

PHYSICAL PRINCIPLES OF CONSTRUCTING HYBRID QM/MM PROCEDURES

A.L. TCHOUGRÉEFF

*Karpov Institute of Physical Chemistry
10 Vorontsovo pole, 105064, Moscow, Russia
andrei@cc.nifhi.ac.ru*

A.M. TOKMACHEV

*Karpov Institute of Physical Chemistry
10 Vorontsovo pole, 105064, Moscow, Russia
tokm@hcc.keldysh.ru*

Abstract Hybrid QM/MM methods are widely used for describing different aspects of behavior of complex molecular systems. The key problem when applying the QM/MM methodology is a substantiated construction/selection of the junction between the parts of the system described at the QM and MM levels, respectively. It is especially important in the case of covalently bound QM and MM subsystems. We pursue here a general approach based on a sequential separation of electronic variables in order to develop a fundamental form of the intersubsystem junction. Special attention is given to construction of frontier one-electron states and renormalization of QM Hamiltonian parameters and MM force fields. From this point of view we consider a series of the junction forms present in the literature and in some cases suggest theoretically more reliable alternatives. General theoretical conclusions are supported by data of numerical experiments.

Keywords: QM/MM methods, junction, physical principles, APSLG scheme, effective Hamiltonian

1. Introduction

Modern computational quantum chemistry tends to cover with acceptable precision molecular systems of real interest. In the framework of *ab initio* methodology achieving a good quality results is usually concerned with extending basis sets of one-electron states and with explicit taking into account a great deal of electron correlation. It leads to ex-

tremely high computational costs for medium and large size systems due to unpleasant scalability of computational resources like $N^4 \div N^7$ (where N is the dimension of a space spanned by one-electron basis functions). Even in the case of semiempirical methods the computational costs grow as N^3 with growth of system size. This is a serious problem for application of quantum chemical procedures for calculating properties of many practically important molecular systems and especially of their chemical transformations [1]. The problem of high computational costs is especially actual when large systems of biological importance are considered or massive calculations of potential energy surfaces (PESs) are necessary, for example, in the framework of molecular dynamics simulations.

In the literature different means are proposed to significantly reduce computational costs without deteriorating the quality of obtained results. The first way is to smooth the dependence of computational costs on the size of the system. This type approaches usually exploit the localization of electronic degree of freedom, based on the "principle of nearsightedness" [2], or the exponential decay of the one-electron density matrix elements [3]. The almost linear scalability is not impossible and, for example, authors of Ref. [4] have shown that the scalability of the order $N^{1.3}$ can be achieved. There is a reasonable number of successful attempts to achieve optimal N -scalability properties for the whole spectrum of quantum chemical schemes – *ab initio*, DFT, and semiempirical. The non-variational schemes are based on the "divide-and-conquer" methodology [5, 6, 7], Fermi operator expansion [8, 9], energy renormalization group [10] and recursion technique [11]. The variational approaches [12, 13, 14, 15] are based on the substitution of energy minimization by the grand canonical potential minimization [16]. These techniques are thoroughly reviewed in Refs. [17, 18]. At the same time these methods are mostly oriented on the tight-binding models of solids or solid clusters. The density matrix renormalization group method [19] is quite important achievement since it competes in quality with most elaborated methods of conventional quantum chemistry [20]. We should stress that the use of local one-electron basis states can significantly reduce computational costs [21]. In this context direct determination of localized Hartree-Fock orbitals can be especially useful and workable [22]. Significant acceleration of computations can be achieved by using pseudodiagonalization procedures [23] or by special choice of the trial electronic wave function [24, 25, 26, 27] alternative to the standard SCF form. The application of the methods with good scalability properties to enzyme reactions, structure and properties of proteins and DNA is straightforward [28, 29].

The molecular mechanics (MM) [30] is another good candidate for using in calculations of properties of large molecular systems. Currently it is extensively used in the field of biochemical simulations. At the same time there are significant limitations for its use. Generally, they are consequences of classical nature of the MM force fields [31]. For example, the application of MM schemes to chemical reactions including the processes of bond formation or cleavage, transition states, or to highly correlated compounds like transition metal complexes is impossible or, at least, questionable since the main prerequisite for the validity of MM procedures is its application to the ground electronic state of a molecule without low-lying excitations. Moreover, the MM schemes usually require rather complex parameterization procedure. The main advantages of the MM schemes are their low cost and high efficiency in the prediction of molecular geometry for organic compounds without significant electron correlation. The principal advantage of the QM procedures is a wide range of problems they potentially apply. The concert exploitation of advantages of the both (quantum and classical) approaches can be achieved by the incorporation of quantum mechanical (QM) description to the MM framework. This methodology, alternative to construction of the QM schemes with the $O(N)$ scalability, leads to hybrid quantum mechanical/molecular mechanical (QM/MM) schemes allowing another way to bypass the bottleneck of N^n -scalability. The QM/MM schemes describe some relatively small part of the system where chemical transformations take place by an appropriate quantum chemical method while the rest (relatively inert environment) is covered by classical force fields (molecular mechanics). The practical usefulness and general validity of these approaches is based on the chemically motivated observation: chemical transformations usually affect only a small part of the whole system (reaction center) while the rôle of surrounding groups and molecules is reduced to modification of the PES due for example to some polarization or steric strains. This situation is characteristic for chemical reactions of biological interest (especially, for catalysis by enzymes).

Since the seminal work by Warshell and Levitt [32] the hybrid QM/MM schemes of calculating large molecular systems acquired an increasing popularity. There is a big variety of different hybrid approaches described in the literature [33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43]. Even more, numerous cases of separating electronic variables like π -electron models [44] or even taking into account explicitly only valence electrons in semiempirical methods [45] can be considered as special cases of hybrid schemes since they also bear the family marks of the QM/MM approach, namely, (i) the separation of the system into parts, and (ii)

treating these parts on quantum or classical levels, respectively. Very close to the QM/MM modeling is a problem of constructing effective fragment (group) potentials which can affect the quantum chemically described part of molecule modeling the influence of environment [46, 47, 48, 49, 50, 52, 53, 54, 51]. In such a broader sense several other problems in the area of computational chemistry seem to be related to the QM/MM context: these are the problem of embedding in the cluster calculations on solids and their surfaces with special attention to adsorption and catalysis problems; the problem of description of solute/solvent effects for reactions in condensed media; the pseudopotential descriptions of molecular and atomic electronic structure to list several most important. Usually all these areas are considered separately without an attempt of establishing relations among them. Also a great variety of different specific schemes referred to as "protocols" implemented in different packages are normally considered only from the point of view of their practical feasibility and their fit for particular applied purpose, rather than in a context of their exact placement among other approximate methods and of evaluation of relative precision of that or another approximation.

The QM/MM modeling can be placed in the general context of combination of different level descriptions. We should mention that the hybrid QM/QM schemes are well known in the literature. Some schemes by construction are far enough from the QM/MM construction since the separation on parts is performed not on the geometrical principle but on the principle of necessity of correlated description to some delocalized orbitals. Well known examples of such an approach is the complete active space self consistent field (CASSCF) scheme [55] and GVBCAS method [56] combining generalized valence bond and CASSCF approaches which can be considered as three-level scheme. The effective crystal field method developed for calculation of d-d spectra [57] and its extension designed to treat catalytical systems [58] can be considered as taking an intermediate position in this hierarchy since they use some strictly localized one-electron states (for example, *d*-orbitals of one transition metal ion) as a high-level subsystem treating it on the full configuration interaction level with delocalization corrections taken into account perturbatively. The QM/QM schemes similar by construction to QM/MM ones (geometric separation on parts) are also quite popular and should be specially mentioned since they have many common problems. For example, the IMOMO method [59, 60, 61, 62] integrating molecular orbital approximations of different level is very similar in the way of construction, main principles and physical contributions taken into account to the QM/MM IMOMM method [38]. Another example

of such a construction is the QM/DIM (diatomics-in-molecules) scheme [63] which takes the interactions between subsystems in the first order of perturbation theory. This scheme can be brought to a type of QM/MM one by relevant significant simplifications [63, 64]: by ignoring, first, the differences in potential curves of diatomic fragments due to angular and spin momentum, second, the couplings between the states of the same symmetry, and multicenter electrostatic contributions. The scheme proposed by Naray-Szabo *et al* [65, 66] is especially interesting here. It uses strictly localized bond orbitals (SLBOs) for the environment and the SCF procedure for the fragment of particular interest. Note that in the case of large systems the calculation of SLBOs can be very time-consuming process, so the transition to the MM scheme for the environment is an important and almost unavoidable step for calculations of biopolymers.

It should be noted that the hybrid quantum/classical schemes apply not only for determination of geometries, energies, and reaction mechanisms. The Monte Carlo [67, 68] and molecular dynamics (MD) [69, 70, 71, 72] simulations are quite popular as frameworks for which various QM/MM procedures serve as "subroutines". Before employing hybrid schemes the large-scale MD simulations were performed only with low-level approximations for force fields. The use of hybrid schemes extends significantly the scope of their application, improve precision of the results that allows to improve the understanding of statistical properties and dynamical processes in liquids and biopolymers.

The area of the QM/MM modeling is not a steady one since there are many unsolved problems mostly caused by *ad hoc* character of construction of junction between subsystems. Despite a general claim of being stipulated by "specific practical needs" the *ad hoc* junction constructs present in the literature frequently lead to more or less serious technical ("practical") problems in construction of hybrid schemes. This point is stressed in Refs. [73, 74]. The most direct application of hybrid schemes is in the case of solute/solvent interactions since the border between quantum and classical parts of the whole system in this case is naturally defined by the division of the whole system into individual molecules or ions. In this case the environment (classical) variables can be chosen in many different ways, ranging from continual models with special properties [72, 75, 76, 77, 78] to the discrete ones explicitly employing the information on the structure of solvent molecules [79, 80, 81, 82] through some intermediate schemes adopting advantages of both extrema [83, 84] are also possible. Nevertheless, even in such a transparent case the problems occur when the formation of complexes between solute metal ions and solvent molecules has to be considered. In this case some solvent

molecules may have to be considered either as part of QM system when form a close contact with the metal ion or as a part of the MM system when they are absorbed by the solvent bulk. Approaches to constructing relevant force fields in this situation are discussed in Ref. [85]. Another important area of applications is solid state and surface chemistry [37, 39, 86, 87] with catalysis and adsorption by oxide systems as the most popular objects. In this case the subsystems are very polar and usually charged, so a special attention should be paid to electrostatic interactions between subsystems. The most important problem in the field of embedding is the unphysical polarization of the QM subsystem (cluster) and the related necessity to adjust effective charges on the ions to obtain sensible results. It should be noted that the most difficult and not trivial case of separating quantum and classical subsystems is that of covalently bound systems since the electrons in chemical bonds connecting two regions are highly correlated. At the same time this case seems to be very actual (especially for biological applications) and requires thorough consideration. In this paper we consider the hybrid QM/MM schemes with covalent linkage between regions in more details, give account of the "state-of-art" in this field, and show how the junction between quantum and classical subsystems can be constructed in a consistent (not *ad hoc*) way by deriving it from an underlying QM description with an emphasis on the choice and refinement of one-electron states related to the inter-subsystem boundary and the related parameterization. We consider the possibility of subsequent derivation (*i.e.* proof) of the model is an argument in favor it (in mathematical sense) while the absence of derivation (or its impossibility) as a contra argument.

The paper is organized as follows. In the next Section we formulate the physical principles of constructing hybrid QM/MM schemes and illustrate their application by separation of electronic variables in the framework of the effective Hamiltonian technique and by derivation of renormalization of QM and MM parameters in the framework of the deductive MM formalism. In the third Section the outline of existing approaches to QM/MM intersubsystem junction is given with numerical examples confirming theoretical conclusions and analysis from the general principles given in the second Section is performed. Finally, a brief conclusion is given.

2. Physical principals of constructing QM/MM schemes

The structure of the area of the QM/MM modeling is obscured by a great deal of the recipes proposed and by the lack of sequential deriva-

tions of them. The situation is partly caused by the structure of the MM itself – the MM scheme is not derived but is taken on the ground of some intuitive concepts of what classical terms must enter the energy energy with adjusted parameters for the latter. We extract here a minimal set of the most essential facts (physical principles) which are necessary prerequisites for constructing hybrid schemes:

- 1 *Chemical transformations are local and the perturbation caused by the environment is small.* This principle is confirmed by all the experimental chemistry – the concept of the reaction center and mechanisms of chemical reactions affecting only several chemical bonds. The very concept of chemical bond is mostly due to the local (in many cases one bond) nature of chemistry.
- 2 *The system can be divided on parts and different level approximations (quantum or classical) can be applied to different parts of the system.* The size of the quantum subsystem should be determined by the locality of transformations and not by the intention to move the boundary far from the reactive center to mask the errors in the junction construction. In the case of three-dimensional structures significant shift of boundary from the reactive center is problematic since the size of the QM system may become very large to be covered by the high-level quantum chemistry.
- 3 *Correct form of intersubsystem junction corresponds to sequential separation of variables.* This principle simply states that the construction should be made not in the *ad hoc* manner.
- 4 *Separation of system into parts should be performed in a manner in which fluctuations of electronic density between subsystems are small, i.e. chemical bonds are not to be cut.* This requirement seems to be quite natural since both QM and MM schemes work properly only in the case of systems with well defined numbers of electrons. The developed QM schemes operate with the many-electron states which are eigenstates of the number of particles operator. Even if the bond to be cut is not polar and the diagonal one-electron densities can be set equal to integers and thus the average numbers of electrons in subsystems are also integer the situation does not change since according to Ref. [88] the one-electron density is separable only for the wave functions where the intergroup electron transfers are projected out. It is definitely not a good approximation for two ends of a chemical bond which can exist only in the case of a nontrivial QM superposition of the states with different number of electrons on its ends. The MM description

for a system with fluctuating number of electrons is not defined as well.

- 5 *There exists a QM description underlying the MM one.* It means that the form of junction is not arbitrary but essentially defined by one-electron states arising from the underlying QM description.

These points are very general. Now we consider in more details the practical implementation of these principles.

