Budapest University of Technology and Economics
Faculty of Chemical Technology and Biotechnology
George A. Olah Doctoral School of Chemistry and Chemical Technology

Development of Multilevel Methods Using Frozen Localized Orbitals

Ph.D. dissertation

Written by: Bence Hégely
Advisor: Prof. Mihály Kállay

Department of Physical Chemistry and Materials Science
MTA-BME Lendület Quantum Chemistry Research Group

2019

Contents

Abbreviations

AO	atomic orbital
B3LYP	Becke's three-parameter hybrid functional with the LYP functional
BP	Boughton–Pulay algorithm
CC	coupled-cluster approach
CCSD	CC method with single and double excitations
CCSD(T)	CCSD with perturbative triples excitations
CIM	clusters-in-molecules
DB	dual basis
DFT	density functional theory
DMB	dual mixed basis
ED	extended domain
EE	electronic embedding
EFP	effective fragment potential
EMFT	embedded mean-field theory
FCI	full configuration interaction
FDE	frozen density embedding
GHO	generalized hybrid orbital
HF	Hartree–Fock
HLSCF	Huzinaga–equation-based LSCF
HMLC	Huzinaga multilevel local correlation
KS	Kohn–Sham method
LA	link atom
LB	large basis
LCCSD(T)	local CCSD(T)
LDA	local (spin) density approximation
LIS	local interacting subspace
LMO	localized MO
LMP2	local MP2
LNO	local natural orbital
LNO-CCSD(T)	LNO-based LCCSD(T)
LRC	long–range correlation
LSCF	local SCF

LYP	correlation functional of Lee, Yang, and Parr
MB	mixed-basis
ME	mechanical embedding
MLC	multilevel local correlation
MM	molecular mechanics
MMH	MM host atom
MO	molecular orbital
MP2	second-order Møller–Plesset method
OEP	optimized effective potential
OS	opposite-spin
PAO	projected atomic orbitals
PBE	Perdew–Burke–Ernzerhof functional
PD	primary domain
PE	polarization embedding
PNO	pair natural orbital
QM	quantum mechanics
QM/MM	mixed quantum mechanics/molecular mechanics
QMH	QM host atom
SCF	self-consistent field
sDFT	subsystem DFT
SLMO	stricly localized MO
vdW	van der Waals
WFT	wave function theory

Symbols

Notations

$\|r\rangle$	column vector r ("ket")
$\langle r\|$	row vector r ("bra")
$(\|r\rangle)^*$	conjugate transpose of r-ket
$(\|r\rangle)^{\mathrm{T}}$	transpose of r-ket
$i_f, j_f, k_f, \ldots$	variables of fragment $I, J, K, \ldots$
i, j, k	indices of the occupied MOs
a, b, c	indices of the virtual MOs
$\mu, \nu, \sigma, \lambda$	indices of the AOs
α, β	indices of the nuclei
$\|\psi\rangle$	notation of MOs
$\|\chi\rangle$	notation of AOs
Z	atomic number
$\boldsymbol{R}$	position vector of the nucleus
$\boldsymbol{r}$	position vector of electron
$\rho(\boldsymbol{r})$	electron density
q	partial charges
T_s	kinetic energy of the KS reference system
V_{nuc}	interaction energy of the electron and the nuclei
J	Coulomb energy
E_{xc}	exchange-correlation energy
V_{nn}	repulsion energy of the nuclei
V_{ee}	repulsion energy of the electrons
T	kinetic energy
T_s^{nad}	nonadditive kinetic energy
f	approximate exchange-correlation functional
$\langle\chi_\mu\chi_\sigma\|\chi_\nu\chi_\lambda\rangle$	two-electron AO integrals (in physicists' notation)
$\mathcal{P}$	primary domain
$\mathcal{E}$	extended domain
P	local interacting subspace

Matrices and operators

$\hat{\nabla} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$	field gradient
$\hat{F}$	Fock/KS operator
$\hat{V}_{\text{eff}}$	effective potential operator
$\hat{V}_{\text{nuc}}$	external potential operator
$\hat{J}$	Coulomb operator
$\hat{V}_{\text{xc}}$	exchange-correlation operator
$\hat{F}^{\text{embed}}$	embedding potential operator
$\hat{V}_{\text{xc}}^{\text{nad}}$	nonadditive exchange-correlation potential operator
$\hat{V}_{\text{kin}}^{\text{nad}}$	nonadditive kinetic potential operator
$\hat{P}$	projection operator
$\mathbf{C}$	MO coefficient matrix
$\mathbf{L}$	LMO coefficient matrix
$\mathbf{W}$	unitary rotation matrix
$\mathbf{D}$	density matrix
$\mathbf{F}$	Fock/KS matrix
$\mathbf{S}$	AO overlap matrix
$\mathbf{E}$	diagonal matrix of eigenvalues
$\mathbf{h}$	core-Hamiltonian
$\mathbf{G}$	electron repulsion matrix
$\mathbf{J}$	Coulomb matrix
$\mathbf{K}$	exact-exchange matrix
$\mathbf{V}_{\text{xc}}$	exchange-correlation potential matrix
a_{xc}	exact-exchange mixing parameter
$\mathbf{P}$	matrix of the projecton operator

Introduction

Computational chemistry grants theoretical explanations for experimental studies to facilitate the understanding of nature. In principle, theoretical models are not only explanatory tools but also able to predict phenomena, however, the power of these forecasts are heavily influenced by the basics of the applied models. Quantum chemistry, a discipline of computational chemistry, offers the most accurate predictions because it relies on quantum mechanics (QM) from first principles.

Despite the exactness of QM, its application is limited because of the complicated form of the wave function. Wave function theory (WFT), a specific field of quantum chemistry, introduces a series of known approximations to the wave function from the beginning, or as more commonly said "*ab initio*", and its high-level variants permit simulations for tens of atoms within 1 kcal/mol accuracy [1,2], the typical accuracy of chemical experiments (shortly referred to chemical accuracy). The predicting power of *ab initio* approaches is marvelous, however, their range of application is quite sobering for most of the chemists who are mostly interested in condensed-phase reactions and properties. The fact that the simulation time of these approaches scales as a high-power with the system size limits the use of accurate *ab initio* techniques even with growing computational capacities.

Numerous efforts have been made over the years to expand the application range by sacrificing the accuracy to some extent in exchange for faster simulations. The Kohn–Sham (KS) model of density functional theory (DFT), where empirical parameters are introduced into the independent particle model, allowed the simulated systems to be up to 1000 atoms large, while the combination of quantum chemistry with molecular mechanics (MM) enabled the simulation of nanoparticles. Both approaches meant groundbreaking developments which pushed the frontiers of quantum chemistry further, however, the price of the trade off, less accuracy for more efficiency, is often in question. Nevertheless, the significance of these works are recognized not only amongst computational chemists but also in the wider scientific community in the forms of Nobel-prizes in 1998 [3] and 2013 [4].

Despite the recent progress of the traditional and composite methods, the need of developing better approaches is still a must because the progressing fields, for instance, computer-aided materials design, command more accurate and efficient models. Therefore, our aim in the last years was to contribute to the field of quantum chemistry by improving existing techniques and developing new ones. This dissertation was written in order to summarize our efforts: after a short overview of the field, our most important developments are discussed in separate chapters.

Chapter 1

Theoretical background

In this chapter, methods which are capable of treating large molecular systems are briefly overviewed. The chapter is structured into three main sections. First, general considerations are discussed regarding how an the extended system should be treated, then the so-called focused models will be thoroughly discussed, which will start with a short outlook to the more simple methods. The remaining parts are about the top techniques of the focused models: the second section is devoted to the quantum mechanics/molecular mechanics (QM/MM) methodology, while the recently emerging quantum embedding and multilevel techniques are presented in the third and final part of the chapter.

The theoretical introduction is aimed to be comprehensive, but those research areas are emphasized which are closely related to the topics of the thesis. The considered methods are classified in the Fig. 1.1. Elementary knowledge of quantum chemistry is assumed throughout the first chapter, while other disciplines of computational chemistry, for example, the basics of molecular mechanics, are presented from first principles.

1.1 Strategies for the accurate description of large systems

The computational time (t) of the simulations can roughly be estimated by the following equation, which is specific to every theoretical method in the asymptotic limit of large molecules:

$$t(N) \approx aN^b, \tag{1.1.1}$$

where N is the number associated with size of the simulated system, a and b are the method-specific prefactor and scaling parameter, respectively. The latter characterizes the scaling behavior (N^b) of a theoretical approach, that is, the increment in computational demand if a larger system is simulated, while the prefactor is a constant which multiplies the scaling behavior. The accurate description of large systems is difficult for traditional

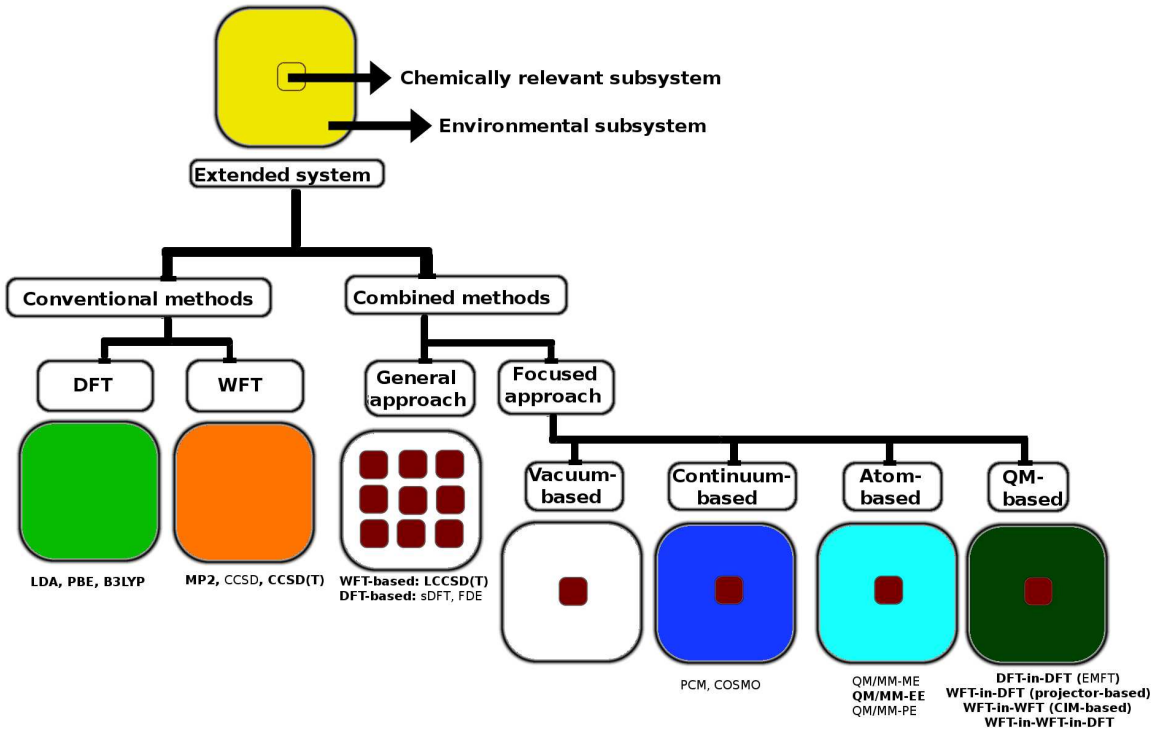

FIGURE 1.1: Illustration of the methods that are considered in the thesis. The different techniques are denoted with different, colors whereas the red stand for the (relatively) high-level methods. The used and mentioned schemes are set as bold and normal, respectively.

quantum chemistry methods because computational expenses steeply grow with the size of the studied systems [5]. The computation time of the Hartree–Fock (HF) and KS DFT calculations using a hybrid functional scales formally as the fourth-power of the systems size, and the scaling of KS calculations with a "pure", i.e., non-hybrid functional is still N^3. The more advanced *ab initio* correlation methods, such as the second-order Møller–Plesset (MP2) approach [6] or the coupled-cluster (CC) models [7], e.g., CC with single and double excitations (CCSD) [8] and CCSD with perturbative triples [CCSD(T)] [9], exhibit a scaling of at least N^5. Consequently, numerous approaches have been developed over the years to break down the expenses of quantum chemical calculations. In the following, two directions are discussed: the first one treats the whole system on the same footing, while the other one takes advantage of the fact that only a small fraction of the system is chemically relevant.

1.1.1 General approaches for extended systems

One way to decrease the computational cost of the traditional techniques is to split up the large system into smaller subsystems or fragments which are more or less

equivalent in size. The conventional methods are only utilized in the fragments, while the interactions among the fragments are described in an approximate manner or completely omitted (see Fig. 1.1 for illustration). In other words, the energy of the whole system, E_{tot}, is decomposed into fragment energies and interaction energies of the fragments:

$$E_{\text{tot}}(i_f, j_f, k_f, ...) = \sum_I E_I(i_f) + E^{\text{int}}(i_f, j_f, k_f, ...), \tag{1.1.2}$$

where the energy of fragment I, E_I, depends on the variables of the fragment (i_f) solely, while the interaction energy of the fragments is denoted as E^{int}.

The motivation behind this idea is that some computational steps which scale as a high-power of N or have large prefactors are only taken separately in the fragments, which can be considered as a space with reduced dimensions. As these costly procedures are constrained, the computation time now scales with the number of fragments, thus the N^b is reduced to linear (N^1).

1.1.2 Techniques exploiting local changes in the electronic structure

Many molecular processes exhibit local electronic structure change hence it is a logical to partition the system into two (or more) subsystems with different sizes: the active site (A) and its environment (E). This separation can also be considered as the specific and simplest partitioning scheme of the previous general approach. The concept is illustrated in Fig. 1.1, while Fig. 1.2 depicts an enzyme which is a typical system with local chemical changes. In this approach, E_{tot} can be separated into three terms:

$$E_{\text{tot}}(a_f, e_f) = E^{\text{E}}(e_f) + E^{\text{A}}(a_f) + E^{\text{int}}(a_f, e_f), \tag{1.1.3}$$

where $E^{\text{E}}(e_f)$, $E^{\text{A}}(a_f)$, and $E^{\text{int}}(a_f, e_f)$ represent the energy of the environment, the active site, and the interaction energy of the subsystems, respectively, while a_f (e_f) denotes the degree of freedom of the active site (environment). Notice that Eq. 1.1.3 is the simplest case of Eq. 1.1.2. Due to the large difference in their size and chemical relevance, it is appealing to apply different methods on the subsystems: the active site is the most relevant part hence it should be handled by highly accurate but costly techniques, while the environment can be described by a more approximate and computationally cheaper method. Theoretical frameworks exploiting this partitioning strategy are called focused models.

The focused models can be constructed in two major ways. The first one is called the additive scheme, where the system energy is decomposed into three parts:

$$E_{\text{tot}}(a_f, e_f) = E_{\text{low}}(e_f) + E_{\text{high}}(a_f) + E^{\text{int}}_{\text{coup}}(a_f, e_f), \tag{1.1.4}$$

where E_{high} represent the energy of the active site obtained with a high-level method, while E_{low} is the energy of the environment computed with a low-level technique. Note that a coupling term, $E^{\text{int}}_{\text{coup}}$, appears explicitly in Eq. 1.1.4, and it is responsible for the interaction energy of the subsystems. The other way of utilizing the kernel of the focused models is the subtraction scheme where the interaction energy of the subsystems is treated implicitly:

$$E_{\text{tot}}(a_f, e_f) = E_{\text{low}}(a_f, e_f) - E_{\text{low}}(a_f) + E_{\text{high}}(a_f), \tag{1.1.5}$$

where the energy of the total system is calculated a bit differently: first, the energy of the full system is computed with a low-level method, $E_{\text{low}}(a_f, e_f)$, and note that the interaction energy is already obtained. Then, the energy of the active site with same level of theory, $E_{\text{low}}(a_f)$, is subtracted, and finally the energy term of the same region, $E_{\text{high}}(a_f)$, is added which is calculated with a high-level technique.

The treatment of the interaction energy sharply distinguishes the two composite schemes. The additive method is defined with an explicit coupling expression between the high- and low-level methods, which can be a complicated to derive. On the other hand, it is more feasible theoretically if the subsystems are mutually polarized due to direct coupling of the different methods. In contrast, the subtractive scheme is more simple and easier to implement despite the fact that the interaction of subsystems is described only with the low-level method. A good example for the simplicity of the latter scheme is the n-layered integrated molecular orbital and molecular mechanics (ONIOM) approach [10], where an arbitrary number of layers can be defined because the last term of Eq. 1.1.5 can be further decomposed, similarly as the total energy.

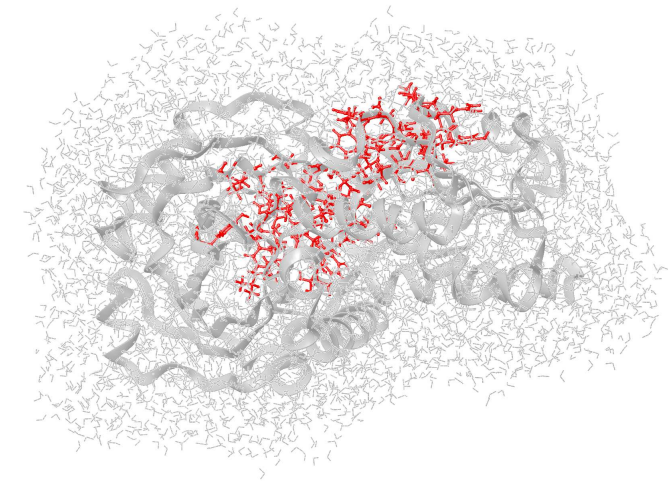

FIGURE 1.2: Partitioning strategy of multilevel methods: the active center of an enzyme (atoms highlighted with red) is treated with accurate and costly methods, while approximate and fast techniques are applied to the environment (grey atoms).

1.1.3 Hierarchy of focused models

In Sect. 1.1.2, two major methods were distinguished based on the handling of the interaction energy of the subsystems, however, another classification is also used regarding the coupling of subsystems. Some focused models, especially those with the subtractive scheme, are called multilevel approach because different levels of theory are used on different regions of the system, but there is no coupling among the methods. In the case of coupled methods, e.g., the electron density of the active subsystem is polarized by its environment, the approach is considered as an embedding method. The focused models can be further grouped, regardless of their multilevel or embedding character, based on the approximation level of the environment (E^{E}) and the interaction of subsystems (E^{int}):

0. **Vacuum-based strategy:** $E^{\mathrm{E}} = 0, E^{\mathrm{int}} = 0$
1. **Continuum-based strategy:** $E^{\mathrm{E}} = 0, E^{\mathrm{int}} \neq 0$
2. **Atom-based strategy:** $E^{\mathrm{E}} = E^{\mathrm{E}}_{\mathrm{MM}}, E^{\mathrm{int}} \neq 0$
3. **Quantum-based strategy** $E^{\mathrm{E}} = E^{\mathrm{E}}_{\mathrm{QM}}, E^{\mathrm{int}} \neq 0$

The above list is sorted with regard to the approximation level of the environment [which can be completely omitted, molecular mechanics-based ($E^{\mathrm{E}}_{\mathrm{MM}}$) (see later), or QM-based ($E^{\mathrm{E}}_{\mathrm{QM}}$)] therefore the computational cost also grows as the approaches are more elaborated. In general, the accuracy also increases with the level of theory, however, this is not guaranteed because of the cancellation of errors. Nevertheless, one can at least expect that the accuracy improves as the active subsystem enlarges.

In the followings, two of the simplest techniques will be presented shortly. The continuum-based strategy is not tightly connected to the research topics of the dissertation, however, the approach is also part of the hierarchy of the focused models thus a short discussion is appropriate for a complete overview.

1.1.3.1 Embedding into vacuum

The vacuum-based strategy (vacuum-embedding) can be regarded as a zeroth order approach because the environment and the interaction of the subsystems are completely neglected, which means that both E^{E} and E^{int} equal zero in Eq. 1.1.3. This technique was the only choice in the early eras of quantum chemistry as neither the computational capacities nor embedding techniques were available. Vacuum embedding might be looked as an outdated approach due to the more and more sophisticated condensed-phase approximations, however, decoupling the region of interest from its surrounding can be a useful tool for understanding environmental effects [11]. If the subsystems are not bonded, such calculations are trivially carried out, otherwise capping atoms, typically hydrogen atoms, are used.

1.1.3.2 Embedding into continuum

In the continuum strategy, the particle–based description of the whole system is bypassed, and the focus is shifted to the subsystem interactions [12,13]. The Hamiltonian of the whole system, $\hat{H}^{\mathrm{AE}}$, can be decomposed as

$$\hat{H}^{\mathrm{AE}}(a_f, e_f) = \hat{H}^{\mathrm{A}}(a_f) + \hat{H}^{\mathrm{E}}(e_f) + \hat{H}^{\mathrm{int}}(a_f, e_f), \tag{1.1.6}$$

where $\hat{H}^{\mathrm{A}}$, $\hat{H}^{\mathrm{E}}$, and $\hat{H}^{\mathrm{int}}$ denote the operator of the active site (A), the environment (E), and the interaction operator, respectively. In the continuum models, the Hamiltonian of the environment, $\hat{H}^{\mathrm{E}}$, is neglected, however, variables describing this subsystem (e_f) are still present in the interaction part. One can introduce the $Q(\boldsymbol{r})$ response function to implicitly include the e_f variables through the $\boldsymbol{r}$ position vector, consequently the Hamiltonian of the full system is now approximated with an effective, $\hat{H}^{\mathrm{AE}}_{\mathrm{eff}}$ operator:

$$\hat{H}^{\mathrm{AE}}_{\mathrm{eff}}(a_f, Q(\boldsymbol{r})) = \hat{H}^{\mathrm{A}}(a_f) + \hat{V}^{\mathrm{int}}(a_f, Q(\boldsymbol{r})), \tag{1.1.7}$$

where $\hat{V}^{\mathrm{int}}$ is a potential operator describing the interaction of subsystem A and E. The approach can be derived by placing the molecule of the active site into a cavity of a dielectric medium filled with vacuum. With these conditions, the electron density of the molecule, $\rho(\boldsymbol{r})$, generates a $V_M(\boldsymbol{r})$ electrostatic potential which polarizes the wall of the cavity. In return, a $V_R(\boldsymbol{r})$ reaction potential is raised by the polarized cavity wall, which effects back to the molecule. Obviously, the electrostatic potential in any point of the space is the sum of the two distinctive, polarizable entities $V(\boldsymbol{r}) = V_M(\boldsymbol{r}) + V_R(\boldsymbol{r})$, which leads to a classical electrostatic (Poission) problem because the electrostatic potential is determined by the mutually polarized molecule and cavity wall. The solution of Poisson's equation is required to obtain the $V(\boldsymbol{r})$ potential:

$$-\hat{\nabla}\left[\varepsilon(\boldsymbol{r})\hat{\nabla}V(\boldsymbol{r})\right] = 4\pi\rho(\boldsymbol{r}), \tag{1.1.8}$$

where $\hat{\nabla}$ stands for the gradient operator. Eq. 1.1.8 can be solved iteratively with the knowledge of the electron density of a molecule, $\rho(\boldsymbol{r})$, and the dielectric constant of a medium, ε. Multiple kernels have been developed over the years for solving the above equation, including the dielectric polarizable continuum model (DPCM), the conductor-like screening model (COSMO), and more recently, the integral equation formalism-based PCM (IEFPCM) (see Ref. 12 and 13 and references therein).

The continuum strategy was originally designed for modeling solvation, however, the exclusive use of the continuum models can lead to inaccurate approximations. The explicit treatment of the first solvation shell can improve the model [14], but this comes with the greater cost. Nevertheless, the continuum models are efficient and often accurate, moreover, they are extended to model various properties which allows the simulation of absorption and emission spectra [15], and non-isotropic environments as well [16].

1.2 Embedding into molecular mechanics

The atom-based strategy neglects none of the interactions of the system, which means that both E^E and E^{int} in Eq. 1.1.3 are non-zero. This can be achieved by utilizing the hybrid QM/MM method, where the classical mechanics-based MM and the quantum chemistry methods are used on the environment and the active site, respectively. The approach was originally developed with semi-empirical QM methods for biochemical reactions by Karplus, Levitt, and Warshel in the 1970s [17, 18]. The QM/MM technique is still one of the most popular and widely used approach in computational chemistry mostly because its efficiency is unbeatable for the modeling of the reactions of nanoparticles. However, its application is not trivial, and unanswered questions still rise regarding its accuracy. In the following, a detailed discussion of the QM/MM method is presented, which starts with the introduction of the basics of MM, with an emphasis on the MM energy functional. Please note that despite the close relation, molecular dynamics will not be discussed.

1.2.1 Molecular mechanics

In quantum chemistry, the electrons and the nuclei are the fundamental building blocks. The motions and interactions of these particles are described by the Schrödinger equation, which yields the energy and the electronic structure of a microsystem. In contrast, the electrons are no longer separate entities in the MM because of the introduction of atoms as new building blocks. Moreover, the motions and interactions of the atoms are described by classical mechanics and static electricity with empirical inclusion of electron correlation by van der Waals (vdW) interactions. The neglect of electrons as individual particles has the consequence that the electronic structure is no longer a result of the simulation, a presumption for the environment of atoms is required, which manifests in the extensive use of, more or less, empirically derived parameters. Nevertheless, if the description of the atomic environments is reasonable, and no chemical rearrangement occurs, then MM can be accurately and efficiently describe interactions of large systems.

1.2.1.1 The MM energy functional

The MM method, which is also called the force field model of molecules, can be visualized as spheres, which can be characterized by their vdW radius, have point-like mass and electric charge in their center, and are bounded by springs. Interactions in this model can be separated as bonded (E_b) and nonbonded (E_nb) interactions, and the energy functional, E_MM, can be written in its most compact form as

$$E_\mathrm{MM} = E_\mathrm{b} + E_\mathrm{nb}. \tag{1.2.1}$$

In the following, the bonded and nonbonded energy terms are briefly discussed and decomposed according to the most widely used force fields, namely the classical force fields [19].

Bonded terms The force field model accounts not only for the stretching of bonds via springs (E_{str}) but also incorporates bending (E_{bend}) and rotation motions (E_{tors}) along the bonds:

$$E_{\mathrm{b}} = E_{\mathrm{str}} + E_{\mathrm{bend}} + E_{\mathrm{tors}}. \tag{1.2.2}$$

These functionals describe the major part of the intramolecular interactions. Expressions can be derived for the stretching and bending energies by expanding Taylor-series around equilibrium bond lengths and angles and terminating them at the second order, which results in a harmonic oscillator model:

$$E_{\mathrm{str}} = \sum_{\mathrm{bonds}} \frac{k_{\mathrm{b}}}{2} \left(l_{\mathrm{b}} - l_{0,\mathrm{b}}\right)^2, \tag{1.2.3}$$

$$E_{\mathrm{bend}} = \sum_{\mathrm{angles}} \frac{k_{\mathrm{a}}}{2} \left(\Theta_{\mathrm{a}} - \Theta_{0,\mathrm{a}}\right)^2, \tag{1.2.4}$$

where k_{b} and k_{a} are force constants of the bond stretching and bending, l_{b} and Θ_{a} are the actual distances and angles defined by the corresponding atoms, whereas $l_{0,\mathrm{b}}$ and $\Theta_{0,\mathrm{a}}$ are the equilibrium parameters of the latter quantities, respectively. Utilizing the simple harmonic oscillator model for the bond stretching is sufficient in most cases where the geometry is close to the equilibrium, but strained systems require higher-order terms of the Taylor-series or the use of the Morse potential, which has correct asymptotic behavior. Higher-order corrections are rarely used for the bending term, however, in specific functional groups, for example, in amino groups, special out-of-plane energy terms are applied to provide correct chemical structures. Bond rotation requires relatively low energy, hence even small distortions result in torsion angles far from the energy minimum, thus expression for E_{tors} can be written as Fourier-series instead of Taylor-series as

$$E_{\mathrm{tors}} = \sum_{\mathrm{torsions}} \sum_{n} \frac{V_n}{2} \left(1 + \cos\left(n\omega - \gamma\right)\right), \tag{1.2.5}$$

where V_n, n, γ, and ω denote the rotation barrier height, the periodicity, the phase, and the torsion angle of a rotation, respectively. In contrast to the stretching and bending terms, bond rotation from the energy minimum can be induced by nonbonded interactions, hence the treatment of E_{tors} requires special attention at the parameterization process.

Nonbonded terms The nonbonded energies account for the intermolecular and some of the intramolecular interactions. E_{nb} consists of the electrostatic interactions (E_{el}) and

the vdW interactions (E_{vdW}):

$$E_{\mathrm{nb}} = E_{\mathrm{el}} + E_{\mathrm{vdW}}. \tag{1.2.6}$$

A part of the intramolecular energies is described by bonded terms, hence nonbonded interactions within 2 bonds are disregarded (1-3 interactions), and nonbonded interactions of atom pairs 3 bonds away (1-4 interactions) are carefully treated, as the latter terms are coupled with torsion motions. In practice, electrostatic contributions are mostly calculated by applying Coulomb's law with partial atomic charges:

$$E_{\mathrm{el}} = \sum_{\alpha} \sum_{\beta>\alpha} \frac{1}{4\pi\epsilon_0} \frac{q_\alpha q_\beta}{|\boldsymbol{R}_\alpha - \boldsymbol{R}_\beta|}, \tag{1.2.7}$$

where the q_α (q_β), $\boldsymbol{R}_\alpha$ ($\boldsymbol{R}_\beta$) mean the partial charge and position vector of atom α (β), respectively, and ϵ_0 is the dielectric constant of vacuum. A more sophisticated approach is to include multipole-multipole and multipole-charge interactions in the electrostatic term, however, except for systems with strong dipole interactions (e.g., ketones containing halogen at the α position), the results are more or less similar in the case of close contacts, and the same at large distances. In the more advanced and recently developed force fields, induced dipoles are added to the fixed charges to model polarization effects [20]. On the other hand, electrostatics cannot account for all distant interactions as even molecules without dipole moments exhibit attraction at large separation (commonly known as dispersion), while repulsion forces act at short ranges. The well-known vdW forces, which originate from electron correlation, are responsible for such effects, and in most force fields a simple 12-6 Lenard-Jones potential is used:

$$E_{\mathrm{vdW}} = \sum_{\alpha} \sum_{\beta\neq\alpha} 4\epsilon_{\alpha\beta} \left[\left(\frac{\sigma_{\alpha\beta}}{|\boldsymbol{R}_\alpha - \boldsymbol{R}_\beta|} \right)^{12} - \left(\frac{\sigma_{\alpha\beta}}{|\boldsymbol{R}_\alpha - \boldsymbol{R}_\beta|} \right)^{6} \right], \tag{1.2.8}$$

where the $\epsilon_{\alpha\beta}$ is the maximum attraction energy of atoms α and β, and $\sigma_{\alpha\beta}$ is a collision parameter which defines the distance where the magnitude of attraction and repulsion energies are the same, i.e., the vdW energy is zero. Other expressions are also introduced in order to better reproduce experimental structures by changing the exponents of the repulsion and attraction terms to 14 and 7, respectively, or by the use of Buckingham [21] or Hill [22] potentials. Nevertheless, the 12-6 Lenard-Jones potential is widely used because it has reasonable accuracy and is computationally the most efficient formula.

It is also useful to note that Eq. 1.2.1 is not general, the equation mainly reflects the functional form of the classical force fields, such as AMBER [23], CHARMM [24], GROMOS [25], and OPLS-AA [26], which are widely used for systems with biological relevance. A general expression can be written as

$$E_{\mathrm{MM}} = E_{\mathrm{str}} + E_{\mathrm{bend}} + E_{\mathrm{tors}} + E_{\mathrm{el}} + E_{\mathrm{vdW}} + E_{\mathrm{cross}}, \tag{1.2.9}$$

where E_{cross} denotes the coupling terms of the aforementioned bonded and nonbonded terms. Theoretically, the coupling of all terms are required for exact energy evaluation but, in practice, the classical form of the energy functional is sufficiently accurate in most cases. The simulation of vibrational spectra is a typical example where the coupling of the stretching and the bending terms is necessary for accurate results.

1.2.1.2 Evaluation of the energy functional

As a consequence of using atoms as building blocks in the force field model, atom types have to be assigned to each atom, and parameters necessary for the evaluation of the MM energy functional have to be determined. In exchange for the efficiency, the cost of using force fields manifests in careful fitting of the necessary parameters to experimental data or high-level QM methods, which can be a complicated task. The number of required constants grows steeply with each new atom type, moreover, the derived numbers are mostly not transferable between systems. For instance, carbohydrates and amino acids are build from similar functional groups, but accurate modeling requires different sets of parameters. Nevertheless, force fields are highly developed for nearly all chemical environments, for example, biomolecules, clays, and metals [27], however, parameter determination is a recurrent problem for new materials.

All in all, the biggest issue with MM goes back to its fundamentals: the need of the assignment of atom types to atoms *a priori* to the simulation. Progress has been made by the development of reacting force fields [28], however, the discussion continues with the overview of QM/MM, which tries to incorporates the advantageous attributes of both the QM and the MM approaches.

1.2.2 QM/MM variants

Systems exhibiting local chemical changes are suitable for QM/MM calculations because (i) the reaction center can be handled by QM methods which are capable of describing the electronic structure, (ii) the locality of electron density change allow for the valid *a priori* assignation of atom types to the atoms of the environment, while limiting the system size treated by QM techniques, and finally (iii) the extensive procedure of parameter derivation for atoms at the reaction center can be bypassed. The fundamental equation of QM/MM methods can be expressed by properly rewriting Eq. 1.1.3 as

$$E_{\text{QM/MM}} = E_{\text{MM}} + E_{\text{QM}} + E_{\text{QM-MM}}, \tag{1.2.10}$$

where $E_{\text{QM/MM}}$, E_{MM}, E_{QM}, and $E_{\text{QM-MM}}$ are the energy of the full system, the environment, the reaction center, and the interactions of the subsystems, respectively. Basically, Eq. 1.2.10 can be evaluated in 3 different ways, with respect to the approximations used for the subsystems and for the interaction of the subsystems [29].

1.2.2.1 Mechanical embedding

The name, mechanical embedding (ME), is somewhat delusive because the method is an ONIOM-like multilevel approach instead of an embedding technique. In this scheme, the QM/MM energy equation can be expressed by rewriting Eq. 1.1.5, the basic equation of the subtractive methods as

$$E^{\mathrm{ME}}_{\mathrm{QM/MM}}(a_f, e_f) = E_{\mathrm{MM}}(a_f, e_f) + E_{\mathrm{QM}}(a_f) - E_{\mathrm{MM}}(a_f), \tag{1.2.11}$$

where $E^{\mathrm{ME}}_{\mathrm{QM/MM}}$ means the QM/MM energy with ME.

There are two merits and dismerits of the approach in the QM/MM framework. The first and foremost advantage is its reasonable accuracy in many cases [30] compared with the other schemes. Secondly, it is easy to implement due to its computational simplicity. The major drawback of using this approach is the insensitive electron density of the QM region to the changes of its environment because the interaction of the subsystems is described only at the MM level. Due to the latter attribute, the method can be considered as an improved version of vacuum embedding. The requirement of the atom type determination in the QM region is another disadvantage because avoiding the atom type assignation and parameter determination can be crucial, for example, in studies where non-standard amino acid residues are used.

1.2.2.2 Electronic embedding

In this QM/MM scheme, the description of the interaction energy is improved by letting the point charges act on the electrons of the active subsystem. The fundamental equation of electronic embedding (EE) can be written as

$$E^{\mathrm{EE}}_{\mathrm{QM/MM}}(a_f, e_f) = E_{\mathrm{MM}}(e_f) + E_{\mathrm{QM}}(a_f, e_f) + E_{\text{QM-MM}}(a_f, e_f), \tag{1.2.12}$$

where the above equation is the reformulation of the additive scheme, though the subtractive ONIOM scheme is also developed with the EE approach [31]. Note that E_{MM} only depends on the variables of the environment, which means that the evaluation of the MM energy functional is not required in those parts of the system where the electronic structure changes. This means, apparently, that the MM parameters of the QM atoms are not necessary, however, this is not entirely true because of the interaction of subsystems:

$$E_{\text{QM-MM}} = E^{\mathrm{b}}_{\text{QM-MM}} + E^{\mathrm{vdW}}_{\text{QM-MM}} + E^{\mathrm{el}}_{\text{QM-MM}}. \tag{1.2.13}$$

The $E^{\mathrm{b}}_{\text{QM-MM}}$ term is the collection of bonded interactions in the interface of the two theories, thus it consists of those terms of E_{str}, E_{bend}, and E_{tors} which require one or more QM atoms. The second term, $E^{\mathrm{vdW}}_{\text{QM-MM}}$, is the vdW interaction of the environment and the active subsystem atoms. This term can be considered problematic at first glance due to the necessary MM parameters, however, the vdW forces decay rapidly, and if the active

site is large enough, then the $E^{\mathrm{vdW}}_{\mathrm{QM\text{-}MM}}$ contribution is insignificant even with inaccurate constants. The third term, $E^{\mathrm{el}}_{\mathrm{QM\text{-}MM}}$, is the electrostatic interaction of subsystems, which can be decomposed into electron-point charge and nucleus-point charge parts:

$$E^{\mathrm{el}}_{\mathrm{QM\text{-}MM}} = -\sum_i \sum_{\alpha \in \mathrm{MM}} \int \psi_i^* \frac{q_\alpha}{|\boldsymbol{r} - \boldsymbol{R}_\alpha|} \psi_i \mathrm{d}\boldsymbol{r} + \sum_{\beta \in \mathrm{QM}} \sum_{\alpha \in \mathrm{MM}} \frac{Z_\beta q_\alpha}{|\boldsymbol{R}_\beta - \boldsymbol{R}_\alpha|}, \tag{1.2.14}$$

where Z_β and $\boldsymbol{R}_\beta$ are the atomic number and the position vector of the QM nucleus β, respectively, ψ_i and ψ_i^* denote occupied molecular orbital (MO) i and its complex conjugate, respectively, and $\boldsymbol{r}$ denotes the position vector of an electron. Note that atomic units are introduced here and will be used in the rest of the thesis. The first term of Eq. 1.2.14 is a one-electron integral quantity, which accounts for the attraction (or repulsion) between the electron density and the point charges of the environment, thus the electron density is polarized by its surroundings in every self-consistent field (SCF) iteration. The second term describes the electrostatic interaction of the point charges and the nuclei of the QM region.

Despite the improvements in EE, such as the polarized electron density and the insignificance of parameters at the active site, it has several technical issues, which will be discussed in Sect. 1.2.3. Nevertheless, its relative simplicity and the incorporation of some of the polarization effects made the method the most widely used QM/MM approaches.

