QM/QM and QM/MM Hybrid Methods

Subsystem and Embedding Methods for Quantum Chemistry

Johannes Neugebauer

ESQC, Sicily, 2022

Lecture 1

Contents

Part 1: Introduction

- general strategies
- intermolecular interactions

Part 2: QM/MM and QM/continuum methods

- partitioning schemes, energy terms
- mechanical, Hamiltonian, polarizable embedding
- continuum solvation, polarizable continuum models

Part 3: QM/QM methods: An introduction

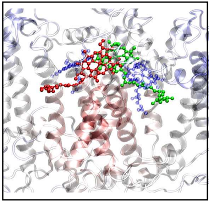
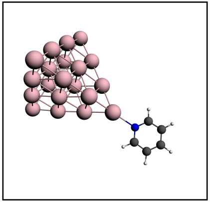
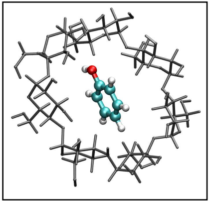
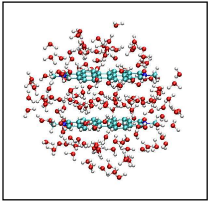
- energy-based partitioning

Introduction

“[A] reliable quantum mechanical description . . . of more than a few atoms was practically impossible for a very long time. . . . The solution to this challenge . . . has emerged from the realization that a description of the properties of complex systems does not require the representation of all parts of the system at the same level of detail.”

A. Warshel, Nobel Lecture, December 8, 2013.

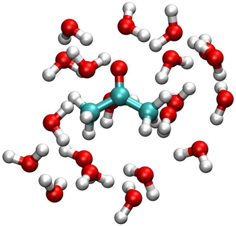
Why Embedding/Subsystem Approaches?

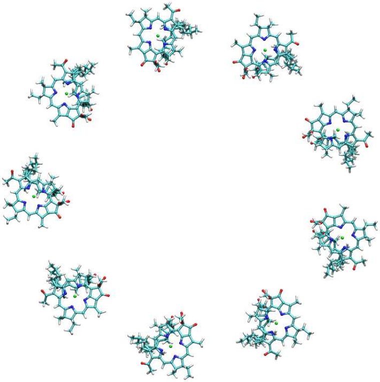





Some answers:

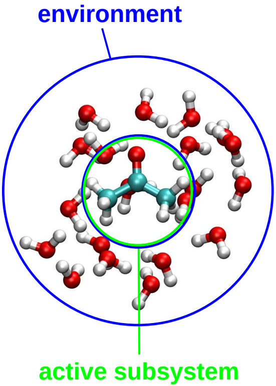
- standard QC calculations on system+environment may be very costly (or even unfeasible)
- huge amount of CPU time would be spent on a presumably small effect of the environment
- conformational sampling shall be avoided if possible
- for analysis, we want to separate solute and solvent properties/effects

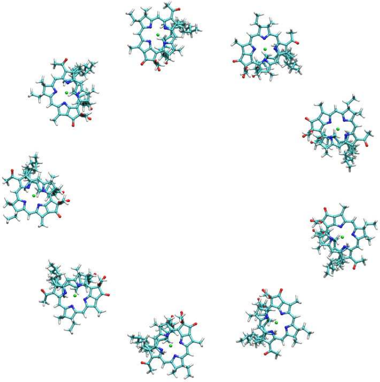
Subsystem vs. Embedding Approaches





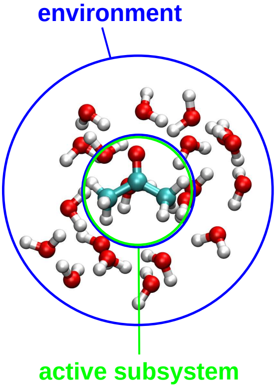
Subsystem vs. Embedding Approaches

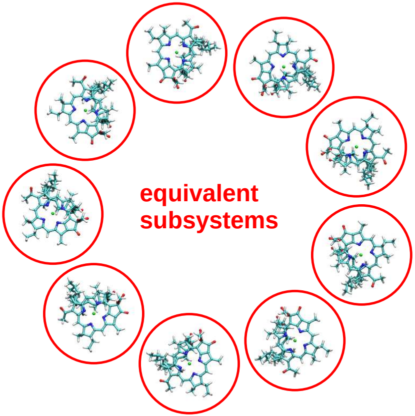




- embedding (“focussed”) approach:
  high accuracy for active part, lower accuracy for environment *and* interaction

Subsystem vs. Embedding Approaches





- embedding (“focussed”) approach:
  high accuracy for active part, lower accuracy for environment *and* interaction
- subsystem (fragmentation) approach:
  high accuracy within fragments, lower accuracy for fragment interaction

How to approach the problem?

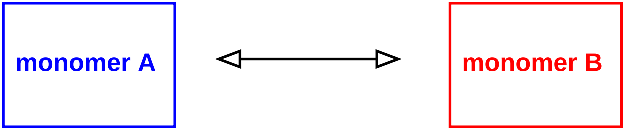
Starting point 1: Isolated subsystems

- treat interaction as perturbation

Starting point 2: Total system

- partition quantum mechanical descriptor into subsystem contributions
  (e.g., wave function, density, density matrix, Green’s function)
- derive optimization conditions for subsystems

Interacting Subsystems: Some initial thoughts



monomer A

monomer B

- isolated monomers can be described by individual Hamiltonians:

$$\begin{aligned}\hat{H}_A\Psi_A &= E_A\Psi_A \\ \hat{H}_B\Psi_B &= E_B\Psi_B\end{aligned}$$

