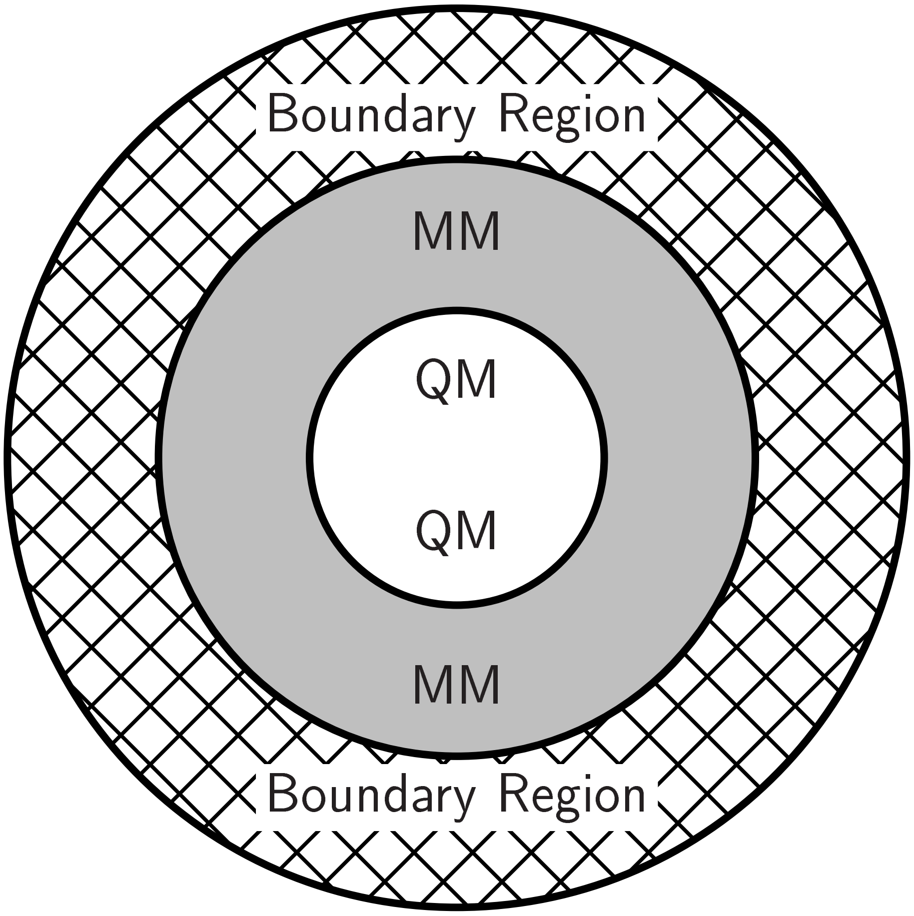
Theory of hybrid Quantum Mechanical Molecular Mechanical methods

An introduction to a way of treating complex chemical systems



$$H_{\text{main}} = H_{\text{QM}} + H_{\text{MM}} + H_{\text{QMMM}}$$

Thomas Lorenzen
Theoretical Chemistry
Department of Chemistry
University of Copenhagen

