JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

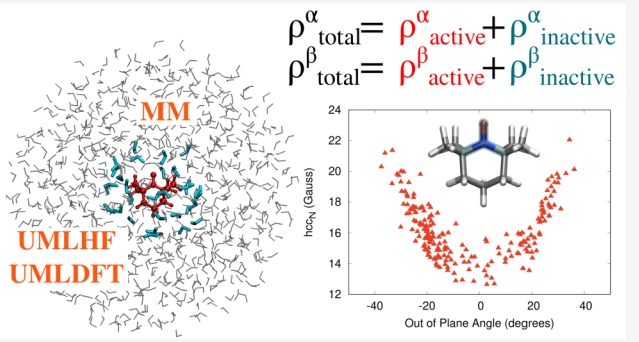
Integrated Multiscale Multilevel Approach to Open Shell Molecular Systems

Tommaso Giovannini,* Gioia Marrazzini, Marco Scavino, Henrik Koch, and Chiara Cappelli*

Cite This: *J. Chem. Theory Comput.* 2023, 19, 1446−1456 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

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 nonpolarizable 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 an aqueous solution.



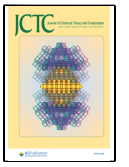
1. INTRODUCTION

The theoretical description of large molecular systems in the condensed phase at a 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.[1−3] 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,[4−8] 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 on 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.[8,9] The interaction between the QM and classical moieties is modeled in terms of classical electrostatics, and in some cases, mutual polarization effects are considered.[5] Only a few examples exist, where purely QM interactions, such as Pauli repulsion and dispersion forces, are coherently introduced in the QM/classical modeling,[10−13] even though they play a crucial role in many systems.[14,15] Finally, the quality of QM/classical methods also depends on the quality of the MM description, which is determined by the availability and reliability of parameter sets.[16]

As an alternative to QM/classical methods, quantum embedding approaches can be exploited.[17−28] There, both subsystems are described quantum-mechanically, generally at different levels of accuracy. The advantage of quantum vs classical embedding is 2-fold: (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)[29] and Multilevel DFT (MLDFT)[30] 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.[29−33] MLHF and MLDFT lie on common theoretical foundations, being based on a decomposition of the total density matrix into active and inactive contributions.[29,30] 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

Received: August 4, 2022
Published: February 13, 2023



(SCF) procedure. The inactive density matrix is kept fixed and gives rise to an effective field, which interacts with the active part (See refs 29 and 30 for more details.).

The partitioning of the system that is featured in MLHF/DFT allows substantial extension of 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.[34,35] 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.[34] As a consequence, the novel class of methods gains advantage from the physicochemical features of both approaches.

In this paper, we integrate classical MM force fields,[8,9] either within the electrostatic or polarizable embedding schemes (the latter based on the Fluctuating Charge (FQ) force field[4−6,36,37]), 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 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 α and β 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 (hccs) of selected molecular spin probes.[38−43] Since hccs are particularly sensitive to the probes' external environment,[44−46] 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 the hcc of the nitrogen atom of PROXYL and TEMPO radicals in an 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.

2. 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.

2.1. 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

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

where $\mathbf{D}^\sigma$ is α- and β-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. Eq 1 is formulated for a generic DFT functional, where $\rho^\sigma(\mathbf{r})$ is α- and β-DFT density functions, ε_x and ε_c indicate exchange and correlation energy densities per unit particle, and c_x is the percentage of HF exact exchange. 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-density matrices as $\mathbf{D} = \mathbf{D}^\alpha + \mathbf{D}^\beta$.

Similar to the closed shell case,[29,30] 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 α and β density matrices into active $\mathbf{D}_A^\sigma$ and inactive $\mathbf{D}_B^\sigma$ contributions:

$$\begin{aligned} \mathbf{D}^\alpha &= \mathbf{D}_A^\alpha + \mathbf{D}_B^\alpha \Rightarrow \rho^\alpha(\mathbf{r}) = \rho_A^\alpha(\mathbf{r}) + \rho_B^\alpha(\mathbf{r}) \\ \mathbf{D}^\beta &= \mathbf{D}_A^\beta + \mathbf{D}_B^\beta \Rightarrow \rho^\beta(\mathbf{r}) = \rho_A^\beta(\mathbf{r}) + \rho_B^\beta(\mathbf{r}) \end{aligned} \quad (2)$$

A similar partitioning applies to DFT 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 2 into eq 1, we obtain

$$E[\mathbf{D}_A^\alpha, \mathbf{D}_B^\alpha, \mathbf{D}_A^\beta, \mathbf{D}_B^\beta] = E[\mathbf{D}_A^\alpha, \mathbf{D}_B^\alpha] + E[\mathbf{D}_A^\beta, \mathbf{D}_B^\beta] \quad (3)$$

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

$$\begin{aligned} & E[\mathbf{D}_A^\sigma, \mathbf{D}_B^\sigma] = \\ & E_A^\sigma\begin{cases} \mathrm{Tr}\mathbf{h}\mathbf{D}_A^\sigma + \frac{1}{2}\mathrm{Tr}\mathbf{D}_A^\sigma\mathbf{J}(\mathbf{D}_A) \\ -\frac{1}{2}c_x\mathrm{Tr}\mathbf{D}_A^\sigma\mathbf{K}(\mathbf{D}_A^\sigma) + \int\rho_A^\sigma(\mathbf{r})\varepsilon_{xc}(\rho_A(\mathbf{r}))\mathrm{d}\mathbf{r} \end{cases} \\ & E_B^\sigma\begin{cases} + \mathrm{Tr}\mathbf{h}\mathbf{D}_B^\sigma + \frac{1}{2}\mathrm{Tr}\mathbf{D}_B^\sigma\mathbf{J}(\mathbf{D}_B) \\ -\frac{1}{2}c_x\mathrm{Tr}\mathbf{D}_B^\sigma\mathbf{K}(\mathbf{D}_B^\sigma) + \int\rho_B^\sigma(\mathbf{r})\varepsilon_{xc}(\rho_B(\mathbf{r}))\mathrm{d}\mathbf{r} \end{cases} \\ & E_{int}^\sigma\begin{cases} + \mathrm{Tr}\mathbf{D}_A^\sigma\mathbf{J}(\mathbf{D}_B) - c_x\mathrm{Tr}\mathbf{D}_A^\sigma\mathbf{K}(\mathbf{D}_B^\sigma) \\ + \int\rho_A^\sigma(\mathbf{r})\varepsilon_{xc}(\rho_B(\mathbf{r}))\mathrm{d}\mathbf{r} + \int\rho_B^\sigma(\mathbf{r})\varepsilon_{xc}(\rho_A(\mathbf{r}))\mathrm{d}\mathbf{r} \end{cases} \\ & E_{non-add}^\sigma\begin{cases} + \int\rho^\sigma(\mathbf{r})\varepsilon_{xc}(\rho(\mathbf{r}))\mathrm{d}\mathbf{r} \\ -\int\rho^\sigma(\mathbf{r})\varepsilon_{xc}(\rho_A(\mathbf{r}))\mathrm{d}\mathbf{r} - \int\rho^\sigma(\mathbf{r})\varepsilon_{xc}(\rho_B(\mathbf{r}))\mathrm{d}\mathbf{r} \end{cases} \end{aligned} \quad (4)$$

