

Using machine learning to find density functionals

Kieron Burke and friends
UC Irvine
Departments of Physics
and of Chemistry
<http://dft.uci.edu>

Physics

- Theoretical physics
- Computational physics
- Quantum mechanics
- Molecules and materials by design

Theoretical physics

Suggested Speakers
(next Fall)
ton Nedelin

$\Theta_{AB} M_{AB} \rightarrow (\tilde{\rho}^i)$
 $\sum_C \lambda_{ABC} = 1 - \frac{S_{AB}}{\Theta_{AD}}$
 $\Phi_{ABC} = \sum \lambda_{ABC}^i$
 $\lambda_{ABC} = \sum \lambda_{ABC}^i$

$\Theta_{MN} \rightarrow \begin{pmatrix} 1 & 1 & 3 & 3 \\ & * & & \\ & & * & \\ & & & * \end{pmatrix}$

$SL(2) \times SO(6,6) \rightarrow \dots$
 $SL(8) \rightarrow \mathbb{R}_M^+ \times SL(7)$
 $W = W_0 + W_A \Phi^A + \frac{1}{2} W_{AB} \Phi^A \Phi^B + \frac{1}{6} W_{ABC} \Phi^A \Phi^B \Phi^C$
 (16 complex)
 $W(S, T, U) = a_0 + \dots + c_3^i S T^3 U^3$
 32 real
 $\begin{pmatrix} 1 \\ S \end{pmatrix} \times \begin{pmatrix} 1 \\ T \end{pmatrix} \times \begin{pmatrix} 1 \\ U \end{pmatrix}$

$V = -3|W|^2 + |DW|^2$
 $DV \neq 0$
 D^2V

$SO(9) \rightarrow SO(8,1)$
 $SO(8,1) \rightarrow T^2 \times T^2 \times T^2$
 $H_{minp} \rightarrow H_{123} \oplus H_{456} \oplus H_{156} \oplus H_{123456}$
 $426 \quad 423$
 $453 \quad 153$

$\langle U \rangle \rightarrow SO(3) \times SO(3) \rightarrow \Lambda_{(ABC)} \equiv \{ \Lambda_{111}, \Lambda_{112}, \Lambda_{122}, \Lambda_{222} \}$

$SO(6,6) \rightarrow SO(2,2) \times SO(3)$
 $SL(6) \rightarrow SL(2) \times SO(3)$

$ds_{11}^2 = \tau^2 ds_4^2 + \rho ds_7^2$
 $R^a_b = d\Omega^a_b + \Omega^a_c \wedge \Omega^c_b$
 $W = \frac{ST+T^2+U^2}{+TU+SU}$
 $ds^2 = d\theta^2 + \sinh^2 \theta dp^2$
 $e^1 = d\theta$
 $e^2 = \sinh \theta dp$
 $N = \frac{ST}{U}$

