SciDAC 4 Annual Meeting: July 16-18, 2019

Advancing Catalysis Modeling:
From Atomistic Chemistry to Whole System Simulation

Martin Head-Gordon,
Alex Bell, Emily Carter (Princeton), Sharon Hammes-Schiffer (Yale), Teresa Head-Gordon, Khaled Ibrahim, Xiaoye Li, David Limmer, Lin Lin, Esmond Ng, Sam Williams, Chao Yang
SciDAC-4 Partnership Team

Alex Bell
Emily Carter
Sharon Hammes-Schiffer
Martin Head-Gordon
Teresa Head-Gordon
Khaled Ibrahim
David Limmer
Sherry Li
Lin Lin
Esmond Ng
Sam Williams
Chao Yang

SciDAC Annual Meeting, July 16-18, 2019
SciDAC 4 Partnership: Advancing catalysis modeling

Physical sciences objectives:
- Advance electronic structure modeling
- Explicit solvent, NQEs and statistical mechanics
- Whole system modeling

Applied math and computer science objectives:
- New algorithms/solvers for this application domain
- New algorithm development (partnerships)
- Improved parallel scaling for supercomputers

23 publications over the past year
This talk will cover some highlights.
See also: poster 9 (Christopher Stein and MHG) and poster 3 (David Williams-Young, P. Beckman and C. Yang)
Advancing catalysis modeling: Outline

1. Electronic structure theory and embedding
   
   Martin Head-Gordon, Emily Carter, Lin Lin and Chao Yang

2. Statistical mechanics & dynamics

3. Applied math and CS/HPC

4. Whole system modeling
ISDF and ACE: Accelerating exact exchange DFT

Wall clock time (s) for Si 1000 system

- **ACE**: [Lin, JCTC, 2016] [Hu, Lin, Banerjee, Vecharynski, Yang, JCTC 2017]
- **ACE-ISDF**: [Hu, Lin, Yang, JCTC 2017] [Dong, Hu, L., JCTC 2018]
- **ACE-ISDF-PCDIIS**: [Hu, Lin, Yang, JCTC 2017a]
- **ACE-GPU-PCDIIS**: [Jia, Lin, in preparation]

Hybrid (conventional): 12425 s

44 x speed up

~ GGA time (more iterations)

Lin Lin, Chao Yang
Hybrid functional RT-TDDFT

For the first time, practical RT-TDDFT with a large basis set

Number of Fock exchange operator applications per orbital within 50 as, 1024 atom silicon

<table>
<thead>
<tr>
<th></th>
<th>S-RK4</th>
<th>PT-CN</th>
<th>PT-CN-ACE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>400</td>
<td>21</td>
<td>5.75</td>
</tr>
<tr>
<td>Reduction</td>
<td>70 times</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**ASCI**: selected CI that adaptively chooses most important configurations up to a specified total. Corrected via PT2. Combined with CASSCF permits very large active space calculations (above (50,50)).

**Ferredoxin core**: 4 irons with oxidation state 2.5. The 20 d orbitals are nominally 2 doubly occupied and 18 singly occupied levels, creating a dense manifold of states. DFT breaks spatial symmetry of this octahedral complex. CASSCF geometry optimization in an active space of 20 d-orbitals, and the 3 orbitals of each sulfur that couple to the d-manifold restores it!

Combined with Emily Carter’s embedding! (see subsequent slide)
Weak correlations: Uncovering the nature of long-range dispersion

Numerical experiments that unfold the long-range $T2$ tensor (right) reveal a striking result (below, right)

Long-range correlation is controlled by only 3 singular values: very low-rank separability! Can be understood on a formal basis (Casimir-Polder for dispersion)

$$C_6^{AB} = \frac{3}{\pi} \int_0^\infty \alpha_A(i\omega)\alpha_B(i\omega)$$

Martin Head-Gordon

Cameron Mackie
Double-bond/double bond correlation (left, molecule above) show the same rapid decay.