where the identity $\mathrm{Tr}\mathbf{D}_A\mathbf{G}(\mathbf{D}_B) = \mathrm{Tr}\mathbf{D}_B\mathbf{G}(\mathbf{D}_A)$ is used.[47] In eq 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 ($E_{A,B}^\sigma$). Note that within this notation, the

https://doi.org/10.1021/acs.jctc.2c00805
J. Chem. Theory Comput. 2023, 19, 1446−1456

dependency of $E[\mathbf{D}_A^\alpha, \mathbf{D}_B^\alpha]$ on β densities, through the exchange-correlation functional and the Coulomb matrix, is not explicitly stated. Also, the coupling terms are divided into a purely interaction energy term, E_{int}^σ, and a nonadditive contribution $E_{non-add}^\sigma$, which originates from the nonlinearity of ε_x and ε_c. It is worth noting that the nonadditive energy terms vanish for UMLHF. As expected, the partitioning of the different terms in eq 4 is coherent with the MLDFT formulation for closed shell systems (See ref 30.). Indeed, as we have previously reported in ref 30, both MLHF(DFT) and their unrestricted formulations differ from other quantum-embedding approaches proposed in the literature. In fact, the partitioning is performed on the density matrix $\mathbf{D}$ and not on the density function $\rho(\mathbf{r})$, as in frozen density embedding (FDE) and embedded density functional theory approaches.[24,48−53] In this way, it is not necessary to resort to an ad hoc definition of nonadditive kinetic potential terms, either approximated or exact.[20,21,25,54,55] In MLDFT, the total density matrix is approximated and then partitioned (See eq 2.), thus avoiding the DFT calculation on the whole system. Furthermore, the active density (matrix) is optimized in the field of the inactive density (matrix), which is kept frozen.

As already discussed in the Introduction, the energy of the active fragment A is optimized, while the inactive density matrix B is kept fixed to the value resulting from the partitioning in eq 2. Therefore, the UMLHF/MLDFT α and β Fock matrix ($F_{\mu\nu}^\sigma$, in the AO basis $\{\chi_\mu\}$) can easily be recovered by differentiating the energy in eq 4 with respect to the active density matrix (D_A^σ), i.e.

$$
\begin{aligned}
F_{\mu\nu}^\sigma = h_{\mu\nu}^\sigma
&+ 2e_A^\sigma \begin{cases} + J_{\mu\nu}(\mathbf{D}_A) - c_x K_{\mu\nu}(\mathbf{D}_A^\sigma) \\ + \int v_{xc}^\sigma(\rho^A(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\mathrm{d}\mathbf{r} \end{cases} \\
&+ 2e_B^\sigma \begin{cases} + J_{\mu\nu}(\mathbf{D}_B) - c_x K_{\mu\nu}(\mathbf{D}_B^\sigma) \\ + \int v_{xc}^\sigma(\rho^B(\mathbf{r}))\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\mathrm{d}\mathbf{r} \end{cases} \\
&+ 2e_{non-add}^\sigma \begin{cases} + \int [v_{xc}^\sigma(\rho(\mathbf{r})) - v_{xc}^\sigma(\rho^A(\mathbf{r})) - v_{xc}^\sigma(\rho^B(\mathbf{r}))] \\ \times \chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\mathrm{d}\mathbf{r} \end{cases}
\end{aligned}
\tag{5}
$$

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 grouped into different $2e_X$ terms, with X = A, B. $2e_{non-add}^\sigma$ is due to the nonlinearity in the DFT functional and again vanishes for UMLHF. Finally, note that $2e_B^\sigma$ accounts for the frozen fragment, of which the density (D_B, $\rho^B(\mathbf{r})$) does not change along SCF cycles. Therefore, $2e_B^\sigma$ is a constant one-electron contribution, which is computed only once, at the beginning of the SCF procedure, similar to MLHF/MLDFT.[29,30]

We further point out that UMLHF and UMLDFT equations directly follow from the partitioning of total α and β densities into active and inactive contributions. Similar to their closed shell counterparts, their accuracy depends on the approach which is exploited to carry out the decomposition in eq 2. Different choices are possible; however, in the present work, the initial set of active occupied molecular orbitals (MOs) is obtained through a partial limited Cholesky decomposition of $\mathbf{D}^\sigma$.[56,57] In ref 33 the procedure is detailed; 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 $\mathbf{D}_A^\sigma$ ($\sigma = \alpha, \beta$) densities. $\mathbf{D}_A^\sigma$ can be written in the AO basis $\{\mu, \nu\}$ as follows[58]

$$
D_{A,\mu\nu}^\sigma = \sum_{IJ} D_{\mu I}^\sigma \tilde{D}_{IJ}^{\sigma\,-1} D_{\nu J}^\sigma \tag{6}
$$

where

$$
\tilde{D}_{IJ}^\sigma = D_{IJ} \tag{7}
$$

$$
D_{A,\mu\nu}^\sigma = \sum_{I} L_{\mu I}^\sigma L_{\nu I}^\sigma \tag{8}
$$

where the diagonal elements I and J are decomposed, the $\tilde{\mathbf{D}}$ submatrix contains the selected diagonal elements, and $L_{\alpha I}$ is the Cholesky orbitals. Diagonal elements are selected so that pivots correspond to the AOs belonging to a predefined set of active atoms. As a result of the decomposition, the active Cholesky occupied MOs are obtained, and the active density matrices $\mathbf{D}_A^\sigma$ are trivially constructed (See eq 8.). The active virtual space can be defined in terms of projected atomic orbitals (PAOs),[59,60] 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 32.).

