SAND2020-9710PE

Predicting the Behavior of Biomolecules with Quantum Physics

UNM Biology Department Seminar: September 2020

Joshua A. Rackers, Truman Fellow

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Who am I?

Physics and Political Philosophy

TEACH FOR AMERICA

High School Physics and Chemistry

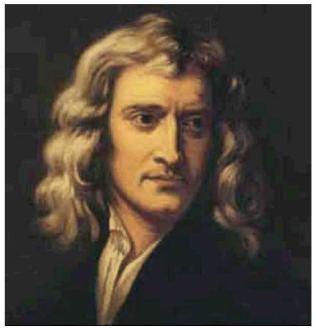
Washington University in St. Louis

Biophysics Ph.D.

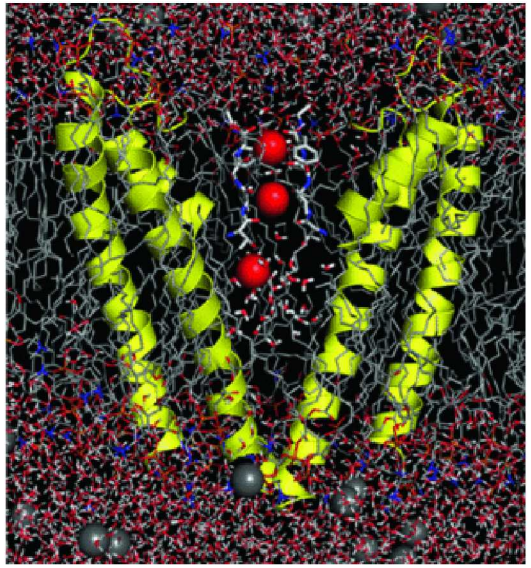
Truman Fellowship

The Grand Challenge of Biophysics

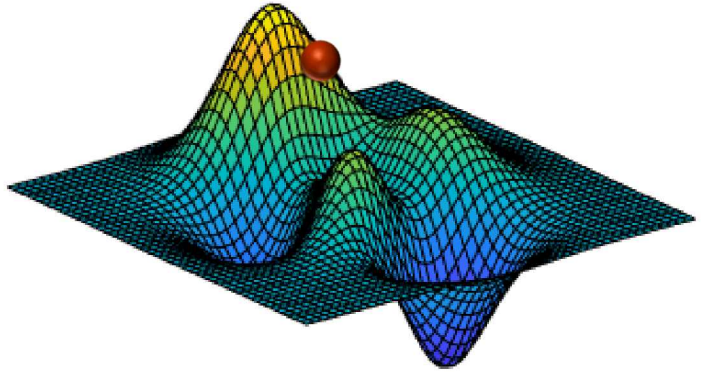
Can we predict the behavior of molecules at the atomic scale?



$$F = ma$$

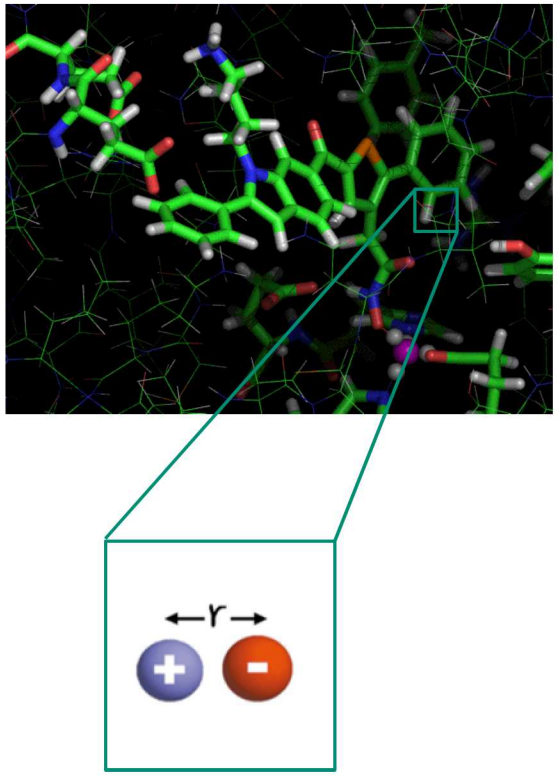


$$H\Psi = E\Psi$$



Why is this a hard problem?

Quantum Mechanics. Even the best QM algorithms scale terribly.

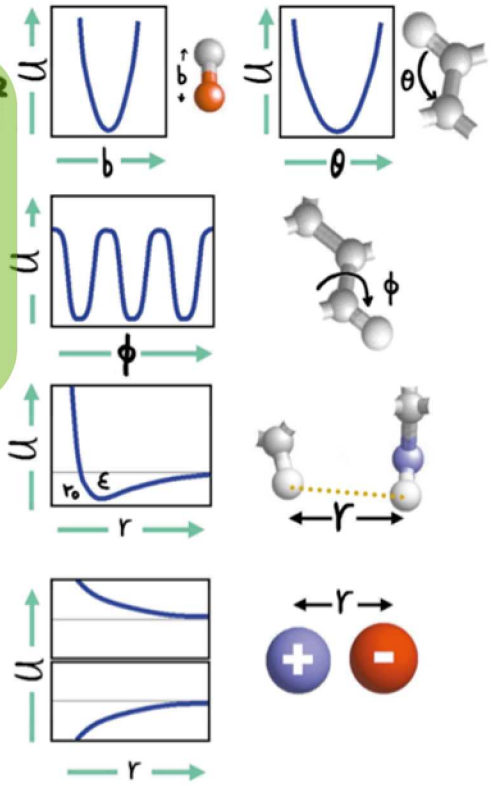


$$U = \sum_{\text{All Bonds}} \tfrac{1}{2}K_b(b-b_0)^2 + \sum_{\text{All Angles}} \tfrac{1}{2}K_\theta(\theta-\theta_0)^2$$

$$+ \sum_{\text{All Torsion Angles}} K_\phi[1-\cos(n\phi+\delta)]$$

$$+ \sum_{\text{All nonbonded pairs}} \varepsilon\left[\left(r_0/r\right)^{12} - 2\left(r_0/r\right)^{6}\right]$$

$$+ \sum_{\text{All partial charges}} 332 q_i q_j / r$$



The Problem.

