New Methods and Models for Condensed



Teresa Head-Gordon

Chemistry, Bioengineering, Chemical & Biomolecular Engineering University of California, Berkeley

> Chemical Sciences Division Lawrence Berkeley National Laboratory



Computational Molecular Sciences (CMS)

CMS is comprised of the sub-fields of quantum chemistry, computational materials science, and biomolecular simulation.

• The history of our field reaches back decades – to the genesis of computational science – CMS is now a "full partner with experiment".

• For an impressive array of chemical, biochemical, and materials challenges, our community has developed simulations and models that directly impact areas such as drug design, macromolecules, advanced materials, etc.

Barriers to Better Scientific Software

- Incredible code complexity arising from long legacies
- Ever-increasing hardware complexity
- Limited education in software development in physical sciences

Molecular Sciences Software Institute (MolSSI) was funded in August 2016 https://nsf.gov/news/news_summ.jsp?cntn_id=189347&org=NSF&from=news

MolSSI Software Community

A principal representative from each of the following community codes has committed to collaboration between their development team and the **MolSSI Software Scientists and leadership:**

- GAMESS
- Molpro
- Q-Chem
- PSI4
- ACESIII
- CFOUR
- Molcas
- Orca
- SISSA

- Gaussian
 CHARMM
 - Amber
 - BOSS
 - Gromacs
 - OpenMM
 - LAMMPS
 - Plumed
 - Turbomole
 - NWChem
 - ONETEP

- NAMD
- Dalton
- Columbus
- Dirac
- DL_POLY
- Tiger-CI
- Schrödinger
- Quantum ESPRESSO
- PARSEC
- APBS

Engaging National Labs, Industry, International

We have commitments to collaborate from following representatives. Workforce development and career paths into dozens of industries. Intel has also agreed to participate in educational aspects of MolSSI, leveraging the company's extensive training materials and experience.

- PNNL (Allison Campbell, Assoc. Lab Director)
- ORNL (Jeff Nichols, Assoc. Lab Director)
- BNL (Robert Harrison, Computational Science Director)
- LBNL (Kathy Yelick, Assoc. Lab Director for Computing Science)
- NVIDIA (David Leubke, Sr. Director of Research)
- Intel (Timothy Mattson, Principal Engineer, Parallel Computing Lab)
- IBM (Jed Pitera, Principal Research Manager)
- Cray, Inc. (Ryan Olsen, Performance Engineering)
 ELS

- EPSRC: ARCHER eCSE
- EU Computational Materials Centers
- EU Center of Excellence on Biomolecular Simulation (BioExcel)
- Our S2I2 Conceptualization workshops prompted the UK's EPSRC to report on how the British CMS community could interface to MolSSI.

MolSSI: Software Scientists



Dr. Paul Saxe

Lead Software Scientist





Doaa Altarawy

Ph.D., Virginia Tech

Computer Science, Machine Learning, Algorithms, Software Engineering, Visualization, Computational Biology



Taylor Barnes

Ph.D., California Institute of Technology

Parallelization, Optimization, Embedding, Batteries



Eliseo Marin-Rimoldi

Ph.D., U. Notre Dame

Monte Carlo, Molecular Simulation, Statistical Mechanics, Advanced Sampling Techniques



Jessica Nash

Ph.D., N.C. State University

Molecular Dynamics Simulations, All-Atom Simulations, Nucleic Acids, DNA Nanotechnology, Nanomaterials



Benjamin Pritchard

Ph.D., U. Buffalo

C++, Software interoperability, Code optimization, High-precision computation



Daniel Smith

Ph.D., Auburn University

Quantum Chemistry, Task-based Workflows, Rapid Prototyping, Databases

MolSSI Board of Directors

Daniel Crawford (VT), Cecilia Clementi (Rice), Teresa Head-Gordon (Berkeley), Vijay Pande (Stanford), Robert Harrison (Stony Brook), Anna Krylov (USC), Theresa Windus (Iowa), Shantenu Jha (Rutgers)



The Molecular Sciences Software Institute

Software Scientists: A team of software engineers who will develop opensource infrastructure, frameworks, and modules for the broad spectrum of community codes. (Open positions available!)
Software Fellows: A cohort of students and postdocs in research groups across the country who will work with the Institute staff to develop new software tools. (Next call coming late August)
Molecular Sciences Consortium: define standards for the community Community Workshops: offer financial and logistical support for workshops that address important scientific areas where software bottlenecks are a challenge

Workshop Outcomes:

Requirements capture Use cases Report My suggestion is to appoint a scribe for each session

Beyond Pairwise Additivity

We know that true energy is not pairwise additive! What are the important (noncovalent) energetic terms that are many-body (not pairwise additive)? Classification per A. J. Stone:

Weak interactions (decay ~exp(-αr)):

- Exchange-repulsion (represented as 1/r¹² term in Lennard-Jones)
- Charge-transfer
- Charge penetration

Strong interactions (decay ~1/rⁿ)

- Dispersion (represented by 1/r⁶ in Lennard-Jones)
- Polarization missing from most broadly used force fields; (decays as 1/r⁴)

Stone, A. J. The Theory of Intermolecular Forces, Clarendon Press, Oxford, 1996.

