The Abdus Salam International Centre for Theoretical Physics

2549-4

ADGLASS Winter School on Advanced Molecular Dynamics Simulations

18 - 21 February 2013

Fundamentals of QM/MM Simulations

Paul Sherwood
STFC Daresbury Laboratory
United Kingdom

Science & Technology
Facilities Council

www.scitech.ac.uk

Fundamentals of QM/MM Simulations

Paul Sherwood
STFC Daresbury Laboratory

paul.sherwood@stfc.ac.uk

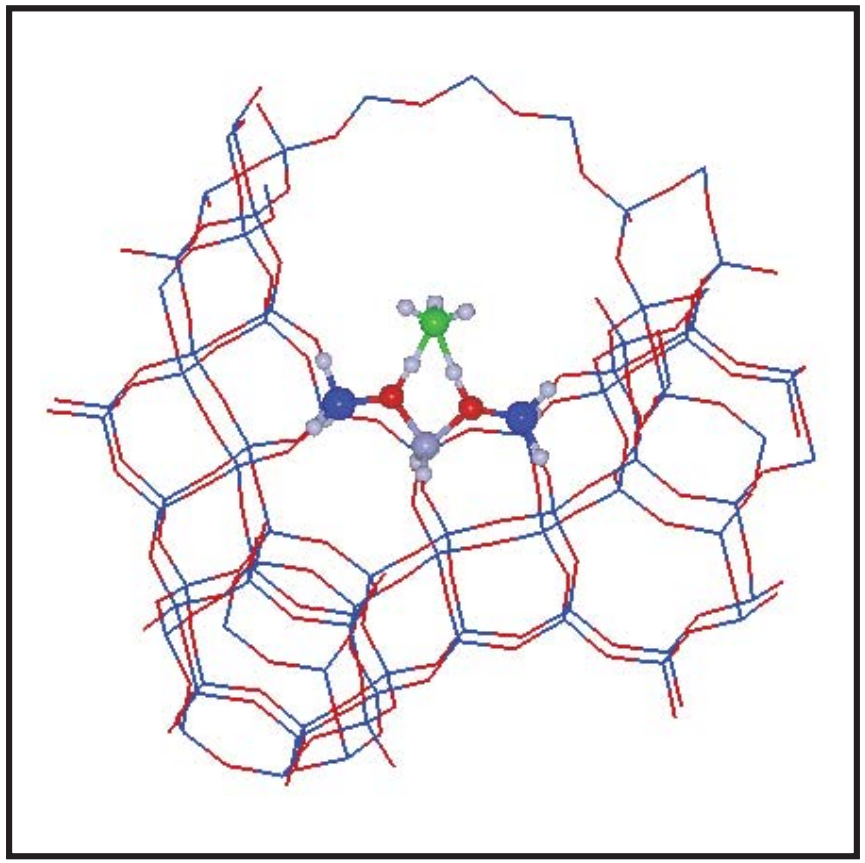
Outline

- Classification and principles of popular QM/MM schemes
  - Links atoms and boundary regions
  - Additive vs Subtractive
  - Electrostatic, Mechanical, Polarised
    - Cluster models for periodic systems

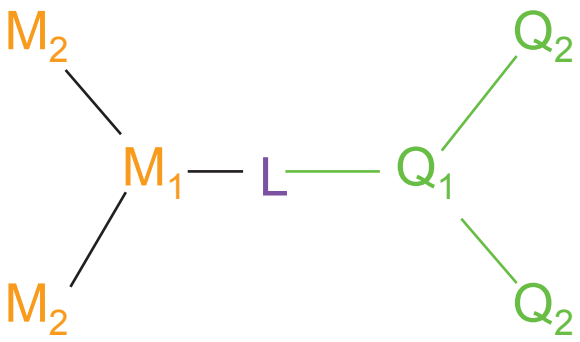
- Additional topics
  - Micro-iterative geometry optimisation
  - DL_FIND library
    - Optimisation, TS search, excited states
  - Periodicity and QM/MM
  - Adaptive QM/MM models

The QM/MM Modelling Approach

- Couple quantum mechanics and molecular mechanics approaches
- QM treatment of the active site
  - reacting centre
  - excited state processes (e.g. spectroscopy)
  - problem structures (e.g. complex transition metal centre)
- Classical MM treatment of environment
  - enzyme structure
  - zeolite framework
  - explicit solvent molecules
  - bulky organometallic ligands



Hybrid Computational Schemes



QM/MM Couplings

Unpolarised or "mechanical embedding"

Polarisation of QM region

"electrostatic embedding"

MM polarisation

shell model or dipole polarisabilities

Termination Scheme

Link Atoms, or Boundary zone

Chemical type (hydrogen atoms, pseudopotentials adjusted connection atoms, localised orbitals)

Charge perturbations (none, charge deletion, charge shift, selection of 1e integrals, double link atoms)

Total Energy Expression

Additive, Uncorrected

E(M,MM) + E(QL,QM) +E(QM/MM)

Additive, Boundary corrected

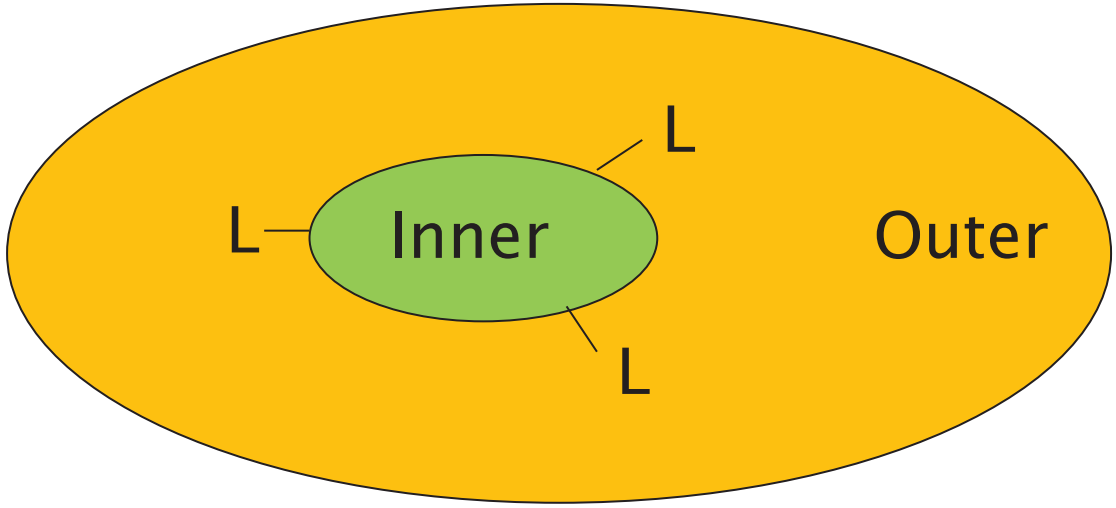
E(M,MM) + E(QL,QM) - E(L,MM)…+

Subtractive

E(MQ,MM) + E(QL,QM) - E(QL,MM)