For far too many applications the point charge force field fails.

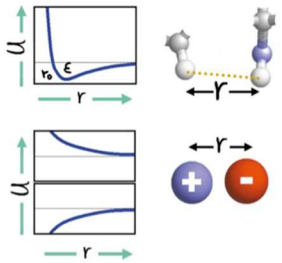
This is not surprising, given that this EXACT functional form was used in the very first MD simulations of a biomolecule, back in 1977

$$+\sum \epsilon\left[\left(r_0/r\right)^{12}-2\left(r_0/r\right)^{6}\right]$$

All nonbonded pairs

$$+\sum 332 q_i q_j / r$$

All partial charges



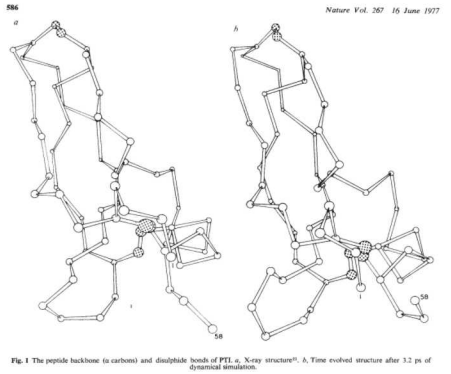
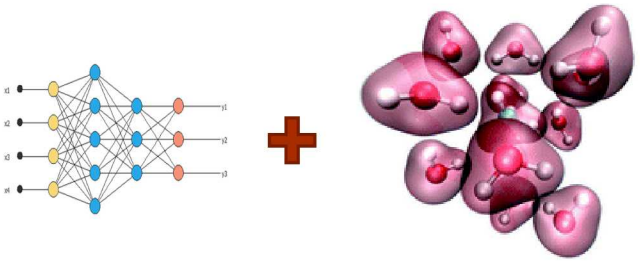


Fig. 1 The peptide backbone (α carbons) and disulphide bonds of PTI. *a*, X-ray structure[10]. *b*, Time evolved structure after 3.2 ps of dynamical simulation.

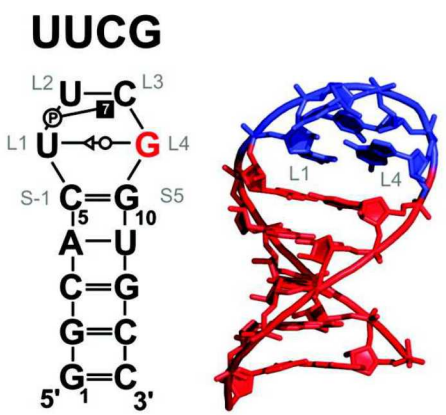
Plan for Today's Talk

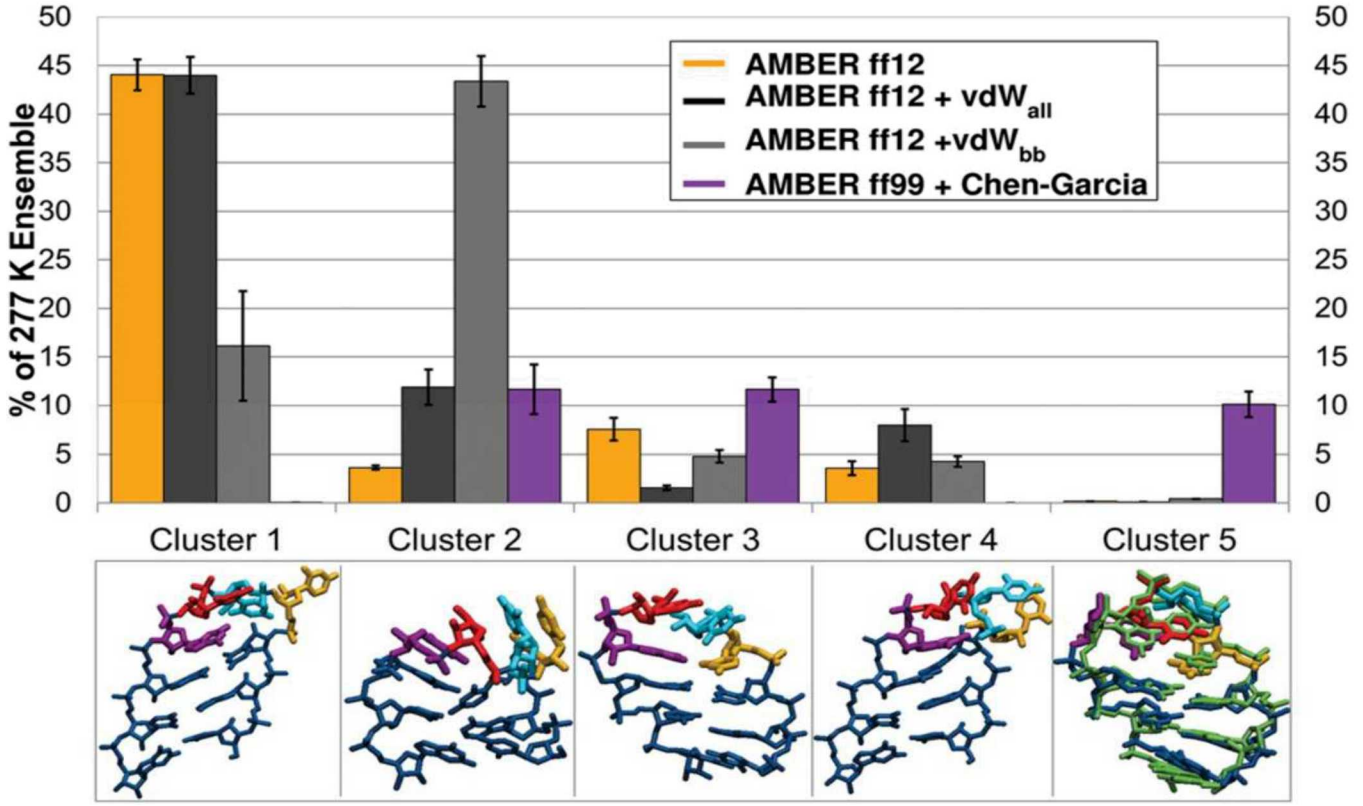
1. Classical: HIPPO (Hydrogen-Like Polarizable Potential)
   A. Why is a new force field necessary?
   B. How HIPPO is different
   C. Simulation results
2. Quantum: *ab initio* Simulations with Deep Learning
   A. OutlookThe importance of the electron density
   B. How Euclidean Deep Neural Networks can help
   C. Trained model results
3. Outlook



