QM/MM models for solvated and embedded systems (ground state energies, properties and reactivity)

Jeremy Harvey, KU Leuven, Belgium
Lecture 1: Introduction to QM/MM

Workshop on Theoretical Chemistry: Modelling of Environment Effects in Quantum Chemistry
Mariapfarr, 20-23 February 2018
Scientific Organizer: Benedetta Mennucci

Multiscale Phenomena in Chemistry

- Chemists need to think about the relation between microscopic properties and macroscale observables
- Usually this implies combining a model for a molecule- or atom-level property with a model for its impact at the macroscale
- E.g. water molecules form strong hydrogen bonds → boiling point of water much higher than that of H_2S
- There are many intermediate length scales: supramolecular constructs (e.g. proteins/micelles), cells, etc
- Also there are many *time* scales: molecular periodic motions (electrons, vibrations) → millions of years

Multiscale Models of Different Types

- The most accurate theoretical models in chemistry tend to be applicable only for small molecules (and short timescales)
- This is because electronic structure theory methods typically scale steeply in terms of computational effort: N^3 or worse
- Size of phase space (number of conformers, timescale for reaching equilibrium in simulations) grows very steeply with N
- Modelling is of necessity always in one sense multi-scale: quantum chemistry is used to predict molecular level properties, and a separate theory (e.g. statistical mechanics) is used to predict or rationalize macroscale observables
- In a fully-fledged multiscale model, a computer program is used to carry out *integrated* modelling of multiple scales

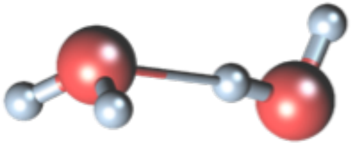
An Aside: Molecular Mechanics Force-Fields

In molecular mechanics, the potential energy of the system for a given arrangement of nuclei is obtained as a ‘mechanical’ expression based on the energy required to distort bond lengths & angles, distort dihedrals, and allow non-bonded atoms to interact via. Coulomb’s Law + Lennard-Jones terms:

$$V(R) = \sum_{i,j}^{non-bonded} \left\{ \frac{1}{4\pi\varepsilon_0} \times \frac{q_i q_j}{r_{ij}} \right\} + \sum_{i,j}^{non-bonded} 4\varepsilon_{ij} \left\{ \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right\} +$$

$$\sum_{i,j}^{bonds} \frac{1}{2} k_{ij} \left(r_{ij} - r_{ij}^{0} \right)^2 + \sum_{i}^{angles} \frac{1}{2} k_i \left(\theta_i - \theta_i^{0} \right)^2 + \sum_{i}^{dihedrals} A_i \{ 1 + \cos(n_i \phi_i) \}$$

θ_i is an angle, ϕ_i a dihedral angle, and k_{ij}, k_i, r_{ij}^0, θ_i^0, A_i and n_i are parameters.



E.g. a dimer of water molecules: the sum includes 4 O–H stretching terms and 2 H–O–H bending terms

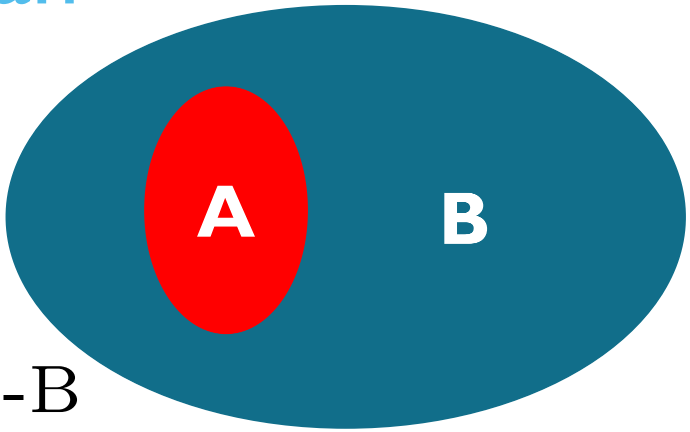
A Commonly Used Multiscale Model: QM/MM

The most fundamental description of molecular systems is based on nuclei and electrons (quantum mechanics, QM). In QM/MM, this level is used for a small "QM" region, with the bulk of the system described as atoms with bonds (molecular mechanics, MM)

MM is efficient → QM/MM applicable with thousands of atoms, only treating the detailed electronic structure of 10-200 atoms in the QM region.

There are also QM/QM methods, with an accurate QM method for the core, and a faster one for the environment.

Hybrid Methods: Hamiltonian



The overall energy includes
an energy term for both subsystems
and a coupling term:

$$V = V^{\text{A}}_{\text{high}} + V^{\text{B}}_{\text{low}} + V^{\text{A-B}}$$

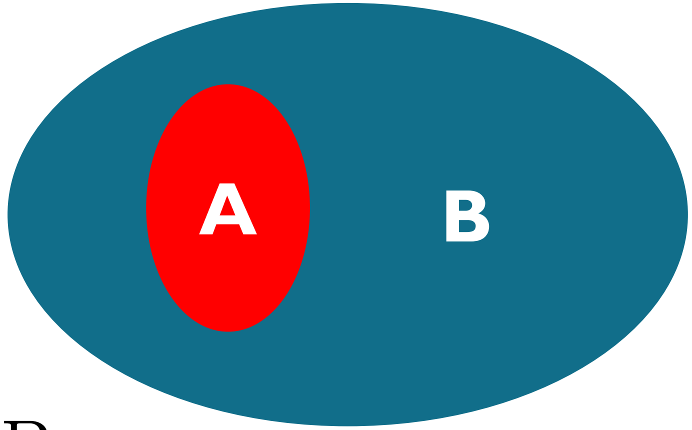
High-level energy of A

Low-level energy of B

This also yields (in principle) an electronic wave-function for the whole system:

$$\Psi = \hat{A} \left| \Psi^{\text{A}}_{\text{high}} \times \Psi^{\text{B}}_{\text{low}} \right\rangle$$