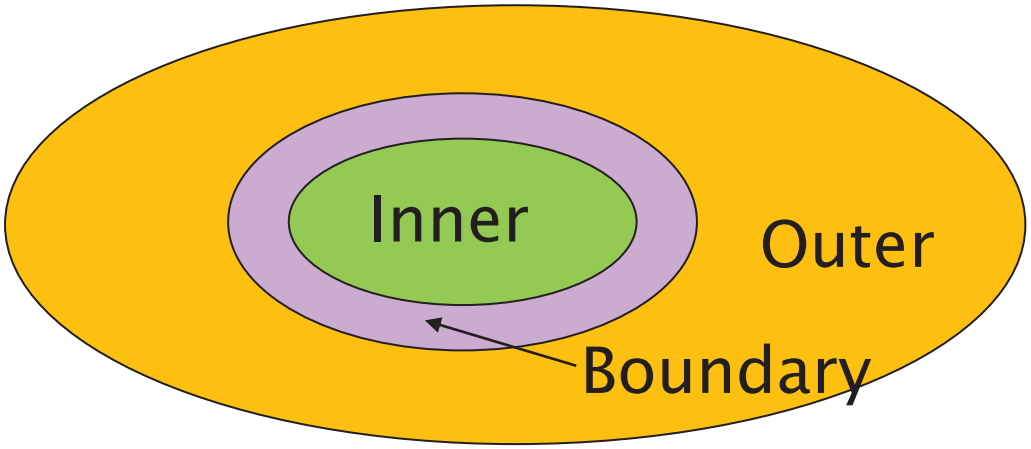
Link Atoms

- QM/MM with extra Link Atoms (L) to terminate broken covalent bonds



Boundary Regions

- Boundary region approaches introduce no new atoms, e.g. Solid State Systems with ionic character – Link atoms are inappropriate



  - Range of representations within QM code inclide modified *ab-initio* atom with model potential, Semi-empirical parameterisation, Frozen orbitals, Design atoms
  - Often associated with re-parameterised MM potentials

Additive Schemes

- Energy Expressions
  - Without link atom correction

    E(O,MM) + E(I,QM) + E(IO,QM/MM)
  - Link atom correction

    E(O,MM) + E(IL,QM) + E(IO,QM/MM) - E(L,MM)
  - Boundary methods

    E(OB,MM) + E(IB,QM) + E(IBO,QM/MM)
- Highly variable in implementation
  - QM/MM couplings,
  - QM termination etc

- Advantages
  - No requirement for forcefield for reacting centre
  - Can naturally build in electrostatic polarisation of QM region - effects of environment of excitations etc
- Disadvantages
  - Electrostatic coupling of the two regions, **E(IO,QM/MM)** is problematic with link atoms
  - Need for boundary atom parameterisation

Subtractive Schemes

- Energy Expression

  E(OI,MM) + E(IL,QM) - E(IL,MM)

  - includes link atom correction
  - can treat polarisation of both the MM and QM regions at the force-field level
- Termination
  - Any (provided a force field model for **IL** is available)
- Advantages
  - Potentially highly accurate and free from artefacts
  - Can also be used for QM/QM schemes (*e.g.* IMOMO, Morokuma *et al*)

- Disadvantages
  - Need for accurate forcefields (mismatch of QM and MM models can generate catastrophes on potential energy surface)
  - Usually no electrostatic influence on QM wavefunction included (e.g. QMPot), (but can be extended to electrostatic embedding: ONIOM-EE)

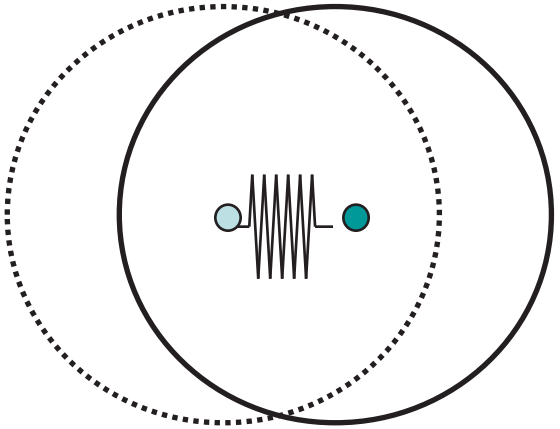
Forcefield Considerations

- Valence FFs
  - e.g. CHARMM, MM2, Dreiding
  - Small fractional charges, sometimes designed to reproduce electrostatic potential
  - Explicit bond , angle, dihedral terms, easy to deal with QM/MM double counting

- Shell Model
  - Do not usually require definition of covalent bonding network (dominated by 2-body terms)
  - Typically based on formal ionic charges
  - Include polarisability

Shell Model Force fields

- Typically used for ionic solids
- Leading terms are non-bonded
  - Electrostatics
    - often based on formal charges
    - polarisability of ions included by splitting total ion charge in
      - Core (often +ve) and Shell (-ve), modelling the valence electrons
      - Shell can shift in response to electrostatic forces, restoring forces from harmonic “spring”
  - van der Waals
    - sometimes compute using shell position
- Can also incorporate 3-body terms
  - some bond angles are preferred over others, introducing some covalent character



Core position

Shell position

Choice of MM Model

- Practical considerations
  - We must be able to remove selected forcefield terms from topology to avoid double counting in both QM and MM
    - handling of link atoms is easier for valence forcefields than for ionic ones
    - ionic forcefields require classification of atoms into layers and defining inter- and intra-layer parameters
  - Always need vdW parameters for interaction of MM atoms with QM
  - For mechanical embedding schemes also require atom partial charges for the QM region
  - Freedom from numerical noise (e.g. MM cutoffs) is important for transition states *etc.*

QM/MM Non-bonded Interactions

- Short-range forces (van der Waals)
  - Typically will follow MM conventions (pair potentials etc), sometimes reparameterisation is performed to reflect replacement of point charges interactions with QM/MM electrostatic terms.
- Electrostatic interactions:
  - Mechanical Embedding
    - *in vacuo* QM calculation coupled classically to MM via point charges at QM nuclear sites
  - Electrostatic Embedding
    - MM atoms appear as centres generating electrostatic contribution to QM Hamiltonian
  - Polarised Embedding
    - MM polarisability is coupled to QM charge density

