Finding density functionals with ML

Kieron Burke and many friends
UC Irvine
Departments of Physics and of Chemistry
Themes

• DFT (electronic structure) in a nutshell

• Along comes ML

• New opportunities?
Standard atomistic simulations: Molecular dynamics

- Solve Newton’s equations for nuclei at given temperature and pressure.
- Use simple force fields between nuclei.
- Can do a million atoms for many picoseconds.
- But, cannot break bonds!
Electronic Structure Problem: Diversity

- For all everyday matter
Basic Electronic Structure Problem

• Just want $E(R)$
Why electronic structure problem is evil

• Quantum mechanics really needed for electrons

• It’s a many-body problem: Every electron sees every other one, as well as nucleus.

• Chemical accuracy is about 1 in $10^7$ for electronic energy of 500 atoms.

• Ab initio quantum chemistry: When # electrons doubles, computer cost increases by 128.
Mathematical form of problem

- **Differential equation in 3N coordinates:**

\[
\{ \hat{T} + \hat{V}_{ee} + \hat{V} \} \psi = E \psi.
\]

Hamiltonian for \( N \) electrons in the presence of external potential \( v(r) \):

\[
\hat{H} = \hat{T} + \hat{V}_{ee} + \hat{V},
\]

where the kinetic and elec-elec repulsion energies are

\[
\hat{T} = -\frac{1}{2} \sum_{i=1}^{N} \nabla_i^2, \quad \hat{V}_{ee} = \frac{1}{2} \sum_{i=1}^{N} \sum_{j \neq i}^{N} \frac{1}{|r_i - r_j|},
\]

and difference between systems is \( N \) and the one-body potential

\[
\hat{V} = \sum_{i=1}^{N} v(r_i)
\]

Often \( v(r) \) is electron-nucleus attraction

\[
v(r) = -\sum_{\alpha} \frac{Z_\alpha}{|r - R_\alpha|}
\]

where \( \alpha \) runs over all nuclei, plus weak applied \( E \) and \( B \) fields.

\[
E = \min_{\psi} \langle \psi | \hat{T} + \hat{V}_{ee} + \hat{V} | \psi \rangle
\]
The greatest free lunch ever: DFT

• 1964: Hohenberg and Kohn proved a theorem showing lowest energy can be found by search over electronic densities (much simpler than wavefunction)

• 1965: Created Kohn-Sham (KS) equations of fake non-interacting electrons (not many-body anymore) which, when solved, yield lowest E and density alone.
KS equations (1965)

Define *fictitious* non-interacting electrons satisfying:

\[
\begin{cases}
-\frac{1}{2}\nabla^2 + v_s(r) \phi_j(r) = \epsilon_j \phi_j(r), \\
\sum_{j=1}^{N} |\phi_j(r)|^2 = n(r).
\end{cases}
\]

where \(v_s(r)\) is *defined* to yield \(n(r)\).

Define \(T_s\) as the kinetic energy of the KS electrons, \(U\) as their Hartree energy and

\[
F = T + V_{ee} = T_s + U + E_{XC}
\]

the remainder is the exchange-correlation energy.

Most important result of exact DFT:

\[
v_s(r) = v(r) + \int d^3r \frac{n(r')}{|r - r'|} + v_{XC}[n](r),
\]

Knowing \(E_{XC}[n]\) gives closed set of self-consistent equations.
The kicker

• Need a formula for a small fraction of electronic energy, called XC energy, in terms of density, containing all quantum many-body effects.

• First formula (1965) good for solids, but not accurate enough for chemistry.

• Next formulas (1990) give useful accuracy for chemistry and materials.
Today’s commonly-used XC approximations

- **Local density approximation (LDA)**
  - Uses only \( n(r) \) at a point.

- **Generalized gradient approx (GGA)**
  - Uses both \( n(r) \) and \( |\nabla n(r)| \)
  - Should be more accurate, corrects overbinding of LDA
  - Examples are PBE and BLYP

- **Hybrid:**
  - Mixes some fraction of HF with GGA
  - Examples are B3LYP and PBE0

\[
E_{\text{LDA}}^L(n) = A_x \int d^3r \ n^{4/3}(r)
\]

\[
E_{\text{GGA}}^{\text{XC}} = \int d^3r \ e_{\text{XC}}^{\text{GGA}}(n(r), |\nabla n(r)|)
\]

\[
E_{\text{hyb}}^{\text{XC}} = a \left( E_X - E_X^{\text{GGA}} \right) + E_{\text{XC}}^{\text{GGA}}
\]
Modern research to find XC energy
Youtube: Teaching the theory in DFT