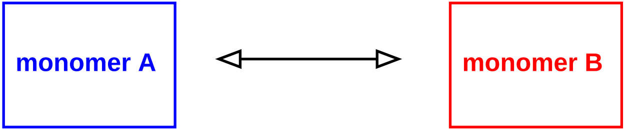
- assuming zero interaction, the total wavefunction can be factorized,

$$(\hat{H}_A + \hat{H}_B)\Psi_A \cdot \Psi_B = (E_A + E_B)\Psi_A \cdot \Psi_B$$

- true dimer Hamiltonian including interaction:

$$\hat{H}_{(A+B)} = \hat{H}_A + \hat{H}_B + \hat{H}_{A\leftrightarrow B}$$

Interacting Subsystems: Some initial thoughts



monomer A

monomer B

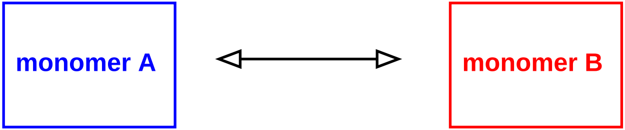
- if interaction is weak: use perturbation theory

$$E^{(1)} = \langle \Psi_A \Psi_B | \hat{H}_{A\leftrightarrow B} | \Psi_A \Psi_B \rangle$$

- form of the interaction Hamiltonian:

$$\hat{H}_{A\leftrightarrow B} = -\sum_{I\in A}^{N_A}\sum_{i\in B}^{n_B} \frac{Z_I}{|\mathbf{R}_I - \mathbf{r}_i|} - \sum_{I\in B}^{N_B}\sum_{i\in A}^{n_A} \frac{Z_I}{|\mathbf{R}_I - \mathbf{r}_i|} + \sum_{i\in A}^{n_A}\sum_{j\in B}^{n_B} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{I\in A}^{N_A}\sum_{J\in B}^{N_B} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}$$

Interacting Subsystems: Some initial thoughts



monomer A

monomer B

- if interaction is weak: use perturbation theory

$$E^{(1)} = \langle \Psi_A \Psi_B | \hat{H}_{A\leftrightarrow B} | \Psi_A \Psi_B \rangle$$

- form of the interaction Hamiltonian:

$$\hat{H}_{A\leftrightarrow B} = -\sum_{I\in A}^{N_A}\sum_{i\in B}^{n_B}\frac{Z_I}{|\mathbf{R}_I - \mathbf{r}_i|} - \sum_{I\in B}^{N_B}\sum_{i\in A}^{n_A}\frac{Z_I}{|\mathbf{R}_I - \mathbf{r}_i|} + \sum_{i\in A}^{n_A}\sum_{j\in B}^{n_B}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{I\in A}^{N_A}\sum_{J\in B}^{N_B}\frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}$$

- conceptual problem: electrons are indistinguishable particles . . .

Interacting Subsystems: Some initial thoughts

- wave function for N independent subsystems:

$$\Psi^{\text{tot}}(\mathbf{r}_{1_1}, \mathbf{r}_{2_1}, \ldots, \mathbf{r}_{n_N}) = \Psi_1(\mathbf{r}_{1_1} \ldots) \cdot \Psi_2(\mathbf{r}_{1_2} \ldots) \ldots \cdot \Psi_N(\mathbf{r}_{1_N} \ldots \mathbf{r}_{n_N})$$

- non-interacting subsystems: product ansatz

Interacting Subsystems: Some initial thoughts

- wave function for N independent subsystems:

$$\Psi^{\text{tot}}(\mathbf{r}_{1_1}, \mathbf{r}_{2_1}, \ldots, \mathbf{r}_{n_N}) = \hat{A}\Psi_1(\mathbf{r}_{1_1}\ldots) \cdot \Psi_2(\mathbf{r}_{1_2}\ldots) \ldots \cdot \Psi_N(\mathbf{r}_{1_N}\ldots\mathbf{r}_{n_N})$$

- non-interacting subsystems: product ansatz
- antisymmetrisation necessary (but no effect for non-overlapping wavefunctions)

Interacting Subsystems: Some initial thoughts

- wave function for N independent subsystems:

$$\Psi^{\text{tot}}(\mathbf{r}_{1_1}, \mathbf{r}_{2_1}, \ldots, \mathbf{r}_{n_N}) = \sum_{\{K_I\}} c_{\{K_I\}} \hat{A} \Psi_{K_1}(\mathbf{r}_{1_1} \ldots) \cdot \Psi_{K_2}(\mathbf{r}_{1_2} \ldots) \ldots \cdot \Psi_{K_N}(\mathbf{r}_{1_N} \ldots \mathbf{r}_{n_N})$$

- non-interacting subsystems: product ansatz
- antisymmetrisation necessary (but no effect for non-overlapping wavefunctions)
- maybe: configuration interaction

Long-Range Interactions: Results from PT

“Polarization approximation”: Antisymmetrization ignored!

- first order: electrostatic interaction (unperturbed monomers)

$$
\begin{aligned}
E^{(1)} &= \langle \Psi_{0,A}\Psi_{0,B}|\hat{H}_{A\leftrightarrow B}|\Psi_{0,A}\Psi_{0,B}\rangle \\
&= \int \rho_A(\mathbf{r})v_{\mathrm{nuc,B}}(\mathbf{r})\mathrm{d}\mathbf{r} + \int \rho_B(\mathbf{r})v_{\mathrm{nuc,A}}(\mathbf{r})\mathrm{d}\mathbf{r} \\
&+ \int \frac{\rho_A(\mathbf{r})\rho_B(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}' + \sum_{I\in A}^{N_A}\sum_{J\in B}^{N_B}\frac{Z_I Z_J}{|\mathbf{R}_I-\mathbf{R}_J|}
\end{aligned}
$$

- possible starting point for further approximations:
  - point-charge representation of total charge density (fitted to ESP)
  - multipole expansion of $\rho_A(\mathbf{r})$ and/or $\rho_B(\mathbf{r})$

Long-Range Interactions: Results from PT

"Polarization approximation": Antisymmetrization ignored!