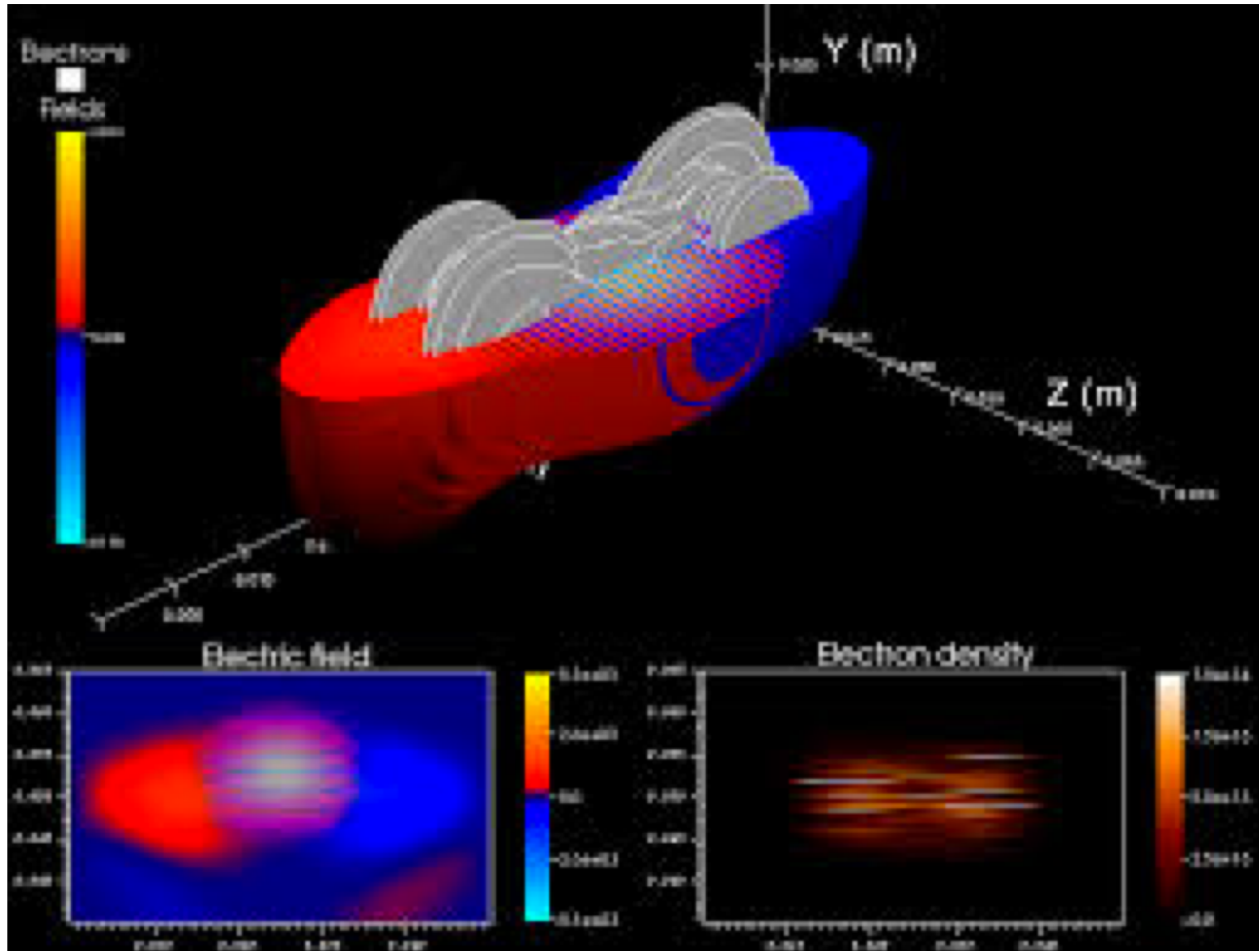
$\int \sqrt{g} R^{(2)} dx = 4\pi(2-2g) < 0$
 Milnor

Peppar & Peppar
 $W = \int (G_{(0)} + G_{(0)} \wedge \Phi + \Phi \wedge d\Phi + \Phi^3)$

$\Omega^a_b = \frac{1}{2} \epsilon^a_{bc} e^b \wedge e^c$
 $\Omega^a_b = \frac{1}{2} \epsilon^a_{bc} e^b \wedge e^c$
 $\Omega^a_b = \frac{1}{2} \epsilon^a_{bc} e^b \wedge e^c$