The Problem: Point force fields fail for many kinds of biomolecules

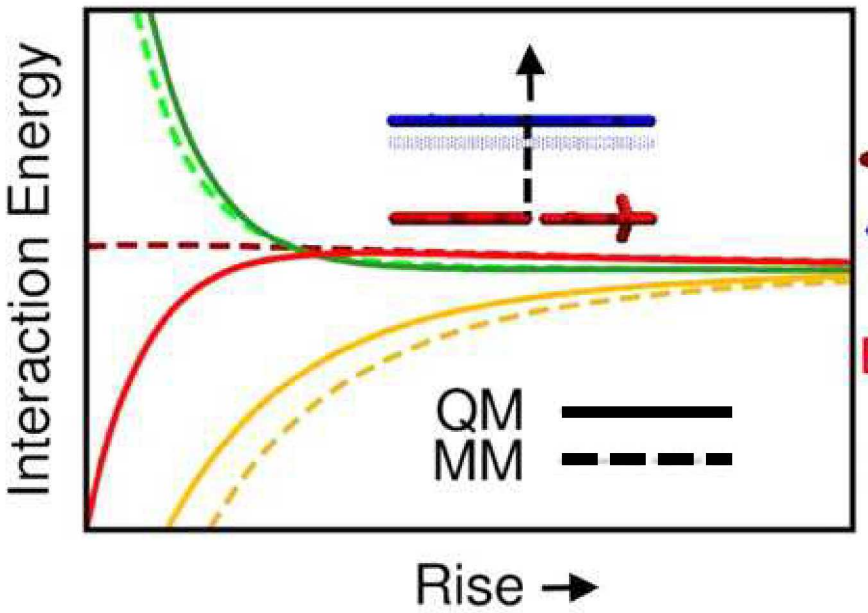
Just one example: Current force fields fail to predict the structure of a simple RNA tetraloop

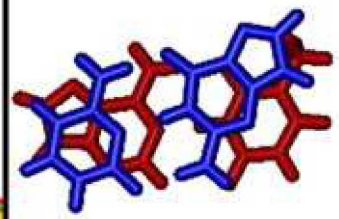




Bergonzo et al., 2015

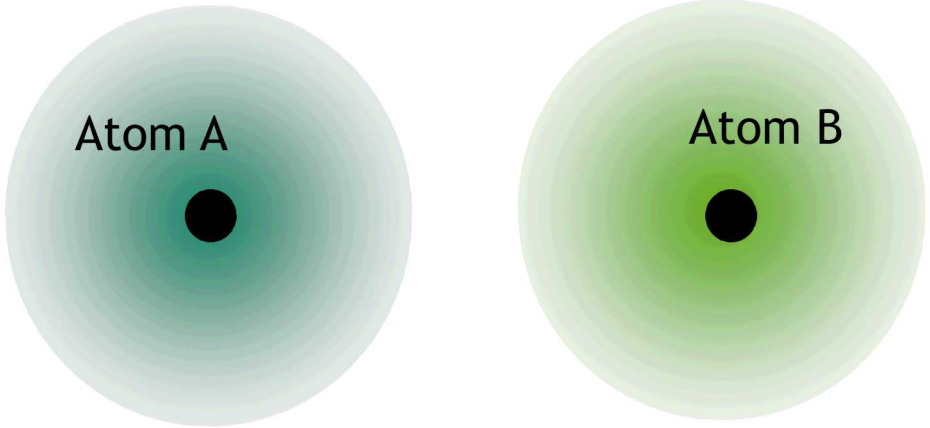
The Problem: Point force fields fail for many kinds of biomolecules





Electrostatics
Exchange
Dispersion

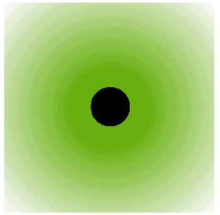
Symmetry Adapted Perturbation Theory (SAPT)



