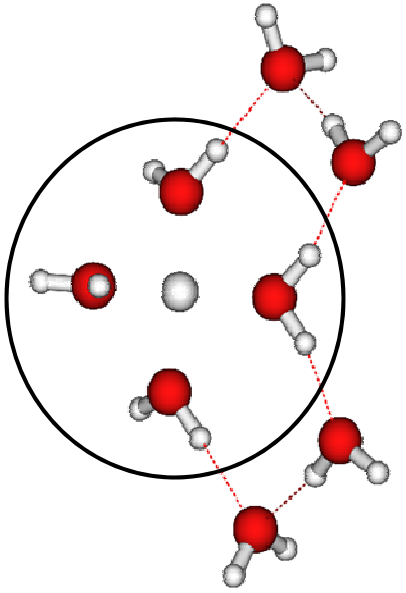
Methods and Code Integration

Ryan M. Olson
University of Minnesota

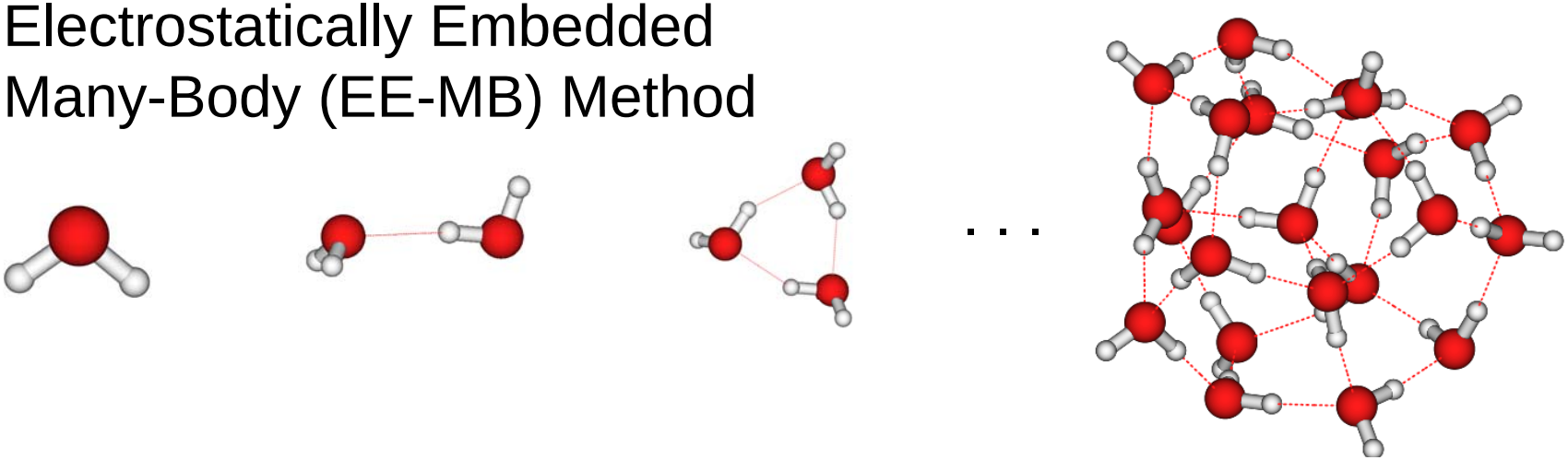
UNIVERSITY OF MINNESOTA
Pacific Northwest National Laboratory
U.S. Department of Energy
Office of Science
U.S. DEPARTMENT OF ENERGY