- Hohenberg-Kohn theorem and constrained search
- Local Density Approximation
- OEP method and hybrids
- Time-dependent density functional theory
Alphabet soup
Applications

- Computers, codes, algorithms always improving
- Making bona fide predictions
- E.g., a new better catalyst for Haber-Bosch process (‘fixing’ ammonia from air) was predicted after about 25,000 failed experiments (Norskov’s group)
- Now scanning chemical and materials spaces using big data methods for materials design (materials genome project).
- World’s hottest superconductor (203K) is hydrogen sulfide, predicted by DFT calculations, then made.
- Latest generation of intel chips (needed for Mac airbook) is half-size and Pb-free with help of DFT.
Route to a Superconducting Phase above Room Temperature in Electron-Doped Hydride Compounds under High Pressure

Ying Sun,1 Jian Lv,1 Yu Xie,1 Hanyu Liu,1,* and Yanming Ma1,2,*

1Innovation Center of Computational Physics Methods and Software and State Key Laboratory for Superhard Materials, College of Physics, Jilin University, Changchun 130012, China
2International Center of Future Science, Jilin University, Changchun 130012, China

(Received 24 December 2018; published 26 August 2019)

The recent theory-orientated discovery of record high-temperature superconductivity ($T_c \sim 250$ K) in sodalite-like clathrate LaH$_{10}$ is an important advance toward room-temperature superconductors. Here, we identify an alternative clathrate structure in ternary Li$_2$MgH$_{16}$ with a remarkably high estimated $T_c$ of $\sim 473$ K at 250 GPa, which may allow us to obtain room-temperature or even higher-temperature superconductivity. The ternary compound mimics a Li- or electron-doped binary hydride of MgH$_{16}$. The parent hydride contains H$_2$ molecules and is not a good superconductor. The extra electrons introduced break up the H$_2$ molecules, increasing the amount of atomic hydrogen compared with the parent hydride, which is necessary for stabilizing the clathrate structure or other high-$T_c$ structures. Our results provide a viable strategy for tuning the superconductivity of hydrogen-rich hydrides by donating electrons to hydrides via metal doping. Our approach may pave the way for finding high-$T_c$ superconductors in a variety of ternary or quaternary hydrides.

DOI: 10.1103/PhysRevLett.123.097001
Projector-Based Embedding Eliminates Density Dependence for QM/MM Calculations of Reactions in Solution

Kara E. Ranaghan, Darya Shchepeanovska, Simon J. Bennie, Narin Jolanta Zurek, Frederick R. Manby, and Adrian J. Mulholland

Centre for Computational Chemistry, School of Chemistry, University of Bristol, Bristol, U.K. BS8 1TS

Supporting Information

ABSTRACT: Combined quantum mechanics/molecular mechanics (QM/MM) methods are increasingly widely utilized in studies of reactions in enzymes and other large systems. Here, we apply a range of QM/MM methods to investigate the Claisen rearrangement of chorismate to prephenate, in solution, and in the enzyme chorismate mutase. Using projector-based embedding in a QM/MM framework, we apply treatments up to the CCSD(T) level. We test a range of density functional QM/MM methods and QM region sizes. The results show that the calculated reaction energetics are significantly more sensitive to the choice of density functional than they are to the size of the QM region in these systems. Projector-based embedding of a wave function method in DFT reduced the 13 kcal/mol spread in barrier heights calculated at the DFT/MM level to a spread of just 0.3 kcal/mol, essentially eliminating the dependence on the functional. Projector-based embedding of correlated ab initio methods provides a practical method for treating reactions in condensed phases.
Electronic Structure Problem: Impact

- Absurdly useful: Nvidia estimates 1/6th world’s supercomputers run DFT calculations.
Difficulties with this research

• Truly multidisciplinary in a deep way – very hard to understand each other’s methods and goals

• Not the latest shiny toy, e.g., not quantum computing or machine learning

• But just a little progress in this field can have enormous impact overnight.
Paying the price of the KS equations

- By using fake electrons, cost increases by factor of 8 when size doubles.
- Limits current sizes to 500 atoms on a single machine in a day.
- If we could avoid this, might be able to do $10^6$ atoms.
- Again, a rule exists, $T_S[n]$, and some folks try to find it, but never accurately and generally enough.
Machine learning
Machine learning in electronic structure

- Explosion of interest in last 5 years
- Machine learning/big data/data science very broad terms
- Some examples:
  - Searching databases of materials calculations to find optimal functionality
  - Searching chemical compound space (SCHNET, ANI,..)
  - Accelerated sampling
  - Designing interatomic potentials (Behler, Csanyi,..)
Original team for ML DFT (2010)