1.2.2.3 Polarization embedding

A logical development over the EE method is the polarization embedding (PE) technique, which aims to improve the one-sided polarization of EE by incorporating the mutual polarization of subsystems. Again, the basic equation of PE can be written as an additive scheme:

$$E^{\mathrm{PE}}_{\mathrm{QM/MM}}(a_f, e_f) = E_{\mathrm{MM}}(a_f, e_f) + E_{\mathrm{QM}}(a_f, e_f) + E_{\mathrm{QM\text{-}MM}}(a_f, e_f). \tag{1.2.15}$$

Note that now every energy term depends on variables of both subsystems. Including polarization in the MM region has to come with the reconsideration of the rigid electrostatic model of classical force fields. In the progressing field of polarizable force fields [20], additional computations have to be carried out other than calculating Eq. 1.2.7 by solving classical electrostatic equations. In the kernel of QM/MM with PE, the QM and MM calculations are carried out in the frozen electric field of the other region repeatedly, until self-consistency is reached for both subsystems. The repeated QM calculations obviously mean extremely high price , however, the computational costs can be lowered if the number of QM calculations are constrained, or with other more efficient approaches [32].

1.2.3 Coupling techniques

The different theoretical foundations of the MM and QM methods imply issues regarding the handling of the interaction energy especially at the subsystem interface. The separation of a system into two subsystems is straightforward in the cases where the different regions are not connected by covalent bonds. On the contrary, partitioning of a more complicated system, for example, a reacting enzyme, is not a trivial task because one has to set the subsystem border across covalent bonds at one point to avoid a QM region with the unmanageable large size. Nevertheless, three major concerns arise regarding the combination of the MM and QM approaches at the border:

1. **The problem of different building blocks: how to treat the bond at the QM and MM border?**
   If the atoms of the system are classified as QM and MM atoms, then the QM region has one or more unpaired electrons (which are more commonly referenced as dangling bonds) due to the QM atoms at the border. Neglecting the problem leads to a model system with radicals, which are not present in the real system.

2. **The problem of electron spill out: how the overpolarization of the density can be avoided in the case of electrostatic embedding?**
   Treating electrostatics as interacting point charges is only a valid approach if the charges are far from each other. This is not the case for the electron-point charge interactions at the border, which can result in unrealistic distortions in the electron density, especially with flexible (diffuse) basis sets.

3. **The problem of double counting: which terms of the MM energy functional can be dropped?**
   It is not a straightforward task to retain bonded terms at the border because stretching, bending, and torsional terms can depend on QM atoms, while some of the intramolecular energy is already counted by the QM method.

Overall, these questions are less relevant as the QM region enlarges, however, recent research in the field suggests that quite big QM subsystem is necessary for accurate QM/MM modeling. Several groups reported [33, 34] that the QM/MM reaction energies slowly reach their asymptotic limit with respect to the QM subsystem size since the border of the QM region should be at least 5-6 Å far from the reacting center. The general approach regarding the minimal QM region is that the reacting atoms should not be included in MM torsional terms [10], which means that the radius of the QM subsystem is at least 3-4 Å large. As the condensed-phases are quite dense with atoms, one has to seriously consider the selection of QM region to minimize the aforementioned problems of QM/MM while maintaining efficiency.

The need for decreasing the computational efforts has been the driving force for the development of coupling methods. The corresponding studies mainly focus on the problem of dangling bonds, the major source of errors, while the 2nd and 3rd questions are often handled in an *ad hoc* manner. Overpolarization is often minimized by simply zeroing or redistributing [35] MM charges near the interface, while more sophisticated approaches are rarely used [36, 37]. The problem of double counting is always neglected and usually those bonded terms are retained which include at least one MM atom. In the following, the two main coupling approaches are discussed: the link atom and the frozen MO schemes, which are illustrated in Fig. 1.3. It should be noted that the effective fragment potential method (EFP) [38], which is also capable of modeling environmental effects, follows a rather different embedding strategy, thus it will not be discussed.

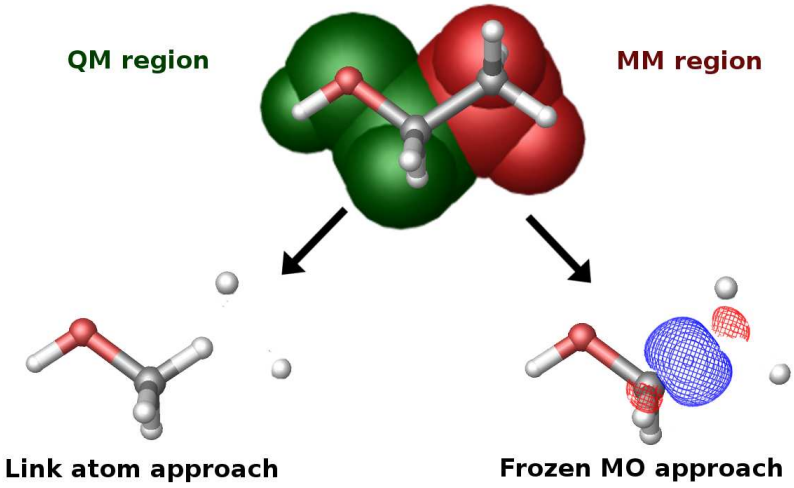

FIGURE 1.3: The link atom and frozen MO coupling schemes are illustrated on ethanol, where the system is separated across the carbon-carbon bond (top figure). In the link atom approach (bottom-left), the hydrogen capping turns ethanol into a methanol, while the carbon-carbon bond is retained in the frozen MO approach (bottom-right). The point charge of the MM host atom (the MM carbon atom) is usually zeroed, and only the electric field of the other point charges effects the QM region.

1.2.3.1 Link atoms

The simplest way to tackle the problem of dangling bonds is to saturate them by inserting atoms at the interface, which are not present in the real system. These atoms add extra degrees of freedom to the system artificially, but constraining the position of link atoms to the QM and MM border atoms solves the issue.

The link atom methods are easy to implement and have acceptable accuracy (if used properly), thus they are extensively (if not exclusively) utilized in QM/MM applications. The largest issue of the ansatz is that the electron density can be largely distorted at the

border. Typically hydrogen atoms are used as link atoms, and, if polar bonds are cut, for instance, the σ bond between an MM oxygen and a QM carbon atom, the charge character of the atoms may change in an artificial way. "Cutting" through σ bonds of functional groups is more troublesome. Thus, the common guide for QM/MM calculations is to cut across nonpolar bonds, e.g., bonds between two sp^3 hybridized carbon atoms, although this constrain may enlarge the QM region significantly. Efforts are made to reproduce the electric field at the border by applying parameterized, monovalent pseudoatoms [39, 40] as link atoms, but this approach has minor improvements in exchange for its limited applicability.

1.2.3.2 Frozen localized MOs

Instead of cutting through bonds and saturating them with non-existing atoms, it would be more feasible to keep the bonds at the border and "freeze" them, i.e., to prevent the remixing of atomic orbital coefficients during the SCF procedure. The approach is theoretically advantageous because it naturally treats bonds of arbitrary order and functional groups, while preserving polarity.

The idea of using localized, frozen bonds to connect the QM and MM subsystems originates from Warshel and Levitt, who used strictly localized orbitals and predefined hybrids [18]. The local SCF (LSCF) approach applied a single hybrid orbital between the subsystems, while other electrons of the MM host atom are not treated explicitly, and a force field description is invoked. In this scheme, the connecting MO is not included in the SCF procedure, rather its coefficients are kept frozen at some predefined values. Different orbital equation and specific parameterization was proposed by Friesner and co-workers [41, 42], while the generalized hybrid orbital (GHO) approach uses frozen hybrids for all the orbitals of the MM host atom, except for the connecting hybrid orbital, which is allowed for orbital relaxation [43]. The early versions of the LSCF and related methods were designed for semi-empirical techniques, where minimal basis set is used, and the orthogonality is guaranteed. Generalization of the LSCF scheme to the DFT and *ab initio* levels was not trivial and it required intense modifications in the orbital equations of the latter approaches to keep atomic orbital coefficients of the frozen MOs fixed while maintaining orthogonality [44]. Recently, a simpler LSCF method was proposed by Ferenczy [45, 46] which will be discussed in the next chapter.

The frozen localized orbital approach is more elegant than the link atom schemes, however, it does not improve the QM/MM ansatz automatically. Transferability issues raise with the use of empirical parameters, the difficult implementation inhibits its widespread use, and the accuracy improvement is not guaranteed. Nevertheless, the approach has great potential in improving the QM/MM technique, and our attempt is presented in Chapter 2.

1.3 Embedding into quantum mechanics

Quantum embedding techniques stand on the top of the hierarchy of the focused models because they include quantum effects into the description of the environment and the interaction of subsystems. The motivation behind such QM/QM approaches is to combine methods which have large prefactor or high scaling behavior with less costly approaches to decrease the overall computational demand. Obviously, these schemes are not as efficient as QM/MM, however, they are free from the previously discussed QM/MM problems, thus their use is more appealing.

Apart from the ONIOM approach using the QM/QM technique, QM-based embedding or multilevel schemes can often be considered as the constrained, multilevel variants of linear-scaling methods, which were developed over the years to tackle the scaling problem of the conventional electronic structure techniques. Two such methods are closely related to the topics of the thesis: the subsystem DFT (sDFT) [47] and the local correlation approaches [48]. The former serves as the theoretical basis for our scheme that is capable of embedding QM methods into DFT, while the latter is fundamental to our multilevel scheme that applies high- and low-level WFT-based approaches to different regions. It is worth mentioning that there are other efficient approaches in the spirit of focused models [49, 50] following a much different embedding philosophy. Nevertheless, this part of the thesis is focused on DFT and WFT embedding schemes, therefore an overview of the related fields will be presented here.

1.3.1 Embedding into density functional theory

The purpose of this subsection is to review the basic concepts required to understand our DFT embedding approach. First, a quick introduction to the KS DFT approach will be presented, which will be useful for understanding sDFT, moreover, the basic notations in quantum chemistry will be introduced. Then, sDFT will be discussed, however, it should be noted that the overview of this active and recent field is far from being comprehensive. It is still useful to put our developments in context, and in the last part, the connection of focused models and the sDFT kernel is described to finally arrive at the parent technique of our DFT embedding scheme.

1.3.1.1 Kohn–Sham DFT

In DFT, the ground-state energy is a unique functional of the electron density of a system, thus there is an unambiguous correspondence between the ground-state properties of a system and its electron density. The theory originates from the Thomas–Fermi model and was further developed by Hohenberg and Kohn, but it was the Kohn–Sham

model that achieved breakthrough in quantum chemistry in the 1990s because some of its variants provided accuracy that is comparable with the MP2 method but at a reduced cost.

The energy of a system in KS DFT is a functional of the electron density, $\rho(\boldsymbol{r})$, which is a probability function expressing the likelihood of finding the electrons in a $\mathrm{d}\boldsymbol{r}$ volume at the position $\boldsymbol{r}$. In the KS DFT kernel, the energy, E_{KS}, can be decomposed as the sum of the following energy functionals (and the nuclear repulsion term):

$$E_{\mathrm{KS}}[\rho] = T_s[\rho] + V_{\mathrm{nuc}}[\rho] + J[\rho] + E_{\mathrm{xc}}[\rho] + V_{\mathrm{nn}}, \tag{1.3.1}$$

where the T_s, V_{nuc}, J, E_{xc}, and V_{nn} is the approximate kinetic energy functional, the nucleus-electron interaction energy, the Coulomb repulsion energy, the exchange-correlation energy, and the repulsion of the nuclei, respectively. Due to the large differences between the masses of the nuclei and the electrons, we can assume that the positions of the nuclei are fixed, thus V_{nn} is a constant. It can be computed using the positions vectors ($\boldsymbol{R}$) and the atomic numbers (Z) of the corresponding atoms:

$$V_{\mathrm{nn}} = \frac{1}{2}\sum_{\alpha}\sum_{\beta\neq\alpha}\frac{Z_\alpha Z_\beta}{|\boldsymbol{R}_\alpha - \boldsymbol{R}_\beta|}. \tag{1.3.2}$$

The interaction energy of the electrons and the nuclei is calculated as the integral of the external potential, $\hat{V}_{\mathrm{nuc}}$, and the electron density:

$$V_{\mathrm{nuc}}[\rho] = -\int \rho(\boldsymbol{r})\hat{V}_{\mathrm{nuc}}\mathrm{d}\boldsymbol{r} = -\int \sum_{\alpha}\frac{\rho(\boldsymbol{r})Z_\alpha}{|\boldsymbol{R}_\alpha - \boldsymbol{r}|}\mathrm{d}\boldsymbol{r}. \tag{1.3.3}$$

The interaction of electrons with their mean field is accounted for by Coulomb's law as

$$J[\rho] = \frac{1}{2}\int\int \frac{\rho(\boldsymbol{r}_1)\rho(\boldsymbol{r}_2)}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|}\mathrm{d}\boldsymbol{r}_1\mathrm{d}\boldsymbol{r}_2, \tag{1.3.4}$$

where $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ stand for the coordinates of electron one and electron two, respectively. The functional form of the $E_{\mathrm{xc}}[\rho]$ term is unknown, but it can be formally expressed with the exact and approximate forms of the electron-electron repulsion ($V_{\mathrm{ee}}[\rho]$) and the kinetic energy functionals ($T[\rho]$):

$$E_{\mathrm{xc}}[\rho] = (V_{\mathrm{ee}}[\rho] - J[\rho]) + (T[\rho] - T_s[\rho])\,. \tag{1.3.5}$$

In practice, an approximate functional is used:

$$E_{\mathrm{xc}}[\rho] = \int f[\rho(\boldsymbol{r})]\mathrm{d}\boldsymbol{r}, \tag{1.3.6}$$

where $f[\rho(\boldsymbol{r})]$ is a function of the electron density. In the KS theory a convenient model system, the noninteracting KS reference is introduced. However, electrons in a real system do interact, but the approximation is still useful to express the density as the sum of single-particle densities. As a result of the non-interacting reference system, the independent

particle model invokes the Kohn–Sham MOs (ψ_i):

$$\rho(\boldsymbol{r}) = \sum_i \rho_i(\boldsymbol{r}) = \sum_i |\psi_i(\boldsymbol{r})|^2. \tag{1.3.7}$$

The approximate kinetic energy functional, $T_s[\rho]$, is exact if one considers the non-interacting KS reference system. The $T_s[\rho]$ functional can be also considered as a good approximation to the exact $T[\rho]$ functional, thus the KS orbitals are used to compute the approximate kinetic energy with the kinetic energy operator, which differentiates with respect to electron coordinates:

$$T_s[\rho] = \sum_i \langle\psi_i| - \frac{1}{2}\hat{\nabla}^2|\psi_i\rangle, \tag{1.3.8}$$

where the bra-ket notation is used: a given column vector r is denoted as $|r\rangle$ and called "ket", while its conjugate transpose row vector is called "bra": $(|r\rangle)^* = \langle r|$. Most practical calculations use real numbers as vector elements hence the following also holds: $(|r\rangle)^{\mathrm{T}} = \langle r|$. After formulating the energy functional, one can search for the ground state density of the system by minimizing the energy through the variation of the KS orbitals under the constrain that the KS orbitals are orthonormal. This can be achieved by solving the Kohn–Sham pseudo-eigenvalue equations iteratively:

$$\hat{F}[\rho]|\psi_i\rangle = \epsilon_i|\psi_i\rangle, \tag{1.3.9}$$

where ϵ_i's are the eigenvalues of the Kohn–Sham operator, $\hat{F}$. The latter can be written as

$$\hat{F}[\rho] = -\frac{1}{2}\hat{\nabla}^2 + \hat{V}_{\mathrm{eff}}[\rho], \tag{1.3.10}$$

where the effective potential operator, $\hat{V}_{\mathrm{eff}}$, is the sum of the external potential ($\hat{V}_{\mathrm{nuc}}$), the Coulomb potential ($\hat{J}$), and the exchange-correlation potential operators ($\hat{V}_{\mathrm{xc}}$):

$$\hat{V}_{\mathrm{nuc}} = \frac{\partial V_{\mathrm{nuc}}[\rho]}{\partial\rho}, \tag{1.3.11}$$

$$\hat{J}[\rho] = \frac{\partial J[\rho]}{\partial\rho}, \tag{1.3.12}$$

$$\hat{V}_{\mathrm{xc}}[\rho] = \frac{\partial E_{\mathrm{xc}}[\rho]}{\partial\rho}. \tag{1.3.13}$$

It is appropriate to reformulate the Kohn–Sham energy and the eigenvalue equation here with the introduction of the linear combination of atomic orbitals (LCAO) ansatz because it is more efficient computationally to vary the molecular orbitals through the expansion coefficients ($C_{\mu i}$) of atomic orbitals (AO), χ_μ:

$$|\psi_i\rangle = \sum_\mu |\chi_\mu\rangle C_{\mu i}. \tag{1.3.14}$$

Using this approach, the density dependency of Eq. 1.3.1 and 1.3.9 can be redefined with the density matrix, which, for closed-shell systems, reads as

$$\mathrm{D}_{\mu\nu} = 2\sum_{i} \mathrm{C}_{\mu i}(\mathrm{C}_{\nu i})^{\mathrm{T}}, \tag{1.3.15}$$

where index i runs over all occupied orbitals, while $\mathrm{D}_{\mu\nu}$ is the element of the density ($\mathbf{D}$) matrix. In the AO basis, the Kohn–Sham equation for a closed-shell system can be rewritten with the matrix notation as

$$\mathbf{F}[\mathbf{D}]\mathbf{C} = \mathbf{SCE}, \tag{1.3.16}$$

where $\mathbf{F}$, $\mathbf{S}$, and $\mathbf{E}$ are the matrix representation of the Kohn–Sham operator, the overlap matrix, and the diagonal matrix of the eigenvalues, respectively. The matrix elements of $\mathbf{S}$ are defined as the scalar product of the overlapping AOs:

$$\mathrm{S}_{\mu\nu} = \langle \chi_\mu | \chi_\nu \rangle, \tag{1.3.17}$$

and $\mathbf{F}[\mathbf{P}]$ can be decomposed as

$$\mathbf{F}[\mathbf{D}] = \mathbf{h} + \mathbf{G}[\mathbf{D}], \tag{1.3.18}$$

where the core-Hamiltonian matrix, $\mathbf{h}$, includes the one-electron integrals incorporating the kinetic energy and the nuclear attraction potential:

$$\mathrm{h}_{\mu\nu} = \langle \chi_\mu | -\frac{1}{2}\hat{\nabla}^2 - \sum_{\alpha} \frac{Z_\alpha}{|\boldsymbol{R}_\alpha - \boldsymbol{r}|} | \chi_\nu \rangle. \tag{1.3.19}$$

The two-electron term, $\mathbf{G}$, includes the Coulomb ($\mathbf{J}$), the exact-exchange ($\mathbf{K}$), and the exchange-correlation ($\mathbf{V}_{\mathrm{xc}}$) potentials:

$$\mathbf{G}[\mathbf{D}] = \mathbf{J}[\mathbf{D}] - a_{\mathrm{xc}}\frac{1}{2}\mathbf{K}[\mathbf{D}] + \mathbf{V}_{\mathrm{xc}}[\mathbf{D}], \tag{1.3.20}$$

where a_{xc} is the exact-exchange mixing parameter in the case of hybrid functionals. The elements of the aforementioned matrices are defined as:

$$\mathrm{J}_{\mu\nu}[\mathbf{D}] = \sum_{\sigma\lambda} \mathrm{D}_{\sigma\lambda} \langle \chi_\mu \chi_\sigma | \chi_\nu \chi_\lambda \rangle, \tag{1.3.21}$$

$$\mathrm{K}_{\mu\nu}[\mathbf{D}] = \sum_{\sigma\lambda} \mathrm{D}_{\sigma\lambda} \langle \chi_\mu \chi_\sigma | \chi_\lambda \chi_\nu \rangle, \tag{1.3.22}$$

$$\left(\mathrm{V}_{\mathrm{xc}}[\mathbf{D}]\right)_{\mu\nu} = \frac{\partial E_{\mathrm{xc}}[\rho]}{\partial \mathrm{D}_{\mu\nu}}, \tag{1.3.23}$$

where the AOs are integrated over the coordinates of electron 1 and 2 in Eq. 1.3.21 and 1.3.22, and the two-electron integrals are expressed in the physicists' notation:

$$\langle \chi_\mu \chi_\sigma | \chi_\nu \chi_\lambda \rangle = \int\int \frac{\chi_\mu^*(\boldsymbol{r}_1)\chi_\sigma^*(\boldsymbol{r}_2)\chi_\nu(\boldsymbol{r}_1)\chi_\lambda(\boldsymbol{r}_2)}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} \mathrm{d}\boldsymbol{r}_1 \mathrm{d}\boldsymbol{r}_2. \tag{1.3.24}$$

The final KS energy expression for a closed-shell system using the matrix notation can be written as

$$E_{\mathrm{KS}}[\mathbf{D}] = \mathrm{Tr}\left[\mathbf{hD} + \frac{1}{2}\mathbf{D}\left(\mathbf{J}[\mathbf{D}] - \frac{1}{2}a_{\mathrm{xc}}\mathbf{K}[\mathbf{D}]\right)\right] + E_{\mathrm{xc}}[\mathbf{D}] + V_{\mathrm{nn}}, \tag{1.3.25}$$

where one should note that using the $\mathbf{V}_{\mathrm{xc}}$ potential (and thus $\mathbf{F}$) is not appropriate for the energy evaluation because the integration of the potential with the density only results in the exchange-correlation energy if the exact form of the potential is known.

The accuracy of KS DFT is determined by the approximations introduced in the E_{xc} functional, thus the major goal of the DFT research is to find more accurate functional forms. As a rule of thumb, the accuracy increases as the approximation of E_{xc} is more sophisticated, which is symbolized by the rungs of "Jacob's ladder" of DFT functionals. However, higher rungs of the ladder also come with increased computational cost, and the accuracy improvement is not guaranteed. The latter feature clearly distinguishes KS DFT from the WFT-based *ab initio* approaches, where better results are expected with the improved description of electron correlation.

1.3.1.2 Subsystem DFT

The theory of sDFT is closely related to the KS DFT framework. The basic idea is that the system is partitioned into subsystems, then modified Kohn–Sham equations are solved for each subsystem in the presence of the embedding potential of the other subsystems. Then, the unique embedding potential of each subsystem is recalculated and the procedure is repeated iteratively until all the subsystem densities reach self-consistency.

The density in sDFT is separated into subsystem densities. For simplicity, consider a closed-shell system with n electrons and $\rho(\boldsymbol{r})$ electron density which is separated into the contribution of two subsystems, A and B:

$$\rho(\boldsymbol{r}) = \rho^{\mathrm{A}}(\boldsymbol{r}) + \rho^{\mathrm{B}}(\boldsymbol{r}), \tag{1.3.26}$$

$$n = \int \rho(\boldsymbol{r})\mathrm{d}\boldsymbol{r} = \int \left[\rho^{\mathrm{A}}(\boldsymbol{r}) + \rho^{\mathrm{B}}(\boldsymbol{r})\right] \mathrm{d}\boldsymbol{r} = n^{\mathrm{A}} + n^{\mathrm{B}}, \tag{1.3.27}$$

where the subsystems are denoted with capital letters, while n^{A} and n^{B} denotes the number of electrons in the subsystems. As in KS DFT, the subsystem densities are defined by the particles of the corresponding subsystem, i.e., the subsystem orbitals ($\{\psi_i^{\mathrm{A}}\}$ and $\{\psi_i^{\mathrm{B}}\}$):

$$\rho^{\mathrm{A}}(\boldsymbol{r}) = \sum_{i\in\mathrm{A}}^{n^{\mathrm{A}}/2} |\psi_i^{\mathrm{A}}(\boldsymbol{r})|^2, \qquad \rho^{\mathrm{B}}(\boldsymbol{r}) = \sum_{i\in\mathrm{B}}^{n^{\mathrm{B}}/2} |\psi_i^{\mathrm{B}}(\boldsymbol{r})|^2, \tag{1.3.28}$$

Note that we are using this approximation for each subsystem but not for the whole system. As a consequence, the approximate kinetic energy of the Kohn–Sham reference

system can be calculated in each subsystem but not for the whole system:

$$T_s[\rho^{\mathrm{A}}] = T_s[\{\psi_i^{\mathrm{A}}\}] = \sum_i^{n_{\mathrm{A}}/2} \langle \psi_i^{\mathrm{A}}| -\frac{1}{2}\hat{\nabla}^2 |\psi_i^{\mathrm{A}}\rangle, \qquad T_s[\rho^{\mathrm{B}}] = \sum_i^{n_{\mathrm{B}}/2} \langle \psi_i^{\mathrm{B}}| -\frac{1}{2}\hat{\nabla}^2 |\psi_i^{\mathrm{B}}\rangle, \quad (1.3.29)$$

$$T_s[\rho] \neq T_s[\rho^{\mathrm{A}}] + T_s[\rho^{\mathrm{B}}]. \quad (1.3.30)$$

The reason for the inequality is that the orbitals (and therefore the density) of the different subsystems overlap:

$$\langle \psi_i^{\mathrm{A}} | \psi_j^{\mathrm{A}} \rangle = \delta_{ij}, \qquad \langle \psi_i^{\mathrm{B}} | \psi_j^{\mathrm{B}} \rangle = \delta_{ij}, \qquad \langle \psi_i^{\mathrm{A}} | \psi_j^{\mathrm{B}} \rangle \neq \delta_{ij}. \quad (1.3.31)$$

The equality can be formally regained with the introduction of the nonadditive kinetic energy functional, T_s^{nad},:

$$T_s[\rho] = T_s[\rho^{\mathrm{A}}] + T_s[\rho^{\mathrm{B}}] + T_s^{\mathrm{nad}}[\rho^{\mathrm{A}}, \rho^{\mathrm{B}}]. \quad (1.3.32)$$

The form of the exact T_s^{nad} functional is not known, which is one of the central problems of sDFT, and this issue will be addressed later. Nevertheless, the total energy of the system with the sDFT approach, E_{sDFT}, can be formulated as

$$E_{\mathrm{sDFT}}[\rho] = T_s[\rho^{\mathrm{A}}] + T_s[\rho^{\mathrm{B}}] + T_s^{\mathrm{nad}}[\rho^{\mathrm{A}}, \rho^{\mathrm{B}}] + V_{\mathrm{nuc}}[\rho] + J[\rho] + E_{\mathrm{xc}}[\rho] + V_{\mathrm{nn}}. \quad (1.3.33)$$

One can search for subsystem density, ρ^{A}, by minimizing the energy expression of Eq. 1.3.33 with the following constrains:

$$\langle \psi_i^{\mathrm{A}} | \psi_j^{\mathrm{A}} \rangle = \delta_{ij} \qquad \frac{\partial E[\rho]}{\partial \rho^{\mathrm{B}}} = 0, \quad (1.3.34)$$

which means that the orbitals within subsystem A, $\{\psi_i^{\mathrm{A}}\}$, are orthonormal, and the density of subsystem B, ρ^{B}, is fixed. This leads to subsystem equations similar to Eq. 1.3.9, but with constrained densities [51]:

$$\left(\hat{F}^{\mathrm{A}}[\rho^{\mathrm{A}}] + \hat{F}^{\mathrm{embed}}[\rho^{\mathrm{A}}, \rho^{\mathrm{B}}]\right) |\psi_i^{\mathrm{A}}\rangle = \epsilon_i^{\mathrm{A}} |\psi_i^{\mathrm{A}}|\rangle, \quad (1.3.35)$$

where $\hat{F}^{\mathrm{A}}$ is the Kohn–Sham operator of subsystem A, and $\hat{F}^{\mathrm{embed}}$ is the embedding potential of subsystem B. The former is very similar to the operator of an isolated system:

$$\hat{F}^{\mathrm{A}}[\rho^{\mathrm{A}}] = -\frac{1}{2}\hat{\nabla}^2 + \hat{V}_{\mathrm{eff}}^{\mathrm{A}}[\rho^{\mathrm{A}}], \quad (1.3.36)$$

$$\hat{V}_{\mathrm{eff}}^{\mathrm{A}}[\rho^{\mathrm{A}}] = \hat{V}_{\mathrm{nuc}}^{\mathrm{A}} + \hat{J}[\rho^{\mathrm{A}}] + \hat{V}_{\mathrm{xc}}[\rho^{\mathrm{A}}] \quad (1.3.37)$$

except that the nuclear potential operator is restricted to those atoms which are assigned to subsystem A:

$$\hat{V}_{\mathrm{nuc}}^{\mathrm{A}} = -\sum_{\alpha \in \mathrm{A}} \frac{Z_\alpha}{|\boldsymbol{R}_\alpha - \boldsymbol{r}|}. \quad (1.3.38)$$

The embedding potential incorporates all potentials that arise from the interactions with subsystem B:

$$\hat{F}^{\mathrm{embed}}[\rho^{\mathrm{A}}, \rho^{\mathrm{B}}] = \hat{V}^{\mathrm{B}}_{\mathrm{nuc}} + \hat{J}[\rho^{\mathrm{B}}] + \hat{V}^{\mathrm{nad}}_{\mathrm{xc}}[\rho^{\mathrm{A}}, \rho^{\mathrm{B}}] + \hat{V}^{\mathrm{nad}}_{\mathrm{kin}}[\rho^{\mathrm{A}}, \rho^{\mathrm{B}}], \tag{1.3.39}$$

where $\hat{V}^{\mathrm{nad}}_{\mathrm{xc}}$ and $\hat{V}^{\mathrm{nad}}_{\mathrm{kin}}$ is the nonadditive exchange-correlation and kinetic potentials, respectively:

$$\hat{V}^{\mathrm{nad}}_{\mathrm{xc}}[\rho^{\mathrm{A}}, \rho^{\mathrm{B}}] = \frac{\partial E_{\mathrm{xc}}[\rho^{\mathrm{A}}, \rho^{\mathrm{B}}]}{\partial \rho^{\mathrm{A}}}, \tag{1.3.40}$$

$$\hat{V}^{\mathrm{nad}}_{\mathrm{kin}}[\rho^{\mathrm{A}}, \rho^{\mathrm{B}}] = \frac{\partial T^{\mathrm{nad}}_{s}[\rho^{\mathrm{A}}, \rho^{\mathrm{B}}]}{\partial \rho^{\mathrm{A}}}. \tag{1.3.41}$$

Note that similar equations can be derived for subsystem B, and the above equations can be generalized to multiple subsystems.

It can be seen from the derived equations that sDFT is a formally exact theory because it becomes equivalent with the conventional KS DFT method in the limit of exact T^{nad}_{s}. The subsystem approach is also suitable for archiving linear-scaling behavior if an appropriate integral screening scheme is used. This also means that the eigenvalue-problem of subsystems is separable, which can be exploited by parallel computing technologies. Besides the computational cost considerations, sDFT allows for the easier interpretation of the calculated results because the kernel of the interacting subsystems is closer to the common chemical sense. On the other hand, issues regarding the traditional KS DFT are inherited by sDFT (the approximation of E_{xc}), and, additionally, problem arises due to the unknown form of the nonadditive kinetic potential and energy. In practice, decomposable, nondecomposable, and reconstructing (see Ref. 47 and references therein) schemes are used to approximate the potential, however, most of these approaches are not sufficiently accurate for arbitrary system partitionings, which results in subsystem densities that are not compatible with the total density, and consequently, sDFT is not applicable if the subsystems are separated across covalent bonds.

Although sDFT treats the whole system on the same basis, it should be emphasized that further approximations can be introduced to the method in the spirit of the focused models. The frozen density embedding (FDE) theory [51] can be considered as a constrained sDFT, where the sDFT equations are used, but the densities of particular subsystems are kept fixed. In FDE, the iterative solution of subsystem eigenvalue equations are only required for the region of interest, and the density of the environment can be obtained from a starting approximation, e.g., calculations for isolated molecules or experiments. Besides the lack of the polarization between subsystems, such approximation often results in negative electron densities [52], which is treated with the "freeze-and-thaw" cycles protocol [53], a procedure which switches the role of the frozen and active subsystems. In the limit of fully variational FDE, the energy of sDFT can be recovered,

while it should also be noted that the freeze-and-thaw procedure is usually more costly than the supermolecular KS DFT approach.

1.3.1.3 WFT-in-DFT and DFT-in-DFT approaches

The framework of sDFT can be easily extended to the focused model approach [54]. The original non self-consistent FDE method is a simple example to extend sDFT, moreover, high-level DFT or WFT schemes can also be applied to the most relevant subsystems, while using the sDFT kernel to the other subsystems and the subsystem interactions. Invoking the subtractive composite scheme, a simple embedding approach can be proposed for our simplified closed-shell system with two regions, where we are interested in the high-level description of subsystem A. In the case of WFT-in-DFT, the energy of the embedding technique can be written as follows:

$$E_{\text{WFT-in-DFT}}[\{\psi^{\text{A}}\},\rho^{\text{A}},\rho^{\text{B}}] = E_{\text{sDFT}}[\rho^{\text{A}},\rho^{\text{B}}] - E^{\text{A}}_{\text{sDFT}}[\rho^{\text{A}}] + E^{\text{A}}_{\text{WFT}}[\{\psi^{\text{A}}_i\}], \tag{1.3.42}$$

where $E^{\text{A}}_{\text{sDFT}}$ is the sDFT energy of subsystem A:

$$E^{\text{A}}_{\text{sDFT}} = T_s[\rho^{\text{A}}] + V^{\text{A}}_{\text{nuc}}[\rho^{\text{A}}] + J[\rho^{\text{A}}] + E_{\text{xc}}[\rho^{\text{A}}], \tag{1.3.43}$$

and $E^{\text{A}}_{\text{WFT}}$ is the energy of subsystem A obtained with the WFT-based correlation method. The subsystem densities are determined by solving the subsystem equations iteratively, and the HF orbitals of the active subsystem, $\{\psi^{\text{A}}_i\}$, can be obtained by minimizing Eq. 1.3.42 with fixed ρ^{B}:

$$\left(\hat{F}^{\text{A}}_{\text{WFT}}[\rho^{\text{A}}] + \hat{F}_{\text{embed}}[\rho^{\text{A}},\rho^{\text{B}}]\right)|\psi^{\text{A}}_i\rangle = \epsilon^{\text{A}}|\psi^{\text{A}}_i\rangle, \tag{1.3.44}$$

where the $\hat{F}^{\text{A}}_{\text{WFT}}$ is the Fock operator. Note that similar equations can be derived for the DFT-in-DFT case. The scope of such embedding schemes, again, is dictated by the accuracy of sDFT, which is rather limited in the case of covalently bounded systems due to the inaccuracy of T_s^{nad} [55]. Therefore, typically adsorption of molecules on surfaces or excited states of solutes are studied [56, 57].

Reconstruction schemes using the optimized effective potential (OEP) method for the nonadditive kinetic potential [58] allows sDFT to treat covalently bounded subsystems, however, Manby and Miller realized the importance of separating the system at the molecular orbital level [59], thus the complicated and unstable OEP procedure can be bypassed. They proposed an embedding approach which avoids the approximation of T_s^{nad} by keeping the orthogonality of subsystem orbitals, that is, $\langle\psi^{\text{A}}_i|\psi^{\text{B}}_j\rangle = \delta_{ij}$. Their embedding technique, which is capable of embedding WFT or DFT methods into a low-level DFT technique, will be discussed in detail in Chapter 3, where our improvements to the method are also presented.

1.3.2 Embedding into wave function theory

In this subsection, a brief introduction of linear-scaling correlation approaches is presented, which will serve as the starting point to our multilevel WFT scheme, furthermore, the existing WFT-in-WFT methods are also overviewed. The reader should note that the term embedding in the context of WFT-in-WFT schemes is often used loosely in the literature which will be avoided in the following texts; the term multilevel approach is an adequate expression. However, it is still important to emphasize that the WFT-in-WFT methods are most accurate ones in the hierarchy of focused models.

1.3.2.1 Linear-scaling coupled-cluster approach

CCSD(T) is considered as the gold standard of quantum chemistry methods because it recovers more than 99% of the correlation energy. However, its N^7 scaling behavior and slow basis set convergence limit the use of the approach: the tractable system size is typically 20-30 atoms with a moderate, triple-ζ basis set. Local correlation methods, exploiting the relatively rapid decay of the electron correlation with the distance, have become particularly effective, affordable, and still accurate tools to study extended systems. Our localized MO (LMO) based approach combines the developments of both main research directions of the field of linear-scaling CC: the fragmentation-based schemes [60] and the "direct methods" [61]. The essential of the former is that the system is separated into smaller, treatable fragments, and the final correlation energy is obtained as the sum of the fragment (and inter-fragment) results, while the latter techniques employ cost-reducing approximations for the CC equations of the whole system.

The total energy of a closed-shell system can be written as a sum of the HF (E_{HF}) and the correlation energy (E_{corr}):

$$E = E_{\mathrm{HF}} + E_{\mathrm{corr}}. \tag{1.3.45}$$

Our local CCSD(T) [LCCSD(T)] calculations starts with the iterative solvation of the HF equations, which can be written in the AO basis as

$$\mathbf{F}[\mathbf{D}]\mathbf{C} = \mathbf{SCE}, \tag{1.3.46}$$

where $\mathbf{F}$ is the Fock-matrix:

$$\mathbf{F}[\mathbf{D}] = \mathbf{h} + \mathbf{J}[\mathbf{D}] - \frac{1}{2}\mathbf{K}[\mathbf{D}], \tag{1.3.47}$$

and the final E_{HF} can be expressed as:

$$E_{\mathrm{HF}}[\mathbf{D}] = \mathrm{Tr}\left[\mathbf{hD} + \frac{1}{2}\mathbf{D}\left(\mathbf{J}[\mathbf{D}] - \frac{1}{2}\mathbf{K}[\mathbf{D}]\right)\right]. \tag{1.3.48}$$

The procedure continues with orbital localization, which rotates the occupied orbitals with a unitary matrix:

$$L_{\mu i} = \sum_j C_{\mu j} W_{ji} \tag{1.3.49}$$

where $L_{\mu i}$ and W_{ji} are elements of the coefficient matrix of the localized orbitals ($\mathbf{L}$) and the unitary matrix ($\mathbf{W}$), respectively, and j runs over occupied orbitals. There are several ways to obtain a unitary matrix, including the Boys [62] and Pipek–Mezey algorithms [63], and it should be noted that the density matrix (and therefore E_{HF}) is unaffected by unitary rotations as $\mathbf{W}(\mathbf{W})^{\mathrm{T}} = \mathbf{1}$. On the contrary of the naive interpretation of the term "linear-scaling", in the case of the LCCSD(T) approaches based on the conventional HF, linear-scaling is only achieved for the evaluation of E_{corr}, and not for the previously discussed two initial steps of the whole method. The evaluation of the HF energy and the orbital localization conventionally scales at least as N^3. The scaling can be reduced to linear [64], yet cost reduction often only appears in the case of unmanageable large systems.