Hybrid Methods: Diversity

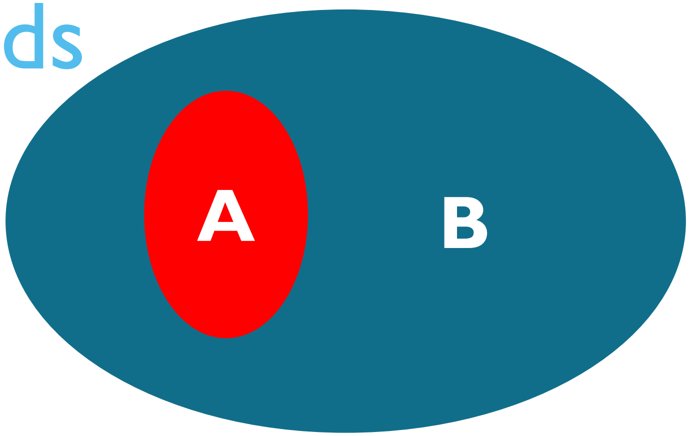


Within the above framework, there are very many variations on how to treat A and B, and the coupling between them.

$$V = V_{\text{high}}^{\text{A}} + V_{\text{low}}^{\text{B}} + V^{\text{A-B}}$$

One frequently use distinction is between *additive* methods and *subtractive* ones. Another distinction is between QM/QM and QM/MM schemes. For QM/MM methods, there can be *mechanical* embedding or *electronic* embedding. Finally the treatment of boundary region bonds can be quite diverse.

Additive/Substractive Methods



In additive methods, the energy of A is computed at one level of theory, that of B at another, and the coupling between them at a third level.

Much diversity: nature of the coupling term (and to some extent the treatment of the environment B) can vary a lot

Subtractive methods treat the *whole* system at the lower level of theory, and correct this at the high level for the active region:

$$V = V_{\text{low}}^{\text{AB}} + (V_{\text{high}}^{\text{A}} - V_{\text{low}}^{\text{A}})$$

This can also be viewed as a high level calculation on the model, to which is added a corrective environment term:

$$V = V_{\text{high}}^{\text{A}} + (V_{\text{low}}^{\text{AB}} - V_{\text{low}}^{\text{A}})$$

ONIOM (NB: original form)

One of the most commonly used subtractive methods is ONIOM: "Our own N-layered Integrated molecular Orbital and molecular Mechanics" method.

ONIOM: A Multilayered Integrated MO + MM Method for Geometry Optimizations and Single Point Energy Predictions. A Test for Diels-Alder Reactions and Pt(P(t-Bu)$_3$)$_2$ + H$_2$ Oxidative Addition

M. Svensson, S. Humbel, R. D. J. Froese, T. Matsubara, S. Sieber and K. Morokuma

J. Phys. Chem. 1996, **100**, 19357-19363.

ONIOM is potentially *n*-layered, hence the acronym: like an onion, there can be many successive layers

ONIOM: Gradients

The gradient (and higher derivatives) of the ONIOM energy can be straightforwardly computed from the gradients of the different methods

$$\frac{\partial V}{\partial X} = \frac{\partial V_{\mathrm{high}}^{\mathrm{A}}}{\partial X} + \left(\frac{\partial V_{\mathrm{low}}^{\mathrm{AB}}}{\partial X} + \frac{-\partial V_{\mathrm{low}}^{\mathrm{A}}}{\partial X}\right)$$

ONIOM: Application

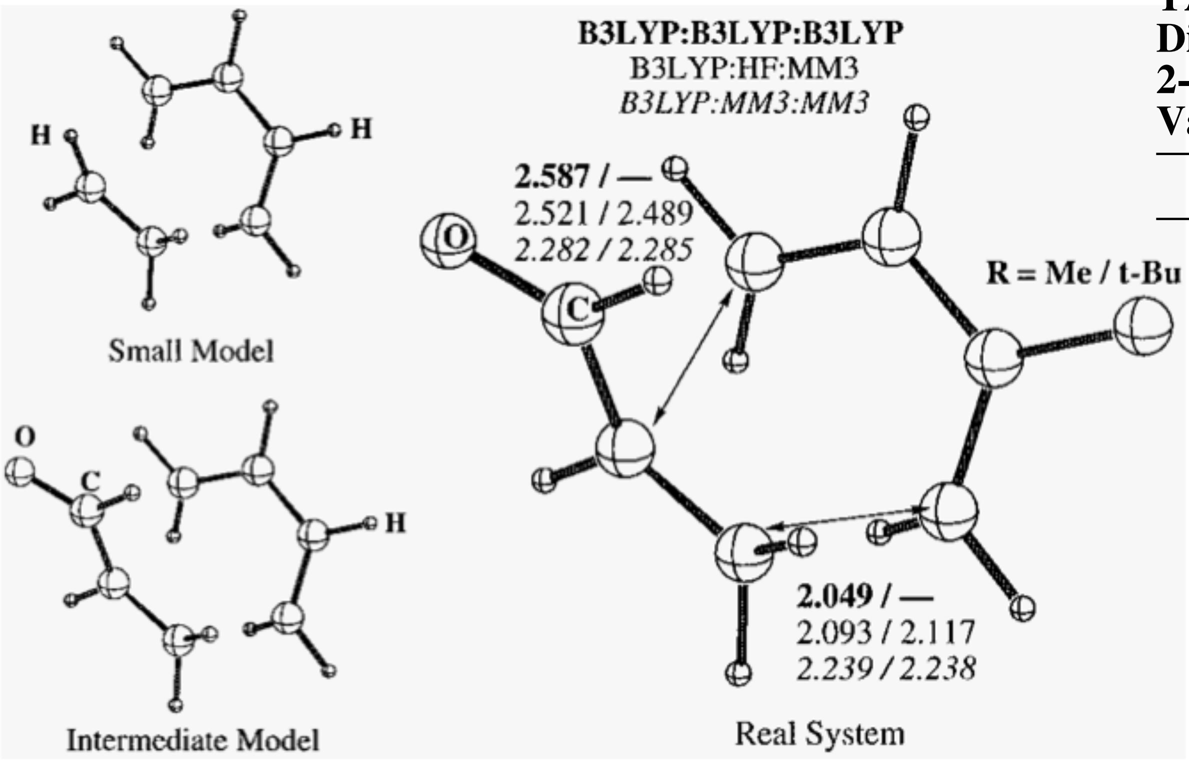


Figure 2. Small model, intermediate model, and real system for the Diels−Alder reactions between acrolein and 2-alkyl-1,3-butadiene. The optimized transition state bond distances (in Å) for the forming bonds are shown for the reaction of acrolein with isoprene and 2-*tert*-butyl-1,3-butadiene (to the right of /) at the nonintegrated (B3LYP:B3LYP:B3LYP) level in bold, those at the ONIOM3(B3LYP:HF:MM3) in normal, and those at the IMOMM(B3LYP:MM3:MM3) level in italic.