In the practical implementation, UML calculations follow this protocol:

1. Generation of the guess AO density matrices ($\mathbf{D}^\alpha$ and $\mathbf{D}^\beta$) by superposition of atomic density (SAD) matrices, obtained through a UHF calculation.
2. Construction of initial Fock matrices from the SAD density matrix and diagonalization, so to obtain initial idempotent α and β densities. At this step, a part of polarization and charge-transfer effects between active and inactive parts is included. Although such terms are introduced by a rather crude approximation of the total density matrix, the numerical results (vide infra) demonstrate that such a choice is indeed able to catch the largest part of these interactions.
3. Partitioning of the density matrices (α and β) into active A and inactive B densities ($\mathbf{D}_A^\alpha$, $\mathbf{D}_A^\beta$ and $\mathbf{D}_B^\alpha$, $\mathbf{D}_B^\beta$). 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 matrix 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 α and β density matrices of the inactive part (See eqs 4 and 5.).
5. Energy minimization (eq 4) in the MO basis of the active part only, until convergence is reached.

2.2. 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.

$$
\mathcal{E} = E_{\mathrm{UMLHF(DFT)}} + E_{\mathrm{MM}} + E_{\mathrm{UMLHF(DFT)/MM}}^{\mathrm{int}} \tag{9}
$$

where $E_{\mathrm{UMLHF(DFT)}}$ is given in eq 4, whereas E_{MM} and $E_{\mathrm{UMLHF(DFT)/MM}}^{\mathrm{int}}$ are MM and UMLHF(DFT)/MM interaction

https://doi.org/10.1021/acs.jctc.2c00805
J. Chem. Theory Comput. 2023, 19, 1446−1456

energies, respectively. Electrostatic and polarizable QM/MM embeddings 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 modeled in polarizable embedding approaches.[5,8] In particular, nonpolarizable embedding methods place fixed charges on MM atoms, which polarize the QM density. Different polarizable QM/MM approaches exist;[13,61−66] in this work, we exploit QM/Fluctuating Charges (FQ),[5,6] 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.[4,5,67−69] Therefore, for both nonpolarizable UMLHF(DFT)/MM or polarizable UMLHF(DFT)/FQ, the UMLHF(DFT)/MM interaction energy can be written as follows

$$E^{\text{int}}_{\text{UMLHF(DFT)/MM}} = \sum_i q_i V_i(\mathbf{D}^\alpha + \mathbf{D}^\beta) \tag{10}$$

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

$$\mathcal{E}[\mathbf{D}^\alpha_A, \mathbf{D}^\beta_A, \mathbf{D}^\alpha_B, \mathbf{D}^\beta_B, \mathbf{q}, \lambda] = E_{\text{UMLHF(DFT)}}[\mathbf{D}^\alpha_A, \mathbf{D}^\alpha_B] + E_{\text{UMLHF(DFT)}}[\mathbf{D}^\beta_A, \mathbf{D}^\beta_B] + \frac{1}{2}\mathbf{q}^\dagger_\lambda \mathbf{M}\mathbf{q}_\lambda + \mathbf{q}^\dagger_\lambda \mathbf{C}_Q + \mathbf{q}^\dagger_\lambda \mathbf{V}(\mathbf{D}^\alpha + \mathbf{D}^\beta) \tag{11}$$

where $E_{\text{UMLHF(DFT)}}[\mathbf{D}^\alpha_A, \mathbf{D}^\alpha_B]$ and $E_{\text{UMLHF(DFT)}}[\mathbf{D}^\beta_A, \mathbf{D}^\beta_B]$ are due to the total α- and β densities, respectively (See eq 4.). In eq 11, $\mathbf{q}_\lambda$ 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 molecule for solvated systems). The $\mathbf{M}$ matrix is the interaction kernel between the FQ charges, which also contains the Lagrangian blocks,[36] 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 11. This procedure yields the following set of linear equations

$$\mathbf{M}\mathbf{q}_\lambda = -\mathbf{C}_Q - \mathbf{V}(\mathbf{D}^\alpha + \mathbf{D}^\beta) \tag{12}$$

In parallel, UMLHF(DFT)/MM α- and β-Fock matrices are defined as follows

$$F^\sigma_{\mu\nu} = F^{\sigma,\text{UMLHF(DFT)}}_{\mu\nu} + \sum_i q_i V^\sigma_{i,\mu\nu} \tag{13}$$

where $F^{\sigma,\text{UMLHF(DFT)}}_{\mu\nu}$ is defined in eq 5. As for energy contributions, the additional QM/MM term is fixed and computed only at the first SCF cycle in case of nonpolarizable QM/MM, whereas it changes in QM/FQ, because FQ charges depend on QM densities. Therefore, the 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.

2.3. 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[70]

$$\mathbf{A}(X) = \mathrm{A}_X \mathbf{1}_3 + \mathbf{A}_{\text{dip}}(\mathrm{X}) \tag{14}$$

In eq 14, the dipolar contribution $\mathbf{A}_{\text{dip}}(\mathrm{X})$ is a zero-trace tensor and vanishes in isotropic media (e.g., solutions). The A_X denotes the so-called hyperfine coupling constant (hcc), which is connected to the spin density function (ρ_X) at nucleus X through the following equation

$$\mathrm{A}_X = \frac{4\pi}{3}\mu_B \mu_X g_e g_X \langle S_Z\rangle^{-1} \rho^{\alpha-\beta}_X \tag{15}$$

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

$$\rho^{\alpha-\beta}_X = \sum_{\mu\nu} D^{\alpha-\beta}_{\mu\nu} \langle \chi_\mu(\mathbf{r})|\delta(\mathbf{r} - \mathbf{r}_X)|\chi_\nu(\mathbf{r})\rangle \tag{16}$$

$D^{\alpha-\beta}$ is the difference between α and β density matrices. UMLHF(DFT) hccs are computed for the active atoms only, and the difference between α and β density matrices in eq 16 refers to both active and inactive spin-density matrices, which are calculated by minimizing UMLHF(DFT) energy in eqs 4-5. When the outer MM layer is considered, the active spin-densities entering eq 16 are obtained by minimizing the UMLHF(DFT)/MM energy (See eqs 11-13.).