Computational physics

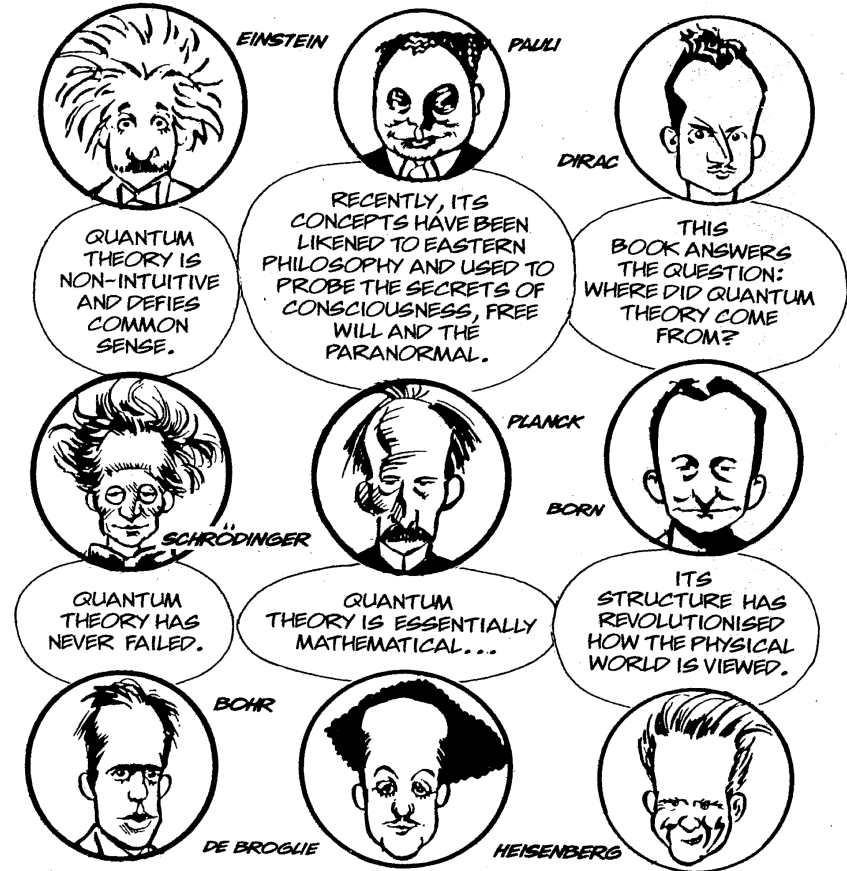
<https://www.theorie.physik.uni-muenchen.de/lruhl/>



Quantum Mechanics

What is Quantum Theory?

Quantum theory is the most successful set of ideas ever devised by human beings. It explains the periodic chart of the elements and why chemical reactions take place. It gives accurate predictions about the operation of lasers and microchips, the stability of DNA and how alpha particles tunnel out of the nucleus.



Niels Bohr's presentation of quantum theory in 1927 remains today's orthodoxy. But Einstein's *thought* experiments in the 1930s questioned the theory's fundamental validity and are still debated today. Could he be right again? Is there something missing?

Let's begin at the beginning . . .