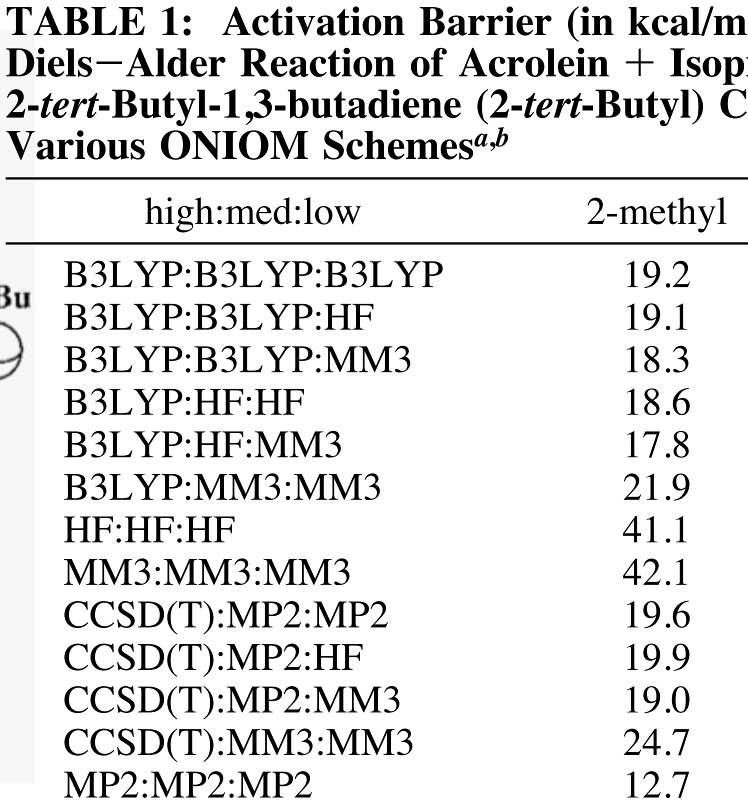


TABLE 1: Activation Barrier (in kcal/m
Diels−Alder Reaction of Acrolein + Isop
2-*tert*-Butyl-1,3-butadiene (2-*tert*-Butyl) C
Various ONIOM Schemes^{a,b}

high:med:low	2-methyl
B3LYP:B3LYP:B3LYP	19.2
B3LYP:B3LYP:HF	19.1
B3LYP:B3LYP:MM3	18.3
B3LYP:HF:HF	18.6
B3LYP:HF:MM3	17.8
B3LYP:MM3:MM3	21.9
HF:HF:HF	41.1
MM3:MM3:MM3	42.1
CCSD(T):MP2:MP2	19.6
CCSD(T):MP2:HF	19.9
CCSD(T):MP2:MM3	19.0
CCSD(T):MM3:MM3	24.7
MP2:MP2:MP2	12.7

J. Phys. Chem. 1996, **100**, 19357-19363.

ONIOM: Application 2

- ONIOM is typically aimed at organometallic cases:

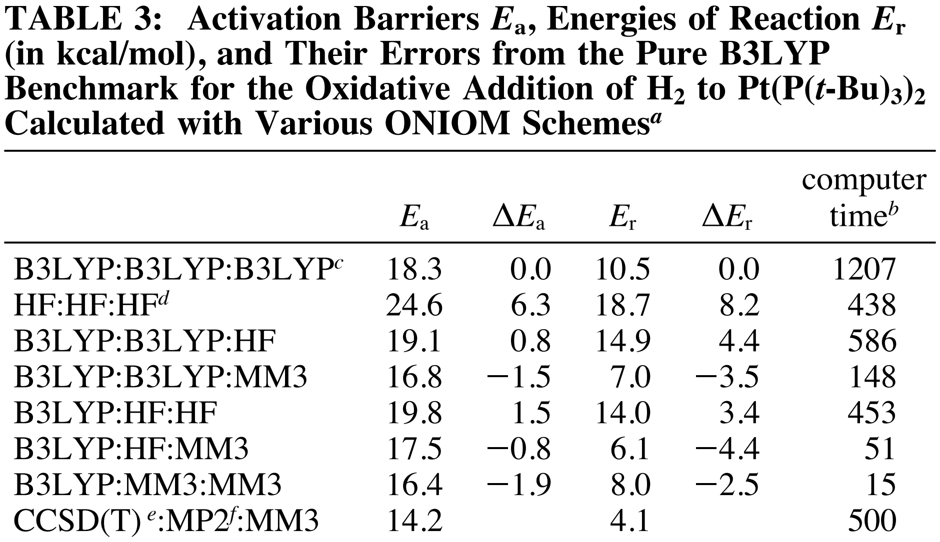
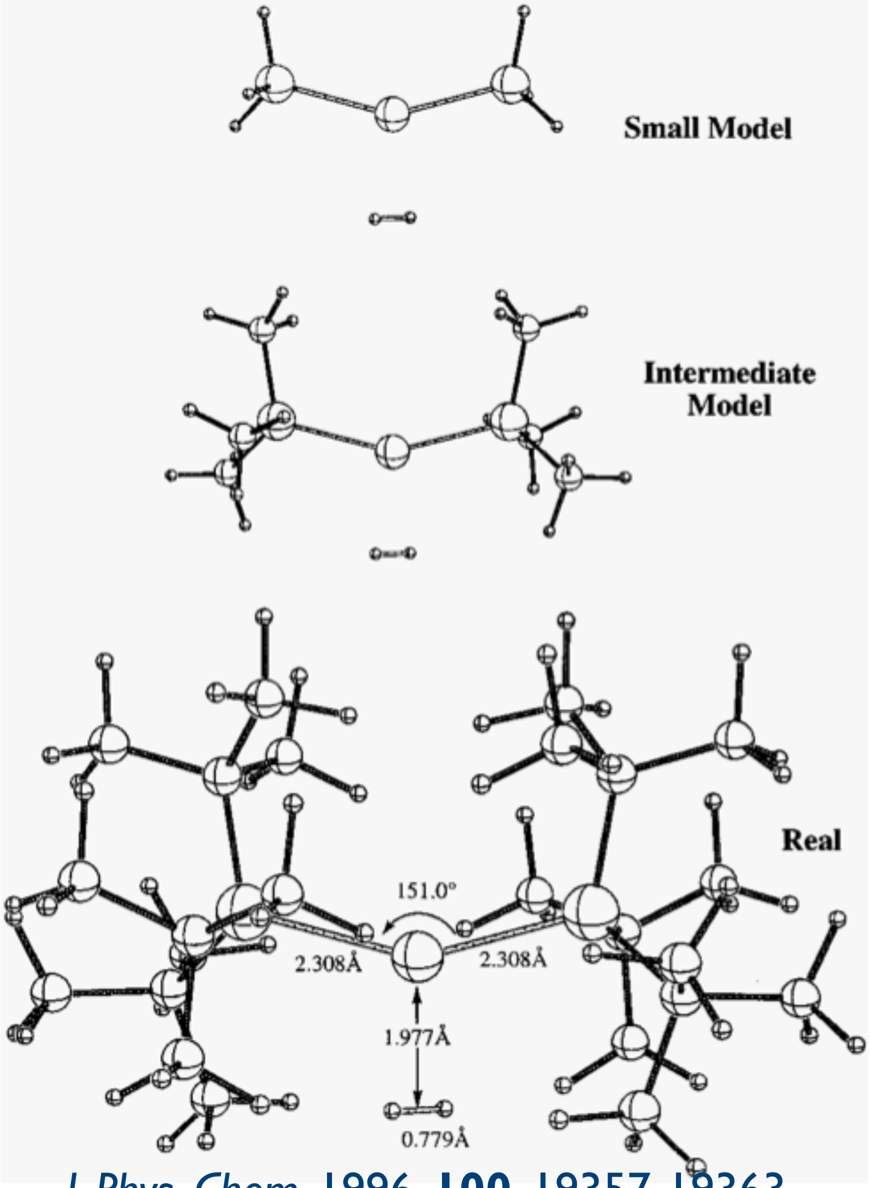