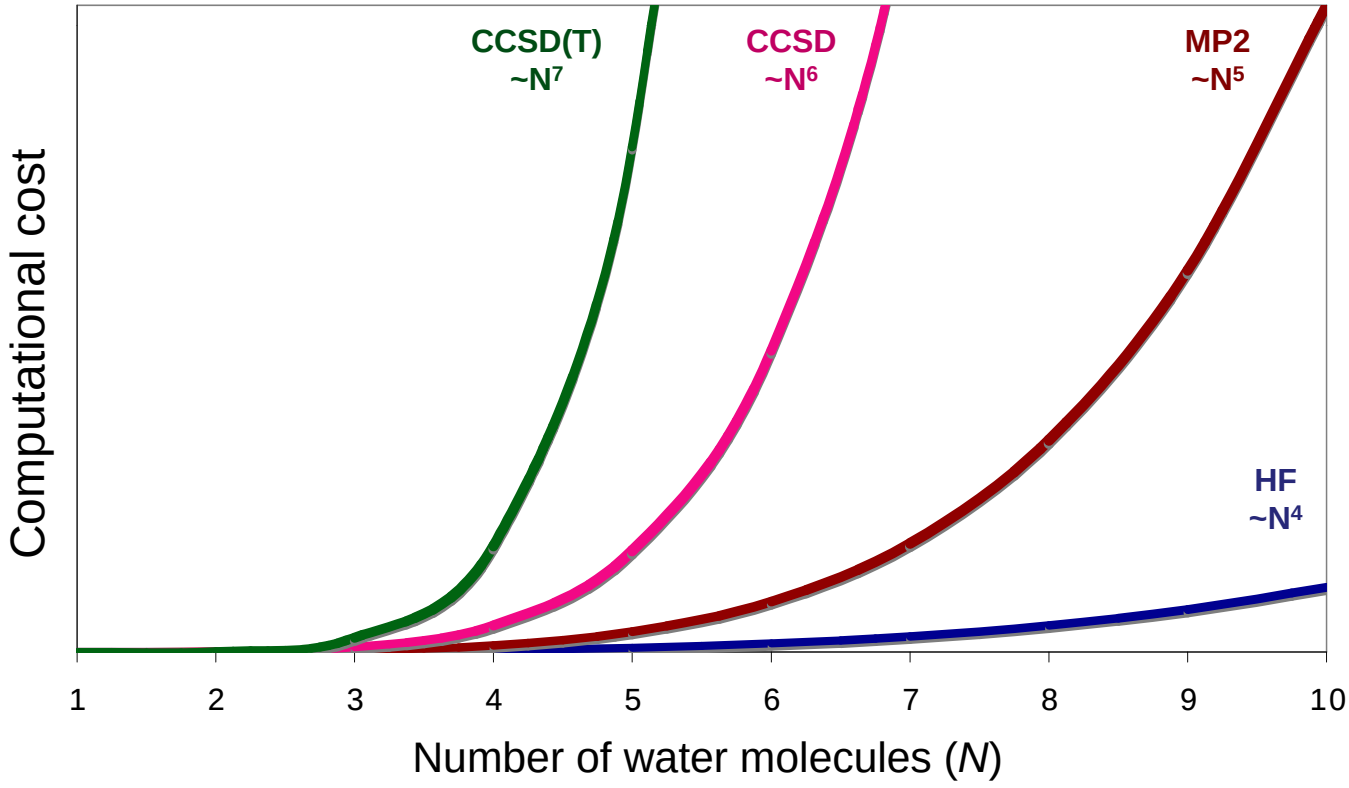
Overview

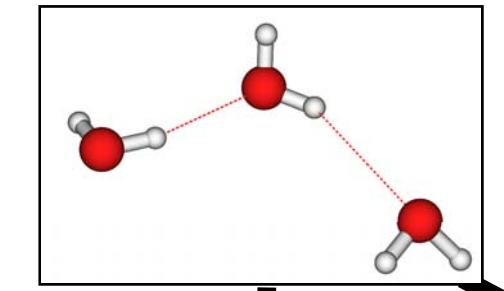
- Methods
  - Electrostatically Embedded Many-Body Method
  - Adaptive Partitioning
  - Configurational-Biased Grand Canonical Monte Carlo
- Applications
- Integration



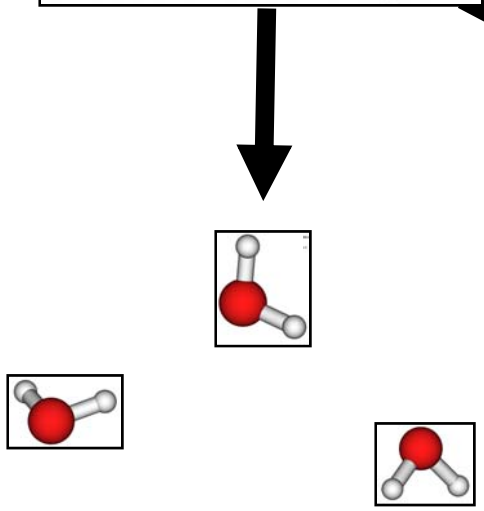
Electrostatically Embedded Many-Body (EE-MB) Method