1. Present results show long-range correlation can be described with 3 dispersion-specific virtuals per occupied.
2. Future: These virtuals could be optimized to extremize the long-range correlation recovery.
3. Future: ISDF (Lin Lin) combined with this rank reduction will be a synergistic combination.
Mean field (Hartree-Fock): defines a reference problem & a perturbation (fluctuation potential)

**MP2:** captures leading electron correlations & usually greatly improves observables

**MP3:** beyond pairwise electron correlations but usually does not improve calculated observables

Mean-field orbitals often exhibit artificial symmetry-breaking (space &/or spin)

... which degrades MP2 results....

... and makes MP3 even less effective....

Abandon mean-field orbitals!

Replace by orbitals optimized with regularized PT2 ($\kappa$-OOMP2) that usually* restore symmetry.

Will the use of these better orbitals affect MP3 results?
MP3 made useful? Exciting test results.

MP3 (& MP2.5) are not much better than MP2

κ-OOMP2 is better than MP3 and OOMP2

MP2.8 (& MP3) using orbitals from κ-OOMP2 surpass CCSD

Advancing catalysis modeling: Outline

1. Electronic structure theory and embedding

2. Statistical mechanics & dynamics
   Teresa Head-Gordon, David Limmer, Lin Lin, Sharon Hammes-Schiffer

3. Applied math and CS/HPC

4. Whole system modeling
Raising Resolution of (Electro) Catalysis Systems

Electrode-electrolyte interface poses a remarkable series of challenges for computation and modeling at all length and time scales.

The goal of our sub-task is to raise the resolution of the interface and for full statistical mechanical modeling of solvent, fluctuations, catalysis, and driven systems with bias with development of efficient simulation models and software.

AIMD:
N~100, \( t_{\text{obs}} \sim 100 \) ps

New reactive potentials with improved accuracy and efficiency

Reactive MD:
N~10000, 100-1000’s ns

Electronegativity equalization method (EEM)

Coulomb energy and charge transfer solved self-consistently

\[ \chi_{eq} = \chi^* + 2\eta^* q_\alpha + \sum_{j \neq i}^{n} \frac{q_j}{R_{ij}} \]

\( \chi^* \) – Effective electronegativity

\( \eta^* \) – Effective hardness

Computationally demanding bottleneck for the ReaxFF potential

\[ \begin{pmatrix}
2\eta_1^* & 1/R_{12} & \cdots & 1/R_{1n} & -1 \\
1/R_{21} & \cdots & \cdots & \cdots & -1 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
1/R_{n1} & \cdots & \cdots & 2\eta_n^* & -1 \\
1 & 1 & 1 & 1 & 0
\end{pmatrix} \begin{pmatrix}
q_1 \\
\vdots \\
q_n \\
\chi_{eq} \\
Q_{tot}
\end{pmatrix} = \begin{pmatrix}
-\chi_1^* \\
\vdots \\
-\chi_n^*
\end{pmatrix} \]

I. Leven, T. Head-Gordon (2019) (submitted)
Stochastic-extended Lagrangian molecular dynamics (Stochastic-XLMD)

Scientific Achievement: Stochastic-XLMD method builds on the iEL/SCF and iEL/0-SCF methods, and can converge robustly and reduce SCF iterations for model many-body potentials for MD simulations.

Significance and Impact:
- rigorous proof of convergence beyond linear response regime and demonstrating numerical effectiveness
- generalized the approach of averaging and techniques for analyzing degenerate elliptic operators and proved the error bound for the new method.
- numerically studied scaling of numerical noise, damping factor and perturbation parameter for efficient simulation of non-quadratic interaction energy forms.


Stochastic-XLMD is currently being extended to AIMD (CP2K) and EEM (ReaxFF in LAMMPs)

Teresa Head-Gordon, Lin Lin
While ReaxFF is sufficient for predicting charge fluctuations in different molecular environments, these models fail to describe redox electrochemical reactions which are accompanied by a flow of charge to/from the anode surfaces.

We have developed a new approach for treating charge flow in reactive force fields through an explicit treatment of electrons. Our model gives improved treatment of charge transfer and also goes beyond the approximation of atomic point charges to account for a more flexible charge distribution in molecules.