- Most with Klaus Mueller of TU Berlin, computer science.
- ML now being applied directly to, e.g., molecular energies from geometries for drug design, many by Matthias Rupp
- Our efforts are focused on finding $T_{s[n]}$ from examples, work by John Snyder (Humboldt fellow at TU Berlin/MPI Halle)
Demo problem in DFT

- \( N \) non-interacting same-spin fermions confined to 1d box

- Define class of potential:
  \[
  v(x) = - \sum_{i=1}^{3} a_i \exp\left(-\frac{(x - b_i)^2}{2c_i^2}\right)
  \]

- Represent the density on a grid with spacing \( \Delta x = 1/(G - 1) \)

- ML-DFA for KE:
  \[
  \hat{T}(n) = \bar{T} \sum_{j=1}^{M} \alpha_j k(n_j, n)
  \]
  \[
  k[n,n'] = \exp \left( - \int dx \left( n(x) - n'(x) \right)^2 / (2\sigma^2) \right)
  \]
Performance for $T_s$

| $N$ | $M$    | $\lambda$   | $\sigma$ | $|\Delta T|$ | $|\Delta T|_{\text{std}}$ | $|\Delta T|_{\text{max}}$ |
|-----|--------|-------------|----------|--------------|---------------------------|--------------------------|
|     | 40     | $2.4 \times 10^{-5}$ | 238      | 3.3          | 3.0                       | 23.                      |
|     | 60     | $1.0 \times 10^{-5}$ | 95       | 1.2          | 1.2                       | 10.                      |
| 1   | 80     | $6.7 \times 10^{-6}$ | 48       | 0.43         | 0.54                      | 7.1                      |
|     | 100    | $3.4 \times 10^{-7}$ | 43       | 0.15         | 0.24                      | 3.2                      |
|     | 150    | $2.5 \times 10^{-7}$ | 33       | 0.060        | 0.10                      | 1.3                      |
|     | 200    | $1.7 \times 10^{-7}$ | 28       | 0.031        | 0.053                     | 0.65                     |
| 2   | 100    | $1.3 \times 10^{-7}$ | 52       | 0.13         | 0.20                      | 1.8                      |
| 3   | 100    | $2.0 \times 10^{-7}$ | 74       | 0.12         | 0.18                      | 1.8                      |
| 4   | 100    | $1.4 \times 10^{-7}$ | 73       | 0.078        | 0.14                      | 2.3                      |
| 1-4†| 400    | $1.8 \times 10^{-7}$ | 47       | 0.12         | 0.20                      | 3.6                      |

LDA $\sim 223$ kcal/mol, Gradient correction $\sim 159$ kcal/mol
We don’t just need the energy

- The KS equations are solving the following equation for us:

\[
\frac{\delta T_S}{\delta n(r)} = -v(r) - v_H[n](r) - v_{XC}[n](r)
\]

- If we had an explicit approximation for \(T_S[n]\), we could solve this directly.
functional derivative?

\[
\frac{\delta T[n]}{\delta n(x)} = \mu - v(x) \quad \xrightarrow{\text{ML-DFA}} \quad \frac{1}{\Delta x} \nabla_n \hat{T}(n) = \sum_{j=1}^{M} \alpha'_j (n_j - n) k(n_j, n) \\
\alpha'_j = \alpha_j / (\sigma^2 \Delta x)
\]

- Functionals are defined on infinite-dimensional spaces
- With finite interpolation, can always find bad directions
- Can we make a cruder definition that will work for our purposes?
Principal component analysis

\[ X = (n_{j_1} - n, \ldots, n_{j_m} - n) \]  
\[ C = \frac{1}{m} X^T X \]  
\[ \lambda_j, x_j \]  
\[ P_{m,\ell}(n) = V^T V \]  
\[ V = (x_1, \ldots, x_\ell)^T \]
Lessons

• Exact noise-free data infinitely available for \( T_s[n] \), every cycle of every KS calculation in the world provides examples.

• Need very accurate derivatives to get accurate density from Euler equation.

• Can find ways to bypass this.

• Functionals can be made arbitrarily accurate with sufficient data.
Road map back to reality

Roadmap to 3d land

1d box
- model selection, projected functional derivatives, OF-DFT
- bond breaking, self-consistent densities

1d diatomics
- dimensionality, basis sets, representation, inversion symmetry

3d atoms, diatomics
- full symmetries, scaling
- scalability, data accumulation
- ab-initio MD, active learning

3d molecules

large systems, real applications

Kieron Burke
IPAM physics and ML 2019
By-passing KS

\[ \frac{\delta T_s}{\delta n(r)} + v_s(r) = 0 \]

Euler equation

Kohn-Sham (KS) Mapping

Hohenberg-Kohn (HK) Mapping

Density

Potential

Total Energy

\[ E_v \]

\[ E[n] \]

\[ E[v] \]
MD simulations testing ML method

Figure 3. Energy errors of ML-HK along MD trajectories. PBE values in blue, ML-HK values in red. 

