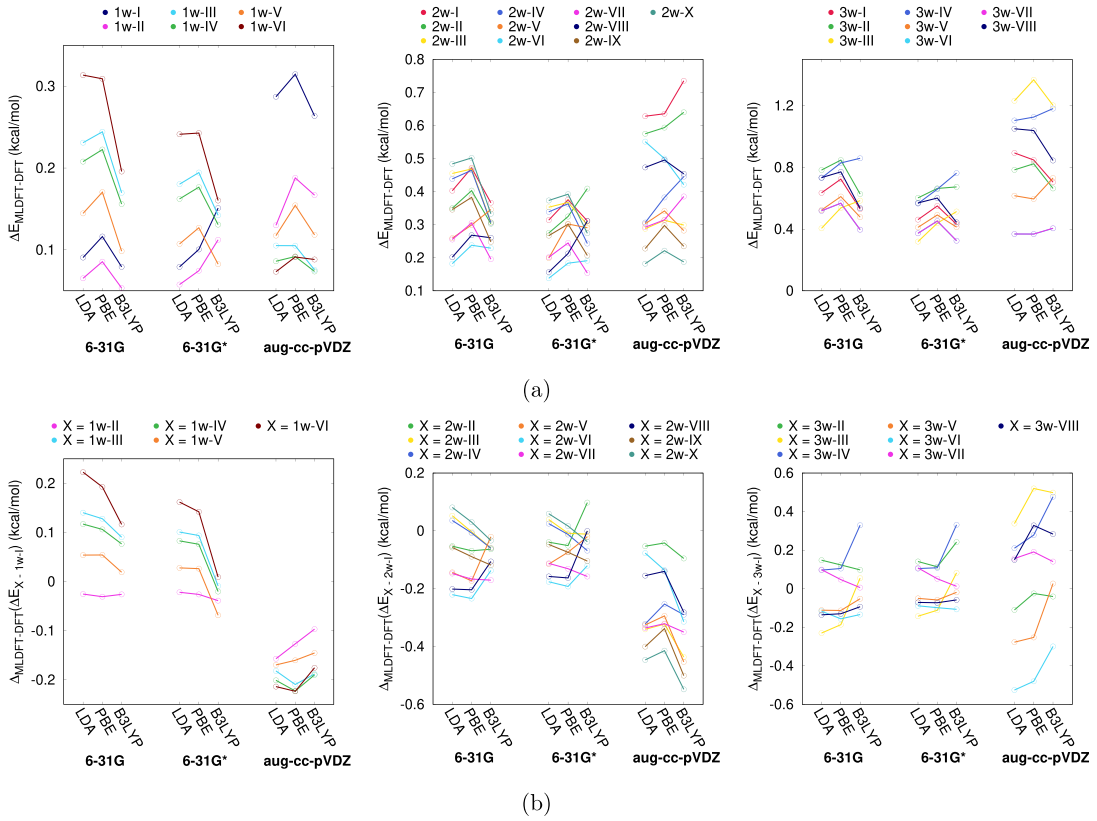


Figure 6. (a) GLY + 1w (left), GLY + 2w (middle), and GLY + 3w (right) GS energy differences between MLDFT and reference DFT values. (b) MLDFT-DFT energy deviations for the energy differences between each conformer of GLY + 1w conformers and **1w-I** (left), GLY + 2w conformers and **2w-I** (middle) and GLY + 3w conformers and **3w-I** (right). All values are reported in kcal/mol.

chemical accuracy. The largest discrepancy is shown by **3w-III** for all the functionals (LDA, PBE, or B3LYP) in combination with aug-cc-pVDZ (∼1.2 kcal/mol). Again, this can be explained by considering the spatial arrangement of water molecules around GLY (see Figure 5). Similar to **2w-I** and **2w-II**, the effect of charge transfer and polarization interactions, which are neglected by the partitioning of the inactive density matrix in MLDFT, may play a relevant role. Such effects are larger for **3w-III**; however, they affect also other conformers which are characterized by a four/five center bridge. It is also worth noticing that the MLDFT error is expected to increase with the size of the studied system because the energy is an extensive quantity. Such a trend is in fact reported for both MOXY and GLY clusters.

Let us now discuss the MLDFT-DFT energy deviations for the energy differences between each conformer of the GLY clusters and **1w-I**, **2w-I**, and **3w-I**, which are reported in Figure 6b. Raw data are given in Tables S3−S5 in the Supporting Information.

For the GLY + 1w system, both MLDFT and DFT predict **1w-I** to be the most stable at all levels of theory, whereas the relative populations of the other conformers strongly depend on the theory level (see Figure 6b, left panel). In particular, the energy differences of each conformer with respect to **1w-I** decrease as larger basis sets are employed and also by moving from LDA to PBE and B3LYP. The error between MLDFT and DFT is instead almost constant (in absolute value) for all different combinations of the basis set and DFT functional, and in all cases, MLDFT correctly reproduces the trends obtained at the reference full DFT level.

The same considerations outlined above for GLY + 1w conformers also apply to GLY + 2w ones (see Figure 6b, middle panel). In fact, by moving from LDA to B3LYP and by including polarization and diffuse functions in the basis set, MLDFT errors with respect to DFT reference values decrease. The largest DFT-MLDFT discrepancy is reported for **2w-X** at the B3LYP/aug-cc-pVDZ level (−0.55 kcal/mol). This is due to the fact that the largest error is associated to the GS energy of the most stable conformer **2w-I** (see left panel of Figure 6a) for this combination of the DFT functional/basis set. However, as already reported for all the other studied systems, the error in the relative energies of the different conformers is always lower than the corresponding error in the total energies.

Finally, also in case of GLY + 3w clusters, the agreement between DFT and MLDFT is almost perfect, with errors ranging from −0.6 to 0.6 kcal/mol. The maximum error is observed for **3w-III** at the PBE/aug-cc-pVDZ level (0.53 kcal/mol), whereas the minimum is reported for **3w-VII** at the

https://dx.doi.org/10.1021/acs.jctc.0c00940
J. Chem. Theory Comput. 2021, 17, 791−803

B3LYP/6-31G* level (error < 0.01 kcal/mol). Therefore, also for these systems, MLDFT provides a reliable description of the relative energies of the different conformers. The only notable exceptions are conformers **3w-III** and **3w-IV** at the LDA/6-31G and B3LYP/6-31G levels, respectively. As a final comment, we note that although the MLDFT error on total GS energy can be larger than 1 kcal/mol, relative energies of the different conformers are accurately predicted, with an error that is always below 1 kcal/mol.

4.2. Toward a Realistic Picture of Solvation. In the previous sections, we have presented and discussed solute−solvent structures obtained by modeling the solvation phenomenon in aqueous solution by means of the so-called cluster approach,[90] in which only the closest water molecules are explicitly treated at the QM level. However, this picture is not realistic, being a strongly approximate way of modeling solvation. In fact, any dynamical aspect of solvation is neglected as well as, more importantly, long range interactions which are especially relevant for polar environments such as water. In this section, we show how MLDFT may be coupled to approaches that have been developed to model solvation more realistically. In particular, we will apply MLDFT to a randomly selected structure extracted from a classical MD simulation performed on both MOXY and GLY in aqueous solution. In this way, the atomistic details of solvation are retained, and dynamical aspects could easily be introduced by repeating the calculations on several structures. A closer investigation of the latter aspect is beyond the scope of our first work on MLDFT and will be the topic of further studies.

Let us start with MOXY. We have selected one random snapshot extracted from a MD simulation, which was previously reported by some of the present authors.[82,91,92] Note that MOXY is a rigid molecule; therefore, a single snapshot well represents its conformational structure (Figure 7).

Figure 7. Selected structure of MOXY + 50 water molecules, as extracted from MD. In MLDFT calculations, MOXY is the active part and water molecules are inactive.

In MLDFT calculations, MOXY is the active fragment and it is treated at the B3LYP/6-31+G* level. The inactive part is constituted by the 50 closest water molecules, which are described at the B3LYP/6-31G level. The reference full DFT calculation is instead performed by using the B3LYP functional, in combination with the 6-31+G* basis set for MOXY and the 6-31G one for water molecules.

In order to quantify the accuracy of MLDFT, we compute the solvation energy E_{solv}, which is defined as

$$E_{solv} = E_{tot} - E_{MOXY} - E_{w} \quad (12)$$

where E_{tot}, E_{MOXY}, and E_{w} are the total, MOXY, and water GS energies, respectively. Note that E_{MOXY} is calculated in the gas phase, and thus, it is the same in both full DFT and MLDFT calculations. E_{tot} and E_{w} are defined differently in the two approaches; in MLDFT, E_{w} is calculated at step 1 of the computational protocol (see Section 3), whereas in full DFT, it refers to the GS energy of the 50 water molecules.

Computed energy values for MOXY are reported in Table 1 for both DFT and MLDFT. We first notice that the MLDFT

Table 1. DFT and MLDFT Total GS Energies (E_{tot}) of MOXY + 50 Water Molecules Snapshot Depicted in Figure 7[a]

	DFT	MLDFT
E_{tot}	−4013.1956	−4013.1660
E_{MOXY}	−193.1079	−193.1079
E_{w}	−3820.0681	−3820.0382
E_{solv}	−0.0196	−0.0199

[a] E_{MOXY}, E_{w}, and E_{solv} are also reported. All values are given in Hartree.

error on the total energy E_{tot} is larger than what is found for clusters (see previous sections). This is not surprising because the error of the method scales with the number of the water molecules in the inactive part. Such discrepancies are primarily due to the neglect of polarization and charge-transfer interactions in the inactive solvent water molecules because their density matrix remains fixed in MLDFT. The largest contribution to the error on total energy is due to E_{w}. In fact, MLDFT E_{w} differs from full DFT of about the same extent as total energies. Such differences between MLDFT and DFT are reflected by the computed solvation energy, which can be taken as a measure of the accuracy of MLDFT. For the studied snapshot, the agreement between MLDFT and DFT is almost perfect (−12.30 vs −12.48 kcal/mol), and the error is of about 0.2 kcal/mol.

The same analysis may be applied to glycidol, for which the snapshots were extracted from MD simulations previously reported by some of us.[79] We recall that GLY is a flexible solute, of which the main conformers may be identified by means of two dihedral angles δ_1 and δ_2 [see Figure 8, panel (a)]. Seven most probable conformers have been selected [see Figure 8 panel (b)].

The MLDFT partition has been done so that GLY is the active fragment and treated at the B3LYP/6-31+G* level, whereas water molecules are inactive and described at the B3LYP/6-31G level. All the reference full DFT calculations are performed by using the B3LYP functional in combination with the 6-31+G* basis set for the solute and the 6-31G one for the water molecules.

The DFT and MLDFT energies (E_{tot}, E_{GLY}, E_{w}, and E_{solv}) are reported in Table S6 in the Supporting Information. Overall, MLDFT total energies are higher than DFT values of about 0.02−0.03 Hartree. The reasons of this discrepancy are the same as reported for MOXY.

The DFT and MLDFT solvation energies are graphically compared in Figure 9. We observe that all MLDFT values are almost in perfect agreement with the reference full DFT data. The average discrepancy is of about 0.7 kcal/mol (∼1 mH), with the largest discrepancy reported for conformer 5 (1.1 kcal/mol). Notice that in this study, we only include the GLY moiety in the active part. Similar calculations performed at the

https://dx.doi.org/10.1021/acs.jctc.0c00940
J. Chem. Theory Comput. 2021, 17, 791−803

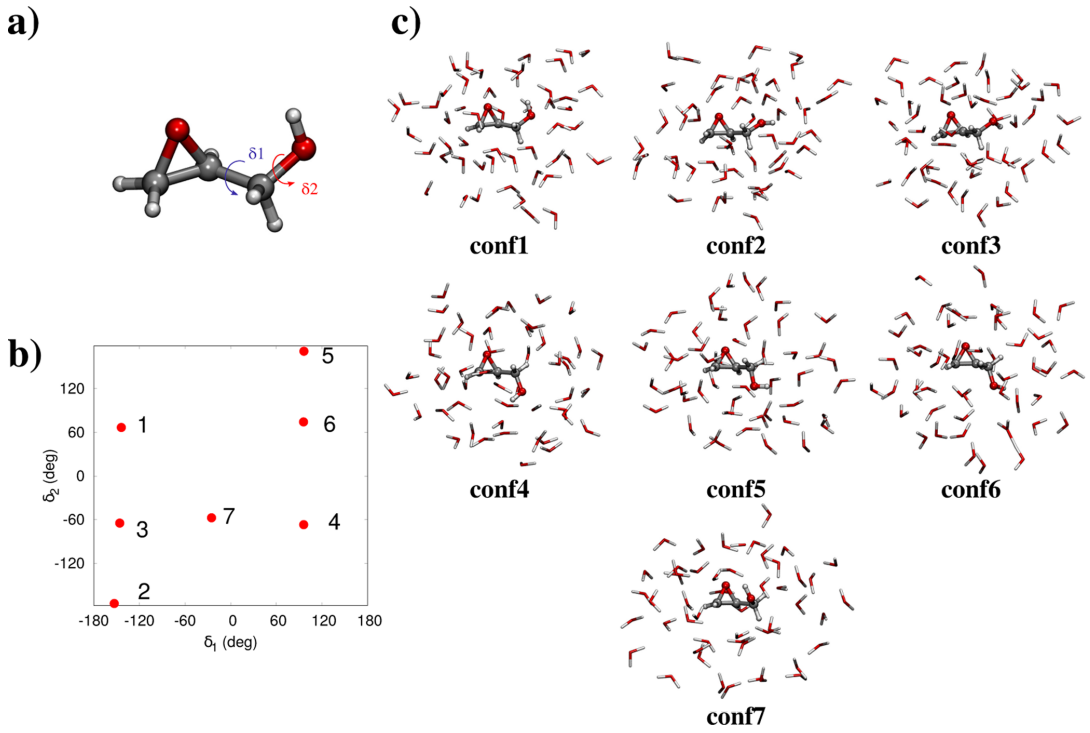


Figure 8. (a) Definition of δ_1 and δ_2 dihedral angles of GLY; (b) δ_1 and δ_2 values for the selected GLY + 50 water molecules snapshots extracted from MD; (c) molecular structures of the seven selected snapshots. In MLDFT calculations, GLY is the active part, whereas water molecules are inactive.

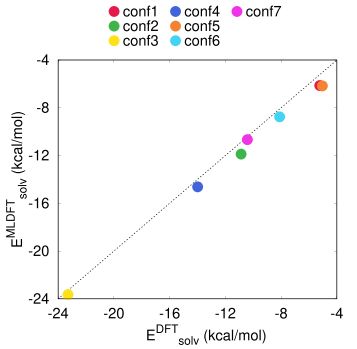


Figure 9. DFT and MLDFT solvation energies (E_{solv}) for the different conformers graphically depicted in Figure 8. Values are given in kcal/mol.

MLHF level[58] needed to insert at least five water molecules in the active fragment to reach the same level of accuracy.

5. SUMMARY AND CONCLUSIONS

In this work, we report a novel density-based multilevel approach based on the DFT treatment of the electronic structure problem. In MLDFT, the full system is partitioned in two layers, one active and one inactive. MLDFT stands apart from methods based on a similar paradigm because of the choice of partitioning an initial density matrix into active and inactive parts, instead of the density function. The MLDFT SCF procedure is then performed in the MO basis of the active subsystem only. This is the source of the reduction in computational cost because the density matrix of the inactive fragments is kept frozen during the optimization of the density.

The MLDFT was applied to aqueous methyloxirane and glycidol, for which two different approaches to solvation were discussed. First, the so-called cluster approach is employed, which models solvation in terms of minimal clusters composed of the solute and a small number of water molecules. Second, a more realistic picture is considered, which focuses on randomly selected snapshots extracted from MD simulations. For all studied structures, the computed data confirm that MLDFT is able to correctly reproduce reference full DFT values, with errors which are always $\lesssim$1 kcal/mol. Because of its favorable computational scaling, MLDFT can be coupled to more realistic approaches to solvation, that is, it can treat a large number of representative snapshots extracted from MD simulations, so to effectively take into account the dynamical aspects of solvation.

In this first presentation of the approach, we have limited the analysis to GS energies. However, MLDFT has the potential to be extended to the calculation of molecular properties and spectra. In particular, the analytical evaluation of molecular gradients will allow for MLDFT geometry optimizations. However, this extension is not trivial because in order to guarantee the continuity of the PES, the same pivots must be

https://dx.doi.org/10.1021/acs.jctc.0c00940
J. Chem. Theory Comput. 2021, 17, 791−803

imposed in the Cholesky decomposition. The calculation of molecular geometries and properties will be the topic of future communications. Also, the active and inactive parts can be described at two different levels of theory, for instance, using two different DFT functionals. Similarly to projection-based approaches, the active and inactive regions can also be treated at the HF and DFT levels, respectively, thus allowing for post-HF calculations on the active part only.[50] Such extensions will be the topic of future communications.

The method will also be further developed by focusing on some technical aspects, which are worth being improved. For instance, in the current implementation, the DFT grid is homogeneous in the whole space. However, it is reasonable to assume that the grid can be downgraded further away from the active part, because within the focused model paradigm, we only seek to accurately model the effect of the environment on the properties of the active subsystem, rather than the intrinsic properties of the environment itself. Technical refinements of the current implementation are in progress and will be discussed in future communications.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.0c00940.