Outline of the talk

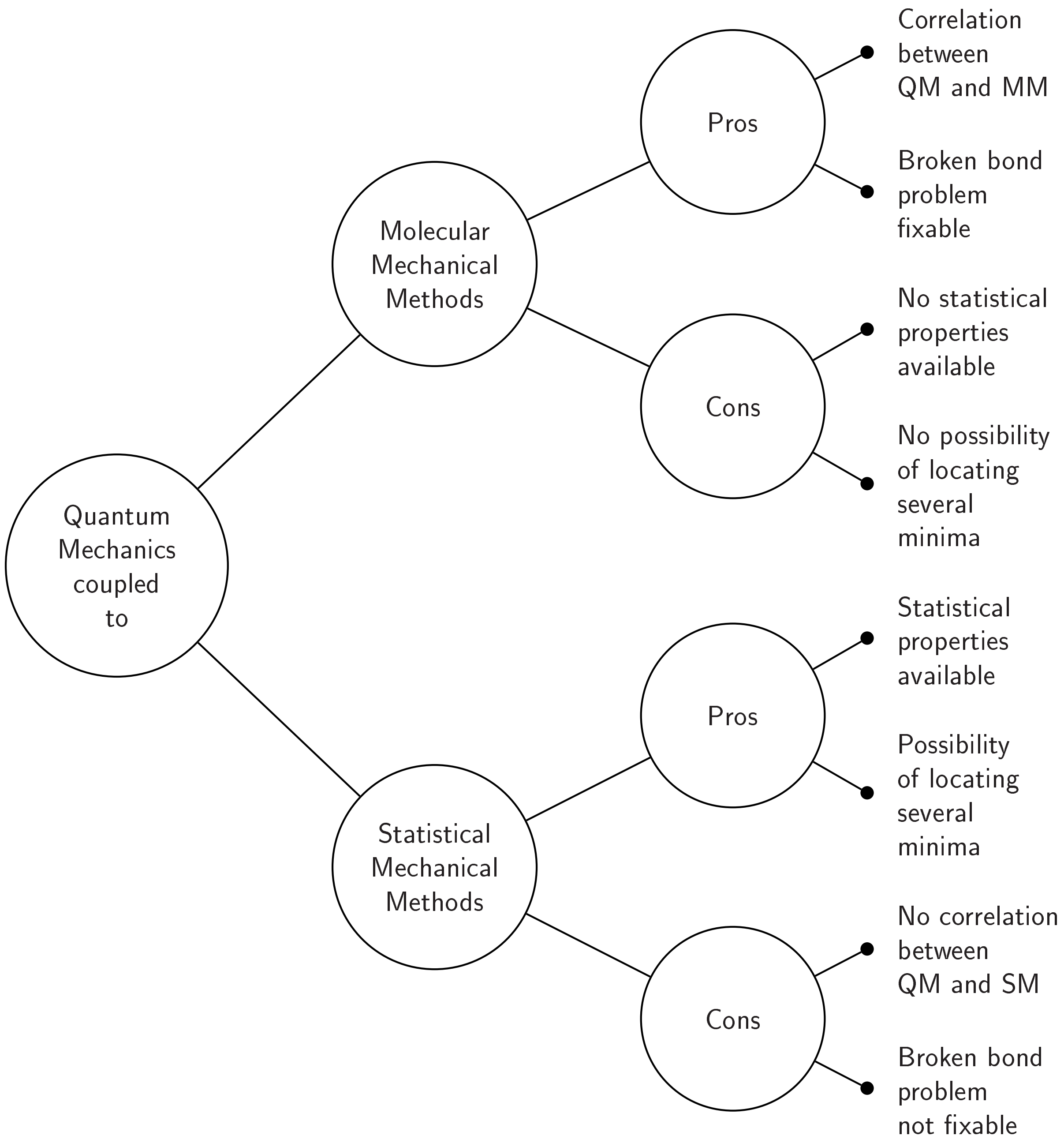
- Why couple QM and MM
- Advantages and disadvantages of QM/MM in comparison to QM/SM
- Aspects of the hybrid QM MM method
- Criteria to fulfill when partitioning a molecular system into a QM part and an MM part
- How to treat broken bonds, if such are present after the partitioning
- Basics
- Establishing the main hamiltonian
- Example
- Acknowledgements
- Bibliography

Why couple QM and MM

- For big molecules QM is to expensive in itself.
- No harm is done in using lower level theory far away from the chemically interesting place.
- If used with care, MM has proven its worth.
- Use QM coupling to MM as an alternative to QM coupling to SM.

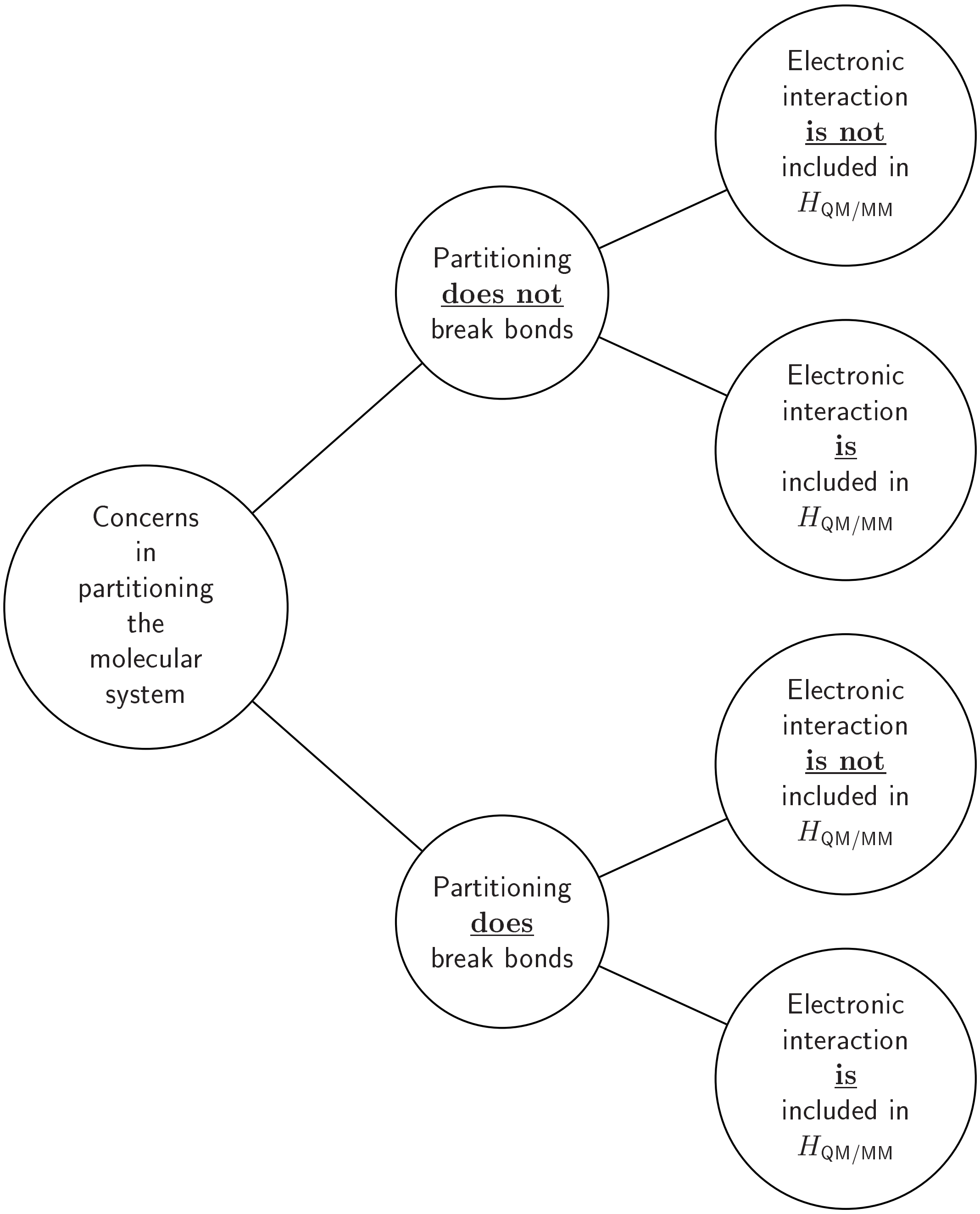
Advantages and disadvantages of QM/MM in comparison to QM/SM