2.1 Effective Hamiltonian technique

As a first step for derivation of junction we consider a general formulation of separation of electronic variables of quantum subsystem from those describing electrons in the classical (MM) subsystem which provide a consistent form of intersubsystem junction. This separation is based on the Löwdin partition technique [89] and the McWeeny group function formalism [90]. Generally, we construct an effective Hamiltonian for the QM subsystem in the presence of classical environment and the PES of combined system. This strategy was proposed and developed in Refs. [73, 74, 91]. Practically it is based on the assumption of existence of a QM scheme underlying the MM one.

We consider two subsystems R - (reactive, quantum) and M - (inert, classical). The electronic Hamiltonian for the whole system is a sum of subsystem Hamiltonians and of their interaction which is taken to comprise the terms of two types – the Coulomb $V^c(q)$ and the resonance (electron transfer) $V^r(q)$ interactions:

$$H = H^R(q) + H^M(q) + V^c(q) + V^r(q). \quad (1)$$

The Hamiltonian for the M -subsystem is a sum of the free M -subsystem Hamiltonian $H_0^M(q)$ and of the attraction of electrons in the M -subsystem to the cores of the R -subsystem $\mathcal{V}^R(q)$. Analogous subdividing is true for the R -subsystem. The "exact" wave function of the system can be represented by generalized group function with number of electrons in subsystems not fixed:

$$\Psi_k = \sum_{\{n_\alpha\}} \sum_{\{i_\alpha\}} C_{\{i_\alpha\}}^k(\{n_\alpha\}) \bigwedge_{\alpha} \Phi_{i_\alpha}^\alpha(n_\alpha), \quad (2)$$

where $\Phi_{i_\alpha}^\alpha(n_\alpha)$ is the i -th n_α -electron wave function of the α -th group and electron distributions $\{n_\alpha\}$ satisfy the equation:

$$\sum_{\alpha} n_\alpha = N_e; \quad \forall \alpha, n_\alpha \geq 0. \quad (3)$$

The coefficients $C_{\{i_\alpha\}}^k(\{n_\alpha\})$ should be determined on the ground of variational principle.

The function Eq. (2) is very general and does not correspond to the assumed wave function of the hybrid QM/MM method. First of all the numbers of electrons in the subsystems must be fixed to apply computational schemes to separate parts on the legal ground. Second, we assume that the M -subsystem is treated with use of the MM, *i.e.* the PES of the M -subsystem is evaluated without explicit invocaion of its wave function. The parameters of the M -subsystem must be transferable, *i.e.* applicable to combination with any R -subsystem and even in the absence of it. For this purpose we should use the wave function of the ground state of the effective Hamiltonian for the M -subsystem since it is in a certain sense close the wave function calculated without any R -subsystem [73]. Thus the required wave function is represented by the antisymmetrized product of electronic wavefunction for the R -subsystem and that of the ground state for the free M -subsystem:

$$\Psi_k = \Phi_k^R \wedge \Phi_{00}^M. \quad (4)$$

It is obtained from the exact wavefunction by two sequential Löwdin projection procedures: the first one to the subspace of the states with fixed number of electrons in the subsystems (projection operator P and its complementary projection operator $Q = 1 - P$) and the second one – to the states with the ground state wavefunction of the free M -subsystem as the multiplier (projection operator \mathcal{P} and its complementary projection operator $\mathcal{Q} = 1 - \mathcal{P}$).

After the first partition we obtain

$$H_{eff}(q, E) = PH^R(q)P + PH^M(q)P + PV^c(q)P + PV^{rr}(q, E)P + \frac{e^2}{2} \sum_{A \neq B} Z_A^R Z_B^M R_{AB}^{-1}, \quad (5)$$

where the second order resolvent contribution coupling two one-electron transfers between the subsystems is:

$$V^{rr}(q, E) = V^r(q)Q(E - QHQ)^{-1}QV^r(q) = V^r(q)QR(q, E)QV^r(q). \quad (6)$$

The Hamiltonian of the type Eq. (5) is typical for the ECF method [57], where projector P corresponds to the SCF wave function of the ligands in transition metal complex. The second projection and subsequent averaging over the ground state of the M -subsystem gives the effective

Hamiltonian for the R -subsystem:

$$\begin{aligned} \mathcal{H}_{eff}^R(q, E, \omega) = & H_0^R(q) + \delta V^M(q) + \\ & + \langle \langle PV^{rr}(q, E)P \rangle \rangle_M + \langle \langle P\mathcal{V}^R(q)P \rangle \rangle_M + \\ & + \langle \langle P\mathcal{W}(q, E)P\mathcal{Q}\mathcal{R}(\omega)\mathcal{Q}P\mathcal{W}(q, E)P \rangle \rangle_M + \frac{e^2}{2} \sum_{A \neq B} Z_A^R Z_B^M R_{AB}^{-1}, \end{aligned} \quad (7)$$

where the symbol $\langle \langle \dots \rangle \rangle_M$ corresponds to the averaging over the ground state of the M -subsystem $\langle \Phi_{00}^M | \dots | \Phi_{00}^M \rangle$, the averaged operator

$$\delta V^M(q) = \mathcal{V}^M(q) + \langle \langle PV^c(q)P \rangle \rangle_M \quad (8)$$

is close to the zero for quite typical case of M -subsystems with zero overall charge and

$$P\mathcal{W}(q, E)P = PV^c(q)P + PV^{rr}(q, E)P + P\mathcal{V}^R(q)P. \quad (9)$$

To perform the averaging explicitly a form of the wavefunction for the M -subsystem should be specified. At the same time some simplifications of the expression Eq. (7) can be made on quite a general level. For example, the second order contribution from the resolvent of first partition expressed through one-electron states of subsystems is:

$$\begin{aligned} \langle \langle PV^{rr}(q, E)P \rangle \rangle_M = & \sum_{rr' \in R} \sum_{mm' \in M} v_{rm}(q)v_{r'm'}(q) \times \\ & \times \left\{ \sum_{\rho \in Im\mathcal{O}_R(N_R-1)} r^+ |\rho\rangle \langle \rho| r' G_{mm'}^{(adv)}(I_\rho - E) + \right. \\ & \left. \sum_{\rho \in Im\mathcal{O}_R(N_R+1)} r |\rho\rangle \langle \rho| r'^+ G_{mm'}^{(ret)}(A_\rho - E) \right\}, \end{aligned} \quad (10)$$

where $G^{(ret)}(\epsilon)$ and $G^{(adv)}(\epsilon)$ are the retarded and advanced one-electron Green's functions for the M -subsystem, respectively.

The renormalization of the bare R -subsystem Hamiltonian leads to addition of some interaction terms to the sum of energies of free subsystems. The PESs for the whole molecule are obtained by formula:

$$\mathcal{E}_k = E_k^R + E_{00}^M, \quad (11)$$

where E_k^R is the k -th eigenvalue of the effective Hamiltonian of the R -subsystem:

$$E_k^R = \langle \Phi_k^R(q) | \mathcal{H}_{eff}^R | \Phi_k^R(q) \rangle \quad (12)$$

and E_{00}^M is the energy of the M -subsystem which is parameterized in the MM form. The detailed analysis [91] of the effective R -subsystem energy E_k^R allows to present it in the form of a sum of contributions of different

physical nature. By construction the general consideration of junction is close to that of intermolecular forces. At the same time there are some important differences: the hybrid schemes require consideration of boundary atoms; the special form of the wave function of inert subsystem is required; the resonance terms cannot be neglected. The correction to the k -th eigenvalue of the bare Hamiltonian (i.e. the difference between E_k^R with and without environment affecting the reaction center) is

$$\Delta E_k \approx \Delta E_k^{elst} + \Delta E_k^{rr} + \Delta E_k^{disp} + \Delta E_k^{coul} + \Delta E_k^{crr} + \Delta E_k^{Rrr}, \quad (13)$$

where the leading contribution ΔE_k^{elst} is of the first order with respect to the perturbation operator and originates from the electrostatic interaction between the subsystems, it can be expressed through the charges on the atoms; the next three contributions are of the second order and follow from the coupling of one-electron transfers between subsystems, of Coulomb electron-electron interactions between them, of interactions between electrons of the M -subsystem and cores of the R -subsystem; and of interaction of the Coulomb electron-electron interaction with the interaction of the electrons of the M -subsystem and cores of the R -subsystem; the second term can be regarded as dispersion interaction between subsystems. Its general form is:

$$\begin{aligned} \Delta E^{disp} = & - \sum_{rr' \in R} \sum_{mm' \in M} (rr|mm)(r'r'|m'm') \times \\ & \times \int_0^\infty d\omega \pi_{rr'}^R(i\omega) \pi_{mm'}^M(i\omega), \end{aligned} \quad (14)$$

where $\pi_{rr'}^R(i\omega)$ and $\pi_{mm'}^M(i\omega)$ are the reduced polarization propagators for the R - and M -subsystems. The sum of the contributions ΔE_k^{disp} and ΔE_k^{coul} has a physical meaning of the second order interaction between the electronic polarization in the M -subsystem and the polarized R -subsystem. The last two contributions to Eq. (13) correspond to the third order in the interaction between the subsystems originating from the coupling two projection procedures. Physically it corresponds to interaction between two one-electron transfers and Coulomb interaction between electrons of the M -subsystem with electrons and cores of the R -subsystem. The explicit form of these contributions is given in Ref. [91].

2.2 Deductive MM formulation and boundary one-electron states

2.2.1 Semiempirical APSLG-MINDO/3 approach.

General physical principles of constructing hybrid QM/MM approaches state

that the subsequent derivation of junction between quantum and classical subsystems requires a QM wave function underlying the MM description of PESs. This QM method is necessary, for example, to perform averaging of the effective Hamiltonian in Eq. (7). At the same time more important that this method should produce in a consistent manner one-electron states necessary for explicit formation of boundary and its response on the changes in molecular geometry of fragments and/or electronic structure of the R -subsystem.

There are some general criteria for the QM method appropriate for underlying the MM description: this method should express molecular electronic structure and electronic energy in relevant local terms (*i.e.* to distinguish bonded and nonbonded contributions) and reproduce molecular properties with sufficient accuracy. The parameters characterizing the wave function of this method (electronic structure parameters or ESPs) have to be transferable in a broad sense of the term "transferability", *i.e.* the form of any bond-related functions (*e.g.* the bond energy dependence on interatomic separation) must be also transferable. As an appropriate method satisfying these conditions we take the trial wave function in the form of antisymmetrized product of strictly localized geminals (APSLG) [92] implemented with slightly modified semiempirical MINDO/3 Hamiltonian [26, 27]. The APSLG wave function is constructed from two-electron building blocks – geminals:

$$|\Phi\rangle = \prod_m g_m^+ |0\rangle. \quad (15)$$

Each geminal corresponds to a chemical bond or to a lone pair and it is expressed as a linear combination of singlet two-electron configurations constructed from two (in the case of chemical bond) or one (in the case of lone pair) hybrid orbitals (HOs) assigned to the geminal at hand:

$$g_m^+ = u_m r_{m\alpha}^+ r_{m\beta}^+ + v_m l_{m\alpha}^+ l_{m\beta}^+ + w_m (r_{m\alpha}^+ l_{m\beta}^+ + l_{m\alpha}^+ r_{m\beta}^+) \quad (16)$$

$$u_m^2 + v_m^2 + 2w_m^2 = 1.$$

The first two configurations are ionic with two electrons at the end of bond while the last is the covalent (Heitler-London) one. In the case of lone pair only first configuration survives with the coefficient $u_m = 1$. The HOs $|r_m\rangle$ and $|l_m\rangle$ correspond to the right and left ends of the bond and are formed by 4×4 orthogonal ($SO(4)$) transformations of atomic orbitals basis sets for each "heavy" (non-hydrogen) atom. The amplitudes of the geminals Eq. (16) and parameters of the $SO(4)$ transformations are determined by optimizing the energy. The variational character of the one-electron states is an essential factor for use of the APSLG-MINDO/3 scheme in the construction of boundary between R - and M -subsystems.

The important feature of the APSLG-MINDO/3 energy is that it is a sum of increments of different form for bonded and nonbonded atoms:

$$\begin{aligned}
 E_A &= \sum_{t_m \in A} [2U_m^t P_m^{tt} + (t_m t_m | t_m t_m)^{T_m} \Gamma_m^{tt}] + \\
 &\quad + 2 \sum_{t_k < t'_m \in A} g_{t_k t'_m}^{T_k} P_k^{tt} P_m^{t't'}, \\
 E_{R_m L_m}^{bond} &= 2\gamma_{R_m L_m} [\Gamma_m^{rl} - 2P_m^{rr} P_m^{ll}] - 4\beta_{r_m l_m}^{R_m L_m} P_m^{rl}, \\
 E_{AB}^{nonbond} &= Q_A Q_B \gamma_{AB}.
 \end{aligned} \tag{17}$$

where U_m^t is the matrix element of attraction of an electron at the t_m -th HO to its own core, $\beta_{r_m l_m}^{R_m L_m}$ is the intrabond resonance integral, Q_A is the Mulliken charge on the atom A , $\gamma_{\alpha\beta}$ is the two-center Coulomb integral, and $g_{t_k t'_m}^{T_k} = 2(t_k t_k | t'_m t'_m)^{T_k} - (t_k t'_m | t'_m t_k)^{T_k}$ is the doubled reduced Coulomb integral. The ESPs of two types enter the expression Eq. (17): (i) the intrabond elements of one- and two-electron density matrices

$$\begin{aligned}
 P_m^{t't'} &= \langle 0 | g_m t_{m\sigma}^+ t'_{m\sigma} g_m^+ | 0 \rangle, \quad \Gamma_m^{t't'} = \langle 0 | g_m t_{m\beta}^+ t'_{m\alpha} t'_{m\alpha} t_{m\beta} g_m^+ | 0 \rangle, \\
 P_m^{rr} &= u_m^2 + w_m^2, \quad P_m^{ll} = v_m^2 + w_m^2, \quad P_m^{rl} = P_m^{lr} = (u_m + v_m)w_m, \\
 \Gamma_m^{rr} &= u_m^2, \quad \Gamma_m^{ll} = v_m^2, \quad \Gamma_m^{rl} = \Gamma_m^{lr} = w_m^2,
 \end{aligned} \tag{18}$$

where t and t' are either r or l , and (ii) the hybridization matrices defining molecular integrals in the basis of HOs.