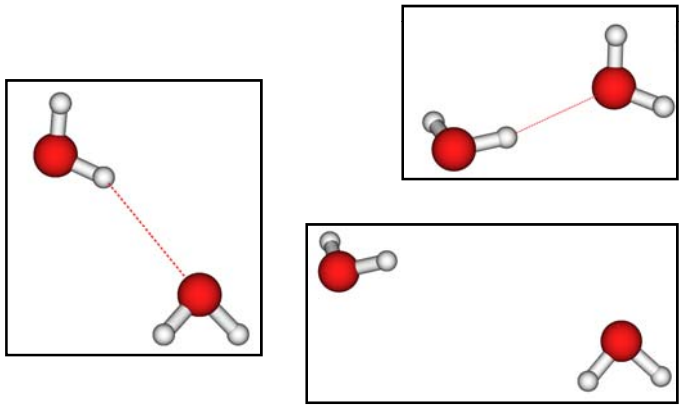






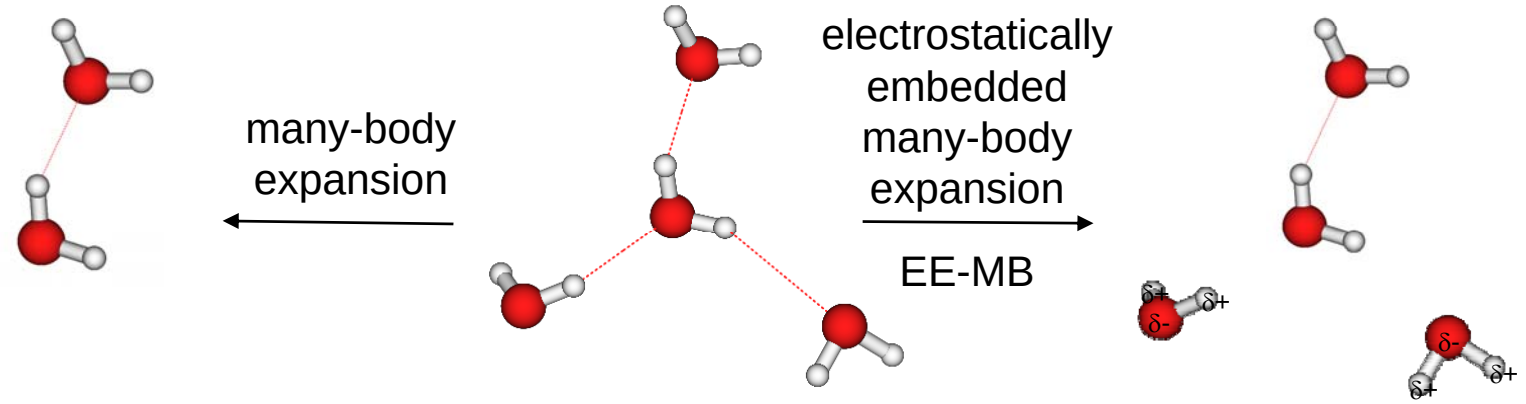
$$V = V_1 + V_2 + V_3 + \cdots + V_N$$

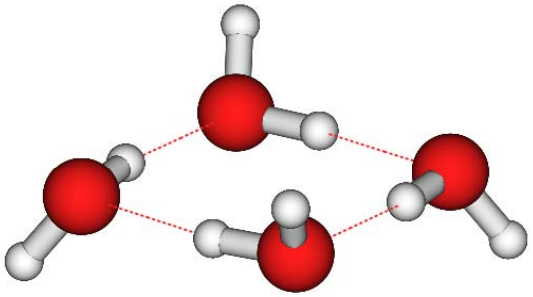




$$V_1 = \sum_{i}^{N} E_i$$

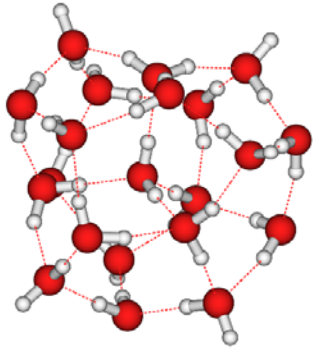
$$V_2 = \sum_{i<j}^{N} (E_{ij} - E_i - E_j)$$



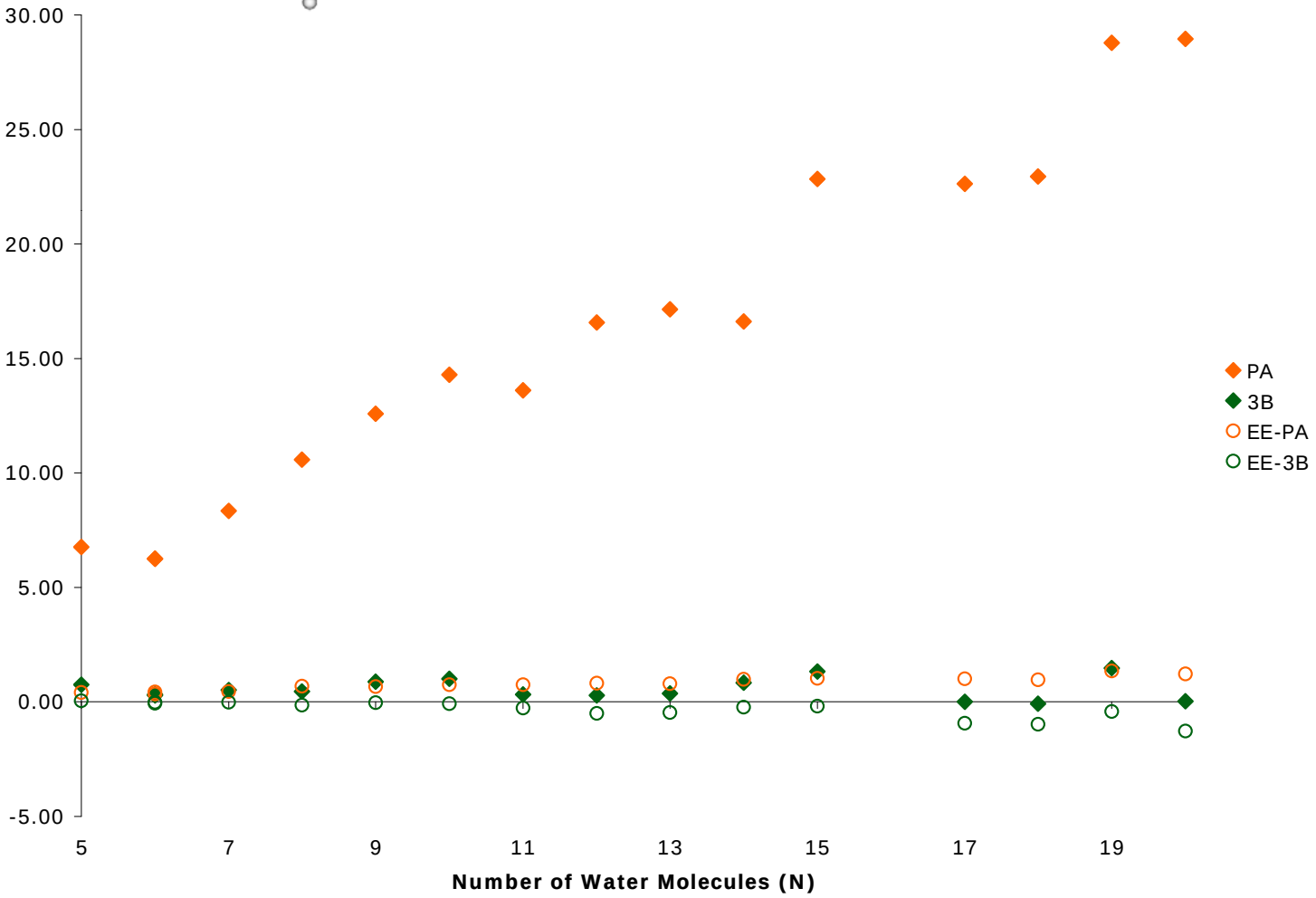


Method	Energy (kcal/mol)	%Error
PA	22.27	22%
EE-PA	28.08	2%
3B	28.14	2%
EE-3B	28.63	0.1%
E_{bind}	**28.65**	