TABLE 3: Activation Barriers E_a, Energies of Reaction E_r (in kcal/mol), and Their Errors from the Pure B3LYP Benchmark for the Oxidative Addition of H_2 to Pt(P(*t*-Bu)$_3$)$_2$ Calculated with Various ONIOM Schemes[a]

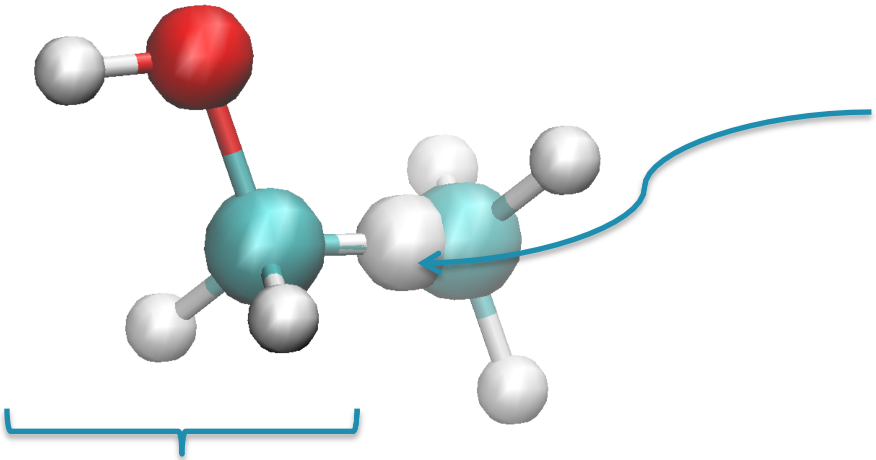
	E_a	ΔE_a	E_r	ΔE_r	computer time[b]
B3LYP:B3LYP:B3LYP[c]	18.3	0.0	10.5	0.0	1207
HF:HF:HF[d]	24.6	6.3	18.7	8.2	438
B3LYP:B3LYP:HF	19.1	0.8	14.9	4.4	586
B3LYP:B3LYP:MM3	16.8	−1.5	7.0	−3.5	148
B3LYP:HF:HF	19.8	1.5	14.0	3.4	453
B3LYP:HF:MM3	17.5	−0.8	6.1	−4.4	51
B3LYP:MM3:MM3	16.4	−1.9	8.0	−2.5	15
CCSD(T)[e]:MP2[f]:MM3	14.2		4.1		500



J. Phys. Chem. 1996, **100**, 19357-19363.

ONIOM: Link Atoms

- The energy for the sub-system A is not well defined when there is a covalent bond crossing the boundary. Instead, in such cases, A is augmented by an appropriate number of link atoms (typically H atoms).



Here an H atom is added
to convert CH_2OH into CH_3OH

High-level region

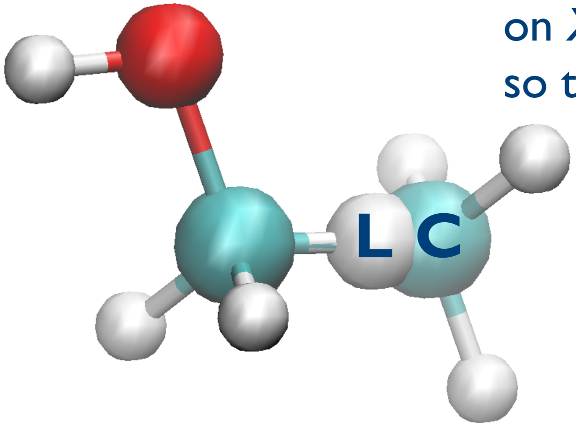
J. Phys. Chem. 1996, **100**, 19357-19363.