3. COMPUTATIONAL DETAILS

In the following sections, UMLDFT is applied to the calculation of the $\mathrm{hcc_N}$ of PROXYL and TEMPO nitroxyl radicals in an aqueous solution (See Figure 1.). Such systems are characterized

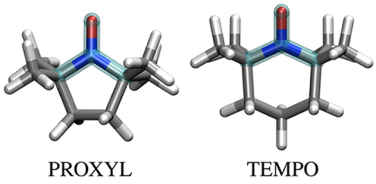


Figure 1. PROXYL (left) and TEMPO (right) molecular structures. The atoms involved in the out-of-plane angle are highlighted in cyan in both radicals.

by the presence of the N−O group, which has been amply exploited as "spin probe" for structural studies of macromolecular systems.[43] 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.[5] First, classical MD simulations of the spin labels in water are performed by using AMBER,[71] according to ref 72 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,[34] 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 described classically by means of the polarizable FQ force field, by exploiting the parameters reported in ref 72, specifically developed to correctly describe the electrostatic and polarization

https://doi.org/10.1021/acs.jctc.2c00805
J. Chem. Theory Comput. 2023, 19, 1446−1456

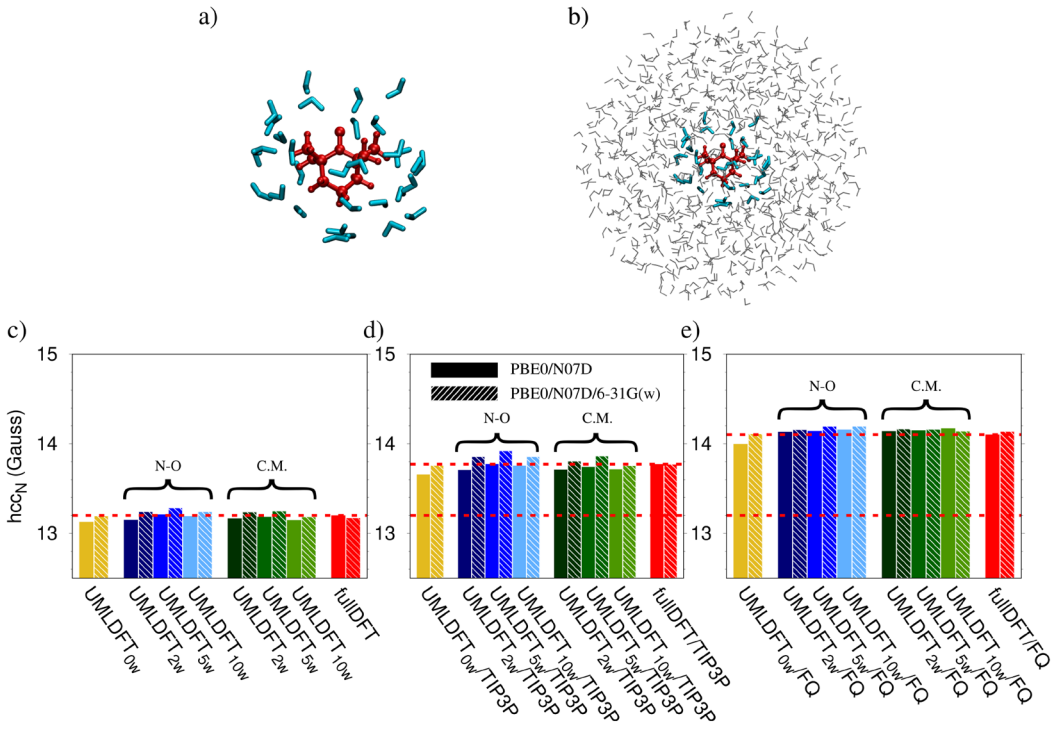


Figure 2. Molecular structure of the randomly selected TEMPO snapshot as described at the UMLDFT (a) and UMLDFT/MM (b) levels. (c-e) hcc$_N$ (Gauss) calculated at the UMLDFT$_{nw}$(/TIP3P,/FQ) and full DFT(/TIP3P, /FQ) levels. Water molecules included in the UMLDFT layer are selected with respect to 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.

interactions of radicals in an aqueous solution. To test the quality of the UMLDFT description, different active/inactive partitions are considered by including 0, 2, 5, or 10 water molecules in the active part (MLDFT$_{nw}$, $n \in [0, 2, 5, 10]$). 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[73] 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$_N$'s are calculated as an average of the 200 uncorrelated snapshots. The structures exploited in this work are available as SI. All QM calculations are performed by using a development version of the electronic structure code e^T.[74]

4. 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$_N$'s are computed. To quantify the role of long-range electrostatics interactions, hcc$_N$'s are also computed by removing the MM layer (See Figure 2a-b.). The model systems exploited in the analysis are depicted in Figure 2a-b; TEMPO is colored in red, UMLDFT water molecules are colored in cyan, and the MM layer is colored in gray. Additional calculations are performed by exploiting the N07D basis set to describe all of the systems, 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 Figure 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 Figure 2 for TEMPO (See also Table 1 for raw data.). The corresponding data for PROXYL are given in Figure S1 in the Supporting Information − SI, together with the associated raw data in Table S1.

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$_N$'s are not particularly affected by *n*w. In fact, a maximum difference of 0.1 G between

Table 1. TEMPO hcc$_N$ (Gauss) Calculated at the UMLDFT$_{nw}$(/TIP3P,/FQ) and Full DFT(/TIP3P, /FQ) Levels for a Randomly Selected Snapshot[a]

		PBE0/N07D			PBE0/N07D/6-31G(w)		
	UMLDFT$_{nw}$	Gas-phase	FQ	TIP3P	Gas-phase	FQ	TIP3P
	0w	13.1	14.0	13.7	13.2	14.1	13.8
N-O	2w	13.2	14.1	13.7	13.2	14.2	13.9
N-O	5w	13.2	14.1	13.8	13.3	14.2	13.9
N-O	10w	13.2	14.2	13.8	13.2	14.2	13.9
CM	2w	13.2	14.1	13.7	13.2	14.2	13.8
CM	5w	13.2	14.2	13.7	13.2	14.2	13.9
CM	10w	13.1	14.2	13.7	13.2	14.1	13.8
	Full DFT (3.5Å)	13.2	14.1	13.8	13.2	14.1	13.8
	Full DFT (6.5Å)	–	–	–	–	14.0	–

[a]Water molecules included in the UMLDFT layer are selected with respect to the N−O group or C.M.

https://doi.org/10.1021/acs.jctc.2c00805
J. Chem. Theory Comput. 2023, 19, 1446−1456

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 G when the water molecules are described by means of the polarizable FQ force field (See Figure 2e.). When the nonpolarizable TIP3P force field is instead used, a minor shift of the computed hcc_N with respect to the values computed for the small cluster is reported (ranging from 0.5 to 0.6 G, see Figure 2d). The different results provided by QM/TIP3P and QM/FQ can be ascribed to the more realistic description of specific solute−solvent interactions interactions which is provided by the fully polarizable QM/FQ.[5] Also, the FQ parametrization which is here exploited is tuned to describe electrostatic and polarization energies in an aqueous solution.[72] As compared to the gas-phase hcc_N, QM/TIP3P and QM/FQ results clearly show the crucial role of including long-range electrostatic and polarization effects in the description of the hcc_N of solvated radical species.