Mutual Polarization: AMOEBA

Consider the N-Body AMOEBA model

 $U = U_{bond} + U_{angle} + U_{b\theta} + U_{oop} + U_{torsion} + U_{7-14buffer} + U_{multipoles}^{perm} + U_{dipoles}^{ind}$

P. Ren & J. Ponder (2003) J. Phys. Chem. B 107, 5933-5947

- Permanent multipole electrostatics: ~5-10X of fixed charge FFs
- Mutual polarization with SCF solver: additional ~5-10X

$$\mu_{i,\gamma} = \alpha_i \left(\sum_{\{j\}} T_{\gamma}^{ij} M_j^{perm} + \sum_{\{j'\}} T_{\gamma\delta}^{ij'} \mu_{j'\delta} \right) \qquad \gamma, \delta = x, y, z$$

Net cost of AMOEBA compared to fixed charge: ~25-100X

J. W. Ponder, C.J. Wu, P. Ren, V.S. Pande, I. Haque, R.A. Distasio Jr., D. Lambrecht, M. Head-Gordon, G.N.I. Clark, M.E. Johnson, T. Head-Gordon (2010). J. Phys. Chem. B (Feature Article and Cover) 114, 2549-2564.

Mutual Polarization: Drude

Consider the PSPC Drude water model

 $U = U_{LJ} + U_{charge}^{perm} + U_{Drude}^{ind}$

Van Belle, D.; Couplet, I.; Prevost, M.; Wodak, S. J. (1987) JMB 198, 721-735.

• Mutual polarization with SCF solver: additional ~10X

 $\frac{\partial U}{\partial r_{D,i}} = k_D d_i - q_{D,i} E_{D,i} (d^N) = 0$ Net cost of PSPC compared to fixed charge: ~10X

• Extended Lagrangian (EL) in NV(T,T*) and NP(T,T*)

 $L(r^{N}, \dot{r}^{N}, r_{D}^{N}, \dot{r}_{D}^{N}) = \frac{1}{2} \sum_{i} (m_{i} - m_{D}) \dot{r}_{i}^{2} + \frac{1}{2} m_{D} \sum_{i} \dot{r}_{D,i}^{2} - U_{LJ}(r^{N}) - \frac{1}{2} \sum_{i} \sum_{j > i} \frac{q_{i}q_{j}}{|r_{i} - r_{j}|} - \frac{1}{2} \sum_{i} \sum_{j} \frac{q_{i}q_{D,j}}{|r_{i} - r_{D,j}|} - \frac{1}{2} \sum_{i} \sum_{j} \frac{q_{D,i}q_{D,j}}{|r_{i} - r_{D,j}|} - \frac{1}{2} k_{D} \sum_{i} |r_{i} - r_{D,i}|^{2}$

Much less expensive! Small mass m_D satisfies BO condition but small time step, larger allows larger time steps with degradation of BO Lamoureux, MacKerell, Roux (2003) J. Chem. Phys.119, 5185; Lamoureux, Roux (2003) J. Chem. Phys. 119 3025. ELSI-MolSSI 2018

EL/SCF for Classical Polarization

Hybrid EL/SCF scheme wherein an extended set of auxiliary electronic DOFs serve as an initial guess of the SCF solver in BOMD:

Loose convergence conserves energy and fewer SCF cycles!

We adapted Niklasson scheme for classical polarization by applying Lagrangian equation

$$\mathcal{L}_{hybrid}^{dipole} = \frac{1}{2} \sum_{i=1}^{N} m_{i} \dot{\vec{r}_{i}}^{2} + \frac{1}{2} \sum_{i=1}^{N} m_{\mu,i} \dot{\vec{\mu}_{i}}^{2} - U(\vec{r}^{N}, \vec{\mu}_{SCF}^{N}) - \frac{1}{2} \omega^{2} \sum_{i=1}^{N} m_{\mu,i} (\vec{\mu}_{SCF,i} - \vec{a}_{i})^{2}$$

in limit that $m_{\mu,i} \rightarrow 0$:

$$\mathbf{m}_{i}\ddot{\mathbf{r}}_{i} = -\frac{\partial U(\mathbf{r}^{N}, \mathbf{\mu}_{SCF}^{N})}{\partial \mathbf{r}_{i}}\Big|_{\mathbf{\mu}^{N}}$$
$$\ddot{\mathbf{a}}_{i} = \omega^{2}(\mathbf{\overline{\mu}}_{SCF} - \mathbf{\overline{\mu}}_{i})$$