MP2 with aug-cc-pVTZ on O and cc-pVTZ on H



Method	E_{bind}	ΔE	%Error
EE-PA	206.61	2.97	1.5%
EE-3B	204.02	0.38	0.2%
Full Cluster	203.64		



Conclusions

- Use of embedding charges reduces the error in the many-body approximation by a factor of 10.
- For moderately-sized systems, calculation of the full Hartree-Fock energy and a MB or EE-MB expansion of the correlation energy can give errors on the order of 0.2% of the net binding energy.
- The EE-MB methods have analytic gradients and analytic Hessians
  - - Geometry optimizations
  - - Frequency calculations
  - - Molecular Simulations

Conclusions

- And …
  - FAST
  - “Trivially” Parallel
- If you are interested:
  - Correlation energy
  - Cutoffs
  - Parallelism

N	Time* (min)	N_{frag}	Total time (min)
21	41168	1	41168
3	4.5	1330	5985
2	2	210	420
1	<1	21	21

Calc	Terms	Time
Full	$V_1 + V_2 + \cdots + V_N$	~ 28.5 days
3B	$V_1 + V_2 + V_3$	~ 4.5 days
PA	$V_1 + V_2$	~ 8 hrs

* Times are for an MP2 calculation with aug-cc-pVTZ on O and cc-pVTZ on H

Mission Accomplished!

UNIVERSITY OF MINNESOTA

Pacific Northwest National Laboratory
U.S. Department of Energy

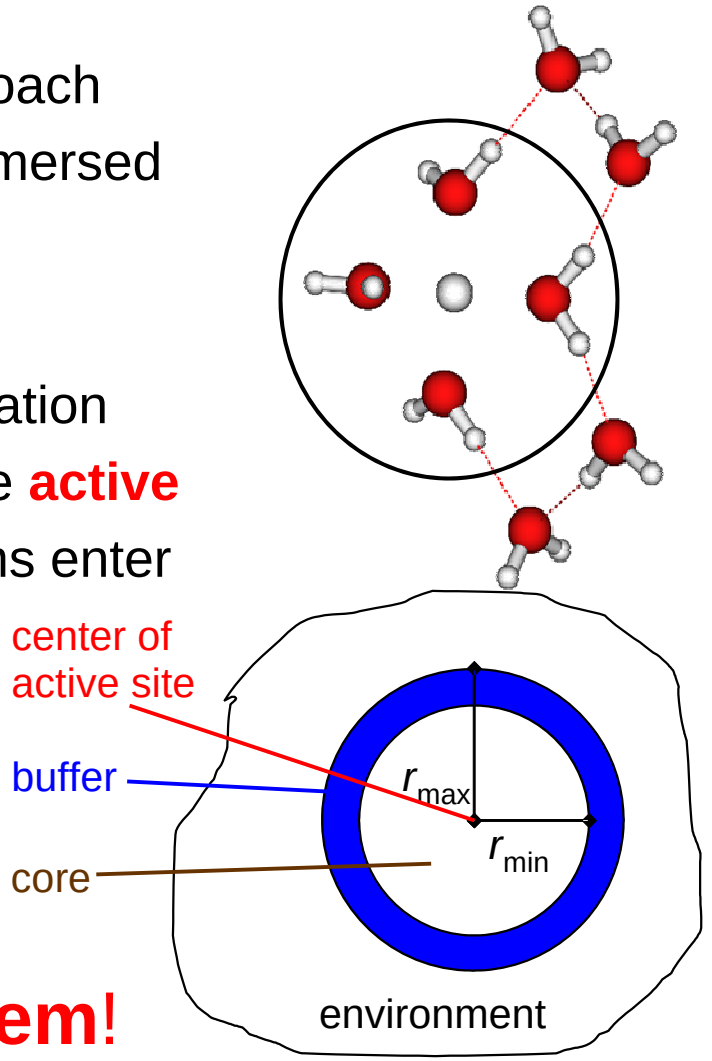
Office of Science
U.S. DEPARTMENT OF ENERGY

Adaptive Partitioning

- Most systems studied with the multilevel approach consist of a small **localized active region** immersed in an extended system, e.g., active site in an enzyme **No problem!**
- For ligand exchange in solution, crack propagation or diffusion and reaction on nanoparticles, the **active region** is **not** necessarily **localized**, and atoms enter or leave active region during the simulation.