To investigate this point in more detail, TEMPO hcc_N's are also computed by increasing the radius of the fullDFT region from 3.5 to 6.5 Å (See Table 1.). A small difference (∼0.1 G) is computed with respect to fullDFT/FQ values obtained by exploiting a cutoff radius of 3.5 Å (See Figure 2.). This further demonstrates that convergence as a function of the size of the fullDFT region is reached and also highlights the better performance of the polarizable FQ with respect to the nonpolarizable TIP3P, for which a larger deviation is reported (See Table 1.). Similar findings are reported for the randomly selected PROXYL snapshot (See Table S1 in the SI.). The data depicted in Figure 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 is less than 0.1 G.

4.1. PROXYL and TEMPO in an 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's of both PROXYL and TEMPO on the whole set of 200 uncorrelated snapshots extracted from MD runs. Computed UMLDFT$_{nw}$/FQ results are reported in Table 2, together with associated standard errors (se) at 67% confidence interval. The latter are computed as

$$\mathrm{se} = \frac{\sigma}{\sqrt{N_{snap}}} \tag{17}$$

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

We first note that the inclusion of water molecules in the active region only marginally affects the computed hcc_N values, in agreement with the benchmark analysis discussed above for the random snapshot. Interestingly, this result shows that a solvent frozen density, which we remark has been polarized by the solute at the second step of our computational protocol, is able to correctly describe hcc_N values. We also investigate the dependence of hcc_N's on the out-of-plane dihedral angle involving C−C−N−O atoms, i.e. the nitrogen atom pyrimidalization (See Figure 1.), which has previously been reported to crucially affect the description of hcc_N's of the studied radical species.[72,78] Computed data at the UMLDFT$_{0w}$/FQ level are depicted in Figure 3 for PROXYL (left) and TEMPO (right). The hcc_N hugely varies as a function of the OOP angle, ranging from about 12 to 26 G. Note that such a variability is the main reason why 200 snapshots are required to converge the property (See the top panel in Figure 3.) and also demonstrates that a reliable sampling of the radical/solvent phase-space is required to reliably model this property.

Table 2. Calculated UMLDFT$_{nw}$/FQ hcc_N (Gauss) Average Values for PROXYL and TEMPO in an Aqueous Solution[a]

		PROXYL	TEMPO
	UMLDFT$_{NO}$/FQ	16.0 ± 0.1	16.5 ± 0.1
	UMLDFT$_{0w}$/FQ	15.3 ± 0.1	16.2 ± 0.1
N-O	UMLDFT$_{2w}$/FQ	15.3 ± 0.1	16.4 ± 0.1
N-O	UMLDFT$_{5w}$/FQ	15.4 ± 0.1	16.2 ± 0.1
N-O	UMLDFT$_{10w}$/FQ	15.4 ± 0.1	16.2 ± 0.1
C.M.	UMLDFT$_{2w}$/FQ	15.3 ± 0.1	16.3 ± 0.1
C.M.	UMLDFT$_{5w}$/FQ	15.3 ± 0.1	16.2 ± 0.1
C.M.	UMLDFT$_{10w}$/FQ	15.3 ± 0.1	16.2 ± 0.1
	"Best" QM/MM[75]	15.5± 0.1	16.3± 0.1
	Exp.	16.4[76]	17.3[77]

[a]Values are averaged over 200 uncorrelated snapshots extracted from MD runs. The "best" QM/MM results are reproduced from ref 75, whereas experimental data are taken from refs 76 and 77. Water molecules included in the UMLDFT layer are selected with respect to the N−O group or C.M. In UMLDFT$_{NO}$, the active space is reduced to the N−O group only.

The data in Table 2 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 Figure 2. Such results can be further investigated by plotting the difference between hcc_N values computed by exploiting the two approaches (Δhcc_N), as a function of the OOP angle (See Figure 4, where UMLDFT$_{10w}$(N−O) is considered.). For most snapshots, the two approaches give similar hcc_N values (Δhcc_N = 0 G). However, large differences, ranging from −0.7 to 1.1 G, with a standard deviation of 0.3 G, are reported for specific values of the out-of-plane angle for both PROXYL and TEMPO. Figure 4 also reports absolute values of the error between the two approaches, together with their Gaussian convolution. For TEMPO (left panel of Figure 4), the largest discrepancies are reported for the region near ±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 shows that such effects need to be accurately investigated when multilevel method-

https://doi.org/10.1021/acs.jctc.2c00805
J. Chem. Theory Comput. 2023, 19, 1446−1456

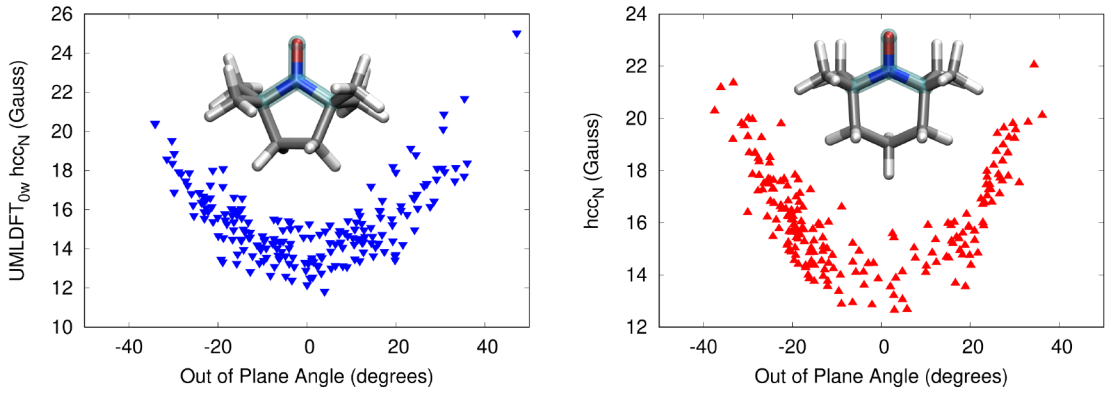


Figure 3. Computed $UMLDFT_{0w}$/FQ hcc_N values (Gauss) for PROXYL (left) and TEMPO (right) in an aqueous solution as a function of the out-of-plane angle (See the inset.).