HIPPO: Hydrogen-Like Intermolecular Polarizable Potential

Point Force Fields

HIPPO



$U_{intermolecular}$ =

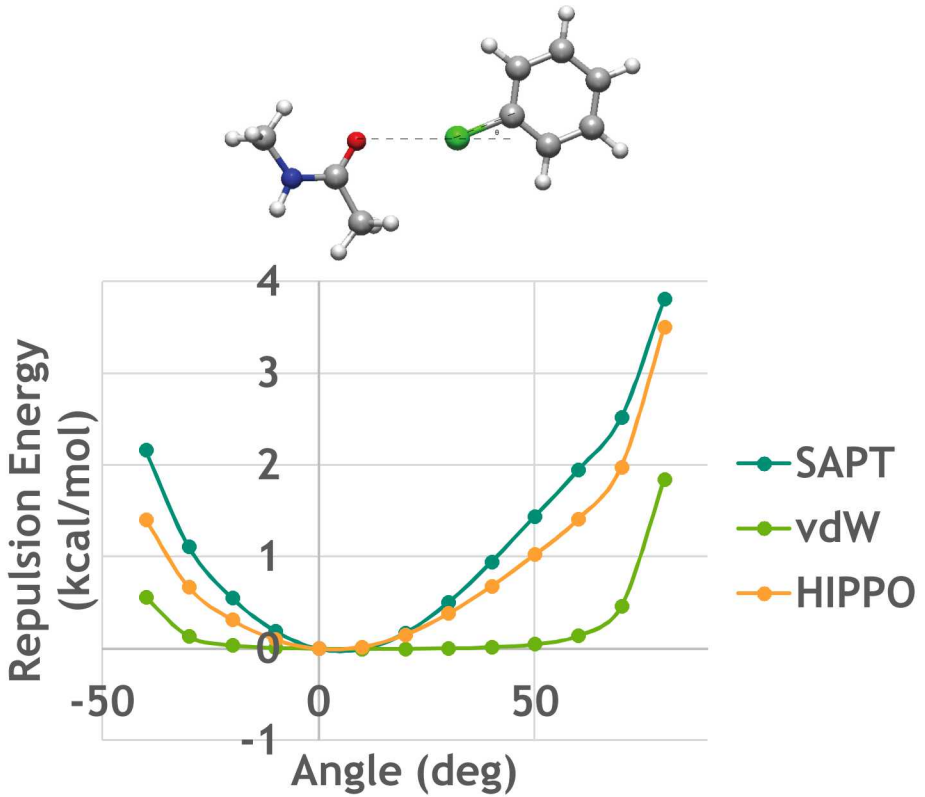
ab initio SAPT

Electrostatics + Polarization +

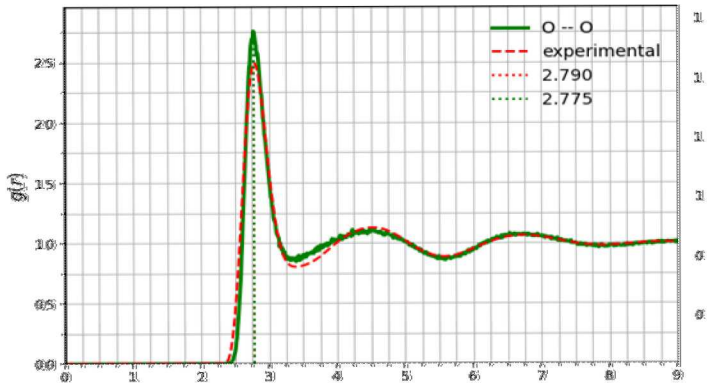
Dispersion + Pauli Repulsion

HIPPO: Hydrogen-Like Intermolecular Polarizable Potential

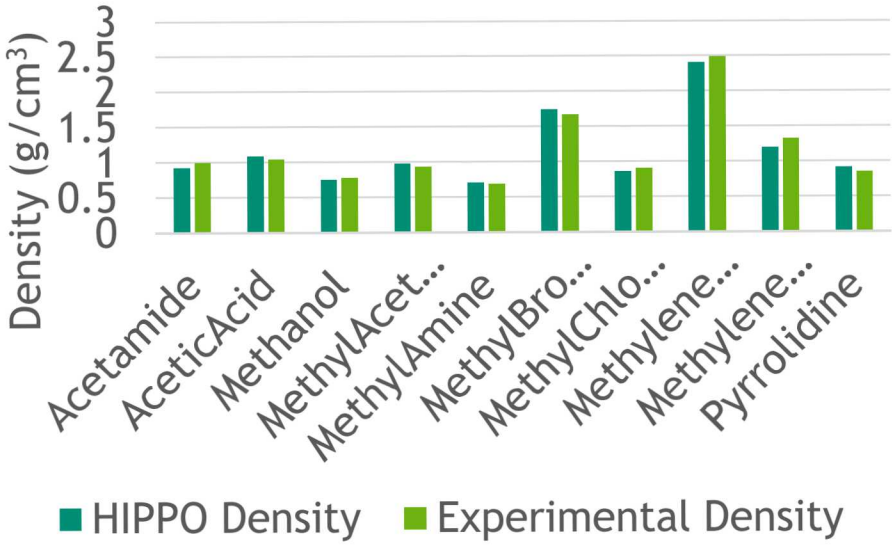
HIPPO is able to reproduce the repulsion energy of drug-like interactions