- second order: $E^{(2)} = E^{(2)}_{\text{ind,A}} + E^{(2)}_{\text{ind,B}} + E^{(2)}_{\text{disp}}$,

$$
\begin{aligned}
E^{(2)}_{\text{ind,A}} &= \sum_K \frac{\langle \Psi_{0,A}\Psi_{0,B}|\hat{H}_{A\leftrightarrow B}|\Psi_{K,A}\Psi_{0,B}\rangle \langle \Psi_{K,A}\Psi_{0,B}|\hat{H}_{A\leftrightarrow B}|\Psi_{0,A}\Psi_{0,B}\rangle}{E_{A,K} - E_{A,0}} \\
E^{(2)}_{\text{ind,B}} &= \sum_K \frac{\langle \Psi_{0,A}\Psi_{0,B}|\hat{H}_{A\leftrightarrow B}|\Psi_{0,A}\Psi_{K,B}\rangle \langle \Psi_{0,A}\Psi_{K,B}|\hat{H}_{A\leftrightarrow B}|\Psi_{0,A}\Psi_{0,B}\rangle}{E_{B,K} - E_{B,0}} \\
E^{(2)}_{\text{disp}} &= \sum_{KL} \frac{\langle \Psi_{0,A}\Psi_{0,B}|\hat{H}_{A\leftrightarrow B}|\Psi_{K,A}\Psi_{L,B}\rangle \langle \Psi_{K,A}\Psi_{L,B}|\hat{H}_{A\leftrightarrow B}|\Psi_{0,A}\Psi_{0,B}\rangle}{E_{A,K} + E_{B,L} - E_{A,0} - E_{B,0}}
\end{aligned}
$$

Short-Range Interactions

Phenomenologically:

- exchange-repulsion (exchange + Pauli repulsion), $\propto \exp[-\eta R]$
- “charge-transfer” (sometimes included in induction)
- charge penetration (effect of non-converging multipole expansions)
- additional terms arise if correlation is considered (SAPT)

Short-Range Interactions

Perturbation theory:

- first-order term (leading contributions):

$$E^{(1)} \quad = \quad \frac{\langle \Psi_{0,A}\Psi_{0,B}|\hat{H}_{A\leftrightarrow B}|\Psi_{0,A}\Psi_{0,B}\rangle + \langle \Psi_{0,A}\Psi_{0,B}|\hat{H}_{A\leftrightarrow B}\hat{P}_{\mathrm{AB}}|\Psi_{0,A}\Psi_{0,B}\rangle}{1 + \langle \Psi_{0,A}\Psi_{0,B}|\hat{P}_{\mathrm{AB}}|\Psi_{0,A}\Psi_{0,B}\rangle}$$

- terms in blue: $\propto S^2$

⇒ models often employ corrections based on overlap

- . . . but antisymmetrization complicates the perturbation analysis . . .

For more details, see, e.g., A. Stone, *The Theory of Intermolecular Forces*, Oxford University Press, 2013.

QM/MM Partitioning

Starting Point for Approximations

- partition Hamiltonian into active system (A) and environment (E)

$$\hat{H}_{A+E} = \hat{H}_A + \hat{H}_E + \hat{H}_{A\leftrightarrow E}$$

- treat A, E, and/or $A \leftrightarrow E$ with different approximations
- maybe several active systems, maybe more layers, ...

Starting Point for Approximations

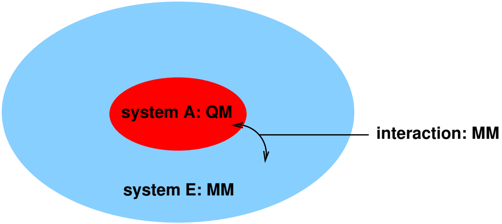
- partition Hamiltonian into active system (A) and environment (E)

$$\hat{H}_{A+E} = \hat{H}_A + \hat{H}_E + \hat{H}_{A\leftrightarrow E}$$

- treat A, E, and/or $A \leftrightarrow E$ with different approximations
- maybe several active systems, maybe more layers, . . .

But how exactly? At which stage shall we introduce approximations?

QM/MM Hybrid Methods



- additive QM/MM:

$$E_{\text{tot}}^{\text{add}} \quad = \quad E_A^{\text{QM}} + E_E^{\text{MM}} + E_{A\leftrightarrow E}^{\text{MM}}$$

- subtractive QM/MM (easier to generalize to QM/QM methods):

$$E_{\text{tot}}^{\text{sub}} \quad = \quad E_{(A+E)}^{\text{MM}} + (E_A^{\text{QM}} - E_A^{\text{MM}})$$

- identical if

$$E_{(A+E)}^{\text{MM}} = E_A^{\text{MM}} + E_E^{\text{MM}} + E_{A\leftrightarrow E}^{\text{MM}} \quad \Leftrightarrow \quad E_{A\leftrightarrow E}^{\text{MM}} = E_{(A+E)}^{\text{MM}} - E_A^{\text{MM}} - E_E^{\text{MM}}$$

QM/MM: Partitioning the Energy

Three energy contributions:

(1) active-system energy

$$E_A^{\mathrm{QM}} = \langle \Psi_A | \hat{H}_A | \Psi_A \rangle$$

(note: $\hat{H}_A$ is the isolated-system Hamiltonian for system A)

(2) environment energy

$$E_E^{\mathrm{MM}} = E_{\mathrm{bonded}}(\mathbf{R}_E) + E_{\mathrm{nonbonded}}(\mathbf{R}_E)$$

where typically

$$\begin{aligned} E_{\mathrm{bonded}}(\mathbf{R}_E) &= E_{\mathrm{stretch}}(\mathbf{R}_E) + E_{\mathrm{bend}}(\mathbf{R}_E) + E_{\mathrm{torsion}}(\mathbf{R}_E) \\ E_{\mathrm{nonbonded}}(\mathbf{R}_E) &= E_{\mathrm{electrostatic}}(\mathbf{R}_E) + E_{\mathrm{vdW}}(\mathbf{R}_E) \end{aligned}$$

E_{vdW} accounts for dispersion and exchange-repulsion

QM/MM: Partitioning the Energy

Prototypical MM energy expressions:

$$E_{\text{stretch}}(\mathbf{R}_E) = \sum_b^{\text{bonds}} k_b (d_b - d_{b,0})^2$$

$$E_{\text{bend}}(\mathbf{R}_E) = \sum_a^{\text{angles}} k_a (\theta_a - \theta_{a,0})^2$$

$$E_{\text{torsion}}(\mathbf{R}_E) = \sum_d^{\text{dihedrals}} k_d [1 + \cos(n\phi + \delta)]^2$$

$$E_{\text{electrostatic}}(\mathbf{R}_E) = \sum_{a<b}^{\text{nb pairs}} \frac{Q_a Q_b}{R_{ab}}$$

$$E_{\text{vdW}}(\mathbf{R}_E) = \sum_{a<b}^{\text{nb pairs}} \epsilon_{ab} \left[\left(\frac{\sigma_{ab}}{r_{ab}} \right)^{12} - \left(\frac{\sigma_{ab}}{r_{ab}} \right)^{6} \right]$$