I. Leven, T. Head-Gordon (2019) In progress

Charge Transfer Models for Reactive Force Fields

- An atom is partitioned into a core containing the nuclei + core electrons and a shell containing valence electron(s).
- Core and shell are treated as Gaussian charges

\[
\rho_c^i(\vec{r}) = q_c \cdot \left( \frac{\alpha_c^i}{\pi} \right)^{3/2} e^{-\alpha_c^i(|\vec{r} - \vec{r}_c^i|)^2} \quad \rho_s(\vec{r}) = q_s \cdot \left( \frac{\alpha_s}{\pi} \right)^{3/2} e^{-\alpha_s(|\vec{r} - \vec{r}_s|)^2}
\]

Resulting interaction potential over core-shell and shell-shell

\[
E_{CT} = \sum_{i} \sum_{j<i} \left( E_{ij}^{\text{elec}}(r_{ij}) + E_{ij}^{\text{Gauss}} \right)
\]

\[
E_{ij}^{\text{elec}}(r_{ij}) = \frac{q_i q_j}{r_{ij}} \text{erf} \left( \sqrt{\frac{\alpha_i \alpha_j}{\alpha_i + \alpha_j}} r_{ij} \right)
\]

\[
E_{ij}^{\text{Gauss}}(r_{ij}) = A_{ij} e^{-\gamma_{ij} r_{ij}^2}
\]

Potential is tuned such that interaction between shell-core and shell-shell yields the ionization potential of core atom type and corresponding electron affinity

I. Leven, T. Head-Gordon (2019) In progress

Teresa Head-Gordon
Nonadiabatic Transition Path Sampling (QTPS)

- Transition path sampling quantum nuclei with quantum master equations

Schile, Addison J., and David T. Limmer. JCP 149, 214109 (2018); 151, 014106 (2019)
Rate Constants in Inhomogeneous Systems from TPS

• Transition path sampling for spatial rate dependence

\[ Z_{AB}(y) = \int dx_0 \rho(x_0) h_A(x_0) h_B(x_t(x_0)) \delta(y - y[x_0]) \propto k_{AB}(y) e^{-\beta F(y)} \]

Schile, Addison J., and David T. Limmer. 150, 191102 (2019)
Artificial Neural Network Reactive Molecular Dynamics Potentials

- Machine learning for reactive FFs

Condensed phase formation and dissociation of nitric acid at the GGA level with neural networks

Neural Networks: reactive potentials from AIMD

Mirza Galib and David T. Limmer. In preparation
Nuclear-Electronic Orbital (NEO) Approach

- Solution of mixed nuclear-electronic time-independent Schrödinger equation with molecular orbital methods
- Treat specified nuclei quantum mechanically on same level as electrons
  - treat only key H nuclei QM
  - retain at least two classical nuclei
- Includes proton delocalization and zero point energy during geometry optimizations, reaction paths, and dynamics rather than as corrections
- Avoids Born-Oppenheimer separation between electrons and protons
- Highly suitable for studying proton-coupled electron transfer (PCET) reactions in electrocatalysis systems
**NEO Coupled-Cluster Singles and Doubles Theory**

- Derived and implemented NEO-CCSD equations

\[ |\text{CCSD}\rangle = e^{\hat{T}} \left| \Psi_{\text{NEO-HF}} \right\rangle \quad \hat{T} = \hat{T}_1^e + \hat{T}_2^{ee} + \hat{T}_1^p + \hat{T}_2^{pp} + \hat{T}_2^{ep} \]

- \( T \) is the sum of single and double cluster operators for electrons and protons
- Calculated proton densities, proton affinities, and optimized geometries with no parameterization
- NEO-CCSD provides qualitatively accurate proton densities & affinities
- NEO-EOM-CCSD is being developed for excited vibrational states

Mean unsigned error wrt experiment for proton affinities: 0.04 eV

NEO-EOM-CCSD for Excited States

- NEO-EOM-CCSD equations derived and implemented
- Applied to PsH, where positron and both electrons quantum mechanical
- Applying to quantum protons