Introducing Quantum theory y J.P. McEvoy

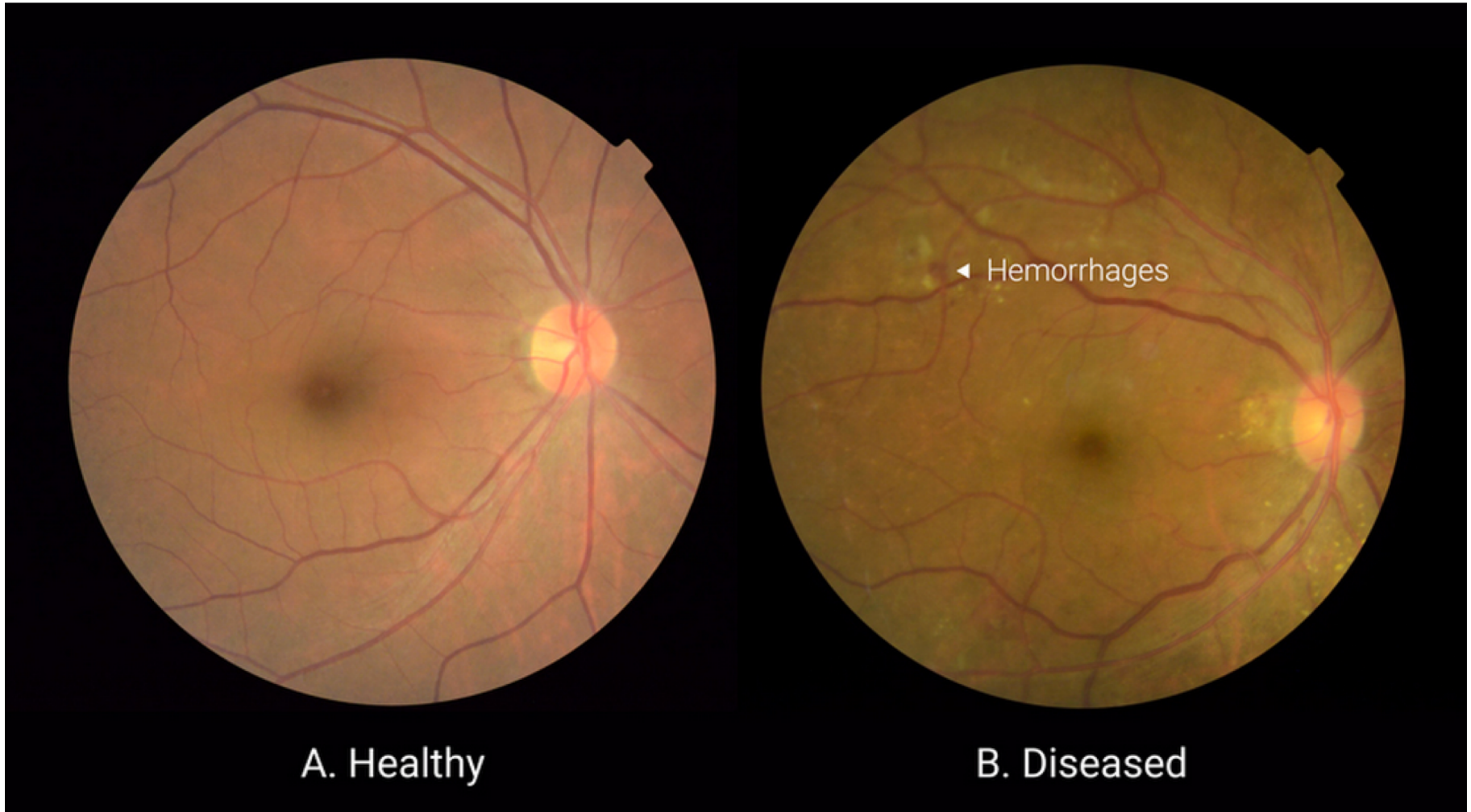
Dirac (1929)

- The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble.
- It therefore becomes desirable that approximate practical methods of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation.

Machine learning

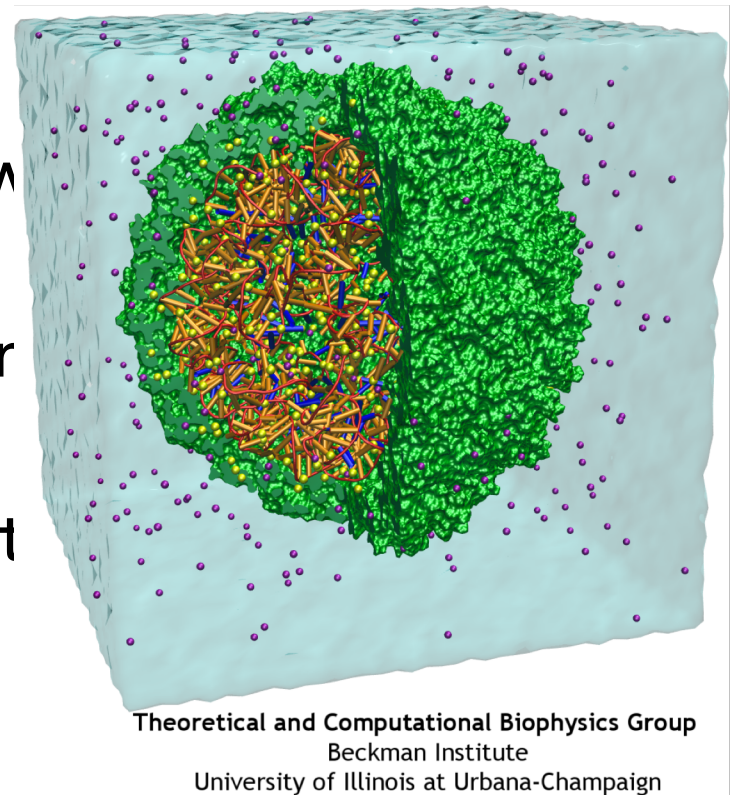
- Newly popular algorithms which, combined with modern computers, are being implemented throughout society
- Examples include Google's page rank
- AlphaGo
- Netflix movie suggestions
- Google translate
- All online ads
- Based on statistics, especially Bayesian analysis