QM/MM: Partitioning the Energy

Three energy contributions (cont'd):

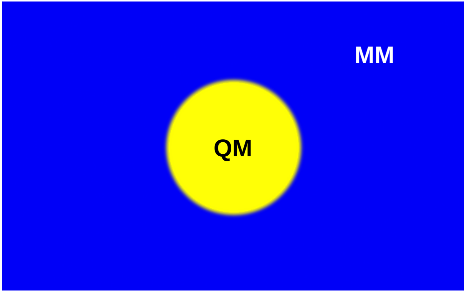
(3) QM↔MM interaction: Apply MM energy expression

$$E^{\mathrm{MM}}_{A\leftrightarrow E} = E_{\mathrm{bonded},A\leftrightarrow E} + E_{\mathrm{vdW},A\leftrightarrow E} + E_{\mathrm{electrostatic},A\leftrightarrow E}$$

Notes:

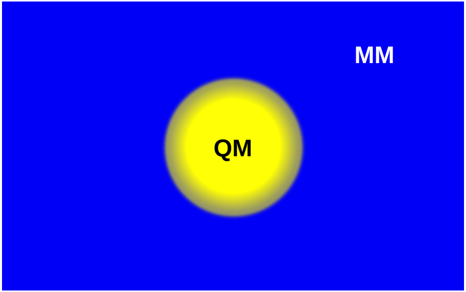
- availability of suitable FF parameters for QM atoms can be problematic
- bonds across the QM-MM boundary require link-atom formalism

QM/MM Interaction



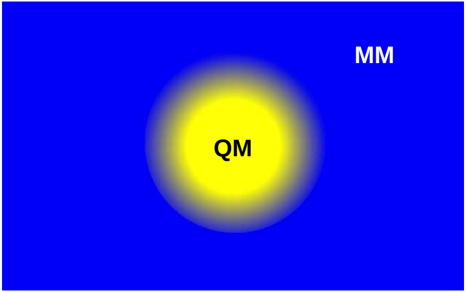
- Mechanical Embedding: Wavefunction of QM part unchanged

QM/MM Interaction



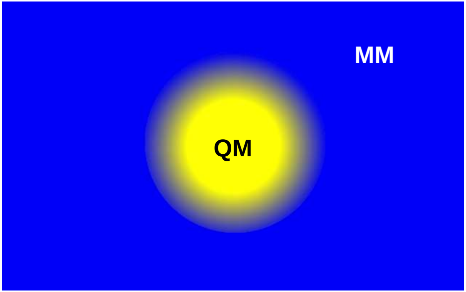
- Mechanical Embedding: Wavefunction of QM part unchanged
- Hamiltonian/Electronic Embedding: QM part polarized by environment

QM/MM Interaction



- Mechanical Embedding: Wavefunction of QM part unchanged
- Hamiltonian/Electronic Embedding: QM part polarized by environment
- Polarizable Embedding: Mutual polarization of QM and MM part

QM/MM Interaction



- Mechanical Embedding: Wavefunction of QM part unchanged
- Hamiltonian/Electronic Embedding: QM part polarized by environment
- Polarizable Embedding: Mutual polarization of QM and MM part

Note: This usually only applies to electrostatics, but could be generalized

QM/MM Interaction

Mechanical Embedding:

- $E_{\text{electrostatic},A\leftrightarrow E}$ evaluated with MM charges for QM atoms

⇒ No polarization of QM wavefunction

⇒ Only indirect effects (change in equilibrium structure)

⇒ Changes in QM charge distribution not reflected in interaction term
(e.g., along reaction coordinates)

QM/MM Interaction

Hamiltonian Embedding:

- electrostatic interaction Hamiltonian:

$$\hat{H}_{\mathrm{elstat},A\leftrightarrow E} = -\sum_{i}^{\mathrm{el.}} \sum_{a}^{\mathrm{MM\ atoms}} \frac{q_a}{|\mathbf{r}_i - \mathbf{R}_a|} + \sum_{I}^{\mathrm{nuc.}} \sum_{a}^{\mathrm{MM\ atoms}} \frac{Z_I q_a}{|\mathbf{R}_I - \mathbf{R}_a|}$$

$$= \sum_{i}^{\mathrm{el.}} v_{\mathrm{emb}}^{\mathrm{elstat}}(\mathbf{r}_i) + \sum_{I}^{\mathrm{nuc.}} \sum_{a}^{\mathrm{MM\ atoms}} \frac{Z_I q_a}{|\mathbf{R}_I - \mathbf{R}_a|}$$

⇒ obtain Ψ_A from

$$\hat{H}'_A \Psi_A = \left(\hat{H}_A + \sum_{i}^{\mathrm{el.}} v_{\mathrm{emb}}^{\mathrm{elstat}}(\mathbf{r}_i) \right) \Psi_A = E'_A \Psi_A$$

⇒ polarizes Ψ_A, changes all energy terms in $E_A^{\mathrm{QM}} = \langle \Psi_A | \hat{H}_A | \Psi_A \rangle$

QM/MM Interaction

Hamiltonian Embedding:

- electrostatic interaction energy:

$$E_{\text{elstat},A\leftrightarrow E} = \left\langle \Psi_A \left| \sum_i^{\text{el.}} v_{\text{emb}}^{\text{elstat}}(\mathbf{r}_i) \right| \Psi_A \right\rangle + \sum_I^{\text{nuc.}} \sum_a^{\text{MM atoms}} \frac{Z_I q_a}{|\mathbf{R}_I - \mathbf{R}_a|}$$