Data related to Figures 2, 6, and 9 (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Tommaso Giovannini − *Department of Chemistry, Norwegian University of Science and Technology, 7491 Trondheim, Norway;* orcid.org/0000-0002-5637-2853; Email: tommaso.giovannini@ntnu.no

Henrik Koch − *Scuola Normale Superiore, 56126 Pisa, Italy;* orcid.org/0000-0002-8367-8727; Email: henrik.koch@sns.it

Authors

Gioia Marrazzini − *Scuola Normale Superiore, 56126 Pisa, Italy*

Marco Scavino − *Scuola Normale Superiore, 56126 Pisa, Italy*

Franco Egidi − *Scuola Normale Superiore, 56126 Pisa, Italy;* orcid.org/0000-0003-3259-8863

Chiara Cappelli − *Scuola Normale Superiore, 56126 Pisa, Italy;* orcid.org/0000-0002-4872-4505

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.0c00940

Author Contributions

[§]G.M. and T.G. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We acknowledge Susi Lehtola for useful comments on the manuscript. We acknowledge funding from the Marie Sklodowska-Curie European Training Network "COSINE—COmputational Spectroscopy In Natural sciences and Engineering", grant agreement no. 765739, and the Research Council of Norway through FRINATEK projects 263110 and 275506.

■ REFERENCES

(1) Dykstra, C.; Frenking, G.; Kim, K.; Scuseria, G. *Theory and Applications of Computational Chemistry: The First Forty Years*; Elsevier, 2011.

(2) Reichardt, C. Solvatochromism, thermochromism, piezochromism, halochromism, and chiro-solvatochromism of pyridinium N-phenoxide betaine dyes. *Chem. Soc. Rev.* **1992**, *21*, 147−153.

(3) Buncel, E.; Rajagopal, S. Solvatochromism and solvent polarity scales. *Acc. Chem. Res.* **1990**, *23*, 226−231.

(4) Reichardt, C. Solvatochromic dyes as solvent polarity indicators. *Chem. Rev.* **1994**, *94*, 2319−2358.

(5) Cannelli, O.; Giovannini, T.; Baiardi, A.; Carlotti, B.; Elisei, F.; Cappelli, C. Understanding the interplay between the solvent and nuclear rearrangements in the negative solvatochromism of a push−pull flexible quinolinium cation. *Phys. Chem. Chem. Phys.* **2017**, *19*, 32544−32555.

(6) Carlotti, B.; Cesaretti, A.; Cannelli, O.; Giovannini, T.; Cappelli, C.; Bonaccorso, C.; Fortuna, C. G.; Elisei, F.; Spalletti, A. Evaluation of Hyperpolarizability from the Solvatochromic Method: Thiophene Containing Push- Pull Cationic Dyes as a Case Study. *J. Phys. Chem. C* **2018**, *122*, 2285−2296.

(7) Cupellini, L.; Calvani, D.; Jacquemin, D.; Mennucci, B. Charge transfer from the carotenoid can quench chlorophyll excitation in antenna complexes of plants. *Nat. Commun.* **2020**, *11*, 1−8.

(8) Bondanza, M.; Cupellini, L.; Lipparini, F.; Mennucci, B. The multiple roles of the protein in the photoactivation of Orange Carotenoid Protein. *Chem* **2020**, *6*, 187−203.

(9) Lomize, A. L.; Pogozheva, I. D.; Lomize, M. A.; Mosberg, H. I. Positioning of proteins in membranes: a computational approach. *Protein Sci.* **2006**, *15*, 1318−1333.

(10) Furse, K. E.; Corcelli, S. A. The dynamics of water at DNA interfaces: Computational studies of Hoechst 33258 bound to DNA. *J. Am. Chem. Soc.* **2008**, *130*, 13103−13109.

(11) Zhao, S.; Wei, G. W. High-order FDTD methods via derivative matching for Maxwell's equations with material interfaces. *J. Comput. Phys.* **2004**, *200*, 60−103.

(12) Myhre, R. H.; Koch, H. The multilevel CC3 coupled cluster model. *J. Chem. Phys.* **2016**, *145*, 044111.

(13) Høyvik, I.-M.; Myhre, R. H.; Koch, H. Correlated natural transition orbitals for core excitation energies in multilevel coupled cluster models. *J. Chem. Phys.* **2017**, *146*, 144109.

(14) Parr, R. G. *Horizons of Quantum Chemistry*; Springer, 1980; pp 5−15.

(15) Burke, K. Perspective on density functional theory. *J. Chem. Phys.* **2012**, *136*, 150901.

(16) Sisto, A.; Stross, C.; van der Kamp, M. W.; O'Connor, M.; McIntosh-Smith, S.; Johnson, G. T.; Hohenstein, E. G.; Manby, F. R.; Glowacki, D. R.; Martinez, T. J. Atomistic non-adiabatic dynamics of the LH2 complex with a GPU-accelerated ab initio exciton model. *Phys. Chem. Chem. Phys.* **2017**, *19*, 14924−14936.

(17) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* **2005**, *105*, 2999−3094.

(18) Mennucci, B. Polarizable Continuum Model. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 386−404.

(19) Tomasi, J.; Persico, M. Molecular interactions in solution: an overview of methods based on continuous distributions of the solvent. *Chem. Rev.* **1994**, *94*, 2027−2094.

(20) Cappelli, C. Integrated QM/Polarizable MM/Continuum Approaches to Model Chiroptical Properties of Strongly Interacting Solute-Solvent Systems. *Int. J. Quantum Chem.* **2016**, *116*, 1532−1542.

(21) Warshel, A.; Karplus, M. Calculation of ground and excited state potential surfaces of conjugated molecules. I. Formulation and parametrization. *J. Am. Chem. Soc.* **1972**, *94*, 5612−5625.

(22) Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **1976**, *103*, 227−249.

(23) Senn, H. M.; Thiel, W. QM/MM methods for biomolecular systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198−1229.
(24) Lin, H.; Truhlar, D. G. QM/MM: what have we learned, where are we, and where do we go from here? *Theor. Chem. Acc.* **2007**, *117*, 185−199.
(25) Curutchet, C.; Muñoz-Losa, A.; Monti, S.; Kongsted, J.; Scholes, G. D.; Mennucci, B. Electronic energy transfer in condensed phase studied by a polarizable QM/MM model. *J. Chem. Theory Comput.* **2009**, *5*, 1838−1848.
(26) Olsen, J. M. H.; Kongsted, J. Molecular properties through polarizable embedding. *Adv. Quantum Chem.* **2011**, *61*, 107−143.
(27) Giovannini, T.; Puglisi, A.; Ambrosetti, M.; Cappelli, C. Polarizable QM/MM approach with fluctuating charges and fluctuating dipoles: the QM/FQFμ model. *J. Chem. Theory Comput.* **2019**, *15*, 2233−2245.
(28) Giovannini, T.; Lafiosca, P.; Cappelli, C. A General Route to Include Pauli Repulsion and Quantum Dispersion Effects in QM/MM Approaches. *J. Chem. Theory Comput.* **2017**, *13*, 4854−4870.
(29) Giovannini, T.; Lafiosca, P.; Chandramouli, B.; Barone, V.; Cappelli, C. Effective yet Reliable Computation of Hyperfine Coupling Constants in Solution by a QM/MM Approach: Interplay Between Electrostatics and Non-electrostatic Effects. *J. Chem. Phys.* **2019**, *150*, 124102.
(30) Giovannini, T.; Ambrosetti, M.; Cappelli, C. Quantum Confinement Effects on Solvatochromic Shifts of Molecular Solutes. *J. Phys. Chem. Lett.* **2019**, *10*, 5823−5829.
(31) Gordon, M. S.; Smith, Q. A.; Xu, P.; Slipchenko, L. V. Accurate first principles model potentials for intermolecular interactions. *Annu. Rev. Phys. Chem.* **2013**, *64*, 553−578.
(32) Gordon, M. S.; Slipchenko, L.; Li, H.; Jensen, J. H. The effective fragment potential: a general method for predicting intermolecular interactions. *Annu. Rep. Comput. Chem.* **2007**, *3*, 177−193.
(33) Sun, Q.; Chan, G. K.-L. Quantum embedding theories. *Acc. Chem. Res.* **2016**, *49*, 2705−2712.
(34) Knizia, G.; Chan, G. K.-L. Density matrix embedding: A strong-coupling quantum embedding theory. *J. Chem. Theory Comput.* **2013**, *9*, 1428−1432.
(35) Chulhai, D. V.; Goodpaster, J. D. Projection-based correlated wave function in density functional theory embedding for periodic systems. *J. Chem. Theory Comput.* **2018**, *14*, 1928−1942.
(36) Chulhai, D. V.; Goodpaster, J. D. Improved accuracy and efficiency in quantum embedding through absolute localization. *J. Chem. Theory Comput.* **2017**, *13*, 1503−1508.
(37) Wen, X.; Graham, D. S.; Chulhai, D. V.; Goodpaster, J. D. Absolutely Localized Projection-Based Embedding for Excited States. *J. Chem. Theory Comput.* **2020**, *16*, 385−398.
(38) Ding, F.; Manby, F. R.; Miller, T. F., III Embedded mean-field theory with block-orthogonalized partitioning. *J. Chem. Theory Comput.* **2017**, *13*, 1605−1615.
(39) Goodpaster, J. D.; Barnes, T. A.; Manby, F. R.; Miller, T. F., III Density functional theory embedding for correlated wavefunctions: Improved methods for open-shell systems and transition metal complexes. *J. Chem. Phys.* **2012**, *137*, 224113.
(40) Goodpaster, J. D.; Barnes, T. A.; Manby, F. R.; Miller, T. F., III Accurate and systematically improvable density functional theory embedding for correlated wavefunctions. *J. Chem. Phys.* **2014**, *140*, 18A507.
(41) Manby, F. R.; Stella, M.; Goodpaster, J. D.; Miller, T. F., III A simple, exact density-functional-theory embedding scheme. *J. Chem. Theory Comput.* **2012**, *8*, 2564−2568.
(42) Goodpaster, J. D.; Ananth, N.; Manby, F. R.; Miller, T. F., III Exact nonadditive kinetic potentials for embedded density functional theory. *J. Chem. Phys.* **2010**, *133*, 084103.
(43) Zhang, K.; Ren, S.; Caricato, M. Multi-state QM/QM Extrapolation of UV/Vis Absorption Spectra with Point Charge Embedding. *J. Chem. Theory Comput.* **2020**, *16*, 4361−4372.
(44) Ramos, P.; Papadakis, M.; Pavanello, M. Performance of frozen density embedding for modeling hole transfer reactions. *J. Phys. Chem. B* **2015**, *119*, 7541−7557.
(45) Pavanello, M.; Neugebauer, J. Modelling charge transfer reactions with the frozen density embedding formalism. *J. Chem. Phys.* **2011**, *135*, 234103.
(46) Tamukong, P. K.; Khait, Y. G.; Hoffmann, M. R. Density differences in embedding theory with external orbital orthogonality. *J. Phys. Chem. A* **2014**, *118*, 9182−9200.
(47) Tamukong, P. K.; Khait, Y. G.; Hoffmann, M. R. Accurate dissociation of chemical bonds using DFT-in-DFT embedding theory with external orbital orthogonality. *J. Phys. Chem. A* **2017**, *121*, 256−264.
(48) Wesolowski, T. A. Hydrogen-bonding-induced shifts of the excitation energies in nucleic acid bases: an interplay between electrostatic and electron density overlap effects. *J. Am. Chem. Soc.* **2004**, *126*, 11444−11445.
(49) Bennie, S. J.; Curchod, B. F. E.; Manby, F. R.; Glowacki, D. R. Pushing the limits of EOM-CCSD with projector-based embedding for excitation energies. *J. Phys. Chem. Lett.* **2017**, *8*, 5559−5565.
(50) Lee, S. J. R.; Welborn, M.; Manby, F. R.; Miller, T. F., III Projection-based wavefunction-in-DFT embedding. *Acc. Chem. Res.* **2019**, *52*, 1359−1368.
(51) Neugebauer, J.; Louwerse, M. J.; Baerends, E. J.; Wesolowski, T. A. The merits of the frozen-density embedding scheme to model solvatochromic shifts. *J. Chem. Phys.* **2005**, *122*, 094115.
(52) Wesolowski, T. A.; Shedge, S.; Zhou, X. Frozen-density embedding strategy for multilevel simulations of electronic structure. *Chem. Rev.* **2015**, *115*, 5891−5928.
(53) Fux, S.; Jacob, C. R.; Neugebauer, J.; Visscher, L.; Reiher, M. Accurate frozen-density embedding potentials as a first step towards a subsystem description of covalent bonds. *J. Chem. Phys.* **2010**, *132*, 164101.
(54) Jacob, C. R.; Neugebauer, J.; Visscher, L. A flexible implementation of frozen-density embedding for use in multilevel simulations. *J. Comput. Chem.* **2008**, *29*, 1011−1018.
(55) Jacob, C. R.; Visscher, L. Calculation of nuclear magnetic resonance shieldings using frozen-density embedding. *J. Chem. Phys.* **2006**, *125*, 194104.
(56) Jacob, C. R.; Neugebauer, J.; Jensen, L.; Visscher, L. Comparison of frozen-density embedding and discrete reaction field solvent models for molecular properties. *Phys. Chem. Chem. Phys.* **2006**, *8*, 2349−2359.
(57) Wesolowski, T. A. Embedding a multideterminantal wave function in an orbital-free environment. *Phys. Rev. A: At., Mol., Opt. Phys.* **2008**, *77*, 012504.
(58) Sæther, S.; Kjærgaard, T.; Koch, H.; Høyvik, I.-M. Density-Based Multilevel Hartree−Fock Model. *J. Chem. Theory Comput.* **2017**, *13*, 5282−5290.
(59) Høyvik, I.-M. Convergence acceleration for the multilevel Hartree−Fock model. *Mol. Phys.* **2020**, *118*, 1626929.
(60) Myhre, R. H.; Sánchez de Merás, A. M. J.; Koch, H. Multi-level coupled cluster theory. *J. Chem. Phys.* **2014**, *141*, 224105.
(61) Wesolowski, T. A.; Warshel, A. Frozen density functional approach for ab initio calculations of solvated molecules. *J. Phys. Chem.* **1993**, *97*, 8050−8053.
(62) Folkestad, S. D.; Koch, H. Equation-of-motion MLCCSD and CCSD-in-HF oscillator strengths and their application to core excitations. *J. Chem. Theory Comput.* **2020**, *16*, 6869.
(63) Giovannini, T.; Koch, H. Energy-Based Molecular Orbital Localization in a Specific Spatial Region. *J. Chem. Theory Comput.* **2020**, *17*, 139−150.
(64) Sánchez de Merás, A. M. J.; Koch, H.; Cuesta, I. G.; Boman, L. Cholesky decomposition-based definition of atomic subsystems in electronic structure calculations. *J. Chem. Phys.* **2010**, *132*, 204105.
(65) He, X.; Merz, K. M., Jr. Divide and conquer Hartree- Fock calculations on proteins. *J. Chem. Theory Comput.* **2010**, *6*, 405−411.
(66) Folkestad, S. D.; Kjønstad, E. F.; Myhre, R. H.; Andersen, J. H.; Balbi, A.; Coriani, S.; Giovannini, T.; Goletto, L.; Haugland, T. S.;