Auxiliary dipoles: initial guess for SCF solved through numerical integrators that are time-reversible/area preserving

SCF.I

Niklasson, *Phys. Rev. Lett.*, **2008**; Niklasson & Cawkwell, *Phys. Rev. B*, **2012**; Niklasson & Souvatzis, *J. Chem. Phys.*, **2013**; Niklasson & Cawkwell, *J. Chem. Phys.*, **2014; A. Albaugh, O. Demerdash, THG (2015) J. Chem. Phys. 143, 174104**

EL/SCF vs. Standard SCF

Compared to standard SCF solver, energy conservation with loose convergence improves under EL/SCF



A. Albaugh, O. Demerdash, THG (2015) J. Chem. Phys. 143, 174104

But EL/SCF is clearly unstable: iterations increase without bound!



ELSI-MolSSI 2018

Origin of Instability of EL/SCF

Auxiliary dipole equations of motion are on faster time scale than that experienced by the real induced dipoles.

Their direct coupling in auxiliary potential

$$- \, \frac{1}{2} \, \omega^2 \sum_{i=1}^N m_{\mu,i} \left(\overline{\mu}_{SCF,i} - \overline{\mu}_i \right)^2$$

leads to corruption of dynamics, manifesting as a "kinetic energy" that builds over time.

This is a classic problem with resonances!

As KE builds, auxiliary dipole initial guess becomes increasingly poor, thereby increasing the number of SCF cycles without bound



A. Albaugh, O. Demerdash, THG (2015) J. Chem. Phys. 143, 174104 ELSI-MolSSI 2018

Solution to Instability: (Inertial) iEL/SCF

We introduce "temperature" control to auxiliary degrees of freedom

Berendsen:

$$= \sqrt{1 + \frac{\Delta t}{\tau} \left(\frac{T_{\mu}}{\langle \dot{\mu}^2 \rangle} - 1\right)}$$

(also Nosé-Hoover chains)



Time-reversibility preserved at loosest convergence

α

Shows stable # iterations (halves SCFs at loose convergence)

A. Albaugh, O. Demerdash, THG (2015). J. Chem. Phys. 143, 174104

Properties with iEL/SCF

0.8

0.75 0

0.2

0.4

t (ps)

0.6

0.8

Dipole magnitude and in-plane and out-plane captured perfectly



Works for water, proteins, ions, etc – no parameterization

A. Albaugh, O. Demerdash, THG (2015). J. Chem. Phys. 143, 174104

iEL/SCF for *ab initio* Molecular Dynamics

iEL/SCF has been extended to linear scaling DFT BOMD

V. Vitale, J. Dziedzic, A. Albaugh, A. Niklasson, THG, C. Skylaris (2017) JCP 146, 124115



Requires ~20 SCF cycles to get to inherent energy drift rates of ~5 K/ps drift rate (noise in linear scaling DFT forces) Using iEL/SCF ~3 SCF cycles to get 1 K/ps; is also compensating for the ~5 K/ps drift rate due to noise in linear scaling DFT forces

EL/0-SCF for Mutual Polarization

Using the following general definition of the polarization energy $U_{pol}(r,\underline{\mu}) = \frac{1}{2} \underline{\mu}^T \underline{\underline{C}} \underline{\mu} - \underline{\mu}^T \underline{\underline{E}}^{perm}$

When $\underline{\mu}$ is different from SCF by δ , error is pushed out to second order in energy and gradient $U_{pol}(r,\underline{\mu}) - \frac{1}{2} \underline{\mu}_{SCF}^T \underline{E}^{perm} = \delta^T \underline{C} \delta$ Challenge for AIMD

> $\frac{dU_{pol}}{dr} - \frac{dU_{pol}^{SCF}}{dr} = \frac{1}{2} \delta^T \frac{\partial C}{\partial r} \delta$ given first order error in gradient

Challenge for AIMD given first order error in gradient

issue now is to control error δ ; using auxiliary dipoles as an initial guess to the real dipoles $\mu = \alpha (\underline{E}^{perm} + T'' \underline{a})$

we perform no SCF cycles but instead define μ_{SCF} from simple linear mixing of real and auxiliary

 $\mu_{SCF,i} \approx \gamma \mu_i + (1 - \gamma) a_i$ This is only used to derive auxiliary equations of motion

$$\ddot{\boldsymbol{a}}_i = \omega^2 \gamma (\boldsymbol{\mu}_i - \boldsymbol{a}_i)$$

A. Albaugh, A. Niklasson, THG (2017). JPC Lett 8, 1714–1723

EL/0-SCF for AMOEBA

Extended system energy is conserved; NVT energies are quantitative





Properties are perfect!