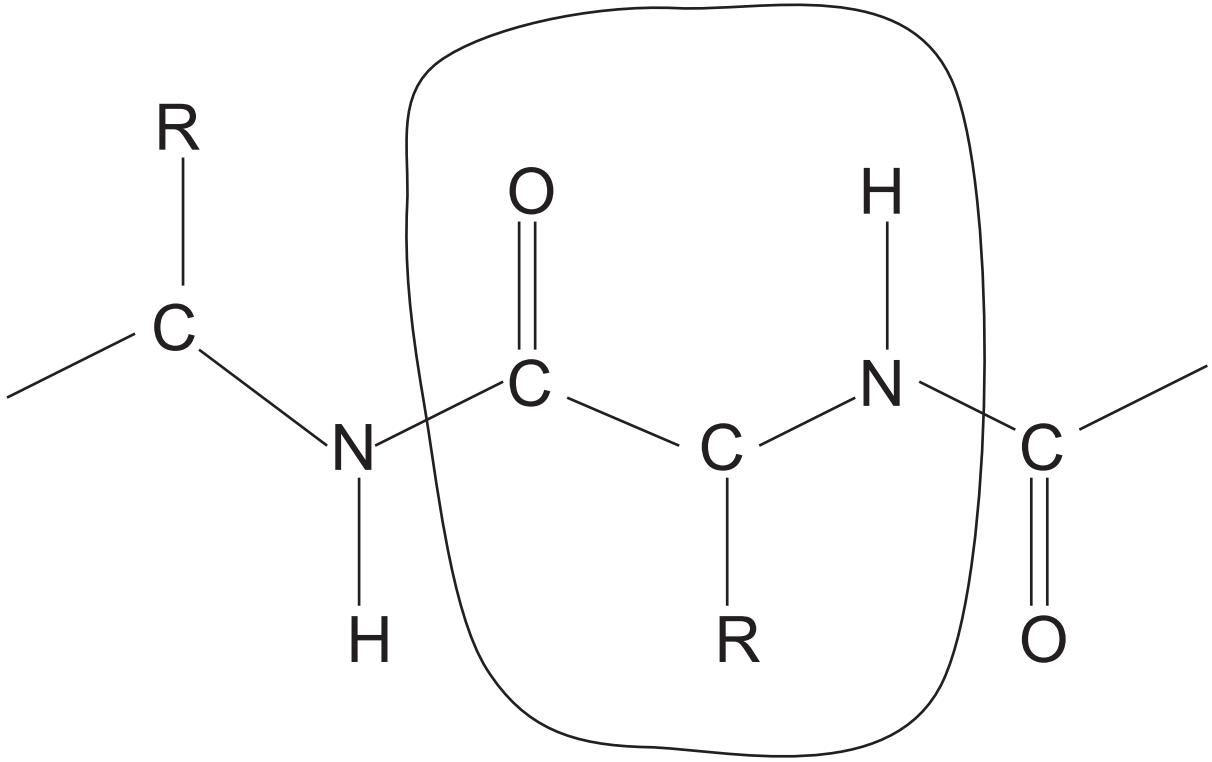
Mechanical Embedding

- Advantages
  - MM and QM energies are separable
    - separate MM relaxation, annealing etc possible
  - QM/MM terms can be integrated directly into the forcefield
  - No interactions between link atoms and MM centres
  - QM energies, gradient, Hessian are the same cost as gas phase
- Drawbacks
  - No model for polarisation of QM region
  - Electrostatic coupling requires atomic charges for QM atoms
    - generally these will be dependent on reaction coordinate
- Examples
    - IMOMM and ONIOM (Morokuma)
    - MNDO/MM (Bakowies and Thiel)

Electrostatic Embedding

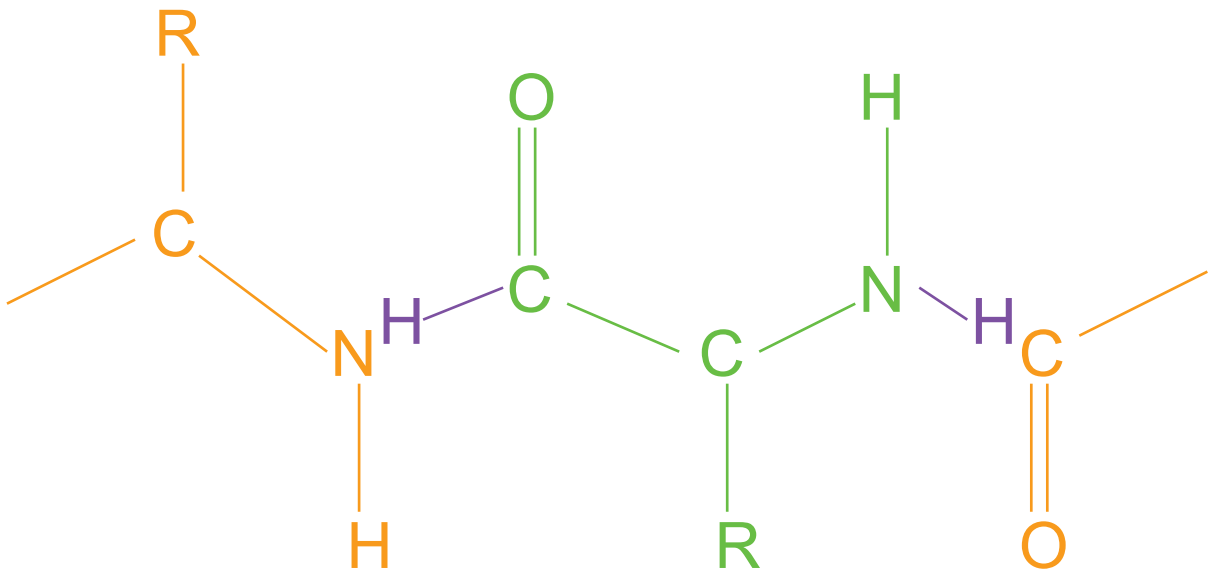
- (i) Assign MM Charges for pure MM system
  - Derived from empirical schemes (e.g. as part of forcefield)
  - Fitted to electrostatic potentials
  - Formal charges (e.g. shell model potentials)
  - Electronegativity equalisation (e.g. QEq)
- (ii) Delete MM charges on atoms in inner region
  - Attempt to ensure that MM "defect" + terminated QM region has
    - correct total charge
    - approximately correct dipole moment
- (iii) Insert charges on MM centres into QM Hamiltonian
  - Explicit point charges
  - Smeared point charges
  - Semi-empirical core interaction terms
  - Make adjustments to closest charges (deletion, shift etc)

Creation of neutral embedding site
(i) Neutral charge groups



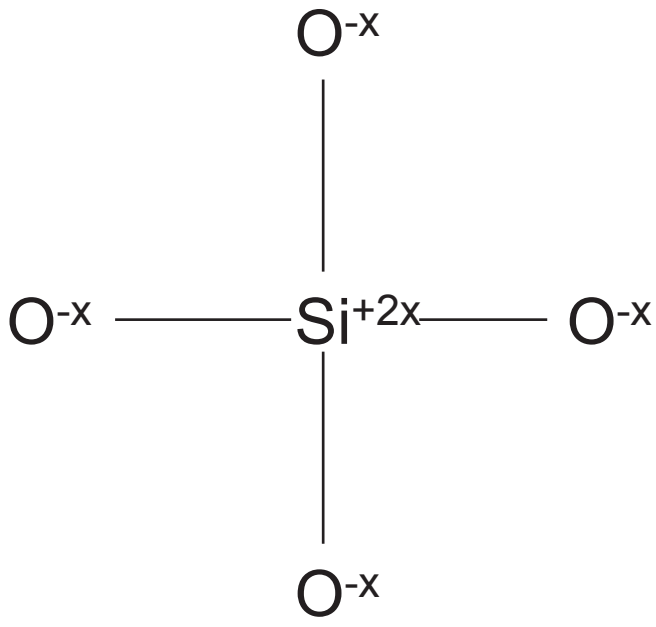
Deletion according to force-field neutral charge-group definitions

Creation of neutral embedding site
(i) Neutral charge groups



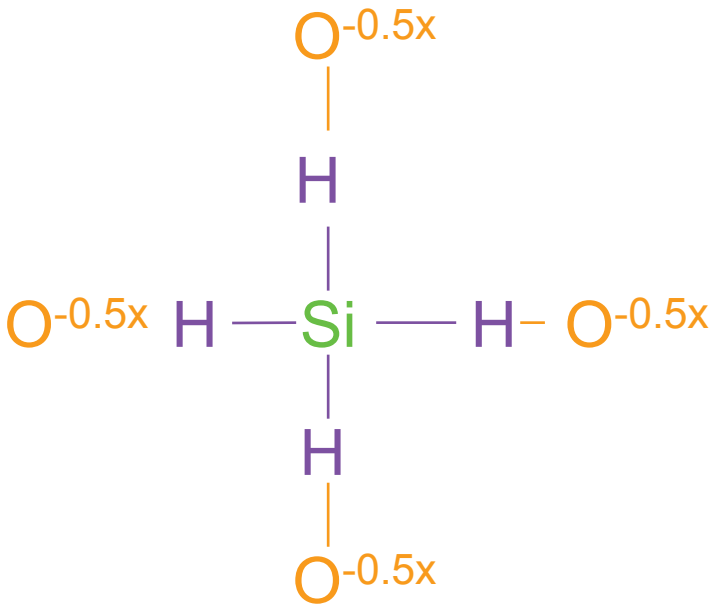
Total charge conserved, poor dipole moments