Hutcheson, A.; Høyvik, I.-M.; Moitra, T.; Paul, A. C.; Scavino, M.; Skeidsvoll, A. S.; Tveten, Å. H.; Koch, H. eT 1.0: An open source electronic structure program with emphasis on coupled cluster and multilevel methods. *J. Chem. Phys.* **2020**, *152*, 184103.
(67) Lebedev, V. I.; Laikov, D. *Doklady Mathematics*; Pleiades Publishing, Ltd., 1999; Vol. *59*, pp 477−481.
(68) Krack, M.; Köster, A. M. An adaptive numerical integrator for molecular integrals. *J. Chem. Phys.* **1998**, *108*, 3226−3234.
(69) Marques, M. A. L.; Oliveira, M. J. T.; Burnus, T. Libxc: A library of exchange and correlation functionals for density functional theory. *Comput. Phys. Commun.* **2012**, *183*, 2272−2281.
(70) Lehtola, S.; Steigemann, C.; Oliveira, M. J. T.; Marques, M. A. L. Recent developments in libxc—A comprehensive library of functionals for density functional theory. *SoftwareX* **2018**, *7*, 1−5.
(71) Aquilante, F.; Boman, L.; Boström, J.; Koch, H.; Lindh, R.; de Merás, A. S.; Pedersen, T. B. *Linear-Scaling Techniques in Computational Chemistry and Physics*; Springer, 2011; pp 301−343.
(72) Koch, H.; Sánchez de Merás, A.; Pedersen, T. B. Reduced scaling in electronic structure calculations using Cholesky decompositions. *J. Chem. Phys.* **2003**, *118*, 9481−9484.
(73) Kohn, W.; Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **1965**, *140*, A1133.
(74) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865.
(75) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(76) Giovannini, T.; Egidi, F.; Cappelli, C. Molecular Spectroscopy of Aqueous Solutions: A The- oretical Perspective. *Chem. Soc. Rev.* **2020**, *49*, 5664−5677.
(77) Giovannini, T.; Egidi, F.; Cappelli, C. Theory and algorithms for chiroptical properties and spectroscopies of aqueous systems. *Phys. Chem. Chem. Phys.* **2020**, *22*, 22864−22879.
(78) Giovannini, T.; Riso, R. R.; Ambrosetti, M.; Puglisi, A.; Cappelli, C. Electronic transitions for a fully polarizable qm/mm approach based on fluctuating charges and fluctuating dipoles: linear and corrected linear response regimes. *J. Chem. Phys.* **2019**, *151*, 174104.
(79) Giovannini, T.; Del Frate, G.; Lafiosca, P.; Cappelli, C. Effective computational route towards vibrational optical activity spectra of chiral molecules in aqueous solution. *Phys. Chem. Chem. Phys.* **2018**, *20*, 9181−9197.
(80) Lipparini, F.; Cappelli, C.; Barone, V. A gauge invariant multiscale approach to magnetic spectroscopies in condensed phase: General three-layer model, computational implementation and pilot applications. *J. Chem. Phys.* **2013**, *138*, 234108.
(81) Giovannini, T.; Grazioli, L.; Ambrosetti, M.; Cappelli, C. Calculation of ir spectra with a fully polarizable qm/mm approach based on fluctuating charges and fluctuating dipoles. *J. Chem. Theory Comput.* **2019**, *15*, 5495−5507.
(82) Giovannini, T.; Olszówka, M.; Egidi, F.; Cheeseman, J. R.; Scalmani, G.; Cappelli, C. Polarizable Embedding Approach for the Analytical Calculation of Raman and Raman Optical Activity Spectra of Solvated Systems. *J. Chem. Theory Comput.* **2017**, *13*, 4421−4435.
(83) Losada, M.; Nguyen, P.; Xu, Y. Solvation of propylene oxide in water: Vibrational circular dichroism, optical rotation, and computer simulation studies. *J. Phys. Chem. A* **2008**, *112*, 5621−5627.
(84) Yang, G.; Xu, Y. Probing chiral solute-water hydrogen bonding networks by chirality transfer effects: a vibrational circular dichroism study of glycidol in water. *J. Chem. Phys.* **2009**, *130*, 164506.
(85) Merten, C.; Bloino, J.; Barone, V.; Xu, Y. Anharmonicity Effects in the Vibrational CD Spectra of Propylene Oxide. *J. Phys. Chem. Lett.* **2013**, *4*, 3424−3428.
(86) Su, Z.; Xu, Y. Hydration of a Chiral Molecule: The Propylene Oxide··· $(Water)_2$ Cluster in the Gas Phase. *Angew. Chem., Int. Ed.* **2007**, *119*, 6275−6278.
(87) Yu, Z.; Xu, L.; Wei, Y.; Wang, Y.; He, Y.; Xia, Q.; Zhang, X.; Liu, Z. A new route for the synthesis of propylene oxide from bio-glycerol derivated propylene glycol. *Chem. Commun.* **2009**, 3934−3936.
(88) Su, Z.; Borho, N.; Xu, Y. Chiral self-recognition: Direct spectroscopic detection of the homochiral and heterochiral dimers of propylene oxide in the gas phase. *J. Am. Chem. Soc.* **2006**, *128*, 17126−17131.
(89) Su, Z.; Wen, Q.; Xu, Y. Conformational Stability of the Propylene Oxide- Water Adduct: Direct Spectroscopic Detection of O- H··· O Hydrogen Bonded Conformers. *J. Am. Chem. Soc.* **2006**, *128*, 6755−6760.
(90) Perera, A. S.; Thomas, J.; Poopari, M. R.; Xu, Y. The clusters-in-a-liquid approach for solvation: new insights from the conformer specific gas phase spectroscopy and vibrational optical activity spectroscopy. *Front. Chem.* **2016**, *4*, 9.
(91) Lipparini, F.; Egidi, F.; Cappelli, C.; Barone, V. The optical rotation of methyloxirane in aqueous solution: a never ending story? *J. Chem. Theory Comput.* **2013**, *9*, 1880−1884.
(92) Giovannini, T.; Olszòwka, M.; Cappelli, C. Effective Fully Polarizable QM/MM Approach To Model Vibrational Circular Dichroism Spectra of Systems in Aqueous Solution. *J. Chem. Theory Comput.* **2016**, *12*, 5483−5492.

https://dx.doi.org/10.1021/acs.jctc.0c00940
J. Chem. Theory Comput. 2021, 17, 791−803

CHAPTER 5

Integrated Multiscale Multilevel Approach to Open Shell Molecular Systems

Tommaso Giovannini[*,†], Gioia Marrazzini[†], Marco Scavino[†], Henrik Koch[†,‡] and Chiara Cappelli[*,†]

[†]*Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy.*
[‡]*Department of Chemistry, Norwegian University of Science and Technology, 7491 Trondheim, Norway.*

E-mail: tommaso.giovannini@sns.it; chiara.cappelli@sns.it

5.1. Abstract

We present a novel multiscale approach to study the electronic structure of open-shell molecular systems embedded in an external environment. The method is based on the coupling of multilevel Hartree-Fock (MLHF) and Density Functional Theory (MLDFT), suitably extended to the unrestricted formalism, to Molecular Mechanics (MM) force fields (FF). Within the ML region, the system is divided into active and inactive parts, thus describing the most relevant interactions (electrostatic, polarization and Pauli repulsion) at the quantum level. The surrounding MM part, which is formulated in terms of non-polarizable or polarizable FFs, permits a

physically consistent treatment of long-range electrostatics and polarization effects. The approach is extended to the calculation of hyperfine coupling constants, and applied to selected nitroxyl radicals in aqueous solution.

5.2. Introduction

The theoretical description of large molecular systems in the condensed phase at high level of accuracy is challenging, due to the substantial number of degrees of freedom (electronic and nuclear) that need to be treated computationally. However, the complexity can be drastically reduced by partitioning the total system into smaller, interacting, subsystems.[237–239] Solvated molecules, drugs interacting with biological matrices (e.g. DNA and proteins), or molecular systems adsorbed on metal surfaces are generally tackled in this way,[8,52,62,132,205] by resorting to "focused" computational approaches. There, the total system is partitioned into layers, which are treated at a different degree of sophistication.

The most used focused approaches are defined in the framework of QM/classical methods, where the attention is focused to the QM layer, whereas the rest of the system is described in terms of classical physics. Generally, the atomistic nature of the whole system is retained, such as in QM/molecular mechanics (QM/MM) approaches.[1,8] The interaction between the QM and classical moieties is modelled in terms of classical electrostatics and, in some cases, mutual polarization effects are considered.[62] Only few examples exist, where purely QM interactions, such as Pauli repulsion and dispersion forces, are coherently introduced in the QM/classical modeling,[58,128,133,176] even though they play a crucial role in many systems.[240,241] Also, the quality of QM/classical methods strongly depends on the quality of the MM description, which is generally determined by the availability and reliability of parameter sets.[242]

As an alternative to QM/classical methods, quantum embedding approaches can be exploited.[67,69,70,75,77,79–81,243–246] There, both subsystems are described quantum-mechanically, generally at different level of accuracy. The advantage of quantum vs classical embedding is twofold: (i) target-environment interactions are treated at the QM level and, (ii) a full QM description does not require any parametrization, therefore quantum embedding approaches

can potentially be applied to any kind of interacting systems at the same level of accuracy. The price to pay consists of a generally larger computational cost of quantum vs classical embedding, that may limit, even substantially, the size of actually treatable systems. For this reason, the development of computationally effective yet physically consistent approaches is mandatory. The recently proposed Multilevel HF (MLHF)[193] and Multilevel DFT (MLDFT)[192] approaches, which are coherently rooted in the context of Hartree-Fock (HF) and Density Functional Theory (DFT), are a remarkable example of this class of methods.[192,193,247–249] MLHF and MLDFT lay on common theoretical foundations, being based on a decomposition of the total density matrix into active and inactive contributions.[192,193] There, HF or DFT equations are reformulated in terms of active and inactive density matrices, and the molecular orbitals (MO) coefficients of the active part only are optimized in the Self Consistent Field (SCF) procedure. The inactive density matrix is kept fixed and gives rise to an effective field, which interacts with the active part (see Refs.[193] and[192] for more details).

The partitioning of the system that is featured in MLHF/DFT allows to substantially extend the size of chemical systems that may be afforded by HF and especially DFT. However, this gain in size may not be sufficient to treat realistic systems in the condensed phase, i.e. surrounded by an external environment. To this end, quantum embedding approaches may benefit from the coupling with an outer layer described in terms of classical physics, e.g. by means of MM force fields.[250,251] Remarkably, the coupling is physically grounded, because long-range interactions are dominated by electrostatics and polarization (and dispersion, which is however described neither by HF nor by DFT). Such forces can effectively be described at low computational cost by polarizable MM force fields. In the resulting approach, electrostatics, polarization and Pauli repulsion interactions are accurately described at the QM level within the MLHF/MLDFT region, whereas long-range interactions are effectively taken into account by means of a polarizable MM level.[250] As a consequence, the novel class of methods gains advantage from the physico-chemical features of both approaches.

In this paper, we integrate classical MM force fields,[1,8] either within the electrostatic or polarizable embedding schemes (the latter based on the Fluctuating Charge (FQ) force field[52,62,110,205,252]), with a novel class of

multilevel approaches, which extend MLHF and MLDFT to open-shell systems. They are based on a unrestricted formulation, which exploits a partition of both α and β spin-densities in active and inactive contributions. Therefore, the computational saving with respect to full HF or DFT descriptions is achieved by keeping frozen the inactive spin-density matrices, which only perturb active densities. The resulting unrestricted MLHF (UMLHF) and UMLDFT are able to account for electrostatic (polarization) forces and Pauli repulsion between active and inactive QM regions, whereas long-range electrostatic and polarization terms are effectively taken into account at a low computational cost through the interaction with the FQ layer. To show the potentialities of the resulting UMLHF(UMLDFT)/MM(FQ) method, it is challenged to compute hyperfine coupling constants (hcc) of selected molecular spin-probes.[253–258] Since hccs are particularly sensitive to the probes' external environment,[259–261] they represent an ideal test bed for the novel multiscale multilevel approach.
The manuscript is organized as follows. In the next section we report the theoretical derivation of UMLHF(UMLDFT)/MM(FQ). The method is then applied to simulate hcc of Nitrogen atom of PROXYL and TEMPO radicals in aqueous solution, by exploiting a hierarchy of classical embedding approaches. A summary and a discussion of the future perspective of this work end the manuscript.

5.3. Theory

In this section, the fundamentals of UMLHF(UMLDFT)/MM(FQ) are reported. We first focus on the extension of MLHF/MLDFT to the unrestricted case, thus defining UMLHF and UMLDFT. Then, the way they are coupled with an outer classical, atomistic layer is detailed, by specifying the method either within the electrostatic or polarizable embedding scheme. In the latter case, the coupling with the FQ force field is discussed. Last, the approach is developed for the calculation of hyperfine coupling constants.

5.3.1. Unrestricted MLHF and MLDFT

The starting point to derive UMLHF/UMLDFT is the expression of the energy $E[\mathbf{D}^\alpha, \mathbf{D}^\beta]$ for open-shell systems in the unrestricted formalism:

$$\begin{aligned} E[\mathbf{D}^\alpha, \mathbf{D}^\beta] &= E[\mathbf{D}^\alpha] + E[\mathbf{D}^\beta] \\ &= \mathrm{Tr}\mathbf{h}\mathbf{D}^\alpha + \frac{1}{2}\mathrm{Tr}\mathbf{D}^\alpha\mathbf{J}(\mathbf{D}^\alpha + \mathbf{D}^\beta) \\ &\quad - c_x\left(\frac{1}{2}\mathrm{Tr}\mathbf{D}^\alpha\mathbf{K}(\mathbf{D}^\alpha)\right) \\ &\quad + (1 - c_x)\int \rho^\alpha(\mathbf{r})\varepsilon_x[\rho(\mathbf{r})]\mathrm{d}\mathbf{r} \\ &\quad + \int \rho^\alpha(\mathbf{r})\varepsilon_c[\rho(\mathbf{r})]\mathrm{d}\mathbf{r} \\ &+ \mathrm{Tr}\mathbf{h}\mathbf{D}^\beta + \frac{1}{2}\mathrm{Tr}\mathbf{D}^\beta\mathbf{J}(\mathbf{D}^\alpha + \mathbf{D}^\beta) \\ &\quad - c_x\left(\frac{1}{2}\mathrm{Tr}\mathbf{D}^\beta\mathbf{K}(\mathbf{D}^\beta)\right) \\ &\quad + (1 - c_x)\int \rho^\beta(\mathbf{r})\varepsilon_x[\rho(\mathbf{r})]\mathrm{d}\mathbf{r} \\ &\quad + \int \rho^\beta(\mathbf{r})\varepsilon_c[\rho(\mathbf{r})]\mathrm{d}\mathbf{r} \end{aligned} \tag{5.1}$$

$E[\mathbf{D}^\alpha, \mathbf{D}^\beta]$ is split into two terms, E^α and E^β, which refer to α and β spin. The $\mathbf{D}^\sigma$ are spin density matrices (with $\sigma = \alpha, \beta$), $\mathbf{h}$ is the one-electron operator, and $\mathbf{J}$ and $\mathbf{K}$ are Coulomb and exchange matrices, respectively. Equation 5.1 is formulated for a generic DFT functional, where $\rho^\sigma(\mathbf{r})$ are spin DFT density functions and $\varepsilon_x, \varepsilon_c$ indicate exchange and correlation energy densities per unit particle. The UHF equations can easily be recovered by imposing $c_x = 1$ and $\varepsilon_c = 0$. The total density matrix $\mathbf{D}$ can be obtained from α and β spin density matrices as $\mathbf{D} = \mathbf{D}^\alpha + \mathbf{D}^\beta$.

Similarly to the closed shell case,[192,193] unrestricted MLHF/MLDFT are formulated by separating the total system into active (A) and inactive (B) parts. From a mathematical point of view, the separation is performed by decomposing spin-density matrices into active $\mathbf{D}^\sigma_A$ and inactive $\mathbf{D}^\sigma_B$ contributions:

$$\begin{aligned} \mathbf{D}^{\alpha} &= \mathbf{D}^{\alpha}_{A} + \mathbf{D}^{\alpha}_{B} \qquad \Rightarrow \qquad \rho^{\alpha}(\mathbf{r}) = \rho^{\alpha}_{A}(\mathbf{r}) + \rho^{\alpha}_{B}(\mathbf{r}) \\ \mathbf{D}^{\beta} &= \mathbf{D}^{\beta}_{A} + \mathbf{D}^{\beta}_{B} \qquad \Rightarrow \qquad \rho^{\beta}(\mathbf{r}) = \rho^{\beta}_{A}(\mathbf{r}) + \rho^{\beta}_{B}(\mathbf{r}) \end{aligned} \tag{5.2}$$

A similar partitioning applies to DFT spin density functions $(\rho^{\alpha}(\mathbf{r}), \rho^{\beta}(\mathbf{r}))$. Note that in general $\mathbf{D}^{\alpha}_{X} \neq \mathbf{D}^{\beta}_{X}$, $\{X = A, B\}$. By substituting Eq. 5.2 into Eq. 5.1, we obtain:

$$E[\mathbf{D}^{\alpha}_{A}, \mathbf{D}^{\alpha}_{B}, \mathbf{D}^{\beta}_{A}, \mathbf{D}^{\beta}_{B}] = E[\mathbf{D}^{\alpha}_{A}, \mathbf{D}^{\alpha}_{B}] + E[\mathbf{D}^{\beta}_{A}, \mathbf{D}^{\beta}_{B}] \tag{5.3}$$

where $E[\mathbf{D}^{\sigma}_{A}, \mathbf{D}^{\sigma}_{B}]$ $(\{\sigma = \alpha, \beta\})$ is given by:

$$\begin{aligned} &E[\mathbf{D}^{\sigma}_{A}, \mathbf{D}^{\sigma}_{B}] = \\ &\qquad E^{\sigma}_{A} \begin{cases} \mathrm{Tr}\mathbf{h}\mathbf{D}^{\sigma}_{A} + \frac{1}{2}\mathrm{Tr}\mathbf{D}^{\sigma}_{A}\mathbf{J}(\mathbf{D}_{A}) \\ -\frac{1}{2}c_{x}\mathrm{Tr}\mathbf{D}^{\sigma}_{A}\mathbf{K}(\mathbf{D}^{\sigma}_{A}) + \int \rho^{\sigma}_{A}(\mathbf{r})\varepsilon_{xc}(\rho_{A}(\mathbf{r}))\mathrm{d}\mathbf{r} \end{cases} \\ &\qquad E^{\sigma}_{B} \begin{cases} +\mathrm{Tr}\mathbf{h}\mathbf{D}^{\sigma}_{B} + \frac{1}{2}\mathrm{Tr}\mathbf{D}^{\sigma}_{B}\mathbf{J}(\mathbf{D}_{B}) \\ -\frac{1}{2}c_{x}\mathrm{Tr}\mathbf{D}^{\sigma}_{B}\mathbf{K}(\mathbf{D}^{\sigma}_{B}) + \int \rho^{\sigma}_{B}(\mathbf{r})\varepsilon_{xc}(\rho_{B}(\mathbf{r}))\mathrm{d}\mathbf{r} \end{cases} \\ &\qquad E^{\sigma}_{int} \begin{cases} +\mathrm{Tr}\mathbf{D}^{\sigma}_{A}\mathbf{J}(\mathbf{D}_{B}) - c_{x}\mathrm{Tr}\mathbf{D}^{\sigma}_{A}\mathbf{K}(\mathbf{D}^{\sigma}_{B}) \\ +\int \rho^{\sigma}_{A}(\mathbf{r})\varepsilon_{xc}(\rho_{B}(\mathbf{r}))\mathrm{d}\mathbf{r} + \int \rho^{\sigma}_{B}(\mathbf{r})\varepsilon_{xc}(\rho_{A}(\mathbf{r}))\mathrm{d}\mathbf{r} \end{cases} \\ E^{\sigma}_{non-add} &\begin{cases} +\int \rho^{\sigma}(\mathbf{r})\varepsilon_{xc}(\rho(\mathbf{r}))\mathrm{d}\mathbf{r} \\ -\int \rho^{\sigma}(\mathbf{r})\varepsilon_{xc}(\rho_{A}(\mathbf{r}))\mathrm{d}\mathbf{r} - \int \rho^{\sigma}(\mathbf{r})\varepsilon_{xc}(\rho_{B}(\mathbf{r}))\mathrm{d}\mathbf{r} \end{cases} \end{aligned} \tag{5.4}$$

In Eq. 5.4, $\varepsilon_{c} + (1 - c_{x})\varepsilon_{x}$ is substituted by ε_{xc} to make the notation compact, and energy terms are separated into active and inactive contributions $(E^{\sigma}_{A,B})$. Also, the coupling terms are divided into a purely interaction energy term, E^{σ}_{int}, and a non-additive contribution $E^{\sigma}_{non-add}$, which originates from the non linearity of ε_{x} and ε_{c}. It is worth noting that the non-additive energy terms vanish for UMLHF. As expected, the partitioning of the different terms in Eq. 5.4 is coherent with the MLDFT formulation for closed shell systems (see Ref.[192]).