<table>
<thead>
<tr>
<th></th>
<th>Benzene</th>
<th>Ethane</th>
<th>Malonaldehyde</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MAE max</td>
<td>MAE max</td>
<td>MAE max</td>
</tr>
<tr>
<td>Training trajectories</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>300K</td>
<td>0.395742</td>
<td>1.92642</td>
<td>0.212117</td>
</tr>
<tr>
<td>300K + 350K</td>
<td>0.265417</td>
<td>1.76190</td>
<td>0.236088</td>
</tr>
<tr>
<td>300K + 400K</td>
<td>0.370876</td>
<td>2.1162</td>
<td>0.101054</td>
</tr>
</tbody>
</table>

Table V. Errors ($\Delta E_D$ in kcal/mol) on the MD datasets for different training trajectory combinations.
By-passing the Kohn-Sham equations with machine learning

By passing the Kohn-Sham equations with machine learning Felix Brockherde, Leslie Vogt, Li Li, Mark E Tuckerman, Kieron Burke, Klaus-Robert Müller, Nature Communications 8, 872 (2017).
Lessons

• Our 1d gradient methods become prohibitively expensive in 3d.
• Instead of using $T_s[n]$, learn $n[v](r)$.
• Much smarter than learning $E[v_s]$
• Works for $H_2$ and $H_2O$ and …
• ..MD of malonaldehyde using ML forces with Leslie Vogt and Mark Tuckerman.
• Created non-local orbital-free density functional running much faster than solving KS equations
Our papers (all on dft.uci.edu)


**Understanding machine-learned density functionals** Li Li, John C. Snyder, Isabelle M. Pelaschier, Jessica Huang, Uma-Naresh Niranjan, Paul Duncan, Matthias Rupp, Klaus-Robert Müller, Kieron Burke, *International Journal of Quantum Chemistry* n/a--n/a (2015).


Our latest trick
Theory

\[ E = E^{\text{DFT}}[n^{\text{DFT}}] + \Delta E[n^{\text{DFT}}], \]

\[ n(r) = \frac{\delta E[v]}{\delta v(r)} \]
Figure 1: Illustration of density-based machine learning for water conformer energies. For all panels, DFT energies (orange) are shown alongside CC energies (blue) for the same molecular conformers, with optimized geometries indicated by open diamonds. 

a) The nuclear potential, represented by an approximate Gaussian potential, is the input to a set of ML models that return the electron density. This learned density is the input for independent ML predictions of molecular energies based on DFT or CC electronic structure calculations.

b) Calculated energies for CC (dark blue) and DFT (dark orange) for 102 sample geometries relative to the lowest training energy (top), along with the relative energy errors for DFT compared to CC for each conformer (bottom).

c) Average out-of-sample prediction errors for the different ML functionals compared to the reference CC energies. The MAE of the DFT energies w.r.t. CC is also shown as a dashed line.

d) The potential energy surface of symmetric water geometries for DFT ML (orange) and CC - DFT (blue) after applying the DFT correction (bottom). For this figure, all DFT calculations use the PBE functional and all CC calculations use CCSD(T) (see Methods for more details).

Moreover, using the results from the Supp Info, one can in principle even construct the exact density from a sequence of such calculations.
Range of PBE errors
Learning curves

Figure 1: Illustration of density-based machine learning for water conformer energies. For all panels, DFT energies (orange) are shown alongside CC energies (blue) for the same molecular conformers, with optimized geometries indicated by open diamonds.

(a) The nuclear potential, represented by an approximate Gaussians potential, is the input to a set of ML models that return the electron density. This learned density is the input for independent ML predictions of molecular energies based on DFT or CC electronic structure calculations.

(b) Calculated energies for CC (dark blue) and DFT (dark orange) for 102 sample geometries relative to the lowest training energy (top), along with the relative energy errors for DFT compared to CC for each conformer (bottom).

(c) Average out-of-sample prediction errors for the different ML functionals compared to the reference E\text{CC} energies. The MAE of the E\text{DFT} energies w.r.t. E\text{CC} is also shown as a dashed line.