Creation of neutral embedding site
(ii) Polar forcefields

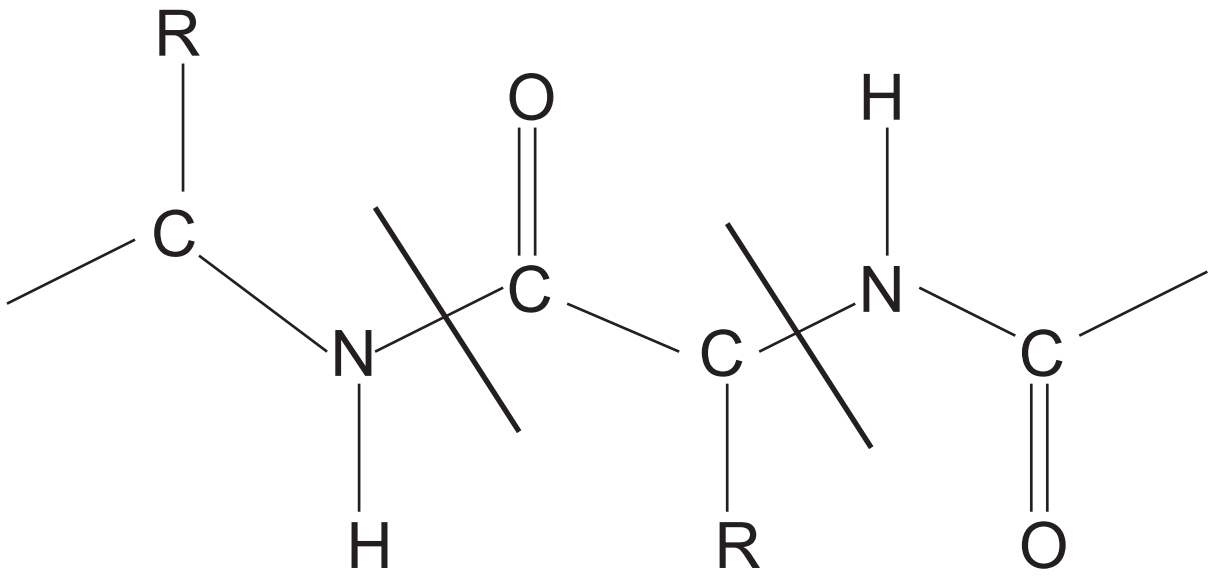


bond dipole models, e.g. for zeolites (Si +0.5x, O -0.5x)

Creation of neutral embedding site
(ii) Polar forcefields

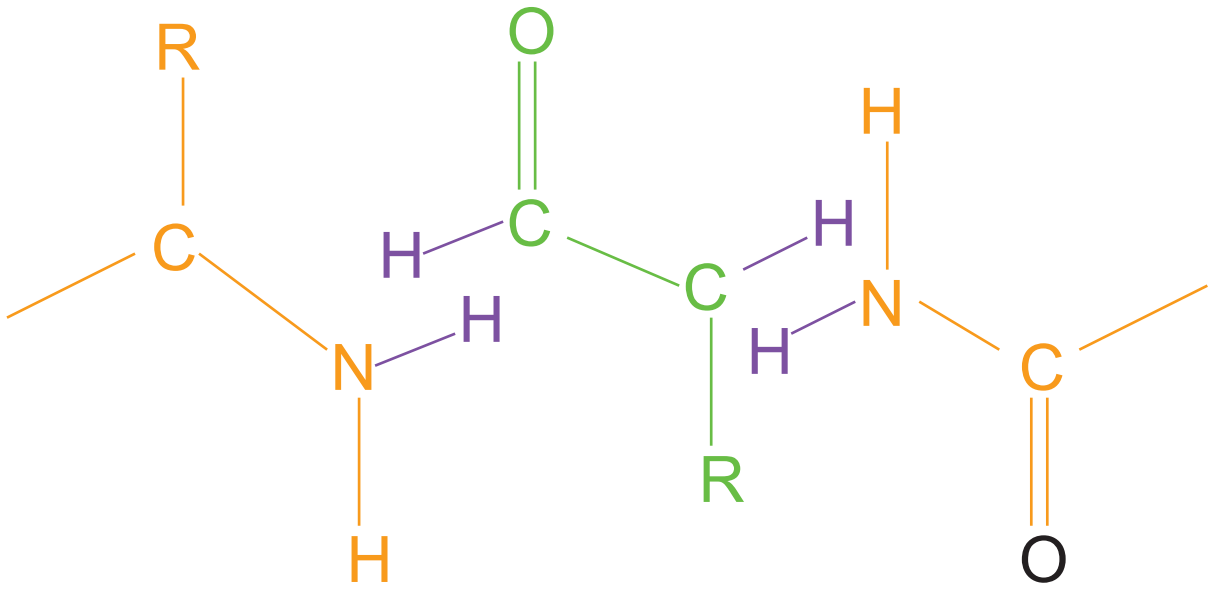


Creation of neutral embedding site
(iii) Double link atoms



Suggestion from Brooks (NIH) for general deletion (not on a force-field neutral charge-group boundary)

Creation of neutral embedding site
(iii) Double link atoms



All fragments are common chemical entities, automatic charge assignment is possible.

QM Termination Schemes

- Boundary schemes
  - Frozen Orbitals
    - Local SCF scheme (Rivail)
    - Generalised hybrid orbital (Gao)
    - ab-initio implementation (Friesner)
  - Pseudopotentials
    - Gaussian basis (Yang), Plane-wave (Rothlisberger), covEPE (Rosch), design atom (Zhang)
  - Adjusted connection atoms (Thiel)
    - semi-empirical mimic for attached methyl group
- Link atom schemes
  - Hydrogen atoms
  - Adjusted electronegativity
    - Hamiltonian shift operator
    - Pseudohalogen
  - Methyl groups

Localised Orbital Approaches (i)

- LSCF (Rivail et al)
  - Semi-empirical
  - Single orbital on QM boundary atom (pointing outwards) is frozen, based on calculation on a fragment (case-by-case set up)

- GHO (Generalised Hybrid Orbital, Gao et al.
  - Semi-empirical (being extended to *ab-initio*)
  - Single orbital (sp^3 hybrid) on MM boundary centre (pointing inwards)
  - Remaining 3 hybrid ("auxiliary orbitals") are populated with fixed density matrix elements to produce correct MM charge
  - Semi-empirical parameters of the MM centre are adjusted based on model compounds (expected to be transferable)

  see www.chem.umn.edu/groups/gao/gho.htm

Localised Orbital Approaches (ii)

- QSite implementation (Friesner et al)
  - *Ab-initio* implementation, in Jaguar package
  - Based on calculations on model fragments, using a particular basis set
  - Local orbitals include contribution from connected atoms (not just the QM and MM centres
  - Adjustment of MM parameters performed on a case-by-case basis, currently being used for protein systems

D.M. Philipp, R.A. Friesner, J. Comput. Chem. 20 (1999) 1468-94

Pseudopotentials

- EPE (elastic polarizing environment) uses
  - a shell model forcefield for oxide materials and
  - effective core potentials (pseudopotentials) on the boundary atoms. No basis functions or electrons are associated with these ionic sites

- covEPE (covalent EPE) method
  - uses a specially parameterised univalent 7-electron atom (based on fluorine) O*, placed at the first MM position
  - adjusts the classical charge on the centre to reproduce ESP, and shell model parameters for O* - Si interactions to match geometry

V.A. Nasluzov, E.A. Ivanova, A.M. Shor, G.N. Vayssilov, U. Birkenheur and N. Rösch, J. Phys. Chem. B, v107 (2003) 2228-2241.