In our approach [48], the linear-scaling behavior of the evaluation of E_{corr} is achieved by utilizing direct techniques and the fragmentation-based clusters-in-molecules (CIM) ansatz of Li [65]. In the CIM approach, the correlation energy is evaluated as the sum of the contributions:

$$E^{\mathrm{CCSD(T)}} = \sum_i \left(\delta E_i^{\mathrm{CCSD}} + \delta E_i^{(\mathrm{T})} \right) = \sum_i \delta E_i^{\mathrm{CCSD(T)}}, \tag{1.3.50}$$

where E_i^{CCSD} and $E_i^{(\mathrm{T})}$ are the CCSD and (T) energy contributions of LMO i, respectively, and the above equation is a formally exact reformulation of the conventional CCSD(T). To reduce the scaling of Eq. 1.3.50, pair and domain approximations are invoked. First, primary domains (PD) are constructed for each LMO i, denoted as $\mathcal{P}_i$. After the virtual space of the whole system is built using projected atomic orbitals (PAOs) [66], the closest PAOs are selected by utilizing the Boughton–Pulay (BP) algorithm [67] (for the detailed description of the BP algorithm, see Sect. 2.2.2.1) to construct a local virtual subspace for each LMO i in its primary domain. After the PAOs are projected onto the AOs of the PD and orthogonalized, the approximate long range correlation energy is calculated for each i-j pair domain, $\mathcal{P}_{ij} = \mathcal{P}_i \cup \mathcal{P}_j$, by evaluating the opposite-spin (OS) MP2 energy [68] (δE_{ij}) within the pair-domain:

$$\delta E_{ij}(\mathcal{P}_{ij}) = -8 \sum_{\tilde{a}_i} \sum_{\tilde{b}_j} \frac{\left[\langle \tilde{a}_i \tilde{b}_j | ij \rangle^{[4]} \right]^2}{\epsilon_{\tilde{a}_i} + \epsilon_{\tilde{b}_j} - \mathrm{F}_{ii} - \mathrm{F}_{jj}}, \tag{1.3.51}$$

where F_{ii} and F_{jj} are Fock-matrix elements calculated with the corresponding LMOs, $\tilde{a}_i$ ($\tilde{b}_j$) and $\epsilon_{\tilde{a}_i}$ ($\epsilon_{\tilde{b}_j}$) are canonicalized PAOs and their orbital energies, respectively, while

the integrals in the numerator of the above equation is a two-electron integral in the canonical basis of the PDs, which is approximated with a multipole expansion up to the fourth order [69]. The calculation of pair correlation energies is also valuable for selecting weakly and strongly interacting pairs. Those pairs are considered as strong (distant) pairs which have higher (lower) pair energy than a given threshold. The rest of the correlation calculation is performed on the linear-scaling number of strong pairs, and linear-scaling is achieved by computing the correlation contribution of each occupied LMOs in its extended domain (ED) separately.

The EDs (denoted as $\mathcal{E}_i$) are constructed for each LMO i by merging the PDs of the LMOs forming a strong pair with LMO i. Then, these LMOs are projected onto the AOs of the ED and the orthogonality is recovered with the Gram–Schmidt–Löwdin [69, 70] procedure. Again, the virtual subspace is constructed using specially selected PAOs, and the MP2 energy in the ED is evaluated as

$$\delta E_i^{\mathrm{MP2}}(\mathcal{E}_i) = \sum_{\tilde{a}\tilde{b}\tilde{j}} \left[2\langle \tilde{a}\tilde{b}|i\tilde{j}\rangle - \langle \tilde{a}\tilde{b}|\tilde{j}i\rangle t_{i\tilde{j}}^{\tilde{a}\tilde{b}[1]} \right], \tag{1.3.52}$$

where $\delta E_i^{\mathrm{MP2}}$, $\tilde{j}$, and $\tilde{a}$ are the MP2 energy, the strong pair LMOs of LMO i, and the PAOs of domain i, respectively, while $t_{i\tilde{j}}^{\tilde{a}\tilde{b}[1]}$ is the MP2 amplitude. Note that utilizing the MP2 domain energy of Eq. 1.3.52 together with the pair energies of Eq. 1.3.51, one can also formulate the local MP2 (LMP2) energy at this stage:

$$E^{\mathrm{LMP2}} = \sum_i \left[\delta E_i^{\mathrm{MP2}}(\mathcal{E}_i) + \frac{1}{2} \sum_j^{\mathrm{distant}} \delta E_{ij}(\mathcal{P}_{ij}) \right], \tag{1.3.53}$$

where the summation over the pair energies is constrained to distant pairs to prevent double counting. As the MP2 amplitudes are already available, these are used to construct MP2 densities, and the corresponding MP2 natural orbitals can be obtained by diagonalizing the density matrix, which results in the local natural orbitals (LNOs) of the extended domains. With proper truncation and selection of the occupied and virtual LNOs, a local interacting subspace (LIS) can be defined (denoted as P_i), which is the subspace of the LNOs of the $\mathcal{E}_i$, and it is used as the basis for calculating the CCSD(T) contributions. The utilization of LNOs is a very efficient way to treat the CC problem: fewer LNOs are necessary compared with the pair-natural orbital (PNO) based schemes, and the basis set convergence is also faster [71]. The discussed LCCSD(T) ansatz can also be named LNO-CCSD(T) because the LNO basis is a specific attribute for the discussed approach. The final correlation energy ($E_{\mathrm{corr}} = E^{\mathrm{LNO\text{-}CCSD(T)}}$) can be written as

$$\begin{aligned} E^{\mathrm{LNO\text{-}CCSD(T)}} = &\sum_i \delta E_i^{\mathrm{CCSD(T)}}(P_i) + \sum_i \left[\delta E_i^{\mathrm{MP2}}(\mathcal{E}_i) - \delta E_i^{\mathrm{MP2}}(P_i) \right] \\ &+ \frac{1}{2} \sum_i \sum_j^{\mathrm{distant}} \delta E_{ij}(\mathcal{P}_{ij}), \end{aligned} \tag{1.3.54}$$

where $\delta E_i^{\mathrm{CCSD(T)}}(P_i)$ and $\delta E_i^{\mathrm{MP2}}(P_i)$ are the CCSD(T) and MP2 energy evaluated in the LIS, respectively, and the subtraction of the latter is needed to prevent double counting of the correlation energy. Note that the discussed kernel for computing the correlation energy is suitable for using arbitrary order CC in the P_i, while the infrastructure of the extended domains can also be used for the spin-scaled variants of MP2 [68], the direct random phase approximation (dRPA) [72], and its second-order screened exchange (SOSEX) extension [73] other than the standard MP2. As a result of the discussed LNO and CIM approaches together with density fitting [74], and direct integral evaluation, the applicability range of the CCSD(T) method can be widened from 20-30 to a few thousand atoms with chemical accuracy, which means that the correlation energy computation is no longer the bottleneck of the calculations in the case of large systems because of the unfavorable scaling and prefactor of the HF equations.

1.3.2.2 WFT-in-WFT embedding

Based on the previously discussed linear-scaling technique, an interaction-based multilevel approach can be developed which exploits the fact that the correlation of the relatively distant orbitals can be sufficiently described by more approximate methods. Various LCCSD(T) methods were also elaborated in other groups, and it was recognized earlier that the CCSD(T) treatment of all orbitals is not a necessity. Mata, Werner, and Schütz used their own version of linear-scaling CCSD(T) for various small molecules to demonstrate that it is adequate to apply CCSD(T) only on the small, most relevant part of the system while utilizing LMP2 or uncorrelated HF on the environment [75]. Li and Piecuch formulated a similar CC-in-LMP2 multilevel technique based on their CIM approach [76], while Rolik and Kállay studied the convergence of CCSD(T)-in-HF errors with respect to the CC subsystem size [77]. More recently, Neese and co-workers proposed CC-in-CC and CC-in-LMP2 schemes [78] invoking their PNO-based LCCSD(T) approach which uses different orbital truncation parameters for different regions and demonstrated the applicability of the method on chemically relevant systems, such as metal complexes and DNA double strands. The newest contribution to the WFT-in-WFT field is the embedding of multireference CC into the complete active space second-order perturbation theory (CASPT2) approach by Köhn *et al.* [79], who presented the high-level and efficient treatment of medium to large systems with multireference character. Overall, numerous multilevel and embedding schemes have been emerged over the years, but the literature lacks an extensive benchmark study that compares the various schemes. In Chapter 3, a comparative study will be presented for the various WFT-in-DFT and WFT-in-WFT approaches, where our WFT-in-WFT schemes will be also discussed.

Chapter 2

QM/MM Using Frozen Localized Orbitals

We started our developments in the field of focused models with the improvement of the QM/MM methodology [80]. As it was mentioned in Sec. 1.2.3.2, in the LSCF framework of Ferenczy [45] provided the use of frozen LMOs were proposed at the border of the QM and MM subsystems with minor modifications on the working equations. Therefore, our first objective was to implement this ansatz and build an easily accessible QM/MM program where the LSCF approach can be utilized and compared with the widely used link atom approach.

We selected the Amber [81, 82] molecular mechanics and the Mrcc [83] quantum chemical codes to develop a versatile and user-friendly QM/MM program which is freely available for academic purposes. The choice of Amber is motivated by its capabilities that include molecular mechanics, molecular dynamics, and techniques that allow the efficient calculation of free energies. The Mrcc code was selected for the QM region owing to its highly efficient implementation of local correlation techniques that are beneficial for calculating reaction mechanisms, a field with primary importance for QM/MM applications. It has to be noted that the link atom method and the frozen localized orbital method require significantly different approaches in their implementation into existing computer codes. The link atom method basically performs standard MM and QM calculations, and it is the generation of the subsystems and the communication between the QM and MM programs that require special attention and coding. This is the reason why several interfaces connecting various MM and QM programs [84–89] have been developed, and only small modifications in the original codes were necessary to be introduced. The situation is different for the frozen localized orbital method, where specific orbital equations are to be used, and this requires more intensive modifications in the QM code. In the forthcoming sections, we present the results of our efforts to combine the Amber and Mrcc codes.

2.1 Theory

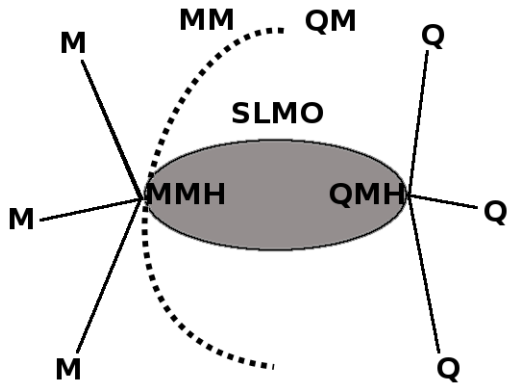

FIGURE 2.1: Separation of the QM and MM subsystems by a frozen strictly localized molecular orbital (SLMO). QM and MM atoms are designated by **Q** and **M**, respectively, and **MMH** **(QMH)** is the MM (QM) host atom.

The QM and MM regions are bound with a strictly localized molecular orbital (SLMO) (see Fig. 2.1). This is realized by defining an MM host (MMH) atom (which is also known as frontier atom in the literature) at the border of the QM and MM subsystems so that it is connected to an QM host (QMH) atom with a bond orbital whose basis functions are located exclusively on the two atoms. The QMH atom contributes to the SLMO with one electron and its other electrons are part of the optimized wave function. The MMH atom of the SLMO also contributes with a single electron to the SLMO, while its other valence electrons are not treated explicitly.

The SLMO is not optimized, rather its coefficients are kept fixed at certain predefined values. These values are determined by calculations performed for model molecules that include a chemically similar bond to the SLMO (see later). Note that the core electrons of the MMH atom are also treated explicitly and their orbitals are taken from the calculations performed for deriving the valence SLMOs.

It is convenient to use optimized orbitals that are orthogonal to the frozen SLMOs, and this is achieved by calculating the optimized orbitals with the Huzinaga-equation [90]:

$$\left[\hat{F} - \hat{P}\hat{F} - \hat{F}\hat{P}\right]|\psi^a\rangle = \epsilon^a|\psi^a\rangle, \tag{2.1.1}$$

where $\hat{P}$ is a projector defined by the frozen MOs (ψ^f) as $\hat{P} = \sum_i |\psi_i^f\rangle\langle\psi_i^f|$, while ψ^a and ϵ^a are the optimized (active) MOs and their eigenvalues, respectively. This can be rewritten in the AO basis as

$$\left[\mathbf{F} - \mathbf{S}\mathbf{P}^f\mathbf{F} - \mathbf{F}\mathbf{P}^f\mathbf{S}\right]\mathbf{C}^{\mathrm{a}} = \mathbf{S}\mathbf{C}^{\mathrm{a}}\mathbf{E}^{\mathrm{a}}, \tag{2.1.2}$$

where $\mathbf{C}^a$ includes the coefficients of the orbitals to be optimized, $\mathbf{E}^a$ is the diagonal matrix of the corresponding eigenvalues, and $\mathbf{P}^f$ projects onto the space of the frozen orbitals. Assuming orthonormal frozen orbitals $\mathbf{P}^f = \mathbf{C}^f(\mathbf{C}^f)^{\mathrm{T}}$, where $\mathbf{C}^f$ contains the coefficients of the frozen orbitals. The role of the last two terms on the left hand side of Eq. 2.1.2 is to shift the eigenvalues of the frozen orbitals to positive values and to guarantee that the lowest-eigenvalue orbitals, those that are used to construct the Fock-matrix

of the next iteration cycle, are orthogonal to the space spanned by the frozen orbitals. Orthogonality between $\mathbf{C}^a$ and $\mathbf{C}^f$ is automatically achieved with the use of the Huzinaga-operator, $\mathbf{F}-\mathbf{S}\mathbf{P}^f\mathbf{F}-\mathbf{F}\mathbf{P}^f\mathbf{S}$, because it is Hermitian, and commutes with $\mathbf{P}^f$, thus the Huzinaga-operator and $\mathbf{P}^f$ share common eigenfunctions (irrespective of the choice of the latter). Three notes are appropriate here. First, the frozen orbitals are expected to appear with positive eigenvalues in Eq. 2.1.2 if they are reasonable approximations to the exact eigenfunctions of $\mathbf{F}$. Second, the basis functions of the optimized orbitals have to include those of the SLMOs since this ensures that orthogonality between the optimized orbitals and the SLMOs could be achieved. Third, the subsystems can be connected by several SLMOs that are typically not orthogonal and then the projector in Eq. 2.1.2 takes the form of $\mathbf{P}^f=\mathbf{C}^f(\sigma^f)^{-1}(\mathbf{C}^f)^{\mathrm{T}}$, where σ^f is the overlap matrix of the SLMOs.

It was found that the calculation of the optimized orbitals by Eq. 2.1.2 is advantageous when the other valence electrons of the MMH atom are represented by point charges placed on the bonds connecting the MMH atom with MM atoms. In case of an sp^3 carbon MMH atom three negative charges, called bond charges, are placed on the three bonds connecting the MMH atom with three MM atoms, while the fourth valence electron is involved in the SLMO. The magnitude of the bond charges (q_{bond}) are determined so that

$$\sum q_{bond} + q_{core}^{MMH} + (-3) = Q_{MM}^{MMH}, \tag{2.1.3}$$

where q_{core}^{MMH} is the MMH atom core charge, -3 accounts for the two core electrons and the single valence electron included in the SLMO, and Q_{MM}^{MMH} is the MM charge of the MMH atom. The magnitude of q_{core}^{MMH} is an adjustable parameter that determines q_{bond} via Eq. 2.1.3. Moreover, the position of the bond charges is, in principle, an additional parameter. These charges are placed at the midpoint of the bond formed by the MMH and the MM atom adjacent to the boundary to provide a sensible representation of the MMH valence electrons which are not accounted for quantum mechanically.

The MM parameters at the boundary, including bonded and nonbonded terms, are used in accordance with the AMBER force field [23], similarly as they are implemented in AMBER for the link atom approach. Detailed description of the retained interactions is discussed in the following section.

2.2 Implementation

In this section, specific features of the QM/MM code obtained by interfacing the MM program Amber and the QM program Mrcc are presented. The coupling of the two codes is mainly based on the interface of Walker and co-workers [88], therefore relevant parts of their work will be reviewed with special emphasis on the similarities and differences of the link atom and the frozen SLMO approaches. In the Amber-Mrcc interface,

as in most implementations of MM and QM program couplings [89], the necessary data exchange is realized by writing and reading of files. All tasks, except the solution of the QM orbital equations, are driven by SANDER, the molecular dynamics engine program of AMBER.

In order to perform a QM/MM calculation, the first step for the user is to prepare a parameter and topology file, a coordinate file and a control file which are required for regular molecular modeling calculations. This procedure may not be straightforward in the case of QM/MM runs, because the QM subsystem may not be represented by classical chemical structures and corresponding force field parameters. Nevertheless, the construction of the topology file facilitates the QM/MM run since it allows the use of the MM program infrastructure with minimal modifications, moreover, it may support system preparations, for example, MM energy minimization or equilibration. It is worth also noting that, in the case of electronic embedding, the only consequence of the topology definition to the QM subsystem is the assignment of vdW parameters. Owing to the short range nature of the vdW interactions an appropriate choice of the QM subsystem can minimize the effect of potentially inconsistent parameters.

Changes in the control file with respect to a standard MM setup consist of marking the QM atoms, specifying the charge and multiplicity of the QM region as well as defining the theory and basis set. For compatibility reasons, only the most general MRCC keywords can be denominated in the control file, however, the user can also set up the full range of options by referring in the control file to an MRCC template file. Having done these arrangements, all other procedures required for a simulation are automatically executed by SANDER for the link atom (LA) approach and for the Huzinaga-equation-based local self-consistent field (HLSCF) method, as well.

2.2.1 Link atom approach

In the case of the link atom approach, the procedure for an AMBER-MRCC calculation closely follows that formerly implemented in SANDER, namely it automatically determines the place where to cut bonds and also positions the hydrogen atoms as a function of the QMH and MMH positions to saturate the dangling bonds of the QM subsystem. It is important to note that link atoms connected to a common host atom are disadvantageous owing to their short spatial separation. The latter criterion is checked by the program, hence the user cannot mark arbitrary QM subsystems. We note that the above procedures both are also applied at the automatic selection of model molecules for the HLSCF method (see later). After the link atoms are defined, bonded terms in the QM region are removed, and, in the case of electronic embedding, the point charges are zeroed on the QM atoms as well as, to avoid overpolarization, on the MMH atoms. In order to

preserve the total charge of the system, residual charges are distributed among the nonzero charges of the MM region, however, there are several schemes for both charge distribution and bonded term removal which can be specified by the user. If the MM energy terms are defined, an input file will be written for MRCC by the AMBER-MRCC interface in every simulation step, which consists of QM atom coordinates, atomic numbers, MM point charges and their coordinates as well as the controlling keywords. MRCC is then executed via a system call and after a successful run, the energy of the QM region and forces acting on atoms are written into an output file which is parsed by the interface. Here, we note that conventional LA calculations are performed during the preparation of a HLSCF run to generate SLMOs in model systems.

2.2.2 Huzinaga-equation-based local self-consistent field approach

In the HLSCF method, just as in the LA approach, the communication between AMBER and MRCC is carried out by writing and reading files. The major difference between the two boundary methods as described earlier is that in the case of HLSCF, the QM and MM regions are connected by a frozen SLMO. The coefficients of a frozen SLMO are taken from a molecule containing a QMH-MMH bond preferably in a chemical environment similar to that in the system studied. Therefore, the first step in order to determine the SLMO coefficients of a QMH-MMH bond is to choose a suitable model system. Model systems are generated automatically in our implementation using selection rules (see Sect. 2.2.2.2), but the user can also select atoms for the model systems manually using the previously mentioned template file. After the model systems are set, the SLMOs are automatically calculated without any user intervention (see Sect. 2.2.2.1) for every QMH-MMH bond, and both the outputs of model calculations and the SLMO coefficients are saved in files. Following this procedure, bonded terms, which are defined only by QM atoms are removed, thus the following MM bonds, angles, and dihedrals are retained:
Bonds = MM-MM, MM-MMH, MMH-QMH
Angles = MM-MM-MM, MM-MM-MMH, MM-MMH-QMH, MMH-QMH-QM
Dihedrals = MM-MM-MM-MM, MM-MM-MM-MMH, MM-MM-MMH-QMH, MM-MMH-QMH-QM, MMH-QMH-QM-QM,
where MM and QM are MM and QM atoms, respectively, while QMH is a QM host atom and MMH is an MM host atom. Note that the MMH atom, which is treated specially in the electrostatic terms is handled as any other MM atom in the bonded terms.

The vdW and electrostatic contributions of the nonbonded terms are calculated as in the LA approach, however, only those electrostatic point charges are zeroed which are defined on QM atoms, furthermore the charge of the full system is preserved by distributing the residual charges over all remaining point charges. Following this procedure, the

charge of the MMH atom is redistributed according to Eq. 2.1.3: the MMH point charge is replaced by the q_{core}^{MMH} core charge of the MMH atom, and the q_{bond} charges are automatically positioned at the midpoint of the bonds formed between the MMH and the MM atoms adjacent to the MMH atom. Interactions of the q_{core}^{MMH} and q_{bond} charges with MM atoms are calculated as in the AMBER force field [23]: q_{core}^{MMH} and q_{bond} interact neither with MM charges adjacent to MMH nor MM atoms adjacent to these atoms (1-2, 1-3 interactions), furthermore interactions with three bonds away from MMH are divided by a factor of 1.2. It is important to mention that the interactions between q_{core}^{MMH} and q_{bond}, just as any interaction including these charges are computed only by direct summation, hence periodic boundary conditions cannot be handled at this stage of development.

When the setup of the MM energy terms is completed, similarly to the LA approach, an input file is written for Mrcc by the interface, and Mrcc is executed. In the case of HLSCF host atoms are flagged in the input file. The aim of these flags is to assign the previously calculated SLMO coefficients to the proper host atoms and to build the $\mathbf{C}^f$ matrix from the strictly localized molecular orbitals. Using these orbitals the overlap matrix of the frozen molecular orbitals are evaluated as $\sigma = (\mathbf{C}^f)^{\mathrm{T}}\mathbf{S}\mathbf{C}^f$. The inverse square root of the orbital overlap matrix is used to calculate the projector $\mathbf{P}^f = \mathbf{C}^f(\sigma^f)^{-1}(\mathbf{C}^f)^{\mathrm{T}}$ and then in each iteration cycle the Fock-matrix and the Huzinaga-matrix $\mathbf{F} - \mathbf{S}\mathbf{P}^f\mathbf{F} - \mathbf{F}\mathbf{P}^f\mathbf{S}$ are built. The diagonalization of the Huzinaga-matrix according to Eq. 2.1.2 produces orbitals orthogonal to the frozen orbitals. When self-consistency is reached, the calculated SCF energy is written in an output file which is parsed by the interface. As to the force evaluation, note that the gradients derived in Ref. 45 for the QM atoms are implemented and tested in Mrcc, however, the QM/MM gradients are not available at this stage of development.

2.2.2.1 Automatic generation of frozen strictly localized orbitals

The coefficients of the frozen SLMOs can be taken from model molecules that contain a chemical bond similar to that of the SLMO. The basic idea for the generation of the orbital coefficients is that after identifying a bond connecting the QM and MM subsystems a molecular graph containing this bond is cut from the system, and a QM/MM calculation is performed for the model molecule generated from the molecular graph using the LA approach. The wave function of the model molecule is calculated by taking into account the electrostatic potential of the point charges of the whole system (all atoms except those included in the model molecule). Note that during the calculation of the model molecule bonded and vdW contributions are not calculated, which is rationalized by the fact that the only aim is to determine the polarized wave function of the model molecule, nevertheless charge neutrality is still assured by various charge distribution schemes set by the user. The orbitals obtained are Boys [62] or Pipek–Mezey [63] localized, which is

followed by the execution of the BP algorithm [67] to select those atoms on which the orbitals are localized.

The first step of the BP algorithm is that, for each localized orbital, the atoms are ranked according to their gross Mulliken population. Starting from the atom with the largest Mulliken population subspaces of atoms are defined by adding atoms of smaller and smaller Mulliken populations. For each subspace, the LMO is truncated to the subspace, and the norm of the square of the difference of the original and the truncated function is calculated. The procedure is terminated when the norm drops below a threshold ϵ, hence the localized orbitals are assigned automatically to atoms whose basis functions span the best subspace for the given localized orbital. In the implementation of the HLSCF method, a completeness criterion of $1 - \epsilon = 0.985$ was chosen.

After the assignation of atoms to localized orbitals the core orbitals of the MMH atom and the single valence orbital connecting the QMH and MMH atoms are selected. SLMOs are generated from selected LMOs by removing contributions from all atoms except the QMH and MMH atoms.

2.2.2.2 Selection of the model molecule

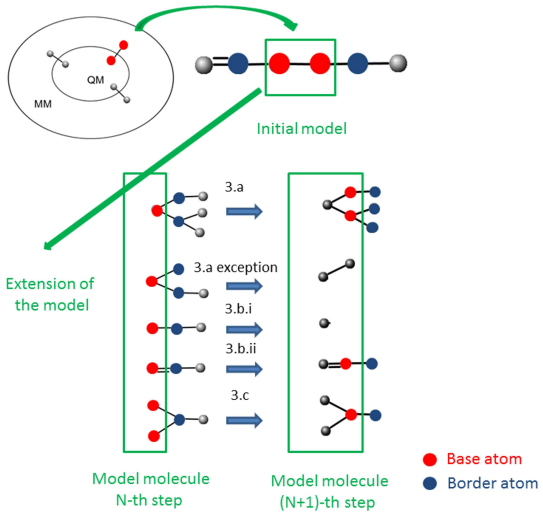

FIGURE 2.2: Schematic representation of the incremental procedure used for the step by step extension of the model system. The rules applied in the N-th step are also identified. See the text for the details.

The model molecules are generated from the list of atom pairs defining those chemical bonds that connect the QM and MM subsystems. Our task is to select a molecular graph for every such atom pair and close their dangling bonds with hydrogens in a way to satisfy the considerations written in the Sect. 2.2.1. Starting from a given atom pair we use an incremental, step by step forwarding procedure to extend the initial model (graph) in both directions (see Fig. 2.2). In this procedure only the non-H atoms are used, the H atoms are added to the model after the selection of heavy atoms is completed. For the easier formulation of the rules of extension, we introduce the following naming convention. At a certain step of the incremental procedure the *base atoms* of the model molecule are those atoms that are connected to other atoms (*border atoms*) not included in the model. The basic idea of the model molecule selection is to find the

smallest molecular subgraph every base atom of which is connected to the remaining part of the system by only one single bond. In the currently implemented algorithm only single bonds can be cut to define the model molecule (*breakable bonds*), all the other bond types are *non-breakable*. However, less restrictive definitions of the breakable bonds can also be applied.

The main steps of the algorithm are demonstrated in Fig. 2.2 and can be summarized as follows.

1. Repeat the model molecule search below for every atom pair defining single bonds between the QM and MM subsystems.
2. Find the base atoms and the border atoms connected to them.
3. Extend the model molecule using the following rules:
   (a) If more than one border atom is connected to a base atom, then all of them are attached to the model molecule, unless all but one base atom-border atom connections are terminal (the border atoms have no further bonds to heavy atoms not included in the model system). In this case only the terminal atoms are added to the model system and the remaining one bond is cut and it is not used in the subsequent base atom-border atom search.
   (b) If only one border atom is connected to a base atom and
      i. this bond is a breakable one, then terminate this extension path
      ii. this bond is a non-breakable one, then add its border atom to the model system.
   (c) If a border atom is connected to two different base atoms (crossing), then the common border atom is also attached to the model system.
4. If every base atom has at the most one non-terminal bond to the border atoms and this bond is breakable, furthermore no crossings occur, then the model molecule is defined, otherwise go to step 2.
5. Add H-atoms to terminate dangling bonds.

The described protocol was applied to generate model molecules in all the calculations presented in the current contribution.

2.3 Sample calculations

In this section results obtained both with the widely used LA approach and with the newly implemented HLSCF method are presented. Both QM/MM boundary methods were tested to model typical applications of hybrid methods. However, the systems selected are sufficiently small so that reference full QM calculations are feasible.

In all the calculations presented below, the wave function of the QM region was calculated with the HF equation for the LA approach and with the Huzinaga-equation (with the HF Fockian) for the HLSCF approach. Pople-type Cartesian 6-31G* basis set [91] was used. The parameters defined in the ff14SB [23] and in the GAFF [92] force fields were used for the peptide and the hexanoic acid calculations, respectively, to evaluate the MM energy function. The MM parameters of the hexanoic acid were obtained by the ANTECHAMBER program [93] using the AM1-BCC charge scheme [94, 95]. Input files were generated by using the Maestro [96] program and the TELEAP program which is part of the AMBERTOOLS package. Figures were prepared using Gnuplot [97] and Marvin [98].

As a first step, the full system was optimized at the HF/6-31G* level in all test cases, then the obtained geometries were modified to perform single point calculations along the defined reaction coordinates. The tests were calculated in vacuum, without periodic boundary conditions, and the electrostatic and vdW terms were not cut off. In the case of the LA approach, the capping hydrogen atom was positioned 1.09 Å away from the QMH atom, furthermore the residual charges which originated from the zeroing of the charge of the QM and MMH atoms were distributed among all MM atoms. Those bonded terms which contained at least one MM or MMH atoms were retained. The SLMOs of the HLSCF method were only calculated for a single configuration and these SLMO coefficients were used for all other configurations. This procedure was found to better reproduce reference results than the recalculation of the SLMOs for each individual configuration. We selected the configuration of the global potential minima for calculating the SLMO coefficients. In this way the SLMO coefficients are treated on an equal footing with the MM parameters in the sense that they both provide a constant environment for the rest of the QM subsystem. The SLMOs were generated with automatically selected model systems, using Boys localization and with the setup described for the LA approach.

In all examples presented below the effect of the QM subsystem size is investigated by varying the subsystem boundary. The boundary is defined by the bond cut at the subsystem separation. Note that “cut” in the LA approach indeed means bond cutting between the QM and MM subsystems and saturating the dangling bond of the QM subsystem by a H-atom. On the other hand, “cut” signifies a frozen SLMO in the HLSCF approach.

2.3.1 Deprotonation energy

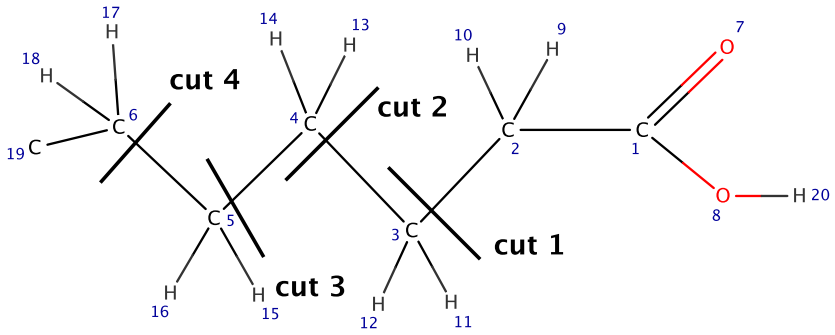

FIGURE 2.3: System separation of the hexanoic acid.

The calculation of the deprotonation energy is a highly sensitive test for the HLSCF method since it requires the evaluation of the energy difference of differently charged systems. Therefore, the deprotonation energy of hexanoic acid was calculated with both the LA and the HLSCF approaches as a first test. Results were compared with the reference full QM calculations. As described above, the geometry of the natural hexanoic acid was first optimized using the reference method, however, the geometry of the deprotonated acid was obtained by simply removing the H_{20} hydrogen atom of the carboxyl group from the neutral form (see Fig. 2.3 for the numbering of atoms). Besides the same geometry of the two species, the MM subsystem and the SLMOs of the neutral hexanoic acid were also used for the deprotonated form to prevent the charge transfer between the QM and the MM subsystems.

TABLE 2.1: QM/MM errors of deprotonation energies for the hexanoic acid molecule with various system separations (see Fig. 2.3).

	LA	HLSCF
System	kcal/mol	kcal/mol
cut1	-1.65	-6.18
cut2	-0.62	-1.51
cut3	-0.19	-1.48
cut4	0.04	-0.02

The convergence of the QM/MM methods to the full QM calculations was investigated by gradually increasing the size of the QM subsystem. The various QM and MM subsystems are shown in Fig. 2.3. The core charge of Eq. 2.1.3 was determined in a trial and error procedure in which the absolute error belonging to the largest QM subsystem, cut4 in Fig. 2.3, was minimized. It was found that the optimal core charge is +5.5 and this parameter was also used in the subsequent calculations.

The results of the LA and the HLSCF methods were compared with the reference, and the errors are presented in Table 2.1. The data show that the error decreases as the QM subsystem increases, hence both QM/MM methods converge to the full QM calculation, moreover, the errors are less than 0.1 kcal/mol if the largest QM subsystem (cut4) is used. For smaller QM subsystems the errors of the LA approach is smaller than those of the HLSCF approach. The larger HLSCF errors stem from the presence of a large positive MMH core charge and negative bond charges that have significant interactions with the site of deprotonation. An appropriate choice of these charges allow a good reproduction of the reference deprotonation energy, but it appears that the optimal charges depend on the subsystem separation. This is not surprising since the position of the charges and their distance from the site of deprotonation depends on the subsystem separation. However, both approaches reproduce the full QM results well if the reaction center is sufficiently far away from the boundary of the QM and MM subsystems.

It is worth emphasizing that the sensitivity of the deprotonation energy on the subsystem separation in the HLSCF approach is a consequence of calculating the energy difference of differently charged system. Similarly, the magnitude of the MMH core charge and the position of the bond charges affect the calculated deprotonation energy for the very same reason. On the other hand, these parameters were found to have much less effect on typical QM/MM calculations where energies of systems with equal total charges are compared. This type of examples are presented in the forthcoming sections. In all these examples an MMH core charge of 5.5 was used, but we note that this charge can also be treated as adjustable parameter.

2.3.2 Conformational energies

Rotational energy curve calculation was performed for the Acetyl(Ace)-Histidine(His)-N-methyl(Nme) system by rotating the imidazole group. Single point energy calculations were performed varying the C_9-C_{11}-C_{14}-N_{15} dihedral angle (see Fig. 2.4 for the numbering of the atoms) by 30°. In order to test the effect of the QM subsystem size on the quality of the QM/MM results, calculations were performed with two QM-MM boundaries; a smaller (cut1) and a larger (cut2) QM region was chosen (Fig. 2.4). In the case of cut1, the boundary is at the bond between C_α (C_9) and C_β (C_{11}), while for cut2, only the methyl groups of the acetyl and N-methyl groups were included in the MM subsystem, hence the boundary is at the bonds between C_2-C_5 and N_{24}-C_{26}.

The relative energies with respect to the lowest energy conformer are shown for the LA, HLSCF, and reference full QM calculations in Fig. 2.5. The QM/MM methods well reproduce the shape of the energy profile of the reference for both system separations. Furthermore the results of both methods improve with larger QM subsystems. In the case

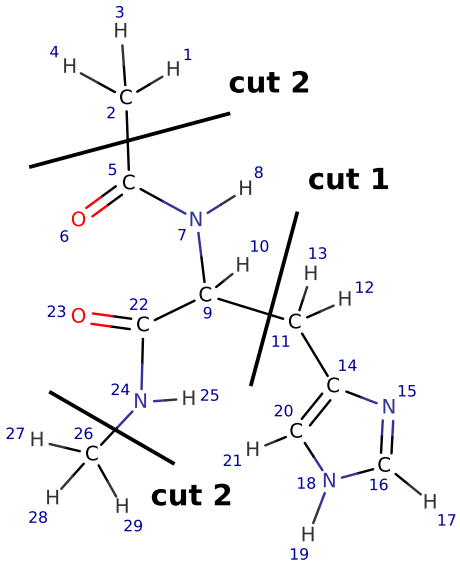

FIGURE 2.4: System separation of the Ace-His-Nme system.

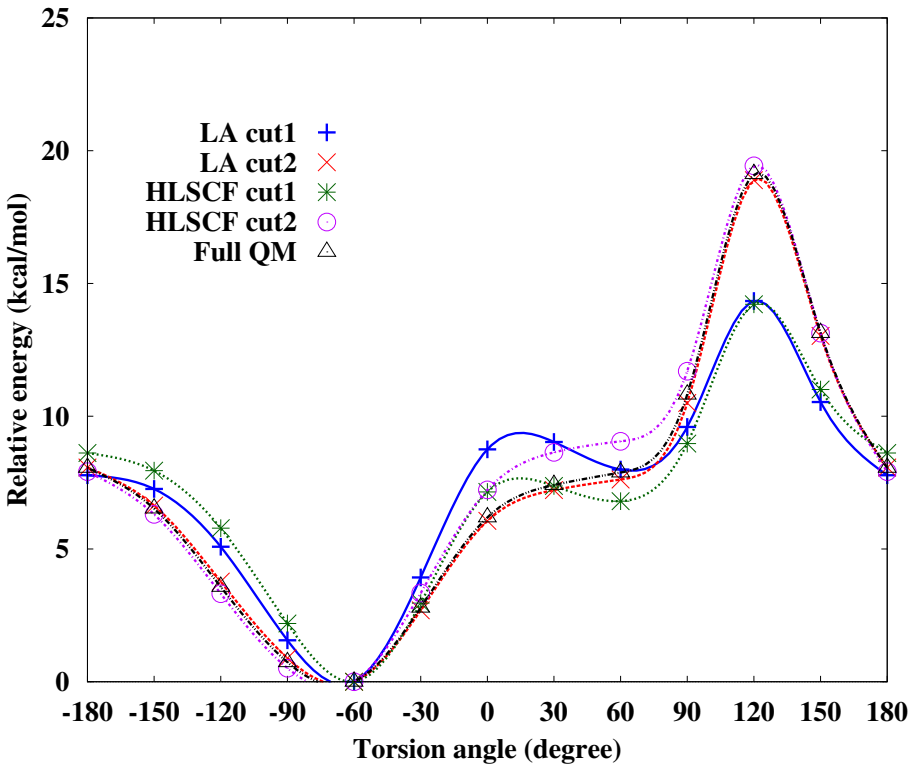

FIGURE 2.5: Energy of the Ace-His-Nme molecule as a function of the rotation angle of the imidazole group.

of the smaller QM subsystem, cut1 in Fig. 2.5, the mean absolute error and the largest error are 1.5 and 4.8 kcal/mol, respectively, for the LA method, while they are 1.5 and

5.3 kcal/mol, respectively, for the HLSCF method. Note that MM dihedral parameters contribute to the obtained energies for this small QM subsystem. Increasing the size of the QM region decreases both the mean and the maximum error; 0.2 and 0.3 kcal/mol for the LA and 1.0 and 1.5 kcal/mol for the HLSCF method are obtained. In summary, it was found that both methods are able to reproduce the reference results for the rotation of the imidazole group within 5 % relative error when the amide bonds are included in the QM subsystems, and thus the subsystem boundaries are separated from the rotating bond by 3 bonds.