The results of Refs. [26, 27] related to heats of formation and equilibrium geometries obtained by the APSLG-MINDO/3 method are somewhat more accurate than those of the standard (SCF-)MINDO/3 method. The APSLG form of the trial wave function also ensures its correct asymptotic behaviour under cleavage of chemical bonds which indirectly justifies some level of *bonae fidelitatis* of the wave function employed. The APSLG form of the wave function also allows to represent a renormalization of the bare R -subsystem Hamiltonian in terms of well-defined characteristics like atomic charges, bond polarizabilities and ionization potentials for the chemical bonds [74, 91].

2.2.2 Deductive molecular mechanics. In the framework of QM/MM junction construction we should determine the structure of boundary one-electron states and their responses to variations of geometric parameters. The problem of constructing optimal boundary HOs is closely related to the more general problem of deduction of the MM description from some consistent QM description of electronic structure. This problem is actual since the form of the force fields and their particular sets of parameters are not justified that leads to a great variety of different MM schemes. The semiempirical APSLG-MINDO/3 method

briefly described above is a perfect candidate for the MM derivation and it was successfully used for this purpose in Refs. [93, 94]. Here we describe the main steps of the MM derivation and their consequences for the QM/MM modeling.

To derive the MM picture we must consider the question of transferability of the ESPs. Above we defined two groups of parameters entering the expression for the energy Eq. (17). Let us consider the first group of parameters. In the case of lone pairs they are perfectly transferable. The SCF approximation for non-polar bonds gives the geometry independent density matrix elements:

$$P_m^{tt'} = \frac{1}{2}, \quad \Gamma_m^{tt'} = \frac{1}{4}. \quad (19)$$

This picture can be termed as FA (*fixed geminal amplitudes*) [93]. It was shown in Ref. [95] by analysis of the bond energy that two small parameters (functions of molecular geometry) for each bond ζ_m^{-1} and μ_m can be defined. They describe intrabond correlation and bond asymmetry, respectively. The extent (degree) of transferability of the elements of density matrices is given by expression of them as power series on the parameters. The contribution of the zeroth order in μ_m is:

$$\Gamma_m^{tt'} = \frac{1}{4} \left(1 - tt' \frac{1}{\Gamma(\zeta_m)} \right), \quad P_m^{tt} = \frac{1}{2}, \quad P_m^{rl} = \frac{\zeta_m}{2\Gamma(\zeta_m)}, \quad (20)$$

where $\Gamma(\zeta_m) = \sqrt{1 + \zeta_m^2}$, while the correction of up to second order in μ_m results in slightly modified values:

$$\begin{aligned} \Gamma_m^{rl} &= \Gamma_{m0}^{rl} \left[1 + \mu_m^2 \frac{2\Gamma(\zeta_m)+1}{2(\Gamma(\zeta_m)+1)^2} (1-\Gamma(\zeta_m)) \right], \\ P_m^{tt} &= P_{m0}^{tt} \left[1 + t\mu_m \frac{\Gamma(\zeta_m)-1}{\Gamma(\zeta_m)+1} \right], \\ P_m^{rl} &= P_{m0}^{rl} \left[1 + \mu_m^2 \frac{2\Gamma(\zeta_m)+1-\Gamma^2(\zeta_m)}{2(\Gamma(\zeta_m)+1)^2} \right], \end{aligned} \quad (21)$$

where subscript 0 corresponds to the estimates by Eq. (20). From these estimates it can be concluded that the bond order ($2P_m^{rl}$) is transferable up to second order with respect to both parameters ζ_m^{-1} and μ_m ; the bond covalency ($2\Gamma_m^{rl}$) is transferable up to second order with respect to μ_m and up to first order with respect to ζ_m^{-1} ; the bond polarity ($P_m^{rr} - P_m^{ll}$) is transferable up to first order with respect to both ζ_m^{-1} and μ_m . The transferability of bond orders is the most important for the MM derivation. The picture of Eqs. (20) and (21) can be termed as TA (*tuned geminal amplitudes*) [93].

The specific bond equilibrium distance can be obtained as the minimum position for the bond-related energy. The core-core interaction in

the MINDO/3 scheme is not purely Coulomb one but is modified by a short-range repulsion term:

$$E_{nn} = \frac{1}{2} \sum_{\alpha \neq \beta} C_{\alpha\beta}, \quad C_{\alpha\beta} = Z_\alpha Z_\beta (\gamma_{\alpha\beta} + D_{\alpha\beta}). \quad (22)$$

In the case of non-polar bond Eq. (20) the minimum condition is:

$$\begin{aligned} \frac{\partial E}{\partial q} &= Z_{R_m} Z_{L_m} \frac{\partial D_{R_m L_m}}{\partial q} - 2 \frac{\zeta_m}{\Gamma(\zeta_m)} \frac{\partial \beta_{r_m l_m}^{R_m L_m}}{\partial q} - \\ &\quad - \frac{1}{2} \left(1 - \frac{1}{\Gamma(\zeta_m)} \right) \frac{\partial \gamma_{R_m L_m}}{\partial q} = 0, \end{aligned} \quad (23)$$

since the derivatives of the ESPs exactly compensate each other. In the limit $\zeta_m \rightarrow \infty$ we recover the FA picture. The elasticity constant for the bond stretching in the FA picture is:

$$k_{R_m L_m} = \left(Z_{R_m} Z_{L_m} \frac{d^2 D_{R_m L_m}}{dr_{R_m L_m}^2} - 2 \frac{\partial^2 \beta_{r_m l_m}^{R_m L_m}}{\partial r_{R_m L_m}^2} - \frac{1}{2} \frac{d^2 \gamma_{R_m L_m}}{dr_{R_m L_m}^2} \right)_{r_{R_m L_m}^0}. \quad (24)$$

It should be noted that the short-range repulsion is due to specific core-core contribution $D_{R_m L_m}$.

General construction of the MM scheme and analysis of adjustment of one-electron states on the QM/MM boundary can be performed using linear response relations between the geometry parameters and the ESPs. These relations can be obtained by expanding the energy up to the second order with respect to both atomic coordinates q and the ESPs x with subsequent minimization:

$$x - x_0 = - (\nabla_x \nabla_x E)^{-1} \nabla_x \nabla_q E (q - q_0). \quad (25)$$

It leads to the following energy expression:

$$\begin{aligned} E &= E_0 + \frac{1}{2} (q - q_0 | \nabla_q \nabla_q E | q - q_0) - \\ &\quad - \frac{1}{2} \left(q - q_0 \left| \nabla_q \nabla_x E (\nabla_x \nabla_x E)^{-1} \nabla_x \nabla_q E \right| q - q_0 \right), \end{aligned} \quad (26)$$

which depends on the geometry variables q only. An important component of the MM construction is determination of the source of angular dependence of the energy which in the APSLG-MINDO/3 approximation is determined by the hybridization. It can be shown [95] that the one-center energy is hybridization dependent only if subtle polarization and correlation effects are taken into account. At the same time the hybridization is a kind of coarse phenomenon which can be reproduced even by methods without intraatomic electron-electron repulsion [96]. Therefore, the resonance energy should be considered as the main source of the hybridization.

Mathematically the hybridization of the sp -basis at a given atom is defined by a $SO(4)$ rotation. Each HO is thus a normalized quaternion with a scalar and a 3-vector parts (s_m, \vec{v}_m) . The set of four vector parts of HOs forms a hybridization tetrahedron with vectors \vec{v}_m pointing to its vertices. In principle, deductive MM can be constructed as a set of classical type equations for these tetrahedra [85]. Here we consider, however, more common case of the MM force fields. The dynamical $SO(4)$ group can be parameterized by a pair of normalized quaternions corresponding to fictitious rotations combining the hybridization (s - p mixing, pseudorotation) and rotation of hybridization tetrahedron as a whole (quasirotation). At the same time the first order variation of the HOs when small quasi- and pseudorotations $\delta\vec{\omega}_l$ and $\delta\vec{\omega}_b$ are applied to the system of HOs can be derived in a very simple form:

$$\begin{aligned}\delta^{(1)}s &= -(\delta\vec{\omega}_b, \vec{v}), \\ \delta^{(1)}\vec{v} &= s\delta\vec{\omega}_b + \delta\vec{\omega}_l \times \vec{v},\end{aligned}\tag{27}$$

where \times stands for the vector product of 3-vectors.

Let us demonstrate an application of formulae Eqs. (25) and (26) on the example of methane molecule with sp^3 hybridization. The stereochemistry of carbon atom is characterized by four unit vectors $\vec{e}_{R_m L_m}$ corresponding to the directions of the CH bonds. In the diatomic coordinate frame with the z axis directed along the $\vec{e}_{R_m L_m}$ vector the resonance integrals are:

$$\beta_{r_m l_m}^{CH} = \beta_{\sigma\sigma}^{CH} s_m^C + \beta_{\zeta\sigma}^{CH} (\vec{v}_m^C, \vec{e}_{CH_m}).\tag{28}$$

The equilibrium conditions for the form and orientation of hybridization tetrahedron can be obtained by zeroing the pseudotorque \vec{N} and quasitorque \vec{K} which are coefficients at $\delta\vec{\omega}_b$ and $\delta\vec{\omega}_l$ in the resonance energy expansion:

$$\begin{aligned}\vec{N} &= 4 \sum_m P_m^r \{ \beta_{\sigma\sigma}^{CH} \vec{v}_m^C - \beta_{\zeta\sigma}^{CH} s_m^C \vec{e}_{CH_m} \} = 0, \\ \vec{K} &= 4 \sum_m P_m^r \beta_{\zeta\sigma}^{CH} \vec{e}_{CH_m} \times \vec{v}_m^R = 0.\end{aligned}\tag{29}$$

The solution of the set of these equations in the case of symmetric hydride (methane) is the hybridization with all vectors \vec{v}_m^C directed along vectors \vec{e}_{CH_m} .

The fixation of parameters $\vec{\omega}_b$ (those of the shape of the hybridization tetrahedron) can be considered as FO (*fixed orbitals*) picture [93]. In this case only the resonance energy is the orientation dependent contribution. The angular dependence of the energy (bending) can be described by introducing small rotation vectors $\delta\vec{\varphi}_m$, which after applying them to

vectors $\vec{e}_{R_m L_m}$ lead to new (distorted) coordination tetrahedron. The final result of substitution of second derivatives of the energy to Eq. (26) is the explicit expression for the bending force field constant which can be written as:

$$k_{HCH} = \beta_{\zeta\sigma}^{CH} \{P_m^{rl}(\vec{v}_m^C, \vec{e}_{CH_m}) + P_{m'}^{rl}(\vec{v}_{m'}^C, \vec{e}_{CH_{m'}})\}, \quad (30)$$

i.e., as a sum of two separate single bond contributions. Moreover, it can be proven [94] that adjustment of hybridization tetrahedron to the geometry deformation does not modify the expression Eq. (30) in the FO picture.

In general both orientation and form of hybridization tetrahedron must be adjusted. This picture can be termed as TO (*tuned orbitals*) [93]. The reaction of hybridization tetrahedron on the changes of local geometry can be considered in the linear response approximation Eq. (25). In the case of valence angles deformation the algebraic structure of the $SO(4)$ manifold ensures quite special structure of geometry deformations which can cause changes in the form of hybridization tetrahedron. These deformations (hybridization compatible) correspond to increasing of a valence angle by certain amount $\delta\theta_{1m}$ with simultaneous decreasing of the opposite (spiro) angle by the same magnitude. Substitution of matrices of second derivatives for the energy into Eq. (25) gives the reaction of the form of hybridization tetrahedron on the angular distortions of molecular geometry in the form:

$$\delta\vec{\omega}_b = -\frac{\beta_{\zeta\sigma}^{CH}}{\sqrt{2}(\sqrt{3}\beta_{\sigma\sigma}^{CH} + \beta_{\zeta\sigma}^{CH})} \left(\delta\chi_{12}\vec{k} + \delta\chi_{13}\vec{j} + \delta\chi_{14}\vec{i} \right), \quad (31)$$

provided the parameters $\delta\chi_{1m}$ describe the hybridization compatible deformations of the coordination tetrahedron. The same considerations also apply to the bond stretching where variation of the resonance parameters can be presented as:

$$\delta\beta_{\mu\nu}^{CH_m} = \gamma_{\mu\nu}^{CH} \delta r_{CH_m}. \quad (32)$$

The response of the form of the hybridization tetrahedron on the change of the bond length can be written as:

$$\delta\vec{\omega}_b = -\frac{\sqrt{3}}{2} \cdot \frac{\gamma_{\sigma\sigma}^{CH} \vec{v}_m^C - \gamma_{\zeta\sigma}^{CH} s_m^C \vec{e}_{R_m L_m}}{\sqrt{3}\beta_{\sigma\sigma}^{CH} + \beta_{\zeta\sigma}^{CH}} \delta r_{CH_m} \quad (33)$$

that leads to the following expression

$$\frac{\partial^2 E}{\partial r_{CH_m} \partial r_{CH_{m'}}} = \frac{1}{2\sqrt{3}} P^{rl} \frac{(\sqrt{3}\gamma_{\sigma\sigma}^{CH} - \gamma_{\zeta\sigma}^{CH})^2}{\sqrt{3}\beta_{\sigma\sigma}^{CH} + \beta_{\zeta\sigma}^{CH}} \quad (34)$$

for the off-diagonal term corresponding to coupling of stretching for two incident C-H bonds.