As already discussed in the Introduction, the energy of the active fragment A is optimized while the inactive density B is kept fixed to the value resulting from the partitioning in Eq. 5.2. Therefore, the UMLHF/MLDFT spin-Fock matrix ($F^{\sigma}_{\mu\nu}$, in the AO basis $\{\chi_\mu\}$) can easily be recovered by differentiating the energy in Eq. 5.4 with respect to the active density ($D^{\sigma}_{A}, \rho^{\sigma}_{A}$), i.e.:

$$
\begin{aligned}
F^{\sigma}_{\mu\nu} &= h^{\sigma}_{\mu\nu} \\
2e^{\sigma}_{A} &\begin{cases} +J_{\mu\nu}(\mathbf{D}_A) - c_x K_{\mu\nu}(\mathbf{D}^{\sigma}_{A}) \\ +\int v^{\sigma}_{xc}(\rho^{A}(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\,\mathrm{d}\mathbf{r} \end{cases} \\
2e^{\sigma}_{B} &\begin{cases} +J_{\mu\nu}(\mathbf{D}_B) - c_x K_{\mu\nu}(\mathbf{D}^{\sigma}_{B}) \\ +\int v^{\sigma}_{xc}(\rho^{B}(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\,\mathrm{d}\mathbf{r} \end{cases} \\
2e^{\sigma}_{non-add} &\begin{cases} +\int [v^{\sigma}_{xc}(\rho(\mathbf{r})) - v^{\sigma}_{xc}(\rho^{A}(\mathbf{r})) - v^{\sigma}_{xc}(\rho^{B}(\mathbf{r}))] \\ \qquad \chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\,\mathrm{d}\mathbf{r} \end{cases}
\end{aligned}
\tag{5.5}
$$

where we have used the compact notation $v_{xc} = (1 - c_x)v_x + v_c$. The two-electron contributions to the Fock matrix can be are grouped into different $2e_X$ terms, with X $= A, B$. $2e^{\sigma}_{non-add}$ is due to the non-linearity in the DFT functional, and again vanishes for UMLHF. Finally, note that $2e^{\sigma}_{B}$ accounts for the frozen fragment, of which the density ($D_B, \rho^{B}(\mathbf{r})$) does not change along SCF cycles. Therefore, $2e^{\sigma}_{B}$ is a constant one-electron contribution, which is computed only once, at the beginning of the SCF procedure, similarly to MLHF/MLDFT.[192,193]

We further point out that UMLHF and MLDFT equations directly follow from the partitioning of total spin-densities into active and inactive contributions. Similarly to their close shell counterparts, their accuracy depends on the approach which is exploited to carry out the decomposition in Eq. 5.2. Different choices are possible, however in the present work the initial set of active occupied molecular orbitals (MOs) are obtained through a partial limited Cholesky decomposition of $\mathbf{D}^{\sigma}$.[195,202] The procedure is detailed in Ref.,[249] however its extension to open-shell systems is not straightforward. In fact, for UMLHF/MLDFT the Cholesky decomposition needs to be performed twice, i.e. for both $\mathbf{D}^{\sigma}_{A}$ ($\sigma = \alpha, \beta$) spin-densities. $\mathbf{D}^{\sigma}_{A}$ can be written in the AO basis $\{\mu, \nu\}$ as follows:[262]

$$D^{\sigma}_{A,\mu\nu} = \sum_{IJ} D^{\sigma}_{\mu I} \left(\tilde{D}^{\sigma}_{IJ}\right)^{-1} D^{\sigma}_{\nu J}$$
$$= \sum_{I} L^{\sigma}_{\mu I} L^{\sigma}_{\nu I} \tag{5.6}$$

where the diagonal elements I and J are decomposed, the $\tilde{\mathbf{D}}$ submatrix contains the selected diagonal elements, and $L_{\alpha I}$ are the Cholesky orbitals. Diagonal elements are selected so that pivots correspond to the AOs belonging to a pre-defined set of active atoms. As a result of the decomposition, the active Cholesky occupied MOs are obtained and the active spin-density matrices $\mathbf{D}^{\sigma}_{A}$ are trivially constructed (see Eq. 5.6). The active virtual space can be defined in terms of projected atomic orbitals (PAOs),[263,264] which are obtained by projecting out occupied components from the subset of AOs centered on the active atoms. Possible linear dependencies are removed by Löwdin orthonormalization (see also Ref.[248]).
In the practical implementation, UML calculations follow this protocol:

1. Generation of the guess AO densities ($\mathbf{D}^{\alpha}$ and $\mathbf{D}^{\beta}$) by superposition of atomic densities (SAD).

2. Construction of initial Fock Matrices from the SAD density and diagonalization, so to obtain initial idempotent spin-densities.

3. Partitioning of the spin-density matrices (α and β) into active A and inactive B spin-densities ($\mathbf{D}^{\alpha}_{A}$, $\mathbf{D}^{\beta}_{A}$ and $\mathbf{D}^{\alpha}_{B}$, $\mathbf{D}^{\beta}_{B}$). As discussed in the previous section, for occupied orbitals this step is performed by partial Cholesky decomposition. Virtual orbitals are instead obtained by means of PAOs or by decomposing the virtual density by using the Cholesky algorithm. Notice that the whole procedure generates orthogonal MOs between the active and inactive parts. For the active part, the resulting MOs are used to transform matrices from the AO to the MO basis of the active part only.

4. Calculation of constant energy terms and one-electron contributions due to the spin-density matrices of the inactive part (see Eqs. 5.4 and 5.5).

5. Energy minimization (Eq. 5.4) in the MO basis of the active part only, until convergence is reached.

5.3.2. Coupling with an outer layer described with MM force fields

The UMLHF(UMLDFT)/MM(FQ) is defined by starting from the total energy of the system, i.e.:

$$\mathcal{E} = E_{\text{UMLHF(DFT)}} + E_{\text{MM}} + E^{\text{int}}_{\text{UMLHF(DFT)/MM}} \ , \tag{5.7}$$

where $E_{UMLHF(DFT)}$ is given in Eq. 5.4, whereas E_{MM} and $E^{\text{int}}_{\text{UMLHF(DFT)/MM}}$ are MM and UMLHF(DFT)/MM interaction energies, respectively. Electrostatic and polarizable QM/MM embedding differ from the way the interaction energy is specified (in our case $E^{\text{int}}_{\text{UMLHF(DFT)/MM}}$); electrostatic embedding approaches limit the description to electrostatic forces only, whereas mutual QM/MM polarization is modelled in polarizable embedding approaches.[8,62] In particular, non-polarizable embedding methods place fixed charges on MM atoms, which polarize the QM density. Different polarizable QM/MM approaches exist;[56,58,127,129,131,203,204] in this work we exploit QM/Fluctuating Charges (FQ),[62,205] where each MM atom is assigned an atomic electronegativity (χ) and chemical hardness (η), which give rise to electric charges (q) as a response to the atomic chemical potential.[52,62,152–154] Therefore, for both non-polarizable UMLHF(DFT)/MM or polarizable UMLHF(DFT)/FQ, the UMLHF (DFT)/MM interaction energy can be written as follows:

$$E^{\text{int}}_{\text{UMLHF(DFT)/MM}} = \sum_i q_i V_i(\mathbf{D}^\alpha + \mathbf{D}^\beta) \ , \tag{5.8}$$

where $V_i(\mathbf{D}^\alpha + \mathbf{D}^\beta)$ is the electric potential generated by the total QM spin-density (i.e. both active and inactive contributions) on the i$-th$ charge (q_i). In the case of non-polarizable QM/MM, q_i values are fixed, whereas in QM/FQ they are obtained by minimizing the following energy expression:

$$\begin{aligned}\mathcal{E}[\mathbf{D}^\alpha_A, \mathbf{D}^\beta_A, \mathbf{D}^\alpha_B, \mathbf{D}^\beta_B, \mathbf{q}, \boldsymbol{\lambda}] = & E_{\text{UMLHF(DFT)}}[\mathbf{D}^\alpha_A, \mathbf{D}^\alpha_B] \\ & + E_{\text{UMLHF(DFT)}}[\mathbf{D}^\beta_A, \mathbf{D}^\beta_B] \\ & + \frac{1}{2}\mathbf{q}^\dagger_\lambda \mathbf{M} \mathbf{q}_\lambda + \mathbf{q}^\dagger_\lambda \mathbf{C}_Q \\ & + \mathbf{q}^\dagger_\lambda \mathbf{V}(\mathbf{D}^\alpha + \mathbf{D}^\beta),\end{aligned} \tag{5.9}$$

where $E_{\text{UMLHF(DFT)}}[\mathbf{D}_A^\alpha, \mathbf{D}_B^\alpha]$ and $E_{\text{UMLHF(DFT)}}[\mathbf{D}_A^\beta, \mathbf{D}_B^\beta]$ due to the total α- and β spin-densities, respectively (see Eq. 5.4). In Eq. 5.9, $\mathbf{q}_\lambda$ indicates a vector collecting FQ charges and a set of Lagrangian multipliers, which ensure charge conservation on each fragment composing the MM layer (e.g. on each solvent molecules for solvated systems). The $\mathbf{M}$ matrix is the interaction kernel between the FQ charges, which also contains the Lagrangian blocks,[110] and the vector $\mathbf{C}_Q$ accounts for the interaction between permanent moments, i.e. χ and charge constraints Q on each FQ moiety.
The FQ charges equilibrated for the UMLHF(DFT)/FQ systems are obtained by minimizing the energy functional in Eq. 5.9. This procedure yields the following set of linear equations:

$$\mathbf{M}\mathbf{q}_\lambda = -\mathbf{C}_Q - \mathbf{V}(\mathbf{D}^\alpha + \mathbf{D}^\beta) \ . \tag{5.10}$$

In parallel, UMLHF(DFT)/MM spin Fock matrices are defined as follows:

$$F_{\mu\nu}^\sigma = F_{\mu\nu}^{\sigma,\text{UMLHF(DFT)}} + \sum_i q_i V_{i,\mu\nu} \ . \tag{5.11}$$

where $F_{\mu\nu}^{\sigma,\text{UMLHF(DFT)}}$ is defined in Eq. 5.5. As for energy contributions, the additional QM/MM term is fixed and computed only at the first SCF cycle in case of non-polarizable QM/MM, whereas it changes in QM/FQ, because FQ charges depend on QM densities. Therefore, UMLHF(DFT)/FQ contribution to the Fock matrix needs to be updated at each SCF cycle, thus introducing mutual polarization effects between UMLHF(DFT) and FQ layers.

5.3.3. Hyperfine Coupling Constants

To highlight the potentialities of UMLHF(DFT)/MM(FQ), in this paper we focus on hyperfine coupling constants. The hyperfine interaction between electron spin $\mathbf{S}$ and nuclear spin $\mathbf{I}$ can be calculated in terms of the hyperfine coupling tensor $\mathbf{A}$, which for a given nucleus X reads:[235,236]

$$\mathbf{A}(X) = \mathrm{A_X}\mathbf{1}_3 + \mathbf{A}_{\text{dip}}(\mathrm{X}) \tag{5.12}$$

In Eq. 5.12, the dipolar contribution $\mathbf{A}_{\text{dip}}(\mathrm{X})$ is a zero-trace tensor and vanishes in isotropic media (e.g. solutions). The $\mathrm{A_X}$ denotes the so-called hyperfine coupling constant (hcc), which is connected to the spin density (ρ_X) at nucleus X through the following equation:

$$A_X = \frac{4\pi}{3}\mu_B\mu_X g_e g_X \langle S_Z\rangle^{-1} \rho_X^{\alpha-\beta} \tag{5.13}$$

where μ_B is the Bohr magneton, g_e is the free electron g-factor (g_e = 2.0022319), whereas μ_X and g_X refers to nucleus X. The $\rho_X^{\alpha-\beta}$ reads:

$$\rho_X^{\alpha-\beta} = \sum_{\mu\nu} D_{\mu\nu}^{\alpha-\beta} \langle\chi_\mu(\mathbf{r})|\delta(\mathbf{r}-\mathbf{r}_X)|\chi_\nu(\mathbf{r})\rangle \tag{5.14}$$

$D^{\alpha-\beta}$ is the difference between α and β density matrices. The UMLHF(DFT) hccs are computed for the active part only. Therefore, the difference between α and β density matrices in Eq. 5.14 only refers to active spin-density matrices, which are calculated by minimizing UMLHF(DFT) energy in Eqs. 5.4-5.5. When the outer MM layer is considered, the active spin-densities entering Eq. 5.14 are obtained by minimizing the UMLHF(DFT)/MM energy (see Eqs. 5.9 -5.11).

5.4. Computational Details

In the following sections, UMLDFT is applied to the calculation of hcc_N of PROXYL and TEMPO nitroxyl radicals in aqueous solution (see Fig. 5.1). Such systems are characterized by the presence of the N-O group, which has been amply exploited as "spin probe" for structural studies of macro molecular systems.[258] In particular, in order to take into account the dynamical aspects of the solvation phenomenon, we resort to the computational protocol suitably designed by us for the study of aqueous systems.[62] First, classical MD simulations of the spin labels in water are performed according to Ref.[177] in order to accurately sample the phase-space. From the classical trajectories, 200 uncorrelated snapshots are extracted and spherical droplets centered on the solutes' center of mass are cut (radius = 13 Å).

In line with previous studies,[250] for each snapshot the radical (TEMPO or PROXYL) and the water shell closer than 3.5 Å with respect to the solute center of mass are described at the UMLDFT level (on average, 30 water molecules are included in the UMLDFT layer). The remaining water molecules are classically described by means of the polarizable FQ force field, by exploiting the parameters reported in Ref.[177] To test the quality of the

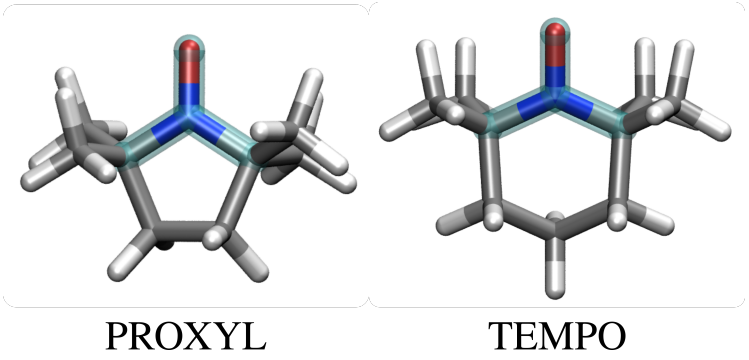


Figure 5.1. PROXYL (left) and TEMPO (right) molecular structures. The atoms involved in the out-of-plane angle are highlighted in cyan in both radicals.