(d) The potential energy surface of symmetric water geometries for E\text{DFT} ML (orange) and E\text{CC} - DFT (blue) after applying the -DFT correction (bottom). For this figure, all DFT calculations use the PBE functional and all CC calculations use CCSD(T) (see Methods for more details).
Figure 1: Illustration of density-based machine learning for water conformer energies. For all panels, DFT energies (orange) are shown alongside CC energies (blue) for the same molecular conformers, with optimized geometries indicated by open diamonds. 

a) The nuclear potential, represented by an approximate Gaussian potential, is the input to a set of ML models that return the electron density. This learned density is the input for independent ML predictions of molecular energies based on DFT or CC electronic structure calculations.

b) Calculated energies for CC (dark blue) and DFT (dark orange) for 102 sample geometries relative to the lowest training energy (top), along with the relative energy errors for DFT compared to CC for each conformer (bottom).

c) Average out-of-sample prediction errors for the different ML functionals compared to the reference CC energies. The MAE of the DFT energies w.r.t. CC energies is also shown as a dashed line.

d) The potential energy surface of symmetric water geometries for DFT ML (orange) and CC - DFT (blue) after applying the -DFT correction (bottom). For this figure, all DFT calculations use the PBE functional and all CC calculations use CCSD(T) (see Methods for more details).

Moreover, using the results from the Supp Info, one can in principle even construct the exact density from a sequence of such calculations.
Learning curves

Figure 1: Illustration of density-based machine learning for water conformer energies. For all panels, DFT energies (orange) are shown alongside CC energies (blue) for the same molecular conformers, with optimized geometries indicated by open diamonds.

- a) The nuclear potential, represented by an approximate Gaussians potential, is the input to a set of ML models that return the electron density $n_{ML}$. This learned density is the input for independent ML predictions of molecular energies based on DFT or CC electronic structure calculations.
- b) Calculated energies for CC (dark blue) and DFT (dark orange) for 102 sample geometries relative to the lowest training energy (top), along with the relative energy errors for DFT compared to CC for each conformer (bottom).
- c) Average out-of-sample prediction errors for the different ML functionals compared to the reference CC energies.
- The MAE of the DFT energies w.r.t. CC energies is also shown as a dashed line.
- d) The potential energy surface of symmetric water geometries for DFT ML (orange) and CC - DFT (blue) after applying the -$\Delta$-DFT correction (bottom). For this figure, all DFT calculations use the PBE functional and all CC calculations use CCSD(T) (see Methods for more details).

Moreover, using the results from the Supp Info, one can in principle even construct the exact density from a sequence of such calculations.
Figure 1: Illustration of density-based machine learning for water conformer energies. For all panels, DFT energies (orange) are shown alongside CC energies (blue) for the same molecular conformers, with optimized geometries indicated by open diamonds. 

a) The nuclear potential, represented by an approximate Gaussians potential, is the input to a set of ML models that return the electron density. This learned density is the input for independent ML predictions of molecular energies based on DFT or CC electronic structure calculations. 

b) Calculated energies for CC (dark blue) and DFT (dark orange) for 102 sample geometries relative to the lowest training energy (top), along with the relative energy errors for DFT compared to CC for each conformer (bottom). 

c) Average out-of-sample prediction errors for the different ML functionals compared to the reference $E_{CC}$ energies. The MAE of the $E_{DFT}$ energies w.r.t. $E_{CC}$ is also shown as a dashed line. 

d) The potential energy surface of symmetric water geometries for $E_{DFT}^{ML}$ (orange) and $E_{CC} - DFT$ (blue) after applying the $-DFT$ correction (bottom). For this figure, all DFT calculations use the PBE functional and all CC calculations use CCSD(T) (see Methods for more details).
Learning curves

- Figure 1: Illustration of density-based machine learning for water conformer energies. For all panels, DFT energies (orange) are shown alongside CC energies (blue) for the same molecular conformers, with optimized geometries indicated by open diamonds.
  - a) The nuclear potential, represented by an approximate Gaussian potential, is the input to a set of ML models that return the electron density $\rho_{\text{ML}}$. This learned density is the input for independent ML predictions of molecular energies based on DFT or CC electronic structure calculations.
  - b) Calculated energies for CC (dark blue) and DFT (dark orange) for 102 sample geometries relative to the lowest training energy (top), along with the relative energy errors for DFT compared to CC for each conformer (bottom).
  - c) Average out-of-sample prediction errors for the different ML functionals compared to the reference $E_{\text{CC}}$ energies. The MAE of the $E_{\text{DFT}}$ energies w.r.t. $E_{\text{CC}}$ is also shown as a dashed line.
  - d) The potential energy surface of symmetric water geometries for $E_{\text{DFT}}$ ML (orange) and $E_{\text{CC}} - E_{\text{DFT}}$ (blue) after applying the DFT correction (bottom). For this figure, all DFT calculations use the PBE functional and all CC calculations use CCSD(T) (see Methods for more details).
Potential energy surfaces

Figure 1: Illustration of density-based machine learning for water conformer energies. For all panels, DFT energies (orange) are shown alongside CC energies (blue) for the same molecular conformers, with optimized geometries indicated by open diamonds.