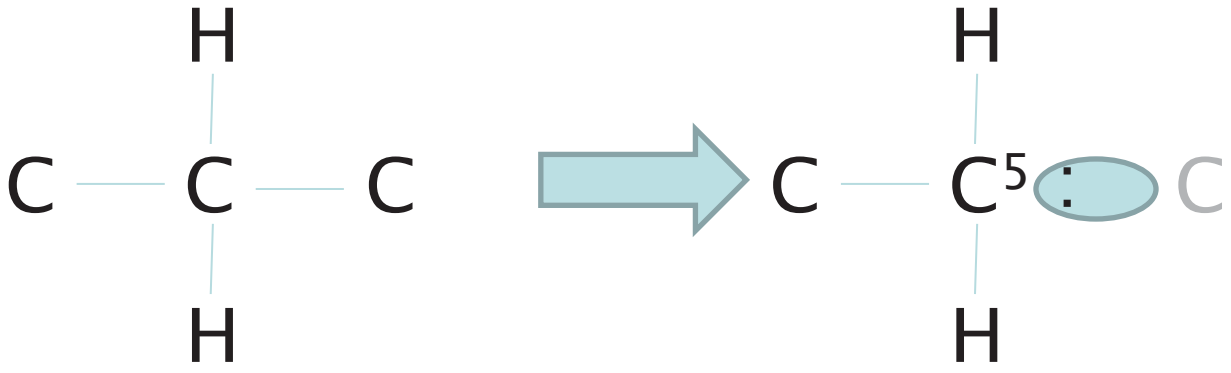
Adjusted Connection Atoms

- Semi-empirical parameterisation of boundary atom
  - Implemented in the MNDO package (Thiel el al)
  - No link atoms needed, a boundary atom is sited at the first MM centre
  - Typically boundary atom is C, parameterised to mimic electronic effects of CH_3

I. Antes, W. Thiel, *J. Phys. Chem. A* 103 (1999) 9290-95.

Zhang's Design Atom

- Boundary atom approach, for covalent type materials (e.g. biological simulations)
- Seeks to change the number of electrons e.g. 5 electron carbon, to replace broken bond with a lone pair



C Xiao, Y Zhang, *J. Chem. Phys.* **2007**, *127*, 124102

Link Atom Schemes

- Hydrogen atoms
  - Most common choice
  - Easily accomodated by regular QM codes
- Adjusted electronegativity
  - Pseudohalogen (Hyperchem)
  - Hamiltonian shift operator - acts only on elements involving basis functions on the link atom
- Methyl groups
  - Used in MOPS code (Cummins, Gready)
  - CH_3 has fixed geometry

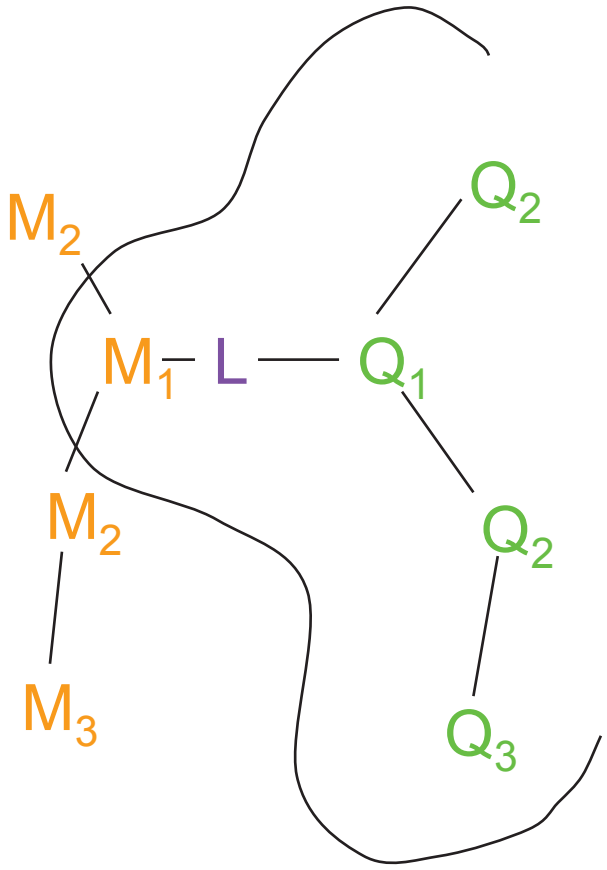
Positioning of link atoms

- Initial placement
  - Usually on terminated bond
- Unconstrained
  - Leads to additional degrees of freedom present in geometry optimisation and MD (no longer favoured)
    - e.g. CHARMM, QUEST
- Constrained
  - Need to take into account forces on link atoms,
    - shared internal coordinate definitions (IMOMM)
    - chain-rule differentiation (QM/Pot, ChemShell)

$$\frac{dE}{dx_{M_1}} = \frac{\partial E}{\partial x_{M_1}} + \frac{\partial E}{\partial x_L} \cdot \frac{\partial x_L}{\partial x_{M_1}}$$

    - use constraint capabilities within optimiser (CHARMM lone pair feature)

Boundary Charge Adjustments



- Some of the classical centres will lie close to link atom (L) or for boundary methods, modified centre Q_1
- Artefacts can result if charge at the M_1 centre is included in Hamiltonian
- Many adjustment schemes have been suggested
  - Adjustments to polarising field can be made independently from specification of MM…MM interactions
  - Similar adjustments may be needed if M_1 is classified as a boundary atom, depending on M_1 treatment.

Boundary Adjustments (i)
Selective deletion of 1e integrals

- L1: Delete integrals for which basis functions i or j are sited on the link atom L

$$V_{ij}^{A} = \left\langle \phi_i(l) \right| \frac{-q_A}{r_{lA}} \left| \phi_j(l) \right\rangle$$

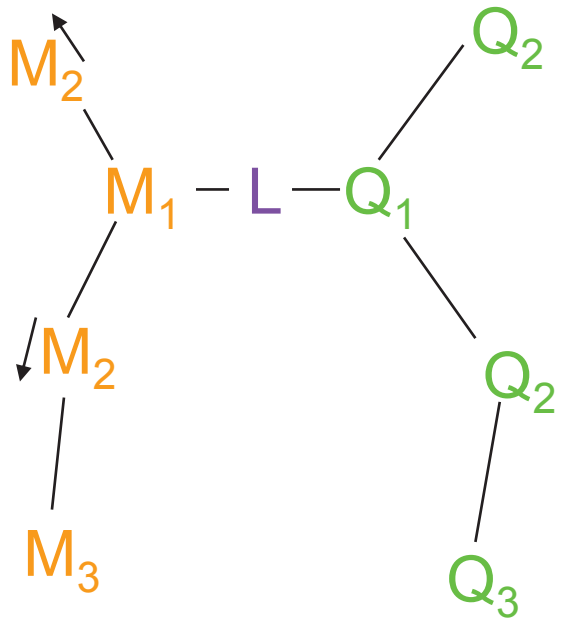
  - found to be effective for semi-empirical wavefunctions
  - difference in potential acting on nearby basis functions causes unphysical polarisation for ab-initio QM models
- L3: Delete integrals for which basis functions i and j are cited on the link atom and q_A is the neighbouring MM atom (M1)
  - less consistent results observed in practice †

† Classification from Antes and Thiel, in Combined Quantum Mechanical and Molecular Mechanical Methods, J. Gao and M. Thompson, eds. ACS Symp. Ser., Washington DC, 1998.