2.3.3 Proton transfer energy curve

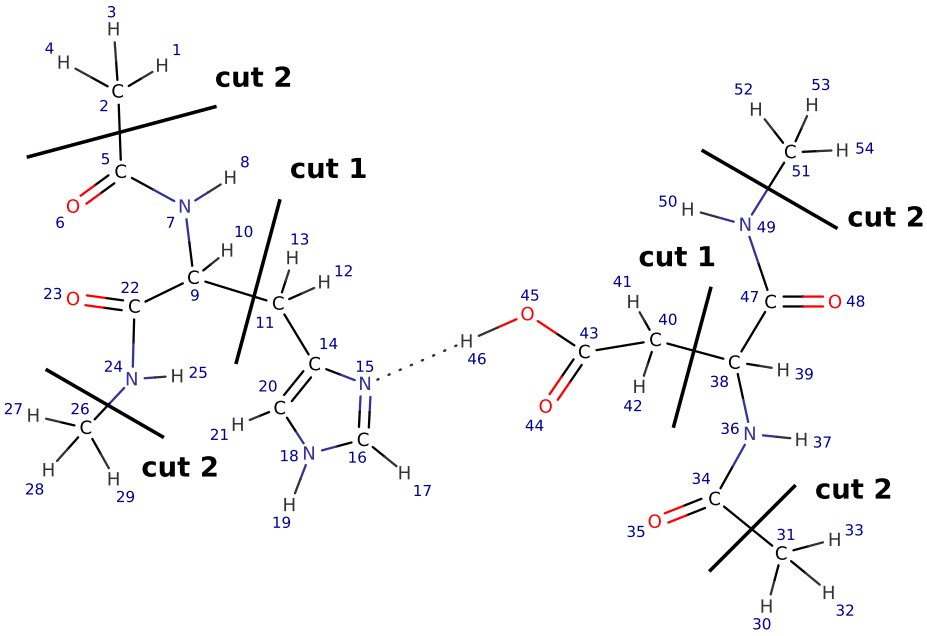

FIGURE 2.6: System separation of the Ace-His-Nme and Ace-Asp-Nme system.

The third example comprises an Ace-His-Nme and a Ace-Aspartic acid(Asp)-Nme peptide, where the proton transfer between the Asp and His residues was investigated. In this reaction, the oxygen(O_{45})-hydrogen(H_{46}) distance in the Asp residue was set as the reaction coordinate (see Fig. 2.6. for the numbering of the atoms). Single point energy calculations were performed as the starting oxygen-hydrogen distance (0.8 Å) was incremented by 0.1 Å. The relative energies with respect to the the minimum of the potential energy surface are reported.

Just as in the second example, the effect of the QM subsystem size was investigated by choosing a smaller (cut1) and a larger (cut2) QM region. In the case of cut1, bonds between the C_α ($C_{9,His}$, $C_{38,Asp}$) and C_β ($C_{11,His}$, $C_{40,Asp}$) connect the QM and MM

subsystems, while in cut2 only the methyl groups were included in the MM region, hence the system was separated along the bonds $C_{2,His}$-$C_{5,His}$ and $N_{24,His}$-$C_{26,His}$, as well as $C_{31,Asp}$-$C_{34,Asp}$ and $N_{49,Asp}$-$C_{51,Asp}$ (Fig. 2.6).

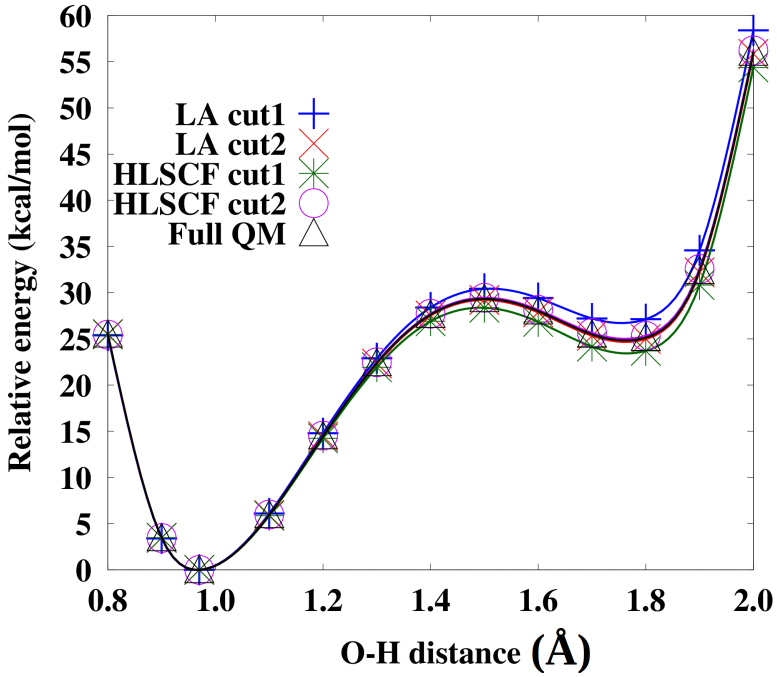

FIGURE 2.7: Energy of the Ace-His-Nme and Ace-Asp-Nme system as a function of the proton from the oxygen of Asp.

Results of the LA and HLSCF approaches together with the reference full QM calculation are shown in Fig. 2.7. As in the previous test, both QM/MM boundary methods well approximate the reference, moreover, the quality of the energy curves also improves as the size of the QM subsystem increases. In the case of cut1 the mean absolute error and the error of barrier height at 1.5 Å are both 1.1 kcal/mol for the LA method, whereas these errors are 0.6 and 0.8 kcal/mol, respectively, for the HLSCF method. The positions of the two minima are also very well reproduced. Applying a larger QM region (cut2) both methods approximate the reference even better: for both methods both the mean error and the error at the potential barrier are 0.1 kcal/mol (well below 1 %). Summarizing the results, it was found that both methods perform well even with the smaller QM subsystem, and they reproduce both the positions and the relative heights of the reference maxima extremely well with the larger QM subsystem.

2.4 Summary

A mixed QM/MM program is developed and tested with sample calculations. The coupling of the MM program AMBER and the QM program MRCC is based on the interface of Walker and co-workers [88]. Besides the traditional LA approach, the HLSCF method was also implemented in the AMBER-MRCC code. In addition to the implementation of the Huzinaga-equation, an automatic model system selector algorithm was also developed, which enables the automatic generation of the required SLMO coefficients without any user intervention. This is an important step towards an effective and user-friendly link orbital-based QM/MM code since the necessity of the generation of system-specific frozen SLMOs was considered as an obstacle that is eliminated by the proposed general method to derive the SLMOs. The current implementation allows the extensive testing of the HLSCF method, and to compare its results with those of the LA and of the full QM calculations.

It was found from the deprotonation energy of the hexanoic acid that the results obtained with both boundary methods monotonously converge to the full QM results with increasing QM subsystem size. Calculations performed for peptides with various QM subsystem sizes also confirm this conclusion. The deprotonation energy was reproduced within a few tenth of kcal/mol by both boundary methods when the subsystem boundary is separated from the site of deprotonation by 5 bonds. The HLSCF results were found to be sensitive to the charge parameters at the boundary, and this sensitivity is attributed to the calculation of the energy difference of differently charged systems. The potential energy curves calculated by rotating the imidazole group in the Ace-Asp-Nme peptide showed good agreement with the reference QM results as both boundary methods have a maximal error not exceeding 5 % when the rotating bond is separated from the QM-MM boundary by 3 bonds. The proton transfer energy curve between an Asp and His residues was very well reproduced considering both the positions and the energies of the minima. Even with a subsystem boundary separated by 2 bonds from the proton acceptors the error is within 2 kcal/mol (7 %) and it reduces significantly when the QM subsystem is further extended by 3 bonds.

The results obtained are in line with former studies [37] that compared the LA and the LSCF approaches at semi-empirical level and concluded that they give comparable results when applied cautiously. However, the LSCF approach is conceptually more appealing and its current Huzinaga-equation-based implementation allows for the first time to perform calculations at the ab initio level without orthogonalizing the basis set to the frozen orbitals. Our study showed that the theoretically more founded HLSCF approach is less accurate in its present form in the case of single point calculations and it requires further developments, such as analytic gradients, to be a useful tool for the community.

Chapter 3

Exact quantum embedding schemes

This chapter is based on Ref. 99 where several developments of the multilevel approaches are presented and extensively tested. The aim of this work was to utilize the Huzinaga-equation in the DFT embedding research, to extend our local correlation approach to the focused models, and to draw a comparison between the DFT-based embedding and the WFT-based multilevel schemes in terms of accuracy and efficiency. Thus, this chapter will start with the discussion of the embedding theory of Miller and Manby [59], and thereafter it will also include our modifications to that scheme. Then, the extension of our linear-scaling LNO-CCSD(T) method to the WFT-in-WFT methods will be presented, and finally the new DFT-in-DFT, WFT-in-DFT, and WFT-in-WFT techniques will be benchmarked.

3.1 Theory

3.1.1 Embedding into DFT: projector-based embedding

When different methods are coupled, or a fragmentation ansatz is applied to a single method (as in the case of sDFT), orthonormality of the orbitals may not be automatically guaranteed. As it was discussed in Sect. 1.3.1.2, the nonadditive kinetic energy functional has to be calculated with an approximate functional or with the complicated OEP method, however, these issues can be avoided by the following exact DFT embedding approach.

3.1.1.1 Embedding scheme of Manby and Miller

Let us divide the system into two subsystems defined by specifying the corresponding atoms: an embedded (active) subsystem hereafter denoted by A and its environment labeled by B. The subsystems are treated with a high- and low-level DFT method, respectively, or alternatively for subsystem A a correlated WFT method can also be used.

Manby and Miller proposed a subtractive scheme for exact DFT embedding, which can be expressed by rewriting Eq. 1.1.5:

$$E_{12}[\tilde{\mathbf{D}}^{\mathrm{A}}; \mathbf{D}^{\mathrm{A}}, \mathbf{D}^{\mathrm{B}}] = E_2[\mathbf{D}^{\mathrm{AB}}] + E_1[\tilde{\mathbf{D}}^{\mathrm{A}}] - E_2[\mathbf{D}^{\mathrm{A}}]. \tag{3.1.1}$$

Note that the subscript of the energy refers to the method, and the superscript of the density refers to the subsystem; thus E_1 is the energy with the higher-level method, $\mathbf{D}^{\mathrm{A}}$ is the density of subsystem A with the lower-level method, $\tilde{\mathbf{D}}^{\mathrm{A}}$ designates the density of subsystem A with the higher-level method, while E_2 and E_{12} is the energy with the low-level method and the energy of the system with the combined approach, respectively. The above embedding scheme is considered exact because if a method is embedded into itself, then $\tilde{\mathbf{D}}^{\mathrm{A}} = \mathbf{D}^{\mathrm{A}}$, and the energy without embedding is recovered.

We suppose that the SCF equations are solved for the entire closed-shell system with the low-level DFT method defined by E_2:

$$E_2[\mathbf{D}^{\mathrm{AB}}] = \mathrm{Tr}\left[\mathbf{hP} + \frac{1}{2}\mathbf{D}^{\mathrm{AB}}\left(\mathbf{J}[\mathbf{D}^{\mathrm{AB}}] - \frac{1}{2}a_{\mathrm{xc},2}\mathbf{K}[\mathbf{D}^{\mathrm{AB}}]\right)\right] + E_{\mathrm{xc},2}[\mathbf{D}^{\mathrm{AB}}] + V_{\mathrm{nn}}, \tag{3.1.2}$$

where $a_{\mathrm{xc},2}$ and $E_{\mathrm{xc},2}$ denotes the exact exchange mixing parameter and the exchange-correlation energy of the low-level technique. Then, the occupied MOs are localized using some localization criterion. For each LMO, the atoms are identified on which the orbital is localized, e.g., an atom is assigned to an orbital if the corresponding Mulliken population of the MO is greater than a threshold. Then, the LMOs can be classified (see Sect. 3.1.1.4) into two groups: the MOs ψ_i^{A}, which are localized on the atoms of the embedded subsystem and the remaining orbitals, ψ_i^{B}, which will be kept frozen in the subsequent calculations. This assignment also fixes the number of electrons in the subsystems, which will be denoted by n^{A} and n^{B}, respectively. We can also compute the density matrices $\mathbf{D}^{\mathrm{A}}$ and $\mathbf{D}^{\mathrm{B}}$ of the A and B subsystems, respectively, from orbitals ψ_i^{A} and ψ_i^{B}, and obviously for the density matrix of the entire system, $\mathbf{D}^{\mathrm{AB}}$, the

$$\mathbf{D}^{\mathrm{AB}} = \mathbf{D}^{\mathrm{A}} + \mathbf{D}^{\mathrm{B}} \tag{3.1.3}$$

identity holds. Now that the corresponding densities are at our disposal, the energy of the subsystem of interest can be calculated with the low-level method:

$$E_2[\mathbf{D}^{\mathrm{A}}] = \mathrm{Tr}\left[\mathbf{hP} + \frac{1}{2}\mathbf{D}^{\mathrm{A}}\left(\mathbf{J}[\mathbf{D}^{\mathrm{A}}] - \frac{1}{2}a_{\mathrm{xc},2}\mathbf{K}[\mathbf{D}^{\mathrm{A}}]\right)\right] + E_{\mathrm{xc},2}[\mathbf{D}^{\mathrm{A}}], \tag{3.1.4}$$

where the constant repulsion of the nuclei is omitted.

In the next step, a second set of SCF equations is solved for the active subsystem of n^{A} electrons with an embedding potential describing the effect of the environment to calculate the embedded energy of the entire system, E_{12}. Provided that the reoptimized orbitals of subsystem A, $\tilde{\psi}_i^{\mathrm{A}}$, are orthogonal to ψ_i^{B}, the Fock-matrix that is compatible

with Eq. 3.1.1 can be defined as

$$\tilde{\mathbf{F}} = \frac{\partial E_{12}}{\partial \tilde{\mathbf{D}}^{\mathrm{A}}} + \frac{\partial E_{12}}{\partial \mathbf{D}^{\mathrm{A}}}, \tag{3.1.5}$$

where $\tilde{\mathbf{F}}$ is the Fock-matrix of the second SCF run. The first term on the right-hand side of the above equation is the Fock-matrix of the high-level DFT or HF method:

$$\frac{\partial E_{12}}{\partial \tilde{\mathbf{D}}^{\mathrm{A}}} = \mathbf{F}_1\left[\tilde{\mathbf{D}}^{\mathrm{A}}\right] = \mathbf{h} + \mathbf{G}_1\left[\tilde{\mathbf{D}}^{\mathrm{A}}\right], \tag{3.1.6}$$

where the $\mathbf{h}$ and $\mathbf{G}$ is the usual core-Hamiltonian and electron repulsion part of the Fockian. The second term of Eq. 3.1.5 is the embedding potential of subsystem B:

$$\frac{\partial E_{12}}{\partial \mathbf{D}^{\mathrm{A}}} = \mathbf{F}_2\left[\mathbf{D}^{\mathrm{AB}}\right] - \mathbf{F}_2\left[\mathbf{D}^{\mathrm{A}}\right] = \mathbf{G}_2\left[\mathbf{D}^{\mathrm{AB}}\right] - \mathbf{G}_2\left[\mathbf{D}^{\mathrm{A}}\right], \tag{3.1.7}$$

thus $\tilde{\mathbf{F}}$ incorporates the Fock-matrix of the high-level method and the constant embedding potential of the environment. To ensure the orthogonality, Manby and co-workers proposed to solve the SCF equations with a modified Fock-matrix to obtain MOs (approximately) orthogonal to those from which the embedding density is computed. The basic idea is the creation of a Hermitian operator whose eigenfunctions include the embedding and the optimized orbitals with different eigenvalues:

$$\left(\tilde{\mathbf{F}} + \mu\mathbf{P}\right)\tilde{\mathbf{C}} = \mathbf{S}\tilde{\mathbf{C}}\tilde{\mathbf{E}}, \tag{3.1.8}$$

where μ is a level-shift parameter (typically a large number), $\tilde{\mathbf{C}}$ is the MO coefficient matrix of the reoptimized orbitals ($\tilde{\psi}_i^{\mathrm{A}}$) of subsystem A, $\tilde{\mathbf{E}}$ is the diagonal eigenvalue matrix of the active orbitals, and $\mathbf{P} = \mathbf{S}\mathbf{D}^{\mathrm{B}}\mathbf{S}$ is the matrix of the projector:

$$\hat{P} = \sum_i |\psi_i^{\mathrm{B}}\rangle\langle\psi_i^{\mathrm{B}}|, \tag{3.1.9}$$

which contains the localized occupied MOs of the environment. When the ψ_i^{B} orbitals are eigenfunctions of $\tilde{\mathbf{F}}$, then $\mathbf{P}$ commutes with $\tilde{\mathbf{F}}$ and also with $\tilde{\mathbf{F}} + \mu\mathbf{P}$, the operator in Eq. 3.1.8. Consequently, there exists a set of orthogonal MOs that are eigenfunctions of both $\tilde{\mathbf{F}} + \mu\mathbf{P}$ and $\mathbf{P}$, and $\tilde{\psi}_i^{\mathrm{A}}$ orbitals orthogonal to ψ_i^{B}'s can be obtained by solving Eq. 3.1.8. However, when $\tilde{\mathbf{F}}$ and $\mathbf{P}$ do not commute, orthogonal ψ_i^{B} and $\tilde{\psi}_i^{\mathrm{A}}$ orbitals cannot be obtained with a finite μ. If μ is infinity, the orbitals $\tilde{\psi}_i^{\mathrm{A}}$ are orthogonal to ψ_i^{B}'s, and the embedding is exact; however, in practice, μ can only take a finite value. To avoid numerical instability, a value of 10^3 E_h was suggested for practical applications, which causes an error of a couple of tens of $\mu\mathrm{E}_h$ in the final energies.

The embedded energy can be further improved if E_{12} is augmented with a correction term, which is a first-order correction to the difference between $E_{12}[\tilde{\mathbf{D}}^{\mathrm{A}}; \tilde{\mathbf{D}}^{\mathrm{A}}, \mathbf{D}^{\mathrm{B}}]$ and

$E_{12}[\tilde{\mathbf{D}}^{\mathrm{A}}; \mathbf{D}^{\mathrm{A}}, \mathbf{D}^{\mathrm{B}}]$. Thus, the final DFT-in-DFT energy expression reads as

$$E_{12}^{\mathrm{DFT}}[\tilde{\mathbf{D}}^{\mathrm{A}}; \mathbf{D}^{\mathrm{A}}, \mathbf{D}^{\mathrm{B}}] = E_{12}[\tilde{\mathbf{D}}^{\mathrm{A}}; \mathbf{D}^{\mathrm{A}}, \mathbf{D}^{\mathrm{B}}] + \mathrm{Tr}\left[(\tilde{\mathbf{D}}^{\mathrm{A}} - \mathbf{D}^{\mathrm{A}})\frac{\partial E_{12}}{\partial \mathbf{D}^{\mathrm{A}})}\right], \tag{3.1.10}$$

where the last term obviously does not influence the exactness of the embedding. In the case of WFT-in-DFT embedding, a calculation is carried out with the WFT method using the above Fock-matrix and keeping orbitals ψ_i^{B} frozen, and the WFT-in-DFT energy is evaluated as

$$\begin{aligned} E_{12}^{\mathrm{WFT}}[\Psi^{\mathrm{A}}; \tilde{\mathbf{D}}^{\mathrm{A}}, \mathbf{D}^{\mathrm{A}}, \mathbf{D}^{\mathrm{B}}] = E_2[\mathbf{D}^{\mathrm{AB}}] + E_1^{\mathrm{WFT}}[\Psi^{\mathrm{A}}; \tilde{\mathbf{D}}^{\mathrm{A}}, \mathbf{D}^{\mathrm{A}}, \mathbf{D}^{\mathrm{B}}] \\ - E_2[\mathbf{D}^{\mathrm{A}}] - \mathrm{Tr}\left[\mathbf{D}^{\mathrm{A}}\frac{\partial E_{12}}{\partial \mathbf{D}^{\mathrm{A}})}\right], \end{aligned} \tag{3.1.11}$$

where E_1^{WFT} is the energy of subsystem A evaluated with the WFT method using Fock-matrix $\mathbf{F}_1$, and Ψ^{A} is the corresponding wave function.

3.1.1.2 Huzinaga-embedding

In our study, most of the constituents of the projector-based embedding theory were retained, but we proposed a theoretical modification, and some algorithmic changes were also put forth to improve the selection of the orbital spaces. The former modification affects the SCF procedure: instead of solving the KS or HF equations with the modified Fock-matrix $\tilde{\mathbf{F}} + \mu\mathbf{P}$, we solve the Huzinaga-equations, Eq. 2.1.1. Thus, in our Huzinaga-equation-based embedding scheme, which is shortly referred as Huzinaga-embedding, the matrix form of the Huzinaga-equation,

$$\left(\tilde{\mathbf{F}} - \mathbf{S}\mathbf{D}^{\mathrm{B}}\tilde{\mathbf{F}} - \tilde{\mathbf{F}}\mathbf{D}^{\mathrm{B}}\mathbf{S}\right)\tilde{\mathbf{C}} = \mathbf{S}\tilde{\mathbf{C}}\tilde{\mathbf{E}} \tag{3.1.12}$$

is solved. Again, the Huzinaga-operator, $\tilde{\mathbf{F}} - \mathbf{S}\mathbf{D}^{\mathrm{B}}\tilde{\mathbf{F}} - \tilde{\mathbf{F}}\mathbf{D}^{\mathrm{B}}\mathbf{S}$, has two distinctive features: it is Hermitian and commutable with projector $\mathbf{P}$. Consequently, the ψ_i^{B} orbitals are eigenfunctions of the Huzinaga-matrix, and the $\tilde{\psi}_i^{\mathrm{A}}$ orbitals will be orthogonal to them, irrespective of the choice of the ψ_i^{B}. It can be shown [90] that the $\tilde{\psi}_i^{\mathrm{A}}$ orbitals satisfying Eq. 3.1.12 make the energy in Eq. 3.1.10 stationary with the condition that they form an orthonormal set and they are orthogonal to the fixed ψ_i^{B} orbitals.

The use of the Huzinaga-equations eliminates the bias caused by the arbitrary level shift parameter μ and makes the embedding "strictly exact". Nevertheless, the difference of the numerical results obtained with the original and the modified approach is expected to be small. The computational complexity of the two schemes is very similar. The extra price we have to pay for the solution of the Huzinaga-equations instead of the level-shifted KS or HF equations is just one matrix multiplication per iteration step in place of a matrix addition, which incurs a negligible overhead with respect to other operations in an SCF procedure.

3.1.1.3 Extensions to the Huzinaga-embedding

An important feature of the above method, which is also peculiar to other embedding schemes, is that the density of the environment, $\mathbf{D}^{\mathrm{B}}$, is kept frozen. One would expect better results if $\mathbf{D}^{\mathrm{B}}$ was also relaxed during the SCF procedure, that is, the electrons of the environment also felt the potential of the active electrons treated at a higher level. The embedded mean-field theory (EMFT) of Miller and co-workers [100] is an elegant approach, which is free of this problem. In EMFT, the density matrix of the entire system, $\mathbf{D}^{\mathrm{AB}}$, is partitioned into subsystem densities $\mathbf{D}^{\mathrm{A}}$ and $\mathbf{D}^{\mathrm{B}}$ depending on whether the corresponding basis functions reside on the atoms of subsystem A or B. The energy is written in the same form as Eq. 1.1.5,

$$E_{12}[\mathbf{D}^{\mathrm{AB}}] = E_2[\mathbf{D}^{\mathrm{AB}}] + E_1[\mathbf{D}^{\mathrm{A}}] - E_2[\mathbf{D}^{\mathrm{A}}], \tag{3.1.13}$$

which now depends on the total density $\mathbf{D}^{\mathrm{AB}} = \mathbf{D}^{\mathrm{A}} + \mathbf{D}^{\mathrm{B}}$, and the Fock-matrix is defined by the derivative of the energy with respect to the latter,

$$\mathbf{F} = \frac{\partial E_{12}}{\partial \mathbf{D}^{\mathrm{AB}}}. \tag{3.1.14}$$

The approach is parameter free and was shown to perform advantageously even for partitions across multiple or aromatic bonds. Its major drawback is that it only enables DFT-in-DFT-type embedding, whereas it is not clear how to carry out WFT-in-DFT-type calculations since the orbitals cannot be unambiguously assigned to the subsystems.

We also attempted to develop a self-consistent generalization of the Huzinaga-embedding scheme along this line. We again suppose that the system is partitioned in the same way as for the projector-based embedding, but now the densities of the two subsystems, $\mathbf{D}^{\mathrm{A}}$ and $\mathbf{D}^{\mathrm{B}}$, are optimized self-consistently in a single SCF run. To that end, in each iteration step of the SCF procedure the occupied orbitals are localized, the LMOs are assigned to the two subsystems, and the $\mathbf{D}^{\mathrm{A}}$ and $\mathbf{D}^{\mathrm{B}}$ density matrices are computed and used to evaluate the energy and the Fock-matrix according to the same equations as for EMFT, Eqs. 3.1.13 and 3.1.14. The converged E_{12} energy is directly our DFT-in-DFT energy, and the definition of the WFT-in-DFT energy is straightforward. This self-consistent approach can also be regarded as an alternative to the EMFT scheme. Compared to EMFT, its major advantage is that it enables both DFT-in-DFT- and WFT-in-DFT-type embedding calculations. We also implemented and tested this approach, but we have found that it does not perform better than the parent Huzinaga-embedding. It can be explained by that the latter benefits from error cancellation, which has been also realized by Miller and co-workers [101]. Since the results are not convincing, and, in addition, the self-consistent scheme is numerically less stable, we refrain from the detailed discussion of the algorithm and the results obtained with this approach.

We also note that there is another simple possibility for the self-consistent generalization of the Huzinaga-embedding scheme. One could utilize the "freeze-and-thaw" cycles procedure (see Sect. 1.3.1.2) for the above Huzinaga-embedding scheme, which is used to improve the densities in the FDE framework. In the DFT embedding approach, the Huzinaga-equations can be solved repeatedly, i.e., after converging the Huzinaga-equations for the active subsystem, we can change the role of the active and the embedding subsystem and solve the Huzinaga-equations for the environment keeping the active orbitals frozen. This procedure can be repeated until self-consistency. Concerning that this scheme is relatively costly, and taking into account the moderate success of the above self-consistent approach, we did not embark on implementing it.

3.1.1.4 Orbital selection schemes

Besides the theoretical improvement, we also tested new selection schemes for the active and frozen orbitals since our experience showed that the simple Mulliken-population-based algorithm, which was introduced by Manby and co-workers, is not satisfactory for relatively complicated systems. In their MO selection scheme, those orbitals are selected as active which have Mulliken-population on any given "active" atom greater than a given threshold. We have found that our improved version of the above scheme, nMul, helps to overcome these problems. In the nMul approach, the orbitals are sorted based on the sums of their Mulliken-populations on the active atoms, and the given number of orbitals with the largest sums are selected. However, the nMul technique still requires the *a priori* knowledge of the number of necessary active orbitals. To avoid the issue of predetermined number of active orbitals, the BP algorithm [67] (see Sect. 2.2.2.1) with a completeness criterion of 0.985 is utilized successfully. With the BP algorithm, the embedding scheme outlined above can be used as a black-box method for most molecules except for extended aromatic systems (see Sect. 3.2), where we found that the number of orbitals may not be compatible for the educt and product, and the nMul scheme can select the wrong type of orbitals. For the latter, we developed a two-level orbital selection strategy, which is a combination of the nMul and BP procedures, and it will be referred to as πMul. First, the core, σ, and the π orbitals are distinguished: an LMO is regarded as a π orbital if more than 4 atoms are assigned to it by the BP scheme, otherwise it is supposed to be core or σ. The latter orbitals are added to the active subsystem based on the atom assignment of the BP algorithm. For the number of π orbitals, an upper limit is set to avoid the incorrect number of embedded orbitals for the reactants and products. If one of the atoms assigned to a π orbital by the BP algorithm is part of the embedded subsystem, the given number of orbitals with the largest Mulliken population is selected by the nMul scheme. This selection scheme clearly requires *a priori* knowledge of the maximum number of π orbitals in subsystem A and is thus not fully black-box.

3.1.2 Embedding into WFT: multilayer approaches

Previously in Sect. 1.3.2.1, our linear-scaling LNO-CCSD(T) technique was discussed, which was considered as an interaction-based multilevel approach hence trivial multilevel schemes, in the sense of the focused models, can be proposed. The LMOs can be classified as active or inactive using the algorithms presented in Sect. 3.1.1.4, and the extended domain and local interacting subspace contributions can be evaluated at different levels of theory depending on the type of the corresponding orbitals.

An obvious possibility is to compute the ED contributions at the CCSD(T) and MP2 levels for an active and an inactive i, respectively, and approximate the pair correlation energies also with MP2. The resulting approach, which will be tested here, can be denoted as LNO-CCSD(T)-in-LMP2:

$$\begin{aligned} E^{\text{LNO-CCSD(T)-in-LMP2}} = &\sum_{i\in\text{A}} \delta E_i^{\text{CCSD(T)}}(P_i) \\ &+ \sum_i \delta E_i^{\text{MP2}}(\mathcal{E}_i) - \sum_{i\in\text{A}} \delta E_i^{\text{MP2}}(P_i) \\ &+ \frac{1}{2}\sum_i \sum_j^{\text{distant}} \delta E_{ij}(\mathcal{P}_{ij}). \end{aligned} \tag{3.1.15}$$

Note the energy formula is the same as Eq. 1.3.54, except the summations of the LIS energy contributions, which are restricted to the LMOs of subsystem A. A further choice is to treat the environment simply at the HF level, that is, to neglect the LMP2 and long range correlation (LRC) energy contributions for the environmental orbitals. Then we have the option to only evaluate the pair correlation energies within the active subsystem, but the resulting LNO-CCSD(T)-in-HF approach does not offer any advantage over the LNO-CCSD(T)-in-DFT-type embedding, which can be easily implemented with the infrastructure described in the previous section. On the other hand, we can take into account the LRC between the embedded subsystem and the environment by also calculating the original $\delta E_{ij}(\mathcal{P}_{ij})$ increments, and restrict the summation of the first three terms of Eq. 3.1.15 to the LMOs of subsystem A. In our study, we also tested this approximation, which is abbreviated as LNO-CCSD(T)-in-HF+LRC.

Concerning the speed of the calculations, it is important to note that, for a medium-sized or large molecule, the evaluation of the LMP2 correlation energy is considerably cheaper than the preceding HF calculation even if reduced-scaling algorithms are used for the calculation of the exchange contribution. The calculation of the MP2 pair correlation energies takes just a fraction of the entire computation time. Consequently, the speed of the LNO-CCSD(T)-in-LMP2 and LNO-CCSD(T)-in-HF+LRC calculations is comparable to that for a LNO-CCSD(T)-in-DFT run, where the density functional includes exact exchange.

3.2 Benchmark calculations

Technical details

The localization-based embedding models proposed by us just as the original projector-based embedding scheme [59, 101] have been implemented in the MRCC suite of quantum chemical programs [83]. For DFT-in-DFT embedding we evaluate the performance of the Huzinaga-embedding in comparison to QM/MM and simple vacuum-embedding for energy differences. Concerning the embedding of WFT methods, in addition to the latter two schemes WFT-in-DFT embedding is also compared to the LNO-CCSD(T)-in-LMP2 and LNO-CCSD(T)-in-HF+LRC approaches.

The Huzinaga-embedding scheme was also benchmarked against the parent projector-based approach, and, as expected, we have found that the two methods behave similarly. In accordance with Ref. 59, the difference between total energies obtained with the two approaches is lower than 50 μE_h, and the resulting difference in reaction energies is significantly smaller than the intrinsic errors caused by the embedding approximations. Consequently, in the following we only present the results of the Huzinaga-embedding.

In our calculations, Pople's 6-31G* and 6-311G** [102, 103] and Dunning's (augmented) correlation-consistent polarized valence triple-ζ basis set [(aug-)cc-pVTZ] were employed [104–106]. If the aug-cc-pVTZ basis is used for the heavy atoms, but the diffuse functions are removed from the hydrogens, the resulting basis set will be termed aug'-cc-pVTZ. The density-fitting approximation was employed throughout both for the SCF and the subsequent WFT calculations. The auxiliary basis sets were those of Weigend and co-workers [107, 108] optimized at the HF and MP2 levels.

The DFT calculations used various functionals: the local (spin) density approximation (LDA) [109, 110], the Perdew–Burke–Ernzerhof (PBE) [111] functional, and B3LYP (Becke's three-parameter hybrid functional including the correlation functional of Lee, Yang, and Parr) [112, 113]. These functionals were chosen because they represent different rungs on the "Jacob's ladder" of the DFT functionals. LDA is the simplest functional in our tests that contains exchange and only uses the electron density as ingredient. PBE is a generalized gradient approximation. It contains the electron density and its gradient as ingredients, and this makes the approach more accurate but also results in somewhat increased computational expenses with respect to the simple LDA. The highest-level functional considered here is B3LYP, which is a hybrid functional, that is, it also contains HF exchange. The inclusion of the latter further improves the performance of the method, but B3LYP is also considerably more costly than the two other functionals due to the expensive exact exchange term.

For the localization of the occupied orbitals the Pipek–Mezey [63] method was invoked. The MO spaces were partitioned by the BP and πMul algorithm discussed

in Sect. 3.1.1.4. For the QM/MM calculations, MRCC was interfaced to the AMBER molecular mechanics code [82, 114], and the link atom coupling of the MM and QM subsystems is utilized. For the evaluation of the MM energy functional, the general AMBER force field (GAFF) [92] was used where the corresponding MM parameters were generated by the ANTECHAMBER program [93] using the AM1-BCC charge scheme [94, 95]. The electrostatic and van der Waals terms were not cut off, and only those bonded terms were retained which contained at least one MM or MM host atom. Residual charges which come from zeroing the charges of the QM and MM host atoms were distributed equally among all MM atoms.

The benchmark test sets were adopted from Ref. 100: the substitution reaction 1-chlorodecane to yield 1-decanol, the deprotonation of decanoic acid, the Diels–Alder reaction between the conjugated octadecanonaene and 1,3-butadiene, and the hydrogenation of pentacene. The test reactions are depicted in Fig. 3.1. The DFT-in-DFT, WFT-in-DFT, and WFT-in-WFT calculations used the geometries of Ref. 100 except for OH^- and H_2, whose bond lengths were reoptimized at the same level as in Ref. 100 and the resulting bond lengths of 0.978 Å and 0.744 Å, respectively, were used. For the QM/MM and vacuum-embedding calculation the C–C bond that is cut when fragmenting the system was replaced by a C–H bond with a standardized bond length of 1.09 Å pointing in the original direction. In the latter calculations, only the single C–C bonds of octadecanonaene were broken to preserve the conjugated system, and similarly, for pentacene the aromatic rings were not split up, that is, benzene, naphthalene, anthracene, and tetracene were taken as model systems.

In Subsections 3.2.1 and 3.2.2, the reference energies are subtracted from the reaction energies, which were calculated by the various embedding techniques, and the error is plotted as a function of embedded carbon atoms, except for the Diels–Alder reaction, where the horizontal axis indicates the number of carbon atoms of the octadecanonaene chain included in the embedded subsystem, and subsystem A enlarges symmetrically from the central C–C double bond, where the reaction takes place. Throughout the discussion of the results, the absolute errors will be simply referred to as errors. Reference energies are obtained by using the corresponding high-level methods and basis sets for the aforementioned reactions. We also adopt the convention of Ref. 100 that the results presented for the case where zero carbon atoms are included in the embedded subsystem are the errors of the reaction energies obtained with the low-level method.

It is worth emphasizing that the Huzinaga-embedding with the BP algorithm-based system partitioning required no user intervention, such as the manual selection of the orbitals included in the embedded subsystem or setting the number of those orbitals, for any system separation even when conjugated π-electron systems were considered. The only exception was the hydrogenation reaction of pentacene, where the number of

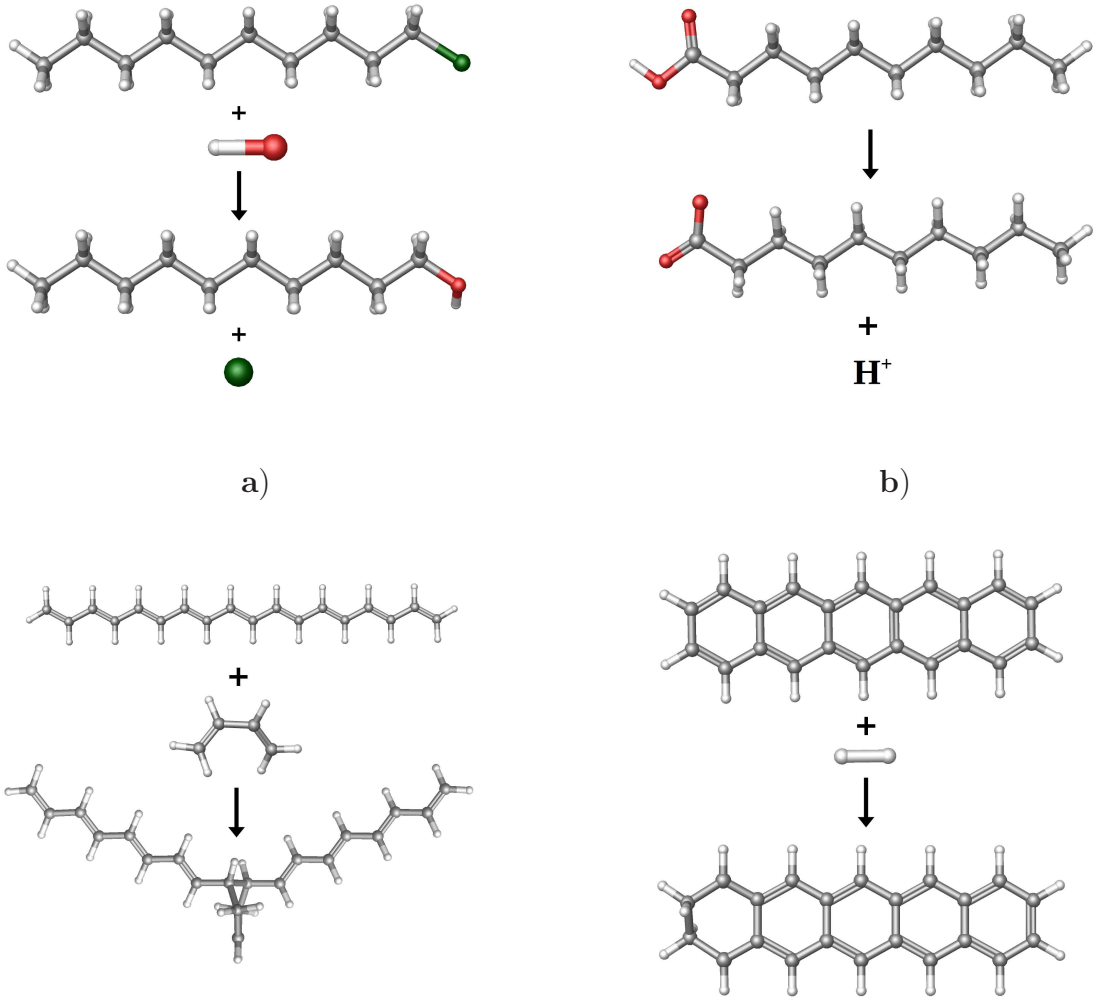

FIGURE 3.1: Visual representation of the benchmark systems. Letters a, b, c, and d refer to the substitution reaction of 1-chlorodecane to yield 1-decanol, the deprotonation of decanoic acid, the Diels–Alder reaction between the conjugated octadecanonaene and 1,3-butadiene, and the hydrogenation of pentacene, respectively. The benchmark examples are designed in order to challenge the embedding techniques by separating the subsystems across various environments.