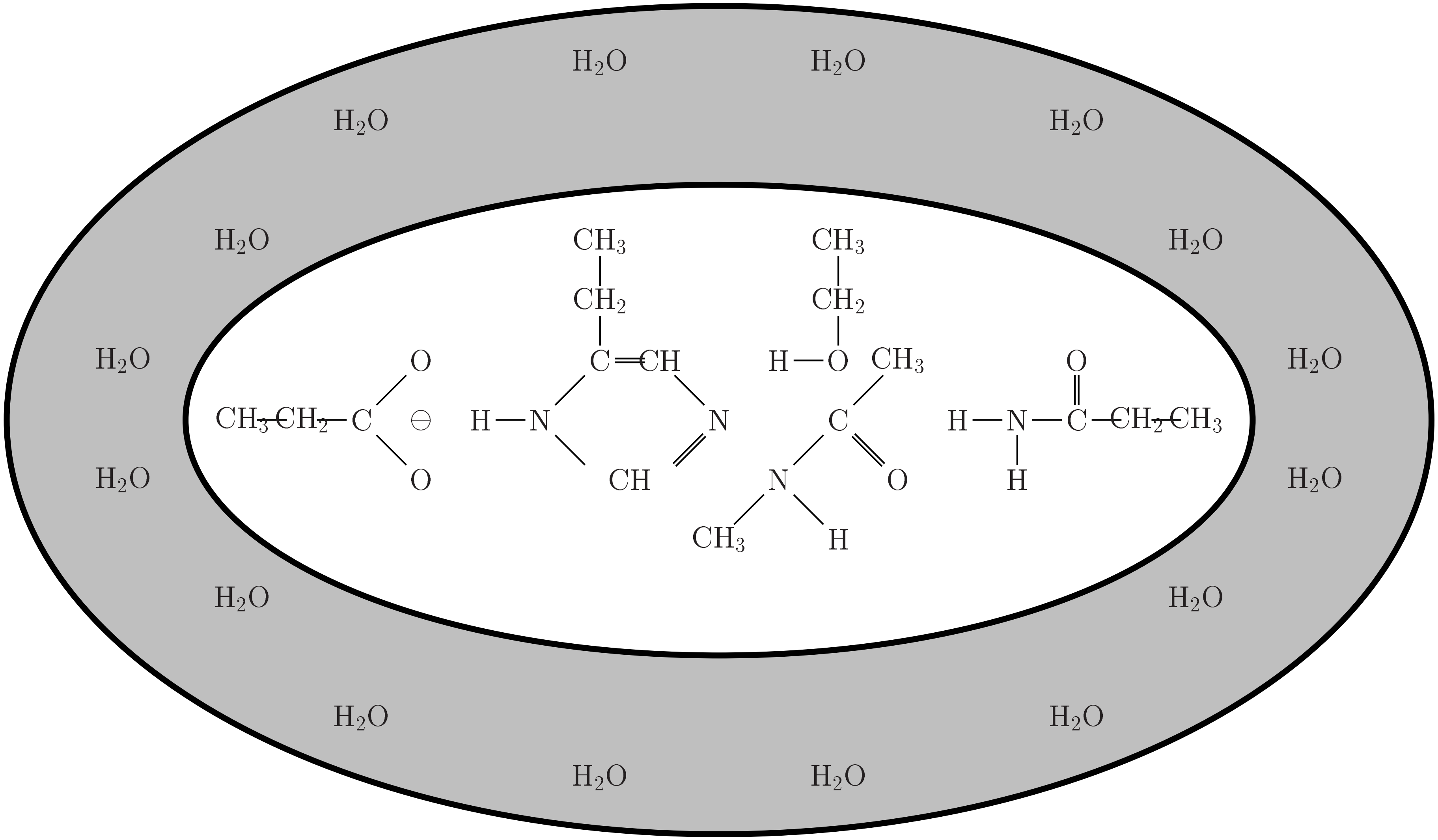
We list some *basic* pros and cons in comparing QM/MM and QM/SM, but stress, that both methods may be refined to annihilate the limitations.