Boundary Adjustments (ii)
Deletion of first neutral charge group

- L2 - Exclude charges on all atoms in the neutral group containing M1
  - Maintains correct MM charge
    - leading error is the missing dipole moment of the first charge group
  - Generally reliable
    - free from artefacts arising from close contacts
  - Limitations
    - only applicable in neutral group case (e.g. AMBER, CHARMM)
    - neutral groups are highly forcefield dependent
    - problematic if a charge group needs to be split
  - Application
    - biomolecular systems

Boundary adjustments (iii)
Charge shift



- Delete charge on M1
- Add an equal fraction of q(M1) to all atoms M2
- Add correcting dipole to M2 sites (implemented as a pair of charges)
  - charge and dipole of classical system preserved
  - Leading sources of residual error is that
    Q---L dipole moment is not equivalent to
    Q------M

Can be combined with GHO approach to give Redistributed Charge and Dipole Scheme (Lin and Truhlar *J Phys Chem A* **2005**, *109*, 3991)

Boundary adjustments (iv)
Gaussian Blur

- Delocalise point charge using Gaussian shape function
  - Large Gaussian width : electrostatic coupling disappears
  - Narrow Gaussian width : recover point charge behaviour
  - Intermediate values
    - short range interactions are attenuated
    - long range electrostatics are preserved
- Importance of balance - apply to entire MM system or to first neutral group
- Particularly valuable for double-link atom scheme where MM link atom charge lies within QM molecular envelope
- Available in GAMESS-UK/CHARMM implemention

Electrostatic Embedding Summary

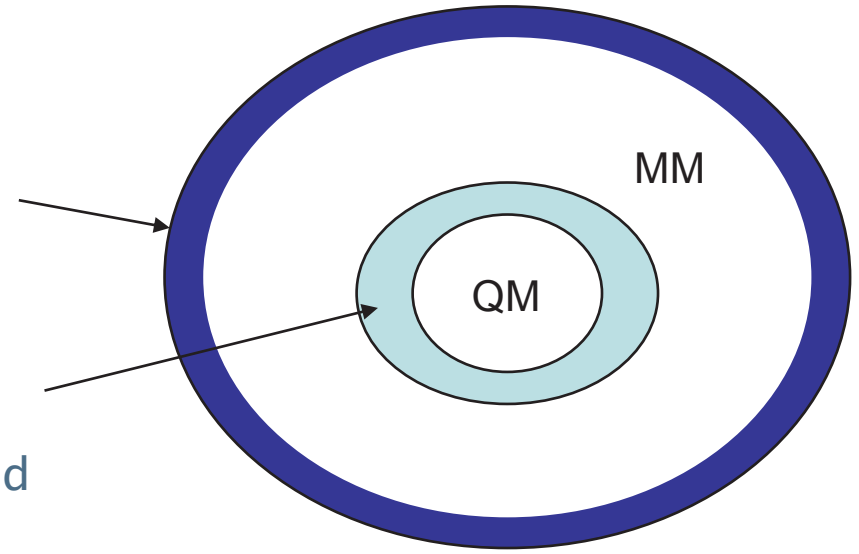
- Advantages
  - Capable of treating changes in charge density of QM
    - important for solvation energies etc
  - No need for a charge model of QM region
    - can readily model reactions that involve charge separation
- Drawbacks
  - Charges must provide a reliable model of electrostatics
    - reparameterisation may be needed for some forcefields
  - Danger of spurious interactions between link atoms and charges
  - QM evaluation needed to obtain accurate MM forces
  - QM energy, gradient, Hessian are more costly than gas phase QM

Polarised embedding schemes

- Incorporate polarisation of classical region
  - most appropriate used when the forcefield itself is based on explicit polarisability
  - back-coupling of polarised charge density to QM calculation can sometimes be omitted
- Approaches
  - Iterative solution of dipole polarisabilities
  - Direct Reaction Field Hamiltonian (van Duijnen, de Vries)
    - solution of coupled polarisabilities using relay matrix
    - possibility of including 2-electron dispersion terms
    - implemented in HONDO and GAMESS-UK
  - Shell model-based schemes
    - atomic charge is split into core and valence electron shell, connected by a harmonic spring
    - e.g. ChemShell solid-state embedding scheme

Solid-state Embedding Scheme

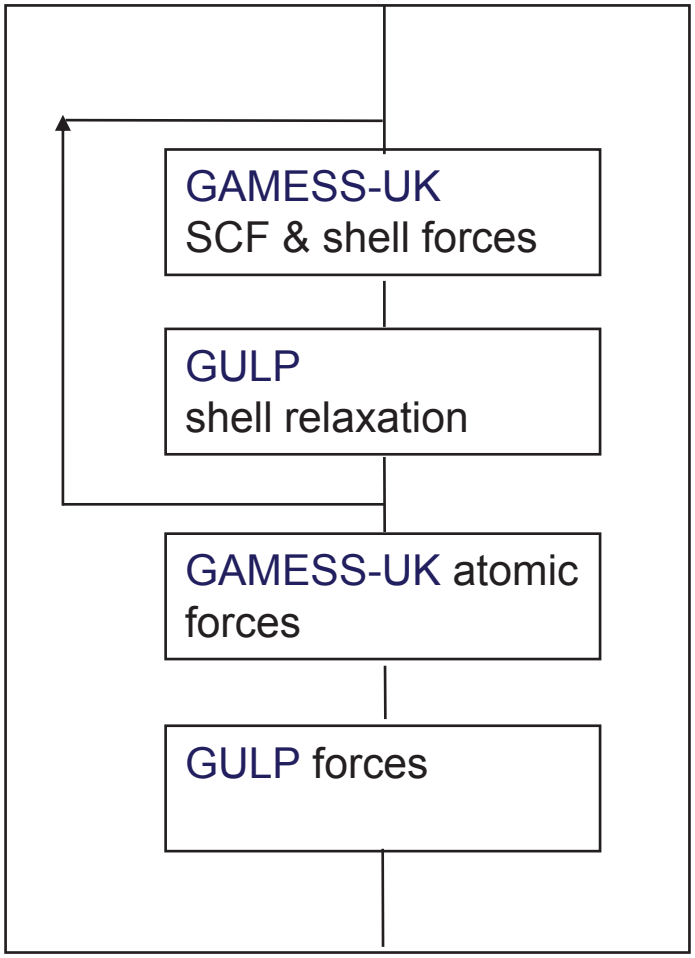
- Classical cluster termination
  - Base model on finite MM cluster
  - QM region sees fitted correction charges at outer boundary
- QM region termination
  - Ionic pseudopotentials (e.g. Zn2+, O2-) associated with atoms in the boundary region
- Forcefield
  - Shell model polarisation
  - Classical estimate of long-range dielectric effects (Mott/Littleton)
- Energy Expression
  - Corrections for boundaries incorporated in parameterisation



- Advantages
  - suitable for ionic materials
- Disadvantages
  - require specialised pseudopotentials
- Applications
  - metal oxide surfaces