<table>
<thead>
<tr>
<th>state</th>
<th>NEO-FCI</th>
<th>NEO-FCC</th>
<th>NEO-EOM-CCSD</th>
<th>NEO-CISD</th>
</tr>
</thead>
<tbody>
<tr>
<td>ground state</td>
<td>-0.090593</td>
<td>-0.090593</td>
<td>-0.085978</td>
<td>-0.085733</td>
</tr>
<tr>
<td>1\textsuperscript{st} excited state</td>
<td>0.156700</td>
<td>0.156700</td>
<td>0.155107</td>
<td>0.179252</td>
</tr>
<tr>
<td>2\textsuperscript{nd} excited state</td>
<td>0.165857</td>
<td>0.165857</td>
<td>0.166041</td>
<td>0.187039</td>
</tr>
<tr>
<td>3\textsuperscript{rd} excited state</td>
<td>0.242055</td>
<td>0.242055</td>
<td>0.239553</td>
<td>0.254243</td>
</tr>
</tbody>
</table>

Energies in au. Excitation energies are relative to the ground state for each method.

EOM: equation-of-motion


Sharon Hammes-Schiffer
Advancing catalysis modeling: Outline

1. Electronic structure theory and embedding

2. Statistical mechanics & dynamics

3. Applied math and CS/HPC
   Chao Yang, Xiaoye Li, Lin Lin, Esmond Ng, Khaled Ibrahim, Sam Williams

4. Whole system modeling
Extending the Scalability of Existing Eigensolvers via Spectrum Slicing

- Existing solvers (ScalAPACK/ELPA) not scalable beyond a few thousands of cores
- Spectrum slicing has better scalability
  - Initial spectrum partition by Lanczos DOS estimation
  - Refined partition by K-means clustering
A new 2D task-based symPACK

- Factorization crucial to PEXSI
- symPACK for symmetric systems
  - 1D supernodal distribution:
    - Balances flops, memory
    - Lacks strong scalability
  - New task-based 2D data distribution
    - Explicit load balancing, not regular block cyclic mapping
    - Balances flops, memory
    - High strong scalability
  - Strong scalability on Cori Haswell:
    - Up to 3x speedup for Serena
    - Up to 2.5x speedup for DG_Phosphorene

www.sympack.org

Sherry Li, Esmond Ng
STRUMPACK – low-rank STRUctured Matrix Package

- Generic direct solvers and preconditioners using hierarchical low-rank compression techniques
  - Nearly-linear complexity $O(N \text{ polylog}(N))$

Recent progress
- In addition to HSS, other formats: HODLR, BLR
- Fast rank-detection using adaptive randomized sampling
  - **Difficulty**: how many random vectors to use?
  - Developed a new stopping criteria based on stochastic norm estimation

Let $A \in \mathbb{R}^{m \times n}$, and $x \in \mathbb{R}^n$ with $x_i \sim N(0,1)$, we proved:

$$E\left[\|Ax\|_2^2\right] = \sigma_1^2 + \cdots + \sigma_r^2 = \|A\|_F^2$$

For $d$ sample vectors: $E\left[\|S\|_F^2\right] = d \|A\|_F^2$

- More accurate than commonly-used upper bound estimation (Halko-Martinsson-Tropp)

<table>
<thead>
<tr>
<th></th>
<th>NEW</th>
<th>OLD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compression</td>
<td>5.50</td>
<td>8.24</td>
</tr>
<tr>
<td>Factorization</td>
<td>1.98</td>
<td>2.83</td>
</tr>
<tr>
<td>Solve</td>
<td>0.14</td>
<td>0.15</td>
</tr>
</tbody>
</table>

**BEM electromagnetic multi-sphere**
Direct solver, $N = 27K$, dense

**Indefinite Maxwell equations, EM diffusion**
Preconditioner to GMRES, $N = 331K$, sparse

State-of-the-art AMS multigrid preconditioner does not converge

Sherry Li, Esmond Ng
Optimization of DFT Calculation on HPC Platforms

Target platforms
- Multi/Many core architectures (Haswell and KNL)
- GPU accelerated architectures

Software tuning challenges:
- CP2K: Complex software stack, thread parallelism across layers, MPI Communication
- DGDFT: MPI Communication, locality in distributed cache, load imbalance, improve the use of external.