UMLDFT description, different active/inactive partitions are considered by including 0, 2, 5 or 10 water molecules in the active part (MLDFT$_{n\mathrm{w}}$, $n \in [0, 2, 5, 10]$). Such molecules are the closest to the radical center of mass (C.M.) or the N-O group (-NO). The remaining water molecules in the UMLDFT region are treated as inactive. The radical is treated at the PBE0/N07D[265] level, whereas the QM water molecules (either active or inactive) are described at the PBE0/6-31G level. We denote this combination as PBE0/N07D/6-31G(w). Moreover, virtual orbitals are determined by using the PAOs algorithm.
Finally, PROXYL and TEMPO hcc$_{\mathrm{N}}$ are calculated as an average of the 200 uncorrelated snapshots. All calculations are performed by using a development version of the electronic structure code e^T.[266]

5.5. Numerical Applications

Testing of the computational approach To first demonstrate the reliability of the computational approach, a random snapshot is extracted from MD simulations, and hcc$_{\mathrm{N}}$ are computed. To quantify the role of long-range electrostatics interactions, hcc$_{\mathrm{N}}$ are also computed by removing the MM layer (see Fig. 5.2a-b). The model systems exploited in the analysis are depicted In Fig. 5.2a-b; TEMPO is colored in red, UMLDFT water molecules in cyan, and the MM layer in grey. Additional calculations are performed by exploiting the N07D basis set to describe all the system, thus quantifying the effect of polarization and diffuse functions on water

molecules, as compared to PBE0/N07D/6-31G(w) calculations (full and patterned boxes in Fig. 5.2c-d). Also, comparison with calculations where the UMLDFT region is treated at the full DFT level is proposed (labeled as fullDFT, fullDFT/TIP3P and fullDFT/FQ, respectively), that in order to quantify the quality of the ML partitioning. All the results obtained with the different models are summarized in Fig. 5.2 for TEMPO (see also Tab. S1 in the Supporting Information - SI). The corresponding data for PROXYL are given in Fig. S1 in the SI, together with the associated raw data in Tab. S2. We first focus on the effect of including additional water molecules in the active fragment. Independently of the basis set exploited, TEMPO and PROXYL hcc_N are not particularly affected by nw. In fact, a maximum difference of 0.1 Gauss between 0w and 10w is reported for both radicals, independently of the method used to select the active water molecules (i.e. with respect to the N-O group or the radical C.M.). These data suggest that the most relevant short-range solute-solvent interactions are correctly taken into account by UMLDFT. Long-range electrostatics (including polarization) plays instead a crucial role, with contributions ranging from 0.85 to 1.3 Gauss when the water molecules are described by means of the polarizable FQ force field (see Fig. 5.2e). When the non-polarizable TIP3P force field is instead used, only a minor shift of the computed hcc_N with respect to the values computed for the small cluster is reported (see Fig. 5.2d). This clearly shows the crucial role of long-range polarization effects in the description of hcc_N of solvated radical species. The data depicted in Fig. 5.2c-e also demonstrate that the water molecules described at the QM level can be treated by using the cheaper 6-31G basis set (i.e. without the need of including polarization and diffuse functions), being the largest difference between full and UML PBE0/N07D or PBE0/N07D/6-31G(w) results less than 0.1 Gauss.

5.5.1. PROXYL and TEMPO in aqueous solution

On the basis of the benchmarking results reported above, UMLDFT/PBE0/N07D/6-31G(w)/FQ is then applied to the calculation of hcc_N of both PROXYL and TEMPO on the whole set of 200 uncorrelated snapshots extracted from MD runs. Computed UMLDFT$_{n\text{w}}$/FQ results are

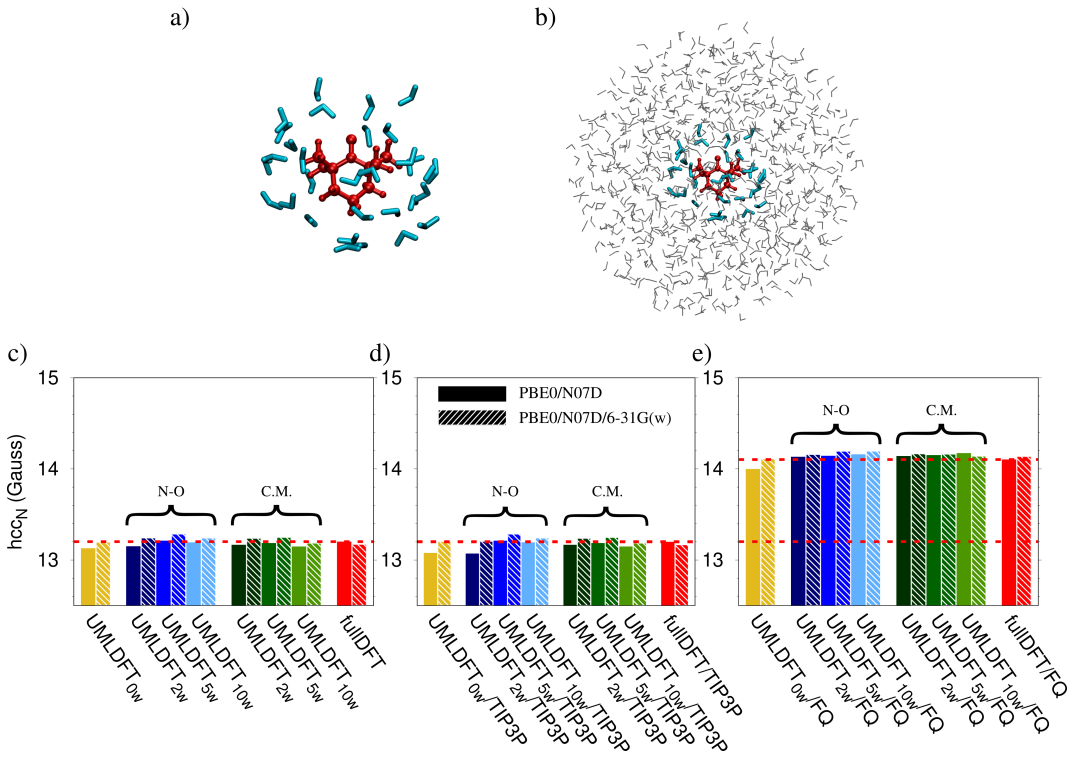


Figure 5.2. Molecular structure of the randomly selected TEMPO snapshot as described at the UMLDFT (a) and UMLDFT/MM (b) levels. (c-e) $\mathrm{hcc_N}$ (Gauss) calculated at the UMLDFT$_{n\mathrm{w}}$(/TIP3P,/FQ) and full DFT(/TIP3P, /FQ) levels. Water molecules included in the UMLDFT layer are selected with respect to the N-O group or C.M. Horizontal red lines correspond to full PBE0/N07D (c-e), PBE0/N07D/TIP3P (d) and PBE0/N07D/FQ (e) results.

reported in Tab. 5.1, together with associated standard errors (*se*) at 67% confidence interval. The latter are computed as:

$$se = \frac{\sigma}{\sqrt{N_{snap}}} \tag{5.15}$$

where σ is the standard deviation for the 200 uncorrelated snapshots (N_{snap}).

		PROXYL	TEMPO
	UMLDFT$_{\mathrm{NO}}$/FQ	16.0 ± 0.1	16.5 ± 0.1
	UMLDFT$_{\mathrm{0w}}$/FQ	15.3 ± 0.1	16.2 ± 0.1
N-O	UMLDFT$_{\mathrm{2w}}$/FQ	15.3 ± 0.1	16.4 ± 0.1
	UMLDFT$_{\mathrm{5w}}$/FQ	15.4 ± 0.1	16.2 ± 0.1
	UMLDFT$_{\mathrm{10w}}$/FQ	15.4 ± 0.1	16.2 ± 0.1
C.M.	UMLDFT$_{\mathrm{2w}}$/FQ	15.3 ± 0.1	16.3 ± 0.1
	UMLDFT$_{\mathrm{5w}}$/FQ	15.3 ± 0.1	16.2 ± 0.1
	UMLDFT$_{\mathrm{10w}}$/FQ	15.3 ± 0.1	16.2 ± 0.1
	"Best" QM/MM[267]	15.5± 0.1	16.3± 0.1
	Exp.	16.4[268]	17.3[269]

Table 5.1. Calculated UMLDFT$_{n\mathrm{w}}$/FQ hcc$_{\mathrm{N}}$ (Gauss) average values for PROXYL and TEMPO in aqueous solution. Values are averaged over 200 uncorrelated snapshots extracted from MD runs. "Best" QM/MM results are reproduced from Ref.,[267] whereas experimental data are taken from Refs.[268,269]

We first note that the inclusion of water molecules in the active region only marginally affects the computed hcc$_{\mathrm{N}}$ values, in agreement with the benchmark analysis discussed above for the random snapshot. We also investigate the dependence of hcc$_{\mathrm{N}}$ on the out of plane dihedral angle involving C-C-N-O atoms, i.e. the Nitrogen atom pyrimidalization (see Fig. 5.1), which has previously been reported to crucially affect the description of hcc$_{\mathrm{N}}$ of the studied radical species.[177,270] Computed data at the UMLDFT$_{\mathrm{0w}}$/FQ level are depicted in Fig. 5.3 for PROXYL (left) and TEMPO (right). hcc$_{\mathrm{N}}$ hugely varies as a function of the OOP angle, ranging from about 12 to 26 Gauss. Note that such a variability is the main reason

why 200 snapshots are required to converge the property (see top panel in Fig. 5.3), and also demonstrates that a reliable sampling of the radical/solvent phase-space is required to reliable model this property.

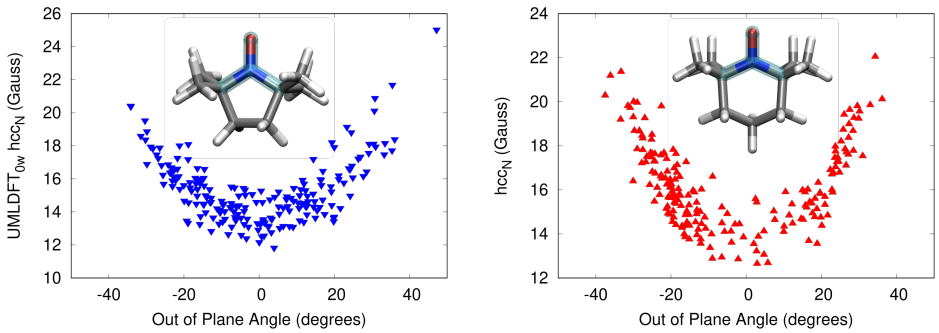


Figure 5.3. Computed UMLDFT$_{0w}$/FQ hcc$_N$ values (Gauss) for PROXYL (left) and TEMPO (right) in aqueous solution as a function of the out of plane angle (see inset).

The data in Tab. 5.1 show that for both radicals, UMLDFT$_{0w}$/FQ and UMLDFT$_{10w}$/FQ provide almost the same computed hcc$_N$, independently of how the active water molecules are selected (i.e. with respect to the N-O group or the C.M.). This is in line with previous observations reported in Fig. 5.2. Such results can be further investigated by plotting the difference between hcc$_N$ values computed by exploited the two approaches (Δhcc$_N$), as a function of the OOP angle (see Fig. 5.4, where UMLDFT$_{10w}$(N-O) is considered). For most snapshots the two approaches give similar hcc$_N$ values (Δhcc$_N$ = 0 Gauss). However, large differences, ranging from -0.7 to 1.1 Gauss, with a standard deviation of 0.3 Gauss, are reported for specific values of the out-of-plane angle for both PROXYL and TEMPO. Fig. 5.4 also reports absolute values of the error between the two approaches, together with their gaussian convolution. For TEMPO (left panel of Fig. 5.4), the largest discrepancies are reported for the region near $\pm$20 degrees, and the gaussian convolution shows a minimum at about 0 degrees. A different situation occurs for PROXYL, for which maxima are located at -20 and 10 degrees. These results show that although average hcc$_N$ values computed by UMLDFT$_{0w}$ and UMLDFT$_{10w}$(N-O) are similar for both radical species, a large fluctuation is reported as a function of the snapshot. This confirms that an appropriate sampling of the phase-space is required, and that the dynamical aspects of the solvation phenomenon need to

accurately be taken into account. In addition, it is worth noting that additional polarization effects which are considered by including the closest water molecules in the active region, do not play a crucial role in determining the final average results. However, the present analysis clearly show that such effects need to be accurately investigated when multilevel methodologies are employed, as already reported by some of us for excitation energies.[250]

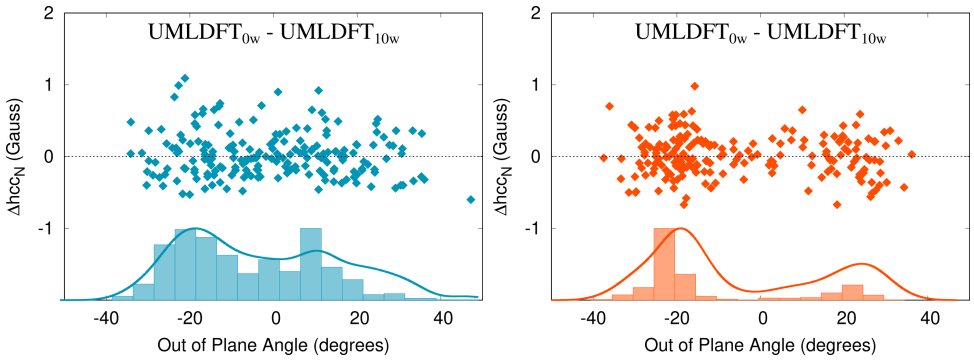


Figure 5.4. Computed UMLDFT$_{0w}$/FQ - UMLDFT$_{10w}$(N-O) Δhcc$_N$ (Gauss) for PROXYL (left) and TEMPO (right) in aqueous solution as a function of the out-of-plane angle. Histograms and gaussian convolutions of the absolute Δhcc$_N$ values are also given.

Covalently bonded fragments The results reported in the previous sections assume active and inactive regions to be non-covalently bonded. However, UMLDFT can in principle be applied to covalently bonded fragments. To demonstrate the method's potentialities, the active space is reduced to the N-O group only, and the remaining atoms of the radicals are included in the inactive UMLDFT region, together with selected water molecules. This approach is denoted as UMLDFT$_{NO}$/FQ, and computed hcc$_N$ are reported in Tab. 5.1. UMLDFT$_{NO}$/FQ results differ of about 0.7 (PROXYL) and 0.3 (TEMPO) Gauss with respect to UMLDFT$_{10w}$/FQ data, which can be taken as reference. However, as already reported above, average values may hinder larger complexity from the solvation dynamics point of view. For this reason, in Fig. 5.5 UMLDFT$_{NO}$/FQ-UMLDFT$_{10w}$(N-O)/FQ differences as a function of the OOP angle are reported for PROXYL (left) and TEMPO (right). Absolute error distributions are also given as histograms and their convolution with a gaussian-type function is plotted.
Both PROXYL and TEMPO Δhcc$_N$ distributions are comparable to those

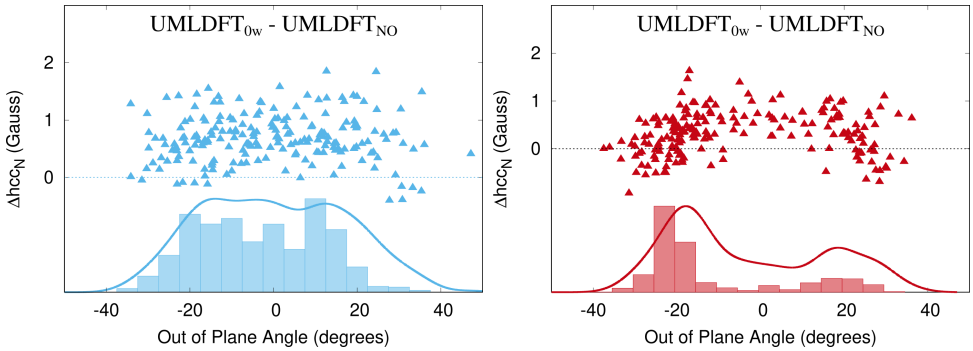


Figure 5.5. Computed UMLDFT$_{10w}$(N-O)/FQ - UMLDFT$_{NO}$ Δhcc$_N$ (Gauss) as a function of the out of plane angle for PROXYL (left) and TEMPO (right) in aqueous solution. Histograms and gaussian convolutions of absolute Δhcc$_N$ values are also given.

already commented in Fig. 5.4. However, in the present case Δhcc$_N$ values show very large variability, from -0.6 to 1.6 Gauss, with a standard deviation of about 0.5 and 0.4 Gauss for PROXYL and TEMPO, respectively. The UMLDFT is able to provide average hcc$_N$ values in good agreement with reference results by only considering a minimal active portion. Such features may be particularly useful for the extension of the approach to correlated Hamiltonians, which can be exploited to describe the active fragment.

Comparison with experimental data We finally move to compare computed and experimental data (see Table 5.1). It has been reported in previous studies that DFT has strong limits at reproducing hcc$_N$ values, due to inappropriate account of electron correlation.[177,271] However UMLDFT results are in good agreement with most accurate QM/MM results reported by some of us,[177] and which are labeled as "best QM/MM" in Table 5.1. Such data were obtained by describing water molecules at the FQ level, with the further (and substantial) inclusion of Pauli repulsion and dispersion contributions, as computed at the QM level.[128,133,176,177] The discrepancy between UMLDFT and experimental data is similar to "best QM/MM" results (1.1 $\pm$ 0.2 Gauss for PROXYL and 1.1 $\pm$ 0.2 Gauss). A proper account of electron correlation, e.g. by resorting to correlated coupled cluster calculations, is expected to reduce the computed error.[177] However, computed UMLDFT/FQ differences between PROXYL and TEMPO hcc$_N$ values (0.9 Gauss) are perfectly in agreement with experimental findings (0.9

Gauss), thus demonstrating that our approach can reliably describe different radical species.

5.6. Summary, Conclusions and Future Perspectives

In this work, we have introduced a novel class of multiscale QM/classical approaches aimed at describing the electronic properties of open shell systems. The methods are based on the coupling of multilevel HF/DFT, which are extended to the unrestricted formalism, with an outer region described at the classical MM level. Similarly to MLHF and MLDFT, UML methodologies are based on a partition of the QM layer into an active and an inactive part. The partitioning is performed on the initial spin-density matrices through a partial Cholesky decomposition of the occupied MOs, and virtual orbitals are obtained by means of PAOs. Note that the active occupied MOs may be further refined by means of a localization procedure targeted to specific molecular regions assigned to the active fragment. To this purpose, the energy-based approach developed by some of us can be extended to the unrestricted formalism.[249] The UML methods substantially reduce the computational cost associated with common ab-initio calculations, because only the active subsystem MOs enter the SCF procedure, whereas the inactive density is kept fixed.