⇒ **level of theory** used to describe such an atom **changes** during simulation

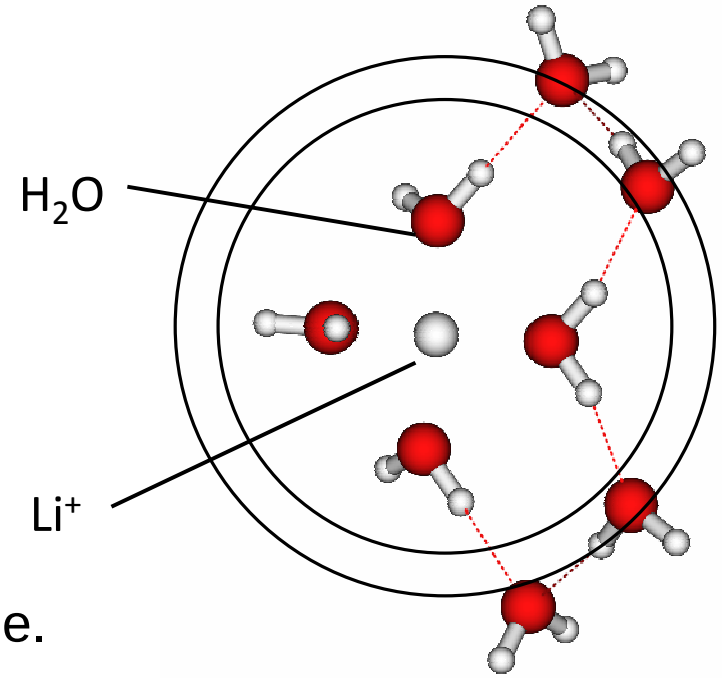
⇒ **discontinuities in the potential energy and forces**



Problem!

The Adaptive Partitioning (AP) method

Need smooth potential energy surface (smooth Hamiltonian) to integrate equations of motion

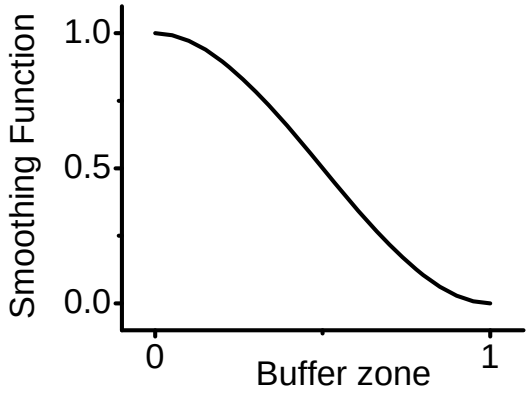


Permuted AP method

- Calculate multilevel energies for a high-level treatment of all possible combinations of the core zone + a subset of the *N* groups in the buffer zone.

- Potential energy is a linear combination of all these multilevel energies, i.e.,

$$V = \text{lin.comb.}\left\{E_C, E_{C,1}, E_{C,2}, E_{C,1,2}\right\}$$



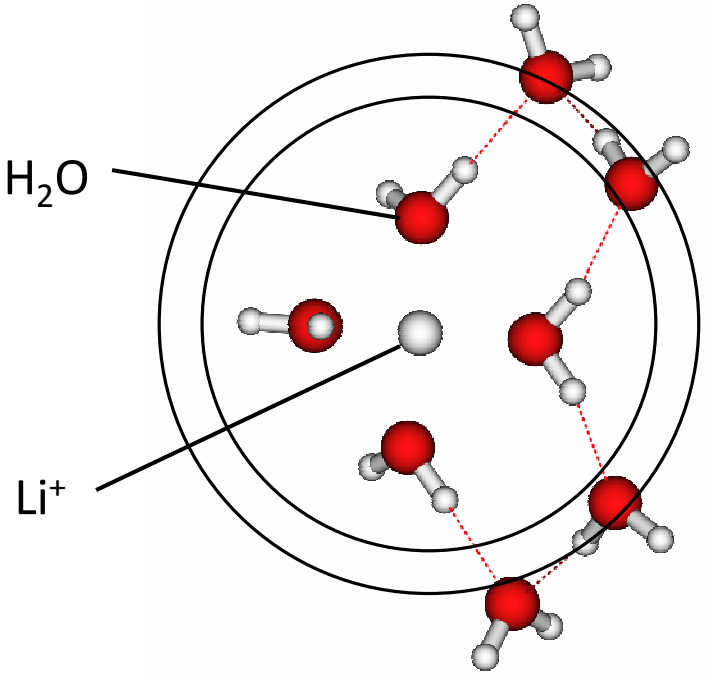
- **Effort scales as *O(2^N).***
- ***Equivalent to ONIOM-XS for 1 group in buffer zone***

The Adaptive Partitioning (AP) method

Sorted AP method

- Computational effort can be reduced if all *N* groups in the buffer zone are sorted and a smart smoothing function is used.

  ⇒ Need to perform only *N+1* multilevel calculations:



$$V = \text{lin.comb.}\{E_C, E_{C,1}, E_{C,1,2}\}$$

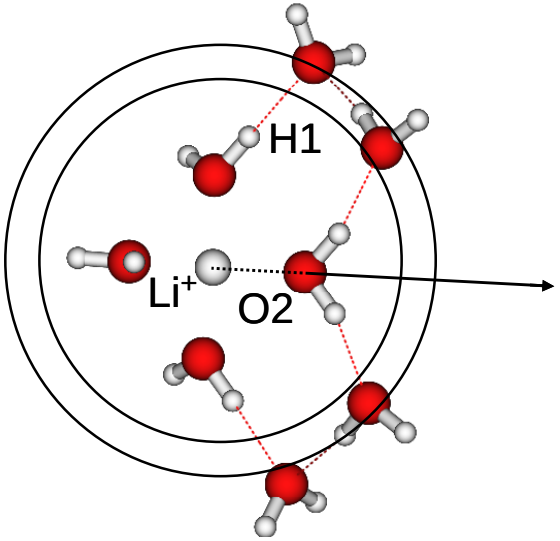
- Potential energy, $V_{c,N}$, is given by the recursion relation:

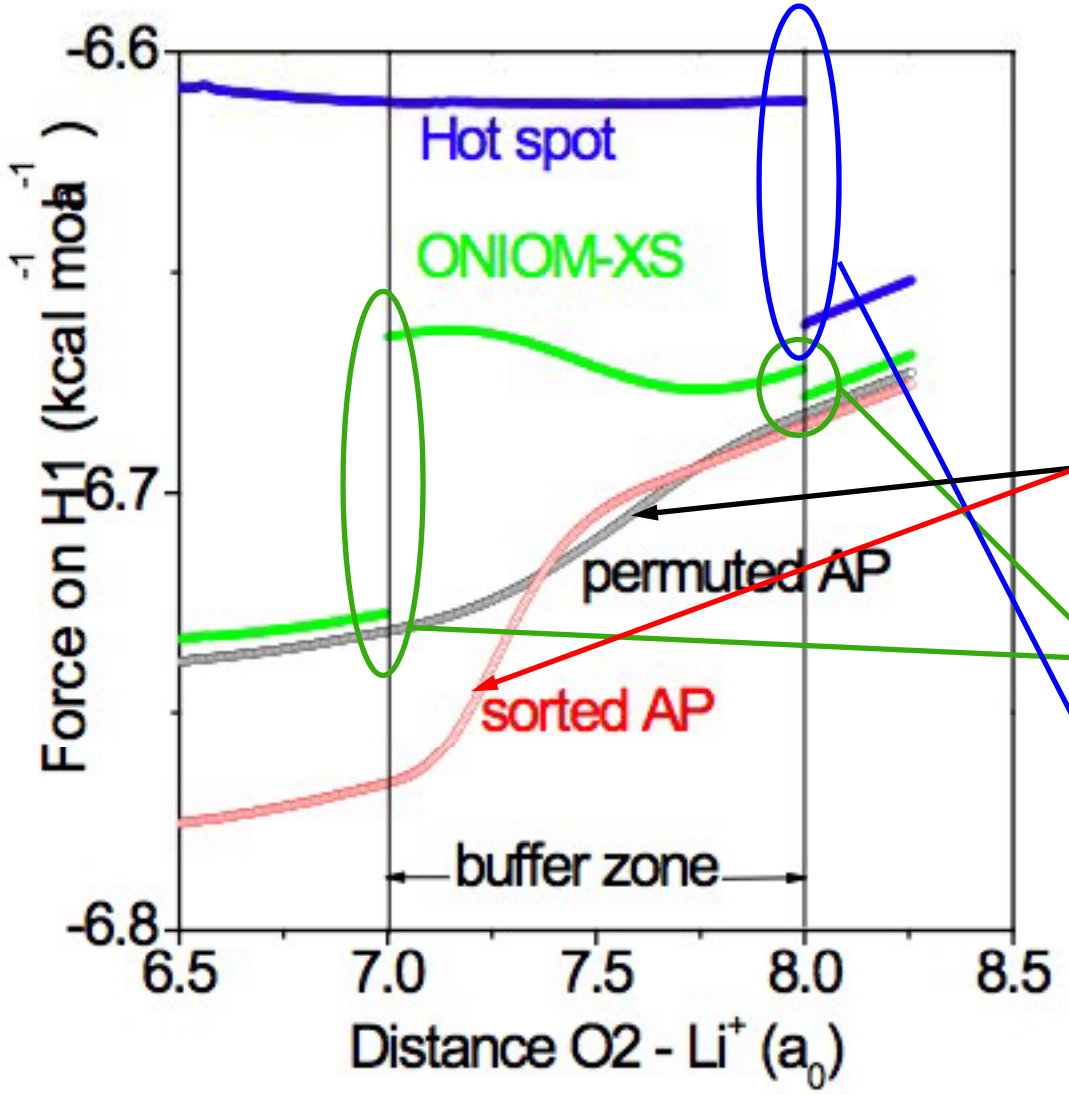
$$V_{C,j} = \tilde{S}_{C,j} E_{C,1,2,\ldots,j} + \left(1 - \tilde{S}_{C,j}\right) V_{C,j-1} \qquad 1 \le j \le N \quad with \quad V_{C,0} \equiv E_C$$

- **Effort scales as *O(N)***

Test simulation

H_2O molecule is leaving the first solvation shell of Li^+





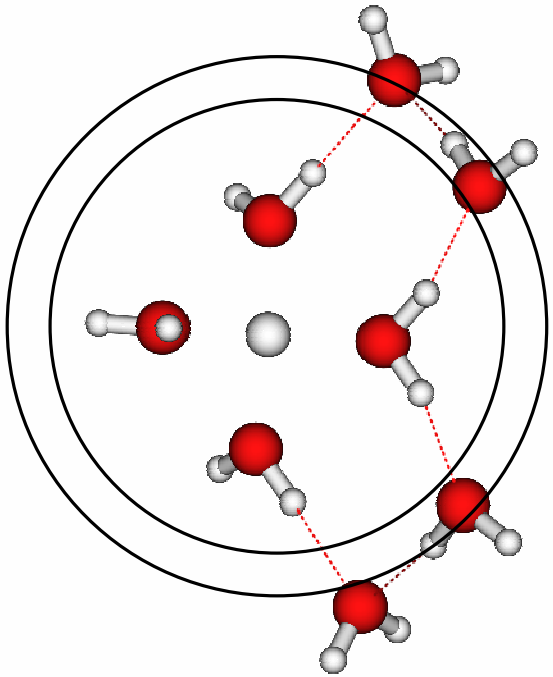
AP methods:
Forces vary smoothly

ONIOM-XS:
Discontinuities

Hot spot method:
Discontinuities

EE-MB + AP

- Natural Adaptation
- Only new n-body interactions need to be calculated
- Allows for larger and higher level QM regions.
- Limitless variety of electronic s... methods which can be applied.