ONIOM: Link Atoms and Gradients

How do gradients work?

$$\frac{\partial V}{\partial X} = \frac{\partial V^{\mathrm{A}}_{\mathrm{high}}}{\partial X} + \left(\frac{\partial V^{\mathrm{AB}}_{\mathrm{low}}}{\partial X} + \frac{-\partial V^{\mathrm{A}}_{\mathrm{low}}}{\partial X}\right)$$

V_{high} implicitly depends on the coordinates X_{C} of atom 'replaced' by link atom L, as X_{L} itself depends on X_{C} (e.g. L is positioned along the same bond axis) so that: $X_{\mathrm{L}} - X_{\mathrm{Q}} = \alpha \times \{X_{\mathrm{C}} - X_{\mathrm{Q}}\}$



$$\Longrightarrow \frac{\partial V^{\mathrm{A}}_{\mathrm{high}}}{\partial X_{\mathrm{C}}} = \frac{\partial V^{\mathrm{A}}_{\mathrm{high}}}{\partial X_{\mathrm{L}}} \times \frac{\partial X_{\mathrm{L}}}{\partial X_{\mathrm{C}}}$$

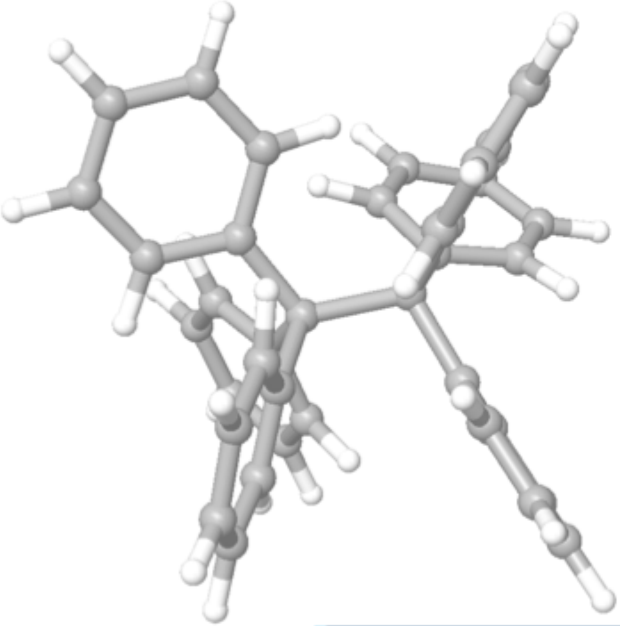
Feliu Maseras and Keiji Morokuma, *J. Comp. Chem.* 1995, **16**, 1170-1179

ONIOM: Mechanical Embedding

When the low level in ONIOM is a forcefield method, the only impact the environment has on the subsystem is to distort it according to the coupling terms implicit in the correction term:

$$V = V^{\mathrm{A}}_{\mathrm{high}} + (V^{\mathrm{AB}}_{\mathrm{low}} - V^{\mathrm{A}}_{\mathrm{low}})$$

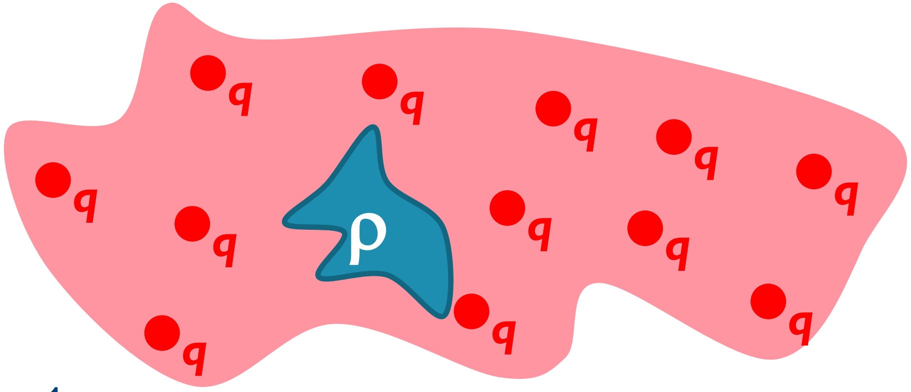
Especially where the forcefield contains no point charges (ball and spring model typical for hydrocarbons), this is a purely "mechanical" effect.



QM/MM and Electronic Embedding

For some systems, electrostatic interactions between the QM and MM regions can be quite strong – and the MM environment may be expected to polarize the QM region somewhat. For these systems, it is desirable to allow the QM wavefunction to relax with respect to the MM point charges. This leads to typical QM/MM electronic embedding methods.

$$V = V_{\mathrm{QM}}^{\mathrm{A}} + V_{\mathrm{MM}}^{\mathrm{B}} + V_{\mathrm{QM\text{-}MM}}^{\mathrm{A\text{-}B}}$$



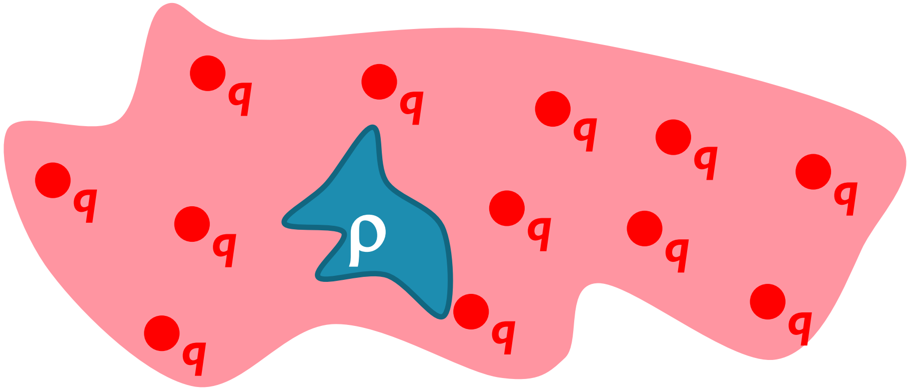
Electronic Embedding in Detail

In practice the energy expression is rather: $V = V_{\mathrm{QM'}}^{\mathrm{AB}} + V_{\mathrm{MM'}}^{\mathrm{AB}}$

$V_{\mathrm{QM'}}^{\mathrm{AB}}$ is the QM energy of A in the presence of the point charges

$$\hat{H}_{\mathrm{elec}} = \frac{-1}{2}\sum_{i=1}^{n}\nabla_i^2 + \sum_{i=1}^{n}\sum_{A=1}^{N}\frac{-Z_A}{r_{iA}} + \sum_{i=1}^{n}\sum_{j>i}^{n}\frac{1}{r_{ij}} + \sum_{i=1}^{n}\sum_{K=1}^{N_{MM}}\frac{-q_K}{r_{iK}}$$