π-orbitals must have been specified for the πMul algorithm: These orbital limits were incremented by 2 at each new aromatic ring broken except for the smallest embedded subsystem, that is, for the product the limits were 0, 2, 2, 4, 4, 6, 6, 8, 8, 10, and 10 for the embedded subsystems of 2, 4, 6, ..., 22 carbon atoms, respectively, while the above limits were increased by one for pentacene. However, other than setting these parameters, no human intervention was required.

3.2.1 DFT-in-DFT

3.2.1.1 The simplest case: substitution of 1-chlorodecane

The first test for the various embedding techniques is the simplest case, where subsystems A and B are connected through a single σ-bond. In this reaction 1-chlorodecane reacts with a hydroxide anion to form 1-decanol and a chloride anion. The results are presented in Fig. 3.2.

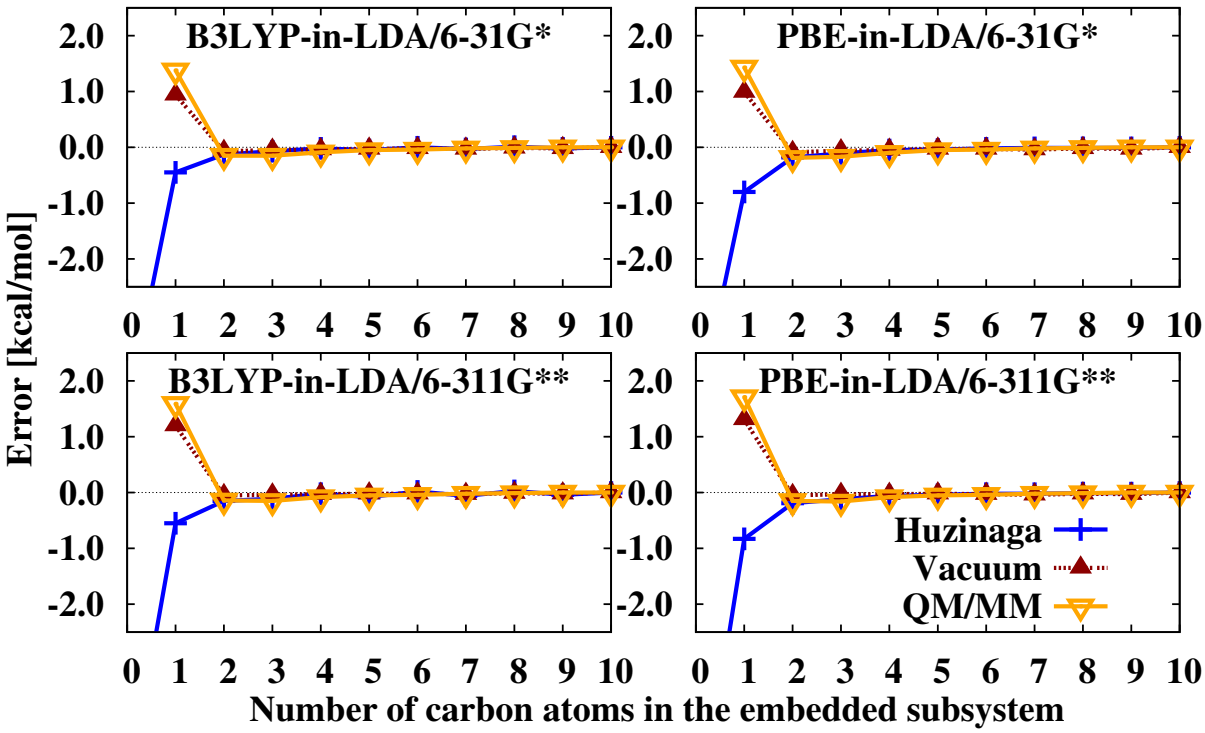

FIGURE 3.2: Error of the reaction energy for the substitution reaction of 1-chlorodecane with various embedding schemes as a function of the number of carbon atoms included in the embedded subsystem using the 6-31G* and 6-311G** basis sets. The B3LYP/6-31G*, B3LYP/6-311G**, PBE/6-31G*, and PBE/6-311G** reference reaction energies are −89.0, −92.1, −88.8, and −91.2 kcal/mol, respectively. The reaction energies of the LDA/6-31G* and LDA/6-311G** low-level methods are −93.7 and −97.5 kcal/mol.

The errors show similar tendencies for the hybrid and the non-hybrid functional, as well as for the smaller and the larger basis set. The error (relative error) of all the embedding techniques is less than 1 kcal/mol (1%) if the embedded subsystem includes at least two carbon atoms. Surprisingly, the cost-efficient vacuum-embedding and QM/MM techniques converge rapidly and produce errors similar to those of the more advanced scheme when sufficiently large subsystem A is applied. The error of the Huzinaga-embedding is still below 1 kcal/mol when the active subsystem includes only one carbon atom, while the error of the two simple approaches is somewhat larger.

3.2.1.2 Dealing with differently charged states: the deprotonation of decanoic acid

For the deprotonation reaction of decanoic acid the subsystems are also separated through a single σ-bond, however, the total charge of the reactant (decanoic acid) and the product (decanoate) is different. This makes the test more sensitive because the effect of error cancellation is likely to be less significant. The results are shown in Fig. 3.3.

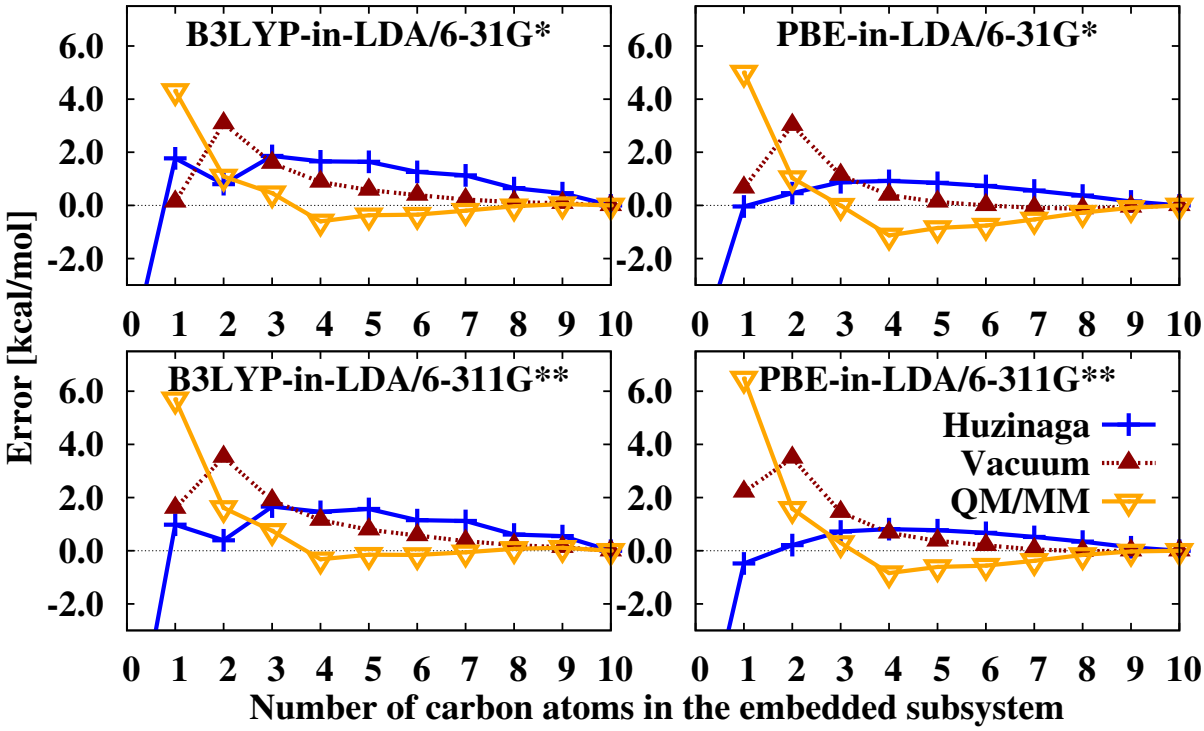

FIGURE 3.3: Error of the reaction energy for the deprotonation of decanoic acid with various embedding schemes as a function of the number of carbon atoms included in the embedded subsystem using the 6-31G* and 6-311G** basis sets. The B3LYP/6-31G*, B3LYP/6-311G**, PBE/6-31G*, and PBE/6-311G** reference reaction energies are 363.3, 359.9, 363.0, and 359.3 kcal/mol, respectively. The reaction energies of the LDA/6-31G* and LDA/6-311G** low-level methods are 357.1 and 351.0 kcal/mol.

As expected, the embedding techniques produce larger errors, and the errors converge more slowly with the enlargement of the embedded subsystem compared with the previous reaction. The Huzinaga-embedding provides satisfactory results even for small active subsystems, but its convergence is not monotonic, and for bigger embedded systems the simplistic methods outperform the Huzinaga scheme. The trends are similar for both basis sets. The PBE-in-LDA reaction energies calculated with the Huzinaga approach are accurate to 1 kcal/mol with active subsystems of any size, while the B3LYP-in-LDA embedding requires quite large embedded systems to reach this accuracy. Nevertheless, if one chooses a sufficiently large embedded region, the vacuum-embedding or the QM/MM method is the most cost-efficient option.

3.2.1.3 Partitioning across double bonds: a Diels–Alder reaction

The Diels–Alder addition of 1,3-butadiene and octadecanonaene is a more difficult test because the subsystems are separated across bonds of order higher than one, moreover, the truncation of the conjugated π orbitals can cause significant errors. The results are depicted in Fig. 3.4.

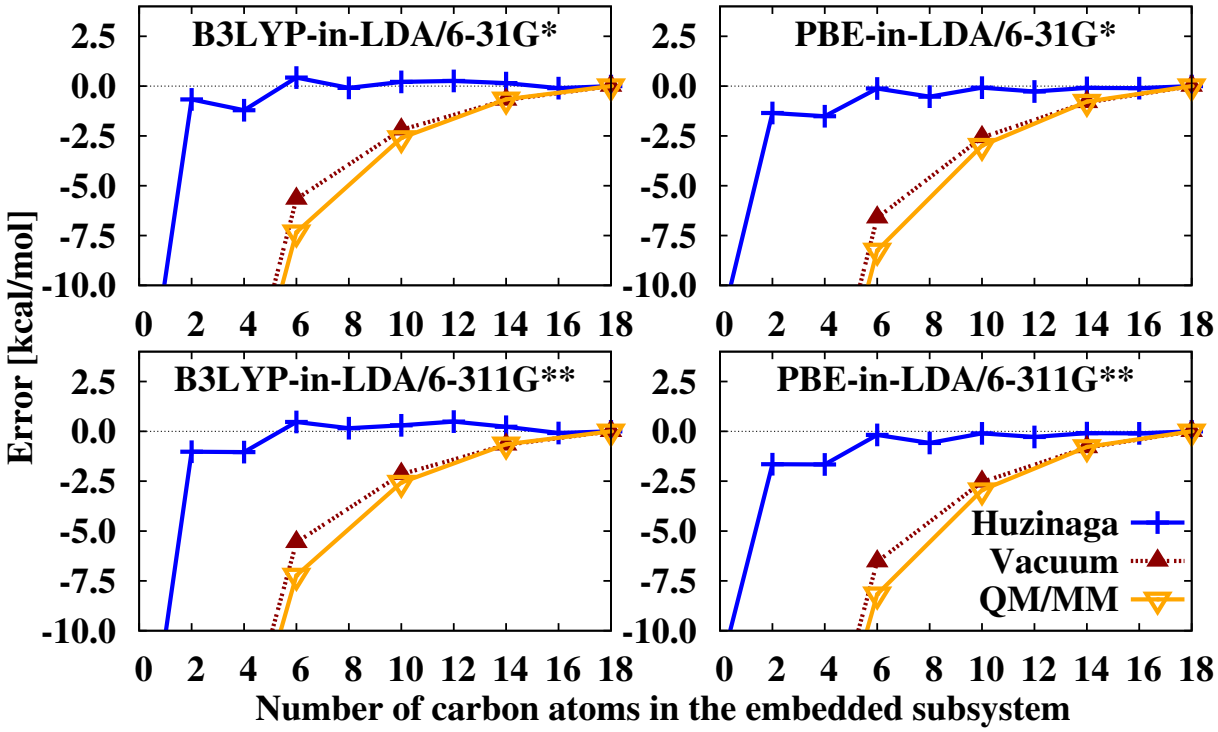

FIGURE 3.4: Error of the reaction energy for the Diels–Alder reaction of 1,4-butadiene and octadecanonaene with various embedding schemes as a function of the number of carbon atoms included in the embedded subsystem using the 6-31G* and 6-311G** basis sets. The B3LYP/6-31G*, B3LYP/6-311G**, PBE/6-31G*, and PBE/6-311G** reference reaction energies are −21.5, −16.3, −27.9, and −23.6 kcal/mol, respectively. The reaction energies of the LDA/6-31G* and LDA/6-311G** low-level methods are −40.1 and −35.6 kcal/mol.

The results show similar trends with both DFT functionals and basis sets. The Huzinaga-embedding outperforms the other embedding methods because the vacuum-embedding and QM/MM requires the entire molecule to be included in the active subsystem for quantitative accuracy, and the error (relative error) of the Huzinaga-embedding is well below 1 kcal/mol (3%) if at least 6 carbon atoms are included in the active subsystem, and it does not exceed 1.7 kcal/mol with any embedded system size. The convergence is not smooth, and the errors usually increase when the system is partitioned across double bonds. It is also worth mentioning that the Huzinaga-embedding probably benefits from the fortunate MO selection of the BP algorithm as it also selects those π orbitals, which are close to the subsystem boundaries but localized mainly on subsystem B.

3.2.1.4 Dividing aromatic systems: the hydrogenation of pentacene

The last test system for the embedding methods is the hydrogenation of pentacene. This is the most challenging reaction because the π orbitals of the aromatic system are even more sensitive to polarization, moreover, the orbitals are also more delocalized, thus the two-level orbital selection scheme is applied for these calculations. The results are presented in Fig. 3.5.

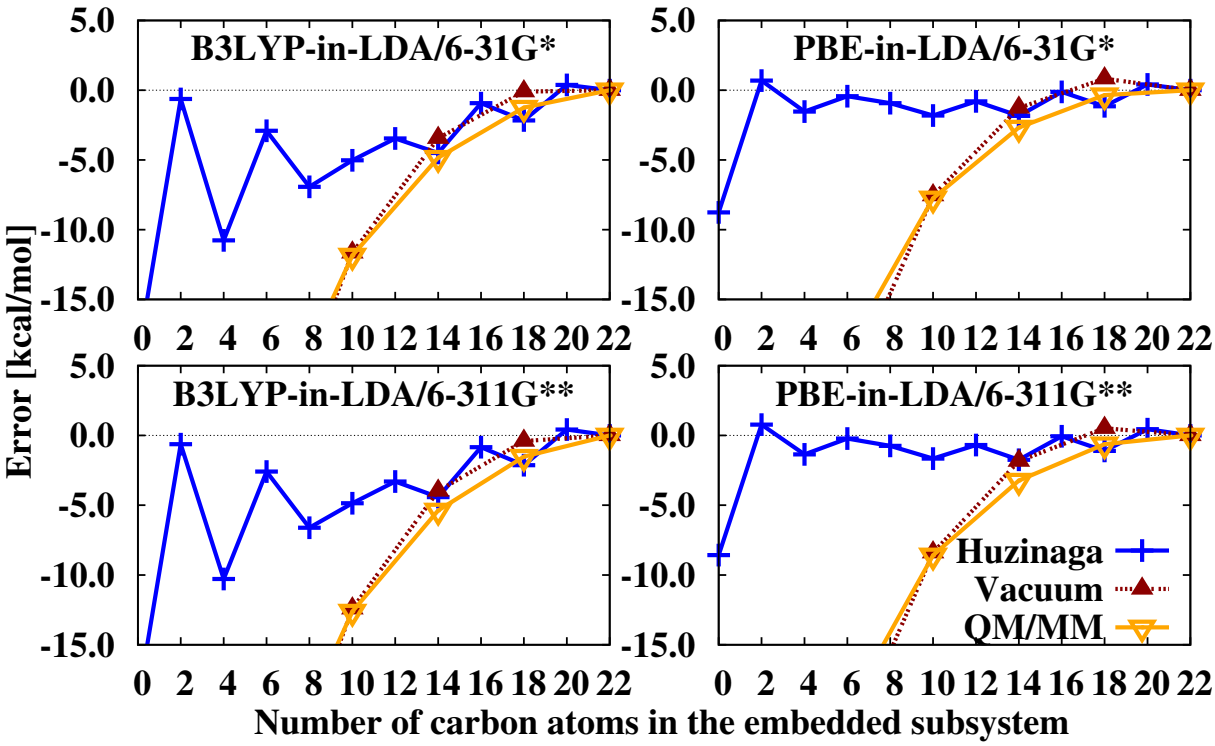

FIGURE 3.5: Error of the reaction energy for hydrogenation of pentacene with the various embedding schemes as a function of the number of carbon atoms included in the embedded subsystem using the 6-31G* and 6-311G** basis sets. The B3LYP/6-31G*, B3LYP/6-311G**, PBE/6-31G*, and PBE/6-311G** reference reaction energies are 32.4, 34.7, 22.2, and 24.3 kcal/mol, respectively. The reaction energies of the LDA/6-31G* and LDA/6-311G** low-level methods are 13.4 and 15.8 kcal/mol.

In general, the behavior of the Huzinaga-embedding is similar with both high-level functionals (and basis sets) but the errors are higher and the convergence is slower for the B3LYP-in-LDA embedding. With the smallest embedded subsystem the error is moderate, and after a big jump at four carbon atoms it starts to decrease but the change is not monotonic, which is a sign of error cancellation effects. The error only drops below 1 kcal/mol, which still corresponds to a considerable relative error, if 20 carbon atoms are part of the embedded subsystem. The vacuum embedding and QM/MM approaches are again not competitive for small embedded subspaces. However, if chemical accuracy is desired, we can also choose these methods since they reach this limit in the same region as the more advanced method.

3.2.2 WFT-in-DFT and WFT-in-WFT

3.2.2.1 The substitution of 1-chlorodecane

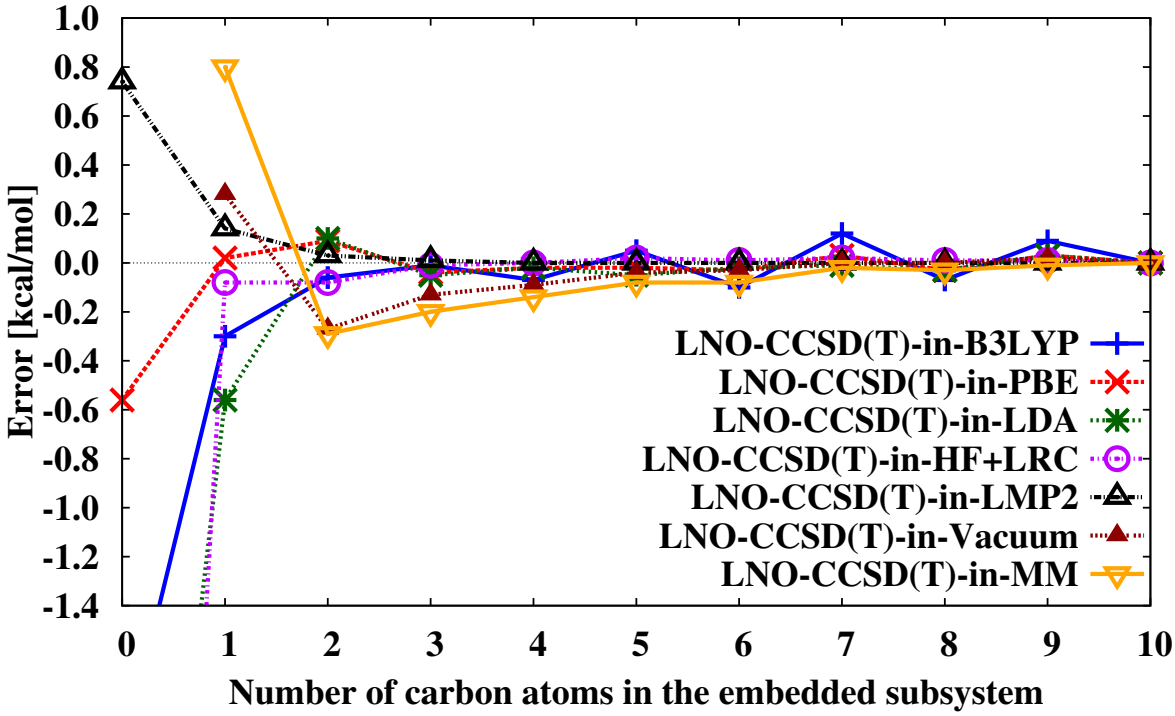

FIGURE 3.6: Error of the reaction energy for the substitution reaction of 1-chlorodecane with various embedding schemes as a function of the number of carbon atoms included in the embedded subsystem using the cc-pVTZ basis set. The LNO-CCSD(T)/cc-pVTZ reference reaction energy is −71.3 kcal/mol. The reaction energies of the B3LYP/cc-pVTZ, PBE/cc-pVTZ, LDA/cc-pVTZ, HF+LRC/cc-pVTZ, and LMP2/cc-pVTZ low-level methods are −73.3, −71.8, −75.4, −78.7, and −70.5 kcal/mol, respectively.

The results for the substitution reaction calculated by the various embedding techniques are shown in Fig. 3.6. The results reveal that the performance of all the embedding methods is good because the errors are below 1 kcal/mol, in addition, the relative error is not greater than 1.1%. The LNO-CCSD(T)-in-HF+LRC approach provides the smallest maximum error, followed by the embedding into PBE, LMP2, vacuum, B3LYP, LDA, and MM. As to the convergence rate, the WFT-in-WFT methods perform the best (LMP2 and HF+LRC embedding), followed by the WFT-in-DFT methods (PBE, B3LYP, and LDA), and the cost-efficient techniques are the slowest to converge with respect to the size of subsystem A. However, in the case of the B3LYP-based embedding, the error fluctuates in the (−0.2, 0.2) kcal/mol interval when larger embedded subsystems are applied. All in all, the considered methods prove their capability of handling this simple test case, but the most accurate technique is the LNO-CCSD(T)-in-LMP2 scheme. If one takes into account the computational cost as well, the vacuum-embedding seems to be the best option here, followed by the QM/MM method and the PBE-based embedding.

3.2.2.2 The deprotonation of decanoic acid

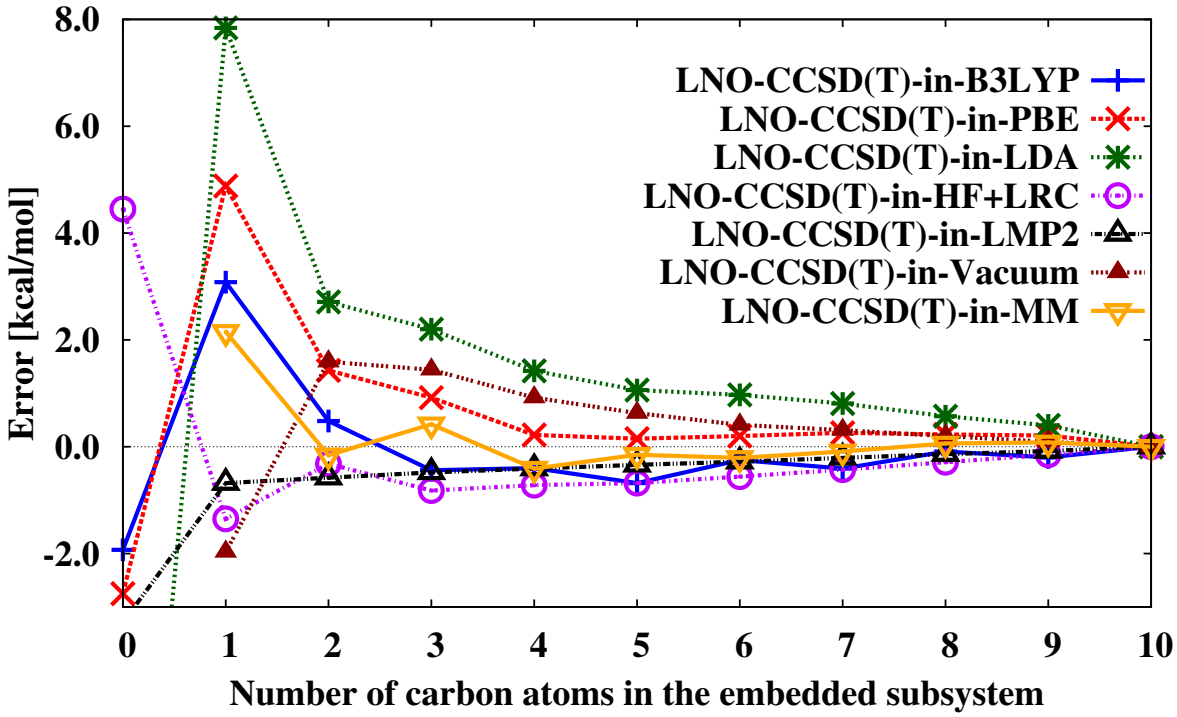

FIGURE 3.7: Error of the reaction energy for the deprotonation reaction of decanoic acid with the various embedding schemes as a function of the number of carbon atoms included in the embedded subsystem using the aug'-cc-pVTZ basis set. The LNO-CCSD(T)/aug'-cc-pVTZ reference reaction energy is 353.2 kcal/mol. The reaction energies of the B3LYP/aug'-cc-pVTZ, PBE/aug'-cc-pVTZ, LDA/aug'-cc-pVTZ, HF+LRC/aug'-cc-pVTZ, and LMP2/aug'-cc-pVTZ low-level methods are 351.2, 350.4, 340.3, 357.6, and 349.9 kcal/mol, respectively.

Fig. 3.7 presents the errors in the deprotonation energies calculated by the embedding techniques investigated. These results show more pronounced differences among the embedding methods than those for the previous reaction: the LNO-CCSD(T)-in-LMP2 method provides the smallest maximum error (0.7 kcal/mol), which is followed by the HF+LRC, the vacuum, the QM/MM, and the DFT embeddings, respectively. Interestingly, the fastest method to converge with respect to the size of subsystem A is the QM/MM, however, its convergence is not monotonic and is only marginally better than that for LNO-CCSD(T)-in-LMP2, which is followed by the HF+LRC, the B3LYP, the PBE, the vacuum, and the LDA embeddings, respectively. That is, the WFT-in-WFT schemes are proven to be more accurate than the LNO-CCSD(T)-in-B3LYP method, which is the most complete WFT-in-DFT embedding but only slightly less expensive than the WFT-in-WFT models considered here. As expected, the WFT-in-DFT methods become more accurate as the quality of the exchange-correlation functional increases, but the vacuum-embedding is still competitive with the simplest WFT-in-DFT method.

3.2.2.3 The Diels–Alder reaction of octadecanonaene

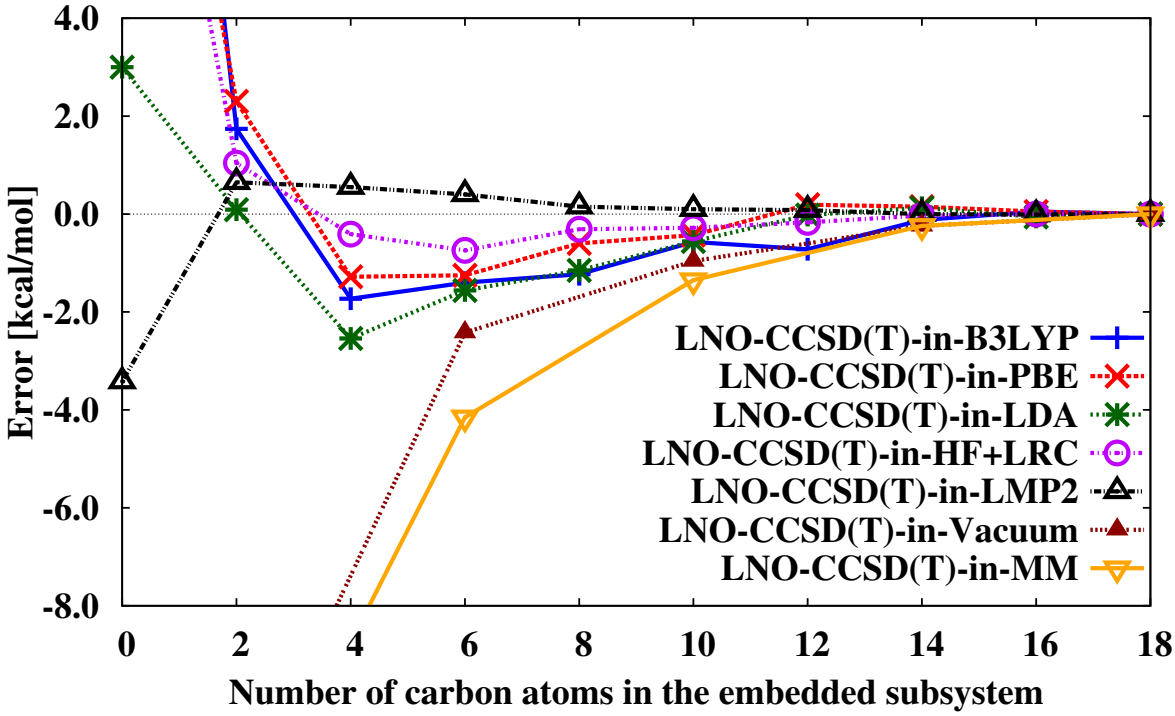

FIGURE 3.8: Error of the reaction energy for the Diels–Alder reaction of 1,4-butadiene and octadecanonaene with various embedding schemes as a function of the number of carbon atoms included in the embedded subsystem using the cc-pVTZ basis set. The LNO-CCSD(T)/cc-pVTZ reference reaction energy is −35.4 kcal/mol. The reaction energies of the B3LYP/cc-pVTZ, PBE/cc-pVTZ, LDA/cc-pVTZ, HF+LRC/cc-pVTZ, and LMP2/cc-pVTZ low-level methods are −13.0, −20.5, −32.4, −22.0, and −38.8 kcal/mol, respectively.

Fig. 3.8 presents the errors of the various embedding schemes calculated for the addition reaction of octadecanonaene and 1,3-butadiene. The results for this reaction show that again the LNO-CCSD(T)-in-LMP2 method produces the smallest maximum error, followed by the HF+LRC, PBE, B3LYP, LDA, vacuum, and the MM embedding, respectively. The latter methods are significantly less accurate than LNO-CCSD(T)-in-LMP2, even if most of them seem to benefit from error cancellation for the smallest embedded subsystems. The convergence of the LNO-CCSD(T)-in-LMP2 is smooth, which is not true for the other WFT-in-WFT and WFT-in-DFT methods. Again, the vacuum-embedding and QM/MM cannot compete with the other methods for this reaction, though it is interesting to observe that vacuum-embedding always outperforms QM/MM. Similar to the previous reactions, the WFT-in-WFT methods perform consistently better than LNO-CCSD(T)-in-B3LYP, which has comparable computational costs. Moreover, the PBE- and LDA-based embeddings have similar accuracy as the LNO-CCSD(T)-in-B3LYP.

3.2.2.4 The hydrogenation of pentacene

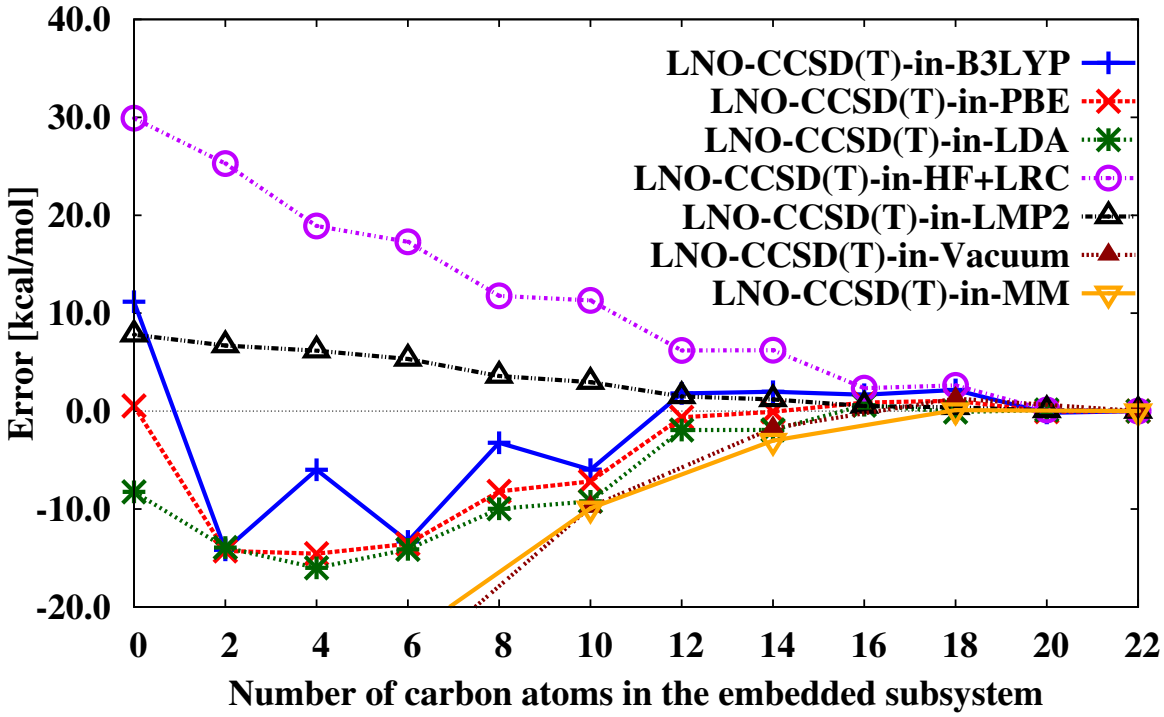

FIGURE 3.9: Error of the reaction energy for the pentacene hydrogenation reaction with various embedding schemes as a function of the number of carbon atoms included in the embedded subsystem using the cc-pVTZ basis set. The LNO-CCSD(T)/cc-pVTZ reference reaction energy is 25.9 kcal/mol. The reaction energies of the B3LYP/cc-pVTZ, PBE/cc-pVTZ, LDA/cc-pVTZ, HF+LRC/cc-pVTZ, and LMP2/cc-pVTZ low-level methods are 37.1, 26.5, 17.7, 55.9, and 33.7 kcal/mol, respectively.

Fig. 3.9 depicts the errors obtained with the embedding techniques studied for the hydrogenation reaction of pentacene. The results show that the most accurate technique is, again, LNO-CCSD(T)-in-LMP2 because this method gives by far the smallest errors among all the embedding schemes. The second best performance is observed for the LNO-CCSD(T)-in-PBE and LNO-CCSD(T)-in-B3LYP schemes, which is followed by the LDA, the HF+LRC, the QM/MM, and the vacuum-embeddings, respectively. Even though LNO-CCSD(T)-in-LMP2 is significantly more accurate than the others, its error (relative error) is still considerable, more than 5 kcal/mol (20%) if the carbon atoms of the first benzene ring are included in the embedded subsystem. The convergence of LNO-CCSD(T)-in-LMP2 is monotonic, though it requires 14 carbons in the active subsystem to reduce the error to 1 kcal/mol, while the fluctuating LNO-CCSD(T)-in-PBE crosses this border earlier. The LNO-CCSD(T)-in-B3LYP errors oscillate and reach convergence more slowly than those of the aforementioned two approaches, which means that it does not offer any advantage even over the cheaper LNO-CCSD(T)-in-PBE scheme.

3.3 Summary

Several embedding and multilevel schemes have been studied. The common feature of the DFT-in-DFT, WFT-in-DFT, and WFT-in-WFT theories is that they utilize orbital localization and are exact, that is, the energy of the whole systems is recovered if the same DFT or WFT is used for the environment as well as for the active site. First, a simple theoretical modification of the projector-based embedding scheme of Manby and co-workers [59] has been introduced. The proposed modification, the use of the Huzinaga-equation, eliminates the bias caused by the arbitrary level shift parameter of the original theory, though it has a very small effect on the numerical results. A modified algorithm has also been suggested for the partitioning of the molecular orbital space, which increases the applicability of the theory. Second, straightforward approaches for extending local correlation methods in the focused models sense have been proposed exploiting their intrinsic feature, that is, they split up the system by default into smaller species, which can be treated at different levels of theory.

Benchmark calculations have been carried out on simplistic, quasi-one dimensional test systems to assess the performance of the embedding and multilevel schemes introduced in comparison to QM/MM and vacuum-embedding. Our results suggest that, if reasonable active subsystems are chosen, the proposed models can be applied as black-box methods except for extended aromatic systems, where the maximum number of selected π orbitals have to be specified. For DFT embedding, the performance of the Huzinaga-equation-based scheme seems to be superior to the other approaches, but QM/MM or even the simple vacuum-embedding performs surprisingly well if the system is partitioned across single bonds. Concerning the utilization of WFT-based techniques as high-level methods, the best choice is to employ a low-level local correlation method, such as LMP2 for the environment. Such multilevel schemes are considerably more accurate than the corresponding WFT-in-DFT approaches, and considering that a LMP2 calculation is significantly less expensive than the preceding HF calculation, WFT-in-DFT embedding can only be competitive if a non-hybrid density functional is used.

Chapter 4

Accelerated quantum embedding

In the previous chapter, the performance of the different multilevel approaches were explored, and the results showed that WFT-based multilevel approaches outperform DFT embedding schemes, while even QM/MM and vacuum embedding were competitive in simple cases. However, combining different levels of theory is not the only way to accelerate calculations. Since the computation time for any electronic structure calculation strongly depends on the size of the AO basis set applied, the most plausible strategy to decrease the computational costs is the reduction of the size of the AO basis. The computational time may decrease with a smaller AO basis, but on the other hand, the predicting power also gets worse as larger errors arise from the incomplete LCAO approximation. Motivated by the fact that the high- and low-level methods often require basis sets of different size for qualitative results, we developed a basis set approximation for our embedding approaches [115]. Thus after an introduction of the existing basis set approximations, our dual basis technique will be discussed in the context of the embedding schemes. Finally, the new approximations are benchmarked on small and realistic test systems.