UMLHF/DFT are coupled to an outer MM layer described in terms of non-polarizable or polarizable force fields. In this way, not only the computational cost is further reduced, but a correct physico-chemical description of the main interactions is preserved. Indeed, the MM part allows for an effective modeling of long-range electrostatics (and polarization forces) in a multiscale fashion.

To test the quality of the approaches, they are applied to compute hcc_N of PROXYL and TEMPO radicals dissolved in aqueous solution. First, the quality of the designed computational protocol is tested on a single snapshot, demonstrating the necessity of including long-range polarization by comparing UMLDFT/TIP3P and UMLDFT/FQ results. Then, PROXYL and TEMPO hcc_N are calculated as an average on a set of uncorrelated snapshots extracted from classical MD runs, which allow to correctly take

into account the dynamical aspect of the solvation phenomenon. The computed data are perfectly in agreement with the best computational estimates proposed in the literature for the same systems and correctly reproduce the experimental findings, thus demonstrating the reliability of the developed methods for real-case systems. Also, we showcase the flexibility of UMLDFT/MM partitioning by including the nitroxyl group only in the active part, i.e. by considering covalently bonded fragments. Although in this case computed hcc_N are not perfectly in agreement with our best estimates, such a flexibility paves the way to the extension to a coupled cluster treatment of the active part, that in order to take into account electron correlation, which may largely affect the electronic properties of open-shell systems.[177,271]

On the other hand, the developed approach can also be extended to treat linear response properties of open-shell systems by means of time-dependent DFT (TD-DFT) formulations. Also, UMLDFT/MM is here tested on aqueous solutions, however the model is general enough to be applied to different solvents,[242] or different embedding environments, such as biological matrices or nanostructured materials. Finally, UMLDFT may also be coupled to more sophisticated polarizable force fields, which improve the description of specific anisotropic interactions.[129]

5.7. Acknowledgments

T.G. and C.C. acknowledge funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 818064). H.K. acknowledges funding through the Research Council of Norway through FRINATEK projects 263110 and 275506, and funding from the ERC under the European Union's Horizon 2020 Research and Innovation Programme (grant agreement No. 101020016). We gratefully acknowledge the Center for High Performance Computing (CHPC) at SNS for providing the computational infrastructure.

5.8. Supporting Information

hcc_N for PROXYL and TEMPO calculated at the UMLDFT$_{nw}$(/TIP3P; /FQ) levels, and associated graphical representation for PROXYL.

		PBE0/N07D			PBE0/N07D/6-31G(w)		
	UMLDFT$_{nw}$	Gas-phase	FQ	TIP3P	Gas-phase	FQ	TIP3P
	0w	13.1	14.0	13.1	13.2	14.1	13.2
N-O	2w	13.2	14.1	13.1	13.2	14.2	13.2
N-O	5w	13.2	13.7	13.2	13.3	14.2	13.3
N-O	10w	13.2	13.7	13.2	13.2	14.2	13.2
CM	2w	13.2	13.7	13.2	13.2	14.2	13.2
CM	5w	13.2	13.7	13.2	13.2	14.2	13.3
CM	10w	13.1	13.6	13.1	13.2	14.1	13.2
	Full DFT	13.2	14.1	13.8	13.2	14.1	13.8

Table 5.2. Calculated UMLDFT$_{nw}$(/TIP3P; /FQ) hcc_N for TEMPO, which different selections of active water molecules.

		PBE0/N07D			PBE0/N07D/6-31G(w)		
	UMLDFT$_{nw}$	Gas-phase	FQ	TIP3P	Gas-phase	FQ	TIP3P
	0w	18.6	18.4	18.7	18.8	19.1	18.8
N-O	2w	18.5	19.0	18.6	18.7	19.0	18.8
N-O	5w	18.4	18.8	18.4	18.5	18.8	18.5
N-O	10w	18.4	18.9	18.5	18.5	18.7	18.5
CM	2w	18.6	19.0	18.7	18.6	18.8	18.7
CM	5w	18.5	19.0	18.6	18.7	18.9	18.7
CM	10w	18.4	18.9	18.4	18.5	18.7	18.5
	Full DFT	18.3	18.8	18.5	18.5	18.9	18.7

Table 5.3. Calculated UMLDFT$_{nw}$(/TIP3P; /FQ) hcc_N for PROXYL, which different selections of active water molecules.

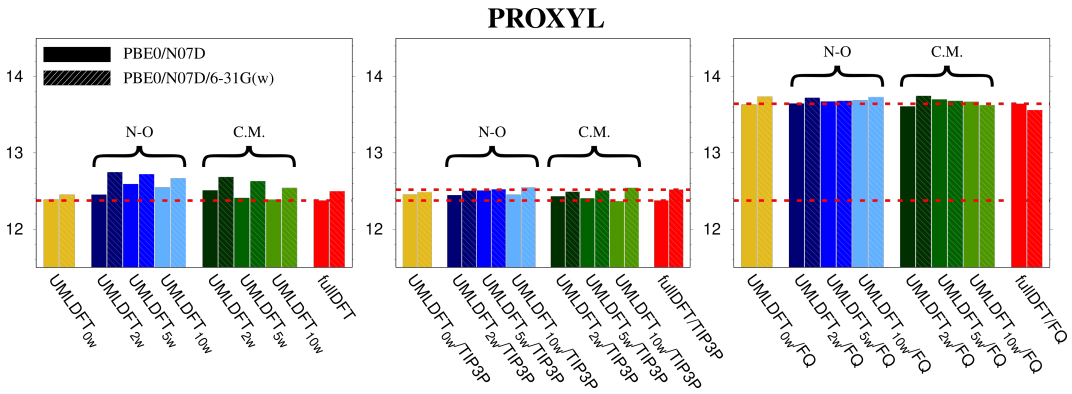


Figure 5.6. hcc_N (Gauss) calculated at the UMLDFT$_{n\mathrm{w}}$(/TIP3P,/FQ) and full DFT(/TIP3P, /FQ) levels for PROXYL. Water molecules included in the UMLDFT layer are selected with respect to the N-O group or C.M. Horizontal red lines correspond to full PBE0/N07D, PBE0/N07D/TIP3P and PBE0/N07D/FQ results.

CHAPTER 6

Summary, Conclusions and Future Perspectives

This thesis focused on the development and application of different multiscale models, which differ on the way the interactions between the target molecule and the environment are defined. The proposed methodologies belong to two different families: (i) QM/MM and (ii) quantum embedding approaches. By focusing on the first class, we have introduced different QM/MM embeddings with the associated interaction potentials, ranging from mechanical to polarizable embeddings, in which the mutual polarization between the QM and MM parts is taken into account. Among them, we deep in the theoretical details of two polarizable QM/MM approaches, namely QM/FQ and QM/FQFμ. In QM/FQ, each MM atom is endowed with a charge, which varies as a response to the external potential. In QM/FQFμ an additional dipole moment is placed at the atoms' positions, and linearly dependent on the external field. In this way, anisotropic interactions can accurately be described. From an algorithmic point of view, the QM/FQFμ response matrix is four times bigger than the QM/FQ one ($\mathbf{T}^{qq}$ block). As a consequence, by inverting the response matrix, QM/FQ can treat 64 times bigger systems than QM/FQFμ at the same computational cost.

As commented above, QM/FQFμ provides a refined description of QM/MM interactions with respect to QM/FQ and other polarizable embedding models. However, also in this case, QM/MM interactions are limited to an electrostatic description, therefore neglecting both Pauli repulsion and

dispersion effects, which are intrinsically of QM nature. The common strategy to include these forces is by calculating the Lennard-Jones potential. Contrary to the electrostatic term, which acts as an external potential in the Hamiltonian, the Lennard-Jones energy is merely added to the total energy of the system, because it does not depend on the QM density. Thus, its inclusion only gives indirect contributions to molecular properties, because it does not directly affect the Hamiltonian, its derivatives and response equations.

Few strategies have been proposed to include dispersion and repulsion effects in a QM/MM framework. Among them, in **Paper 1**, the QM/MM quantum Pauli-repulsion framework developed by our research group,[176,177] and based on the explicit calculation of the exchange integral between the QM and a ficticious MM density, has been applied to the calculation of polarizabilities and first-hyperpolarizability of organic molecules in aqueous solutions. In this work, QM/MM electrostatics is classically described by using both FQ and FQFμ force fields. The robustness of the approach is demonstrated for different organic acids dissolved in aqueous solution. The computed results clearly show that, as expected, polarizabilities and first-hyperpolarizabilities are particularly sensitive to the external environment. Indeed, the final numerical value is the result of a delicate balance between electrostatics, hydrogen-bonding, and non-electrostatic target-embedding interactions. Only by a physically consistent description of these interactions, as provided by our hybrid QM/MM approach, reliable results can be obtained. The reported reliability of the approach paves the way for its extensive application to different properties and spectroscopies, as for instance Raman and Hyper-Raman. In fact, the two aforementioned spectroscopies explicitly depend on the geometrical derivatives of polarizabilities and hyper-polarizabilities. Therefore, an accurate description of these two properties is mandatory to yield a reliable modeling of such spectroscopies.

In **Paper 2**, we have introduced a novel quantum embedding approach within the framework of density functional theory, named multilevel DFT. Similarly to its HF counterpart, in MLDFT the system is partitioned into an active and an inactive fragment. While the active density is optimized in the field of the inactive fragment, the inactive density remains frozen. In this way, the computational cost is intrinsically reduced. Also, additional saving

is obtained by solving the SCF equations in the MO basis of the active part only. MLDFT shares many similarities with other quantum embedding methods based on DFT, such as the FDE method. However, in MLDFT the active-inactive partitioning is performed on the density matrix, whereas in the FDE formalism it is obtained at the density function level. In FDE, this causes the need of defining a non-additive kinetic potential term that enforces Pauli exclusion between the electrons of the various subsystems. On the other hand, in MLDFT the density matrix decomposition needs to be performed by means of a decomposition algorithm, which may be mathematically well-defined, but may not have any particular physical meaning. However, if the decomposition is applied to the fully converged DFT density matrix, then, the MLDFT energy corresponds to the exact DFT energy of the full system, thus being completely independent of the decomposition algorithm.

The introduction of MLDFT paves the way for an accurate description of many different properties and spectroscopies of molecular systems embedded in generic external environments. In order to apply the method to open-shell subsystem, in **Paper 3**, ML-based methods (MLHF and MLDFT) are extended to the unrestricted formalism. Also, in order to further extend the applicability of ML-based methods, we have proposed their coupling with an additional MM layer, which can be either polarizable (at the FQ level) or non-polarizable. Such an extension is indeed physically-grounded: while short-range interactions may be characterized by purely quantum effects, such as Pauli repulsion, and a quantum description is therefore a natural choice, long-range interactions are usually dominated by electrostatics, which can accurately be described at the MM level. The results presented in **Paper 3** clearly show that the proposed UMLHF(DFT)/MM approach can accurately describe the properties of open-shell systems, such as hyperfine coupling constants, which are particularly sensitive to the external environment.

To conclude, in this Thesis, we have laid the foundations of different theoretical approaches based either on QM/MM or on multilevel partitioning of embedded system. Such methods paves the way for a physically-consistent and cost-effective description of the properties of complex systems. In particular, the MLDFT/MM approach developed in **Paper 3** is particularly

promising, because it retains the most relevant physical interactions at an accurate level of theory in all regions. In fact, within the core region described at the MLDFT level, the interactions (electrostatics, polarization and Pauli repulsion) are quantum-mechanically described. Long-range terms, which can be reasonably approximated by electrostatics (and polarization), are accurately described at the MM level. Dispersion is the only missing interaction energy term, which is not accounted for in our modeling. The further inclusion of such a term is particularly challenging, because it is related to a second-order expansion of the intermolecular interaction. However, its consideration is expected to further increase the accuracy of the proposed methods to describe real systems.

Bibliography

[1] Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227–249.

[2] Warshel, A.; Karplus, M. *J. Am. Chem. Soc.* **1972**, *94*, 5612–5625.

[3] Warshel, A. *Ann. Rev. Bioph. Biom.* **2003**, *32*, 425–443.

[4] Gao, J.; Xia, X. *Science* **1992**, *258*, 631–635.

[5] Vreven, T.; Morokuma, K.; Farkas, Ö.; Schlegel, H. B.; Frisch, M. J. *J. ComputĊhem* **2003**, *24*, 760–769.

[6] Steindal, A. H.; Ruud, K.; Frediani, L.; Aidas, K.; Kongsted, J. *J. Phys. Chem. B* **2011**, *115*, 3027–3037.

[7] Senn, H. M.; Thiel, W. *Curr. Opin. Chem. Biol.* **2007**, *11*, 182–187.

[8] Senn, H. M.; Thiel, W. *Angew. Chem. Int. Ed.* **2009**, *48*, 1198–1229.

[9] Lin, H.; Truhlar, D. G. *Theor. Chem. Acc.* **2007**, *117*, 185–199.

[10] Vreven, T.; Byun, K. S.; Komaromi, I.; Dapprich, S.; Montgomery, J. A. J.; Morokuma, K.; Frisch, M. J. *J.Chem. Theory and Comput.* **2006**, *2*, 815–826.

[11] Dapprich, S.; Komaromi, I.; Byun, K.; Morokuma, K.; Frisch, M. *J. Mol. Struct. - Theochem* **1999**, *461*, 1–21.

[12] Vreven, T.; Mennucci, B.; da Silva, C.; Morokuma, K.; Tomasi, J. *J. Chem. Phys.* **2001**, *115*, 62–72.

[13] Rega, N.; Cossi, M.; Barone, V. *J. Am. Chem. Soc.* **1998**, *120*, 5723–5732.

[14] Barone, V.; Bloino, J.; Monti, S.; Pedone, A.; Prampolini, G. *Phys. Chem. Chem. Phys.* **2010**, *12*, 10550–10561.

[15] Barone, V.; Biczysko, M.; Brancato, G. In *Combining Quantum Mechanics and Molecular Mechanics. Some Recent Progresses in QM/MM Methods*; Sabin, J. R., Canuto, S., Eds.; Advances in Quantum Chemistry; Academic Press, 2010; Vol. 59; pp 17–57.

[16] Rega, N.; Brancato, G.; Barone, V. *Chem. Phys. Lett.* **2006**, *422*, 367–371.

[17] Rega, N.; Brancato, G.; Petrone, A.; Caruso, P.; Barone, V. *J. Chem. Phys.* **2011**, *134*, 074504.

[18] Thompson, M. A. *J. Phys. Chem.* **1996**, *100*, 14492–14507.

[19] Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700–733.

[20] Gao, J.; Freindorf, M. *J. Phys. Chem. A* **1997**, *101*, 3182–3188.

[21] Martín, M. E.; Sánchez, M. L.; del Valle, F. J. O.; Aguilar, M. A. *J. Chem. Phys.* **2000**, *113*, 6308–6315.

[22] Cui, Q.; Karplus, M. *J. Chem. Phys.* **2000**, *112*, 1133–1149.

[23] Cui, Q.; Karplus, M. *J. Phys. Chem. B* **2000**, *104*, 3721–3743.

[24] Cappelli, C.; Mennucci, B.; Monti, S. *J. Phys. Chem. A* **2005**, *109*, 1933–1943.

[25] Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, *100*, 10580–10594.

[26] Schwabe, T.; Olsen, J. M. H.; Sneskov, K.; Kongsted, J.; Christiansen, O. *J. Chem. Theory Comput.* **2011**, *7*, 2209–2217.

[27] Nielsen, C. B.; Christiansen, O.; Mikkelsen, K. V.; Kongsted, J. *J. Chem. Phys.* **2007**, *126*, 154112.

[28] Kongsted, J.; Osted, A.; Mikkelsen, K. V.; Christiansen, O. *J. Chem. Phys.* **2003**, *118*, 1620–1633.

[29] Sneskov, K.; Schwabe, T.; Christiansen, O.; Kongsted, J. *Phys. Chem. Chem. Phys.* **2011**, *13*, 18551–18560.

[30] Steindal, A. H.; Olsen, J. M. H.; Ruud, K.; Frediani, L.; Kongsted, J. *Phys. Chem. Chem. Phys.* **2012**, *14*, 5440–5451.

[31] Marini, A.; Muñoz-Losa, A.; Biancardi, A.; Mennucci, B. *J. Phys. Chem. B* **2010**, *114*, 17128–17135.

[32] Jacobson, L. D.; Herbert, J. M. *J. Chem. Phys.* **2010**, *133*, 154506.

[33] Curutchet, C.; Muñoz-Losa, A.; Monti, S.; Kongsted, J.; Scholes, G. D.; Mennucci, B. *J. Chem. Theory Comput.* **2009**, *5*, 1838–1848.

[34] Scholes, G. D.; Curutchet, C.; Mennucci, B.; Cammi, R.; Tomasi, J. *J. Phys. Chem. B* **2007**, *111*, 6978–6982.

[35] Curutchet, C.; Kongsted, J.; Muñoz-Losa, A.; Hossein-Nejad, H.; Scholes, G. D.; Mennucci, B. *J. Am. Chem. Soc.* **2011**, *133*, 3078–3084.

[36] Bryce, R. A.; Vincent, M. A.; Malcolm, N. O. J.; Hillier, I. H.; Burton, N. A. *J. Chem. Phys.* **1998**, *109*, 3077–3085.

[37] Bryce, R. A.; Vincent, M. A.; Hillier, I. H. *J. Phys. Chem. A* **1999**, *103*, 4094–4100.

[38] Hillier, I. H. *J. Mol. Struct. - Theochem* **1999**, *463*, 45 – 52.