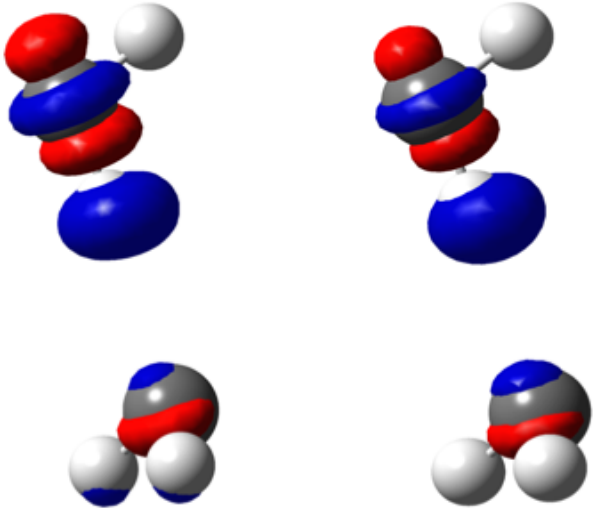
(n electrons, N nuclei,
N_{MM} MM point charges)



$V_{\mathrm{MM'}}^{\mathrm{AB}}$ is more straightforward:
it contains the MM energy
of MM atoms + MM-like
interactions between MM and QM atoms

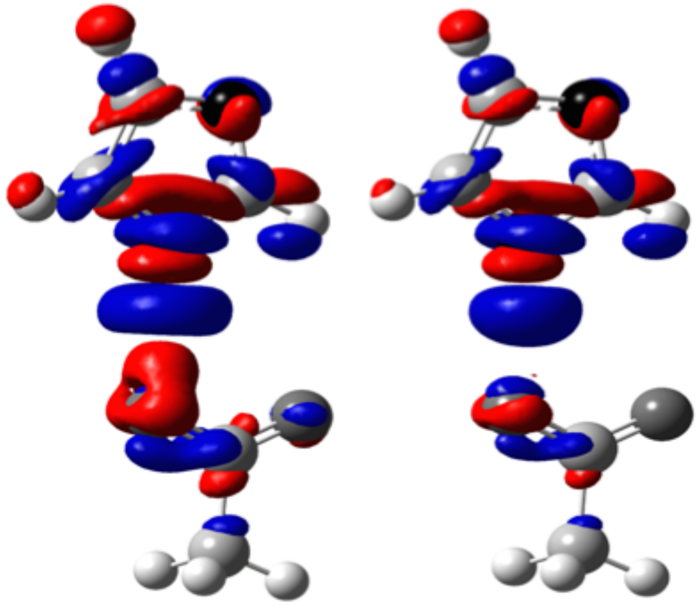
Electronic Polarization

Point charges indeed lead to similar polarization compared to all-QM calculations:



(a) (b)

4. Water dimer



(a) (b)

5. Imidazole – Acetate

Analysis of polarization in QM/MM modelling of biologically relevant hydrogen bonds, Senthilkumar, Mujika, Ranaghan, Manby, Mulholland and Harvey, *J. Roy. Soc. Interface*, 2008, **5**, S207 – S216

Charge Leakage

Regions of high electron density in the QM region could in principle 'leak out' onto positive point charges due to lack of Pauli repulsion by these charges. Not a big problem for modest-sized Gaussian basis sets, but significant with plane waves.

Solution: replace 1/*r* term with a shielded interaction.

$$\sum_{i=1}^{n} \sum_{K=1}^{N_{MM}} \frac{-q_K}{r_{iK}} \mapsto \sum_{i=1}^{n} \sum_{K=1}^{N_{MM}} \left(-q_K \times f(r_{iK})\right)$$

See e.g. *A regularized and renormalized electrostatic coupling Hamiltonian for hybrid quantum-mechanical–molecular-mechanical calculations*, P. K. Biswas and V. Gogonea, *J. Chem. Phys.* 2005, **123**, 164114.

Covalent bonds and the QM-MM boundary

In QM/MM calculations, covalent bonds between the QM and MM boundary can be treated in several different ways:

- Link atoms, as in mechanical embedding
  Note that the electrostatic treatment of MM atoms near the MM boundary can be varied, and this can have a big effect on results
- Pseudo-atoms: instead of an H atom to replace the bonded MM atom, use some kind of 1-electron atom with a pseudopotential designed to yield the correct bond length and electronic structure. (See discussion in Xiao and Zhang, *J. Chem. Phys.* 2007, 127, 124102)
- Frozen orbitals: constrain one hybrid atomic orbital of the QM atom to have exactly the same shape it has in a reference compound, with two electrons – behaves a bit like a lone pair
  (see e.g. Murphy, Philipp and Friesner, *J. Comp. Chem.* 2000, **21**, 1442.)

QM/MM Gradients

There are four types of gradient elements in QM/MM with electronic embedding:

$$\frac{\partial V_{\mathrm{QM}}}{\partial X_{\mathrm{QM}}}$$

Like standard QM gradient except for modified w.f.

$$\frac{\partial V_{\mathrm{QM}}}{\partial X_{\mathrm{MM}}}$$

Not routinely computed by QM codes; but for SCF methods, just equal to $-q_K \times E_{\mathrm{QM}}$

$$\frac{\partial V_{\mathrm{MM}}}{\partial X_{\mathrm{QM}}}$$

Standard MM-like terms, for interactions (e.g. vdW) with the QM atoms

$$\frac{\partial V_{\mathrm{MM}}}{\partial X_{\mathrm{MM}}}$$

Standard MM

Exploring QM/MM Potential Energy Surfaces

- Large systems place high demands on the Hamiltonian
- QM/MM as well as (near-)linear scaling methods can now routinely cope with (very) large systems
- Nowadays, a *larger* challenge of big systems is the complexity of their potential energy surfaces
- Can be explored using geometry optimization and reaction path methods
- Or by simulation with QM/MM
- Free energies from Umbrella Sampling/Metadynamics and other related approaches
- Combining MM and QM/MM or $\mathrm{QM^{low}}$/MM and $\mathrm{QM^{high}}$/MM

QM/MM Geometry Optimization: Microiterations

Time-consuming for large systems (*N* atoms). Even efficient algorithms tend to take ~*N* steps until acceptable convergence.