diabetic retinopathy



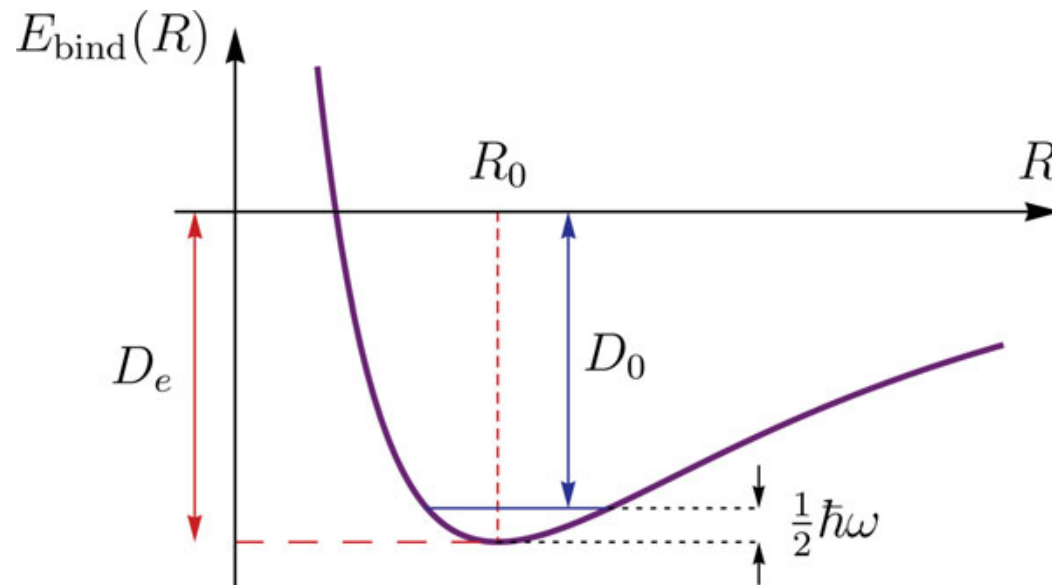
Molecular dynamics

- Solve Newton's equations for nuclei at given temperature and pressure.
- Use simple force fields between atoms
- Can do a million atoms for microseconds
- Simulate medicines and materials
- But, cannot break bonds!