- total QM/MM energy:

$$\begin{aligned} E_{\text{tot}}^{\text{QM/MM}} &= E_A^{\text{QM}} + E_E^{\text{MM}} + E_{\text{vdW},A\leftrightarrow E} \\ &\quad + \left\langle \Psi_A \left| \sum_i^{\text{el.}} v_{\text{emb}}^{\text{elstat}}(\mathbf{r}_i) \right| \Psi_A \right\rangle + \sum_I^{\text{nuc.}} \sum_a^{\text{MM atoms}} \frac{Z_I q_a}{|\mathbf{R}_I - \mathbf{R}_a|} \\ &= E'_A + E_E^{\text{MM}} + E_{\text{vdW},A\leftrightarrow E} + \sum_I^{\text{nuc.}} \sum_a^{\text{MM atoms}} \frac{Z_I q_a}{|\mathbf{R}_I - \mathbf{R}_a|} \end{aligned}$$

(assuming no bonding interactions across the QM/MM boundary)

QM/MM Interaction

Polarizable Embedding:
Introduce flexible MM charge model, polarized by QM charge distribution

Induced point dipoles:

- assign polarizabilities to all MM atoms (or general expansion points)
- determine induced dipoles as:

$$\boldsymbol{\mu}_a^{\mathrm{ind}} \quad = \quad \boldsymbol{\alpha}_a \cdot \vec{F}(\mathbf{R}_a)$$

- electric field at MM atom a originates from MM point charges, other induced dipoles, and QM charge distribution

Alternatives:

- fluctuating charges
- drude oscillators (charge-on-spring models)

An Example: The DRF Model

Discrete Reaction Field Operator

$$\hat{H}_{\mathrm{QM/MM}} = \sum_i v^{\mathrm{DRF}}(\mathbf{r}_i)$$

$$v^{\mathrm{DRF}}(\mathbf{r}) = \sum_a^{\text{MM atoms}} \frac{q_a}{|\mathbf{r}-\mathbf{R}_a|} + \sum_{\mathrm{a}}^{\text{MM atoms}} \boldsymbol{\mu}_a^{\mathrm{ind}} \cdot \frac{(\mathbf{r}-\mathbf{R}_s)}{|\mathbf{r}-\mathbf{R}_s|}^3$$

- induced dipoles are computed self-consistently from

$$\mu_a^{\mathrm{ind}} = \alpha_a \left[\boldsymbol{F}^{\text{QM, nuc}}(\mathbf{R}_a) + \boldsymbol{F}^{\text{QM, el}}(\mathbf{R}_a) + \boldsymbol{F}^{\text{MM charges}}(\mathbf{R}_a) + \boldsymbol{F}^{\text{MM ind. dip}}(\mathbf{R}_a) \right]$$

T.D. Poulsen, P.R. Ogilby, K.V. Mikkelsen, *J. Chem. Phys.* 2002, **116**, 3730;

L. Jensen, P.Th. van Duijnen, J.G. Snijders, *J. Chem. Phys.* 2003, **118**, 514.

Other Polarizable Embedding Schemes

- PE model (Kongsted and co-workers)
- QM/MMpol (Mennucci and co-workers)
- EFP (Gordon, Slipchenko and co-workers)
- . . .

Parameters for charges/polarizabilities

- either calculated for solvent molecules/env. fragments (e.g., CHELPG charges, LOPROP polarizabilities)
- or fitted, e.g., to experimental polarizabilities

QM/MM Interaction

Hamiltonian Embedding:

- effects of dispersion/Pauli repulsion: often not included in $\hat{H}'_A$

⇒ MM atoms can act as electron traps (*overpolarization*, *electron spill-out*, or *electron leaking*)

- one solution strategy: replace Coulombic potential $(-1/r)$ by

$$v_a(r) = -\frac{r^n_{a,\mathrm{cov}} - r^n}{r^{n+1}_{a,\mathrm{cov}} - r^{n+1}}$$

A. Laio, J. VandeVondele, U. Rothlisberger, *J. Chem. Phys.* 2002, **116**, 6941.

- strategies to include Hamiltonian embedding for Pauli repulsion/dispersion based on “model MM density matrices”:

T. Giovannini, P. Lafiosca, C. Cappelli, *J. Chem. Theory Comput.* 2017, **13**, 4854.

QM-derived Embedding Potentials

The Effective Fragment Potential (EFP) method

- effective-fragment orbitals are obtained from DFT/HF (+ localization)
- interaction Hamiltonian:

$$\hat{H}_{A\leftrightarrow E} = \hat{V}_{\mathrm{elstat}} + \hat{V}_{\mathrm{pol}} + \hat{V}_{\mathrm{ex-rep}}$$

- $\hat{V}_{\mathrm{ex-rep}}$ denotes the exchange-repulsion interaction,

$$\hat{V}_{\mathrm{ex-rep}} \quad = \quad \sum_i \sum_j \beta_j e^{-\alpha_j (r_i - R_j)^2}$$

(local potential; R_j is coordinate vector of an LMO centroid; α_j and β_j are parameters)

C.I. Viquez Rojas, L.V. Slipchenko, J. Chem. Theory Comput. 2020, **16**, 6408.

QM-derived Embedding Potentials

Electrostatic potential in EFP:

- based on a *distributed multipolar analysis* of fragment charge distributions
- expansion points k:
  - each nuclear center
  - each bond midpoint
- multipoles up to octopoles

$$\hat{V}_k^{\text{elstat}}(\mathbf{r}) = q_k T(\mathbf{r}_k) - \sum_a^{x,y,z} \mu_a^k T_a(\mathbf{r}_k) + \frac{1}{3} \sum_{a,b}^{x,y,z} \Theta_{a,b}^k T_{a,b}(\mathbf{r}_k) - \frac{1}{15} \sum_{a,b}^{x,y,z} \Omega_{a,b,c}^k T_{a,b,c}(\mathbf{r}_k)$$

- general definition of full interaction tensor: $\boldsymbol{T}^{(k)}(\mathbf{r}_k) = \nabla^k \frac{1}{|\mathbf{r}-\mathbf{r}_k|}$
- here: $T(\mathbf{r}_k) = 1/|\mathbf{r} - \mathbf{r}_k|$, T_a: element of $\boldsymbol{T}^{(1)}(\mathbf{r}_k) = \nabla \frac{1}{|\mathbf{r}-\mathbf{r}_k|}$ etc.