2.2.3 Implications for QM/MM methods. The above results constitute a basis for construction of intersubsystem junction in hybrid QM/MM schemes. There are two different effects arising from separation of system on parts and consideration of one part as a MM one – renormalization of the QM Hamiltonian parameters and imposing additional forces and torques on the MM subsystem. Let us exemplify the technique of short-range intersubsystem junction construction by one special but quite characteristic case of a boundary sp^3 carbon atom with one HO pointing to the QM region (bond $m = 1$) and others related to the MM one (this case can be classified as the MM boundary atom). The transition to the MM picture is performed by setting the FA and TO approximations. To determine the effect of the QM part on the MM part we consider the perturbation of molecular energy due to changes of electron densities in the QM region. It affects both one-center energy of boundary atom and resonance energy between HO r_1 of the QM bond with number $m = 1$ and all other HOs in the QM subsystem. The perturbation sets up quasi- and pseudotorques for the boundary atom:

$$\begin{aligned}
\vec{K}' &= 2 \sum_A \{ (\delta P_{r_1\sigma} \beta_{\zeta\sigma}^{R_1A} + \delta P_{r_1\zeta} \beta_{\zeta\zeta}^{R_1A}) \vec{e}_{R_1A} + \\
&\quad + \beta_{\pi\pi}^{R_1A} (\delta P_{r_1\xi} \vec{e}_{R_1A}^\xi + \delta P_{r_1v} \vec{e}_{R_1A}^v) \} \times \vec{v}_1^{R_1}, \\
\vec{N}' &= -s_1^{R_1} \vec{v}_1^{R_1} \{ (P_1^{rr} - P_1^{ll}) [2U_s - 2U_p + C_2 + C_3 + 2C_5] + \\
&\quad + (1/2 - \Gamma_1^{rl}) [C_2 + 2C_3 (s_1^{R_1})^2] \} + \\
&\quad + 2 \sum_A \{ (\delta P_{r_1\sigma} \beta_{\sigma\sigma}^{R_1A} + \delta P_{r_1\zeta} \beta_{\sigma\zeta}^{R_1A}) \vec{v}_1^{R_1} - \\
&\quad - (\delta P_{r_1\sigma} \beta_{\zeta\sigma}^{R_1A} + \delta P_{r_1\zeta} \beta_{\zeta\zeta}^{R_1A}) s_1^{R_1} \vec{e}_{R_1A} - \\
&\quad - \beta_{\pi\pi}^{R_1A} s_1^{R_1} (\delta P_{r_1\xi} \vec{e}_{R_1A}^\xi + \delta P_{r_1v} \vec{e}_{R_1A}^v) \},
\end{aligned} \tag{35}$$

where $\vec{e}_{R_1A}^\xi$, $\vec{e}_{R_1A}^v$, and $\vec{e}_{R_1A} = \vec{e}_{R_1A}^\zeta$ are the orthonormal basis of the DCF defined by the R_1A pair of atoms and C_2 , C_3 , and C_5 are linear combinations of the Slater-Condon parameters:

$$\begin{aligned}
C_2^A &= 2F_0^A(sp) + 4G_1^A(sp) - 2F_0^A(pp) - 8F_2^A(pp), \\
C_3^A &= F_0^A(ss) - 2F_0^A(sp) - 4G_1^A(sp) + F_0^A(pp) + 4F_2^A(pp), \\
C_5^A &= 2F_0^A(sp) - G_1^A(sp) - 2F_0^A(pp) + 7F_2^A(pp).
\end{aligned} \tag{36}$$

The additional pseudo- and quasitorques produce the pseudo- and quasirotations of the hybridization tetrahedron on the boundary atom R_1 . In the linear response approximation it corresponds to the treatment of the corresponding pseudo- and quasitorques by the $(\nabla_{\vec{\omega}}^2 E)^{-1}$ matrix

which is simple (diagonal in the basis of $\delta\vec{\omega}_b$ and $\delta\vec{\omega}_l$) in the case of symmetric hydride:

$$\begin{aligned}\delta\vec{\omega}_b &= -\frac{\vec{N}'}{4P^{rl}(\beta_{\sigma\sigma} + \beta_{\zeta\sigma}/\sqrt{3})}, \\ \delta\vec{\omega}_l &= -\frac{\sqrt{3}\vec{K}'}{8P^{rl}\beta_{\zeta\sigma}},\end{aligned}\quad (37)$$

and, therefore, the hybridization tetrahedron acquires new form and orientation which can be inconsistent with the MM atoms' positions. It corresponds to additional classical forces and torques acting on the MM neighbours of the boundary atom at hand. The forces are directed along the $\vec{e}_{R_m L_m}$ vectors and stem from two sources: the variation of the shape of HOs:

$$\begin{aligned}f_{bm} &= \frac{1}{2(\beta_{\sigma\sigma} + \beta_{\zeta\sigma}/\sqrt{3})} \{ -\gamma_{\sigma\sigma}^{R_m L_m} s_m^{L_m} \vec{v}_m^{R_m} - \\ &- \gamma_{\sigma\zeta}^{R_m L_m} v_{m\zeta}^{L_m} \vec{v}_m^{R_m} + \gamma_{\zeta\sigma}^{R_m L_m} s_m^{R_m} s_m^{L_m} \vec{e}_{R_m L_m} + \gamma_{\pi\pi}^{R_m L_m} s_m^{R_m} \vec{v}_m^{L_m} + \\ &+ (\gamma_{\zeta\zeta}^{R_m L_m} - \gamma_{\pi\pi}^{R_m L_m}) s_m^{R_m} v_{m\zeta}^{L_m} \vec{e}_{L_m R_m} \}, \vec{N}'\end{aligned}\quad (38)$$

and the variation of the hybridization tetrahedron orientation:

$$\begin{aligned}f_{lm} &= \frac{\sqrt{3}}{4\beta_{\zeta\sigma}} \{ -\gamma_{\zeta\sigma}^{R_m L_m} s_m^{L_m} \vec{e}_{R_m L_m} \times \vec{v}_m^{R_m} - \gamma_{\pi\pi}^{R_m L_m} \vec{v}_m^{L_m} \times \vec{v}_m^{R_m} - \\ &- (\gamma_{\zeta\zeta}^{R_m L_m} - \gamma_{\pi\pi}^{R_m L_m}) v_{m\zeta}^{L_m} \vec{e}_{R_m L_m} \times \vec{v}_m^{R_m} \}, \vec{K}'\end{aligned}\quad (39)$$

Analogously, torques acting upon the neighbouring to boundary MM atom L_m are due to variation of the form of hybridization tetrahedron:

$$\begin{aligned}\vec{t}_{bm} &= -\frac{1}{2(\beta_{\sigma\sigma} + \beta_{\zeta\sigma}/\sqrt{3})} \{ \beta_{\sigma\zeta}^{R_m L_m} (\vec{e}_{R_m L_m} \times \vec{v}_m^{L_m}) (\vec{v}_m^{R_m}, \vec{N}') - \\ &- s_m^{R_m} [\beta_{\zeta\sigma}^{R_m L_m} s_m^{L_m} + (\beta_{\zeta\zeta}^{R_m L_m} - \beta_{\pi\pi}^{R_m L_m}) v_{m\zeta}^{L_m}] \vec{e}_{R_m L_m} \times \vec{N}' - \\ &- (\beta_{\zeta\zeta}^{R_m L_m} - \beta_{\pi\pi}^{R_m L_m}) s_m^{R_m} (\vec{e}_{R_m L_m} \times \vec{v}_m^{L_m}) (\vec{e}_{R_m L_m}, \vec{N}') \}\end{aligned}\quad (40)$$

and its orientation:

$$\begin{aligned}\vec{t}_{lm} &= -\frac{\sqrt{3}}{4\beta_{\zeta\sigma}} \{ \beta_{\zeta\sigma}^{R_m L_m} s_m^{L_m} (\vec{v}_m^{R_m} \otimes \vec{e}_{R_m L_m} - v_{m\zeta}^{R_m} \mathcal{I}) + \\ &+ (\beta_{\zeta\zeta}^{R_m L_m} - \beta_{\pi\pi}^{R_m L_m}) [v_{m\zeta}^{L_m} (\vec{v}_m^{R_m} \otimes \vec{e}_{R_m L_m} - v_{m\zeta}^{R_m} \mathcal{I}) + \\ &+ (\vec{e}_{R_m L_m} \times \vec{v}_m^{L_m}) \otimes (\vec{e}_{R_m L_m} \times \vec{v}_m^{R_m})] \} \vec{K}'\end{aligned}\quad (41)$$

It should be noted that the MM subsystem also affects the parameters of the QM subsystem since any geometry deformation in the MM subsystem induces changes in pseudo- and quasirotaion angles defining hybridization of the boundary atom. It renormalizes the QM one-center Hamiltonian parameters according to:

$$\begin{aligned}\delta U_1^r &= 2(U_s - U_p) s_1^{R_1} (\delta\vec{\omega}_b, \vec{v}_1^{R_1}), \\ \delta(r_1 r_1 | r_1 r_1)^{R_1} &= 2(C_2 s_1^{R_1} + 2C_3 (s_1^{R_1})^3) (\delta\vec{\omega}_b, \vec{v}_1^{R_1}).\end{aligned}\quad (42)$$

The resonance integrals in the QM subsystem are also renormalized. In the DCF their modification can be expressed as:

$$\begin{aligned}
\delta\beta_{r_1\sigma}^{R_1A} &= \beta_{\sigma\sigma}^{R_1A}\delta^{(1)}s_1^{R_1} + \beta_{\zeta\sigma}^{R_1A}\delta^{(1)}v_{1\zeta}^{R_1}, \\
\delta\beta_{r_1\zeta}^{R_1A} &= \beta_{\sigma\zeta}^{R_1A}\delta^{(1)}s_1^{R_1} + \beta_{\zeta\zeta}^{R_1A}\delta^{(1)}v_{1\zeta}^{R_1}, \\
\delta\beta_{r_1\xi}^{R_1A} &= \beta_{\pi\pi}^{R_1A}\delta^{(1)}v_{1\xi}^{R_1}, \\
\delta\beta_{r_1v}^{R_1A} &= \beta_{\pi\pi}^{R_1A}\delta^{(1)}v_{1v}^{R_1},
\end{aligned} \tag{43}$$

where variations of HOs are given by Eq. (27).

To illustrate the formulae given above we performed the estimates of the magnitude of the renormalization of the QM parameters and changes in the QM forces and torques due to boundary taken as sp^3 carbon atom. The changes in the QM one-center Hamiltonian parameters due to elongation of one of the boundary MM bonds are:

$$\frac{\partial U_1}{\partial r_2} = -1.162 \frac{eV}{A}; \quad \frac{\partial(t_1t_1|t_1t_1)}{\partial r_2} = 0.537 \frac{eV}{A}, \tag{44}$$

while the corresponding effects due to change of the bond angle between two boundary MM bonds are:

$$\frac{\partial U_1}{\partial \chi_{23}} = 0.755 \frac{eV}{rad}; \quad \frac{\partial(t_1t_1|t_1t_1)}{\partial \chi_{23}} = -0.349 \frac{eV}{rad}. \tag{45}$$

At the other hand the changes in the one- and two-electron densities of the QM boundary bond lead to the following forces and torques acting on the neighbouring MM atoms:

$$\begin{aligned}
f_{bm} &= [-2.225\delta P + 0.465\delta\Gamma] \frac{eV}{A}; \\
\vec{t}_{bm} &= [2.891\delta P - 0.604\delta\Gamma] \frac{\vec{e}_m \times \vec{e}_1}{|\vec{e}_m \times \vec{e}_1|} \frac{eV}{rad}.
\end{aligned} \tag{46}$$

The numerical estimates show that in the case of variational determination of one-electron states the effect of the boundary (besides electrostatic, van-der-Waals *etc.* contributions) can be considered as a relatively weak perturbation.

3. Characteristic recipes of QM/MM modeling

In the present Section we employ the theoretical framework developed above to rationale the state of art in the field of hybrid QM/MM modeling. It is instructive to invoke different classifications of QM/MM schemes developed. The simplest but not very much informative classification can be based on the types of quantum mechanical and molecular mechanical schemes used. There are no fundamental restrictions on the choice of the QM schemes and in fact very large diversity can

be found in the literature. All types of QM procedures – *ab initio* [41, 67, 86, 97, 98, 99, 100], DFT [68, 79, 101, 102, 103, 104, 105], and semiempirical [36, 106, 107, 108] – are quite popular and widely used as well as in usual quantum chemical studies. The usual form of wave function implied is the SCF one. At the same time methods based on the valence bond approximation are also well known [35, 99, 109]. The choice of the MM scheme can be quite important since it affects the structure of bonding terms near (or in some schemes on) the border between the QM and MM parts and the electrostatic polarization of the QM part by the MM-treated environment depends on the particular charge scheme employed in the MM procedure. Practically, it is more convenient to work with the force field with electrostatics based on the atomic charges than with the force field using bond dipole interactions for this purpose since it can cause significant errors in the description of polar species. This notion have led the authors of Ref. [110] employing the MM3 force field [111] to replacing the the bond dipoles with potential-derived atomic point charges. We note that the sequential derivation of junction in the framework of the effective Hamiltonian approach reported in leads to representing electrostatic interactions in terms of atomic charges [112]. If the force field contains many explicit couplings between bonded interactions (like the MM3 force field) it may cause additional problems with construction of junction. Some words should be said about ionic force fields using formal ionic charges and employing electrostatic and short-range force fields. In the case of these force fields short-range interactions arising from the MM charges can not be separated from the long-range interactions [43]. These schemes often lead to incorrect electrostatic potentials so the most popular choice (especially in description of covalently bound QM/MM systems) is the valence force fields. Different implementations use different force fields: for example, the MM3 force field [111] is used in Refs. [110, 113], the CHARMM force field [114] is used in Refs. [34, 115], the AMBER force field [116] is used in Refs. [33, 117].

In the previous Section we obtained the formula for junction between quantum and classical subsystems Eq. (13). The control for the types of interactions which are taken into account is an important characteristic of particular QM/MM scheme. The authors of Ref. [110] have proposed a classification of hybrid schemes based on the interaction between fragments. According to it, the simplest type of model is mechanical embedding (examples of this type of modeling are the IMOMM [38] and IMOMO [59] schemes by Morokuma) when both QM and MM systems are not polarized by each other and their interaction is represented by classical force fields only. In this context the choice of parameters of in-

tersystem interaction can be crucially important, so, they are frequently optimized [97, 118]. More elaborated model is that including polarization of the QM subsystem. This polarization can be covered by including the MM charges into one-electron part of the Hamiltonian of the QM subsystem:

$$h_{\mu\nu}^{pol} = h_{\mu\nu} - \sum_M q^M V_{\mu\nu}^M, \quad (47)$$

where M is the atom of the MM subsystem and q is its charge. This type of modeling is quite popular [34, 119, 120]. At the same there is an objection against such schemes since they violate the principle that *actio* should be equal to *reactio* [121] (no effect on the part of the QM system upon the MM system is assumed). The most elaborated is the scheme including classical treatment of the MM system polarization [122]. This type of approaches is rarely used though the MM polarization was taken into account even in the early scheme [32] by using atomic polarizabilities. Practically intersubsystem polarization corresponds to dispersion interaction Eq. (14). We should note that using atomic polarizabilities and adding the corresponding diatomic terms to the PES is not always substantiated. For example, in the case of the chemically transforming R -subsystem the different electronic terms change their relative energies and thus the positions of the poles of the $\pi_{r,r'}^R(i\omega)$ reduced polarization propagator change along the reaction path. This affects the dispersion energy Eq. (14) but is not reflected by the atom-atom scheme, which is nevertheless better than nothing.