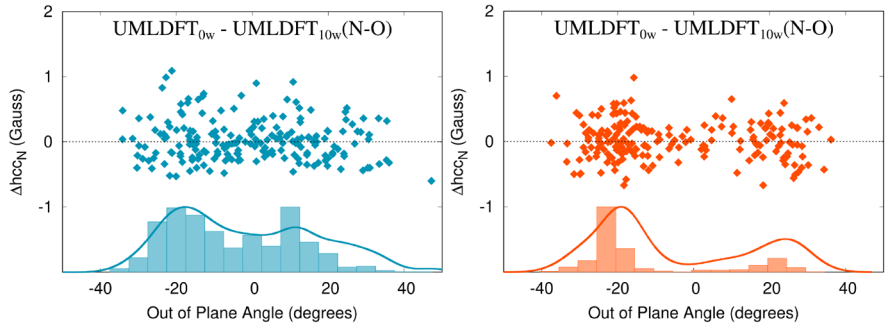


Figure 4. Computed $UMLDFT_{0w}$/FQ - $UMLDFT_{10w}$(N−O) Δhcc_N (Gauss) for PROXYL (left) and TEMPO (right) in an aqueous solution as a function of the out-of-plane angle. Histograms and Gaussian convolutions of the absolute Δhcc_N values are also given.

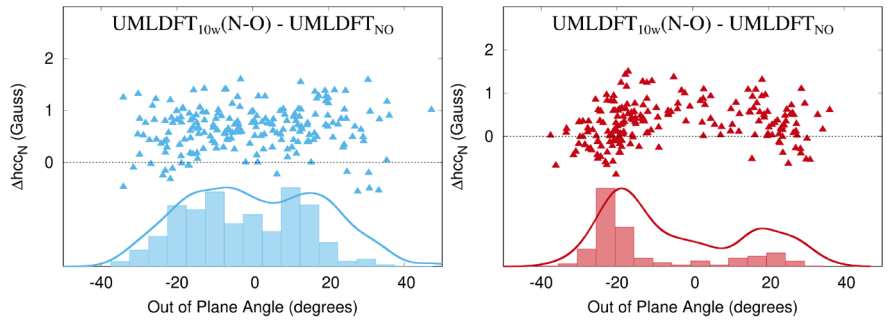


Figure 5. Computed $UMLDFT_{10w}$(N−O)/FQ - $UMLDFT_{NO}$ Δhcc_N (Gauss) as a function of the out-of-plane angle for PROXYL (left) and TEMPO (right) in an aqueous solution. Histograms and Gaussian convolutions of absolute Δhcc_N values are also given.

ologies are employed, as already reported by some of us for excitation energies.[34]

Covalently Bonded Fragments. The results reported in the previous sections assume active and inactive regions to be noncovalently 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's are reported in Table 2. $UMLDFT_{NO}$/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 a reference. However, as already reported above, in order to better analyze the computed results, in Figure 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 already commented on in Figure 4. However, in the present case, Δhcc_N values show very large variability, from −0.6 to 1.6 G, with a standard deviation of about 0.5 and 0.4 G 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 2.). It has been reported in previous studies that DFT has strong limits at

https://doi.org/10.1021/acs.jctc.2c00805
J. Chem. Theory Comput. 2023, 19, 1446−1456

reproducing hcc_N values, due to an inappropriate account of electron correlation.[72,79] However, UMLDFT results are in good agreement with most accurate QM/MM results reported by some of us[72] and which are labeled as "best QM/MM" in Table 2. 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.[10−12,72] The discrepancy between UMLDFT and experimental data is similar to "best QM/MM" results (1.1 ± 0.2 G for PROXYL and 1.1 ± 0.2 G). It is worth remarking that the "best QM/MM" requires a suitable parametrization of electrostatic (polarization) and Pauli repulsion terms for the specific case of aqueous solutions. Although such a "best QM/MM" formulation is indeed general,[10] to be efficiently applied to generic environments, parametrization is always needed. In contrast, UMLDFT resorts to a full QM treatment (at the DFT level) of both the target system and its surroundings. Therefore, its application to any kind of external environment is straightforward.

Finally note that a proper account of electron correlation, e.g. by resorting to correlated coupled cluster calculations, is expected to reduce the computed error.[72] However, computed UMLDFT/FQ differences between PROXYL and TEMPO hcc_N values (0.9 G) are perfectly in agreement with experimental findings (0.9 G), thus demonstrating that our approach can reliably describe different radical species.

5. 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. Similar 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 α and β 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.[33] The UML methods proposed here may 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 matrix is kept fixed. To further reduce such a computational cost, linear scaling implementations recently proposed for the restricted formulation can be extended to the unrestricted case too.[32] UMLHF/DFT are coupled to an outer MM layer described in terms of nonpolarizable or polarizable force fields. In this way, not only the computational cost is further reduced but also a correct physicochemical 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 the hcc_N's of PROXYL and TEMPO radicals dissolved in an aqueous solution. First, the quality of the designed computational protocol is tested on a single snapshot, demonstrating the necessity of including long-range electrostatics and polarization effects by comparing UMLDFT/TIP3P and UMLDFT/FQ results. The results also show that, at least for the selected systems and property, the rather crude description of active-inactive polarization effects which is achieved at the second step of our computational protocol is sufficient to correctly reproduce full DFT results. However, to consistently include such terms, "freeze-and-thaw" cycles, similar to what has been proposed in the context of FDE, would be necessary.[80] PROXYL and TEMPO hcc_N's are then 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's are not perfectly in agreement with our best estimates, such flexibility paves the way to the extension of a coupled cluster treatment of the active part, in order to take into account electron correlation, which may largely affect the electronic properties of open shell systems.[72,79]

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 tested here on aqueous solutions; however, the model is general enough to be applied to different solvents[16] 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.[64]

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.2c00805.

PROXYL hcc_N calculated at $UMLDFT_{nw}$(/TIP3P,/FQ) levels and associated graphical representation for PROXYL (PDF)

Molecular geometries of PROXYL and TEMPO in aqueous solution exploited in this work (ZIP)

■ AUTHOR INFORMATION

Corresponding Authors

Tommaso Giovannini − *Scuola Normale Superiore, 56126 Pisa, Italy*; orcid.org/0000-0002-5637-2853; Email: tommaso.giovannini@sns.it

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

Authors

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

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

Henrik Koch − *Scuola Normale Superiore, 56126 Pisa, Italy; Department of Chemistry, Norwegian University of Science and Technology, 7491 Trondheim, Norway*; orcid.org/0000-0002-8367-8727

https://doi.org/10.1021/acs.jctc.2c00805
J. Chem. Theory Comput. 2023, 19, 1446−1456

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jctc.2c00805

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

T.G. and C.C. acknowledge funding from the European Research Council (ERC) under the European Union Horizon 2020 research and innovation programme (grant agreement No. 818064). H.K. acknowledges funding through the Research Council of Norway through FRINATEK project 275506 and funding from the ERC under the European Union 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.