M. Gordon *et al.*, J. Phys. Chem. A 2001, **105**, 293.

C.I. Viquez Rojas, L.V. Slipchenko, J. Chem. Theory Comput. 2020, **16**, 6408.

QM-derived Embedding Potentials

Polarization in EFP:

- based on *distributed polarizabilities* on LMO centroids

⇒ localized “orbital” polarizabilities from finite field calculations

M. Gordon *et al.*, J. Phys. Chem. A 2001, **105**, 293.

C.I. Viquez Rojas, L.V. Slipchenko, J. Chem. Theory Comput. 2020, **16**, 6408.

Continuum Solvation Models

Continuum Solvation Models

Overall Goal: Solvation Free Energies

$$\Delta G_{\mathrm{solv}} = \Delta G_{\mathrm{cavity}} + \Delta G_{\mathrm{Dispersion}} + \Delta G_{\mathrm{elstat}}$$

- cavity formation and dispersion terms:
  often assumed proportional to “solute surface”
- only electrostatic term enters the QM Hamiltonian
- solvent: homogeneous medium with dielectric constant ϵ_{out}
- electrostatic potential $\varphi(\mathbf{r})$: from Poissson’s equation,

$$\boldsymbol{\nabla} \cdot [\epsilon(\mathbf{r}) \boldsymbol{\nabla} \varphi(\mathbf{r})] = -4\pi\rho(\mathbf{r})$$

- note: $\epsilon(\mathbf{r}) = \epsilon_{\mathrm{in}} = 1$ within QM region, ϵ_{out} outside

J. Tomasi, B. Mennucci, R. Cammi, *Chem. Rev.* 2005, **105**, 2999;
J.M. Herbert, *WIREs Comput. Mol. Sci.* 2021, **11**, e1519;
F. Jensen, *Introduction to Compututational Chemistry*, 2nd ed., Wiley, 2007, Section 14.7.

Continuum Solvation Models

- reaction potential: $\varphi_{\text{reac}}(\mathbf{r}) = \varphi(\mathbf{r}) - \varphi_{\text{mol}}(\mathbf{r})$, with

  $$\varphi_{\text{mol}}(\mathbf{r}) = \int \frac{\rho_{\text{mol}}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \mathrm{d}\mathbf{r}'$$

  (electrostatic potential generated by molecular charge density)
- electrostatic solvation energy:

  $$\Delta G_{\text{elstat}} = \frac{1}{2} \int \varphi_{\text{reac}}(\mathbf{r}) \rho_{\text{mol}}(\mathbf{r}) \mathrm{d}\mathbf{r}$$

- analytic solutions available for
  - charge monopole in a spherical cavity (Born model)
  - point dipole in a spherical cavity (Onsager model)
  - general multipoles in a spherical cavity (Kirkwood model)
  - . . . and in ellisoidal cavities (Kirkwood–Westheimer model)

Polarizable Continuum Models

- 3D polarization problem is turned into a surface charge problem
- reaction potential can be expressed exactly as

$$\varphi_{\text{reac}}(\mathbf{r}) = \int_{\text{surface}} \frac{\sigma(\mathbf{s})}{|\mathbf{r}-\mathbf{s}|} \mathrm{d}\mathbf{s}$$

- surface charge $\sigma(\mathbf{s})$: defined through *jump* condition for E-field,

$$\epsilon_{\text{out}}(\vec{n}_s \cdot \vec{\nabla})\varphi(\mathbf{s})|_{\text{out}} \quad = \quad \epsilon_{\text{in}}(\vec{n}_s \cdot \vec{\nabla})\varphi(\mathbf{s})|_{\text{in}}$$

- in practice: surface charge is discretized,

$$\varphi_{\text{reac}}(\mathbf{r}) \approx \sum_k \frac{\sigma(\mathbf{s}_k)A_k}{|\mathbf{r}-\mathbf{s}_k|}$$

Conductor-Like Screening Model (COSMO)

- dielectric constant is set to $\epsilon = \infty$
- most important consequence:

$$\varphi(\mathbf{r}) \quad = \quad \varphi_{\text{reac}}(\mathbf{r}) + \varphi_{\text{mol}}(\mathbf{r}) = 0$$

  at boundary between molecule and dielectric environment

- this condition fixes the values of the surface charges
- the same condition can be derived by minimizing the electrostatic energy of the molecule–surface charge system w.r.t. the surface charges
- ideal, unscreened charges are finally scaled by

$$f(\epsilon) \quad = \quad \frac{\epsilon - 1}{\epsilon + k}$$

  ($k = 0.5$ in original work by Klamt, later versions often use $k = 0$)

A. Klamt, G. Schüürmann, *J. Chem. Soc. Perkin Trans.* **2** (1993), 799;

Conductor-Like Screening Model (COSMO)

- electrostatic part of solvation energy in COSMO

$$E_{\mathrm{el,COSMO}} = \frac{1}{2}\sum_{\mu}\sum_{\nu} q_\mu A_{\mu\nu} q_\nu + \sum_{A}\sum_{\mu} q_\mu B_{A\mu} Z_A + \sum_{\mu} q_\mu C_\mu$$

$q_\mu =$ apparent surface charges (ASCs), from minimization of E_{solv},

$$\mathbf{A}\vec{q} = -(\mathbf{B}\vec{Z} + \vec{C})$$

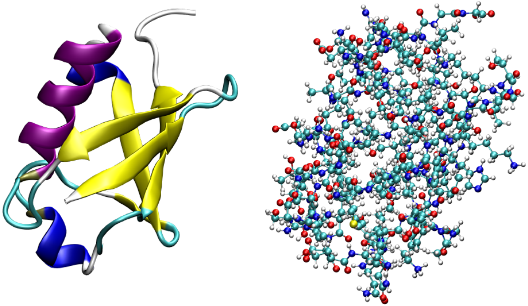
with

$$\begin{aligned} A_{\mu\nu} &= \begin{cases} |\vec{r}_\mu - \vec{r}_\nu|^{-1} & \text{for} \quad \mu \neq \nu \\ 1.07\sqrt{\frac{4\pi}{S_\mu}} & \text{for} \quad \mu = \nu \end{cases} \\ B_{A\mu} &= |\vec{r}_\mu - \vec{r}_A|^{-1} \\ C_\mu &= v_{\mathrm{Coul}}[\rho](\vec{r}_\mu) \end{aligned}$$