It should be noted that the classification given by authors of Ref. [110] is not complete since, for example, it does not include some self-consistent schemes like that of Ref. [40] and does not consider the possibility of charge transfer between subsystems [74, 91]. The authors of Ref. [123] have imposed a requirement of intersubsystem self-consistency on the construction of junction. It means that the charge transfer between subsystems should be taken into account. Practical implementation of this requirement was performed by using special iterative procedure of double (intrafragment and interfragment) self-consistent (DSC) calculations. It leads to explicit taking into account the electron transfer between subsystems (and also the polarization of the QM subsystem). This methodology, however, can not be justified since the self-consistent field procedures are applied to systems with strongly fluctuating numbers of electrons that also leads to poor definition of the fragments themselves (and of their quantum numbers). According to results of the previous Section the electron transfers should be considered as virtual ones and taken into account in the perturbative fashion. This point is confirmed numerically since the application of the DSC scheme to the

iron picket-fence porphyrin has led to improbably large intersubsystem charge transfer of 3.6 electrons [123].

In the case of intersubsystem junction represented by classical bonding terms an important question arises: which terms should be included and which should not? The most popular way is to include classical bonded force fields when at least one MM atom is involved in it [33, 34]. At the same time it does not allow to avoid double counting of interactions computed quantum mechanically. To smooth this inconsistency the authors of Ref. [124] proposed to calculate only those classical bonding force fields where at least one central atom is from the MM subsystem or in the case of improper dihedral terms only those with both outer atoms from the MM subsystem.

The ways to construct the boundary between subsystems and general ideas concerning the structure of the "grey" zone between quantum and classical fragments proposed in the literature are miscellaneous. For example, authors of Ref. [123] pose the requirement of so called "QM-MM continuity" that leads to introducing intermediate fragments which are treated by both QM and MM methods. Since the early times of quantum chemistry and until now it was popular to make the problem of molecular structure investigation tractable by neglecting large substituents and saturating free valencies by hydrogen atoms. The QM/MM methodology tries to take the bulky substituents into account explicitly. The most straightforward way to treat covalently bound QM and MM parts is the link atom method stemming from saturating dangling bonds produced while cutting the system in parts across the bonds by additional atoms (usually, hydrogens). This methodology requires deleting the terms in the expression for the total energy of the system related to the link atoms. Practically it is very difficult since in the framework of this scheme serious artefacts appear quite expectedly. It should be noted that the polarization of the MM subsystem is often reproduced incorrectly. To cure this error some interaction terms for atoms near the boundary are typically artificially eliminated. Originally [34], all the interactions (electrostatic and van-der-Waals) between link atoms and MM subsystem were not taken into account. Later, different recipes to omit some real physical interactions and by this compensate unphysical ones were proposed: the authors of Ref. [124] neglect the Coulomb interactions between the QM subsystem and the MM group most close to it, the authors of Ref. [125] force the charge on the boundary MM atom to be zero. It should be noted that such eliminations can not be justified and must be considered as special methods for masking the errors caused by the inconsequence in junction construction. In fact, the problem is often a consequence of the small distance between the link

atom and the first MM atom. The more complex scheme [126] does not neglect the interactions of link atom with the MM subsystem but sets charge on the boundary MM atom to be equal to zero. Another recipe used is shifting the values of the charges with subsequent compensation of this perturbation by introducing fictitious dipoles [127]. Typically, the possibility to manipulate by interactions of the artificial link atom are considered as an advantage of special flexibility of the link atom approach [128]. However, to our opinion the price of this flexibility is too high since it leads to a great uncertainty in the results obtained and thus marks down possible predictive capacity of the QM/MM approach.

The saturation of dangling bonds by hydrogen is not the only possible way proposed in the literature. The main reason of using other types of saturation groups is the intention to reproduce the polarity and other properties of the broken bond more correctly. The choices (suggestions) for saturation groups well known in the literature are pseudohalogens [129, 130] and "dummy groups" [131]. The important drawback of link atom method (and, especially of the "dummy groups" method) is introducing of additional nuclear degrees of freedom for which no reasonable equation of motion (or equivalently no energy minimum condition) can be derived. It requires special link atom corrections. To cope with problems incurred by introducing link atoms an adjusted connection atom approach [132] was developed. The connection atom is a QM one but its interactions with the environment include some classical terms. So, the bonded interactions of the connection atom with MM atoms are covered by force field for a carbon atom. It is assumed that the C-C bond is cut and the connection atom mimics the methyl group. This method was calibrated against the NDDO calculations for methyl compounds. It should be noted that this scheme is valid only for equilibrium bond distances since the molecules used for calibration were taken in their equilibrium geometries.

The same problems can be addressed by using effective group potential techniques [51, 52]. For example, in Ref. [52] a special potential reproduces the interactions between the QM subsystem electrons and the missing valence electrons of one-electron boundary atom. In the pseudobond approach [133, 134] the free valence of a QM atom is saturated with a special atom located exactly at the position of the neighbour MM atom. The basis set and number of electrons of this pseudoatom are set to be equal to those of fluorine atom. The C-C bond is mimicked by a specially adjusted effective core potential. Another approach based on the use of effective potentials is proposed in Ref. [135]. A series of one-electron quantum capping potentials replacing the link atom is formed by modification of conventional effective core potentials: the

spherical shielding and Pauli terms replace the excess valence electrons. The capping potential is adjusted to reproduce all-electron geometries and charge distributions. The analysis of this scheme shows that the error induced by capping potential is significant (especially for angular distortions) but generally less than that in the simple link atom scheme. It is noteworthy that this scheme can be extended to construct many-electron quantum capping potentials.

The important problem of the link atom schemes is geometry optimization since as it is shown in Ref. [110] the difference between QM and MM interaction energies for link atom enters into the total energy and strongly depends on the link atom position since the forms of the QM and MM force fields are very much different. Practically it leads to the collapse of the fictitious link atom with the boundary atom. The characteristic result is that the equilibrium position of the link atom is badly defined and can not be reasonably rationalized. One of possible recipes for avoiding the collapse of the link atom is to use for geometry optimization a potential energy function not coinciding with the total energy of the molecular system [110]:

$$E_{pot} = E_{total} - E_{link}. \quad (48)$$

At the same time such an approach seems quite artificial. Moreover, in Ref. [136] (cited by Ref. [121]) it is stated that the strong deviation of the link atom equilibrium position from the line connecting atoms forming covalent bond leads to serious problems. Moreover, the vibrational spectra calculated by the optimization of the link atom position method are worse than even the MM-force field derived. Also the QM/MM calculated proton affinity for small gas phase aluminosilicate clusters is very sensitive to the length of the bond between boundary QM atom and the hydrogen atom introduced [137]. The problems with positioning of link atoms are considered by numerous authors. For example, authors of Ref. [38] proposed a special procedure for geometry optimization with rigid restrictions on the position of link atom. More complex is the so-called scaled position link atom method (SPLAM) [121]. It requires to introduce bonding, dipole, and van-der-Waals corrections. In some cases it works significantly better than the simple link atom method but the status of result is still unclear. Moreover some numerical estimates performed with use of the SPLAM in Ref. [121] can not be considered as successful. We can see the typical failure of the QM/MM schemes since it gives the results which are quite close to the pure MM ones. In this case the use of the hybrid QM/MM approach becomes senseless. For example, consideration of the water dimer have shown that the BLYP QM method predicts the OO distance and the HO...H angle (2.98

Å and 123° , respectively) very close to the experiment (2.98 Å and 122°), while the MM scheme works significantly worse (2.77 Å and 162°) and corresponding hybrid scheme gives values almost coinciding with those of the MM (2.78 Å and 163°). It witnesses the fact that some important contributions are missed in such approach.

It is seen that the implementation of the link atom method is generally not a trivial task. The most simple for implementation model closely related to the link atom method [110] is the so-called "subtractive scheme" implemented for example in Refs. [59, 61]. In the framework of these schemes the dangling bonds of the QM fragment are saturated by some groups (hydrogens, methyl groups *etc.*) and form the so-called model system. The whole system is calculated by a MM method of choice (or, more generally, by whatever low-level method, may be a QM one) and the energy is obtained by adding the QM (high-level) calculated energy of the model subsystem and subtracting the MM calculated energy of the model subsystem:

$$E_{tot} = E_{low}(real) + E_{high}(model) - E_{low}(model), \quad (49)$$

where "high" and "low" refers to the levels of approximation while "real" and "model" – to the type of the calculated system. The above expression obscures the need for explicit formulation of the properties of the boundary between subsystems. This scheme has obvious drawbacks. It takes into account interaction between subsystems only on the classical level and thus the electrostatic polarization of the QM region by the MM-treated environment (and *vice versa*) can not be taken into account in this approach. Moreover, description of reaction center by molecular mechanics in many cases can be quite problematic (even for the purpose of relative error evaluation), for example, due to absence of required force field parameters (especially, if transition states are considered). It should be noted that the errors behave irregularly with changes of the geometry: the MM scheme can perfectly describe the system near equilibrium and totally fail near the saddle point. Saturation of broken bonds by hydrogens (or other groups) can essentially affect the results of electronic structure calculations. Therefore, QM subsystem should be chosen to be quite large. The validity of the subtractive scheme is almost explicitly based on more or less accidental compensation of errors. The most simple implementation of this scheme is the IMOMM methodology [38]. The IMOMO scheme is the analogous procedure of combining two QM (MO-based) methods. According to Ref. [138] these approaches combining two subsystems are called two-layered. The subtractive approach combining more than two regions (i.e., including some intermediate buffers) is called ONIOM [138]. The expression for the total

energy in the framework of the ONIOM scheme is written analogously to that of Eq. (49) and in the case of three-level model is:

$$E_{tot} = E_{low}(real) + E_{high}(S - model) - E_{med}(S - model) + E_{med}(I - model) - E_{low}(I - model), \quad (50)$$

where "med" corresponds to the medium level of approximation, while S-model and I-model correspond to the small and intermediate model systems, respectively. While using the ONIOM scheme it is explicitly prescribed by its authors [138] to estimate the energy difference incurred by transition from a model molecule to a more realistic one and to use that pair of methods as high-/low-level ones for which these differences are close enough. This procedure is of course nothing else but a check whether the errors are likely to compensate. It should be noted that the authors of Ref. [139] made the QM/MM boundary in their implementation of the IMOMM method for surfaces more flexible to smooth the consequences of the molecule partitioning.

The main features of the "subtractive" hybrid schemes can be illustrated by numerical examples. To understand better the problems caused by the *ad hoc* way of junction construction we consider characteristic examples where these schemes fail. In the most cases the sources of failures can be seen on the simple molecular objects which serve as tests for one or another hybrid scheme. As an example we consider an application of the IMOMM method to conformational properties of cis-butane performed in Ref. [113]. It was shown that marking two C end atoms of the molecule as QM ones leads to valence angles 129.9° , while pure QM calculation gives 117° and pure MM calculation predicts the value 116.1° for these angles. Therefore, we see the case when transition to QM/MM procedure spoils even the results of pure MM calculations. Problems also arise when multilayered schemes are used. For example, the energy of reaction of the oxidative addition of H_2 to $Pt(P(t-Bu)_3)_2$ calculated by the B3LYP:HF:MM3 scheme is 7.9 kcal/mol smaller than that calculated by the B3LYP:HF:HF scheme [138]. The same difference in 7.9 kcal/mol is obtained for schemes B3LYP:B3LYP:MM3 and B3LYP:B3LYP:HF. Therefore, the choice of the description for the third layer (the most inert subsystem) turns out to be crucial for the description of the energy of this reaction which is estimated to be ~ 4 kcal/mol [138]. It speaks on the incorrect construction of junction in this scheme.

Principally different and more justified type of construction of junction between QM and MM parts of the molecule is based on the use of local one-electron functions. Only this type junctions cutting the nuclei and not the bonds can be justified from the general point of view. In the local self-consistent field (LSCF) approach [41, 107, 140, 141] the chemi-

cal bonds between QM and MM regions are represented by SLBOs. The BOs can be obtained by whatever of *a posteriori* localization procedures known in the literature. At the same time the localized orbitals obtained in the localization procedures have some degree of delocalization, *i.e.* they have non-zero contributions for the atoms not constituting a given bond (or a lone pair) mimicked by this particular BO. These contributions are called the "tails" of the localized orbitals and their neglect lead to the SLBOs which are used in the LSCF scheme. An important assumption made in the LSCF construction is that of the transferability of the SLBOs in very wide ranges. The QM part of the system is described by a set of delocalized MOs while the boundary is modeled by the frozen SLBOs. The frozen character of the boundary SLBO may cause the sensitivity of the LSCF results to the size of the QM region. The determination of the electronic structure of the QM region is based on the diagonalization of the modified Fock operator which includes Coulomb and exchange interaction with boundary orbitals and interaction with the MM charges. In its original implementation the LSCF method was developed by fixing the positions of the atoms of the environment and thus it was not suitable for geometry optimization procedures. This restriction was avoided in Ref. [142] by special adjustment of the force field parameters. The authors of Ref. [142] have noticed that in the framework of original LSCF scheme the ion-nuclei interactions are underestimated and the variation of the overlap between boundary basis functions with respect to the boundary bond length is not taken into account. To cure these problems they defined a 5-components boundary bond potential

$$E_{X-Y} = (A + Br + Cr^2)e^{Dr} + \frac{E}{r} \quad (51)$$

where the first contribution describes an overlap and the second contribution describes an interaction of effective charges with parameters A - E numerically optimized. In fact, the first contribution should correspond to the intrabond resonance and one has to consider the potential Eq. (51) mainly as a correction for the bonding (the overlap dependent contribution). At the same time the particular values of the parameters A - E seem not to correspond to their declared physical meaning. The same conclusion can be drawn from the energy profile for the process of the boundary bond elongation – for large values of bond length the difference between LSCF and SCF curves drastically increases since the unphysically large Coulomb contribution becomes prevailing. Moreover, optimized bond lengths in the SCF and the LSCF/MM approaches can differ significantly and also remarkably depend on the environment the

bond is assigned to (for example, C(QM)-N(MM) bond is by 0.024 Å longer than the bond N(QM)-C(MM) with the same hybridization). The problems with the correction of bond description by the potential Eq. (51) are caused by the number and type of factors it is designed to reproduce – character of boundary orbitals is a function of geometry (see Eqs. (31) and (33)), elements of intrabond density matrices (especially, of the two-electron ones) also depend on the geometry (see Eqs. (20) and (21)) as important examples. Moreover, the form of the bond energy is not arbitrary and can be defined from the analysis of particular QM expression (see Eq. (17)). The angular dependence of the bond potential was not considered is totally neglected in the potential Eq. (51), while it naturally appears in the subsequent derivation of intersubsystem junction due to angular dependence of resonance interactions. It should be stated that the LSCF approach has two important drawbacks by construction: the procedure of separation of the system in parts is not flexible and not justified; the parameters for the SLBOs should be determined from model molecules for each new system. Construction of extensive database of SLBO parameters is considered as a strategy within this approach.