■ REFERENCES

(1) Gordon, M. S.; Fedorov, D. G.; Pruitt, S. R.; Slipchenko, L. V. Fragmentation methods: A route to accurate calculations on large systems. *Chem. Rev.* **2012**, *112*, 632−672.
(2) Collins, M. A.; Cvitkovic, M. W.; Bettens, R. P. The combined fragmentation and systematic molecular fragmentation methods. *Acc. Chem. Res.* **2014**, *47*, 2776−2785.
(3) Collins, M. A.; Bettens, R. P. Energy-based molecular fragmentation methods. *Chem. Rev.* **2015**, *115*, 5607−5642.
(4) 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.
(5) Giovannini, T.; Egidi, F.; Cappelli, C. Molecular spectroscopy of aqueous solutions: a theoretical perspective. *Chem. Soc. Rev.* **2020**, *49*, 5664−5677.
(6) 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.
(7) Gómez, S. A.; Rojas-Valencia, N.; Gómez, S.; Egidi, F.; Cappelli, C.; Restrepo, A. Binding of SARS-CoV-2 to Cell Receptors: A Tale of Molecular Evolution. *ChemBioChem.* **2021**, *22*, 724−732.
(8) Senn, H. M.; Thiel, W. QM/MM methods for biomolecular systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198−1229.
(9) 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.
(10) 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.
(11) Giovannini, T.; Ambrosetti, M.; Cappelli, C. Quantum Confinement Effects on Solvatochromic Shifts of Molecular Solutes. *J. Phys. Chem. Lett.* **2019**, *10*, 5823−5829.
(12) Marrazzini, G.; Giovannini, T.; Egidi, F.; Cappelli, C. Calculation of linear and non-linear electric response properties of systems in aqueous solution: A polarizable quantum/classical approach with quantum repulsion effects. *J. Chem. Theory Comput.* **2020**, *16*, 6993−7004.
(13) 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.
(14) Grimme, S.; Diedrich, C.; Korth, M. The Importance of Inter- and Intramolecular van der Waals Interactions in Organic Reactions: the Dimerization of Anthracene Revisited. *Angew. Chem., Int. Ed.* **2006**, *118*, 641−645.
(15) Stöhr, M.; Van Voorhis, T.; Tkatchenko, A. Theory and practice of modeling van der Waals interactions in electronic-structure calculations. *Chem. Soc. Rev.* **2019**, *48*, 4118−4154.
(16) Ambrosetti, M.; Skoko, S.; Giovannini, T.; Cappelli, C. Quantum Mechanics/Fluctuating Charge Protocol to Compute Solvatochromic Shifts. *J. Chem. Theory Comput.* **2021**, *17*, 7146−7156.
(17) 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.
(18) Sun, Q.; Chan, G. K.-L. Quantum embedding theories. *Acc. Chem. Res.* **2016**, *49*, 2705−2712.
(19) Knizia, G.; Chan, G. K.-L. Density matrix embedding: A strong-coupling quantum embedding theory. *J. Chem. Theory. Comput.* **2013**, *9*, 1428−1432.
(20) 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.
(21) 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.
(22) 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.
(23) Ramos, P.; Papadakis, M.; Pavanello, M. Performance of frozen density embedding for modeling hole transfer reactions. *J. Phys. Chem. B* **2015**, *119*, 7541−7557.
(24) Pavanello, M.; Neugebauer, J. Modelling charge transfer reactions with the frozen density embedding formalism. *J. Chem. Phys.* **2011**, *135*, 234103.
(25) 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.
(26) 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.
(27) 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.
(28) Neugebauer, J.; Louwerse, M. J.; Belanzoni, P.; Wesolowski, T. A.; Baerends, E. J. Modeling solvent effects on electron-spin-resonance hyperfine couplings by frozen-density embedding. *J. Chem. Phys.* **2005**, *123*, 114101.
(29) Sæther, S.; Kjærgaard, T.; Koch, H.; Høyvik, I.-M. Density-Based Multilevel Hartree−Fock Model. *J. Chem. Theory Comput.* **2017**, *13*, 5282−5290.
(30) Marrazzini, G.; Giovannini, T.; Scavino, M.; Egidi, F.; Cappelli, C.; Koch, H. Multilevel density functional theory. *J. Chem. Theory Comput.* **2021**, *17*, 791−803.
(31) Høyvik, I.-M. Convergence acceleration for the multilevel Hartree−Fock model. *Mol. Phys.* **2020**, *118*, 1626929.
(32) Goletto, L.; Kjønstad, E. F.; Folkestad, S. D.; Høyvik, I.-M.; Koch, H. Linear-Scaling Implementation of Multilevel Hartree−Fock Theory. *J. Chem. Theory Comput.* **2021**, *17*, 7416−7427.
(33) Giovannini, T.; Koch, H. Energy-Based Molecular Orbital Localization in a Specific Spatial Region. *J. Chem. Theory Comput.* **2021**, *17*, 139−150.
(34) Goletto, L.; Giovannini, T.; Folkestad, S. D.; Koch, H. Combining multilevel Hartree-Fock and multilevel coupled cluster approaches with molecular mechanics: a study of electronic excitations in solutions. *Phys. Chem. Chem. Phys.* **2021**, *23*, 4413−4425.
(35) Bennie, S. J.; van der Kamp, M. W.; Pennifold, R. C.; Stella, M.; Manby, F. R.; Mulholland, A. J. A projector-embedding approach for multiscale coupled-cluster calculations applied to citrate synthase. *J. Chem. Theory Comput.* **2016**, *12*, 2689−2697.
(36) 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.