A. Klamt, G. Schüürmann, *J. Chem. Soc. Perkin Trans.* **2** (1993), 799;

C.C. Pye, T. Ziegler, *Theor. Chem. Acc.* **101** (1999), 396

A Local COSMO Variant



Ubiquitin, > 1200 atoms; (a) Cartoon representation, (b) atomistic representation

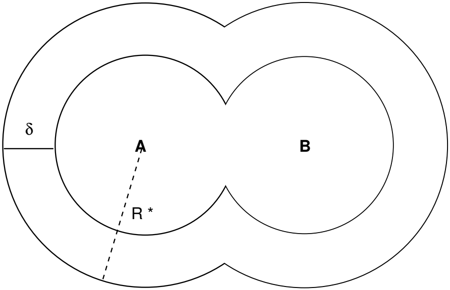
- Problem: Size of matrix $\mathbf{A}$ for large systems

Matrix	Dimension	H_2O	Ubiquitin	FMO
$\mathbf{A}$	$\mathcal{O}(n^2_{\text{ASCs}})$	$12\,769$	$4.2 \cdot 10^8$	$7.0 \cdot 10^9$

- solution: update only *active* ASCs (around active fragment)

A. Goez, J. Neugebauer, *J. Chem. Theory Comput.* 2015 **11**, 5277.

Polarizable Continuum Models



$(R^* = R_A + R_{\mathrm{vdW}}^{\mathrm{solv}})$

Further important aspects:

- shape of molecular cavity
- outlying charge corrections
- extensions for dispersion, exchange–repulsion

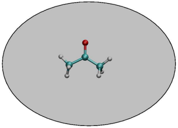
J. Tomasi, B. Mennucci, R. Cammi, *Chem. Rev.* 2005, **105**, 2999;
J.M. Herbert, *WIREs Comput. Mol. Sci.* 2021, **11**, e1519.

QM/MM: Conclusions/Further Comments

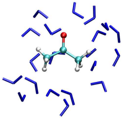
- various ways to include electrostatics (and systematically improve it)
- polarizable force fields also directly affect response properties
- alternatively, polarizable force fields can result in state-specific embedding potentials
- repulsive short-range correction needed to avoid overpolarization
- continuum models solve sampling problem in addition

QM/QM Hybrid Methods:
An Introduction

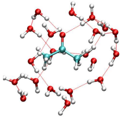
Methods for Environmental Effects



continuum models



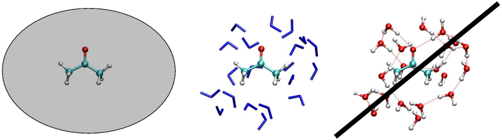
QM/MM methods



pure QM methods

Wish list for an “ideal” environmental model:

Methods for Environmental Effects

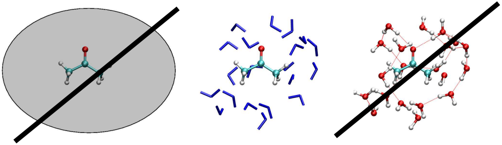


continuum models **QM/MM methods** **pure QM methods**

Wish list for an “ideal” environmental model:

- should be efficient (comparable to isolated molecule)

Methods for Environmental Effects

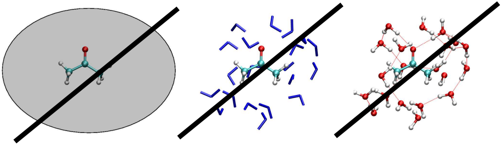


continuum models **QM/MM methods** **pure QM methods**

Wish list for an “ideal” environmental model:

- should be efficient (comparable to isolated molecule)
- should be able to model specific effects (atomistic structure)

Methods for Environmental Effects

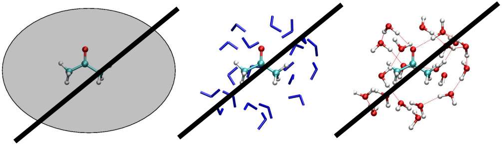


continuum models **QM/MM methods** **pure QM methods**

Wish list for an "ideal" environmental model:

- should be efficient (comparable to isolated molecule)
- should be able to model specific effects (atomistic structure)
- should be transferable (parameter-free)

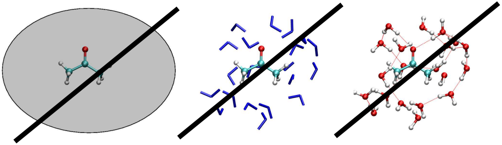
Methods for Environmental Effects



Wish list for an “ideal” environmental model:

- should be efficient (comparable to isolated molecule)
- should be able to model specific effects (atomistic structure)
- should be transferable (parameter-free)
- should focus on the embedded system

Methods for Environmental Effects



continuum models QM/MM methods pure QM methods

Wish list for an “ideal” environmental model:

- should be efficient (comparable to isolated molecule)
- should be able to model specific effects (atomistic structure)
- should be transferable (parameter-free)
- should focus on the embedded system

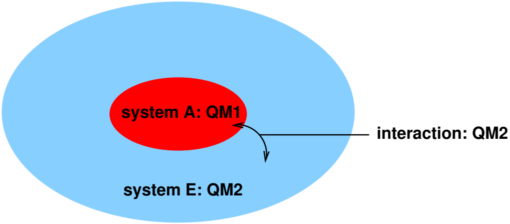
⇒ QM/QM embedding methods

QM/QM Hybrid Methods

Partitioning is possible at different levels:

- energy partitioning
- wavefunction partitioning
- density-matrix partitioning
- Green's function partitioning
- density partitioning