4.1 Introduction to AO basis set approximations

A trivial way to reduce the AO basis is the use of the mixed-basis set approach. In this simplistic approximation the atoms of the chemically relevant part of the system are described using a large AO basis set, whereas a smaller basis set is employed for the remaining atoms [116–124]. Although this approach is conceptually simple and trivial to implement, it is not elegant from the theoretical point of view. The use of basis sets of different sizes gives rise to the artificial polarization of the electronic structure since the electron density grows in the region where more basis functions can be found. These artifacts considerably worsen the efficiency of the approach for chemical processes where the electronic structure significantly changes, such as charge manifestation processes [125].

In addition, the use of mixed-basis sets might cause convergence problems for SCF algorithms and also increases the probability that the SCF procedure converges to an unstable solution (see later).

The so-called dual basis (DB) set approach, which is free from the aforementioned problems, is another related strategy to decrease the basis set size. In a DB HF or KS SCF calculation the SCF equations are first solved in a relatively small basis set, the resulting electron density or MOs are some way expanded in a bigger AO basis, and usually an approximate SCF energy is also evaluated with the latter basis set, hereafter referred to as the large basis. For the computation of the correlation energy in a DB framework the simplest approach is to use the larger basis set together with the MOs constructed in the preceding DB HF calculation, but more complicated correlated DB methods also exist, in which two basis sets are employed also at the correlation calculation. The DB approximation was first employed in 1983 at the HF level by Havriliak and King [126], who studied the Rydberg states of the ammonium radical. For the first time, DB sets were applied in correlated calculations by Jurgens-Lutovsky and Almlöf [127]. This approach was later also adapted in the DB MP2 method of Wolinski and Pulay [128, 129].

Significant progress has been made on DB methods by Head-Gordon and co-workers [130–139]. In their DB SCF approach an SCF calculation is converged in a small basis set, the density matrix is projected into the large basis set, a single Roothaan step is taken, and an energy correction is evaluated [132, 133]. Alternatively, the expression for the energy lowering can be reformulated to avoid the diagonalization of the Fock-matrix, which makes the method suitable for a linear-scaling implementation [132]. Based on the new DB HF approach, a DB MP2 method was also developed utilizing the density-fitting (DF) MP2 technique [133]. Later analytic gradients were also implemented for the DB SCF [134] and the DF-MP2 [136] methods. The utility of these approaches was demonstrated in ab initio molecular dynamics simulations [138] and for noncovalent interactions [137, 140].

Further promising DB HF approximations were introduced by Gill *et al.* deriving perturbative corrections that estimate the complete basis set HF energy by diagonalizing the Fock potential from a small basis in a large one [141, 142]. New DB MP2 methods were also proposed by these authors, where the MOs are obtained from the large basis HF calculation, and the DB approximation is invoked at the integral transformation [143, 144]. An interesting dual-level DFT approach was published by Nakajima and Hirao, where, in the primary SCF calculation, not only a smaller basis set but also a less sophisticated functional was used [145]. This approach was also extended to time-dependent DFT [146] and relativistic DFT [147], while Gill and co-workers generalized the method by also applying a cruder quadrature grid at the first SCF run [148]. The DB approximation was used in reduced-scaling correlation methods by Kobayashi and

Nakai [149] and Friedrich and Dolg [150–153], and very recently Røeggen and Gao proposed a scheme closely related to the DB approach for the calculation of the correlation energy of molecules and solids [154]. We also note that the complete auxiliary basis set singles correction employed frequently for improving the HF reference energy in explicitly correlated methods [155–158] can also be regarded as a special DB approximation.

The mixed-basis approach has also been combined with quantum embedding schemes, where the use of mixed-basis sets seems to be a natural way to reduce the computational costs because the system is anyway partitioned into high- and low-level regions [100, 159]. Of course, the major disadvantage of the mixed-basis methods, the imbalanced description of the high and low-level densities, is also inherited by the combined approaches and can be even amplified for embedding theories where the environmental densities are not frozen, such as EMFT [125], especially when using a minimal basis set for the environment. For the latter method it was recently demonstrated that the electron density artifacts can be largely eliminated by the semi-empirical Fock-corrected DFT approach [160], where the energy evaluated with an inexpensive functional and a minimal basis is corrected to recover the KS energy obtained with an accurate functional and a larger basis set. Nevertheless, to our knowledge, embedding schemes have not yet been combined with the genuine DB approaches presented above.

In Ref. 115, we proposed a slightly modified version of the DB method of Head-Gordon and co-workers [133]. The new approach might be interesting on its own right but is particularly suitable for combination with embedding theories. We applied the new approach to our Huzinaga-embedding method and WFT-in-WFT schemes. The performance of the resulting DB embedding theories was evaluated in extensive benchmark calculations and also compared to that of the corresponding mixed-basis embedding schemes.

4.2 Theory

4.2.1 Dual basis set approach

Let us assume that we have a closed-shell system of n electrons and two basis sets, a small and a large one; the former is not necessarily a subset of the latter. We are interested in the HF or KS energy evaluated with the large basis set. The DB calculation starts with solving the SCF equations in the small basis set. The resulting occupied MOs are projected into the large basis and orthonormalized. Since the method is invariant with respect to the orthogonalization algorithm, any one can be chosen. From the orthonormal MOs a density matrix, $\mathbf{D}_0$, is computed. In the case the initial SCF is converged with a density-based algorithm, i.e., no MOs are available, we can directly

project the density matrix onto the large basis and use some purification technique to construct an n-representable $\mathbf{D}_0$. If MOs are needed, they can be obtained as the natural orbitals of $\mathbf{D}_0$ transformed to the orthogonalized AO basis, that is, by diagonalizing matrix $\mathbf{S}^{1/2}\ \mathbf{D}_0\ \mathbf{S}^{1/2}$.

Subsequently, one SCF iteration is performed with $\mathbf{D}_0$ in the large basis set. The Fock-matrix $\mathbf{F}[\mathbf{D}_0]$ is constructed, the $E[\mathbf{D}_0]$ energy is calculated, the Fock-matrix is diagonalized, and an improved density matrix, $\mathbf{D}_1$, is computed from the new MOs. To obtain a good approximation to the converged SCF energy with the large basis set, a first-order estimate for $E[\mathbf{D}_1]$, $E^{[1]}[\mathbf{D}_1]$, is derived. If we expand $E[\mathbf{D}]$ into Taylor series around $\mathbf{D}_0$, drop the quadratic and higher-order terms, and utilize the Fock-matrix definition $\mathbf{F} = \frac{\partial E}{\partial \mathbf{D}}$, we arrive at the

$$E^{[1]}[\mathbf{D}_1] = E[\mathbf{D}_0] + \mathrm{Tr}[\mathbf{F}[\mathbf{D}_0](\mathbf{D}_1 - \mathbf{D}_0)] \tag{4.2.1}$$

approximation, which also defines our DB SCF energy. If a DB correlation calculation is carried out, the DB SCF energy is evaluated as described above, and the correlation energy is calculated with Fock-matrix $\mathbf{F}[\mathbf{D}_0]$ and the MOs obtained by the diagonalization of $\mathbf{F}[\mathbf{D}_0]$. Obviously, both the DB SCF and the DB correlation approaches are exact in the sense that, if the small basis and large basis sets are identical, the SCF or the correlation energy computed with the given basis set is retrieved.

The DB approach outlined here differs at two important points from that presented in Ref. 133. First, in our energy expression, Eq. 4.2.1, the zeroth-order term, $E[\mathbf{D}_0]$, is an approximate SCF energy evaluated in the large basis, while Head-Gordon and co-workers use the converged small basis SCF energy instead. In this respect our method is analogous to the approach of Gill *et al.* [141] Second, in contrast to Ref. 133 and most of the other DB schemes our density matrix $\mathbf{D}_0$ is n-representable. On the one hand, these changes make the theory somewhat more accurate and theoretically more satisfactory. On the other hand, the calculation of energy gradients is more complicated with our approach, and our purified $\mathbf{D}_0$ is less sparse than the density matrix obtained simply by projecting the small basis density matrix into the large basis, which worsens the efficiency of the Fock-matrix construction in the large basis. Anyway, the n-representability of the density matrix or, equivalently, the orthonormality of the orbitals used to evaluate the large-basis Fock-matrix is a must for embedding theories, which are of primary concern in this study. Furthermore, as it will be obvious soon, the generalization of Eq. 4.2.1 to our embedding method would be problematic if $E[\mathbf{D}_0]$ was replaced by the converged small basis SCF energy.

4.2.2 Huzinaga-embedding with dual basis sets

We seek the generalization of the DB approach presented in Sect. 4.2.1 for the Huzinaga-embedding outlined in Chapter 3, which is not unambiguous. Of the several possibilities, we explored two ones.

In Ansatz 1 (A1), which is probably the most straightforward DB variant of the Huzinaga-embedding, E_{12}^{DFT} is computed with the small basis set as described in Sect. 3.1.1. The occupied ψ_i^{A} and ψ_i^{B} orbitals are projected into the large basis set. The projected ψ_i^{B} orbitals are orthonormalized, and the resulting orbitals, ψ_{0i}^{B}, are used to compute the n^{B}-representable density matrix $\mathbf{D}_0^{\mathrm{B}}$. We suppose that the ψ_{0i}^{B} orbitals and the $\mathbf{D}_0^{\mathrm{B}}$ density matrix are good enough approximations for the corresponding large basis quantities, or at least their quality is less important compared to that of the orbitals and densities of subsystem A. Hence, we do not endeavor to improve them, and they will be kept fixed in the following. The projected ψ_i^{A} orbitals are also orthonormalized, but beforehand orbitals ψ_{0i}^{B} are projected out by multiplying the corresponding MO coefficient matrix with the projector $\mathbf{Q} = 1 - \mathbf{D}_0^{\mathrm{B}}\mathbf{S}$ to ensure the orthogonality of the MOs of the active subsystem and the environment. From the ψ_{0i}^{A} orbitals constructed in this way, the n^{A}-representable density matrix $\mathbf{D}_0^{\mathrm{A}}$ is computed, and we also calculate the n-representable $\mathbf{D}_0^{\mathrm{AB}} = \mathbf{D}_0^{\mathrm{A}} + \mathbf{D}_0^{\mathrm{B}}$ density matrix. The $\tilde{\psi}_i^{\mathrm{A}}$ orbitals are processed similarly to ψ_i^{A}: they are projected onto the large basis, orbitals ψ_{0i}^{B} are projected out of them using $\mathbf{Q}$, the resulting MOs are orthonormalized to construct orbitals $\tilde{\psi}_{0i}^{\mathrm{A}}$, which are used to evaluate the n^{A}-representable $\mathbf{D}_0^{\mathrm{A}}$. Notice again that $\tilde{\psi}_{0i}^{\mathrm{A}}$ and ψ_{0i}^{B} are orthogonal to each other.

Utilizing the new density matrices estimates for the large basis values of the three terms on the right-hand-side of Eq. 3.1.1 are calculated relying on our DB approach. First, the Fock-matrix corresponding to the low-level method is constructed with $\mathbf{D}_0^{\mathrm{AB}}$ and diagonalized. A new approximation to $\mathbf{D}^{\mathrm{AB}}$, $\mathbf{D}_1^{\mathrm{AB}}$, is computed from the MOs, and an estimate for $E_2[\mathbf{D}^{\mathrm{AB}}]$ is evaluated as

$$E_2^{[1]}[\mathbf{D}_1^{\mathrm{AB}}] = E[\mathbf{D}_0^{\mathrm{AB}}] + \mathrm{Tr}[\mathbf{F}_2[\mathbf{D}_0^{\mathrm{AB}}](\mathbf{D}_1^{\mathrm{AB}} - \mathbf{D}_0^{\mathrm{AB}})]. \tag{4.2.2}$$

Second, to calculate an approximate $E_2[\mathbf{D}^{\mathrm{A}}]$ in a similar way, an appropriate $\mathbf{D}_1^{\mathrm{A}}$ is required. For that purpose one possibility would be to localize the orbitals obtained in the first step and assign the LMOs to the subsystems A and B as described in Sect. 3.1.1.4. In this case, however, orbitals ψ_{0i}^{B} and density matrix $\mathbf{D}_0^{\mathrm{B}}$ would be redefined, and we should avoid this to preserve the consistency with the other contribution. Instead, we project out the ψ_{0i}^{B} orbitals from the occupied MOs calculated in the first step using projector $\mathbf{Q}$. This results in n^{A} pieces of orbitals similar to ψ_{0i}^{A} and n^{B} MOs of close to zero norm. These orbitals are orthonormalized employing canonical orthogonalization dropping n^{B} eigenvectors of the overlap matrix with small eigenvalues. The remaining orbitals are the updated MOs of the active subsystem, ψ_{1i}^{A}, which are used to compute $\mathbf{D}_1^{\mathrm{A}}$. The $F_2[\mathbf{D}_0^{\mathrm{A}}]$

Fock-matrix is constructed, and the approximate $E_2[\mathbf{D}^{\mathrm{A}}]$ contribution is evaluated as

$$E_2^{[1]}[\mathbf{D}_1^{\mathrm{A}}] = E[\mathbf{D}_0^{\mathrm{A}}] + \mathrm{Tr}[\mathbf{F}_2[\mathbf{D}_0^{\mathrm{A}}](\mathbf{D}_1^{\mathrm{A}} - \mathbf{D}_0^{\mathrm{A}})]. \tag{4.2.3}$$

Note that the diagonalization of $\mathbf{F}_2[\mathbf{D}_0^{\mathrm{A}}]$ is avoided here as $\mathbf{D}_1^{\mathrm{A}}$ is calculated in an indirect way. Third, the $\mathbf{F}_1[\tilde{\mathbf{D}}_0^{\mathrm{A}}]$ Fock-matrix is computed, which is then used to construct $\tilde{\mathbf{F}}[\tilde{\mathbf{D}}_0^{\mathrm{A}}; \mathbf{D}_0^{\mathrm{A}}, \mathbf{D}_0^{\mathrm{B}}]$ and the $\tilde{\mathbf{F}} - \mathbf{S}\mathbf{D}_0^{\mathrm{B}}\tilde{\mathbf{F}} - \tilde{\mathbf{F}}\mathbf{D}_0^{\mathrm{B}}\mathbf{S}$ Huzinaga-matrix. The latter is diagonalized to get orbitals $\tilde{\psi}_{1i}^{\mathrm{A}}$, which are still orthogonal to orbitals ψ_{0i}^{B}. The updated density matrix of the active subsystem, $\tilde{\mathbf{D}}_1^{\mathrm{A}}$, is evaluated using orbitals $\tilde{\psi}_{1i}^{\mathrm{A}}$. The approximate large basis $E_1[\tilde{\mathbf{D}}^{\mathrm{A}}]$ contribution is calculated similarly to the other ones as

$$E_1^{[1]}[\tilde{\mathbf{D}}_1^{\mathrm{A}}] = E[\tilde{\mathbf{D}}_0^{\mathrm{A}}] + \mathrm{Tr}[\mathbf{F}_1[\tilde{\mathbf{D}}_0^{\mathrm{A}}](\tilde{\mathbf{D}}_1^{\mathrm{A}} - \tilde{\mathbf{D}}_0^{\mathrm{A}})]. \tag{4.2.4}$$

After introducing the above three contributions the large basis estimate for E_{12} can be defined by the

$$E_{12}^{[1]}[\tilde{\mathbf{D}}_1^{\mathrm{A}}; \mathbf{D}_1^{\mathrm{A}}, \mathbf{D}_0^{\mathrm{B}}] = E_2^{[1]}[\mathbf{D}_1^{\mathrm{AB}}] + E_1^{[1]}[\tilde{\mathbf{D}}_1^{\mathrm{A}}] - E_2^{[1]}[\mathbf{D}_1^{\mathrm{A}}] \tag{4.2.5}$$

expression. To derive our large basis embedding energy we should consider the last term of Eq. 3.1.10. This is a first-order correction for the relaxation of the density of the active subsystem, thus we do not attempt to add another first-order correction for this term for the basis-set extension, but it is simply evaluated with the available $\tilde{\mathbf{D}}_1^{\mathrm{A}}$, $\mathbf{D}_1^{\mathrm{A}}$, and $E_{12}[\tilde{\mathbf{D}}_1^{\mathrm{A}}; \mathbf{D}_0^{\mathrm{A}}, \mathbf{D}_0^{\mathrm{B}}]$ quantities. Consequently, our final DB DFT-in-DFT (HF-in-DFT) energy of A1 reads as

$$E_{12}^{\mathrm{DFT}[1]}[\tilde{\mathbf{D}}_1^{\mathrm{A}}; \mathbf{D}_1^{\mathrm{A}}, \mathbf{D}_0^{\mathrm{B}}] = E_{12}^{[1]}[\tilde{\mathbf{D}}_1^{\mathrm{A}}; \mathbf{D}_1^{\mathrm{A}}, \mathbf{D}_0^{\mathrm{B}}] + \mathrm{Tr}\left[(\tilde{\mathbf{D}}_1^{\mathrm{A}} - \mathbf{D}_1^{\mathrm{A}})\frac{\partial E_{12}}{\partial \mathbf{D}_0^{\mathrm{A}}}\right] \tag{4.2.6}$$

In the case of a WFT-in-DFT approach the WFT calculation is carried out with the $\tilde{\mathbf{F}}[\tilde{\mathbf{D}}_0^{\mathrm{A}}; \mathbf{D}_0^{\mathrm{A}}, \mathbf{D}_0^{\mathrm{B}}]$ Fock-matrix in the basis of the $\tilde{\psi}_{1i}^{\mathrm{A}}$ MOs keeping orbitals ψ_{0i}^{B} frozen, and the A1 DB WFT-in-DFT energy is evaluated according to the

$$\begin{aligned} E_{12}^{\mathrm{WFT}[1]}[\Psi^{\mathrm{A}}; \tilde{\mathbf{D}}_1^{\mathrm{A}}, \mathbf{D}_1^{\mathrm{A}}, \mathbf{D}_0^{\mathrm{B}}] = {} & E_2^{[1]}[\mathbf{D}_1^{\mathrm{AB}}] \\ & + E_1^{\mathrm{WFT}[1]}[\Psi^{\mathrm{A}}; \tilde{\mathbf{D}}_1^{\mathrm{A}}, \mathbf{D}_1^{\mathrm{A}}, \mathbf{D}_0^{\mathrm{B}}] \\ & - E_2^{[1]}[\mathbf{D}_1^{\mathrm{A}}] - \mathrm{Tr}\left[\mathbf{D}_1^{\mathrm{A}}\frac{\partial E_{12}}{\partial \mathbf{D}_0^{\mathrm{A}}}\right] \end{aligned} \tag{4.2.7}$$

equation.

It is easy to see that Eq. 4.2.6 reduces to Eq. 4.2.1 in the limit that the entire system is active, and A1 is exact in the sense that the DB DFT energy evaluated for the entire system with the high-level DFT method is recovered in this limit. A1 also inherits the exactness of the DB approach, that is, independently of the partitioning of the system, the DFT-in-DFT energy, Eq. 3.1.10, calculated with the particular basis set is reproduced if the small basis and large basis sets are the same. These also imply that similar statements hold for the DB WFT-in-DFT energy obtained with A1.

We have also designed a somewhat simpler DB embedding scheme, Ansatz 2 (A2). Here the SCF equations are solved with the small basis set and the low-level method for the entire system, the orbitals are localized and assigned to the subsystems, but the further steps of the embedding calculation with the small basis are not taken. With the large basis set the ψ_{0i}^{A}, ψ_{0i}^{B}, and ψ_{1i}^{A} orbitals as well as the $\mathbf{D}_0^{\mathrm{A}}$, $\mathbf{D}_0^{\mathrm{B}}$, $\mathbf{D}_0^{\mathrm{AB}}$, $\mathbf{D}_1^{\mathrm{A}}$, $\mathbf{D}_1^{\mathrm{B}}$, and $\mathbf{D}_1^{\mathrm{AB}}$ density matrices are constructed, and the low-level energy contributions, $E_2^{[1]}[\mathbf{D}_1^{\mathrm{AB}}]$ and $E_2^{[1]}[\mathbf{D}_1^{\mathrm{A}}]$, are calculated as in the case of A1. In the third step, however, the Huzinaga-equations are iterated to convergence with the $\tilde{\mathbf{F}}[\tilde{\mathbf{D}}^{\mathrm{A}}; \mathbf{D}_0^{\mathrm{A}}, \mathbf{D}_0^{\mathrm{B}}]$ Fock and the $\tilde{\mathbf{F}} - \mathbf{S}\mathbf{D}_0^{\mathrm{B}}\tilde{\mathbf{F}} - \tilde{\mathbf{F}}\mathbf{D}_0^{\mathrm{B}}\mathbf{S}$ Huzinaga-matrix, where $\tilde{\mathbf{D}}^{\mathrm{A}}$ is the self-consistent density matrix formed from orbitals $\tilde{\psi}_i^{\mathrm{A}}$. The DFT-in-DFT energy is calculated according to Eq. 4.2.5 and 4.2.6 with replacing $\tilde{\mathbf{D}}_1^{\mathrm{A}}$ by $\tilde{\mathbf{D}}^{\mathrm{A}}$ and $E_1^{[1]}[\tilde{\mathbf{D}}_1^{\mathrm{A}}]$ by $E_1^{[1]}[\tilde{\mathbf{D}}^{\mathrm{A}}]$. If a WFT-in-DFT calculation is carried out, the correlation energy is computed using the Fock-matrix $\tilde{\mathbf{F}}[\tilde{\mathbf{D}}^{\mathrm{A}}; \mathbf{D}_0^{\mathrm{A}}, \mathbf{D}_0^{\mathrm{B}}]$ with the ψ_{0i}^{B} and $\tilde{\psi}_i^{\mathrm{A}}$ orbitals, and the energy is evaluated as $E_{12}^{\mathrm{WFT}[1]}[\Psi^{\mathrm{A}}; \tilde{\mathbf{D}}^{\mathrm{A}}, \mathbf{D}_1^{\mathrm{A}}, \mathbf{D}_0^{\mathrm{B}}]$.

One can simply prove that the A2 DFT-in-DFT (WFT-in-DFT) approach, in contrast to A1, is exact in the sense that the high-level DFT (WFT) energy of the whole system is retrieved if subsystem A incorporates the entire system. In turn, it is also true for A2 that the DFT-in-DFT (WFT-in-DFT) energy of the original Huzinaga-embedding is reproduced with identical small and large basis sets.

Concerning the computational expenses of our DB DFT-in-DFT embedding scheme, we should take into account that, in the small basis, both the low- and the high-level SCF equations must be solved with A1, whereas only the low-level SCF iterations have to be converged with A2. With the large basis set, however, altogether three Fock-matrix builds - one with n and two with n^{A} electrons - are necessary for A1, while a Fock-matrix of n^{A} electrons has to be constructed in each SCF cycle with A2. This suggests that A2 is only competitive if the embedded subsystem is small with respect to the entire system but turns expensive when enlarging the active subsystem, even though the energy converges to a more accurate value. Concerning our WFT-in-DFT method, the DB approximation only reduces the costs of the SCF part, but those of the subsequent correlation calculation are not affected. Thus, the DB approximation can efficiently decrease the expenses of the WFT-in-DFT embedding calculation if the SCF run is a bottleneck, that is, the size of subsystem A is small with respect to that of the entire system.

As mentioned above, there are further possibilities to introduce DB embedding scheme which are compatible with our DB approach and the Huzinaga-embedding. For example, one could directly project the small-basis $\mathbf{D}^{\mathrm{AB}}$ density onto the large basis, build the $\mathbf{F}_2$ Fock-matrix from the projected and purified density matrix, and diagonalize it. The resulting orbitals could be localized to obtain orbitals ψ_{1i}^{A} and ψ_{1i}^{B}, and then the calculations could proceed along the route similar to either A1 or A2. Alternatively, the

small-basis orbitals projected into the large basis could also be localized before constructing the corresponding density matrices. We did not explore these possibilities since these are all more complicated than the two schemes presented above, which seem the simplest combinations of the Huzinaga-embedding and the DB approximations.

Finally, we also note that the DB embedding schemes introduced can be combined with the mixed-basis approach. In the resulting dual mixed-basis (DMB) scheme the small basis set of the DB embedding calculation is also the small basis of the mixed-basis approximation, while the large basis set is a mixed-basis with the same small basis for the environment and a bigger one for the active subsystem. The major advantage of this scheme is that the initial SCF calculation is carried out with a homogeneous basis, and the density obtained will not be artificially polarized due to the mixing of basis sets. This density is frozen in the subsequent steps of the embedding calculation and expected to generate a more balanced embedding potential than the ones computed with mixed-basis sets. Moreover, these steps of the embedding calculation will be faster because of the decreased size of the basis set.

4.2.3 Further embedding approaches

Our WFT-in-WFT approach, which was discussed in Sect. 3.1.2, can also be combined with our Huzinaga-embedding outlined in Chapter 3 to define WFT-in-WFT-in-DFT embedding schemes, which will be referenced as the Huzinaga multilevel local correlation approximation. In these three-layer approaches, first, a HF-in-DFT calculation is performed, and the WFT-in-WFT energy is calculated on the basis of the HF orbitals keeping the KS orbitals of the environment frozen. We also note that the embedding approaches discussed herein can also be integrated with the QM/MM method [80] to introduce further three- or four-level embedding schemes, such as DFT-in-DFT-in-MM, WFT-in-DFT-in-MM, or WFT-in-WFT-in-DFT-in-MM.

The DB approximation can also accelerate the above embedding methods and the multilevel local correlation schemes as well. In the case of WFT-in-WFT approaches the original DB scheme of Sect. 4.2.1 can be used for the HF calculation. This is only efficient if the HF is the rate determining step, e.g., if a LNO-CCSD(T)-in-LMP2 calculation is carried out, and the embedded subsystem is relatively small. To speed up the three- or four-layer embedding calculations the DB method of Sect. 4.2.2 is applicable. Again, for WFT-in-WFT-in-DFT, WFT-in-DFT-in-MM, or WFT-in-WFT-in-DFT-in-MM embedding considerable gain in speed can only be expected if the subsystem treated at the highest level is compact. Finally, the DMB approach can also be combined with our multilayer methods. Its major advantage is that, in the case of WFT-in-WFT or WFT-in-WFT-in-DFT schemes, it speeds up not only the SCF iterations but also the correlation calculations.

4.2.4 Improved orbital selection schemes

Selecting which orbitals are treated with high- and low-level methods is a crucial point of multilevel approaches. As we noted in Sect. 3.1.1.4, the different number of the active orbitals in the educt and product or the different characteristics of these orbitals (e.g., the total number of active orbitals is the same, but the number of π-orbitals is not) can cause large errors in the computed reaction energies. We developed three orbital selection schemes, which proved useful for our benchmark calculations, however, later we experienced in particular chemical applications that these schemes are not robust enough. The BP algorithm tends to incompatibly determine the total number of active orbitals for the reactant and the product in the cases where the subsystem boundary is close to delocalized orbitals. Calculations using basis sets with diffuse functions also suffer from this anomaly because localized σ orbitals with diffuse functions have longer localization tails. This problem can be bypassed by fixing the total number of active orbitals with the nMul scheme, however, it requires the *a priori* knowledge of the number of the active orbitals, and there is still a chance that the selection may result in different number of active σ and π orbitals for the educt and the product. To overcome this problem, we also presented a two-level selection method, πMul, however, the disadvantages of the BP and nMul selection schemes are obviously inherited by the πMul technique, moreover, the selection criteria for the atom assignment to LMOs was also proved sensitive to the basis set and the exchange-correlation functionals used in the KS calculation. Because of the failure of the orbital selection algorithms, human intervention was required to select the appropriate orbitals. Thus, we proposed improved selection schemes in Ref. 115, which could be used as a black-box approach in all the problematic cases.

The new Mulliken population-based scheme is a very simple but robust automatic orbital selection algorithm. The basics of the aMul technique are similar to the previously discussed nMul scheme: the sum of the Mulliken populations on the active atoms is calculated for each LMO, and those LMOs are considered as active for which this value is greater than an threshold ϵ. The πMul scheme for aromatic systems was also improved. Instead of classifying the MOs based on the number of the assigned atoms, an MO completeness [67] is calculated for up to two atoms by the BP algorithm, that is, the calculation of the BP completeness is finished at two atoms, and an MO is considered as π orbital if its completeness is under a threshold κ. After sorting all the MOs, the σ and π orbitals are selected with the aMul and nMul algorithms, respectively.

Our numerical experience shows that the new techniques are basis set and exchange-correlation functional insensitive, and, the appropriate number of σ and π orbitals are selected for the reactants and products. The aMul algorithm can be used as a black-box scheme if the subsystems are separated through a single bond or an isolated multiple bond, while the πMul scheme is recommended if aromatic systems are broken.

4.3 Validation on small molecules

In this section, after the technical details, first, the accuracy of our DB scheme is discussed briefly with an emphasis on choosing the small basis set. Second, the results of benchmark DFT embedding calculations are presented using our two DB set approaches, namely DB-A1 and DB-A2, and their DMB set alternatives, DMB-A1 and DMB-A2. Finally, the performance of the WFT-in-DFT, WFT-in-WFT, and WFT-in-WFT-in-DFT methods is assessed employing the corresponding DB approaches. The accuracy and efficiency of these techniques are compared to the high-level methods using the large basis set, as well as to the original Huzinaga-embedding employing the large basis and mixed-basis sets. We note that the efficiency of the various techniques will be characterized here by the size of the active subsystem required for a particular accuracy. The efficiency of the methods will also be quantified by presenting factual computation times for a larger example in Sect. 4.4.

Computational details

All the schemes presented here have been implemented in the MRCC program suite [83]. In our calculations Pople's double- (6-31G, 6-31G*, 6-31G**, 6-31+G*, 6-31+G**, 6-31++G**) and triple-ζ (6-311G, 6-311G*, 6-311+G*, 6-311G**, 6-311+G**, 6-311++G**) basis sets [102, 103] as well as Dunning's (augmented) correlation consistent polarized double-, triple-, and quadruple-ζ bases [(aug-)cc-pVDZ, (aug-)cc-pVTZ, and (aug-)ccpVQZ] [104–106] were used. Both the SCF and the correlation calculations employed the density fitting approximation using the auxiliary bases developed by Weigend and co-workers [107, 108]. In all calculations spherical harmonic Gaussians were employed.

The DFT-in-DFT, WFT-in-DFT, and WFT-in-WFT-in-DFT embedding methods were tested, again, with the LDA [109, 110], PBE [111], and B3LYP [112, 113] exchange-correlation functionals, which were chosen in order to represent the various rungs on the Jacob's ladder of density functionals. Since we have found that the convergence of the errors with the subsystem size is similar for these functionals with the Huzinaga-embedding scheme, only the B3LYP-in-LDA, WFT-in-PBE, and WFT-in-WFT-in-PBE results are shown in this thesis. The rest of the results can be found in the Supporting Information of Ref. 115.

The WFT-in-DFT, WFT-in-WFT, and WFT-in-WFT-in-DFT computations used the LNO-CCSD(T) [71, 161, 162] method as the high-level approach with tight truncation thresholds except for aromatic systems, where the default thresholds were applied. In the multilevel local correlation calculations our LMP2 approximation [69] was invoked as the low- or medium-level method.

The Pipek–Mezey [63] localization scheme was employed for the occupied orbitals throughout. Unless otherwise noted, the aMul orbital selection technique was utilized with the default $\epsilon = 0.30$ population threshold. The improved πMul algorithm was used for the aromatic systems with the $\kappa = 0.95$ completeness criterion and the default population threshold for the σ orbitals.

The benchmark systems were, again, taken from Ref. 100: the substitution reaction of 1-chlorodecane with a hydroxide ion to yield 1-decanol and a chloride ion, the deprotonation of decanoic acid, the Diels–Alder reaction between the conjugated octadecanonaene and 1,3-butadiene, and the hydrogenation of pentacene (see Figure 3.1). These test systems will shortly be referenced as a, b, c, and d, respectively.

The reference reaction energies for our benchmark tests were calculated using the high-level methods with the large basis sets. All the conventions which were introduced in Sect. 3.2 are also adopted in the followings. This also means that the presented errors in the zero carbon atom limit are the errors of the reaction energies calculated with the low-level method using the basis set compatible with the embedding scheme, namely, the small basis for the DMB and mixed-basis variants, the DB with the DB ansatz, and the large basis for the original Huzinaga-embedding or multilevel local correlation approaches. For the WFT-in-WFT-in-DFT methods, the HF and DFT regions at the initial HF-in-DFT calculations were kept fixed, and only the border of the two correlation approaches was varied. At DMB and mixed-basis WFT-in-WFT-in-DFT calculations, the large basis set was only used for the region treated by the LNO-CCSD(T) method, while the small basis set was employed for the rest of the molecule.

4.3.1 Dual basis set approximation

To assess the performance of our DB approach and to select the possible small basis and large basis pairings we carried out benchmark calculations for the test reactions with the B3LYP, PBE, and LDA DFT approaches and the LNO-CCSD(T) method without any embedding approximation. Since in our DB scheme the small basis is not necessarily the subspace of the large basis set, one can choose an appropriate small basis from various basis set families. For the DFT theories various combinations of the Pople basis sets were tested, while for the LNO-CCSD(T) method Dunning's (aug-)cc-pVTZ and aug-cc-pVQZ basis sets were employed as the large basis set combined with several Pople and Dunning basis sets as the small basis to investigate the effects of the DB approximation on the correlation energies. The basis sets augmented with diffuse functions were used as the large basis set if the phenomenon considered requires the presence of such functions, that is, for reactions a and b, where one of the reactants or products is an anion.

Our findings are in line with the conclusions of Head-Gordon *et al.* [133, 137], who found that accurate reaction energies can be achieved by eliminating the functions with

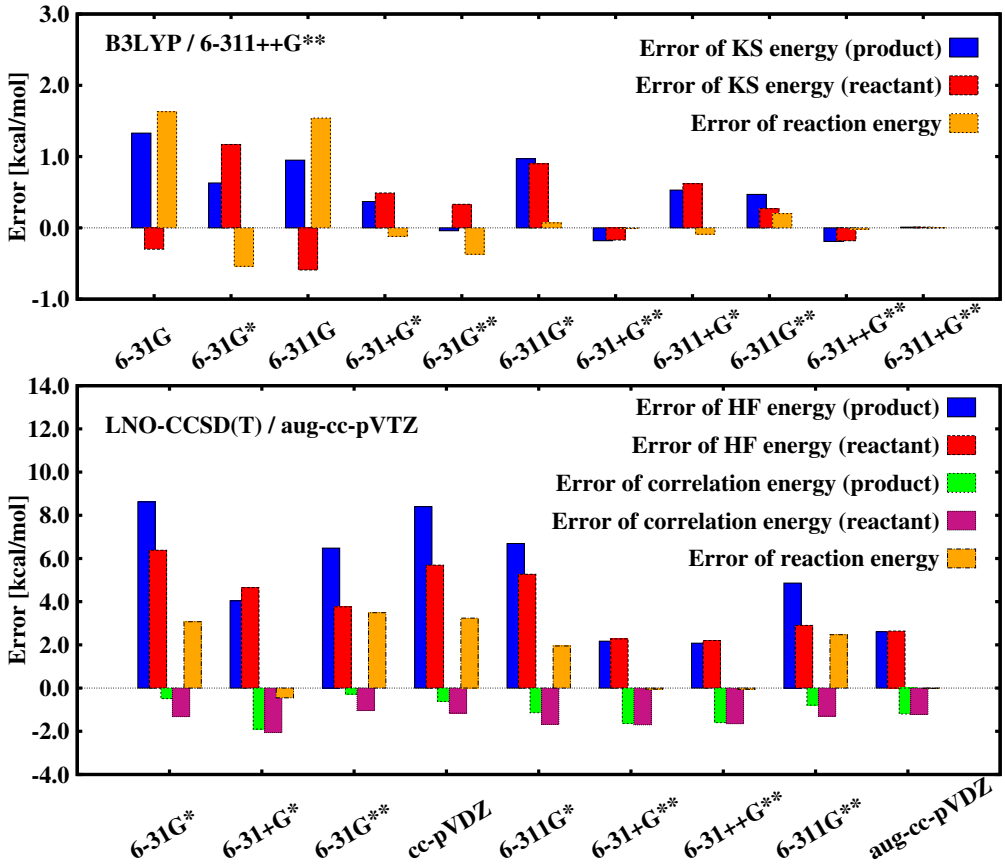

FIGURE 4.1: Error of B3LYP (top) and LNO-CCSD(T) (bottom) reaction energies calculated with the DB approximation using various small basis sets in the case of test reaction b. The KS, HF, and correlation energy errors are decomposed for the educt and the product. The 6-311++G** and aug-cc-pVTZ large basis sets were used for the B3LYP and the LNO-CCSD(T) calculations, respectively. The large basis B3LYP and LNO-CCSD(T) reference reaction energies are 350.63 and 359.03 kcal/mol, respectively.

high angular momenta from the large basis set. As an example, the errors for the B3LYP and LNO-CCSD(T) approaches with the DB approximation for the deprotonation of decanoic acid are shown in Fig. 4.1, while the other tests can be found in the Supporting Information of Ref. 115.

The results for B3LYP reveal that the DB approach can reproduce the reaction energies within a half kcal/mol when at least a double-ζ basis set with polarization functions on heavy atoms is employed. Moreover, the errors can be reduced to a few tenths of a kcal/mol with the addition of diffuse functions on heavy atoms, and the accuracy can be further improved to a couple of hundredths of a kcal/mol when the aforementioned functions are also placed on the hydrogens. These findings suggest that diffuse functions in the small basis are less important for DB DFT methods. However, we should keep in mind that the electronic structure of the decanoate anion is less diffuse as the negative charge is delocalized on the carboxylate group. For small anions with rather diffuse electron distribution, such as the chloride and hydroxide ions, diffuse functions

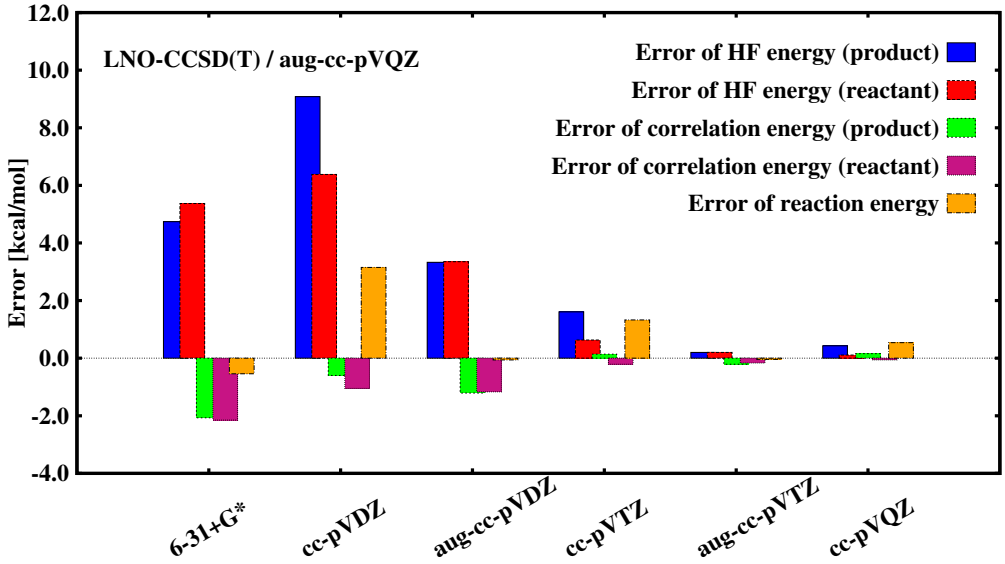

FIGURE 4.2: Error of LNO-CCSD(T) reaction energies (in kcal/mol) calculated with the DB approximation using various small basis sets and aug-cc-pVQZ as large basis in the case of test reaction b. The HF and correlation energy errors are decomposed for the educt and the product. The LNO-CCSD(T)/aug-cc-pVQZ reference reaction energy is 352.94 kcal/mol.

are essential in the small basis, which can be seen from the results for test reaction a (see the Supporting Information of Ref. 115).