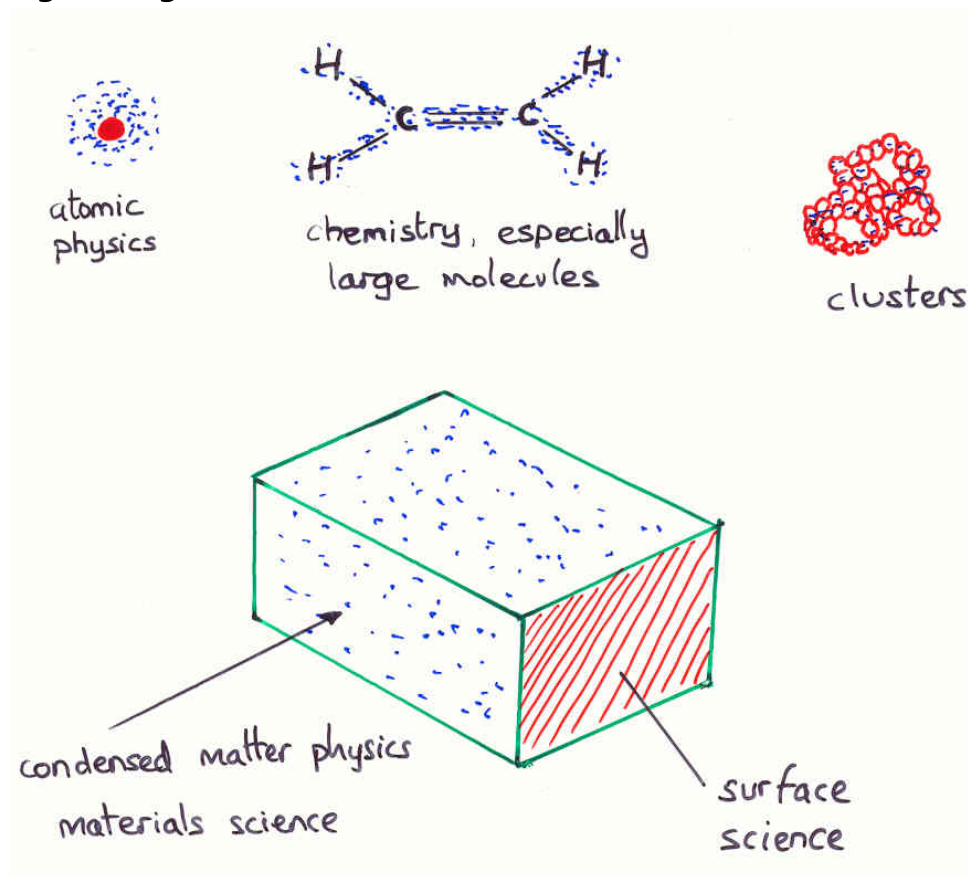
Basic Electronic Structure Problem

- Just want $E(R)$



Electronic Structure Problem: Diversity

- For all everyday matter



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.
- Required accuracy is about 1 in 10^7 for electronic energy of 500 atoms, in chemistry or materials science.
- When # electrons doubles, computer cost increases by 128 if you solve S-eqn.

Mathematical form of problem

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

$$\hat{H} = \hat{T} + \hat{V}_{\text{ee}} + \hat{V},$$

where the kinetic and elec-elec repulsion energies are

$$\hat{T} = -\frac{1}{2} \sum_{i=1}^N \nabla_i^2, \quad \hat{V}_{\text{ee}} = \frac{1}{2} \sum_{i=1}^N \sum_{j \neq i}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|},$$

and difference between systems is N and the one-body potential

$$\hat{V} = \sum_{i=1}^N v(\mathbf{r}_i)$$

Often $v(\mathbf{r})$ is electron-nucleus attraction

$$v(\mathbf{r}) = - \sum_{\alpha} \frac{Z_{\alpha}}{|\mathbf{r} - \mathbf{R}_{\alpha}|}$$

where α runs over all nuclei, plus weak applied \mathbf{E} and \mathbf{B} fields.

$$\{\hat{T} + \hat{V}_{\text{ee}} + \hat{V}\} \Psi = E \Psi, \quad E = \min_{\Psi} \langle \Psi | \hat{T} + \hat{V}_{\text{ee}} + \hat{V} | \Psi \rangle$$

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:

$$\left\{ -\frac{1}{2}\nabla^2 + v_s(\mathbf{r}) \right\} \phi_j(\mathbf{r}) = \epsilon_j \phi_j(\mathbf{r}), \quad \sum_{j=1}^N |\phi_j(\mathbf{r})|^2 = n(\mathbf{r}).$$

where $v_s(\mathbf{r})$ is *defined* to yield $n(\mathbf{r})$.