So far we have shown …

Hot spot and **ONIOM-XS** do **NOT** remove discontinuities

⇒ Sampling of undefined ensemble

⇒ Simulation system heats up / becomes unstable

⇒ Simulation results have artifacts

⇒ Should not be used for challenging applications!

Hot spot method worse than a method that does not alter forces

Adaptive partitioning method removes discontinuities and can be **extended to *multiple high-level cores***, Li^+ *and* Cl^- in H_2O.

Permuted AP: Effort scales as $O(2^N)$

⇒ applicable to systems with few groups in buffer zone

Sorted AP: Smart smoothing functions.

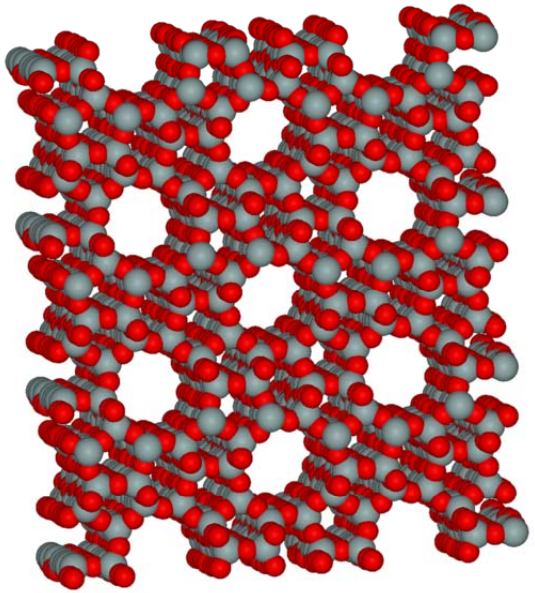
⇒ Effort scales as $O(N)$

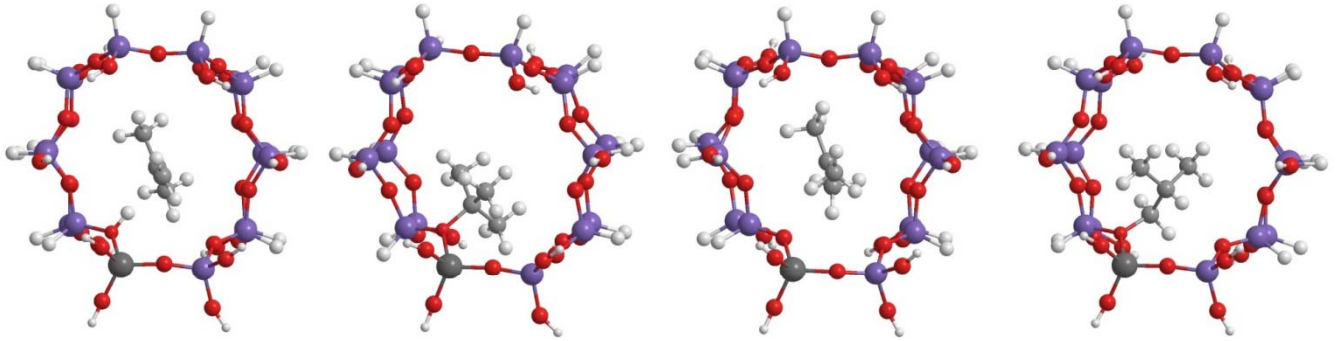
⇒ applicable to large systems

Main message: To avoid problems in molecular simulations always start with Hamiltonian or Lagrangian

Zeolite Study

- Applications of zeolites
  - Heterogeneous catalysis for oil cracking
  - Separation of a variety gas mixtures
    - Alkane/alkene (petroleum industry)
    - Removal of H_2S from biogas mixtures
- Often, the adsorption process is complicated
  - Not just van der Waals interactions (steric factors)
  - But interactions of adsorbate π-electrons, hydrogen-bonding, and chemisorption play a role in hydrophilic zeolites