The LNO-CCSD(T) DB calculations require at least a double-ζ small basis set with polarization and diffuse functions on heavy atoms for an accuracy of less than a half kcal/mol in the reaction energies. Similarly to the B3LYP results, the addition of the aforementioned functions to hydrogens reduces the reaction energy errors to a couple of hundredths of a kcal/mol. This trend holds on for other reactions as well, which means that the correlated calculations seem to be more sensitive to diffuse functions. If the reactants or products of the reaction require diffuse functions, the small basis must also include them at least on the heavy atoms for a tolerable error. The decomposition of the reaction energy errors also reveals that individual HF energies may be more accurately approximated using a smaller basis set (for example, 6-31+G**) than a bigger one (aug-cc-pVDZ); however, the error cancellation is more beneficial when the small basis and large basis are from the same basis set family. It is also interesting that the error of the correlation energy is less dependent on the small basis when polarization functions are placed at least on the heavy atoms, which can be rationalized by the fact that the correlation energy is more sensitive to the underlying AO basis than the quality of the occupied orbitals. The conclusions for the DB LNO-CCSD(T) results with the aug-cc-pVQZ large basis set are similar to the triple-ζ case (see Fig. 4.2) since the smallest errors are obtained when aug-cc-pVTZ is used as the small basis set. The calculations

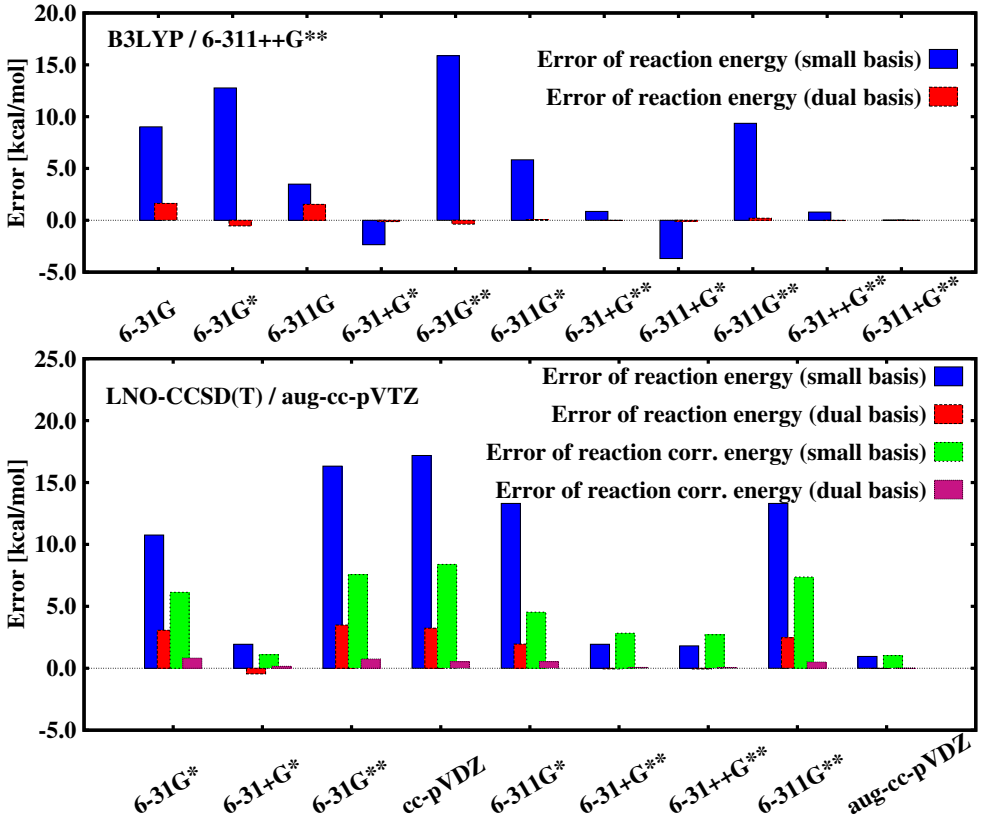

FIGURE 4.3: Error of B3LYP (top) and LNO-CCSD(T) (bottom) reaction energies calculated with the small basis and with the DB approximation using various small basis sets in the case of test reaction b. In the case of the LNO-CCSD(T) errors, the error of reaction correlation energy is also indicated. The 6-311++G** and aug-cc-pVTZ large basis sets were used for the B3LYP and the LNO-CCSD(T) calculations, respectively. The large basis B3LYP and LNO-CCSD(T) reference reaction energies are 350.63 and 359.03 kcal/mol, respectively.

employing the aug-cc-pVDZ basis as the small basis set are also remarkably accurate, and the reaction energy errors are still a kcal/mol or smaller even with the 6-31+G* small basis set. It is also instructive to compare the results of the DB calculations and the simple calculations which are only performed with the corresponding small basis set without any DB approximation. Our results suggest (see Fig. 4.3) that the DB calculations are roughly by an order of magnitude more accurate than the corresponding small basis only calculations if reliable small bases are used.

The DB approximation of Head-Gordon *et al.* was already successfully applied in catalytic reactions with zeolites [163], and, in light of these findings, we also utilized our scheme in modelling the complexation of antioxidants in a very cost-effective manner [164]. Based on the DB results presented above, we choose the double-ζ counterparts of the triple-ζ large basis sets as small basis in the following embedding calculations to maximize the accuracy of our DB embedding ansatz.

4.3.2 DFT-in-DFT

The errors for the B3LYP-in-LDA calculations are presented in Fig. 4.4. For better visualization, the results calculated with ansatz A2 are not shown because they are very close to those obtained by A1. The A2 results together with the details of the DFT embedding calculations employing other functionals can be found in the Supporting Information of Ref. 115.

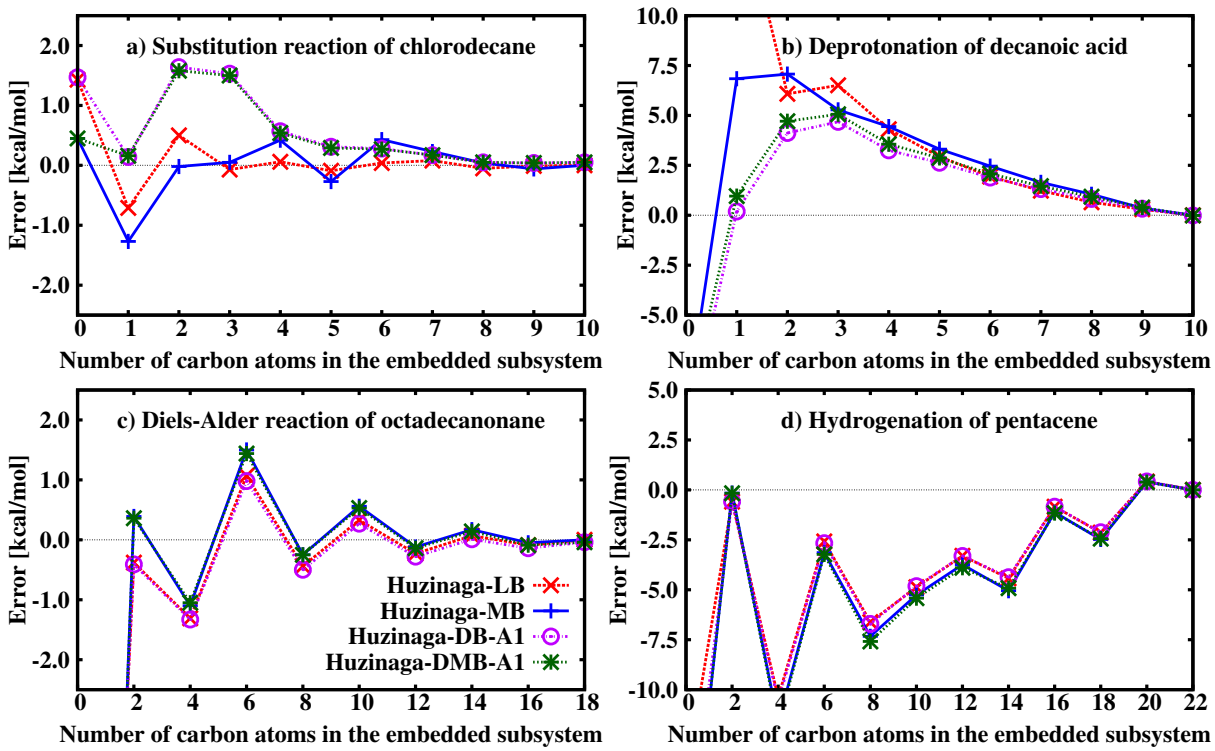

FIGURE 4.4: Error of the reaction energies using various embedding approaches as a function of the number of carbon atoms included in the embedded subsystem. The LB and MB acronyms stand for large basis and mixed-basis, respectively. The 6-311++G** (a), 6-311++G** (b), 6-311G** (c), and 6-311G** (d) large basis sets were used in the calculations, while the corresponding small basis sets in the mixed-basis, DB, and DMB calculations were 6-31+G**, 6-31+G**, 6-31G*, and 6-31G*, respectively. The large basis reference reaction energies are −56.71, 350.63, −16.05, and 34.76 kcal/mol, respectively.

The DB error for the high-level method, which is equivalent to the error of the DB Huzinaga-embedding at the maximum number of carbon atoms, is less than 0.05 kcal/mol for all the reactions, which suggests that the selection of the small basis sets is appropriate. The results show that the errors of the DB, DMB, mixed-basis, and large basis schemes are similar, except for reactions a and b at smaller embedded subsystems. For the former reaction the DB and DMB variants produce errors of 1.0 to 1.5 kcal/mol larger compared to the parent method for the smaller subsystems. The opposite trends can be observed for test reaction b, where the DB and DMB approaches yield 1.5 to 2.0

kcal/mol smaller errors compared with the large basis approach, but the difference of the errors vanishes again as the embedded subsystem enlarges. These findings show that the error cancellation may have an important role in the DB calculations with small active regions but is not consistent. Nevertheless, the absolute errors are still moderate even at the small embedded subsystems, and the various DB schemes perform similarly well when the subsystem boundary is 5 to 6 Å away from the reaction center. Note that this distance approximately coincides with the distance where the error fluctuations of all the embedding variants disappear or reduce to a low level. Interestingly, for the complicated test systems, reactions c and d, the DB Huzinaga-embedding is practically indistinguishable from the parent Huzinaga-embedding method for any reasonable embedded subsystem. The simple mixed-basis approach performs remarkably well for the considered reactions, especially for reactions c and d, but its slower convergence is conspicuous for reaction b, where a charged species is treated with the mixed-basis. The performance of the DMB approach is as good as that of the pure DB scheme.

The errors obtained with the DB-A1 and DB-A2 methods are very similar, and this trend can also be observed for the performances of the DMB-A1 and DMB-A2 techniques. Taking into account that the A2 ansatz is more expensive because a high-level SCF calculation is carried out with the large basis, the A1 ansatz is clearly more beneficial.

All in all, considering also the computational costs, all the approximate embedding schemes are more efficient than the original Huzinaga-embedding. Using larger embedded subsystems, which are anyway required for accurate results, the DB approach is at least as accurate as the mixed-basis scheme. However, increasing the size of the embedded subsystem the costs of the latter are getting more expensive, and the DB approach tends to be more efficient. The most efficient variant seems to be the DMB approach, which inherits the accuracy of the DB technique but is less expensive than both the DB and the mixed-basis methods.

4.3.3 WFT-in-DFT, WFT-in-WFT, and WFT-in-WFT-in-DFT

In this subsection, the results of LNO-CCSD(T)-in-PBE, LNO-CCSD(T)-in-LMP2, and LNO-CCSD(T)-in-LMP2-in-PBE calculations using the large basis as well as the mixed-basis, DB, and DMB approximations are presented for our benchmark systems. The results of further embedding calculations with the LDA and B3LYP functionals can be found in the Supporting Information of Ref. 115. The convergence of the Huzinaga multilevel local correlation method was tested with a fixed DFT and HF partitioning, and only the LNO-CCSD(T) and LMP2 regions were varied. The HF and DFT subsystems were chosen so that the size of the HF region will be minimal while also keeping the error of the LNO-CCSD(T)-in-DFT method using the same HF region as small as possible. In

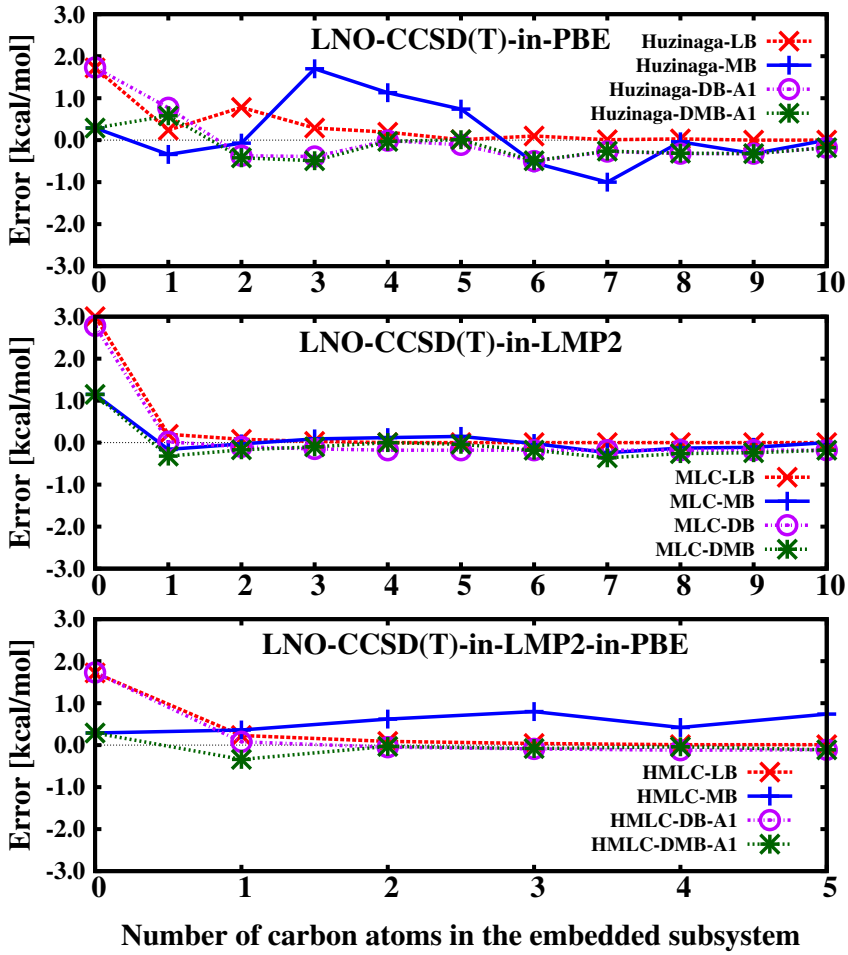

FIGURE 4.5: Error of the reaction energies for the substitution reaction of 1-chlorodecane using various embedding approaches as a function of the number of carbon atoms included in the embedded subsystem. The MLC abbreviation refers to the multilevel local correlation approach, while HMLC means the Huzinaga-embedding based MLC scheme. See Fig. 4.4 for further abbreviations. The aug-cc-pVTZ and aug-cc-pVDZ basis sets were used as large basis and small basis sets, respectively. In the Huzinaga-embedding based multilevel local correlation calculations the boundary of DFT and HF regions was fixed at the fifth carbon atom. The reference LNO-CCSD(T)/aug-cc-pVTZ energy is -52.12 kcal/mol, and the error of the DB approximation is -0.18 kcal/mol.

the case of Huzinaga multilevel local correlation with DMB or mixed-basis calculations, the large basis and small basis boundary also varies, which means that the large basis set is only placed on the LNO-CCSD(T) region, whereas the small basis set is used for the rest of the system.

The results of the various embedding schemes for the benchmark systems a, b, c, and d are presented in Fig. 4.5, 4.6, 4.7, and 4.8, respectively. Note that the accuracy of the DB approximation for the LNO-CCSD(T) method without any embedding (see the caption of the figures) somewhat decreases compared to the DFT results for test reactions

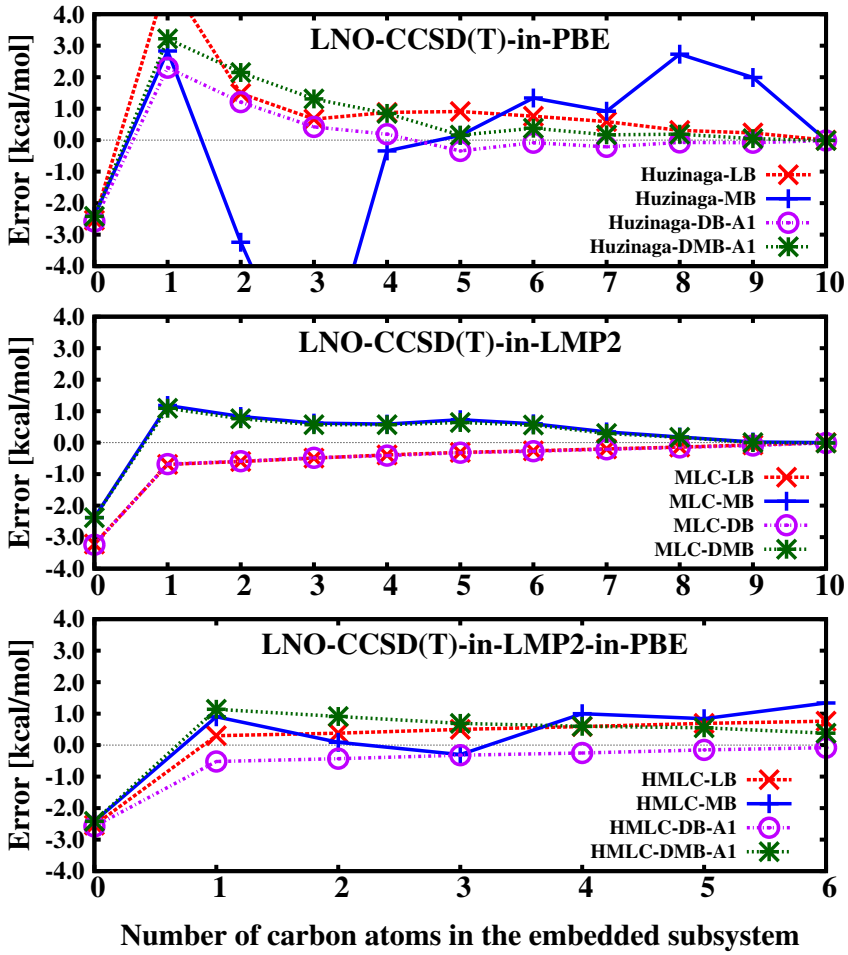

FIGURE 4.6: Error of the reaction energies for the deprotonation of decanoic acid using various embedding approaches as a function of the number of carbon atoms included in the embedded subsystem. See Figs. 4.4 and 4.5 for the abbreviations. The aug-cc-pVTZ and aug-cc-pVDZ basis sets were used as large basis and small basis sets, respectively. In the Huzinaga-embedding based multilevel local correlation calculations the boundary of DFT and HF regions was fixed at the sixth carbon atom. The reference LNO-CCSD(T)/aug-cc-pVTZ energy is 359.03 kcal/mol, and the error of the DB approximation is −0.01 kcal/mol.

a and d. The increased error of the DB ansatz is likely due to the greater sensitivity of the HF method to the quality of the small basis in the case of challenging electronic structures, such as aromatic systems and anions with diffuse charge distributions (cf. Fig. 4.5). Nevertheless, these errors are still tolerable and could be consistently decreased with a less aggressive truncation of the large basis set [133, 137].

Despite the worse performance of the DB approximation in some cases, the errors of the Huzinaga-embedding with the DB and DMB schemes follow similar trends as we have seen for the DFT embedding. The DB and DMB approximations are capable of reproducing the reaction energies computed with the large basis scheme within 1 kcal/mol

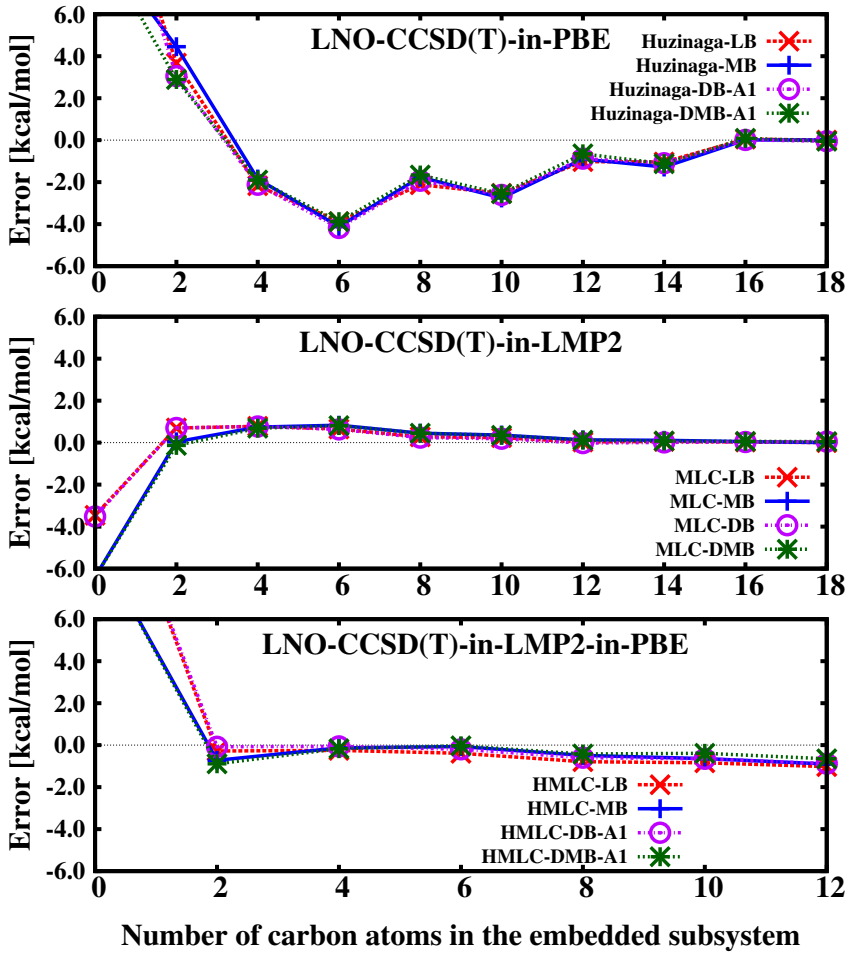

FIGURE 4.7: Error of the reaction energies for the Diels–Alder reaction of octadecanonaene and 1,3-butadiene using various embedding approaches as a function of the number of carbon atoms included in the embedded subsystem. See Figs. 4.4 and 4.5 for the abbreviations. The cc-pVTZ and cc-pVDZ basis sets were used as large basis and small basis sets, respectively. In the Huzinaga-embedding based multilevel local correlation calculations the boundary of DFT and HF regions was fixed at the fourth carbon atom from the reaction center (that is, 12 carbon atoms were included in the embedded subsystem). The reference LNO-CCSD(T)/cc-pVTZ energy is −35.55 kcal/mol, and the error of the DB approximation is −0.05 kcal/mol.

for test systems a and c, while even better convergence with respect to the active subsystem size can be seen for reaction b. For reaction d, the DB (DMB) technique performs worse by 1.0 (3.0 to 7.0) kcal/mol compared with the large basis scheme if the active subsystem is small, but the differences vanish (or significantly decrease) when the embedded subsystem consists of at least 3 carbon rings. This active subsystem size is also necessary for the original method to produce meaningful reaction energies. In contrary to the DFT results, the mixed-basis technique is unreliable for reactions a and b because large error fluctuations can be observed even at extended embedded subsystems. This

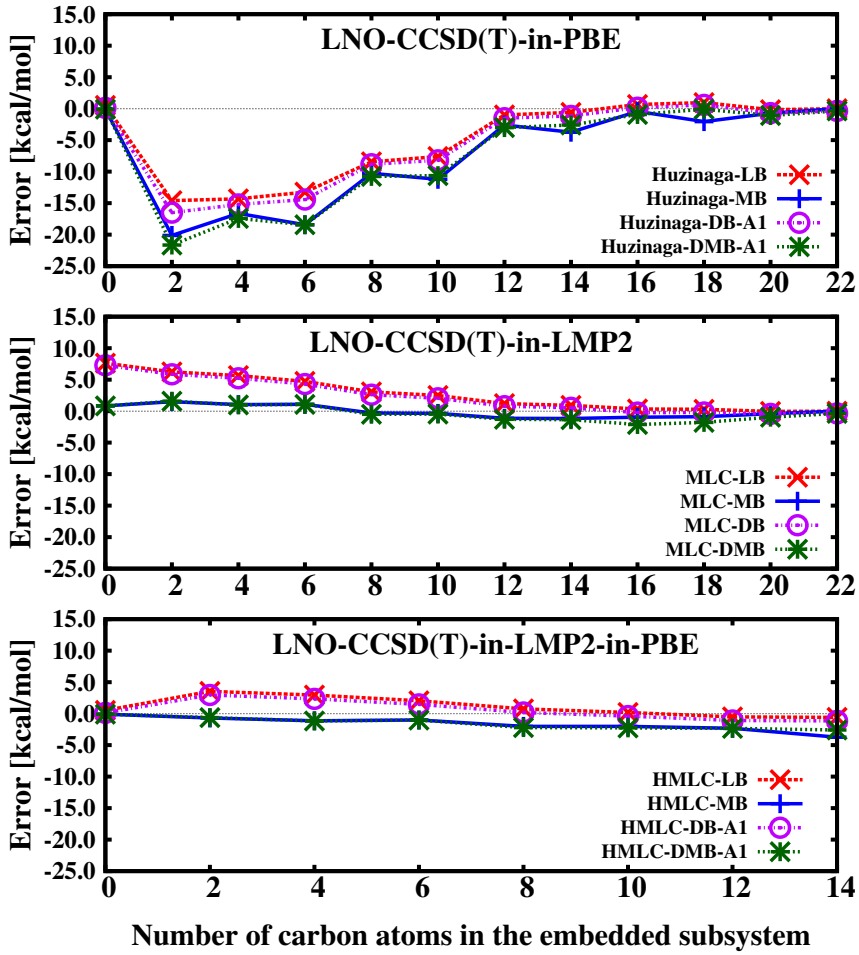

FIGURE 4.8: Error of the reaction energies for the hydrogenation of pentacene using various embedding approaches as a function of the number of carbon atoms included in the embedded subsystem. See Figs. 4.4 and 4.5 for the abbreviations. The cc-pVTZ and cc-pVDZ basis sets were used as large basis and small basis sets, respectively. In the Huzinaga-embedding based multilevel local correlation calculations the boundary of DFT and HF regions was fixed at the sixth carbon atom from the reaction center (that is, 14 carbon atoms of 3 aromatic rings were included in the embedded subsystem). The reference LNO-CCSD(T)/cc-pVTZ energy is 26.02 kcal/mol, and the error of the DB approximation is −0.36 kcal/mol.

is a consequence of the fact that an unstable HF-in-PBE reference is found in particular points with the default settings of the SCF algorithm. This phenomenon disappears for the most difficult test systems, reactions c and d, where no charged species are included in the reaction, but also less accurate results are produced compared to the large basis method.

The multilevel local correlation approach is not only more accurate than the Huzinaga-embedding but also more expensive, which makes the application of DB approaches even more desirable. The results show that the accuracy of the multilevel local correlation

approaches is practically unaffected for any test reaction by the use of the DB approximation. The reaction energies of the DB and DMB schemes are within 0.2 kcal/mol of the large basis and mixed-basis approaches, respectively, independently of the size of the embedded subsystem.

The combination of the DFT embedding and the multilevel local correlation approach has even more potential for accurate and efficient calculations. The results of the Huzinaga multilevel local correlation technique show that, if the HF and DFT subsystems are carefully chosen, then for any basis set combination the LNO-CCSD(T) region has to cover only 2 to 3 bonds from the reaction center to reproduce the corresponding LNO-CCSD(T)-in-DFT reaction energy within 0.5 kcal/mol. The only exception is reaction d, where larger LNO-CCSD(T) layers are needed. The convergence pattern of the errors with the size of the LNO-CCSD(T) region is reminiscent of that for the multilevel local correlation approach, except for the mixed-basis scheme, which inherits the large error fluctuations from the parent Huzinaga-embedding. These tests indicate that it is unnecessary to correlate all the occupied orbitals of the HF region at the LNO-CCSD(T) level as the LMP2 subsystem serves as a cost-effective intermediate layer.

In conclusion, the WFT embedding test calculations revealed that the DB and DMB schemes, either with the A1 or the A2 approaches, are capable of accelerating the Huzinaga calculations over the original large basis schemes, while the mixed-basis technique shows unacceptable inaccuracies in specific cases. Again, the A2 approach is less favorable over ansatz A1 because no improvement is provided in the accuracy in exchange for the increased computational cost. All the basis set reduction approaches considered here effectively accelerate the multilevel local correlation calculations; however, the DMB method is the clear winner because it saves computation time not only for the HF step but also for the correlation calculation. The most promising technique is the combination of the Huzinaga DFT embedding Ansatz A1 with the DMB and multilevel local correlation, which effectively breaks down the cost of all computational steps while maintaining high accuracy.

4.4 Application: Catalytic reaction of Catechol-*O*-Methyltransferase

To check if the conclusions drawn for the small, quasi-one dimensional molecules are valid for real-life 3D systems and to demonstrate the applicability of our multilevel approaches, in this section, the performance of the methods was tested for an enzyme reaction of biological relevance. In these calculations our methods are also combined with the link atom-based QM/MM technique. The projector-based embedding approach of Manby and co-workers was already applied to such systems [165, 166]; however, this is the first time when the accuracy of the WFT-in-WFT-in-MM and WFT-in-WFT-in-DFT-in-MM approximations as well as multilayer DB techniques is explored with respect to the active subsystem size. Here, we choose the catechol-*O*-methyltransferase (COMT) enzyme as our model system, which has an important role in neurotransmitter regulation by transferring a methyl group from S-adenosyl-L-methionine (SAM) to deprotonated catechol (CAT) amines (see Fig. 4.9). An important feature of the COMT system is that the reacting species, the SAM and the CAT molecules, are noncovalently bound to the protein, thus the boundary effects of the QM/MM technique are small.

In the upcoming subsections, the computational details and the system separation approaches are discussed. Then, the accuracy of our most efficient methods, namely the Huzinaga-A1, the multilevel local correlation method, and the combination of the latter with DMB, is compared to the reference large basis and the DB LNO-CCSD(T) method. Finally, timings for the most relevant calculations are presented.

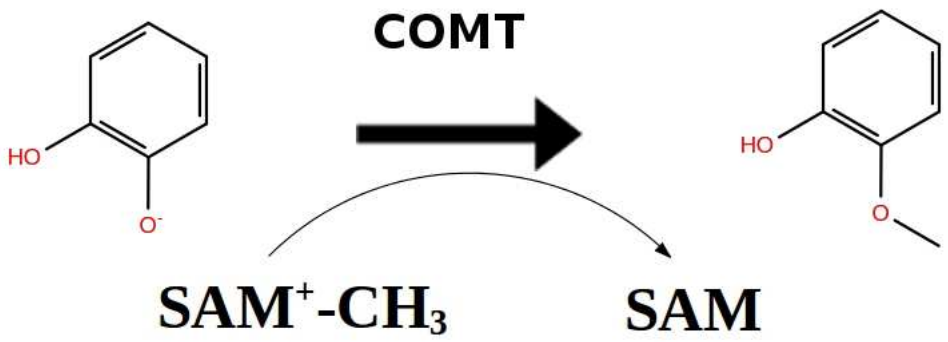

FIGURE 4.9: Methytransfer reaction of the CAT and the SAM molecules catalyzed by the COMT enzyme.

Computational details

The reactant and the product structure of the COMT system were taken from Ref. 34. The structures were unsolvated, optimized at the QM/MM level with the largest QM region by the Martínez group, and we did not attempt to change these structures. The Antechamber [93] and the Tleap utilities of the Ambertools17 program package [81] were invoked to set up the force field parameters for the MM calculations. The ff14SB [23] parameters were used for the protein, while the TIP3P [167] model was employed for the water molecules and the catalytic Mg^{2+} ion. For the CAT and SAM molecules, the General Amber Force Field 2 (GAFF2) [92] was utilized with the AM1-BCC charge scheme [94, 95]. The nonbonded energy terms were not cut off, and only those bonded energy terms were retained that included at least one MM or MM host atom. The residual charges, which are the results of zeroing the charges of the QM and MM host atoms due to electrostatic embedding, were equally distributed among all MM atoms.

The QM/MM calculations were also carried out with Ambertools17 invoking the Amber-Mrcc interface [80] and using the aforementioned force fields for the MM region. The QM calculations were performed with the LNO-CCSD(T) [71,161,162] and LMP2 [69] local correlation methods and the PBE [111] density functional employing Grimme's D3 dispersion correction [168, 169] together with Dunning's cc-pVTZ and cc-pVDZ basis sets [104] as the large basis and small basis sets, respectively. Again, the QM calculations employed the density fitting approximation with the auxiliary basis sets of Weigend and co-workers [107, 108]. The exchange contributions in the HF calculations were approximated using local fitting domains [69, 170] with the default domain settings proposed in Ref. 69. All the calculations used the Boys localization technique [171], whereas the Huzinaga-embedding, the multilevel local correlation approach, and the Huzinaga multilevel local correlation method employed the aMul orbital selection scheme. The reference reaction energy was computed at the LNO-CCSD(T)-in-MM level with the cc-pVTZ basis set. The presented errors were calculated by subtracting the reference reaction energy from the reaction energies computed with the various approaches. Other conventions that were mentioned in Sect. 4.3 are also adopted here.

The selected QM region is illustrated in Fig. 4.10, while the details about the selection of the various subsystems is discussed in the following.

4.4.1 The selection and partitioning of the QM region

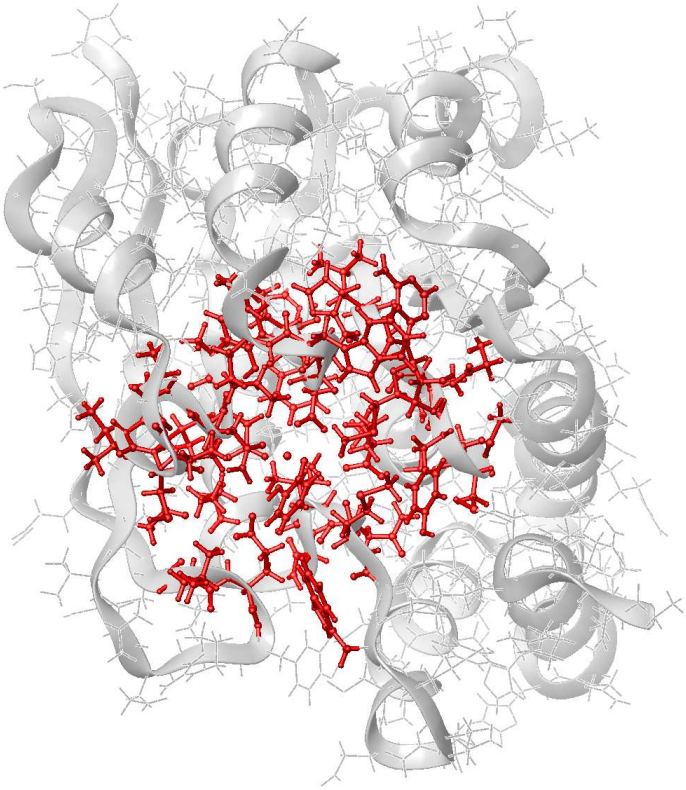

FIGURE 4.10: COMT protein including 3420 atoms. The QM region (highlighted in red) consists of 36 amino acid residues, 3 water molecules, an Mg^{2+} ion, the SAM, and the CAT molecules, which results in 571 QM atoms including 18 link atoms. See the text for more details regarding our selection approach.

At the construction and partitioning of the QM region we considered the distance of the atoms from the atoms involved in the reaction and the Mg^{2+} ion located near the reaction center, which is essential to the reaction. The amino acid residues were added to the QM part if one of their atoms was within 7.5 Å measured from the two atoms where the methyl group is transferred, the sulfur atom of the SAM or the oxygen atom of the CAT, or from the Mg^{2+} ion. Furthermore, these selections were modified by adding atoms with the intent to avoid the separation of peptide bonds in the cases where the selected residues are adjacent, and removing atoms otherwise to cut across bonds between the C_α and C_β atoms. Those amino acid residues were also included in the QM region that were out of the cutoff distance but both of their neighboring residues were added previously. This approach is resulted in 571 QM atoms including 18 link atoms with a net charge of -1. The system is illustrated in Fig. 4.10. Note that this scheme produced a residual charge of -0.619, which was transferred from the QM region to the MM part, nevertheless, the individual MM charges were practically unchanged.