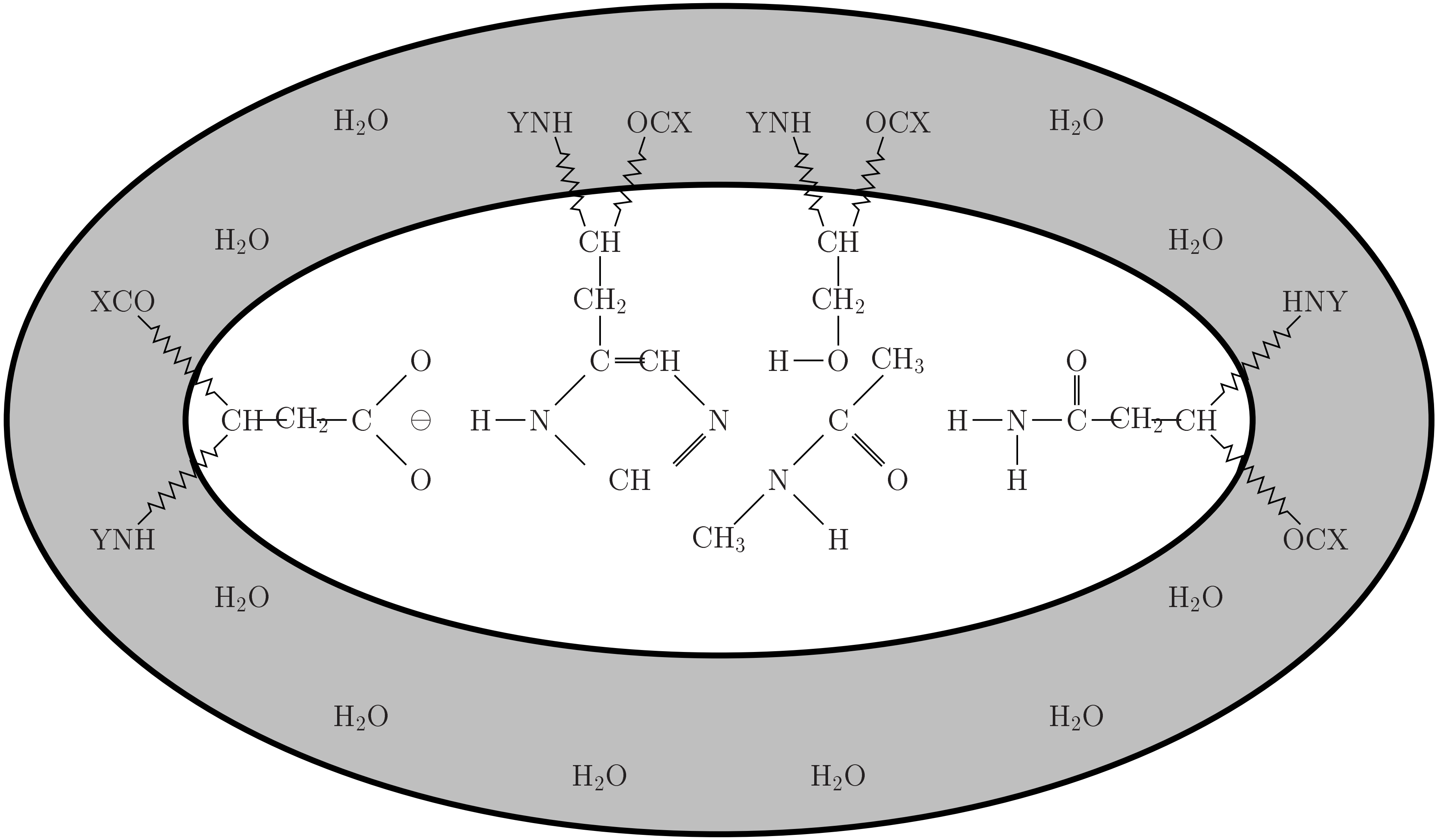


Aspects of the hybrid QM MM method

We note the most *essential* aspects in the combined QM MM method, but others of course exist.






H2O
YNH
OCX
YNH
OCX
H2O
H2O
H2O
XCO
HNY
CH
CH
CH2
CH2
O
C
CH
H
O
CH3
CH
CH2
C
⊖
H
N
N
C
H
N
C
CH2
CH
O
O
CH
N
O
H
YNH
OCX
CH3
H
H2O
H2O
H2O
H2O
H2O
H2O

Criteria to fulfill when partitioning a molecular system into a QM part and an MM part

- Fundamental knowledge of the system (Structure info in the case of enzymes).
- Largest possible size of QM part as opposed to level of accuracy necessary to treat QM part satisfactory.
- In case of calculating molecular properties, how will a broken bond between the QM region and the MM region affect the result.
- In case of modelling chemical reactions, how will the reaction affect the electronic distribution at the boundary (The goal is, that there should be no change in the electronic distribution).