Solid-state Embedding – Microiterations

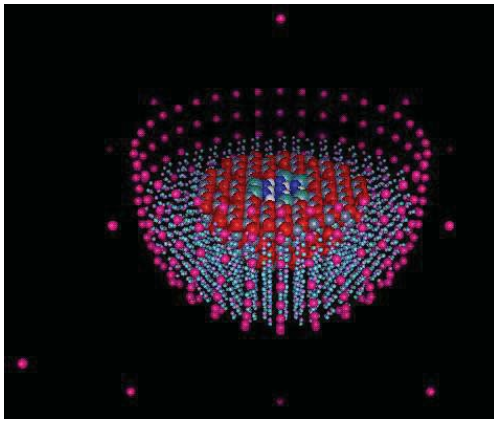
- ChemShell implementation is based on shell model code GULP, (Julian Gale)
- Both shell and core positions appear as point charges in QM code (GAMESS-UK)
- Self-consistent coupling of shell relaxation
  - compute electrostatic forces on shells in GAMESS-UK
  - relax shell positions in GULP

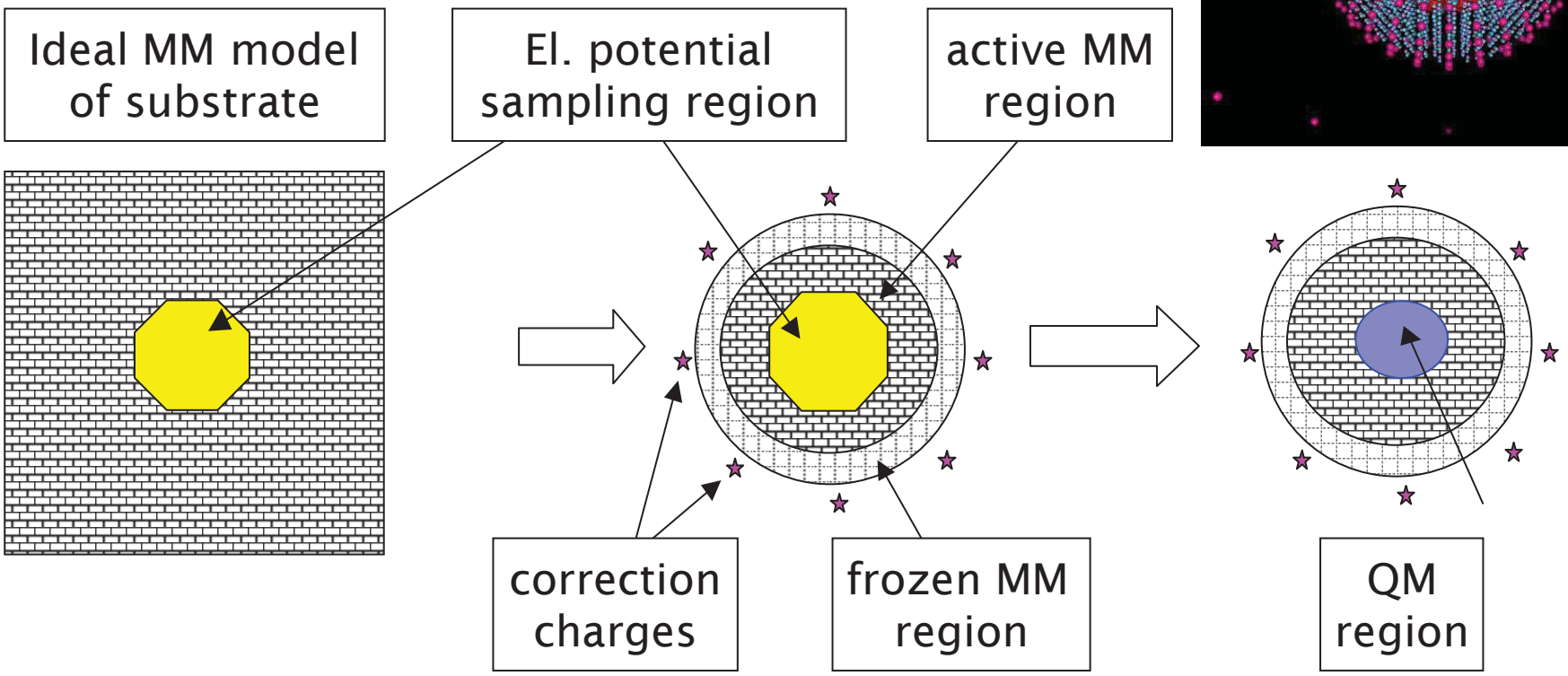


A. A. Sokol, S. T. Bromley, S. A. French, C. R. A. Catlow and P. Sherwood, *Int. J. Quantum. Chem*, 2004, 99, 695

Solid State Embedding
Fitted MM Boundary Correction

- electrostatic potential for periodic MM system or large nanoparticle
- cut out (hemi-) spherical cluster
- correction charges reproduce the potential





Polarised Embedding Schemes Summary

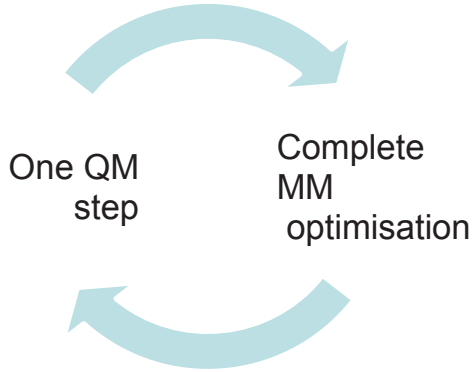
- Advantages
  - More accurate treatment of solvation effects
  - Allows coupling to systems where the best forcefields are based on polarisation (e.g. shell model potentials for metal oxide systems)

- Drawbacks
  - Additional cost
    - solution of coupled polarisabilities
    - some schemes will require additional SCF iterations
  - Requirement for polarised force-field
  - Danger of electrostatic instabilities close to boundaries
    - difficult to apply reliably when using link atoms

Defining the QM Region

- Things to consider
  - "Charge conservation" .. Will the QM cluster you are going to use have the same charge as the MM atoms it will replace?
  - Size of QM cluster.. If the cluster is too small there will be effects from the boundaries (a C-H bond is not the same as C-C)
  - Charge can't flow across QM/MM boundary - so think about charge transfer effects
  - Availability of a suitable method for terminating the QM cluster
    - If using hybrid orbitals: are they available for the type of bond in question?
    - Suitability of link atoms
      - replacing a C-C bond with C-H is usually OK (both are low polarity bonds)
      - replacing C-O or C-N with C-H less satisfactory.

Exploiting QM/MM capabilities: Micro-iterative QM/MM optimisation



- Electrostatic embedding: ESP charges calculated on the fly, fitting potential at MM sites
- Optimisation effort becomes more or less independent of the MM system size
- Saves a factor of 2–10 in CPU time

Kästner, S. Thiel, Senn, Sherwood, W. Thiel,
J. Chem. Theory Comput. **3**, 1064-1072, (2007).