- a) The nuclear potential, represented by an approximate Gaussians potential, is the input to a set of ML models that return the electron density. This learned density is the input for independent ML predictions of molecular energies based on DFT or CC electronic structure calculations.

- b) Calculated energies for CC (dark blue) and DFT (dark orange) for 102 sample geometries relative to the lowest training energy (top), along with the relative energy errors for DFT compared to CC for each conformer (bottom).

- c) Average out-of-sample prediction errors for the different ML functionals compared to the reference E\text{CC} energies. The MAE of the E\text{DFT} energies w.r.t. E\text{CC} is also shown as a dashed line.

- d) The potential energy surface of symmetric water geometries for E\text{DFT} ML (orange) and E\text{CC} - DFT (blue) after applying the the -DFT correction (bottom). For this figure, all DFT calculations use the PBE functional and all CC calculations use CCSD(T) (see Methods for more details).
Resorcinol
Figure 3: Resorcinol dynamics from an initial condition near a conformational change showing a) the atomic positions explored during 100 fs NVE MD trajectories run with standard DFT (dark orange), RESPA-corrected forces (light blue), and $E_{\text{CC}}^{\text{DFT}}[n_{\text{DFT}} \text{sML}]$ (blue), b) the conformer energy along each trajectory (solid lines), with the error relative to CC shown as a shaded line width, and c) the evolution of the C-C-O-H dihedral angle for each trajectory with dashed grey lines indicating the barrier between conformers. For this figure, all DFT calculations use PBE and all CC energies are from CCSD(T).
Opportunities for ML in physics using DFT

DFT is a working procedure for coarse-graining!
But it needs good functional approximations.
Classical DFT – faster than MD

• Can prove HK theorem for equilibrium classical statistical mechanics (Bob Evans, *Advances in Physics* 28 (2), 143-200, 1979.)

• Thus can solve self-consistent equations directly for equilibrium density.

• But every liquid has a different interaction and so a different functional.

• Can look at wetting, phase transitions, nucleation, etc.

• But often very crude for realistic systems due to crude functional approximations.
Published in: Robert Evans; Daan Frenkel; Marjolein Dijkstra; Physics Today 72, 38-39 (2019)
DOI: 10.1063/PT.3.4135
Copyright © 2019 American Institute of Physics
Density-functional fluctuation theory of crowds

J. Felipe Méndez-Valderrama, Yunus A. Kinkhabwala, Jeffrey Silver, Itai Cohen & T.A. Arias

Extraction of functionals for model system of walking flies. To test whether this approach applies to actual populations, we consider a model crowd consisting of wild-type male *Drosophila melanogaster* from an out-bred laboratory stock. It is well known that flies exhibit complex spatial preferences and social behaviors. Here we seek to determine whether a large crowd of individuals with such complex behaviors indeed can be described within our vexation and frustration framework. The flies are confined in 1.5 mm tall transparent chambers where they can walk freely but cannot fly or climb on top of each other. We record overhead videos of the flies, bin the arena, and use custom Matlab-based tracking algorithms (Methods) to measure the individual bin counts \( N_b \) in each video frame. To explore a variety of behaviors, we use arenas of different shapes and apply heat gradients across the arenas to generate different spatial preferences. We find that the flies fully adjust to such changes in their environments after 5 min. We also find that the behavior of the flies changes slowly over a time scale of hours (Methods). We thus take care to make our observations over 10 minute windows during time periods where the behavior is stable.
DFT of nuclear forces

Density functional theory (DFT) in nuclear physics and astrophysics (I)

G. Colò

IPNL
February 2015
Warm dense matter

- Controlled fusion at NIF
- Planetary interiors
- Matter under extreme temperature and pressure
- Take theorems of DFT and heat up
- At high enough temperatures, electrons behave classically

Impacts of Quantum Chemistry Calculations on Exoplanetary Science, Planetary Astronomy, and Astrophysics

A white paper submitted in response to Astro2020 call to Astronomy and Astrophysics

Der-you Kao1,2*, Marko Gacesa1,3, Renata M. Wentzcovitch4, Shawn Domagal-Goldman1, Ravi K. Kopparapu1, Stephen J. Klippenstein5, Steven B. Charnley1, Wade G. Henning6, Joe Renaud7, Paul Romani9, Yuni Lee1,10, Conor A. Nixon1, Koblar A. Jackson8, Martin A. Cordiner1,9, Nicholas A. Lombardo11, Scott Wieman11,1, Vladimir Airapetian12, Veronica Allen12, Daria Pidhorodetska13, Erika Kohler1,2, Julianne Moses13, Timothy A. Livengood6, Danielle N. Simkus1,2, Noah J. Planavsky14, Chuanfei Dong15, David A. Yuen1, Arie van den Berg17, Alexander A. Pavlov1, Jonathan J. Fortney17