Simple QM/QM Hybrid Methods



- ONIOM: a generalization of subtractive QM/MM

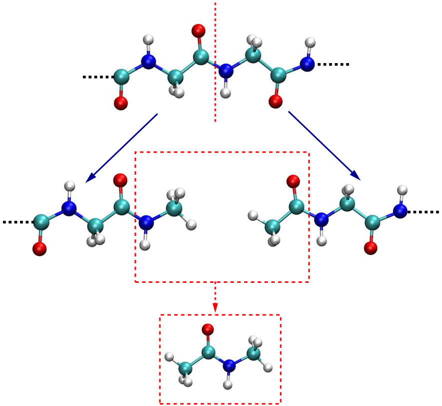
$$E_{\text{tot}}^{\text{QM/QM}} = E_{(A+E)}^{\text{QM2}} + (E_A^{\text{QM1}} - E_A^{\text{QM2}})$$

  interaction defined as

$$E_{A\leftrightarrow E}^{\text{QM2}} = E_{(A+E)}^{\text{QM2}} - E_A^{\text{QM2}} - E_E^{\text{QM2}}$$

- generalizations for many layers (QM1/QM2/MM/continuum) and electrostatic embedding
- properties require corresponding definitions

Fragmentation Methods



Molecular Fractionation with Conjugated Caps (MFCC)

- developed for proteins; partition into oligopeptides
- use model for neighboring fragments as caps
- sum up results for all capped fragments, subtract concaps
- no electrostatic embedding; better with larger caps

D.W. Zhang, J.Z.H. Zhang, *J. Chem. Phys.* **119** (2003), 3599.

Increment Methods

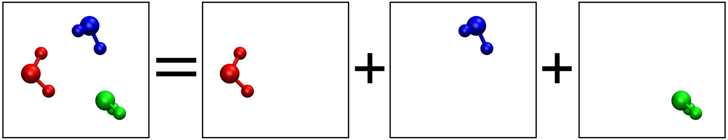
General idea:

- largest part of E_{tot} is due to energies of (isolated) subsystems
- interaction energy: mainly due to pair interactions
- general energy expression:

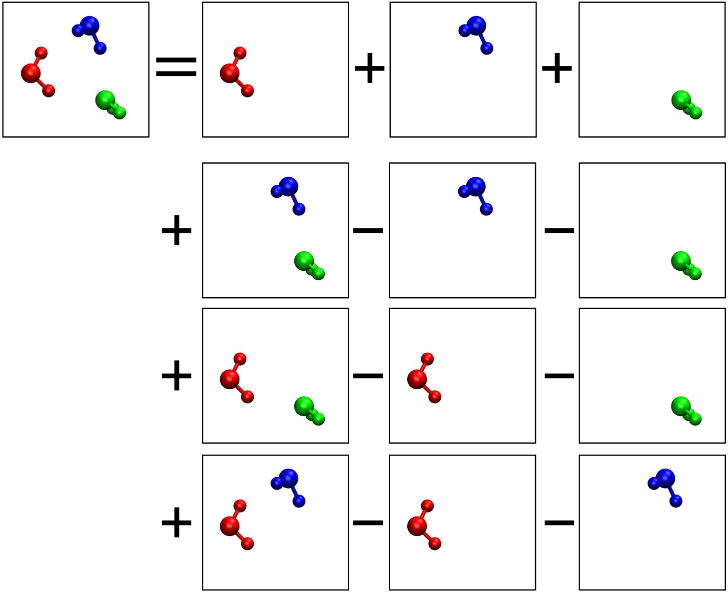
$$\begin{aligned}
E_{\mathrm{tot}} &= \sum_{I} E_I + \sum_{I<J} E_{I\leftrightarrow J} + \sum_{I<J<K} E_{I\leftrightarrow J\leftrightarrow K} + \dots \\
E_{I\leftrightarrow J} &= E_{IJ} - E_I - E_J \\
E_{I\leftrightarrow J\leftrightarrow K} &= E_{IJK} - E_I - E_J - E_K \\
&\quad -(E_{IJ} - E_I - E_J) - (E_{IK} - E_I - E_K) - (E_{JK} - E_J - E_K)
\end{aligned}$$

- nth-order increment method: exact for n subsystems

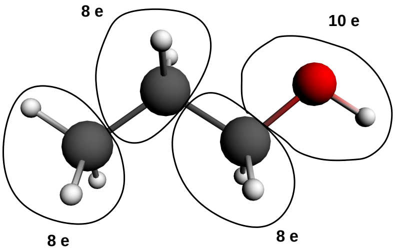
Increment Methods



Increment Methods



Fragment Molecular Orbital Method (1)

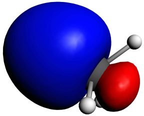


FMO-1

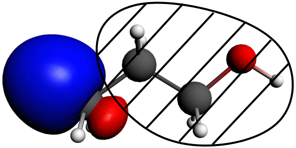
- fragmentation without separation of bond electron pairs
- fragment calculations under full electrostatic embedding
- no capping applied

K. Kitaura *et al.*, *Chem. Phys. Lett.* **313** (1999), 701.

Fragment Molecular Orbital Method (1)



generate localized MOs
for capped fragment

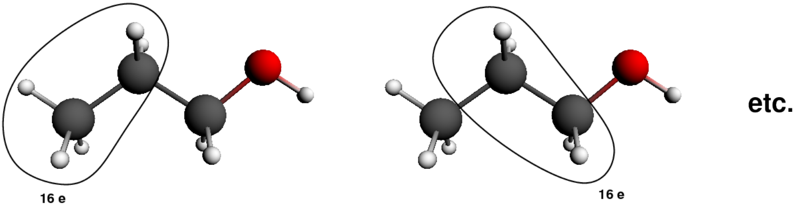


use these LMOs as basis
for "true" fragments

FMO-1

- fragment orbitals are expressed through localized orbitals obtained from reference compound (capped fragment)
- minimal basis is applied ⇒ bond electron pairs have no freedom to change compared to reference compound

Fragment Molecular Orbital Method (2)



FMO-2

- 2nd order increment method
- monomers and dimers are obtained with full electrostatic embedding potential, based on localized orbitals

⇒ FMO-2 implicitly contains higher-order terms

More about QM/QM tomorrow . . .