How to treat broken bonds, if such are present after the partitioning

- For each broken bond, a link atom should be added along the broken bond in order to be able to perform a QM calculation. The link atom is often chosen as a hydrogen atom and should not be allowed to interact with any MM atom.
- In addition to *nonbonded interaction terms* there are now also *bonded interaction terms* present.
- *Bonded interaction terms* : For each *bond* or *angle* or *dihedral* crossing the boundary, a force field *stretch term* or *bend term* or *torsional term* is added.
- *Nonbonded interaction terms* : Comparing [1] and [2] there seems to be some difference in approach.

  [1] **:** Nonbonded interaction terms (van der walls interaction plus electrostatic interaction) between a QM atom and an MM atom are added <u>if</u> these are separated by more than three bonds!!!

  [2] **:** Nonbonded interaction terms (van der walls interaction plus electrostatic interaction) between a QM atom and an MM atom are added <u>whether or not</u> these are separated by more than three bonds!!!

Basics

Choosing how to interact between QM and MM

- Comparing [1] and [2] there seems to be some difference in approach.

  [1] **:** Method seems to be mainly based on treating
  MM-partial-charge QM-partial-charge interactions!!!
  (Consequence : No terms from H_{QMMM} will enter the HF SCF equations).

  [2] **:** Method seems to be mainly based on treating
  MM-partial-charge QM-electron-nuclei interactions!!!
  (Consequence : Some terms from H_{QMMM} will enter the HF SCF equations).

Notation

- Indices i, j corresponds to QM electrons.
- Indices k, l corresponds to QM nuclei.
- Indices s, t corresponds to QM atoms.
- Indices m, n corresponds to MM atoms.
- Indices α, β corresponds to basis functions.

Establishing the main hamiltonian

Partitioning of the hamiltonian

$$H_{\mathsf{main}}\Psi = E\Psi$$

$$H_{\mathsf{main}} = H_{\mathsf{QM}} + H_{\mathsf{MM}} + H_{\mathsf{QMMM}}$$

$$E = \frac{\langle\Psi|H_{\mathsf{main}}|\Psi\rangle}{\langle\Psi|\Psi\rangle}$$

$$E = E_{\mathsf{QM}} + E_{\mathsf{MM}} + E_{\mathsf{QMMM}}$$

Treating the separated parts

$$\begin{aligned} H_{\mathsf{QM}} &= -\frac{1}{2}\sum_i \nabla_i^2 - \sum_{i,k}\frac{Z_k}{r_{ik}} + \sum_{i,j}\frac{1}{r_{ij}} \\ &\quad + \sum_{k,l}\frac{Z_k Z_l}{r_{kl}} \end{aligned}$$

$$\begin{aligned} H_{\mathsf{MM}} &= \sum_{\mathsf{bond}}^{\mathsf{MM\,MM}} \frac{k_{\mathsf{b}}}{2}(r-r_0)^2 + \sum_{\mathsf{angle}}^{\mathsf{MM\,MM}} \frac{k_{\mathsf{a}}}{2}(\theta-\theta_0)^2 + \sum_{\mathsf{dihedral}}^{\mathsf{MM\,MM}} \frac{k_{\mathsf{d}}}{2}(1+\cos\Theta) \\ &\quad + \sum_{m,n}\left(\frac{A_{mn}}{r_{mn}^{12}} - \frac{B_{mn}}{r_{mn}^{6}}\right) + \sum_{\mathsf{Hydrogen\ bond}}\left(\frac{C_{mn}}{r_{mn}^{12}} - \frac{D_{mn}}{r_{mn}^{10}}\right) \\ &\quad + \sum_{m,n}\frac{q_m q_n}{r_{mn}} \\ &\quad + \frac{1}{2}\sum_{m<n} q_m \frac{\boldsymbol{\mu}_n\cdot\boldsymbol{r}_{mn}}{r_{mn}^3} + \frac{1}{2}\sum_{m>n} q_m \frac{\boldsymbol{\mu}_n\cdot\boldsymbol{r}_{mn}}{r_{mn}^3} - \frac{1}{2}\sum_m \alpha_m \boldsymbol{E}_m\cdot\boldsymbol{E}_m \end{aligned}$$

Establishing the main hamiltonian

H_{QMMM} - Approach I - Based on [1] and [3]

One might at first try to not explicitly include electronic interactions in H_{QMMM} thus getting

$$\begin{aligned}
H_{\mathsf{QMMM}} =& \sum_{\mathsf{bond}}^{\mathsf{QM\,MM}} \frac{k_{\mathsf{b}}}{2}(r-r_0)^2 + \sum_{\mathsf{angle}}^{\mathsf{QM\,MM}} \frac{k_{\mathsf{a}}}{2}(\theta-\theta_0)^2 + \sum_{\mathsf{dihedral}}^{\mathsf{QM\,MM}} \frac{k_{\mathsf{d}}}{2}(1+\cos\Theta) \\
&+ \sum_{s,m}\left(\frac{A_{sm}}{r_{sm}^{12}} - \frac{B_{sm}}{r_{sm}^{6}}\right) + \sum_{\mathsf{Hydrogen\ bond}}\left(\frac{C_{sm}}{r_{sm}^{12}} - \frac{D_{sm}}{r_{sm}^{10}}\right) \\
&\boxed{+\sum_{s,m}\frac{q_s q_m}{r_{sm}}} \\
&\boxed{+\frac{1}{2}\sum_{s,m} q_s \frac{\boldsymbol{\mu}_m \cdot \boldsymbol{r}_{sm}}{r_{sm}^3}}
\end{aligned}$$

This approach will make H_{QMMM} <u>independent</u> on electronic coordinates, and thus H_{QMMM} will just be a number to be added to the total energy.