Author Affiliations:

*Corresponding Author: der-you.kao@nasa.gov, 301-614-6710
The standard approach to EOS calculations relies on density functional theory (DFT), which targets the tridimensional electronic density rather than the \(N_e\) many-body wave function. Its success and simplicity have led to a widespread use in material science and to the development of several software packages that allow fast and reproducible calculations [3]. Although DFT is formally exact, the explicit functional form to describe the exchange and correlation (XC) effects between electrons remains approximated [6]. Indeed, a systematic and efficient route to improve the XC functional is still lacking. Therefore, in practical solid-state calculations, benchmarks against experimental data are often required to validate the XC functional used to describe the system in a satisfactory manner.

**Phase Diagram of Hydrogen and a Hydrogen-Helium Mixture at Planetary Conditions by Quantum Monte Carlo Simulations**

Kieron Burke

IPAM physics and ML 2019

Interior of Jupiter
Relations between WDM and classical DFT

• Mermin theorem applies to equilibrium electrons at all temperatures
• Implies entropic contributions to XC and non-interacting functionals
• As $T \rightarrow 0$, becomes ground-state theory
• As $T \rightarrow \infty$, becomes classical theory of charged liquid!
Essence of HK theorem

• Prior to 1964, essentially believed one needs the many-body wavefunction to find properties of system.
• HK theorem says density alone, in principle, determines all properties.
• Extreme compression of information
• KS scheme leverages this for electronic problem.
Atmospheric CO₂ at Mauna Loa Observatory

Scripps Institution of Oceanography
NOAA Earth System Research Laboratory
The community atmospheric chemistry box model
CAABA/MECCA-4.0

Rolf Sander, Andreas Baumgaertner, David Cabrera-Perez, Franziska Frank, Sergey Gromov, Jens-Uwe Groß, Hartwig Harder, Vincent Huijnen, Patrick Jöckel, Vlassis A. Karydis, Kyle E. Niemeyer, Andrea Pozzer, Hella Riede, Martin G. Schultz, Domenico Taraborrelli, and Sebastian Tauer

1 Air Chemistry Department, Max-Planck Institute of Chemistry, P.O. Box 3060, 55020 Mainz, Germany
2 Deutsches Zentrum für Luft- und Raumfahrt (DLR), Project Management Agency, 53227 Bonn, Germany
3 Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institut für Physik der Atmosphäre, Oberpfaffenhofen, 82234 Weßling, Germany
4 IEK-7, Forschungszentrum Jülich, Jülich, Germany
5 Royal Netherlands Meteorological Institute (KNMI), De Bilt, the Netherlands
6 IEK-8, Forschungszentrum Jülich, Jülich, Germany
7 School of Mechanical, Industrial & Manufacturing Engineering, Oregon State University, Corvallis, Oregon, USA
8 also at: Institute of Global Climate and Ecology (Roshydromet and RAS), Moscow, Russia
9 now at: Research and Development, Method Development in Remote Sensing, German Weather Service (DWD), Offenbach, Germany
10 now at: Jülich Supercomputing Center (JSC), Forschungszentrum Jülich, Jülich, Germany

Correspondence: Rolf Sander (rolf.sander@mpic.de)

Received: 8 August 2018 – Discussion started: 3 September 2018
Revised: 11 March 2019 – Accepted: 12 March 2019 – Published: 5 April 2019

Abstract. We present version 4.0 of the atmospheric chemistry box model CAABA/MECCA that now includes a number of new features: (i) skeletal mechanism reduction, (ii) the Mainz Organic Mechanism (MOM) chemical mechanism for volatile organic compounds, (iii) an option to include reactions from the Master Chemical Mechanism (MCM) and other chemical mechanisms, (iv) updated isotope tagging, and (v) improved and new photolysis modules (JVAL, RADJIMT, DISSOC). Further, when MECCA is connected to a global model, the new feature of coexisting multiple chemistry mechanisms (PolyMECCA/CHEMGLUE) can be used. Additional changes have been implemented to make the code more user-friendly and to facilitate the analysis of the model results. Like earlier versions, CAABA/MECCA-4.0 is a community model published under the GNU General Public License.