DL-FIND

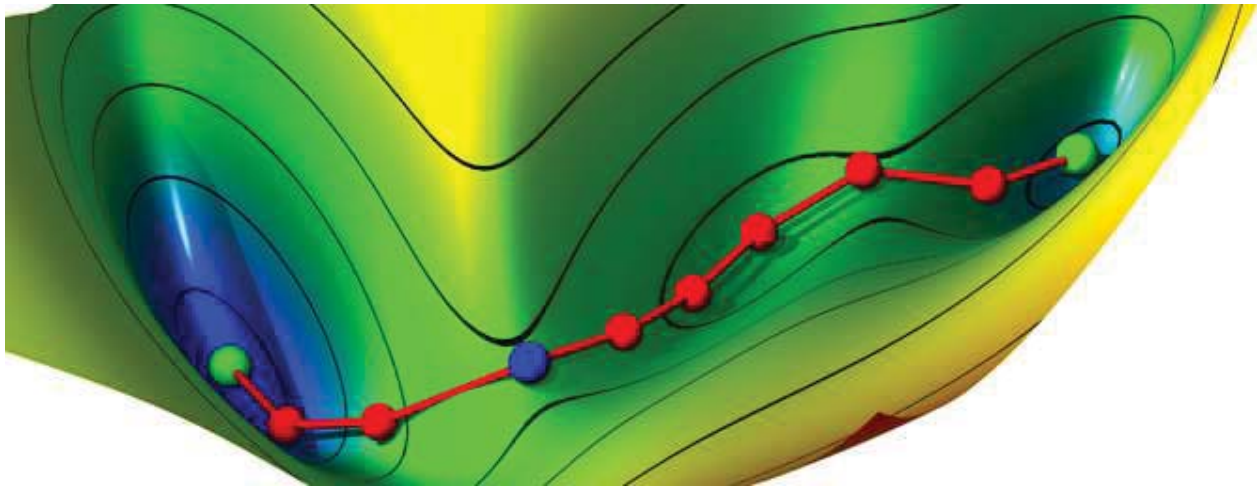
- For transition states in complex systems, offers
  - A selection of coordinate systems
    - Cartesians (including frozen atoms and components)
    - Internals (including all constraints):
      - DLC (delocalised internal coordinates)
      - DLC-TC (total connection)
      - HDLC (hybrid DLC+cartesian) and HDLC-TC
  - and a selection of search methods
    - Standard hessian based methods (e.g. P-RFO)
    - Nudged Elastic Band
    - Dimer method
    - Growing String method
  - Modular architecture aims to make all these combinations available
  - Support for micro-iterations underway
  - Open Source (L-GPL)

http://ccpforge.cse.rl.ac.uk/projects/dl-find/

Science & Technology Facilities Council
Daresbury Laboratory

DL-FIND Nudged Elastic Band

- Multiple images, connected by “springs”
- Converges to the minimum-energy path
- Climbing image: transition state
- Costly, but can cover difficult reactions
- Particularly well suited to massive parallelism (under development)

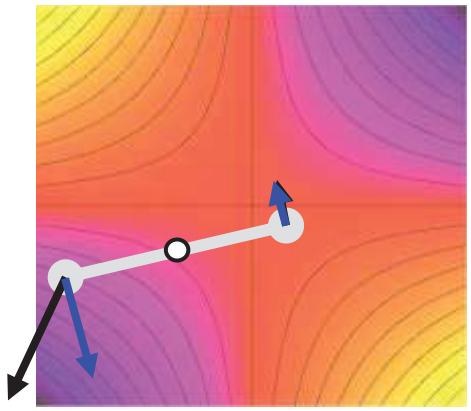


DL-Find - Dimer Method

- Converges to first-order saddle points without calculation of a Hessian
- Suitable for large systems
- TS search is converted in two minimization problems
- Available in DL-FIND in Cartesian, redundant internal, and HDLC coordinates

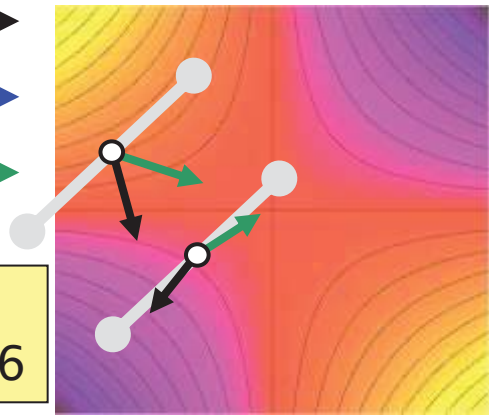
Henkelman, Jónsson, *J. Chem. Phys.* **111**, 7010 (1999)

rotation



translation

true force
rotational force
translational force



Enhancements to reduce optimisation cost by ~50%
Kaestner and Sherwood, *J Chem Phys*, 128 (2008) 014106

Excited State Developments I – DL-Find

- Conical intersection - the geometry where two electronic states are degenerate. Likely point of radiationless decay.
- Algorithms to find the lowest-energy point on a conical intersection:
  - Penalty function
  - Gradient projection method
  - Lagrange-Newton method

T Keal, A. Koslowski & W. Thiel. Comparison of algorithms for conical intersection optimisation using semiempirical methods. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)* **118**, 837-844 (2007).

Periodicity and QM/MM

- Periodic QM/MM treatments must incorporate
  - Ewald sum of MM charges for polarizing potential
  - QM…..QM image interactions
- When using standard QM methods, periodic Ewald calculations from MM code can be used (with point charge model for the QM region) and subtract double-counted terms
  - Nam, Gao and York, *J. Chem. Theory Comput.* **2005**, *1*, 2.

  or use faster Particle Mesh Ewald scheme, (in AMBER)
  - Walker, Crowley and Case, *J Comp Chem* **2008**, *29*, 1019
- Different approaches are available for QM methods which deal with charged density on real-space grid, e.g. CP2K
  - Laino, Mohamed, Laio, and Parrinello, *J. Chem. Theory Comput.* **2006**, *2*, 1370–1378.

  or Siesta
  - Crespo, Scherlis, Martí, Ordejón, Roitberg and Estrin. *J. Phys. Chem. B* **2003**, *107*, 13728

Adaptive QM/MM Schemes (i)

- Change of the QM region during the simulation
  - Potential for discontinuity in energy and forces
  - Generally based on principle that forces on atoms in a buffer region are *interpolated* between QM values and MM values, depending on the distance from QM zone
- Rode's "Hot Spot" method
  - Kerdcharoen, Liedl, and Rode, *Chem Phys* **1996**, *211*, 313.
- ONIOM-XS
  - Kerdcharoen, and Morokuma, *Chem Phys Lett* **2002**, *355*, 257.
- LOTF schemes (MM with on-the-fly parameterisation)
  - Csanyi, Albaret, Payne, De Vita, *Phys. Rev. Lett.* **2004**, *93*, 175503.

Adaptive QM/MM Schemes (ii)

- Truhlar's schemes define a conserved potential energy by performing multiple QM/MM calculations (permuting boundary molecules between QM and MM zones) with geometry dependent weights to apply interpolation
- Number of possible contributions is 2^N (where N molecules in the boundary zone)
- Schemes linear in N are also possible, weighting functions are quite complex

  Heyden, Lin, and Truhlar, *J. Phys. Chem. B,* **2007**, *111*, 2231.

  Bulo, Ensing, Sikkema, and Visscher, *J. Chem. Theory Comput.* **2009**, *5*, 2212.

  Takenaka, Kitamura, Koyano and Nagaoka, M. *Chem Phys Lett* **2012**, *524*, 56.
- Much more later from Gabor

Summary

- We have reviewed QM/MM approaches
  - Links atoms vs boundary regions
  - Additive vs Subtractive
  - Electrostatic vs Mechanical
    - Details of Electrostatic embedding
  - A polarised QM/MM model based on shell model FF + boundary atoms, fitted electrostatic corrections for cluster models of periodic systems
- Additional topics
  - Micro-iterative geometry optimisation
  - DL_FIND library
    - Optimisation TS search, excited states
  - Periodicity and QM/MM
  - Introduction to Adaptive Schemes

Acknowledgements

- Development of ChemShell has been a collaboration with the groups of Richard Catlow and Walter Thiel

- This talk includes contributions from many people including
  - Early QM/MM work at Daresbury:
    - Alex de Vries
  - ChemShell, and DL_FIND, including optimisation and microiterative schemes
    - Johannes Kaestner, Thomas Keal and Alex Turner
  - Solid-state embedding scheme
    - Alexey Sokol, Sam French, Stefan Bromley
- Funding – EPSRC, EU, Shell SIOP Amsterdam