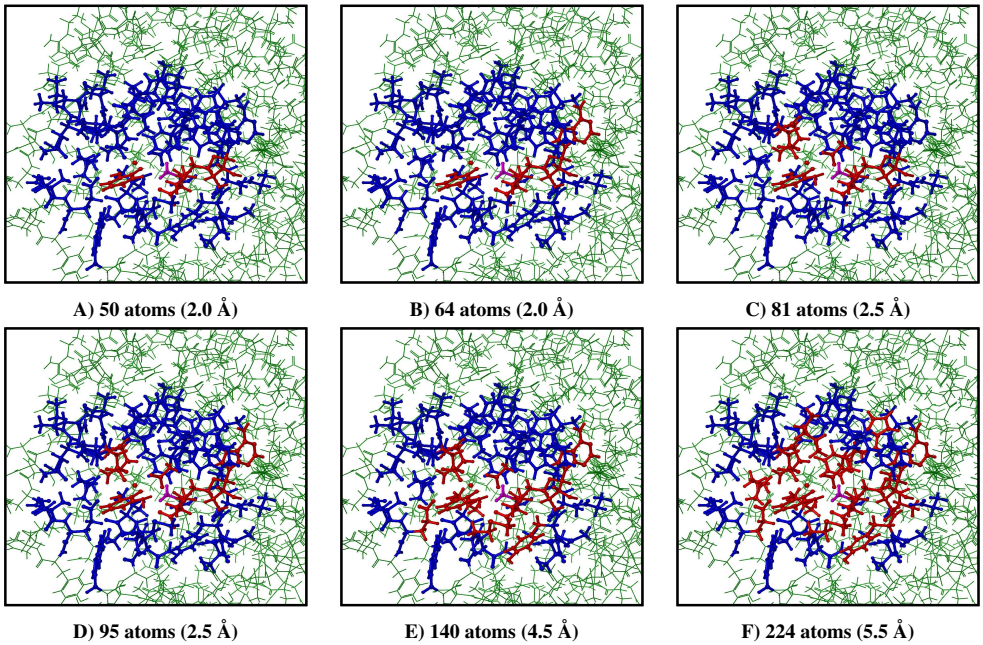

FIGURE 4.11: Visualization of the selected QM subsystems, which are shortly referenced as the corresponding capital letters. The MM, PBE-D3 (LMP2), and LNO-CCSD(T) methods are applied to the regions highlighted in green, blue, and red, respectively, in the case of the Huzinaga (multilevel local correlation) approach. As a reference point, the reacting methyl group is highlighted in purple. The region treated with LNO-CCSD(T) is gradually expanded from the reaction center toward the full QM region (see the text for details).

The partitioning of the QM region followed similar principles with slight modifications. If the selected residues were next to each other, the subsystem borders were set at the C_α atoms, because, in contrast to QM/MM, the residual charge distribution is not a factor in determining the QM/QM border. The applied cutoff distances were 2.0, 2.0, 2.5, 2.5, 4.5, and 5.5 Å, which resulted in 50, 64, 81, 95, 140, and 224 embedded QM atoms, respectively. The corresponding active subsystems will shortly be marked by letters A to F, respectively. Subsystems A and C are exceptions of the above partitioning scheme because the adenosyl group of the SAM molecule was eliminated from the selection to make these regions more compact. The QM subsystems are shown in Fig. 4.11, while the detailed list of QM atoms and embedded QM atoms can be found in the Supporting Information of Ref. 115.

4.4.2 Results

The errors of the various schemes are presented in Fig. 4.12. The DB approximation proves very accurate for this system as the DB error is only a few hundredths of a kcal/mol. The results obtained with the Huzinaga-embedding follow similar trends as we have seen in Sect. 4.3.3. The errors are less than 1.5 kcal/mol even if small embedded subsystems are used. Some error fluctuation can be observed in this region for the DMB scheme, but, for most subsystems, the errors are smaller compared to the parent large basis scheme. Interestingly, the errors of the DMB scheme are also much smaller in the case of subsystems A and C compared to the B and D partitionings, which suggests that the error cancellation is more fortunate if the embedded subsystem is more compact. Nevertheless, the errors decrease to a low level if the subsystem border is at least 5.5 Å away from the reaction center, which is in line with our previous findings. It is also important to note that the neglect of the D3 correction (not shown) increases the errors by about 1 kcal/mol for the small embedded regions, which proves the importance of the inclusion of dispersion interactions for embedding calculations even at the low-level method.

The performance of the multilevel local correlation scheme with the large basis set is excellent, its maximum error is very low, 0.3 kcal/mol, for any embedded subsystem, and the errors monotonically decrease beyond subsystem C. On the other hand, moderate error fluctuations occur in the case of the multilevel local correlation technique with DMB, but the approach is still very accurate as the maximum error does not exceed 0.5 kcal/mol. These attributes are also inherited by the Huzinaga multilevel local correlation method because their errors approximately equal those of the corresponding multilevel local correlation schemes shifted by the error of the Huzinaga-embedding with active subsystem F. The errors of the Huzinaga multilevel local correlation methods are still within 0.5 kcal/mol compared to the reference reaction energy, which shows that accurate reaction energies can be achieved by applying the highest-level technique only to a small region around the reaction center.

These results, in accordance with the conclusions for the small molecules, reveal that all the considered methods have tolerable accuracy even with relatively small embedded subsystems. Note also that the subsystem boundaries cross only one or two bonds for the small molecules, while here the subsystems are separated by many more bonds, but the error does not increase with the number of such bonds. The DFT/HF border is recommended to be at least 5 Å away from the reaction center in the case of the Huzinaga-embedding to minimize the ambiguity stemming from the error cancellation effects and to produce close to converged results. The errors of the Huzinaga multilevel local correlation method converge faster with respect to the subsystem size, but a domain of radius of 2.5 Å around the reaction center still has to be treated with the high-level method for accurate results. Taking into account these findings, to demonstrate what we can gain

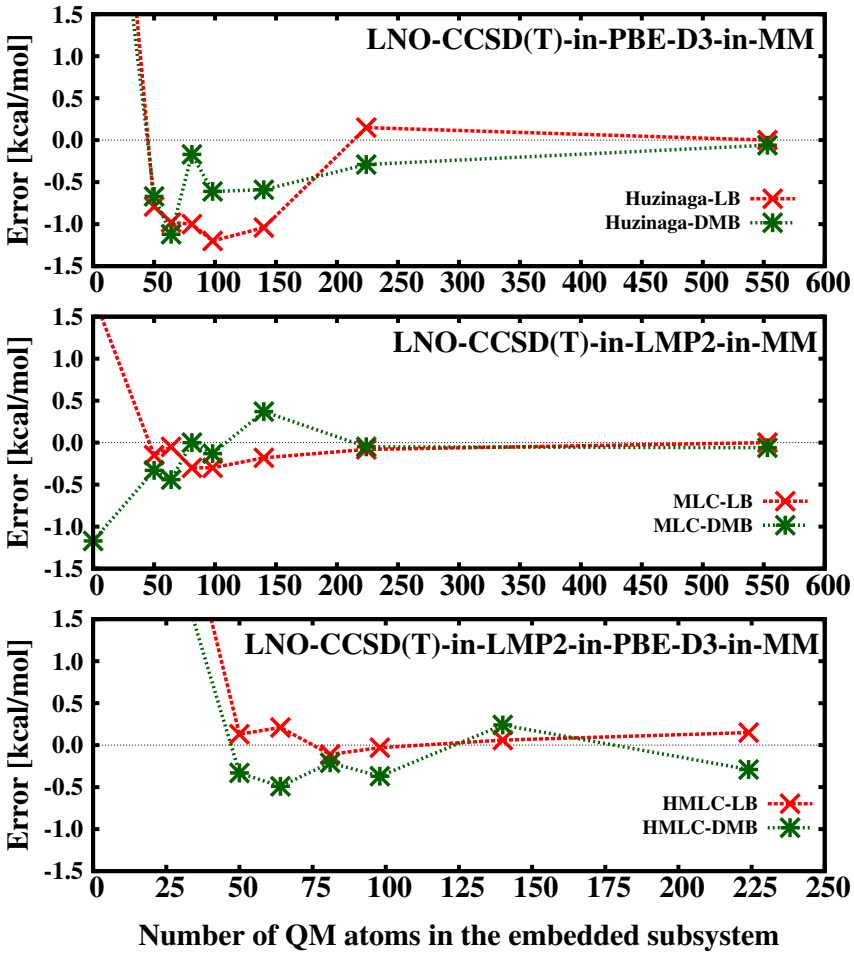

FIGURE 4.12: Error of the reaction energies for the methyl transfer reaction catalyzed by the COMT enzyme using various multilevel approaches as a function of the number of QM atoms included in the embedded subsystem. See Figs. 4.4 and 4.5 for the abbreviations. The cc-pVTZ and cc-pVDZ basis sets were used as large basis and small basis sets, respectively. In the Huzinaga-embedding based multilevel local correlation calculations the boundary of the DFT and HF regions was fixed, and the latter included the atoms of subsystem F. The reference LNO-CCSD(T)/cc-pVTZ reaction energy is -23.18 kcal/mol, and the error of the DB approximation is -0.06 kcal/mol.

with the various approximations, computation times were measured for the Huzinaga and multilevel local correlation approaches using the F and C partitionings, respectively. The timings for the Huzinaga multilevel local correlation approach were determined employing the LMP2 and LNO-CCSD(T) methods for subsystems F and C, respectively.

TABLE 4.1: Representetive timings for the various multilevel approaches. The reported computation times are wall-clock times measured on a 6-core 3.5 GHz Intel Xeon E5-1650 processor. The total computation times are decomposed into the DFT, HF, and correlation contributions. The errors and speedups are given with respect to the LNO-CCSD(T)-in-MM method using the cc-pVTZ basis set. See Figs. 4.4 and 4.5 for the abbreviations.

Method	LNO-CCSD(T)		Huzinaga		MLC		HMLC	
Basis set approach	LB	DB	LB	DMB	LB	DMB	LB	DMB
HF region	full	full	F	F	full	full	F	F
CCSD(T) region	full	full	F	F	C	C	C	C
DFT [day]	–	–	0.43	0.16	–	–	0.43	0.16
HF [day]	2.64	1.61	1.32	0.66	2.64	1.24	1.32	0.66
Correlation [day]	3.65	3.65	1.08	0.90	1.16	0.61	0.67	0.39
Total [day]	6.29	5.26	2.83	1.71	3.81	1.85	2.41	1.21
Speedup	–	1.20	2.23	3.68	1.65	3.39	2.61	5.20
Error [kcal/mol]	–	−0.06	0.15	−0.29	−0.30	0.00	−0.11	−0.21

The timings of the various schemes are presented in Table 4.1. The results show that a speedup of 1.6 can be achieved for the reference HF calculation by using the DB approximation without embedding approaches. At first glance this can be surprising because the cc-pVTZ basis set consists of more than twice as many AOs as the cc-pVDZ basis. The reason for the relatively low efficiency of the DB approximation is that small local fitting domains are used during the SCF iterations, and finally, to get an accurate HF energy, one iteration is performed with much larger fitting domains [69, 170]. This final step usually requires computation time comparable to that for the preceding SCF iterations altogether. In a DB HF calculation employing the local fitting approximation the single iteration in the large basis set must be carried out with the large fitting domains, thus only the expenses of the SCF iterations performed with small fitting domains can be saved. Nevertheless, the speedup for the HF step can be increased by the DMB ansatz to 2.1, while the speedup for the SCF part is 3.2 if the DMB HF approach is also combined with DFT via the Huzinaga-embedding. Concerning the overall speedup of the local correlation calculation, we can see that the DB approximation alone does not significantly affect the costs of the calculations, though it only results in a negligible loss of accuracy. The speedups are more pronounced for the multilevel schemes at the expense of 0.3 kcal/mol loss in accuracy. The DMB approach roughly doubles the speedup with respect to the corresponding multilevel scheme applying the large basis set. The most efficient but reliable method is clearly the Huzinaga multilevel local correlation technique, followed by the Huzinaga and the multilevel local correlation approach. We note that the potential speedups for the aforementioned methods can be even greater if larger errors can be afforded and a smaller domain is selected for the highest-level method.

4.5 Summary

In this chapter, an alternative dual basis scheme has been presented, which performs similarly to the previous approaches when applied to conventional electronic structure methods but is suitable for the generalization to embedding schemes. The approach has been utilized in practice to model the complexation of antioxidants in a cost-efficient way. Relying on the new approach two possible DB approximations have been proposed for the projector-based embedding method using the Huzinaga-equation. We have demonstrated that the DB approximation considerably increases the efficiency of embedding calculations in exchange for negligible loss of accuracy and also outperforms the notorious mixed-basis set approach. However, if the DB and mixed-basis approximations are combined, the resulting DMB scheme inherits the advantageous features of both and further improves the efficiency for both the WFT and the DFT embeddings. The DB and DMB schemes have also been tested for our multilevel correlation approach, where various local correlation methods are used for the different regions of the system. Our results show that the accuracy of the multilevel local correlation method is practically unaffected by the basis set approximations. The combination of the Huzinaga-embedding and the multilevel local correlation schemes has also been introduced, namely the Huzinaga multilevel local correlation approach. The benchmark calculations show that it is sufficient to only treat a small subsystem of the molecule at the CCSD(T) level if an MP2 layer is also inserted between the CCSD(T) and the DFT regions. For the CCSD(T) layer, if no aromatic system is broken, it is sufficient to include atoms separated by $2-3$ bonds from the reactive center of the molecule, while the MP2 region is recommended to consist of further atoms that are separated by $5-6$ bonds from the bonds affected in the reaction. In addition, the DMB approach can also be safely employed together with the Huzinaga multilevel local correlation method using the large basis set only for the atoms of the CCSD(T) layer. The proposed approaches have also been combined with QM/MM techniques and tested for the biologically relevant COMT system. The trends observed for the enzyme reaction verify our conclusions drawn for the small benchmark systems. This application also demonstrates that the proposed methods are useful for the accurate treatment of large molecular systems including thousands of atoms.

Chapter 5

Conclusions

In this thesis, the development of several approaches have been presented that are capable of modeling extended systems. The aim of our research was to develop new methods or improve the existing techniques. The work was centered around the atom- and quantum-based multilevel schemes.

The research began in the QM/MM methodology, where an important problem is the coupling of the MM and QM methods. The simple but heuristic link atom scheme is essentially the only choice in todays practical calculations, even though its use is limited to cases where the system is partitioned across single bonds, preferably carbon-carbon bonds. The frozen orbital-based LSCF approach is free from such constrains, however, its utilization is obstructed due to complicated orbital equations and parameter dependency. For the first time [80], we presented the implementation of the Huzinaga-equation-based HLSCF approach in the Amber and Mrcc codes, which makes it possible to apply WFT and DFT methods without orthogonalizing the basis set to the frozen orbitals separating the QM and MM subsystem [45]. A significant novelty of the program is the automatic generation of the frozen localized orbitals obtained from calculations performed for model molecules cut out of the system. The Amber-Mrcc code also allows the use of the link atom approach, thus the benchmarking of the different couplings are now possible. Sample calculations, utilizing the electronic embedding scheme, were performed to check the performance of both the HLSCF and the link atom approaches. It was found that the conceptually appealing HLSCF method is less accurate than the link atom method and further developments are necessary for practical applications.

Later, our attention was turned to the quantum-based multilevel approaches [99] to improve the efficiency of the QM calculations. Combining different levels of theory is difficult because the orbital orthogonality is not ensured automatically. Without it, the kinetic energy contains a nonadditive term which has to be approximated. Manby *et al.* [59] proposed a simple, exact, and projector-based embedding scheme which enforces orbital orthogonality and allows for the embedding of DFT and WFT methods

into lower-level DFT approaches, utilizing orbital localization. Instead of augmenting the Fock-matrix with the somewhat arbitrary level-shift operator, we proposed the use of the Huzinaga-equation, which strictly enforces the Pauli exclusion principle. Our local correlation method is also extended in the spirit of focused models: since its kernel splits up the system into local domains, simple multilevel theories can be defined if the domains of the active subsystem and the environment are treated at a different level. Several orbital selection algorithms were also designed to enable the use of the aforementioned techniques as black-box approaches. The considered projector-based Huzinaga-embedding and WFT-in-WFT schemes are benchmarked on simplistic, quasi-one dimensional test systems for reaction energies and compared with QM/MM and vacuum-embedding. We concluded that in the case of DFT-in-DFT, the Huzinaga-embedding scheme is more efficient than the other approaches, but QM/MM or even simple vacuum embedding is still competitive in particular tests. Concerning the embedding of WFT methods, the clear winner is the LNO-CCSD(T)-in-LMP2 approach, and the Huzinaga-embedding can only be more advantageous if a non-hybrid density functional is employed.

Based on their success, the embedding and multilevel schemes were further developed to increase their efficiency [115]. A dual basis approach, similar to the DB approximation of Head-Gordon [132], was designed for embedding calculations with the intent to reduce the computational expenses of the the DFT- and WFT-based methods. The new approach was applied to the Huzinaga-embedding, our WFT-in-WFT method, the combined Huzinaga multilevel local correlation approximations, and it was also used in an application [164]. The performance of the resulting DB DFT and WFT embedding methods was evaluated in extensive benchmark calculations on the aforementioned test set and also compared to that of the corresponding embedding schemes exploiting the mixed-basis approximation. Our results show that, with an appropriate combination of basis sets, the DB approach significantly speeds up the embedding calculations, and, for chemical processes where the electronic structure considerably changes, it is clearly superior to the mixed-basis approximation. The results also demonstrate that the DB approach, if integrated with the mixed-basis approximation, efficiently eliminates the major weakness of the latter, and the combination of the DB and mixed-basis schemes is the most efficient strategy to accelerate embedding calculations. The above findings were also confirmed by the test calculations on a realistic three-dimensional system with biological importance.

The presented approaches are available in the Amber and Mrcc programs, and they are already useful for the efficient validation of DFT-based QM/MM results and the highly accurate estimation of various properties, such as reaction energies and reaction barrier heights. The implementation of further molecular properties is in progress to widen the applicability of the methods.

Bibliography

[1] A. Ganyecz, M. Kállay and J. Csontos, J. Chem. Theory Comput. **13**, 4193 (2017).

[2] A. D. Boese, M. Oren, O. Atasoylu, J. M. L. Martin, M. Kállay and J. Gauss, J. Chem. Phys. **120**, 4129 (2004).

[3] The Nobel Prize in Chemistry 1998. Nobelprize.org. Nobel Media AB 2014. Web. 3 Aug 2018. http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1998.

[4] The Nobel Prize in Chemistry 2013. Nobelprize.org. Nobel Media AB 2014. Web. 3 Aug 2018. http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2013/.

[5] T. Helgaker, P. Jørgensen and J. Olsen. *Molecular Electronic Structure Theory*. Wiley, Chichester, 2000.

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

[7] J. Gauss, Coupled-cluster theory, in *Encyclopedia of Computational Chemistry*, edited by P. R. Schleyer, W. L. Jorgensen, H. F. Schaefer III, P. R. Schreiner and W. Thiel. Wiley, New York, 1998, p. 615.

[8] G. D. Purvis III and R. J. Bartlett, J. Chem. Phys. **76**, 1910 (1982).

[9] K. Raghavachari, G. W. Trucks, J. A. Pople and M. Head-Gordon, Chem. Phys. Lett. **157**, 479 (1989).

[10] T. Vreven, K. S. Byun, I. Komáromi, S. Dapprich, J. A. Montgomery Jr., K. Morokuma and M. J. Frisch, J. Chem. Theory Comput. **2**, 815 (2006).

[11] C. M. Isborn, A. W. Götz, M. A. Clark, R. C. Walker and T. J. Martínez, J. Chem. Theory Comput. **8**, 5092 (2012).

[12] J. Tomasi, B. Mennucci and R. Cammi, Chem. Rev. **105**, 2999 (2005).

[13] B. Mennucci, WIREs Comput Mol Sci. **2**, 386 (2012).

[14] D. Hessz, B. Hégely, M. Kállay, T. Vidóczy and M. Kubinyi, J. Phys. Chem. A. **118**, 5238 (2014).

[15] R. Improta, V. Barone, G. Scalmani and M. J. Frisch, J. Chem. Phys. **125**, 054103 (2006).

[16] G. Parisio, A. Marini, A. Biancardi, A. Ferrarini and B. Mennucci, J. Phys. Chem. B. **115**, 9980 (2011).

[17] A. Warshel and M. Karplus, J. Am. Chem. Soc. **94**, 5612 (1972).

[18] A. Warshel and M. Levitt, J. Mol. Biol. **103**, 227 (1976).

[19] A. R. Leach. *Molecular Modelling: Principles and Applications. Second Edition.* Dorling Kindersley (India) Pvt. Ltd., 14 Local Shopping Centre, Panchsheel Park, New Delhi 110017, India, 2010.

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

[21] R. A. Buckingham, Proc. Royal Soc. A. **168**, 264 (1938).

[22] T. L. Hill, J. Chem. Phys. **16**, 399 (1948).

[23] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, J. Chem. Theory Comput. **11**, 3696 (2015).

[24] J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. de Groot, H. Grubmüller and A. D. MacKerell Jr., Nat. Methods. **14**, 71 (2017).

[25] M. M. Reif, P. H. Hünenberger and C. Oostenbrink, J. Chem. Theory Comput. **8**, 3705 (2012).

[26] M. J. Robertson, J. Tirado-Rives and W. L. Jorgensen, J. Chem. Theory Comput. **11**, 3499 (2015).

[27] H. Heinz, T.-J. Lin, R. K. Mishra and F. S. Emami, Langmuir. **29**, 1754 (2013).

[28] A. C. T. van Duin, S. Dasgupta, F. Lorant and W. A. Goddard, J. Phys. Chem. A. **105**, 9396 (2001).

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

[30] L. Cao and U. Ryde, Front. Chem. **6**, 89 (2018).

[31] L. W. Chung, W. M. C. Sameera, R. Ramozzi, A. J. Page, M. Hatanaka, G. P. Petrova, T. V. Harris, X. Li, Z. Ke, F. Liu, H.-B. Li, L. Ding, , and K. Morokuma, Chem. Rev. **115**, 5678 (2015).

[32] Y. Zhang and H. Lin, J. Chem. Theory Comput. **4**, 414 (2008).

[33] G. Jindal and A. Warshel, J. Phys. Chem. B. **120**, 9913 (2016).

[34] H. J. Kulik, J. Zhang, J. P. Klinman and T. J. Martínez, J. Phys. Chem. B. **120**, 11381 (2016).

[35] H. Lin and D. G. Truhlar, J. Phys. Chem. A. **109**, 3391 (2005).

[36] D. Das, K. P. Eurenius, E. M. Billings, P. Sherwood, D. C. Chatfield, M. Hodošček and B. R. Brooks, J. Chem. Phys. **117**, 10534 (2002).

[37] N. Reuter, A. Dejaegere, B. Maigret and M. Karplus, J. Phys. Chem. A. **104**, 1720 (2000).

[38] M. S. Gordon, M. A. Freitag, P. Bandyopadhyay, J. H. Jensen, V. Kairys and W. J. Stevens. **105**, 293 (2001).

[39] G. A. DiLabio, M. M. Hurley and P. A. Christiansen, J. Chem. Phys. **116**, 9578 (2002).

[40] C. Xiao and Y. Zhang, J. Chem. Phys. **127**, 124102 (2007).

[41] D. M. Philipp and R. A. Friesner, J. Comput. Chem. **20**, 1468 (1999).

[42] R. B. Murphy, D. M. Philipp and R. A. Friesner, J. Comput. Chem. **21**, 1442 (2000).

[43] J. Pu, J. Gao and D. G. Truhlar, J. Phys. Chem. A. **108**, 632 (2004).

[44] X. Assfeld and J.-L. Rivail, Chem. Phys. Lett. **263**, 100 (1996).

[45] G. G. Ferenczy, J. Comput. Chem. **34**, 854 (2013).

[46] G. G. Ferenczy, J. Comput. Chem. **34**, 862 (2013).

[47] C. R. Jacob and J. Neugerbauer, WIREs Comput Mol Sci. **4**, 325 (2014).

[48] P. R. Nagy, G. Samu and M. Kállay, J. Chem. Theory Comput. (2018).

[49] A. M. Elsohly, C. L. Shaw, M. E. Guice, B. D. Smith and G. S. Tschumper, Mol. Phys. **105**, 2777 (2007).

[50] Q. Sun and G. K.-L. Chan, Acc. Chem. Res. **49**, 2705 (2016).

[51] T. A. Wesolowski and A. Warshel, J. Phys. Chem. **97**, 8050 (1993).

[52] K. Kiewisch, G. Eickerlinga, M. Reiher and J. Neugebauer, J. Chem. Phys. **128**, 044114 (2008).

[53] T. A. Wesolowski and J. Weber, Chem. Phys. Lett. **248**, 71 (1996).

[54] A. S. P. Gomes and C. R. Jacob, Annu. Rep. Prog. Chem. C. **108**, 222 (2012).

[55] A. W. Götz, S. M. Beyhan and L. Visscher, J. Chem. Theory Comput. **5**, 3161 (2009).

[56] T. Klüner, N. Govind, Y. A. Wang and E. A. Carter, Phys. Rev. Lett. **86**, 5954 (2001).

[57] J. Neugebauer, C. Curutchet, A. Muñoz-Losa and B. Mennucci, J. Chem. Theory Comput. **6**, 1843 (2010).

[58] J. D. Goodpaster, T. A. Barnes and T. F. Miller III, J. Chem. Phys. **134**, 164108 (2011).

[59] F. R. Manby, M. Stella, J. D. Goodpaster and T. F. Miller III, J. Chem. Theory Comput. **8**, 2564 (2012).

[60] K. Raghavachari and A. Saha, Chem. Rev. **115**, 5643 (2015).

[61] H.-J. Werner and M. Schütz, J. Chem. Phys. **135**, 144116 (2011).

[62] S. F. Boys, Rev. Mod. Phys. **32**, 296 (1960).

[63] J. Pipek and P. Mezey, J. Chem. Phys. **90**, 4916 (1989).

[64] C. Ochsenfeld, J. Kussmann and D. S. Lambrecht, Reviews in Computational Chemistry (eds K. B. Lipkowitz and T. R. Cundari). (2007).

[65] W. Li, P. Piecuch, J. R. Gour and S. Li, J. Chem. Phys. **131**, 114109 (2009).

[66] P. Pulay, Chem. Phys. Lett. **100**, 151 (1983).

[67] J. W. Boughton and P. Pulay, J. Comput. Chem. **14**, 736 (1993).

[68] Y. Jung, Y. Shao and M. Head-Gordon, J. Comput. Chem. **28**, 1953 (2007).

[69] P. R. Nagy, G. Samu and M. Kállay, J. Chem. Theory Comput. **12**, 4897 (2016).

[70] Z. Tóth, P. R. Nagy, P. Jeszenszki and Á. Szabados, Theor. Chem. Acc. **134**, 100 (2015).

[71] Z. Rolik, L. Szegedy, I. Ladjánszki, B. Ladóczki and M. Kállay, J. Chem. Phys. **139**, 094105 (2013).

[72] H. Eshuis, J. E. Bates and F. Furche, Theor. Chem. Acc. **131**, 1084 (2012).

[73] A. Grüneis, M. Marsman, J. Harl, L. Schimka and G. Kresse, J. Chem. Phys. **131**, 154115 (2009).

[74] G. Samu and M. Kállay, J. Chem. Phys. **146**, 204101 (2017).

[75] R. A. Mata, H.-J. Werner and M. Schütz, J. Chem. Phys. **128**, 144106 (2008).

[76] W. Li and P. Piecuch, J. Phys. Chem. A. **114**, 8644 (2010).

[77] Z. Rolik and M. Kállay, J. Chem. Phys. **135**, 104111 (2011).

[78] M. Sparta, M. Retegan, P. Pinski, C. Riplinger, U. Becker and F. Neese, J. Chem. Theory Comput. **13**, 3198 (2018).

[79] D. J. Coughtrie, R. Giereth, D. Kats, H.-J. Werner and A. Köhn, J. Chem. Theory Comput. **14**, 693 (2018).

[80] B. Hégely, F. Bogár, G. G. Ferenczy and M. Kállay, Theor. Chem. Acc. **134**, 132 (2015).

[81] D. A. Case, D. S. Cerutti, T. E. Cheatham III, T. A. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Götz, D. Greene, N. Homeyer, S. Izadi, A. Kovalenko, T. S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. Mermelstein, K. M. Merz, G. Monard, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D. R. Roe, A. Roitberg, C. Sagui, C. L. Simmerling, W. M. Botello-Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, L. Xiao, D. M. York, and P. A. Kollman. Amber 2017. University of California, San Francisco, 2017.

[82] R. Salomon-Ferrer, D. A. Case and R. C. Walker.

[83] Mrcc, a quantum chemical program suite written by M. Kállay, Z. Rolik, J. Csontos, P. Nagy, G. Samu, D. Mester, J. Csóka, B. Szabó, I. Ladjánszki, L. Szegedy, B. Ladóczki, K. Petrov, M. Farkas, P. D. Mezei, and B. Hégely. See also Ref. 71 as well as http://www.mrcc.hu/.

[84] H. L. Woodcock III, M. Hodošček, A. T. B. Gilbert, P. M. W. Gill, H. F. Schaefer III and B. R. Brooks, J. Comput. Chem. **28**, 1485 (2007).

[85] J. Torras, G. de M. Seabra, E. Deumens, S. B. Trickey and A. E. Roitberg, J. Comput. Chem. **29**, 1564 (2008).

[86] Y. Hagiwara, T. Ohta and M. Tateno, J. Phys. Condens. Matter. **21**, 064234 (2009).

[87] S. Metz, J. Kästner, A. A. Sokol, T. W. Keal and P. Sherwood, WIREs Comput. Mol. Sci. **4**, 101 (2014).

[88] A. Götz, M. A. Clark and R. C. Walker, J. Comput. Chem. **35**, 95 (2014).

[89] S. Riahi and C. N. Rowley, J. Comput. Chem. **35**, 2076 (2014).

[90] S. Huzinaga and A. A. Cantu, J. Chem. Phys. **55**, 5543 (1971).

[91] W. J. Hehre, R. Ditchfield and J. A. Pople, J. Chem. Phys. **56**, 2257 (1972).

[92] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, J. Comput. Chem. **25**, 1157 (2004).

[93] J. Wang, W. Wang, P. A. Kollman and D. A. Case, J. Mol. Graph. Model. **25**, 247 (2006).

[94] A. Jakalian, B. L. Bush, D. B. Jack and C. I. Bayly, J. Comput. Chem. **21**, 132 (2000).

[95] A. Jakalian, D. B. Jack and C. I. Bayly, J. Comput. Chem. **23**, 1623 (2002).

[96] Schrödinger Release 2015–1: Maestro, version 10.1, Schrödinger, LLC, New York, NY, 2015.

[97] T. Williams and C. Kelley (2013) Gnuplot: an interactive plotting program. http://www.gnuplot.info.

[98] Marvin was used for drawing and displaying chemical structures, Marvin 9.22.0, 2014, ChemAxon. http://www.chemaxon.com.

[99] B. Hégely, P. R. Nagy, G. G. Ferenczy and M. Kállay, J. Chem. Phys. **145**, 064107 (2016).

[100] M. E. Fornace, J. Lee, K. Miyamoto, F. R. Manby and T. F. Miller III, J. Chem. Theory Comput. **11**, 568 (2015).

[101] J. D. Goodpaster, T. A. Barnes, F. R. Manby and T. F. Miller III, J. Chem. Phys. **140**, 18A507 (2014).

[102] P. C. Hariharan and J. A. Pople, Theor. Chim. Acta. **28**, 213 (1973).

[103] R. Krishnan, J. S. Binkley, R. Seeger and J. A. Pople, J. Chem. Phys. **72**, 650 (1980).

[104] T. H. Dunning Jr., J. Chem. Phys. **90**, 1007 (1989).

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

[106] D. E. Woon and T. H. Dunning Jr., J. Chem. Phys. **98**, 1358 (1993).

[107] F. Weigend, J. Comput. Chem. **29**, 167 (2008).

[108] F. Weigend, A. Köhn and C. Hättig, J. Chem. Phys. **116**, 3175 (2002).

[109] P. A. M. Dirac, Proc. Roy. Soc. (London) A. **123**, 714 (1929).

[110] J. C. Slater, Phys. Rev. **81**, 385 (1951).

[111] J. P. Perdew, K. Burke and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[112] A. D. Becke, Phys. Rev. A. **38**, 3098 (1988).

[113] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).

[114] D. A. Case, V. Babin, J. T. Berryman, R. M. Betz, Q. Cai, D. S. Cerutti, T. E. Cheatham III, T. A. Darden, R. E. Duke, H. Gohlke, A. W. Göetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T. S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K. M. Merz, F. Paesani, D. R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C. L. Simmerling, W. Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, and P. A. Kollman. AMBER 14. University of California, San Francisco, 2014.

[115] B. Hégely, P. R. Nagy and M. Kállay, J. Chem. Theory Comput. **14**, 4600 (2018).

[116] C. W. Bauschlicher Jr. and P. S. Bagus, Chem. Phys. Lett. **90**, 355 (1982).

[117] D. B. Chesnut and K. D. Moore, J. Comput. Chem. **10**, 648 (1989).

[118] M. Orozco and F. J. Luque, Chem. Phys. Lett. **160**, 305 (1989).

[119] J. F. Hinton, P. Guthrie, P. Pulay and K. Wolinski, J. Am. Chem. Soc. **114**, 1604 (1992).

[120] R. A. Kirby and A. E. Hansen, Int. J. Quantum Chem. **57**, 199 (1996).

[121] G. A. DiLabio and J. S. Wright, Chem. Phys. Lett. **297**, 181 (1998).

[122] G. A. DiLabio, J. Phys. Chem. A. **103**, 11414 (1999).

[123] S. Moon and D. A. Case, J. Comput. Chem. **27**, 825 (2006).

[124] D. M. Reid, R. Kobayashi and M. A. Collins, J. Chem. Theory Comput. **10**, 146 (2014).

[125] S. J. R. Lee, K. Miyamoto, F. Ding, F. R. Manby and T. F. Miller III, Chem. Phys. Lett. **683**, 375 (2017).

[126] S. Havriliak and H. F. King, J. Am. Chem. Soc. **105**, 4 (1983).

[127] R. Jurgens-Lutovsky and J. Almlöf, Chem. Phys. Lett. **178**, 451 (1991).

[128] K. Wolinski and P. Pulay, J. Chem. Phys. **118**, 9497 (2003).

[129] A. Ksiazek and K. Wolinski, Mol. Phys. **106**, 769 (2008).

[130] M. S. Lee and M. Head-Gordon, J. Chem. Phys. **107**, 9085 (1997).

[131] M. S. Lee and M. Head-Gordon, Comput. Chem. **24**, 295 (2000).

[132] W. Liang and M. Head-Gordon, J. Phys. Chem. A. **108**, 3206 (2004).

[133] R. P. Steele, R. A. DiStasio Jr., Y. Shao, J. Kong and M. Head-Gordon, J. Chem. Phys. **125**, 074108 (2006).

[134] R. P. Steele, Y. Shao, R. A. DiStasio and M. Head-Gordon, J. Phys. Chem. A. **110**, 13915 (2006).

[135] R. P. Steele and M. Head-Gordon, Mol. Phys. **105**, 2455 (2007).

[136] R. A. DiStasio Jr., R. P. Steele and M. Head-Gordon, Mol. Phys. **105**, 2731 (2007).

[137] R. P. Steele, R. A. DiStasio Jr. and M. Head-Gordon, J. Chem. Theory Comput. **5**, 1560 (2009).

[138] R. P. Steele, M. Head-Gordon and J. C. Tully, J. Phys. Chem. A. **114**, 11853 (2010).

[139] Y. Mao, P. R. Horn, N. Mardirossian, T. Head-Gordon, C.-K. Skylaris and M. Head-Gordon, J. Chem. Phys. **145**, 044109 (2016).

[140] M. S. Marshall, J. S. Sears, L. A. Burns, J.-L. Brédas and C. D. Sherrill, J. Chem. Theory Comput. **6**, 3681 (2010).

[141] J. Deng, A. T. B. Gilbert and P. M. W. Gill, J. Chem. Phys. **130**, 231101 (2009).

[142] J. Deng, A. T. B. Gilbert and P. M. W. Gill, J. Chem. Phys. **133**, 044116 (2010).

[143] J. Deng and P. M. W. Gill, J. Chem. Phys. **134**, 081103 (2011).

[144] J. Deng, A. T. B. Gilbert and P. M. W. Gill, J. Chem. Theory Comput. **11**, 1639 (2015).

[145] T. Nakajima and K. Hirao, J. Chem. Phys. **124**, 184108 (2006).

[146] S. Tokura, T. Sato, T. Tsuneda, T. Nakajima and K. Hirao, J. Comput. Chem. **29**, 1187 (2008).

[147] W. Mizukami, T. Nakajima, K. Hirao and T. Yanai, Chem. Phys. Lett. **508**, 177 (2011).

[148] J. Deng, A. T. B. Gilbert and P. M. W. Gill, Phys. Chem. Chem. Phys. **12**, 10759 (2010).

[149] M. Kobayashi and H. Nakai, Int. J. Quantum Chem. **109**, 2227 (2009).

[150] J. Friedrich and M. Dolg, J. Chem. Theory Comput. **5**, 287 (2009).

[151] J. Zhang and M. Dolg, J. Chem. Phys. **140**, 044114 (2014).

[152] J. Zhang and M. Dolg, J. Chem. Theory Comput. **11**, 962 (2015).

[153] T. Anacker, D. P. Tew and J. Friedrich, J. Chem. Theory Comput. **12**, 65 (2016).

[154] I. Røeggen and B. Gao, J. Chem. Phys. **148**, 134118 (2018).

[155] T. B. Adler, G. Knizia and H.-J. Werner, J. Chem. Phys. **127**, 221106 (2007).

[156] J. Noga, S. Kedžuch and J. Šimunek, J. Chem. Phys. **127**, 034106 (2007).

[157] A. Köhn and D. P. Tew, J. Chem. Phys. **132**, 024101 (2010).

[158] L. Kong and E. F. Valeev, J. Chem. Phys. **133**, 174126 (2010).

[159] F. Ding, F. R. Manby and T. F. Miller III, J. Chem. Theory Comput. **13**, 1605 (2017).

[160] K. Miyamoto, T. F. Miller III and F. R. Manby, J. Chem. Theory Comput. **12**, 5811 (2016).

[161] M. Kállay, J. Chem. Phys. **142**, 204105 (2015).

[162] P. R. Nagy and M. Kállay, J. Chem. Phys. **146**, 214106 (2017).

[163] J. Gomes, M. Head-Gordon and A. T. Bell. **118**, 21409 (2014).

[164] B. Kirschweng, D. M. Tilinger, B. Hégely, G. Samu, D. Tátraaljai, E. Földes and B. Pukánszky, Eur. Polym. J. **103**, 228 (2018).

[165] S. J. Bennie, M. W. van der Kamp, R. C. R. Pennifold, M. Stella, F. R. Manby and A. J. Mulholland, J. Chem. Theory Comput. **12**, 2689 (2016).

[166] X. Zhang, S. J. Bennie, M. W. van der Kamp, D. R. Glowacki, F. R. Manby and A. J. Mulholland, Royal Soc. Open Sci. **5**, 171390 (2018).

[167] W. L. Jorgensen, J. Chandrasekhar, J. Madura and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).

[168] S. Grimme, J. Antony, S. Ehrlich and H. Krieg, J. Chem. Phys. **132**, 154104 (2010).

[169] S. Grimme, S. Ehrlich and L. Goerigk, J. Comput. Chem. **32**, 1456 (2011).

[170] R. Polly, H.-J. Werner, F. R. Manby and P. J. Knowles, Mol. Phys. **102**, 2311 (2004).

[171] J. M. Foster and S. F. Boys, Rev. Mod. Phys. **32**, 300 (1960).

Declaration

I, Bence Hégely hereby declare that this PhD dissertation is the result of my own work. All content that was taken from other sources –either literally or reworded– is clearly indicated, and its source is presented in the reference list.

Budapest, February 25, 2019 Bence Hégely