HIPPO Molecular Dynamics Simulations



	Experiment	HIPPO
Density	0.997 g/cm^3	1.007 g/cm^3
Diffusion Coeff.	2.3 cm^2/s	2.3 cm^2/s
Dielectric	78.3	83.1



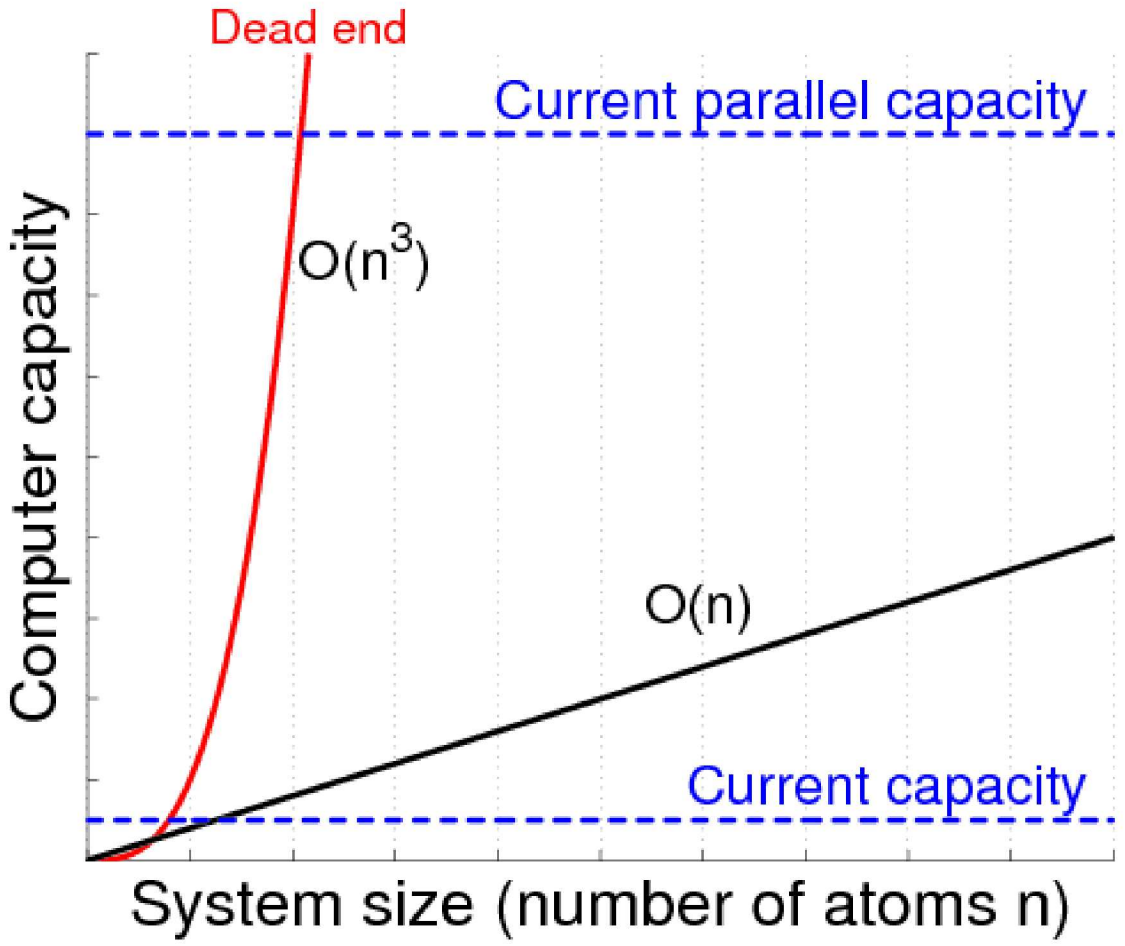
Outlook

- Current: Highly accurate simulations of water, ions and other small molecules
- Future: Turn HIPPO into a full protein, DNA and RNA force field

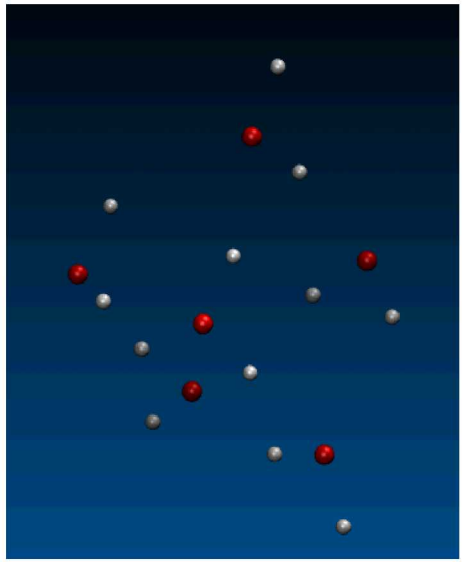
- Find the code at Tinker and OpenMM

Why can't we just do Quantum Mechanics?