Define T_S as the kinetic energy of the KS electrons, U as their Hartree energy and

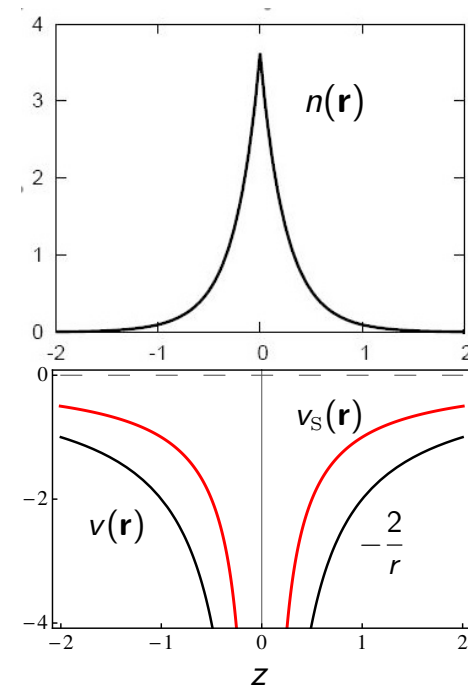
$$F = T + V_{ee} = T_S + U + E_{XC}$$

the remainder is the exchange-correlation energy.

Most important result of exact DFT:

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

Knowing $E_{XC}[n]$ gives closed set of self-consistent equations.



DFT in a nutshell, Kieron Burke, Lucas O. Wagner, *Int. J. Quant. Chem.* **113**, 96-101 (2013).

The Hubbard dimer: a density functional case study of a many-body problem D J Carrascal, J Ferrer, J C Smith, K Burke, *J Physics: Cond Mat* **27**, 393001 (2015)

PHET

- Online simulations from U Colorado

The kicker

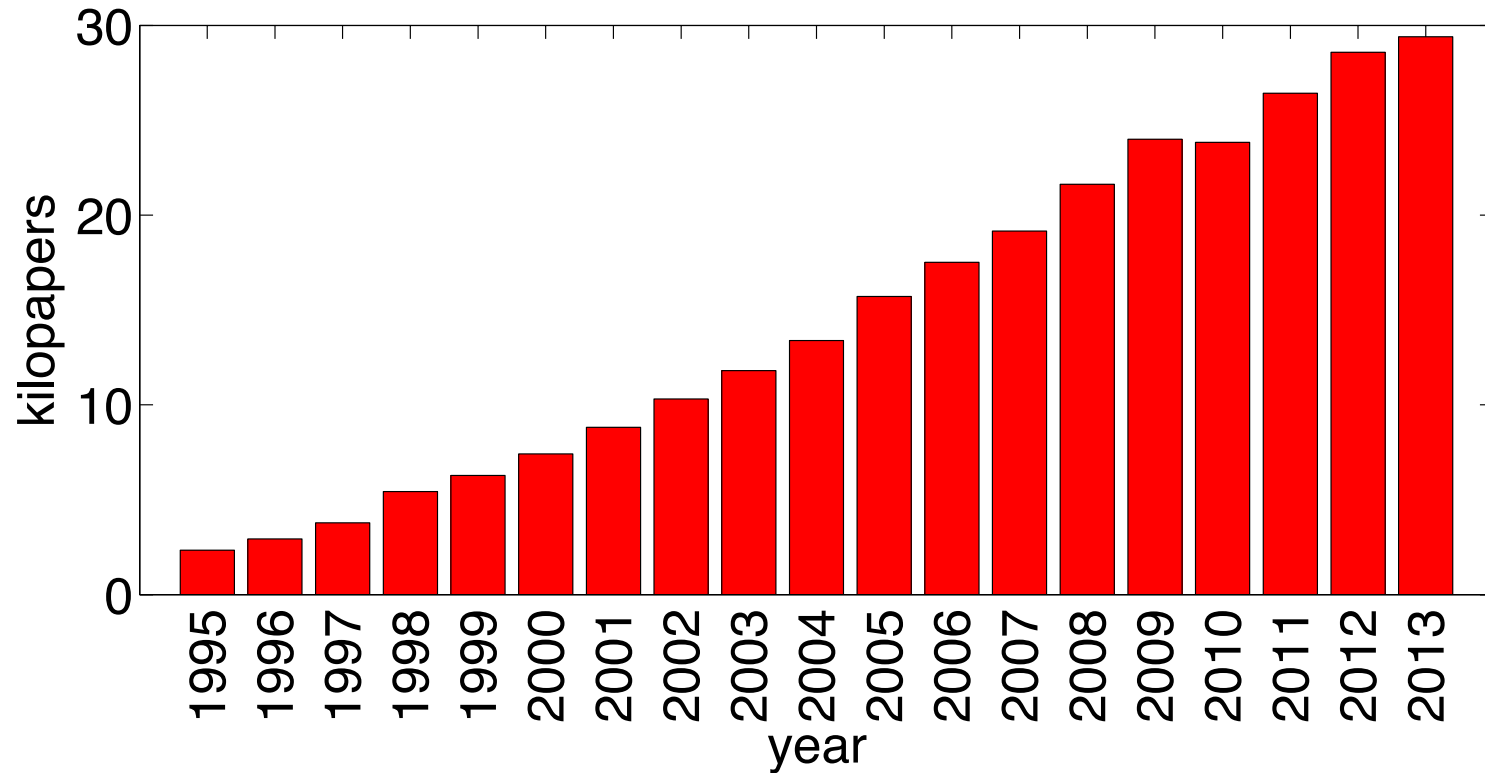
- Need a formula for a small fraction of electronic energy, called XC energy, in terms of density, containing all quantum many-body effects.
- First formula (1965) good for solids, but not accurate enough for chemistry.
- Next formulas (1990) give useful accuracy for chemistry and materials.