[39] Arora, P.; Slipchenko, L. V.; Webb, S. P.; DeFusco, A.; Gordon, M. S. *J. Phys. Chem. A* **2010**, *114*, 6742–6750.

[40] Li, H.; Gordon, M. S. *J. Chem. Phys* **2007**, *126*, 124112.

[41] Li, H. *J. Chem. Phys* **2009**, *131*, 184103.

[42] Gao, J. *J. Phys. Chem. B* **1997**, *101*, 657–663.

[43] Gao, J. *J. Chem. Phys.* **1998**, *109*, 2346–2354.

[44] Xie, W.; Song, L.; Truhlar, D. G.; Gao, J. *J. Phys. Chem. B* **2008**, *112*, 14124–14131.

[45] Xie, W.; Song, L.; Truhlar, D. G.; Gao, J. *J. Chem. Phys.* **2008**, *128*, 234108.

[46] Xie, W.; Orozco, M.; Truhlar, D. G.; Gao, J. *J. Chem. Theory Comput.* **2009**, *5*, 459–467.

[47] Song, L.; Han, J.; Lin, Y.-l.; Xie, W.; Gao, J. *J. Phys. Chem. A* **2009**, *113*, 11656–11664.

[48] Cembran, A.; Payaka, A.; Lin, Y.-l.; Xie, W.; Mo, Y.; Song, L.; Gao, J. *J. Chem. Theory Comput.* **2010**, *6*, 2242–2251.

[49] Gao, J.; Cembran, A.; Mo, Y. *J. Chem. Theory Comput.* **2010**, *6*, 2402–2410.

[50] Lipparini, F.; Cappelli, C.; Barone, V. *J. Chem. Theory Comput.* **2012**, *8*, 4153–4165.

[51] Lipparini, F.; Cappelli, C.; Scalmani, G.; De Mitri, N.; Barone, V. *J. Chem. Theory Comput.* **2012**, *8*, 4270–4278.

[52] Cappelli, C. *Int. J. Quantum Chem.* **2016**, *116*, 1532–1542.

[53] Boulanger, E.; Harvey, J. N. *Curr. Opin. Struc. Biol.* **2018**, *49*, 72–76.

[54] Loco, D.; Polack, É.; Caprasecca, S.; Lagardere, L.; Lipparini, F.; Piquemal, J.-P.; Mennucci, B. *J. Chem. Theory Comput.* **2016**, *12*, 3654–3661.

[55] Boulanger, E.; Thiel, W. *J. Chem. Theory Comput.* **2012**, *8*, 4527–4538.

[56] Olsen, J. M. H.; Steinmann, C.; Ruud, K.; Kongsted, J. *J. Phys. Chem. A* **2015**, *119*, 5344–5355.

[57] Nåbo, L. J.; Olsen, J. M. H.; Holmgaard List, N.; Solanko, L. M.; Wüstner, D.; Kongsted, J. *J. Chem. Phys.* **2016**, *145*, 104102.

[58] Reinholdt, P.; Kongsted, J.; Olsen, J. M. H. *J. Phys. Chem. Lett.* **2017**, *8*, 5949–5958.

[59] Nåbo, L. J.; Olsen, J. M. H.; Martínez, T. J.; Kongsted, J. *J. Chem. Theory Comput.* **2017**, *13*, 6230–6236.

[60] Giovannini, T.; Egidi, F.; Cappelli, C. *Chemical Society Reviews* **2020**, *49*, 5664–5677.

[61] Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027–2094.

[62] Giovannini, T.; Egidi, F.; Cappelli, C. *Chem. Soc. Rev.* **2020**, *49*, 5664–5677.

[63] Mennucci, B. *WIREs Comput. Mol. Sci.* **2012**, *2*, 386–404.

[64] Miertuš, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117–129.

[65] Orozco, M.; Luque, F. J. *Chem. Rev.* **2000**, *100*, 4187–4226.

[66] Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999–3094.

[67] Gordon, M. S.; Smith, Q. A.; Xu, P.; Slipchenko, L. V. *Annu. Rev. Phys. Chem.* **2013**, *64*, 553–578.

[68] Gordon, M. S.; Slipchenko, L.; Li, H.; Jensen, J. H. *Annu. Rep. Comput. Chem.* **2007**, *3*, 177–193.

[69] Sun, Q.; Chan, G. K.-L. *Acc. Chem. Res.* **2016**, *49*, 2705–2712.

[70] Knizia, G.; Chan, G. K.-L. *J. Chem. Theory. Comput.* **2013**, *9*, 1428–1432.

[71] Chulhai, D. V.; Goodpaster, J. D. *J. Chem. Theory Comput.* **2018**, *14*, 1928–1942.

[72] Chulhai, D. V.; Goodpaster, J. D. *J. Chem. Theory Comput.* **2017**, *13*, 1503–1508.

[73] Wen, X.; Graham, D. S.; Chulhai, D. V.; Goodpaster, J. D. *J. Chem. Theory Comput.* **2020**, *16*, 385–398.

[74] Ding, F.; Manby, F. R.; Miller III, T. F. *J. Chem. Theory Comput.* **2017**, *13*, 1605–1615.

[75] Goodpaster, J. D.; Barnes, T. A.; Manby, F. R.; Miller III, T. F. *J. Chem. Phys.* **2012**, *137*, 224113.

[76] Goodpaster, J. D.; Barnes, T. A.; Manby, F. R.; Miller III, T. F. *J. Chem. Phys.* **2014**, *140*, 18A507.

[77] Manby, F. R.; Stella, M.; Goodpaster, J. D.; Miller III, T. F. *J. Chem. Theory Comput.* **2012**, *8*, 2564–2568.

[78] Goodpaster, J. D.; Ananth, N.; Manby, F. R.; Miller III, T. F. *J. Chem. Phys.* **2010**, *133*, 084103.

[79] Zhang, K.; Ren, S.; Caricato, M. *J. Chem. Theory Comput.* **2020**, *16*, 4361–4372.

[80] Ramos, P.; Papadakis, M.; Pavanello, M. *J. Phys. Chem. B* **2015**, *119*, 7541–7557.

[81] Pavanello, M.; Neugebauer, J. *J. Chem. Phys.* **2011**, *135*, 234103.

[82] Kratz, E. G.; Walker, A. R.; Lagardère, L.; Lipparini, F.; Piquemal, J.-P.; Andrés Cisneros, G. *J. Comput. Chem.* **2016**, *37*, 1019–1029.

[83] Lipparini, F.; Scalmani, G.; Lagardère, L.; Stamm, B.; Cancès, E.; Maday, Y.; Piquemal, J.-P.; Frisch, M. J.; Mennucci, B. *The Journal of chemical physics* **2014**, *141*, 184108.

[84] Vennelakanti, V.; Nazemi, A.; Mehmood, R.; Steeves, A. H.; Kulik, H. J. *Current opinion in structural biology* **2022**, *72*, 9–17.

[85] Mennucci, B., Cammi, R., Eds. *Continuum Solvation Models in Chemical Physics*; Wiley, New York, 2007.

[86] Cances, E.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3032–3041.

[87] Mennucci, B.; Cancès, E.; Tomasi, J. *J. Phys. Chem. B* **1997**, *101*, 10506–10517.

[88] Mennucci, B. *J. Phys. Chem. Lett.* **2010**, *1*, 1666–1674.

[89] Barone, V.; Cossi, M.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3210–3221.

[90] Barone, V.; Cossi, M. *J. Phys. Chem A* **1998**, *102*, 1995–2001.

[91] Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Comput. Chem* **2003**, *24*, 669–681.

[92] Cossi, M.; Barone, V. *J. Chem. Phys.* **1998**, *109*, 6246–6254.

[93] Cammi, R.; Cossi, M.; Tomasi, J. *J. Chem. Phys.* **1996**, *104*, 4611–4620.

[94] Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. *J. Chem. Phys.* **2002**, *117*, 43–54.

[95] Scalmani, G.; Frisch, M. J. *J. Chem. Phys.* **2010**, *132*, 114110.

[96] Lipparini, F.; Scalmani, G.; Mennucci, B.; Cancès, E.; Caricato, M.; Frisch, M. J. *J. Chem. Phys.* **2010**, *133*, 014106.

[97] Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *Chem. Phys. Lett* **1995**, *246*, 122 – 129.

[98] Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1996**, *100*, 19824–19839.

[99] Marenich, A. V.; Olson, R. M.; Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 2011–2033.

[100] Chamberlin, A. C.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2008**, *112*, 3024–3039.

[101] Liu, J.; Kelly, C. P.; Goren, A. C.; Marenich, A. V.; Cramer, C. J.; Truhlar, D. G.; Zhan, C.-G. *J. Chem. Theory Comput.* **2010**, *6*, 1109–1117.

[102] Mennucci, B. *Phys. Chem. Chem. Phys.* **2013**, *15*, 6583–6594.

[103] Frediani, L.; Cammi, R.; Corni, S.; Tomasi, J. *The Journal of chemical physics* **2004**, *120*, 3893–3907.

[104] Lipparini, F.; Egidi, F.; Cappelli, C.; Barone, V. *J. Chem. Theory Comput.* **2013**, *9*, 1880–1884.

[105] Egidi, F.; Carnimeo, I.; Cappelli, C. *Opt. Mater. Express* **2015**, *5*, 196–209.

[106] Egidi, F.; Russo, R.; Carnimeo, I.; D'Urso, A.; Mancini, G.; Cappelli, C. *J. Phys. Chem. A* **2015**, *119*, 5396–5404.

[107] Giovannini, T.; Olszowka, M.; Cappelli, C. *J. Chem. Theory Comput.* **2016**, *12*, 5483–5492.

[108] Giovannini, T.; Olszòwka, M.; Egidi, F.; Cheeseman, J. R.; Scalmani, G.; Cappelli, C. *J. Chem. Theory Comput.* **2017**, *13*, 4421–4435.

[109] Giovannini, T.; Del Frate, G.; Lafiosca, P.; Cappelli, C. *Phys. Chem. Chem. Phys.* **2018**, *20*, 9181–9197.

[110] Giovannini, T.; Ambrosetti, M.; Cappelli, C. *Theor. Chem. Acc.* **2018**, *137*, 74.

[111] Giovannini, T.; Macchiagodena, M.; Ambrosetti, M.; Puglisi, A.; Lafiosca, P.; Lo Gerfo, G.; Egidi, F.; Cappelli, C. *Int. J. Quantum Chem.* **2018**, e25684.

[112] Loco, D.; Cupellini, L. *Int. J. Quantum Chem.* **2018**, DOI: 10.1002/qua.25726.

[113] Jurinovich, S.; Curutchet, C.; Mennucci, B. *ChemPhysChem* **2014**, *15*, 3194–3204.

[114] Lamoureux, G.; MacKerell Jr, A. D.; Roux, B. *J. Chem. Phys.* **2003**, *119*, 5185–5197.

[115] Lamoureux, G.; Roux, B. *J. Chem. Phys.* **2003**, *119*, 3025–3039.

[116] Anisimov, V. M.; Lamoureux, G.; Vorobyov, I. V.; Huang, N.; Roux, B.; MacKerell, A. D. *J. Chem. Theory Comput.* **2005**, *1*, 153–168.

[117] Lemkul, J. A.; Huang, J.; Roux, B.; MacKerell Jr, A. D. *Chem. Rev.* **2016**, *116*, 4983–5013.

[118] Geerke, D.; van Gunsteren, W. F. *Mol. Phys.* **2007**, *105*, 1861–1881.

[119] Geerke, D. P.; Thiel, S.; Thiel, W.; van Gunsteren, W. F. *J. Chem. Theory Comput.* **2007**, *3*, 1499–1509.

[120] Boulanger, E.; Thiel, W. *J. Chem. Theory Comput.* **2014**, *10*, 1795–1809.

[121] Thole, B. T. *Chem. Phys.* **1981**, *59*, 341–350.

[122] Ren, P.; Ponder, J. W. *J. Phys. Chem. B* **2003**, *107*, 5933–5947.

[123] Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. *J. Phys. Chem. B* **2010**, *114*, 2549–2564.

[124] Lipparini, F.; Cappelli, C.; Barone, V. *J. Chem. Phys.* **2013**, *138*, 234108.

[125] Carnimeo, I.; Cappelli, C.; Barone, V. *J. Comput. Chem.* **2015**, *36*, 2271–2290.

[126] Egidi, F.; Lo Gerfo, G.; Macchiagodena, M.; Cappelli, C. *Theor. Chem. Acc.* **2018**, *137*, 82.

[127] Di Remigio, R.; Giovannini, T.; Ambrosetti, M.; Cappelli, C.; Frediani, L. *J. Chem. Theory Comput.* **2019**, *15*, 4056–4068, PMID: 31244130.

[128] Giovannini, T.; Ambrosetti, M.; Cappelli, C. *J. Phys. Chem. Lett.* **2019**, *10*, 5823–5829.

[129] Giovannini, T.; Puglisi, A.; Ambrosetti, M.; Cappelli, C. *J. Chem. Theory Comput.* **2019**, *15*, 2233–2245.

[130] Giovannini, T.; Grazioli, L.; Ambrosetti, M.; Cappelli, C. *Journal of Chemical Theory and Computation* **2019**, *15*, 5495–5507.

[131] Giovannini, T.; Riso, R. R.; Ambrosetti, M.; Puglisi, A.; Cappelli, C. *J. Chem. Phys.* **2019**, *151*, 174104.

[132] Gómez, S. A.; Rojas-Valencia, N.; Gómez, S.; Egidi, F.; Cappelli, C.; Restrepo, A. *ChemBioChem* **2021**, *22*, 724–732.

[133] Marrazzini, G.; Giovannini, T.; Egidi, F.; Cappelli, C. *J. Chem. Theory Comput.* **2020**, *16*, 6993–7004.

[134] Huang, J.; Simmonett, A. C.; Pickard IV, F. C.; MacKerell Jr, A. D.; Brooks, B. R. *J. Chem. Phys.* **2017**, *147*, 161702.

[135] Caprasecca, S.; Jurinovich, S.; Viani, L.; Curutchet, C.; Mennucci, B. *J. Chem. Theory Comput.* **2014**, *10*, 1588–1598.

[136] Donati, G.; Wildman, A.; Caprasecca, S.; Lingerfelt, D. B.; Lipparini, F.; Mennucci, B.; Li, X. *J. Phys. Chem. Lett.* **2017**, *8*, 5283–5289.

[137] Jurinovich, S.; Pescitelli, G.; Di Bari, L.; Mennucci, B. *Phys. Chem. Chem. Phys.* **2014**, *16*, 16407–16418.

[138] Loco, D.; Jurinovich, S.; Cupellini, L.; Menger, M. F.; Mennucci, B. *Photochem. Photobiol. Sci.* **2018**, *17*, 552–560.

[139] Caprasecca, S.; Cupellini, L.; Jurinovich, S.; Loco, D.; Lipparini, F.; Mennucci, B. *Theor. Chem. Acc.* **2018**, *137*, 84.

[140] Ren, S.; Lipparini, F.; Mennucci, B.; Caricato, M. *Journal of Chemical Theory and Computation* **2019**, *15*, 4485–4496.

[141] Jensen, L.; van Duijnen, P. T.; Snijders, J. G. *J. Chem. Phys.* **2003**, *119*, 12998–13006.

[142] Poulsen, T. D.; Ogilby, P. R.; Mikkelsen, K. V. *The Journal of chemical physics* **2002**, *116*, 3730–3738.

[143] Kongsted, J.; Osted, A.; Mikkelsen, K. V.; Christiansen, O. *The Journal of chemical physics* **2003**, *118*, 1620–1633.

[144] Jensen, L.; Swart, M.; van Duijnen, P. T. *The Journal of chemical physics* **2005**, *122*, 034103.

[145] Jensen, L.; Van Duijnen, P. T. *International journal of quantum chemistry* **2005**, *102*, 612–619.

[146] Jensen, L.; van Duijnen, P. T. *J. Chem. Phys.* **2005**, *123*, 074307.

[147] Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J. W.; Ren, P. *J. Chem. Theory Comput.* **2013**, *9*, 4046–4063.

[148] Dziedzic, J.; Mao, Y.; Shao, Y.; Ponder, J.; Head-Gordon, T.; Head-Gordon, M.; Skylaris, C.-K. *J. Chem. Phys.* **2016**, *145*, 124106.

[149] Loco, D.; Buda, F.; Lugtenburg, J.; Mennucci, B. *J. Phys. Chem. Lett.* **2018**, *9*, 2404–2410.

[150] Mao, Y.; Shao, Y.; Dziedzic, J.; Skylaris, C.-K.; Head-Gordon, T.; Head-Gordon, M. *J. Chem. Theory Comput.* **2017**, *13*, 1963–1979.

[151] Nottoli, M.; Mennucci, B.; Lipparini, F. *Physical Chemistry Chemical Physics* **2020**, *22*, 19532–19541.

[152] Rick, S. W.; Stuart, S. J.; Berne, B. J. *J. Chem. Phys.* **1994**, *101*, 6141–6156.

[153] Rick, S. W.; Stuart, S. J.; Bader, J. S.; Berne, B. *J. Mol. Liq.* **1995**, *65*, 31–40.

[154] Rick, S. W.; Berne, B. J. *J. Am. Chem. Soc.* **1996**, *118*, 672–679.

[155] Mortier, W. J.; Van Genechten, K.; Gasteiger, J. *J. Am. Chem. Soc.* **1985**, *107*, 829–835.

[156] Rappe, A. K.; Goddard III, W. A. *J. Phys. Chem* **1991**, *95*, 3358–3363.

[157] Sanderson, R. *Science* **1951**, *114*, 670–672.