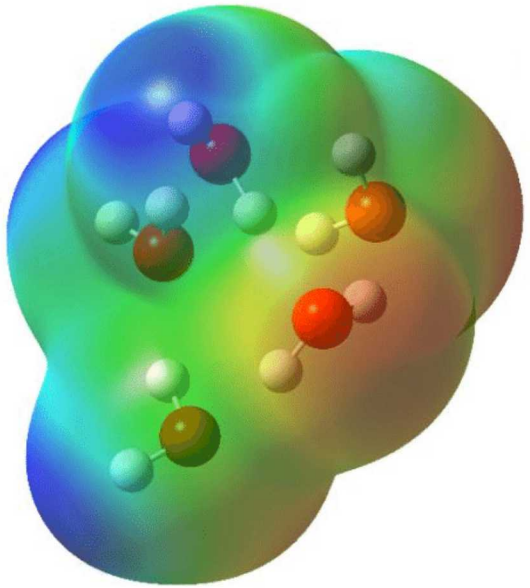
- QM methods scale terribly!
  - Best: N^3
  - Worst: N^7
  - Exact: exponential



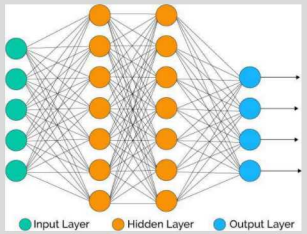
Big idea: Use Deep Learning to predict electron densities



CCSD(T)



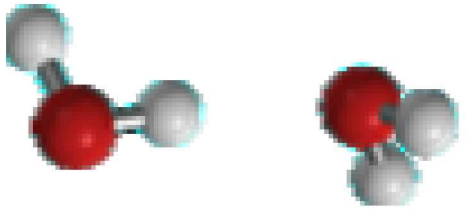
Physics-Informed
Neural Network Model



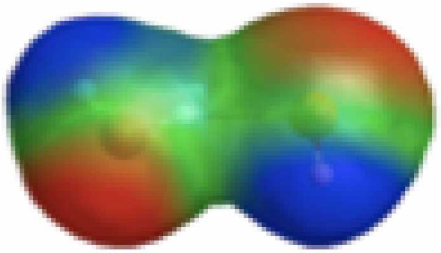
1. The Hellmann-Feynman Theorem

$$H\Psi = E\Psi \qquad E = \langle \Psi^* | H | \Psi \rangle$$

$$\frac{\delta E}{\delta \lambda} = \frac{\delta}{\delta \lambda} \langle \Psi^* | H | \Psi \rangle$$



$$\frac{\delta E}{\delta \lambda} = \left\langle \Psi^* \middle| \frac{\delta H}{\delta \lambda} \middle| \Psi \right\rangle$$



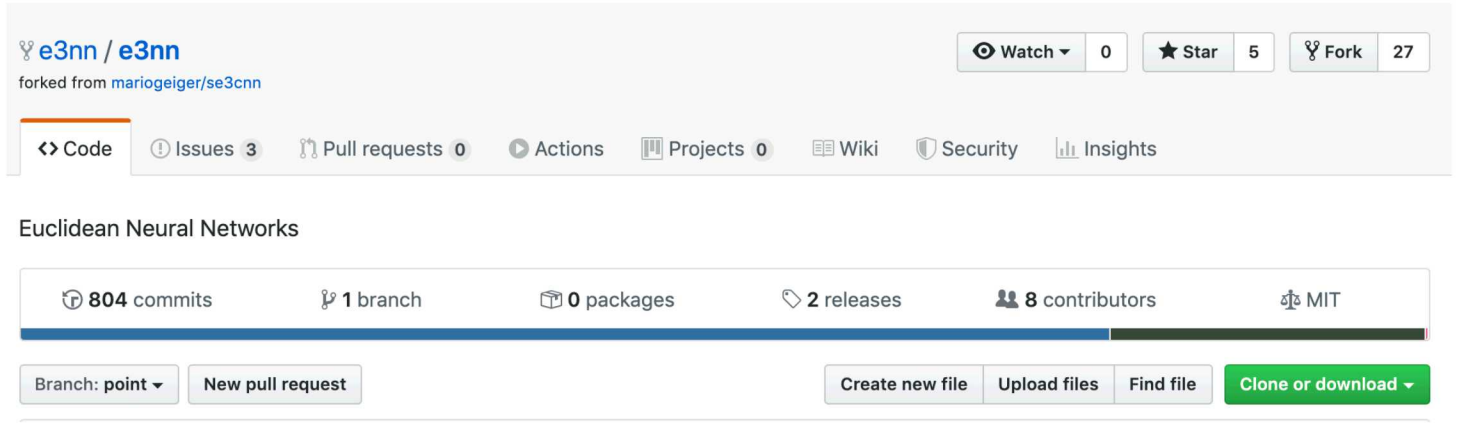
$$\frac{\delta E}{\delta \boldsymbol{R}_A} = \left\langle \Psi^* \middle| \frac{\delta V}{\delta \boldsymbol{R}_A} \middle| \Psi \right\rangle \quad (H = \mathrm{K} + \mathrm{V})$$

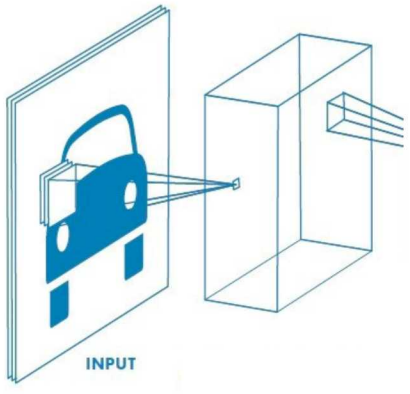
$$\frac{\delta V}{\delta \boldsymbol{R}_A} = \boldsymbol{F}(\boldsymbol{R}_A) = -Z_A \sum \frac{Z_B(\boldsymbol{R}_B - \boldsymbol{R}_A)}{|\boldsymbol{R}_B - \boldsymbol{R}_A|^3} + Z_A \int \frac{\rho(\boldsymbol{r})(\boldsymbol{r} - \boldsymbol{R}_A)d\boldsymbol{r}}{|\boldsymbol{r} - \boldsymbol{R}_A|^3}$$