One should in optimizations of the QM part add gradients from the MM part, as optimization of the QM part otherwise will proceed without correlation to the MM part.

Establishing the main hamiltonian

H_{QMMM} - Approach II - Based on [2] and [3]

One might then explicitly include electronic interactions in H_{QMMM} thus getting

$$
\begin{aligned}
H_{\mathsf{QMMM}} = & \sum_{\mathsf{bond}}^{\mathsf{QM\,MM}} \frac{k_{\mathsf{b}}}{2}(r-r_0)^2 + \sum_{\mathsf{angle}}^{\mathsf{QM\,MM}} \frac{k_{\mathsf{a}}}{2}(\theta-\theta_0)^2 + \sum_{\mathsf{dihedral}}^{\mathsf{QM\,MM}} \frac{k_{\mathsf{d}}}{2}(1+\cos\Theta) \\
& + \sum_{s,m}\left(\frac{A_{sm}}{r_{sm}^{12}} - \frac{B_{sm}}{r_{sm}^{6}}\right) + \sum_{\text{Hydrogen bond}}\left(\frac{C_{sm}}{r_{sm}^{12}} - \frac{D_{sm}}{r_{sm}^{10}}\right) \\
& \boxed{\boxed{-\sum_{i,m}\frac{q_m}{r_{im}}} + \sum_{k,m}\frac{Z_k q_m}{r_{km}}} \\
& \boxed{\boxed{-\frac{1}{2}\sum_{i,m}\frac{\boldsymbol{\mu}_m\cdot\boldsymbol{r}_{im}}{r_{im}^3}} + \frac{1}{2}\sum_{k,m} Z_k \frac{\boldsymbol{\mu}_m\cdot\boldsymbol{r}_{km}}{r_{km}^3}}
\end{aligned}
$$

This approach will make H_{QMMM} <u>dependent</u> on electronic coordinates, and thus H_{QMMM} will not just be a number to be added to the total energy.

The terms in H_{QMMM} involving electronic coordinates must be added to the HF SCF procedure, and in this case these terms in H_{QMMM} will make the QM part feel the MM part also in the HF SCF procedure.

Establishing the main hamiltonian

Look of the extra terms entering the Hartree Fock equations

Due to the double boxed expressions involving electronic coordinates the HF SCF procedure must be modified to include the resulting one electron contributions. The one electron integrals to be included in the HF SCF procedure are the following.

$$H_{\alpha\beta}^{\mathsf{QMMM}} = -\sum_m \left(\int \phi_\alpha \frac{q_m}{r_{im}} \phi_\beta \, dr + \frac{1}{2} \int \phi_\alpha \frac{\boldsymbol{\mu}_m \cdot \boldsymbol{r}_{im}}{r_{im}^3} \phi_\beta \, dr \right)$$

As these extra terms are all one electron terms we may readily add them to the pure quantum mechanical Fock operator $F_{\alpha\beta}^{\mathsf{QM}}$ thus producing another Fock operator $F_{\alpha\beta}$ looking like the following.

$$F_{\alpha\beta} = F_{\alpha\beta}^{\mathsf{QM}} + H_{\alpha\beta}^{\mathsf{QMMM}}$$

Comment on inclusion of polarizability terms

It should be noted, that inclusion of polarizability terms in the HF SCF scheme induces a significant increase in computation time.
This is because the total restricted Hartree Fock wave function Ψ and the induced dipoles $\boldsymbol{\mu}$ are coupled and thus should be determined self consistently within the HF SCF scheme.

Establishing the main hamiltonian

Approach I - The total energy

Assuming, that we are working in a closed shell system and thus applying a restricted Hartree Fock approach we end up with the following energy expression. Note, that $E_{\mathsf{MM}} = H_{\mathsf{MM}}$ as well as $\underline{E_{\mathsf{QMMM}} \equiv H_{\mathsf{QMMM}}}$.