It is interesting to compare the possibilities and errors of different hybrid QM/MM schemes. The careful examination and comparison of link atom and LSCF techniques was performed in Ref. [128] using the CHARMM force field [114] and the AM1 method [143] as a quantum chemical procedure. In the case of the link atom procedure two options were used: QQ – the link atom does not interact with the MM subsystem and HQ – link atom interacts with all MM atoms. The main conclusion of this consideration is that the LSCF and the link atom schemes are of similar quality. The error in the proton affinity determination induced by these schemes is several kcal/mol. It is noteworthy that all the schemes work rather badly in description of conformational properties of n-butane. The large charge on the MM atoms in the proximity of the QM subsystem (especially on the boundary atom) cause significant errors in the proton affinity estimates for all methods (especially, in the case of the LSCF approach where the error can be of tens of kcal/mol). This is not surprising since the stability and transferability of intrabond one- and two-electron density matrix elements Eq. (19) is broken here. It proves that the simple electrostatic model is not well appropriate for these schemes and that a detailed analysis of the junction form is necessary in general case. Moreover, the numerical analysis shows that the error induced by the HQ model is less than that induced by the QQ model. Since the HQ model explicitly includes unphysical interactions with artificial link atom it means that these interactions partly compen-

sate errors in junction construction. Practically, the link atom interacts with the MM atoms even in the case of the QQ model but indirectly via the interactions of QM subsystem with the MM atoms. In the case of the QQ model non-compensated charge on the QM subsystem (without link atom) interacts with MM atomic charges. It causes significant polarization of the QM subsystem. It is confirmed by numerical estimates of charges on the QM subsystem [128].

An attempt to escape the need of constructing large databases in the LSCF approach was made by Gao *et al* in the generalized hybrid orbital (GHO) method [144, 145]. This approach is intended to interpolate the ESPs related to the shape and orientation of the HOs residing on the frontier atoms. The first step in the GHO method is dividing the hybrid orbitals into active and auxiliary ones – the first ones are added to the QM subsystem. In the original LSCF approach three orbitals of the boundary atom are included into the self-consistent procedure. In the GHO approach only one active orbital of the boundary atom participates in the set of orbitals subjected to the SCF procedure while the auxiliary orbitals mimic the effective core potential. Therefore, we can characterize the boundary atom in the GHO scheme as the MM one while the boundary atom in the LSCF scheme is the QM one. It allows to substitute the parameterization of the one-electron density on the HO for each molecule by choosing semiempirical parameters for boundary atom with assumption of their transferability. At the same time the possibility to choose reasonable and transferable semiempirical parameters for boundary atom must rely upon correct determination of the HOs. Conversely, the GHO scheme uses very crude assumptions [145] about the form and the direction of HOs based on the pure geometry considerations which are equivalent to assumptions of (i) C_{3v} symmetry of the local MM environment and (ii) coincidence of all the HOs directions with the directions of bonds. Practically, these conditions are satisfied only for methane molecule. Moreover, even for these assumptions the s -/ p -ratio for active orbital is determined by incorrect formula working only for equivalent active and auxiliary orbitals or for purely p active orbital. The real structure of the HOs as a function of the geometry is given by the expressions of the previous Section and we note that neither of the assumptions accepted in Ref. [145] is fulfilled. Neither asymmetry of the geometry of the boundary atom environment nor the chemical nature of neighbouring groups are taken into account in Ref. [145]. In practical implementation of Refs. [144, 145] all the HOs are determined by directions of the bonds between the boundary atom and its three MM neighbours. The coincidence of HOs and bond directions is a rather crude and uncontrollable assumption. In the case of molecules

with large asymmetry or significant geometry constraints the difference between directions of bonds and HOs can exceed 20° with cyclopropane as the most characteristic example. Since the s -/ p - ratio and direction of the active HO in the GHO method are totally the functions of only MM atoms positions the form of this HO may be far from being optimal to say nothing of its behaviour with geometry variations. In our opinion the non-variational form of the HOs has led the authors of Ref. [144] to significant and hardly justifiable renormalization of the Hamiltonian parameters and the MM force field to reproduce correct bond lengths, direction of auxiliary orbitals along the corresponding bonds, and the Mulliken charges on the atoms, so, that the boundary could not be considered as a weak perturbation anymore. For example, in the case of carbon atom the resonance parameter β_s should be changed by more than 10 eV; the MM C-C ideal bond length parameter r_0 should be changed by 0.05 Å. It should be noted that the change of the delocalized description of electronic structure to the localized one does not lead to significant changes in the parameters of the semiempirical Hamiltonian as it was exemplified on the examples of MINDO/3 [26], MNDO, AM1, and PM3 [27] schemes.

The principles similar to those of the LSCF are used for junction construction in Ref. [146] based on the fragment SCF method. Another model thoroughly elaborated to consider effect of motions of environment atoms on the *ab initio* level is that by Philipp and Friesner [147]. In the framework of this model the procedure of freezing the SLBO was refined and the degree of delocalization of SLBOs is extended from two boundary atoms to two boundary groups. At the same time an attempt to obtain numerical results of good quality has led authors to introducing some very artificial procedures in the construction of the junction like placement of very large charges on the bond. Moreover, this model requires extensive parameterization. To reproduce energetics of alanine dipeptide and tetrapeptide 27 parameters describing interface between subsystems were adjusted. Essential parameterization process is quite important problem of this model. It would be fair to say that reproduction of conformational energies for a polar system like polypeptide is a difficult test for any computational methodology. The authors of Ref. [147] give the essential requirements for the bond at which molecule is cut: (i) this bond should not have significant multiple bonding character, and (ii) this bond should be far away from the region where significant electron re-distribution occurs. The first requirement seems to be quite reasonable for description of bond by single localized orbital, while the second one reflects a lack of some contributions to the energy which are important for small QM subsystems chosen. The estimate of these

contributions is not possible due to *ad hoc* way of construction of intersubsystem junction.

The method based on the effective fragment potential (EFP) construction implemented in Ref. [54] is close to the general LSCF methodology. In this case the boundary is modeled by a buffer region consisting of several localized molecular orbitals which are obtained by a QM calculation on all or a subset of the system. These orbitals are set frozen in the EFP calculation. The orbitals of the QM part are forced to be orthogonal to those of the buffer region and the environment is represented by an EFP. The important idea of this approach is to make the distance between the QM and EFP regions significant with the purpose to present the corresponding intersubsystem interactions as nonbonded ones. At the same time the frozen character of the buffer one-electron states can be important since the changes in polarization contributions from the buffer region are neglected during a geometry optimization. The numerical example given by authors of Ref. [54] is quite characteristic. They calculated the proton affinities of lysine and H-bonded and non-H-bonded tripeptide Gly-Lys-Gly by the QM/buffer/EFP method. If the buffer region is chosen as γ - and δ -CH₂ groups of lysine chain (*i.e.*, quite far from the reaction center) the QM/buffer/EFP calculation gives the value of the proton affinity 2.2 kcal/mol higher than the reference QM one for all these molecules. It unequivocally testifies that these 2.2 kcal/mol constitute the error of junction construction in this case which seems to be quite large. Moreover, the difference between QM/buffer and QM/buffer/EFP results for proton affinities of lysine and non-H-bonded tripeptide Gly-Lys-Gly is only 0.2 kcal, *i.e.* the effect of environment described by the EFP is by an order of magnitude smaller than the error produced by the junction.

4. Conclusion

The hybrid QM/MM modeling is quite promising for study of large molecules especially in the rapidly growing field of computational biological chemistry. The case of covalent linkage between the QM and MM parts is especially important since it includes the enzymatic catalysis. At the same time this case is the most complex one since the boundary between subsystems is not well defined and construction of intersubsystem junction is not straightforward. In this case many *ad hoc* recipes are proposed in the literature. We have shown the state-of-the-art in this field in somewhat critical manner with special attention to the problems arising during the QM/MM modeling. At the same time the sequential derivations of QM/MM junctions are shown to be possible. We have

given a series of physical principles which in our opinion should govern the construction of hybrid QM/MM schemes. As particular implementations of these principles we have demonstrated how the effective electronic Hamiltonian approach can be applied to the purpose of junction construction. It has required the use of quantum chemical description underlying the MM one. The APSLG-type trial wave function was taken for this description.

It should be noted that the important question of structure of boundary one-electron states (especially as a function of molecular geometry) remained unclear. We proposed a way of determination of these states based on the derivation of the MM description from the QM (APSLG) one. The derivation of the deductive MM allows to determine the effects of the MM subsystem on the QM one (renormalization of parameters) and of the QM subsystem on the MM one (torques and forces acting on the MM atoms). Explicit expressions are given for example of boundary sp^3 carbon atom. Numerical estimates obtained for different QM/MM schemes illustrate the general points stressed in the text.

Acknowledgments

Financial support for AMT on the part of the Haldor Topsoe A/S is gratefully acknowledged. Authors are grateful to Professor J.-L. Rivail for the reprints and preprints of his works on the hybrid QM/MM methods, to Prof. I.V. Abarenkov for permission to use his work prior of publication and to Dr. X. Assfeld and Dr. A.A. Sokol for valuable discussions.

References

- [1] J. A. Pople Quantum chemical models (Nobel lecture), *Angew. Chem. Int. Ed.* **38**, 1894-1902 (1999).
- [2] W. Kohn Density functional and density matrix method scaling linearly with the number of atoms, *Phys. Rev. Lett.* **76**, 3168-3171 (1996).
- [3] S. Goedecker Decay properties of the finite-temperature density matrix in metals, *Phys. Rev. B* **58**, 3501-3502 (1998).
- [4] K. N. Kudin and G. Scuseria A fast multipole method for periodic systems with arbitrary unit cell geometries, *Chem. Phys. Lett.* **283**, 61-68 (1998).
- [5] W. Yang Direct calculation of electron density in density-functional theory, *Phys. Rev. Lett.* **66**, 1438-1441 (1991).
- [6] S. L. Dixon, K. M. Merz Jr. Fast, accurate semiempirical molecular orbital calculations for macromolecules, *J. Chem. Phys.* **107**, 879-893 (1997).
- [7] J. J. P. Stewart Application of localized molecular orbitals to the solution of semiempirical self-consistent field equations, *Int. J. Quantum Chem.* **58**, 133-146 (1996).

- [8] S. Goedecker and L. Colombo Efficient linear scaling algorithm for tight-binding molecular dynamics, *Phys. Rev. Lett.* **73**, 122-125 (1994).
- [9] S. Goedecker and M. Teter Tight-binding electronic-structure calculations and tight-binding molecular dynamics with localized orbitals, *Phys. Rev. B* **51**, 9455-9464 (1995).
- [10] R. Baer and M. Head-Gordon Electronic structure of large systems: Coping with small gaps using the energy renormalization group method, *J. Chem. Phys.* **109**, 10159-10168 (1998).
- [11] R. Haydock, V. Heine and M. J. Kelly Electronic structure based on the local atomic environment for tight-binding bands, *J. Phys. C* **8**, 2591-2605 (1975).
- [12] W. Yang Absolute-energy-minimum principles for linear-scaling electronic-structure calculations, *Phys. Rev. B* **56**, 9294-9297 (1997).
- [13] F. Mauri, G. Galli and R. Car Orbital formulation for electronic structure calculations with linear system size scaling, *Phys. Rev. B* **47**, 9973-9976 (1993).
- [14] P. Ordejón, D. Drabold, M. Grumbach and R. M. Martin Unconstrained minimization approach for electronic computations that scales linearly with system size, *Phys. Rev. B* **48**, 14646-14649 (1993).
- [15] M. Challacombe A simplified density matrix minimization for linear scaling self-consistent field theory, *J. Chem. Phys.* **110**, 2332-2342 (1999).
- [16] X.-P. Li, R. W. Nunes and D. Vanderbilt Density matrix electronic structure method with linear system size scaling, *Phys. Rev. B* **47**, 10891-10894 (1993).
- [17] S. Goedecker Linear scaling electronic structure methods, *Rev. Mod. Phys.* **71**, 1085-1123 (1999).
- [18] S. Y. Wu and C. S. Jayanthi Order-N methodologies and their applications, *Physics Reports* **358**, 1-74 (2002).
- [19] S. R. White Density matrix formulation for quantum renormalization groups, *Phys. Rev. Lett.* **69**, 2863-2866 (1992).
- [20] S. R. White and R. L. Martin *Ab initio* quantum chemistry using the density matrix renormalization group, *J. Chem. Phys.* **110**, 4127-4130 (1999).
- [21] M. Schütz, M. Hetzer and H.-I. Werner Low-order scaling local electron correlation methods. I. Linear scaling local MP2, *J. Chem. Phys.* **111**, 5691-5705 (1999).
- [22] J. Rubio, A. Povill, J. P. Malrieu and P. Reinhardt Direct determination of localized Hartree-Fock orbitals as a step toward N scaling procedures, *J. Chem. Phys.* **107**, 10044-10050 (1997).
- [23] M. C. Strain, G. E. Scuseria and M. J. Frisch Achieving linear scaling for the electronic quantum Coulomb problem, *Science* **271**, 51-53 (1996).
- [24] A. Derecskei-Kovacs, D. E. Woon and D. S. Marynick Nonempirical wave functions for very large molecules. II. The PRDDO/M/FCP method, *Int. J. Quantum Chem.* **61**: 67-76 (1997).
- [25] J. M. Cullen A rapid generalized valence-bond algorithm for semiempirical NDDO calculations, *Int. J. Quantum Chem.* **56**, 97-113 (1995).
- [26] A. M. Tokmachev and A. L. Tchougréeff Semiempirical implementation of strictly localized geminals for analysis of molecular electronic structure, *J. Comp. Chem.* **22**, 752-764 (2001).