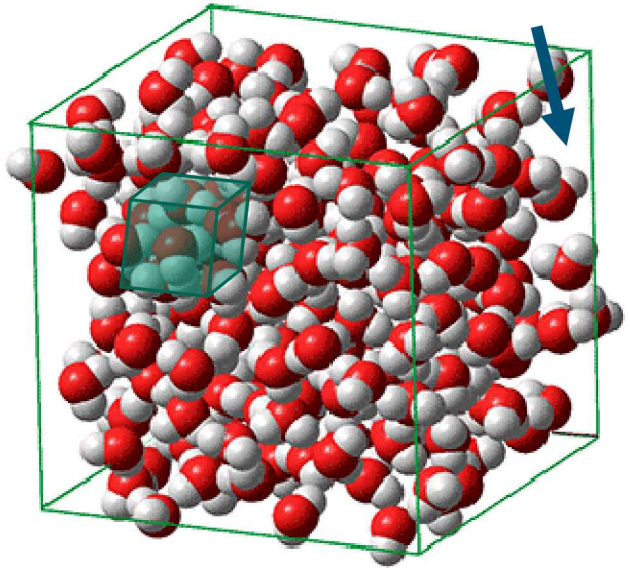
2. Euclidean Neural Networks

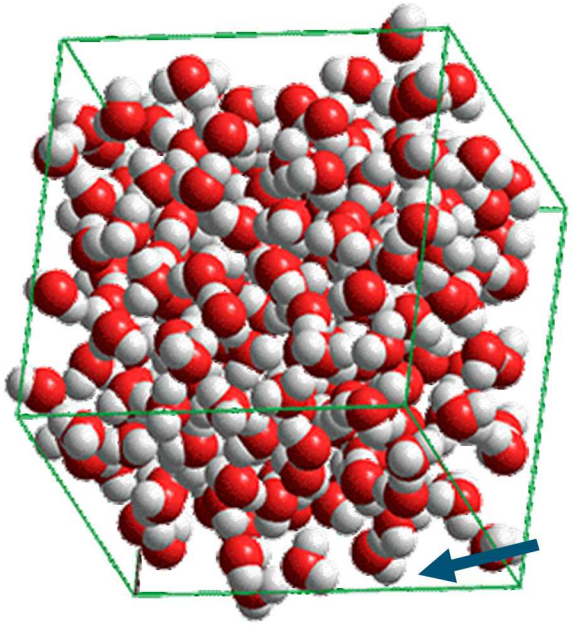
Tess Smidt

Mario Geiger

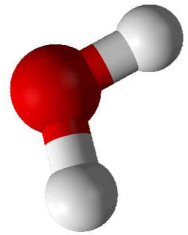








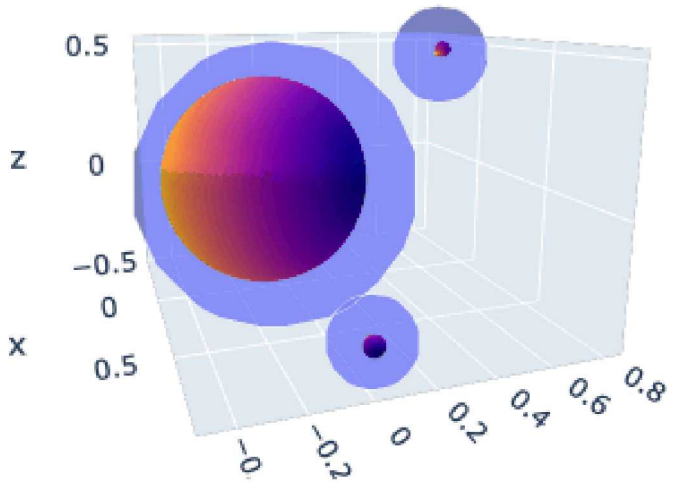
Toy Example: Water Dimer



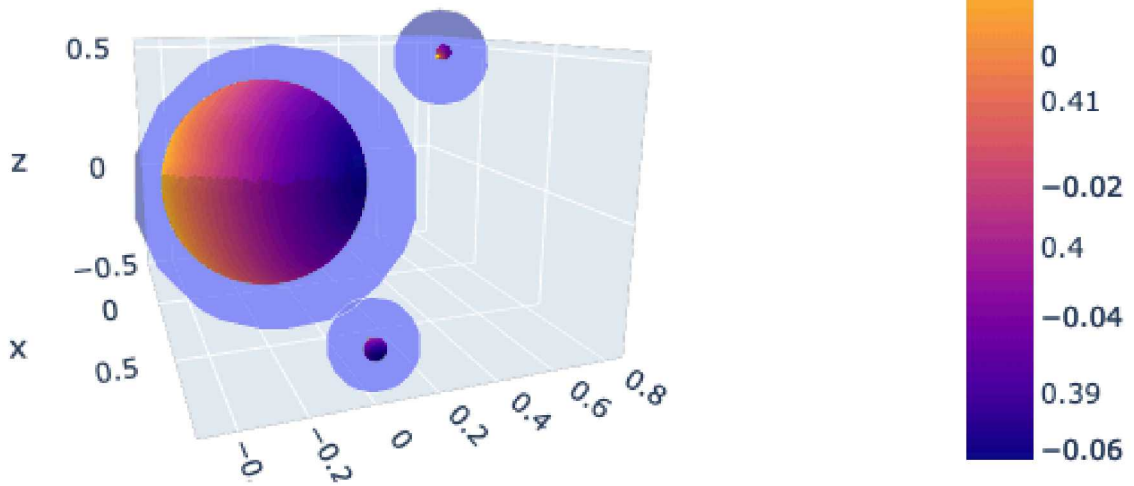
Hartree Fock calculation; 1s + 1p on O and H auxiliary density fitting basis set