[158] Geerlings, P.; De Proft, F.; Langenaeker, W. *Chemical reviews* **2003**, *103*, 1793–1874.

[159] Moens, J.; Geerlings, P.; Roos, G. *Chemistry–A European Journal* **2007**, *13*, 8174–8184.

[160] Geerlings, P.; De Proft, F. *Physical Chemistry Chemical Physics* **2008**, *10*, 3028–3042.

[161] Geerlings, P.; Ayers, P. W.; Toro-Labbé, A.; Chattaraj, P. K.; De Proft, F. *Accounts of chemical research* **2012**, *45*, 683–695.

[162] Geerlings, P.; Fias, S.; Boisdenghien, Z.; De Proft, F. *Chemical Society Reviews* **2014**, *43*, 4989–5008.

[163] Geerlings, P.; Chamorro, E.; Chattaraj, P. K.; De Proft, F.; Gázquez, J. L.; Liu, S.; Morell, C.; Toro-Labbé, A.; Vela, A.; Ayers, P. *Theoretical Chemistry Accounts* **2020**, *139*, 1–18.

[164] Lipparini, F.; Barone, V. *J. Chem. Theory Comput.* **2011**, *7*, 3711–3724.

[165] Lennard-Jones, J. E. *Proc. Phys. Soc.* **1931**, *43*, 461.

[166] Stone, A. *The theory of intermolecular forces*; OUP Oxford, 2013.

[167] Gokcan, H.; Kratz, E. G.; Darden, T. A.; Piquemal, J.-P.; Cisneros, G. A. *J. Phys. Chem. Lett.* **2018**, *9*, 3062–3067.

[168] Cisneros, G. A.; Piquemal, J.-P.; Darden, T. A. *J. Chem. Phys.* **2005**, *123*, 044109.

[169] Piquemal, J.-P.; Cisneros, G. A.; Reinhardt, P.; Gresh, N.; Darden, T. A. *J. Chem. Phys.* **2006**, *124*, 104101.

[170] Cisneros, G. A.; Piquemal, J.-P.; Darden, T. A. *J. Chem. Phys.* **2006**, *125*, 184101.

[171] Tkatchenko, A.; Scheffler, M. *Phys. Rev. Lett.* **2009**, *102*, 073005.

[172] Tkatchenko, A.; Romaner, L.; Hofmann, O. T.; Zojer, E.; Ambrosch-Draxl, C.; Scheffler, M. *MRS bulletin* **2010**, *35*, 435–442.

[173] Tkatchenko, A.; DiStasio Jr, R. A.; Car, R.; Scheffler, M. *Phys. Rev. Lett.* **2012**, *108*, 236402.

[174] Hermann, J.; DiStasio, R. A.; Tkatchenko, A. *Chem. Rev.* **2017**, *117*, 4714–4758.

[175] Curutchet, C.; Cupellini, L.; Kongsted, J.; Corni, S.; Frediani, L.; Steindal, A. H.; Guido, C. A.; Scalmani, G.; Mennucci, B. *J. Chem. Theory Comput.* **2018**, *14*, 1671–1681.

[176] Giovannini, T.; Lafiosca, P.; Cappelli, C. *J. Chem. Theory Comput.* **2017**, *13*, 4854–4870.

[177] Giovannini, T.; Lafiosca, P.; Chandramouli, B.; Barone, V.; Cappelli, C. *J. Chem. Phys.* **2019**, *150*, 124102.

[178] McWeeny, R. *Methods of molecular quantum mechanics*; Academic press: London, 1992.

[179] Amovilli, C.; McWeeny, R. *Chem. Phys.* **1990**, *140*, 343–361.

[180] Kitaura, K.; Morokuma, K. *Int. J. Quantum Chem.* **1976**, *10*, 325–340.

[181] Morokuma, K.; Kitaura, K. *Chemical applications of atomic and molecular electrostatic potentials*; Springer, 1981; pp 215–242.

[182] Su, P.; Li, H. *J. Chem.Phys.* **2009**, *131*, 014102.

[183] Mennucci, B.; Amovilli, C.; Tomasi, J. *Chem. Phys. Lett.* **1998**, *286*, 221–225.

[184] Bennie, S. J.; Curchod, B. F.; Manby, F. R.; Glowacki, D. R. *J. Phys. Chem. Lett.* **2017**, *8*, 5559–5565.

[185] Lee, S. J.; Welborn, M.; Manby, F. R.; Miller III, T. F. *Acc. Chem. Res.* **2019**, *52*, 1359–1368.

[186] Neugebauer, J.; Louwerse, M. J.; Baerends, E. J.; Wesolowski, T. A. *J. Chem. Phys.* **2005**, *122*, 094115.

[187] Wesolowski, T. A.; Shedge, S.; Zhou, X. *Chem. Rev.* **2015**, *115*, 5891–5928.

[188] Fux, S.; Jacob, C. R.; Neugebauer, J.; Visscher, L.; Reiher, M. *J. Chem. Phys.* **2010**, *132*, 164101.

[189] Jacob, C. R.; Neugebauer, J.; Visscher, L. *J. Comput. Chem.* **2008**, *29*, 1011–1018.

[190] Jacob, C. R.; Visscher, L. *J. Chem. Phys.* **2006**, *125*, 194104.

[191] Jacob, C. R.; Neugebauer, J.; Jensen, L.; Visscher, L. *Phys. Chem. Chem. Phys.* **2006**, *8*, 2349–2359.

[192] Marrazzini, G.; Giovannini, T.; Scavino, M.; Egidi, F.; Cappelli, C.; Koch, H. *J. Chem. Theory Comput.* **2021**, *17*, 791–803.

[193] Sæther, S.; Kjærgaard, T.; Koch, H.; Høyvik, I.-M. *J. Chem. Theory Comput.* **2017**, *13*, 5282–5290.

[194] Aquilante, F.; Boman, L.; Boström, J.; Koch, H.; Lindh, R.; de Merás, A. S.; Pedersen, T. B. *Linear-Scaling Techniques in Computational Chemistry and Physics*; Springer, 2011; pp 301–343.

[195] Sánchez de Merás, A. M.; Koch, H.; Cuesta, I. G.; Boman, L. *J. Chem. Phys.* **2010**, *132*, 204105.

[196] Koch, H.; Sánchez de Merás, A.; Pedersen, T. B. *J. Chem. Phys.* **2003**, *118*, 9481–9484.

[197] Helgaker, T.; Jorgensen, P.; Olsen, J. *Molecular electronic-structure theory*; John Wiley & Sons, 2014.

[198] Van Lenthe, J.; Zwaans, R.; Van Dam, H. J.; Guest, M. *Journal of computational chemistry* **2006**, *27*, 926–932.

[199] McWeeny, R. *Reviews of Modern Physics* **1960**, *32*, 335.

[200] Pino, R.; Scuseria, G. E. *Chemical physics letters* **2002**, *360*, 117–122.

[201] Niklasson, A. M.; Tymczak, C.; Challacombe, M. *The Journal of chemical physics* **2003**, *118*, 8611–8620.

[202] Aquilante, F.; Bondo Pedersen, T.; Sánchez de Merás, A.; Koch, H. *J. Chem. Phys.* **2006**, *125*, 174101.

[203] Olsen, J. M.; Aidas, K.; Kongsted, J. *J. Chem. Theory Comput.* **2010**, *6*, 3721–3734.

[204] Olsen, J. M. H.; Kongsted, J. *Adv. Quantum Chem.* **2011**, *61*, 107–143.

[205] Giovannini, T.; Egidi, F.; Cappelli, C. *Phys. Chem. Chem. Phys.* **2020**, *22*, 22864–22879.

[206] Kerr, J. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **1875**, *50*, 337–348.

[207] Papadopoulos, M. G.; Sadlej, A. J.; Leszczynski, J., et al. *Non-linear optical properties of matter*; Springer, 2006.

[208] Franken, P.; Hill, A. E.; Peters, C. e.; Weinreich, G. *Physical Review Letters* **1961**, *7*, 118.

[209] Bishop, D. M. *Reviews of Modern Physics* **1990**, *62*, 343.

[210] Bishop, D. M. *Advances in quantum chemistry* **1994**, *25*, 1–45.

[211] Bishop, D. M. *Advances in Chemical Physics (John Wiley & Sons, Inc., 2007)* **1998**, 1–40.

[212] Nalwa, H. S. *Handbook of advanced electronic and photonic materials and devices, Ten-volume Set*; Academic Press, 2000; Vol. 1.

[213] Bredas, J. L.; Adant, C.; Tackx, P.; Persoons, A.; Pierce, B. *Chemical reviews* **1994**, *94*, 243–278.

[214] Buckingham, A.; Orr, B. *Quarterly Reviews, Chemical Society* **1967**, *21*, 195–212.

[215] Kirtman, B.; Champagne, B. *International Reviews in Physical Chemistry* **1997**, *16*, 389–420.

[216] Prasad, P. N.; Williams, D. J., et al. *Introduction to nonlinear optical effects in molecules and polymers*; Wiley New York, 1991; Vol. 1.

[217] Shelton, D. P.; Rice, J. E. *Chemical Reviews* **1994**, *94*, 3–29.

[218] Ward, J. *Reviews of Modern Physics* **1965**, *37*, 1.

[219] Wolff, J. J.; Wortmann, R. *Adv. Phys. Org. Chem.* **1999**, *32*, 121–217.

[220] Ferrighi, L.; Frediani, L.; Cappelli, C.; Sałek, P.; Ågren, H.; Helgaker, T.; Ruud, K. *Chem. Phys. Lett.* **2006**, *425*, 267–272.

[221] Frediani, L.; Ågren, H.; Ferrighi, L.; Ruud, K. *The Journal of chemical physics* **2005**, *123*, 144117.

[222] Born, M.; Oppenheimer, R. *Annalen der Physik* **1927**, *389*, 457–484.

[223] Born, M.; Huang, K. *Dynamical theory of crystal lattices*; Oxford university press, 1954.

[224] Egidi, F.; Giovannini, T.; Piccardo, M.; Bloino, J.; Cappelli, C.; Barone, V. *J. Chem. Theory Comput˙* **2014**, *10*, 2456–2464.

[225] Bishop, D. M.; Kirtman, B. *J. Chem. Phys.* **1991**, *95*, 2646–2658.

[226] Bishop, D. M.; Kirtman, B. *The Journal of chemical physics* **1992**, *97*, 5255–5256.

[227] Flytzanis, C. In Rabin, H., Tang, CL,(eds) Quantum Electronics, vol. IA. 1975.

[228] Rice, J. E.; Amos, R. D.; Colwell, S. M.; Handy, N. C.; Sanz, J. *J. Chem. Phys.* **1990**, *93*, 8828–8839.

[229] Rice, J. E.; Handy, N. C. *Int. J. Quantum Chem.* **1992**, *43*, 91–118.

[230] Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2009**, *113*, 6378–6396.

[231] Casida, M. E. In *Recent Advances in Density Functional Methods Part I*; Chong, D. P., Ed.; World Scientific, Singapore, 1995; pp 155–192.

[232] Giovannini, T.; Ambrosetti, M.; Cappelli, C. *JPCL* **2019**, *10*, 5823–5829.

[233] Gilbert, B. C.; Davies, M.; Murphy, D. M. *Electron paramagnetic resonance*; Royal Society of Chemistry, 2002; Vol. 18.

[234] Barone, V.; Cimino, P.; Pavone, M. *Continuum Solvation Models in Chemical Physics: From Theory to Applications* **2008**, 145.

[235] Malkina, O. L.; Vaara, J.; Schimmelpfennig, B.; Munzarová, M.; Malkin, V. G.; Kaupp, M. *J. Am. Chem. Soc.* **2000**, *122*, 9206–9218.

[236] Kaupp, M.; Remenyi, C.; Vaara, J.; Malkina, O. L.; Malkin, V. G. *J. Am. Chem. Soc.* **2002**, *124*, 2709–2722.

[237] Gordon, M. S.; Fedorov, D. G.; Pruitt, S. R.; Slipchenko, L. V. *Chem. Rev.* **2012**, *112*, 632–672.

[238] Collins, M. A.; Cvitkovic, M. W.; Bettens, R. P. *Acc. Chem. Res.* **2014**, *47*, 2776–2785.

[239] Collins, M. A.; Bettens, R. P. *Chem. Rev.* **2015**, *115*, 5607–5642.

[240] Grimme, S.; Diedrich, C.; Korth, M. *Angew. Chem. Int. Ed.* **2006**, *118*, 641–645.

[241] Stöhr, M.; Van Voorhis, T.; Tkatchenko, A. *Chem. Soc. Rev.* **2019**, *48*, 4118–4154.

[242] Ambrosetti, M.; Skoko, S.; Giovannini, T.; Cappelli, C. *J. Chem. Theory Comput.* **2021**, *17*, 7146–7156.

[243] Tamukong, P. K.; Khait, Y. G.; Hoffmann, M. R. *The Journal of Physical Chemistry A* **2014**, *118*, 9182–9200.

[244] Tamukong, P. K.; Khait, Y. G.; Hoffmann, M. R. *J. Phys. Chem. A* **2017**, *121*, 256–264.

[245] Wesolowski, T. A. *J. Am. Chem. Soc.* **2004**, *126*, 11444–11445.

[246] Neugebauer, J.; Louwerse, M. J.; Belanzoni, P.; Wesolowski, T. A.; Baerends, E. J. *J. Chem. Phys.* **2005**, *123*, 114101.

[247] Høyvik, I.-M. *Mol. Phys.* **2020**, *118*, 1626929.

[248] Goletto, L.; Kjønstad, E. F.; Folkestad, S. D.; Høyvik, I.-M.; Koch, H. *J. Chem. Theory Comput.* **2021**, *17*, 7416–7427.

[249] Giovannini, T.; Koch, H. *J. Chem. Theory Comput.* **2021**, *17*, 139–150.

[250] Goletto, L.; Giovannini, T.; Folkestad, S. D.; Koch, H. *Phys. Chem. Chem. Phys.* **2021**, *23*, 4413–4425.

[251] Bennie, S. J.; van der Kamp, M. W.; Pennifold, R. C.; Stella, M.; Manby, F. R.; Mulholland, A. J. *J. Chem. Theory Comput.* **2016**, *12*, 2689–2697.

[252] Egidi, F.; Giovannini, T.; Del Frate, G.; Lemler, P. M.; Vaccaro, P. H.; Cappelli, C. *Phys. Chem. Chem. Phys.* **2019**, *21*, 3644–3655.

[253] Stendardo, E.; Pedone, A.; Cimino, P.; Menziani, M. C.; Crescenzi, O.; Barone, V. *Phys. Chem. Chem. Phys.* **2010**, *12*, 11697–11709.

[254] Barone, V.; Cimino, P.; Pedone, A. *Magn. Reson. Chem.* **2010**, *48*, S11–S22.

[255] Rinkevicius, Z.; Frecus, B.; Murugan, N. A.; Vahtras, O.; Kongsted, J.; Ågren, H. *J. Chem. Theory Comput.* **2011**, *8*, 257–263.

[256] de Almeida, K. J.; Rinkevicius, Z.; Hugosson, H. W.; Ferreira, A. C.; Ågren, H. *Chem. Phys.* **2007**, *332*, 176–187.

[257] Giner, E.; Tenti, L.; Angeli, C.; Ferré, N. *J. Chem. Theory Comput.* **2017**, *13*, 475–487.

[258] Pavone, M.; Sillanpää, A.; Cimino, P.; Crescenzi, O.; Barone, V. *J. Phys. Chem. B* **2006**, *110*, 16189–16192.

[259] Berliner, L. J. *Spin labeling: theory and applications*; Academic, New York, 1976.

[260] Kocherginsky, N.; Swartz, H. M. *Nitroxide spin labels: reactions in biology and chemistry*; CRC Press, New York, 1995.

[261] Buchaklian, A. H.; Klug, C. S. *Biochemistry* **2005**, *44*, 5503–5509.

[262] Myhre, R. H.; Sánchez de Merás, A. M.; Koch, H. *J. Chem. Phys.* **2014**, *141*, 224105.

[263] Pulay, P. *J. Chem. Phys.* **1983**, *78*, 5043–5051.

[264] Saebo, S.; Pulay, P. *Annu. Rev. Phys. Chem.* **1993**, *44*, 213–236.

[265] Barone, V.; Cimino, P. *Chem. Phys. Lett.* **2008**, *454*, 139–143.

[266] Folkestad, S. D.; Kjønstad, E. F.; Myhre, R. H.; Andersen, J. H.; Balbi, A.; Coriani, S.; Giovannini, T.; Goletto, L.; Haugland, T. S.; Hutcheson, A., et al. *J. Chem. Phys.* **2020**, *152*, 184103.

[267] Giovannini, T.; Lafiosca, P.; Chandramouli, B.; Barone, V.; Cappelli, C. *The Journal of chemical physics* **2019**, *150*, 124102.

[268] Keana, J. F.; Lee, T. D.; Bernard, E. M. *J. Am. Chem. Soc.* **1976**, *98*, 3052–3053.

[269] Rockenbauer, A.; Korecz, L.; Hideg, K. *J. Chem. Soc., Perkin Trans. 2* **1993**, 2149–2156.

[270] Improta, R.; Barone, V. *Chem. Rev.* **2004**, *104*, 1231–1254.

[271] Sharma, B.; Tran, V. A.; Pongratz, T.; Galazzo, L.; Zhurko, I.; Bordignon, E.; Kast, S. M.; Neese, F.; Marx, D. *J. Chem. Theory Comput.* **2021**, *17*, 6366–6386.