$$E = E_{\mathsf{QM}} + E_{\mathsf{MM}} + E_{\mathsf{QMMM}}$$

$$E_{\mathsf{QM}} = \frac{1}{2}\sum_{\alpha\beta} P_{\alpha\beta}(H^{\mathsf{QM}}_{\alpha\beta} + F^{\mathsf{QM}}_{\alpha\beta}) + \sum_{k,l}\frac{Z_k Z_l}{r_{kl}}$$

$$\begin{aligned}
E_{\mathsf{MM}} = & \sum_{\mathsf{bond}}^{\mathsf{MM\,MM}} \frac{k_{\mathsf{b}}}{2}(r-r_0)^2 + \sum_{\mathsf{angle}}^{\mathsf{MM\,MM}} \frac{k_{\mathsf{a}}}{2}(\theta-\theta_0)^2 + \sum_{\mathsf{dihedral}}^{\mathsf{MM\,MM}} \frac{k_{\mathsf{d}}}{2}(1+\cos\Theta) \\
& + \sum_{m,n}\left(\frac{A_{mn}}{r^{12}_{mn}} - \frac{B_{mn}}{r^{6}_{mn}}\right) + \sum_{\mathsf{Hydrogen\ bond}}\left(\frac{C_{mn}}{r^{12}_{mn}} - \frac{D_{mn}}{r^{10}_{mn}}\right) \\
& + \sum_{m,n}\frac{q_m q_n}{r_{mn}} \\
& + \frac{1}{2}\sum_{m<n} q_m \frac{\boldsymbol{\mu}_n \cdot \boldsymbol{r}_{mn}}{r^3_{mn}} + \frac{1}{2}\sum_{m>n} q_m \frac{\boldsymbol{\mu}_n \cdot \boldsymbol{r}_{mn}}{r^3_{mn}} - \frac{1}{2}\sum_{m}\alpha_m \boldsymbol{E}_m \cdot \boldsymbol{E}_m
\end{aligned}$$

$$\begin{aligned}
E_{\mathsf{QMMM}} = & \sum_{\mathsf{bond}}^{\mathsf{QM\,MM}} \frac{k_{\mathsf{b}}}{2}(r-r_0)^2 + \sum_{\mathsf{angle}}^{\mathsf{QM\,MM}} \frac{k_{\mathsf{a}}}{2}(\theta-\theta_0)^2 + \sum_{\mathsf{dihedral}}^{\mathsf{QM\,MM}} \frac{k_{\mathsf{d}}}{2}(1+\cos\Theta) \\
& + \sum_{s,m}\left(\frac{A_{sm}}{r^{12}_{sm}} - \frac{B_{sm}}{r^{6}_{sm}}\right) + \sum_{\mathsf{Hydrogen\ bond}}\left(\frac{C_{sm}}{r^{12}_{sm}} - \frac{D_{sm}}{r^{10}_{sm}}\right) \\
& \boxed{+ \sum_{s,m}\frac{q_s q_m}{r_{sm}}} \\
& \boxed{+ \frac{1}{2}\sum_{s,m} q_s \frac{\boldsymbol{\mu}_m \cdot \boldsymbol{r}_{sm}}{r^3_{sm}}}
\end{aligned}$$

Establishing the main hamiltonian

Approach II - The total energy

Assuming, that we are working in a closed shell system and thus applying a restricted Hartree Fock approach we end up with the following energy expression. Note, that $E_{\mathsf{MM}} = H_{\mathsf{MM}}$ but $\underline{E_{\mathsf{QMMM}} \neq H_{\mathsf{QMMM}}}$.

$$E = E_{\mathsf{QM}} + E_{\mathsf{MM}} + E_{\mathsf{QMMM}}$$

$$E_{\mathsf{QM}} = \frac{1}{2}\sum_{\alpha\beta} P_{\alpha\beta}(H_{\alpha\beta}^{\mathsf{QM}} + F_{\alpha\beta}^{\mathsf{QM}}) + \sum_{k,l}\frac{Z_k Z_l}{r_{kl}}$$

$$\begin{aligned} E_{\mathsf{MM}} = & \sum_{\mathsf{bond}}^{\mathsf{MM\,MM}} \frac{k_{\mathsf{b}}}{2}(r-r_0)^2 + \sum_{\mathsf{angle}}^{\mathsf{MM\,MM}} \frac{k_{\mathsf{a}}}{2}(\theta-\theta_0)^2 + \sum_{\mathsf{dihedral}}^{\mathsf{MM\,MM}} \frac{k_{\mathsf{d}}}{2}(1+\cos\Theta) \\ & + \sum_{m,n}\left(\frac{A_{mn}}{r_{mn}^{12}} - \frac{B_{mn}}{r_{mn}^{6}}\right) + \sum_{\mathsf{Hydrogen\ bond}}\left(\frac{C_{mn}}{r_{mn}^{12}} - \frac{D_{mn}}{r_{mn}^{10}}\right) \\ & + \sum_{m,n}\frac{q_m q_n}{r_{mn}} \\ & + \frac{1}{2}\sum_{m<n} q_m \frac{\boldsymbol{\mu}_n \cdot \boldsymbol{r}_{mn}}{r_{mn}^3} + \frac{1}{2}\sum_{m>n} q_m \frac{\boldsymbol{\mu}_n \cdot \boldsymbol{r}_{mn}}{r_{mn}^3} - \frac{1}{2}\sum_m \alpha_m \boldsymbol{E}_m \cdot \boldsymbol{E}_m \end{aligned}$$