Even if each energy and gradient evaluation for the whole QM/MM system is cheap (roughly the cost of the QM energy + gradient for the QM region), this makes optimization expensive.

Hence *microiterations*: carry out many cycles of optimization of MM atom positions for each calculation of QM energy/gradient. Possible because QM and MM degrees of freedom are almost uncoupled:

$$\frac{\partial V_{\mathrm{MM}}}{\partial X_{\mathrm{MM}}}$$

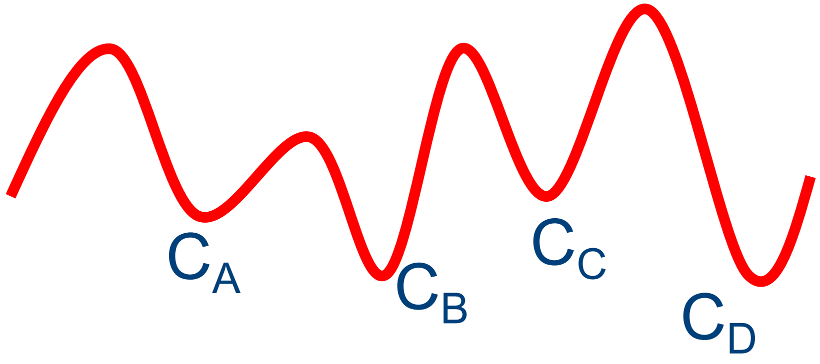
Does not depend on the QM system

$$\frac{\partial V_{\mathrm{QM}}}{\partial X_{\mathrm{MM}}}$$

Quite small; depends only weakly on the QM wavefunction – can be approximated by electrostatics.

See e.g. *A Mixed Quantum Mechanics/Molecular Mechanics (QM/MM) Method for Large-Scale Modeling of Chemistry in Protein Environments*, Murphy, Philipp and Friesner, *J. Comp. Chem.* 2000, **21**, 1442.

Energy Minimization Reaches Local Minima



- Starting from a given structure, one reaches a *local* minimum (typically the 'closest' one)
- A system with *N* atoms typically has a potential energy surface with $\sim\exp(\alpha N)$ local minima – a large number!!
- There are also large numbers of saddle points
- One typically *cannot* find the global minimum, or the lowest energy saddlepoint – unlike for systems with ~10s of atoms
- This is *not* a difficulty limited to QM/MM – it is rather a property of large systems (MM, semiempirical, large QM)

Characterizing Minima

- For molecular systems, optimization is typically followed by calculating frequencies.
- These require the Hessian, i.e. the terms:

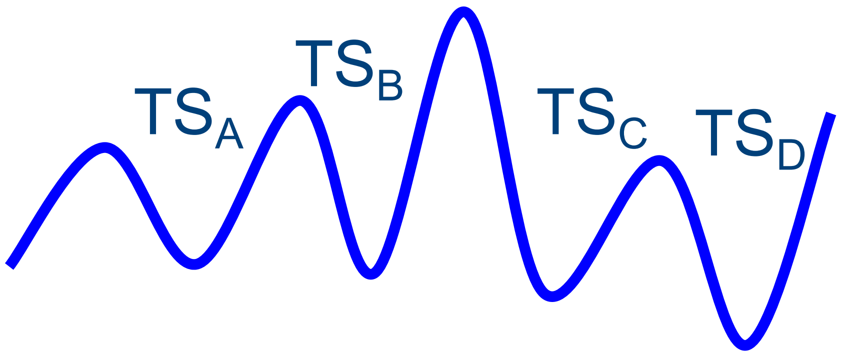
$$\frac{\partial V_{\mathrm{QM}}^{2}}{\partial X_i \partial X_j} \qquad \frac{\partial V_{\mathrm{QM}}^{2}}{\partial x_i \partial x_j} \qquad \frac{\partial V_{\mathrm{QM}}^{2}}{\partial x_i \partial X_j}$$

$$\frac{\partial V_{\mathrm{MM}}^{2}}{\partial X_i \partial X_j} \qquad \frac{\partial V_{\mathrm{MM}}^{2}}{\partial x_i \partial x_j} \qquad \frac{\partial V_{\mathrm{MM}}^{2}}{\partial x_i \partial X_j}$$

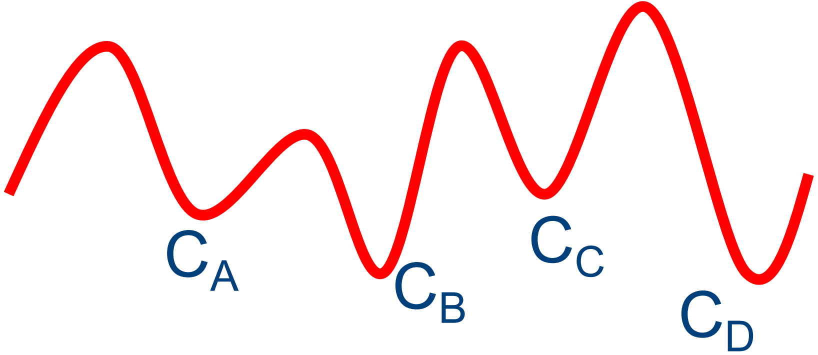
Full and partial Hessian

- The full Hessian may be unavailable, or too large to store or diagonalize (e.g. 10,000 atoms => 30,000 x 30,000)
- Also, typically in QM/MM, some atoms may be frozen during optimization => typically one does not have a minimum
- Gradient norm not arbitrarily close to zero, so ‘frequencies’ for rotational modes may not be that close to zero.
- It can be useful to diagonalize a block sub-Hessian (see Efficient Calculation of QM/MM Frequencies with the Mobile Block Hessian, Ghysels *et al.*, *JCTC* 2011, **7**, 496-514.)
- Provided that eigenvectors for vibrational modes of interest (e.g. the reaction coordinate) are well within the block of atoms for which the Hessian is generated, then should be OK.