- [27] A. M. Tokmachev and A. L. Tchougréeff Fast NDDO method for molecular structure calculations based on strictly localized geminals, *J. Phys. Chem. A*, accepted.
- [28] D. York, T. S. Lee and W. Yang Quantum mechanical study of aqueous polarization effects on biological macromolecules, *J. Am. Chem. Soc.* **118**, 10940-10941 (1996).
- [29] J. P. Lewis, C. W. Carter Jr., J. Herman, W. Pan, T. S. Lee and W. Yang Active species for the ground state complex of cytidine deaminase: A linear scaling quantum mechanical investigation, *J. Am. Chem. Soc.* **120**, 5407-5410 (1998).
- [30] U. Burkert and N. Allinger, *Molecular Mechanics* (ACS, Washington, DC, 1982).
- [31] M. B. Darkhovskii and A. L. Tchougréeff On unification of quantum chemistry and molecular mechanics. Potential energy surfaces of transition metal complexes as exemplified by spin transition in cis- [Fe(NCS)2(bipy)2], *Khim. Fiz.* **18**, 73-79 (1999) [*Chem. Phys. Reports* **18**, 149 (1999)].
- [32] A. Warshel and M. Levitt Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme, *J. Mol. Biol.* **103**, 227-249 (1976).
- [33] U. C. Singh and P. A. Kollman A combined *ab initio* quantum mechanical and molecular mechanical method for carrying out simulations on complex molecular systems: Application to the CH₃Cl+Cl⁻ exchange reaction and gas-phase protonation of polyethers, *J. Comp. Chem.* **7**, 718-730 (1986).
- [34] M. J. Field, P. A. Bash and M. Karplus A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulation, *J. Comp. Chem.* **11**, 700-733 (1990).
- [35] F. Bernardi, M. Olivucci and M. A. Robb Simulation of MC-SCF results on covalent organic multi-bond reactions: Molecular mechanics with valence bond (MM-VB), *J. Am. Chem. Soc.* **114**, 1606-1616 (1992).
- [36] D. Bakowies and W. Thiel Semiempirical treatment of electrostatic potentials and partial charges in combined quantum mechanical and molecular mechanical approaches, *J. Comp. Chem.* **17**, 87-108 (1996).
- [37] I. H. Hillier Chemical reactivity studied by hybrid QM/MM methods, *J. Mol. Struct. (Theochem)* **463**, 45-52 (1999).
- [38] F. Maseras and K. Morokuma IMOMM: A new integrated *ab initio* + molecular mechanics geometry optimization scheme of equilibrium structures and transition states, *J. Comp. Chem.* **16**, 1170-1179 (1995).
- [39] J. Sauer and M. Sierka Combining quantum mechanics and interatomic potential functions in *ab initio* studies of extended systems, *J. Comp. Chem.* **21**, 1470-1493 (2000).
- [40] M. A. Thompson QM/MMpol: A consistent model for solute/solvent polarization. Application to the aqueous solvation and spectroscopy of formaldehyde, acetaldehyde, and acetone, *J. Phys. Chem.* **100**, 14492-14507 (1996).
- [41] X. Assfeld and J. L. Rivail Quantum chemical computations on parts of large molecules: The *ab initio* local self consistent field method, *Chem. Phys. Lett.* **263**, 100-106 (1996).
- [42] J. Gao "Methods and Applications of Combined Quantum Mechanical and Molecular Mechanical Potentials." In *Reviews in Computational Chemistry* (VCH Publishers, Inc., New York, 1996).

- [43] P. Sherwood, "Hybrid Quantum Mechanics/Molecular Mechanics Approaches." In *Modern Methods and Algorithms of Quantum Chemistry*, Proceedings, Second edition, J. Grotendorst, ed., NIC Series, 3 (John von Neumann Institute for Computing, Jülich, 2000).
- [44] T. E. Peacock, *Electronic Properties of Aromatic and Heterocyclic Molecules* (Academic Press, London, 1965).
- [45] K. F. Freed, "Theoretical Basis for Semiempirical Theories." In *Semiempirical Methods of Electronic Structure Calculations*. Part A. Techniques. G.A. Segal, ed. (Plenum Press, New York, 1977).
- [46] K. Ohta, Y. Yoshioka, K. Morokuma and K. Kitaura The effective fragment potential method. An approximate *ab initio* MO method for large molecules, *Chem. Phys. Lett.* **101**, 12-17 (1983).
- [47] J. A. Mejias and J. F. Sanz Unrestricted compact model potentials for *ab initio* embedded cluster calculations: Magnetic interactions in KNiF_3 , *J. Chem. Phys.* **102**, 850-858 (1995).
- [48] R. Colle, A. Curioni and O. Salvetti A non-local representation of the effective potential due to a molecular fragment, *Theoret. Chim. Acta* **86**, 451-465 (1993).
- [49] I. Frank, S. Grimme, M. von Armin and S. D. Peyerimhoff The solvent shift in the $n \rightarrow p^*$ excitation of $\text{CH}_2\text{O} \cdot \text{H}_2\text{O}$: An MRD-CI investigation using effective potentials for the representation of the water molecules, *Chem. Phys.* **199**, 145-153 (1995).
- [50] S. Katsuki The spectral representation technique for an active-electron-only molecular orbital calculation applied to the case of nonspherical frozen charge, *J. Chem. Phys.* **98**, 496-501 (1993).
- [51] I. V. Abarenkov and I. M. Antonova Chemical bond modeling with the energy driven orbital localization, *Int. J. Quantum Chem.*, Submitted.
- [52] R. Poteau, I. Ortega, F. Alary, A. Ramirez Solis, J.-C. Barthelat and J. P. Daudey Effective group potentials. I. Method, *J. Phys. Chem. A* **105**, 198-205 (2001).
- [53] V. Kairys and J. H. Jensen QM/MM boundaries across covalent bonds: A frozen localized molecular orbital-based approach for the effective fragment potential method, *J. Phys. Chem. A* **104**, 6656-6665 (2000).
- [54] M. S. Gordon, M. A. Freitag, P. Bandyopadhyay, J. H. Jensen, V. Kairys and W. J. Stevens The effective fragment potential method: A QM-based MM approach to modeling environmental effects in chemistry, *J. Phys. Chem. A* **105**, 293-307 (2001).
- [55] B. O. Roos The complete active space self-consistent field method and its applications in electronic structure calculations, *Adv. Chem. Phys.* **69**, 399-446 (1987).
- [56] S. Clifford, M. J. Bearpark and M. A. Robb A hybrid MC-SCF method: generalised valence bond (GVB) with complete active space SCF (CASSCF), *Chem. Phys. Lett.* **256**, 320-326 (1996).
- [57] A. V. Soudackov, A. L. Tchougréeff and I. A. Misurkin Ground-state multiplicities and d-d excitations of transition-metal complexes by effective Hamiltonian method, *Int. J. Quantum Chem.* **58**, 161-173 (1996).
- [58] A. M. Tokmachev, A. L. Tchougréeff and I. A. Misurkin Effective Hamiltonian approach to catalytic activity of transition metal complexes, *Int. J. Quantum Chem.* **84**, 99-109 (2001).

- [59] S. Humbel, S. Sieber and K. Morokuma The IMOMO method: Integration of different levels of molecular orbital approximations for geometry optimization of large systems: Test for n-butane conformation and S_N2 reaction: $RC1+Cl^-$, *J. Chem. Phys.* **105**, 1959-1967 (1996).
- [60] M. Svensson, S. Humbel and K. Morokuma Energetics using the single point IMOMO (integrated molecular orbital + molecular orbital) calculations: Choices of computational levels and model system, *J. Chem. Phys.* **105**, 3654-3661 (1996).
- [61] R. D. J. Froese, S. Humbel, M. Svensson and K. Morokuma IMOMO(G2MS): A new high-level G2-like method for large molecules and its applications to Diels-Adler reactions, *J. Phys. Chem. A* **101**, 227-233 (1997).
- [62] Th. Vreven and K. Morokuma On the application of the IMOMO (integrated molecular orbital + molecular orbital) method, *J. Comp. Chem.* **21**, 1419-1432 (2000).
- [63] A. V. Nemukhin, B. L. Grigorenko, E. Ya. Skasyrskaya, I. A. Topol and S. K. Burt A new hybrid approach for modeling reactions in molecular clusters: Application for the hydrogen bonded systems, *J. Chem. Phys.* **112**, 513-521 (2000).
- [64] B. L. Grigorenko, A. V. Nemukhin and V. A. Apkarian Many-body potentials and dynamics based on diatomics-in-molecules: Vibrational frequency shifts in ArnHF (n=112,62) clusters, *J. Chem. Phys.* **104**, 5510-5516 (1996).
- [65] G. Naray-Szabo and P. Surjan Bond orbital framework for rapid calculation of environmental effects on molecular potential energy surfaces, *Chem. Phys. Lett.* **96**, 499-501 (1983).
- [66] G. G. Ferenczy, J. L. Rivail, P. R. Surjan and G. Naray-Szabo NDDO fragment SCF approximation for large electronic systems, *J. Comp. Chem.* **13**, 830-837 (1992).
- [67] Th. N. Truong and E. V. Stefanovich Development of a perturbative approach for Monte Carlo simulations using a hybrid *ab initio* QM/MM method, *Chem. Phys. Lett.* **256**, 348-352 (1996).
- [68] I. Tunon, I. T. C. Martins-Costa, C. Millot, M. F. Ruiz-Lopez and J. L. Rivail A coupled density functional-molecular mechanics Monte Carlo simulation method, *J. Comp. Chem.* **17**, 19-29 (1996).
- [69] Ch. D. Berweger, W. F. van Gunsteren and F. Muller-Plathe Finite element interpolation for combined classical/quantum mechanical molecular dynamics simulations, *J. Comp. Chem.* **18**, 1484-1495 (1997).
- [70] O. A. Sharafeddin, K. Hinsien, T. Carrington Jr. and B. Roux Mixing quantum-classical molecular dynamics methods applied to intramolecular proton transfer in acetylacetone, *J. Comp. Chem.* **18**, 1760-1772 (1997).
- [71] P. Grochowski, B. Lesyng, P. Bala and J. A. McCammon Density-functional based parametrization of a valence bond method and its applications in quantum-classical molecular dynamics simulations of enzymatic reactions, *Int. J. Quantum Chem.* **60**, 1143-1164 (1996).
- [72] S. Chalmet, D. Rinaldi and M. F. Ruiz-Lopez A QM/MM/continuum model for computations in solution: Comparison with QM/MM molecular dynamics simulations, *Int. J. Quantum Chem.* **84**, 559-564 (2001).
- [73] A. L. Tchougreeff Group functions, Lowdin partition, and hybrid QC/MM methods for large molecular systems, *Phys. Chem. Chem. Phys.* **1**, 1051-1060 (1999).

- [74] A. M. Tokmachev, A. L. Tchougréeff and I. A. Misurkin Effective electronic Hamiltonian for quantum subsystem in hybrid QM/MM methods as derived from APSLG description of electronic structure of classical part of molecular system, *J. Mol. Struct. (Theochem)* **506**, 17-34 (2000).
- [75] M. V. Basilevsky and G. E. Chudinov Dynamics of charge transfer chemical reactions in a polar medium within the scope of the Born-Kirkwood-Onsager model, *Chem. Phys.* **157**, 327-344 (1991).
- [76] J. Tomasi and M. Perisco Molecular interactions in solution: An overview of methods based on continuous distributions of the solvent, *Chem. Rev.* **94**, 2027-2094 (1994).
- [77] T. Mineva, N. Russo and E. Sicilia Solvation effects on reaction profiles by the polarizable continuum model coupled with the Gaussian density functional method, *J. Comp. Chem.* **19**, 290-299 (1998).
- [78] V. Barone, M. Cossi and J. Tomasi Geometry optimization of molecular structures in solution by the polarizable continuum model, *J. Comp. Chem.* **19**, 404-417 (1998).
- [79] D. Wei and D. R. Salahub A combined density functional and molecular mechanics simulation of a quantum water molecule in aqueous solution, *Chem. Phys. Lett.* **224**, 291-296 (1994).
- [80] M. Orozco, J. M. López, C. Colomines, C. Alhambra, M. A. Busquets and F. J. Luque Theoretical representation of solvation in biochemical systems: From discrete solute-solvent interactions to bulk solvation, *Int. J. Quantum Chem.* **60**, 1179-1187 (1996).
- [81] H. Takahashi, T. Hori, H. Hashimoto and T. Nitta A hybrid QM/MM method employing real space grids for QM water in the TIP4P water solvents, *J. Comp. Chem.* **22**, 1252-1261 (2001).
- [82] J. Gao Energy components of aqueous solution: Insight from hybrid QM/MM simulations using a polarizable solvent model, *J. Comp. Chem.* **18**, 1061-1071 (1997).
- [83] P. Th. van Duijnen and A. H. de Vries Direct reaction field force field: A consistent way to connect and combine quantum-chemical and classical descriptions of molecules, *Int. J. Quantum Chem.* **60**, 1111-1132 (1996).
- [84] G. L. Heard and B. F. Yates Hybrid supermolecule-polarizable continuum approach to solvation: Application to the mechanism of the Stevens rearrangement, *J. Comp. Chem.* **17**, 1444-1452 (1996).
- [85] A. L. Tchougréeff $SO(4)$ group and deductive molecular mechanics, *J. Mol. Struct. (Theochem)*, submitted.
- [86] U. Eichler, Ch. M. Kölmel and J. Sauer Combining *ab initio* techniques with analytical potential functions for structure predictions of large systems: Method and application to crystalline silica polymorphs, *J. Comp. Chem.* **18**, 463-477 (1997).
- [87] S. P. Greatbanks, I. H. Hillier and P. Sherwood Comparison of embedded cluster models to study zeolite catalysis: Proton transfer reactions in acidic chabazite, *J. Comp. Chem.* **18**, 562-568 (1997).
- [88] R. Constanciel, In *Localization and Delocalization in Quantum Chemistry*, O. Chalvet, R. Daudel, S. Diner, J.P. Malrieu, eds. (D. Reidel Publishers, Dordrecht, 1975).