$$\begin{aligned} E_{\mathsf{QMMM}} = & \sum_{\mathsf{bond}}^{\mathsf{QM\,MM}} \frac{k_{\mathsf{b}}}{2}(r-r_0)^2 + \sum_{\mathsf{angle}}^{\mathsf{QM\,MM}} \frac{k_{\mathsf{a}}}{2}(\theta-\theta_0)^2 + \sum_{\mathsf{dihedral}}^{\mathsf{QM\,MM}} \frac{k_{\mathsf{d}}}{2}(1+\cos\Theta) \\ & + \sum_{s,m}\left(\frac{A_{sm}}{r_{sm}^{12}} - \frac{B_{sm}}{r_{sm}^{6}}\right) + \sum_{\mathsf{Hydrogen\ bond}}\left(\frac{C_{sm}}{r_{sm}^{12}} - \frac{D_{sm}}{r_{sm}^{10}}\right) \\ & \boxed{\boxed{-\sum_{\alpha\beta} P_{\alpha\beta}\left(\sum_m \int \phi_\alpha \frac{q_m}{r_{im}} \phi_\beta \, dr\right)} + \sum_{k,m}\frac{Z_k q_m}{r_{km}}} \\ & \boxed{\boxed{-\frac{1}{2}\sum_{\alpha\beta} P_{\alpha\beta}\left(\sum_m \int \phi_\alpha \frac{\boldsymbol{\mu}_m \cdot \boldsymbol{r}_{im}}{r_{im}^3} \phi_\beta \, dr\right)} + \frac{1}{2}\sum_{k,m} Z_k \frac{\boldsymbol{\mu}_m \cdot \boldsymbol{r}_{km}}{r_{km}^3}} \end{aligned}$$

Example

QM $(6-31G*)$ and QMMM data from [1]		
$r_{\mathsf{C-Cl}}$ — Å	ΔE_{QM} — kcal/mole[a]	ΔE_{QMMM} — kcal/mole[b]
6.014	0.0	0.0
4.571	−2.6	−5.4
3.274	−6.7	0.5
2.682	−1.0	10.3
2.383	7.2	18.3

QM $(6-31G*)$ and QMSM data from [4]		
$r_{\mathsf{C-Cl}}$ — Å	ΔE_{QM} — kcal/mole[c]	ΔE_{QMSM} — kcal/mole[d]
∞	0.0	
7.500	−2.4	0.0
6.000	−3.6	
4.500	−6.4	
3.269	−10.3	
2.974	−9.2	
2.678	−4.5	
2.530	−0.1	
2.383	3.6	28.0

[a] Energy differences scaled according to the $r_{\mathsf{C-Cl}} = 6.014$ structure
[b] Energy differences scaled according to the $r_{\mathsf{C-Cl}} = 6.014$ structure
[c] Energy differences scaled according to the $r_{\mathsf{C-Cl}} = \infty$ structure
[d] Energy differences scaled according to the $r_{\mathsf{C-Cl}} = 7.500$ structure

Example

QM and QMMM data from [1]		
$r_{\mathsf{C-Cl}}$ — Å	ΔE_{QM} — kcal/mole^a	ΔE_{QMMM} — kcal/mole^b
6.014	0.0	0.0
4.571	−2.6	−5.4
3.274	−6.7	0.5
2.682	−1.0	10.3
2.383	7.2	18.3

E_{QMMM} and $E_{\mathsf{QMMM}} + E_{\mathsf{MM}}$ data from [1]		
$r_{\mathsf{C-Cl}}$ — Å	E_{QMMM} — kcal/mole	$E_{\mathsf{QMMM}} + E_{\mathsf{MM}}$ — kcal/mole
6.014	−131.1	−1208.2
4.571	−126.5	−1211.0
3.274	−109.3	−1201.0
2.682	−101.5	−1196.9
2.383	−92.3	−1197.1

[a] Energy differences scaled according to the $r_{\mathsf{C-Cl}} = 6.014$ structure
[b] Energy differences scaled according to the $r_{\mathsf{C-Cl}} = 6.014$ structure

Acknowledgements

- Acknowledgements go to the people initiating this course
  - Kurt Mikkelsen $\left\{ \begin{array}{l} \text{University} \\ \text{of} \\ \text{Copenhagen} \end{array} \right.$
  - Lars Hemmingsen $\left\{ \begin{array}{l} \text{Royal Veterinary} \\ \text{and} \\ \text{Agricultural University} \end{array} \right.$
- Acknowledgements also go to all you listeners

Bibliography

[1] 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 : Applications to the $CH_3Cl + Cl^-$ Exchange Reaction and Gas Phase Protonation of Polyethers. *Journal of Computational Chemistry*, 7(6):718–730, 1986.

[2] M. J. Field, P. A. Bash, and M. Karplus. A Combined Quantum Mechanical and Molecular Mechanical Potential for Molecular Dynamics Simulations. *Journal of Computational Chemistry*, 11(6):700–733, 1990.

[3] J. Gao. Methods and Applications of Combined Quantum Mechanical and Molecular Mechanical Potentials. In K. B. Libkowitz and D. B. Boyd, editors, *Reviews in Computational Chemistry*, volume 7, pages 119–185. VCH Publishers, 1996.

[4] J. Chandrasekhar, S. F. Smith, and W. L. Jorgensen. Theoretical examination of the S_N2 Reaction involving Cloride Ion and Methyl Chloride in the Gas Phase and Aqueous Solution. *Journal of the American Chemical Society*, 107:154–163, 1985.