Accomplishments:
- CP2K: Improve CP2K build on Haswell and KNL (up to 1.7x)
- DGDFT: Improve communication, load balancing and interaction with external libraries (up to 2x).

Khaled Ibrahim
Optimizing Density Calculation

Loop optimizations:
- Loop reordering to improve the load balancing
- Native indexing, no error checking
- Loop parallelization

Performance:
- 10-37x speedup.
- KNL 15x speedup

Khaled Ibrahim
Advancing catalysis modeling: Outline

1. Electronic structure theory and embedding
2. Statistical mechanics & dynamics
3. Applied math and CS/HPC
4. Whole system modeling

Alex Bell, Emily Carter, Teresa Head-Gordon, Martin Head-Gordon
Collaboration within SciDAC on the Vibrational Stark Effect

Vibrational Stark Effect in Heterogenous Catalysis

Vibrational Stark Effect on Noble Metal Atoms

Treatment of Electric Fields in Bio-systems

Opportunity to investigate the same phenomena in the context of different physical systems.
Results: Experimental vs. computed vibrational frequencies

- Calculations: SCAN functional
- Theory predicts red-shift upon adsorption on Ag(110)
- Theory predicts further red-shifting upon aqueous environment
- Origin of discrepancy between experiment and theory: anharmonicity, surface termination, surface coverage, and functional.
- Further experiments underway to assess theoretical predictions.

Experiments: Kun Jiang (LBL) + Xianyin Ma (Fudan)
**Origin of the Stark effect: Red and blue shifting in M-CO complexes**

**Physical Stark effect:** electric fields due to local charges control frequency shifts via dipole moment changes.

**Chemical Stark effect:** charge flow due to CO binding controls frequency shifts via dative interactions

Probe by finding the frequency shifts using constraints to prevent charge transfer (and polarization)

<table>
<thead>
<tr>
<th>Species</th>
<th>$\delta_{FRZ}$</th>
<th>$\delta_{POL}$</th>
<th>$\delta_{CT}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>TiCO$^-$</td>
<td>1.69</td>
<td>-6.19</td>
<td>-383.87</td>
</tr>
<tr>
<td>VCO</td>
<td>-0.81</td>
<td>40.25</td>
<td>-237.00</td>
</tr>
<tr>
<td>CrCO$^+$</td>
<td>86.94</td>
<td>25.01</td>
<td>-34.92</td>
</tr>
<tr>
<td>VCO$^-$</td>
<td>-1.03</td>
<td>-6.51</td>
<td>-369.18</td>
</tr>
<tr>
<td>CrCO</td>
<td>-0.18</td>
<td>0.04</td>
<td>-119.21</td>
</tr>
<tr>
<td>MnCO$^+$</td>
<td>56.05</td>
<td>33.29</td>
<td>8.57</td>
</tr>
</tbody>
</table>

Above results identify charge transfer (“chemical effects”) as the origin of red shifting (strongest in metal anions)

Frozen response (“physical effects”) is the origin of blue shifting in the cations for above cases: the $\delta$-CO$\delta^+$ dipole increases as bond-length gets shorter.

E. Rossomme, M. Head-Gordon (2019) In progress
SciDAC 4 Partnership: Advancing catalysis modeling

Physical sciences objectives:

Advance electronic structure modeling
Explicit solvent, NQEs and statistical mechanics
Whole system modeling

Applied math and computer science objectives:

New algorithms/solvers for this application domain
New algorithm development (partnerships)
Improved parallel scaling for supercomputers

With Alex Bell, Emily Carter, Sharon Hammes-Schiffer, Teresa Head-Gordon, Khaled Ibrahim, Xiaoye Li, David Limmer, Lin Lin, Esmond Ng, Sam Williams, Chao Yang