Parallel timings at 1fs are as good as 2fs RESPA using standard SCF!



A. Albaugh, A. Niklasson, THG (2017). JPC Lett 8, 1714–1723 ELSI-MolSSI 2018

EL(T,T*) scheme for Drude Oscillators

Drude model: Mobile charge (m_D) attached to parent atom (m_i-m_D) by harmonic spring. Strength of spring constant determined by polarizability.

Pushing the time step: EL(T,T*) stable until 4 fs but 25% mass repartitioning: incorrect linear and angular momentum....and polarization properties are degrading



A. Albaugh, A. Niklasson, THG (2017). JPC Lett 8, 1714–1723 ELSI-MolSSI 2018

EL/0-SCF for Drude Oscillators

EL/0-SCF stable until 6-7 fs: no mass repartitioning; NVT but cold Drude auxiliary oscillators; same quality as tightly converged SCF solution....and polarization properties are perfect





A. Albaugh, A. Niklasson, THG (2017). JPC Lett 8, 1714–1723

Why does EL/0-SCF work?

Single evaluation of $\underline{\mu} = \alpha (\underline{E}^{perm} + T''\underline{a})$ using current auxiliaries, updated every 2-6 fs



is performing a tightly converged SCF calculation on the fly given the time decay of real dipoles/Drudes.

ELSI-MolSSI 2018

A. Albaugh and T. Head-Gordon (2017). *JCTC* 13(11):5207-5216.

iEL/0-SCF and Isokinetic Scheme

We have combined isokinetic and RESPA (multi-time stepping) methods to push outer time step to 90 fs – polarization in middle timescale inexpensive now!





Our timings are ~10X faster than reported for SCF version

A. Albaugh, M. Tuckerman, THG (2018) In preparation

iEL/0-SCF and Isokinetic Scheme

Isokinetic method gives rise to small errors in molecular

dipole distributions



But standard structural and thermodynamic properties are in very good agreement with standard MD/SCF approaches







A. Albaugh, M. Tuckerman, THG (2018) In preparation

ELSI-MolSSI 2018

Summary

•iEL/SCF improves calculation of mutual polarization, dropping SCF cycles by half; for linear scaling DFT AIMD, SCF drops from ~20 cycles to ~3-6.

• iEL/0-SCF evaluates an SCF quality solution to mutual polarization at cost of the chosen permanent electrostatics;

• working on approaches to AIMD!

Acknowledgments



Dr. Alex Albaugh, Dr. Saurabh Belsare, Dr. Asmit Bhowmick, Dr. Lisa Felberg, James Lincoff, Dr. Sukanya Sasmal. (Ugrads: Taoran Zhang, Tianyou Liu, Michael Vincent, Sun Hajela)

Dr. Kochise Bennett, Dr. Omar Demerdash, Dr. Luis Ruiz Pestana, Dr. Sudhir Sharma, Dr. Valerie Vassier Welborn, Dr. Akshaya Das

R. Saykally, D. Wemmer, D. Neumark, E. Williams, L. Lammers, M. Salmeron, M. Head-Gordon (Berkeley), C. Skylaris, J. Dziedzic, J. Essex, R. Bradshaw (Southampton), D. Case (Rutgers), M. Tuckerman (NYU), A.NN Niklasson (LANL), M. Havenith, D. Marx (Bochum-Ruhr) J. Forman-Kay (Toronto), J. Dyson, P. Wright (Scripps), B. Swope, J. Rice, (IBM), J. Ponder (WashU), P. Ren (UT Austin), J. P. Piquemal (Sorbonne, Paris); L.-P. Wang (UC Davis), P. Eastman, V. Pande, T. Martinez, T. Markland (Stanford), MolSSI Board of Directors, Software Scientists

\$DOE/CPIMS (Enzymes), ALCC/NERSC/DOE (Cycles), DOE/SCIDAC-4 NSF-TMC (Polarization), NSF-ACI (MolSSI), NIH-NIGMS (NMR); NIH-NIGMS (IDPS), NIH-Biophysics TG, Ruhr Bochum & UC Berkeley: ReSolv/CalSolv \$ *ELSI-MolSSI 2018*