(37) 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.
(38) Stendardo, E.; Pedone, A.; Cimino, P.; Menziani, M. C.; Crescenzi, O.; Barone, V. Extension of the AMBER force-field for the study of large nitroxides in condensed phases: an ab initio parameterization. *Phys. Chem. Chem. Phys.* **2010**, *12*, 11697−11709.
(39) Barone, V.; Cimino, P.; Pedone, A. An integrated computational protocol for the accurate prediction of EPR and PNMR parameters of aminoxyl radicals in solution. *Magn. Reson. Chem.* **2010**, *48*, S11−S22.
(40) Rinkevicius, Z.; Frecus, B.; Murugan, N. A.; Vahtras, O.; Kongsted, J.; Ågren, H. Encapsulation influence on EPR parameters of spin-labels: 2, 2, 6, 6-tetramethyl-4-methoxypiperidine-1-oxyl in cucurbit [8] uril. *J. Chem. Theory Comput.* **2012**, *8*, 257−263.
(41) de Almeida, K. J.; Rinkevicius, Z.; Hugosson, H. W.; Ferreira, A. C.; Ågren, H. Modeling of EPR parameters of copper (II) aqua complexes. *Chem. Phys.* **2007**, *332*, 176−187.
(42) Giner, E.; Tenti, L.; Angeli, C.; Ferré, N. Computation of the Isotropic Hyperfine Coupling Constant: Efficiency and Insights from a New Approach Based on Wave Function Theory. *J. Chem. Theory Comput.* **2017**, *13*, 475−487.
(43) Pavone, M.; Sillanpää, A.; Cimino, P.; Crescenzi, O.; Barone, V. Evidence of variable H-bond network for nitroxide radicals in protic solvents. *J. Phys. Chem. B* **2006**, *110*, 16189−16192.
(44) Berliner, L. J. *Spin labeling: theory and applications*; Academic: New York, 1976; DOI: 10.1016/C2013-0-07177-0.
(45) Kocherginsky, N.; Swartz, H. M. *Nitroxide spin labels: reactions in biology and chemistry*; CRC Press: New York, 1995.
(46) Buchaklian, A. H.; Klug, C. S. Characterization of the Walker A motif of MsbA using site-directed spin labeling electron paramagnetic resonance spectroscopy. *Biochemistry* **2005**, *44*, 5503−5509.
(47) McWeeny, R. *Methods of molecular quantum mechanics*; Academic Press: London, 1992.
(48) Wesolowski, T. A.; Shedge, S.; Zhou, X. Frozen-density embedding strategy for multilevel simulations of electronic structure. *Chem. Rev.* **2015**, *115*, 5891−5928.
(49) 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.
(50) Wesolowski, T. A.; Warshel, A. Frozen density functional approach for ab initio calculations of solvated molecules. *J. Phys. Chem.* **1993**, *97*, 8050−8053.
(51) 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.
(52) Goodpaster, J. D.; Barnes, T. A.; Miller, T. F., III Embedded density functional theory for covalently bonded and strongly interacting subsystems. *J. Chem. Phys.* **2011**, *134*, 164108.
(53) Chulhai, D. V.; Jensen, L. Frozen Density Embedding with External Orthogonality in Delocalized Covalent Systems. *J. Chem. Theory Comput.* **2015**, *11*, 3080−3088.
(54) 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.
(55) Lee, S. J.; Welborn, M.; Manby, F. R.; Miller, T. F., III Projection-based wavefunction-in-DFT embedding. *Acc. Chem. Res.* **2019**, *52*, 1359−1368.
(56) Aquilante, F.; Bondo Pedersen, T.; Sánchez de Merás, A.; Koch, H. Fast noniterative orbital localization for large molecules. *J. Chem. Phys.* **2006**, *125*, 174101.
(57) Sánchez de Merás, A. M.; Koch, H.; Cuesta, I. G.; Boman, L. Cholesky decomposition-based definition of atomic subsystems in electronic structure calculations. *J. Chem. Phys.* **2010**, *132*, 204105.
(58) Myhre, R. H.; Sánchez de Merás, A. M.; Koch, H. Multi-level coupled cluster theory. *J. Chem. Phys.* **2014**, *141*, 224105.
(59) Pulay, P. Second and third derivatives of variational energy expressions: Application to multiconfigurational self-consistent field wave functions. *J. Chem. Phys.* **1983**, *78*, 5043−5051.
(60) Saebo, S.; Pulay, P. Local treatment of electron correlation. *Annu. Rev. Phys. Chem.* **1993**, *44*, 213−236.
(61) Olsen, J. M.; Aidas, K.; Kongsted, J. Excited states in solution through polarizable embedding. *J. Chem. Theory Comput.* **2010**, *6*, 3721−3734.
(62) Olsen, J. M. H.; Kongsted, J. Molecular properties through polarizable embedding. *Adv. Quantum Chem.* **2011**, *61*, 107−143.
(63) 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.
(64) 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.
(65) 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.
(66) 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.
(67) 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.
(68) Rick, S. W.; Stuart, S. J.; Bader, J. S.; Berne, B. Fluctuating charge force fields for aqueous solutions. *J. Mol. Liq.* **1995**, *65*, 31−40.
(69) Rick, S. W.; Berne, B. J. Dynamical Fluctuating Charge Force Fields: The Aqueous Solvation of Amides. *J. Am. Chem. Soc.* **1996**, *118*, 672−679.
(70) Eriksson, L. A.; Malkina, O. L.; Malkin, V. G.; Salahub, D. R. The hyperfine structures of small radicals from density functional calculations. *J. Chem. Phys.* **1994**, *100*, 5066−5075.
(71) Sorin, E. J.; Pande, V. S. Exploring the Helix-Coil Transition via All-Atom Equilibrium Ensemble Simulations. *Biophys. J.* **2005**, *88*, 2472−2493.
(72) 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.
(73) Barone, V.; Cimino, P. Accurate and feasible computations of structural and magnetic properties of large free radicals: The PBE0/N07D model. *Chem. Phys. Lett.* **2008**, *454*, 139−143.
(74) 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. e T 1.0: An open source electronic structure program with emphasis on coupled cluster and multilevel methods. *J. Chem. Phys.* **2020**, *152*, 184103.
(75) 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.
(76) Keana, J. F.; Lee, T. D.; Bernard, E. M. Side-chain substituted 2, 2, 5, 5-tetramethylpyrrolidine-N-oxyl (proxyl) nitroxides. A new series of lipid spin labels showing improved properties for the study of biological membranes. *J. Am. Chem. Soc.* **1976**, *98*, 3052−3053.
(77) Rockenbauer, A.; Korecz, L.; Hideg, K. Ring pseudorotation in pyrrolidine N-oxyl radicals: an analysis of 13C-hyperfine structure of EPR spectra. *J. Chem. Soc., Perkin Trans.* **1993**, *2*, 2149−2156.
(78) Improta, R.; Barone, V. Interplay of electronic, environmental, and vibrational effects in determining the hyperfine coupling constants of organic free radicals. *Chem. Rev.* **2004**, *104*, 1231−1254.
(79) Sharma, B.; Tran, V. A.; Pongratz, T.; Galazzo, L.; Zhurko, I.; Bordignon, E.; Kast, S. M.; Neese, F.; Marx, D. A Joint Venture of Ab Initio Molecular Dynamics, Coupled Cluster Electronic Structure Methods, and Liquid-State Theory to Compute Accurate Isotropic

Hyperfine Constants of Nitroxide Probes in Water. *J. Chem. Theory Comput.* **2021**, *17*, 6366−6386.
(80) Wesolowski, T. A.; Weber, J. Kohn-Sham equations with constrained electron density: an iterative evaluation of the ground-state electron density of interacting molecules. *Chem. Phys. Lett.* **1996**, *248*, 71−76.

https://doi.org/10.1021/acs.jctc.2c00805
J. Chem. Theory Comput. 2023, 19, 1446−1456