1 Introduction

MECCA (Module Efficiently Calculating the Chemistry of the Atmosphere) is an atmospheric chemistry module that contains a comprehensive chemical mechanism with tropospheric and stratospheric chemistry of both the gas and the aqueous phases. For the numerical integration, MECCA uses the KPP (Kinetic PreProcessor) software (Sandu and Sander, 2006).

To apply the MECCA chemistry to atmospheric conditions, MECCA must be connected to a base model via the MESSy (Modular Earth Submodel System) interface (Jöckel et al., 2010). This base model can be a complex 3-dimensional model but it can also be a simple box model. CAABA (Chemistry As A Boxmodel Application) is such a box model, simulating the atmospheric environment in which the MECCA chemistry takes place.
Direct Kinetic and Atmospheric Model Interactions of the CH₂OO + (CH₃)₂CO Reaction with Acetone

Rabi Chhantyal-Pun, M. Anwar H. Khan, Rebecca Zachary J. Buras, Carl J. Percival, Dudley E. Shallcross

School of Chemistry, University of Bristol, Cantock’s Close, Bristol BS8 1TS, United Kingdom
Department of Chemical Engineering, Massachusetts Institute of Technology, 77 Mass Ave, Cambridge, MA 02139, United States
Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr, Pasadena, CA 91109, United States

Supporting Information

ABSTRACT: Mounting evidence suggests that Criegee intermediate reactions are occurring as key tropospheric oxidants of both organic and inorganic gases, supporting photochemical and oxidation chemistry initiated by OH radicals. Here, the rate coefficient of the reaction of the simplest Criegee intermediate CH₂OO with acetone (CH₃)₂CO, was measured using laser flash photolysis and cavity ring-down spectroscopy methods under tropospherically relevant conditions of pressure and temperature. The pressure dependence of \( k(\text{CH}_2\text{OO} + (\text{CH}_3)_2\text{CO}) = 10^{-13} \left[\text{N}_2\right]/((3.7 \pm 1.4) \times 10^{16} + [\text{N}_2]) \) cm³ molecule⁻¹ s⁻¹ was measured from 50 to 100 Torr range, returning a high-pressure limit value of \((4.7 \pm 0.1) \times 10^{-13}\) cm³ molecule⁻¹ s⁻¹ at 293 K. A temperature dependence of \( k(\text{CH}_2\text{OO} + (\text{CH}_3)_2\text{CO}) = 3.7 \times 10^{-13} \left(\text{N}_2\right)/((3.7 \pm 1.4) \times 10^{16} + [\text{N}_2]) \) cm³ molecule⁻¹ s⁻¹ was observed in the 250–310 K range. The global modeling results were used to model the speciated Criegee intermediate field using the recently reported energy landscapes of Criegee intermediates. The incorporation of the rate coefficient into global models predicts decreases in acetone concentration of as much as 10–40 ppbv. The formation of a van der Waals complex before passing over a sub-barrier to form a SOZ, as shown in Figure 4. Only one cycloaddition pathway is possible because of the symmetric substitution of the carbonyl carbon atoms in both reactants. The molecular geometries and energies, calculated by Jalan et al. using RCCSD(T)-F12a/VTZ-F12//B3LYP/MG3S methods, were input to Arkane, a new one-dimensional ME solver previously named Cantherm that is part of the RMG-Py package. The \( k(T, P) \) values output by

Figure 4. Minimum energy pathway for the reaction of \( \text{CH}_2\text{OO} + (\text{CH}_3)_2\text{CO} \). The stationary point energies were calculated at the RCCSD(T)-F12a/VTZ-F12//B3LYP/MG3S level of theory and were obtained from the previous study by Jalan et al.¹¹
Conformers and Charge Distribution of Different Protonation States by Quantum Chemical Calculations

We used quantum chemical calculations to obtain the energetically most favorable model conformers for the two protonation states of interest and assess differences in their molecular electrostatic potential (MEP), which describes the charge distribution around the molecule. Starting from the structures published by Rak
Maybe we can?

• If we make $T_S$ much more accurate and reliable, can suddenly do much larger simulations, so much more relevant
• If we make $E_{XC}$ much more accurate and reliable, we can do many more simulations of more challenging systems
• Huge impact if ML improves functionals for classical DFT, nuclear DFT, or thermal DFT
• Can we do this in 10 years, not 100?
• Maybe we must!
Summary

• Direct solution of Schrödinger equation too expensive to be practical in many situations.
• Density functional theory (DFT) allows larger and more diverse systems to be modelled, but with uncontrollable errors.
• DFT chemical and materials modeling already plays important role in modern materials and drug design.
• Two very important possibilities:
  – Much bigger DFT simulations, more realistic
  – Much more accurate, systematic DFT calculations
• Other (smaller) areas need better functionals
• Thanks to IPAM and NSF and DOE.