- [89] P.-O. Löwdin Studies in perturbation theory. IV. Solution of eigenvalue problem by projection operator formalism, *J. Math. Phys.* **3**, 969-982 (1962).
- [90] R. McWeeny, *Methods of Molecular Quantum Mechanics*, Second edition (Academic Press, London 1992).
- [91] A. M. Tokmachev and A. L. Tchougréeff Potential energy surfaces in hybrid quantum mechanical/molecular mechanical methods, *Int. J. Quantum Chem.* **84**, 39-47 (2001).
- [92] P. R. Surján, In *Theoretical Models of Chemical Bonding*, Part 2, Maksić, Z.B., ed., (Springer, Heidelberg, 1989).
- [93] A. M. Tokmachev and A. L. Tchougréeff Generic molecular mechanics as based on local quantum description of molecular electronic structure, *Int. J. Quantum Chem.* **88**, 403-413 (2002).
- [94] A. L. Tchougréeff and A. M. Tokmachev Deductive molecular mechanics of sp^3 carbon atom, *Int. J. Quantum Chem.*, submitted.
- [95] A. L. Tchougréeff and A. M. Tokmachev, in preparation.
- [96] R. Hoffmann An extended Hückel theory. I. Hydrocarbons, *J. Chem. Phys.* **39**, 1397-1412 (1963).
- [97] M. Freindorf and J. Gao Optimization of the Lennard-Jones parameters for a combined *ab initio* quantum mechanical and molecular mechanical potential using the 3-21G basis set, *J. Comp. Chem.* **17**, 386-395 (1996).
- [98] F. J. Luque, M. Bachs, C. Alemán and M. Orozco Extension of MST/SCRF method to organic solvents: *Ab initio* and semiempirical parametrization for neutral solutes in CCl_4 , *J. Comp. Chem.* **17**, 806-820 (1996).
- [99] Y. Mo and J. Gao *Ab initio* QM/MM simulations with a molecular orbital-valence bond (MOVb) method: Application to an S_N2 reaction in water, *J. Comp. Chem.* **21**, 1458-1469 (2000).
- [100] T. Yoshida, N. Koga and K. Morokuma A combined *ab initio* MO-MM study on isotacticity control in propylene polymerization with silylene-bridged group 4 metallocenes. C_2 symmetrical and asymmetrical catalysts, *Organometallics* **15**, 766-777 (1996).
- [101] M. Sierka and J. Sauer Structure and reactivity of silic and zeolite catalysts by a combined quantum mechanics-shell-model potential approach based on DFT, *Faraday Discussions* **106**, 41-62 (1997).
- [102] R. V. Stanton, D. S. Hartsough and K. M. Merz Calculation of solvation free energies using a density functional/molecular dynamics coupled potential, *J. Phys. Chem.* **97**, 11868-11870 (1993).
- [103] P. D. Lyne, M. Hodoscek and M. Karplus A hybrid QM-MM potential employing Hartree-Fock or density functional methods in the quantum region, *J. Phys. Chem. A* **103**, 3462-3471 (1999).
- [104] X. P. Long, J. B. Nicholas, M. F. Guest and R. L. Ornstein A combined density functional theory/molecular mechanics formalism and its application to small water clusters, *J. Mol. Struct.* **412**, 121-133 (1997).
- [105] S. A. French, A. A. Sokol, S. T. Bromley, C. R. A. Catlow, S. C. Rogers, F. King and P. Sherwood From CO_2 to methanol by hybrid QM/MM embedding, *Angew. Chem.* **113**, 4569-4572 (2001).

- [106] P. L. Cummins and J. E. Gready Coupled semiempirical molecular orbital and molecular mechanics model (QM/MM) for organic molecules in a aqueous solution, *J. Comp. Chem.* **18**, 1496-1512 (1997).
- [107] V. Théry, D. Rinaldi, J.-L. Rivail, B. Maigret and G. G. Ferenczy Quantum mechanical computations on very large molecular systems: The local self-consistent field method, *J. Comp. Chem.* **15**, 269-282 (1994).
- [108] G. G. Ferenczy and J. G. Ángyán Intra- and intermolecular interactions in crystals of polar molecules. A study by the mixed quantum mechanics/molecular mechanical SCMP-NDDO method, *J. Comp. Chem.* **22**, 1679-1690 (2001).
- [109] J. Åqvist and A. Warshel Simulation of enzyme reactions using valence bond force fields and other hybrid quantum/classical approaches, *Chem. Rev.* **93**, 2523-2544 (1993).
- [110] D. Bakowies and W. Thiel Hybrid models for combined quantum mechanical and molecular mechanical approaches, *J. Phys. Chem.* **100**, 10580-10594 (1996).
- [111] N. L. Allinger, Y. H. Yuh and J.-H. Lii Molecular mechanics. The MM3 force field for hydrocarbons. 1, *J. Am. Chem. Soc.* **111**, 8551-8565 (1989).
- [112] A.L. Tchougréeff Separation of electronic variables in quantum chemical problems on potential energy surfaces of large molecular systems, *Khim. Fiz.* **16**, 62-77 (1997) [*Chem. Phys. Reports* **16**, 1035 (1997)].
- [113] T. Matsubara, S. Sieber and K. Morokuma A test of the new "integrated MO + MM" (IMOMM) method for the conformational energy of ethane and n-butane, *Int. J. Quantum Chem.* **60**, 1101-1109 (1996).
- [114] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus CHARMM - A program for macromolecular energy, minimization, and dynamics calculations, *J. Comp. Chem.* **4**, 187-217 (1983).
- [115] A. J. Turner, V. Moliner and I. H. Williams Transition-state structural refinement with GRACE and CHARMM: Flexible QM/MM modeling for lactate dehydrogenase, *Phys. Chem. Chem. Phys.* **1**, 1323-1331 (1999).
- [116] W. D. Cornell, P. Ciepak, C. I. Bayly, I. R. Gould, K. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. Kollman A 2nd generation of force field for the simulation of proteins, nucleic acids, and organic molecules, *J. Am. Chem. Soc.* **117**, 5179-5197 (1995).
- [117] B. Waszkowycz, I. H. Hillier, N. Gensmantel and D. W. Payling Combined quantum mechanical-molecular mechanical study of catalysis by the enzyme phospholipase-A2 - an investigation of the potential energy surface for amide hydrolysis, *J. Chem. Soc., Perkin Trans.* **2**, 2025-2032 (1991).
- [118] D. Yin and A. D. MacKerell Jr. Combined *ab initio*/empirical approach for optimization of Lennard-Jones parameters, *J. Comp. Chem.* **19**, 334-348 (1998).
- [119] V. V. Vasilyev, A. A. Bliznyuk and A. A. Voityuk A combined quantum chemical/molecular mechanical study of hydrogen-bonded systems, *Int. J. Quantum Chem.* **44**, 897-930 (1992).
- [120] M. A. Thompson, E. D. Glendening and D. Feller The nature of K⁺/crown ether interactions: A hybrid quantum mechanical-molecular mechanical study, *J. Phys. Chem.* **98**, 10465-10476 (1994).
- [121] M. Eichinger, P. Tavan, J. Hutter and M. Parrinello A hybrid method for solutes in complex solvents: Density functional theory combined with empirical force fields, *J. Chem. Phys.* **110**, 10452-10467 (1999).

- [122] M. A. Thompson and G. K. Schenter Excited states of the bacteriochlorophyll b dimer of rhodospseudomonas viridis: A QM/MM study of the photosynthetic reaction center that includes MM polarization, *J. Phys. Chem.* **99**, 6374-6386 (1995).
- [123] I. B. Bersuker, M. K. Leong, J. E. Boggs and R. S. Pearlman A method of combined quantum mechanical (QM)/molecular mechanics (MM) treatment of large polyatomic systems with charge transfer between the QM and MM fragments, *Int. J. Quantum Chem.* **63**, 1051-1063 (1997).
- [124] K. P. Eurenus, D. C. Chatfield, B. R. Brooks and M. Hodoscek Enzyme mechanisms with hybrid quantum and molecular mechanical potentials. I. Theoretical considerations, *Int. J. Quantum Chem.* **60**, 1189-1200 (1996).
- [125] M. J. Harrison, N. A. Burton and I. H. Hillier Catalytic mechanism of the enzyme papain: Predictions with a hybrid quantum mechanical/molecular mechanical potential, *J. Am. Chem. Soc.* **119**, 12285-12291 (1997).
- [126] V. V. Vasilyev *J. Mol. Struct. (Theochem)* **304**, 129 (1994).
- [127] A. H. de Vries, P. Sherwood, S. J. Colins, A. M. Rigby, M. Rigutto and G. J. Kramer Zeolite structure and reactivity by combined quantum-chemical-classical calculations, *J. Phys. Chem. B* **103**, 6133-6141 (1999).
- [128] N. Reuter, A. Dejaegere, B. Maigret and M. Karplus Frontier bonds in QM/MM methods: A comparison of different methods, *J. Phys. Chem. A* **104**, 1720-1735 (2000).
- [129] Hyperchem, Inc. *Hyperchem Users Manual*; Computational Chemistry; Hypercube, Inc.: Ontario, Canada, 1994.
- [130] P. Cummins and J. Gready Combined quantum and molecular mechanics (QM/MM) study of the ionization state of 8-methylpterin substrate bound to dihydrofolate reductase, *J. Phys. Chem. B* **104**, 4503-4510 (2000).
- [131] S. Ranganathan and J. Gready Hybrid quantum and molecular mechanical (QM/MM) studies on the pyruvate to L-lactate interconversion in L-lactate dehydrogenase, *J. Phys. Chem. B* **101**, 5614-5618 (1997).
- [132] I. Antes and W. Thiel Adjusted connection atoms for combined quantum mechanical and molecular mechanical methods, *J. Phys. Chem. A* **103**, 9290-9295 (1999).
- [133] Y. Zhang, T.S. Lee and W. Yang A pseudobond approach to combining quantum mechanical and molecular mechanical methods, *J. Chem. Phys.* **110**, 46-54 (1999).
- [134] Y. Zhang, H. Liu and W. Yang Free energy calculation on enzyme reactions with an efficient iterative procedure to determine minimum energy paths on a combined *ab initio* QM/MM potential energy surface, *J. Chem. Phys.* **112**, 3483-3492 (2000).
- [135] J. A. DiLabio, M. M. Hurley and P. A. Christiansen Simple one-electron quantum capping potentials for use in hybrid QM/MM studies of biological molecules, *J. Chem. Phys.* **116**, 9578-9584 (2002).
- [136] D. Bakowies, PhD Thesis, Universität Zürich, 1994.
- [137] G. J. Kramer and R. A. van Santen Theoretical determination of proton affinity differences in zeolites, *J. Am. Chem. Soc.* **115**, 2887-2897 (1993).

- [138] M. Svensson, S. Humbel, R. D. J. Froese, T. Matsubara, S. Sieber and K. Morokuma ONIOM: A multilayered integrated MO + MM method for geometry optimizations and single point energy predictions. A test for Diels-Alder reactions and $\text{Pt}(\text{P}(\text{t-Bu})_3)_2 + \text{H}_2$ oxidative addition, *J. Phys. Chem.* **100**, 19357-19363 (1996).
- [139] J. R. Shoemaker, L. W. Burggraf and M. S. Gordon SIMOMM: An integrated molecular orbital/molecular mechanics optimization scheme for surfaces, *J. Phys. Chem. A* **103**, 3245-3251 (1999).
- [140] G. Monard, M. Loos, V. Théry, K. Baka and J. L. Rivail Hybrid classical quantum force field for modeling very large molecules, *Int. J. Quantum Chem.* **58**, 153-159 (1996).
- [141] S. Antonczak, G. Monard, M. F. Ruiz-López and J. L. Rivail Modeling of peptide hydrolysis by thermolysin. A semiempirical and QM/MM study, *J. Am. Chem. Soc.* **120**, 8825-8833 (1998).
- [142] N. Ferré, X. Assfeld and J.-L. Rivail Specific force field parameters determination for the hybrid *ab initio* QM/MM LSCF method, *J. Comp. Chem.* **23**, 610-624 (2002).
- [143] M. J. S. Dewar, E. G. Zoebisch, E. Healy and J. J. P. Stewart AM1: A new general purpose quantum mechanical molecular model, *J. Am. Chem. Soc.* **107**, 3902-3909 (1985).
- [144] J. Gao, P. Amara, C. Alhambra and M. J. Field A generalized hybrid orbital (GHO) method for the treatment of boundary atoms in combined QM/MM calculations, *J. Phys. Chem. A* **102**, 4714-4721 (1998).
- [145] P. Amara, M. J. Field, C. Alhambra and J. Gao The generalized hybrid orbital method for combined quantum mechanical/molecular mechanical calculations: formulation and tests of the analytic derivatives, *Theoret. Chem. Acc.* **104**, 336-343 (2000).
- [146] G. Náray-Szabó Chemical fragmentation in quantum mechanical models, *Computers & Chemistry* **24**, 287-294 (2000).
- [147] D. M. Philipp and R. A. Friesner Mixed *ab initio* QM/MM modeling using frozen orbitals and tests with alanine dipeptide and tetrapeptide, *J. Comp. Chem.* **20**, 1468-1494 (1999).