1 layer e3nn CNN; Max L = 3

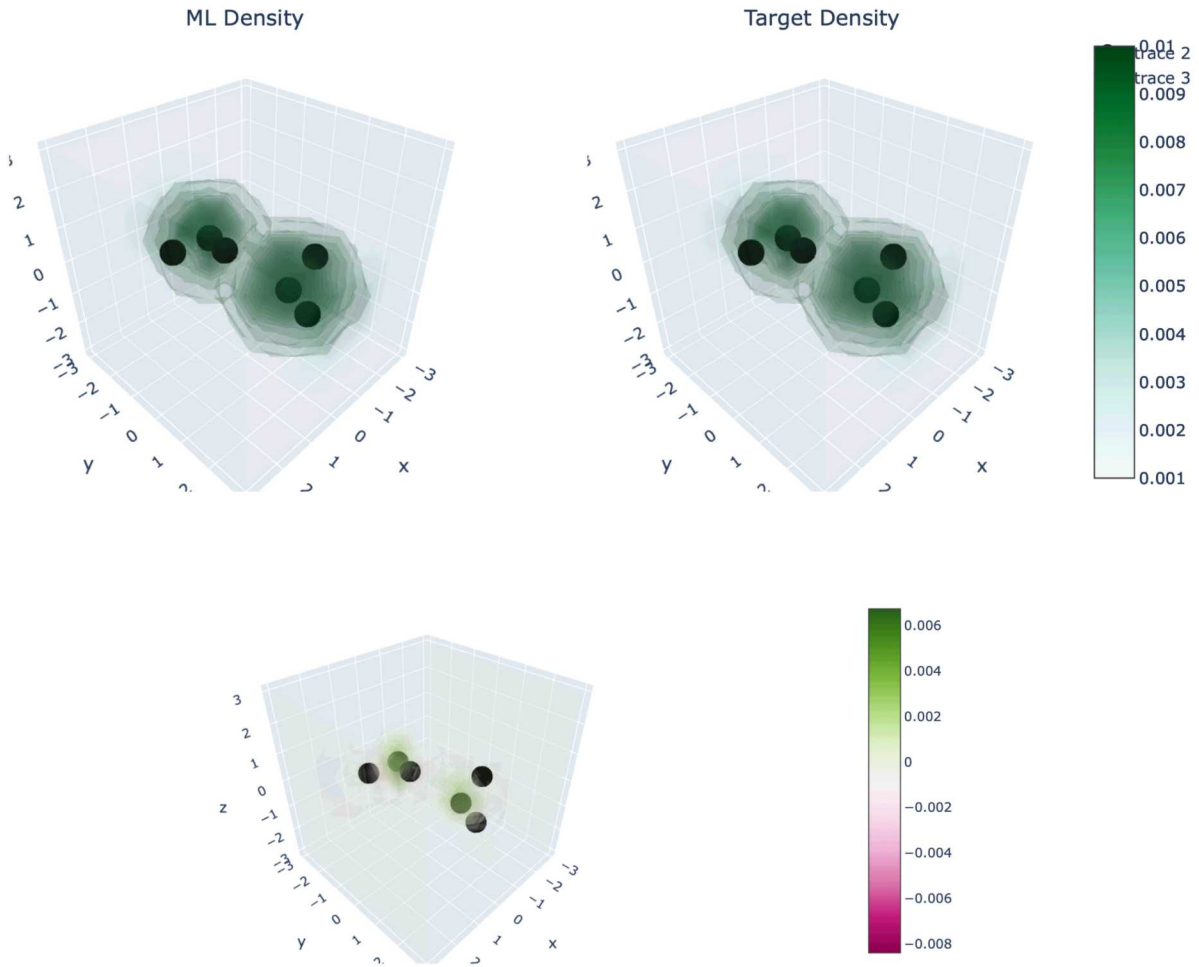
Ground Truth



Trained Model after 100 steps



Results for Water Clusters



Current and Future Work

- Current:
  - Training neural network to predict densities of water clusters.
  - Predicting electron densities of very large water clusters
  - Training neural network on databases of organic molecules (building blocks)

- Future:
  - Forces!
  - With forces, we will be able to do *ab initio* molecular dynamics at the biological scale!

Outlook

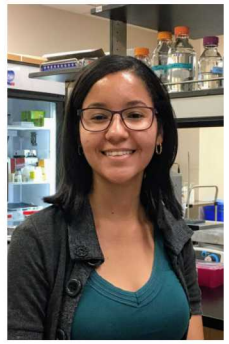
1. A leap in accuracy for biomolecular simulations is available right now!
   A. HIPPO dramatically improves the standard point model
   B. Full biomolecular simulations are in the works!
2. Quantum-accurate simulations are on the horizon
   A. We are leveraging the power of the electron density
   B. And the new technology of Euclidean Neural Networks
   C. The potential result could shatter the "Quantum Scaling Wall"

Acknowledgements

Steve Plimpton

Aidan Thompson

Susan Rempe



Roseane Silva
(Washington University in St. Louis)

Tess Smidt
(Lawrence Berkeley National Lab)

Austin Ellis
(Sandia)

LDRD LABORATORY DIRECTED RESEARCH & DEVELOPMENT