Reaction path study

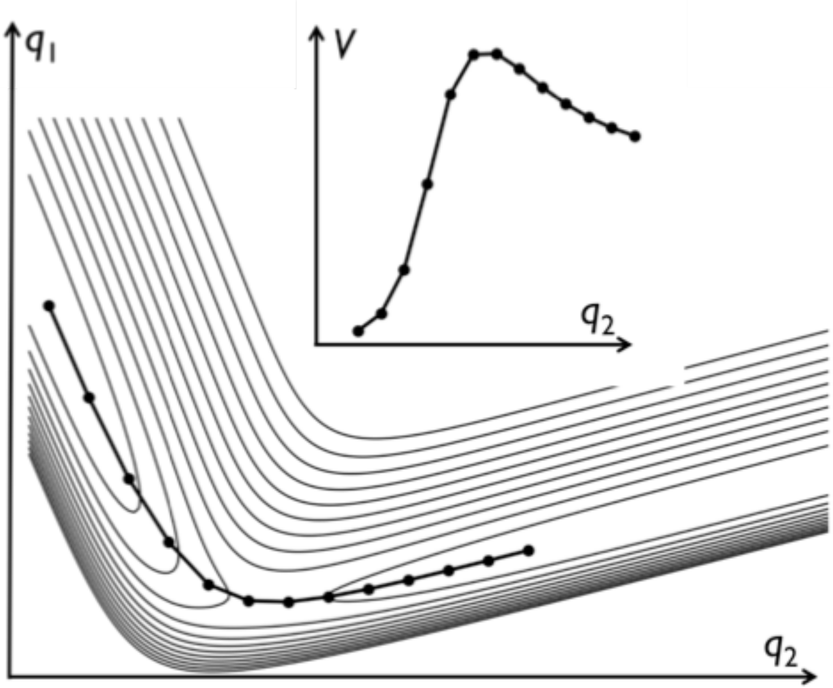


- In molecular systems, a typical approach is to locate relevant minima for reactants, intermediates, products, then TSs separating them
- This can be done also with QM/MM: algorithms for TS searching can be used more or less as such for QM/MM systems (using e.g. a partial block Hessian to guide the search)
- Two difficulties: many coordinates, and hard to guarantee that the located minimum and TS are connected



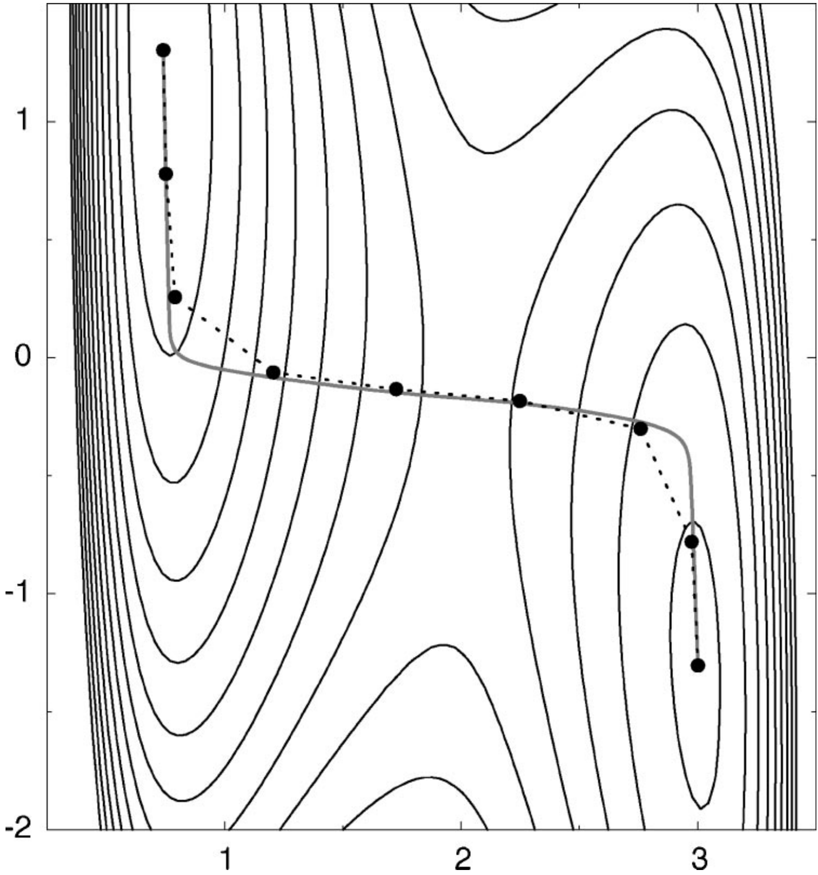
- The variance in energies of local minima / TSs typically exceeds the target relative energy.

Reaction paths: Adiabatic mapping



- In order to obtain meaningful relative energies along e.g. a reaction path, one can use techniques such as adiabatic mapping
- Key coordinate (here 'q_2') chosen
- Minima found for set of values of this coordinate $q_{2,i}$ (enforced e.g. by adding harmonic term $k(q_2-q_{2,i})^2$.
- This method works well when the reaction path remains highly parallel to q_2 along the whole reaction – but in practice *reaction path curvature* is a major issue

Nudged Elastic Bands



Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points, Henkelmann and Jónsson, *J. Chem. Phys.* 2000, **113**, 9978-9985

$$X_0, X_1, X_2, \ldots, X_{n-1}, X_n$$

For each image, one follows a modified gradient, the sum of a harmonic spring force in the direction tangent to reaction path, and the true gradient orthogonal to the tangent. This tangent direction is given by:

$$\tau_i = \frac{X_i - X_{i-1}}{|X_i - X_{i-1}|} + \frac{X_{i+1} - X_i}{|X_{i+1} - X_i|}$$

Lecture 1 Conclusions

- The QM/MM method provides an attractive and efficient method to compute energies and wavefunctions for large systems
- Many tests suggest that QM/MM is reasonably accurate for describing many physical effects at the QM:MM boundary
- The *gradient* (and higher derivatives) of the energy are well defined, so geometry optimization is possible
- Not because we use a QM/MM hamiltonian, but simply due to the large size of the typical systems studied, geometry optimization has some new difficulties compared to smaller ('QM') systems
- Global vs. Local minima
- Connectivity of reaction paths: need adiabatic mapping or NEB