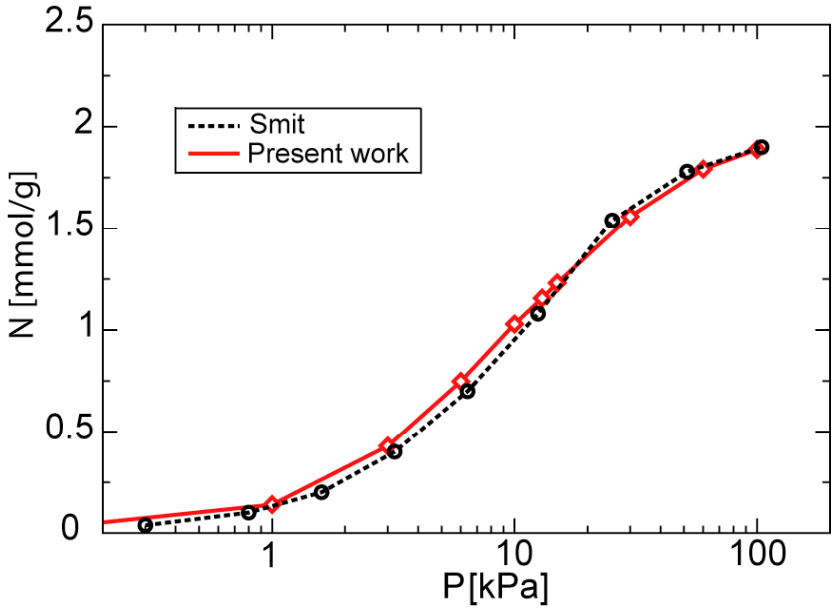
QM vs. MM

- Standard molecular mechanics (MM) force fields are not capable of accurately describing these processes
  - In acidic zeolites, alkenes show much larger isosteric heats of adsorption than alkanes, but MM parameters are quite similar in most force fields
- M06-2X is the most capable of the functionals tested.

Method	π complex	tert-butoxide	tert-butyl carbenium ion	isobutoxide	MUE
Best estimate	15.1	13.9	-9.8	13.9	
M06-2X	12.7	16.6	-9.1	15.6	1.9
M06-L	14.4	15.6	-2.3	13.5	2.6
M06	13.3	16.0	-3.9	14.4	2.6
M06-HF	12.4	18.6	-12.5	18.8	3.7
PBEh	2.9	4.7	-15.6	4.7	9.1
TPSSh	-0.9	1.8	-17.7	1.6	12.1
B3LYP	-2.5	-5.5	-20.7	-4.8	16.6

Validation of QM/MM 1.0

- A module for coupled-decoupled CBMC simulations in the grand canonical ensemble (CBMC-GC) has been implemented in the QM/MM 1.0 Tool.
- Using the MM force field of Smit, the adsorption isotherm of ethane (and of other alkanes) in silicalite was computed using QM/MM 1.0 and found to match the literature data.



- Currently, the isosteric heats of adsorption for propane and propene in zeolite NaA are computed (low pressure limit ⇒ single QM region)

Algorithm Development and Projects for Next Project Period

- To compute the complete adsorption isotherm using CBMC-GC, the QM/MM approach needs to be extended to allow for
  - Multiple QM regions depending on the number of adsorbates near acidic sites
  - Ability to move, merge, create, and destroy QM regions in an automated fashion

- Multipole representation of periodic MM region

- Benchmark systems:
  - Adsorption of propene/propane mixtures in zeolite NaA where experiment[a] shows a separation factor of 16 at 100 kPa
  - Adsorption of $H_2S/CH_4/CO_2$ mixtures in zeolites FAU NaX (Si/Al = 1) and FAU NaY (Si/Al = 2.5) that are candidate structures for H_2S removal[b]

[a] Silva *et al.*, *Ind. Eng. Chem. Res.* **1999,** *38*, 2051

[b] Mauge *et al.*, *J. Catalysis* **2002**, *207*, 353

Tools

- New Methods
  - M06 Density Functionals - Yan
  - Charge Model 4 (CM4) - Alek
  - Solvation Model 8 (SM8) - Alek
  - QM/MM & Other Hybrid Techniques - Marat
  - Electrostatically Embedded Many-Bodied (EE-MB) Method
  - Adaptive Partition method for free floating QM regions within Hybrid Methods
  - Hybrid Configurational-Biased Monte-Carlo using a Grand Canonical Ensemble (Hybrid CBMC-GC)

Integrated Tools

- Compatibility / Interoperability
  - Combine these newly developed technologies with the existing technologies in popular computational chemistry codes.

NWChem

Schrodinger

GAMESS

Gaussian

Q-Chem

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

ACES

Levels of Integration

- Fine Grained Integration
  - Source code modification of core routines
  - Highly technical
  - Varying degrees of difficulty
    - Program dependent

- Course Grained Integration
  - Using core routines without modification
  - Commonly coded as an external program that makes use of methods within electronic structure codes.
    - Scripting

Python Interface to NWChem

- Access internal objects of NWChem from a high-level scripting language.
- Simplify coarse grained integrations
- Direct access to NWChem parallelism.
  - No need to write external parallel programs!

State of Fine Grained Integration

NWChem

Schrodinger

GAMESS

Gaussian

Q-Chem

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

ACES

M06 Integration

NWChem

Schrodinger

Q-Chem

GAMESS

Gaussian

ACES

CM4/SM8 Integration

GAMESS Gaussian

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

Density-Based Solvation Models

NWChem

Schrodinger

GAMESS

Gaussian

Q-Chem

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

ACES

Coarse Grained Methods

- Use the Python interface to NWChem to efficient port coarse grained algorithms
  - Adaptive Partition
  - EE-MB
  - GCMC
- Do so in a way that they can talk to each other!

Acknowledgments

- Donald Truhlar
- Yan Zhao
- Christopher Cramer
- Ilya Siepmann
- Jiali Gao
- Darrin York
- Alek Marenich
- Erin Dahlke
- Andreas Heyden
- Bruce Garrett
- Marat Valiev
- Michel Dupuis
- Shawn Kathmann
- Greg Schenter
- Peng-Dong Fan

ONR!

UNIVERSITY OF MINNESOTA

Pacific Northwest National Laboratory
U.S. Department of Energy

Office of Science
U.S. DEPARTMENT OF ENERGY