Modern research to find XC energy



Electronic Structure Problem: Impact

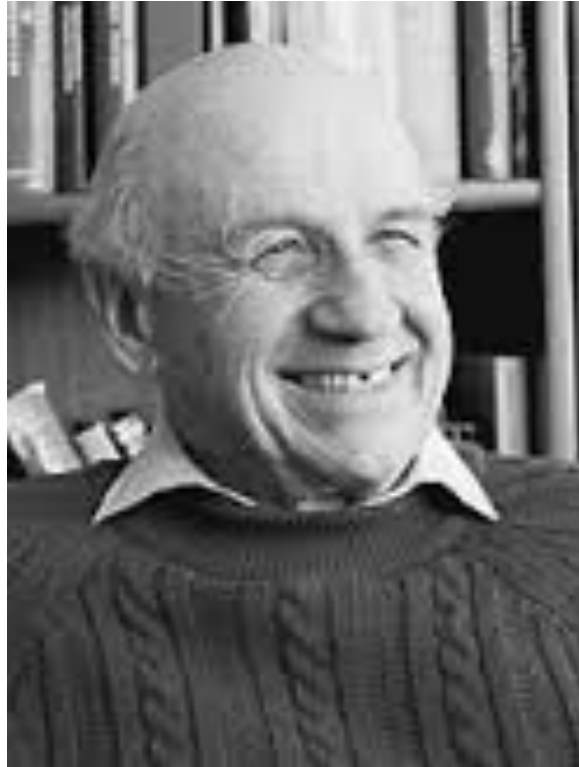
- Absurdly useful



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.

Walter Kohn (1923-2016)



Two problems in paradise

- Problem 1:
 - By using fake electrons, cost increases by factor of 8 when size doubles.
 - Limits 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, and some folks try to find it, but never accurately enough.
- Problem 2: Most interesting exotic phenomena in solids are too strongly quantum mechanical, so all today's approximations fail

Big data in materials

- Materials genome project
 - Endless workshops and meetings
 - Topics include data mining for specific functionality and machine-learned classical forces
-
- Special issue of *Journal of Chemical Physics*, Data-enabled Chemistry



Kernel ridge regression

- Kernel ridge regression (KRR). Given $\{\mathbf{x}_j, f_j\}$

$$\hat{f}(\mathbf{x}) = \sum_{j=1}^M \alpha_j k(\mathbf{x}_j, \mathbf{x})$$

$$k(\mathbf{x}, \mathbf{x}') = \exp(-\|\mathbf{x} - \mathbf{x}'\|^2 / (2\sigma^2))$$

length scale

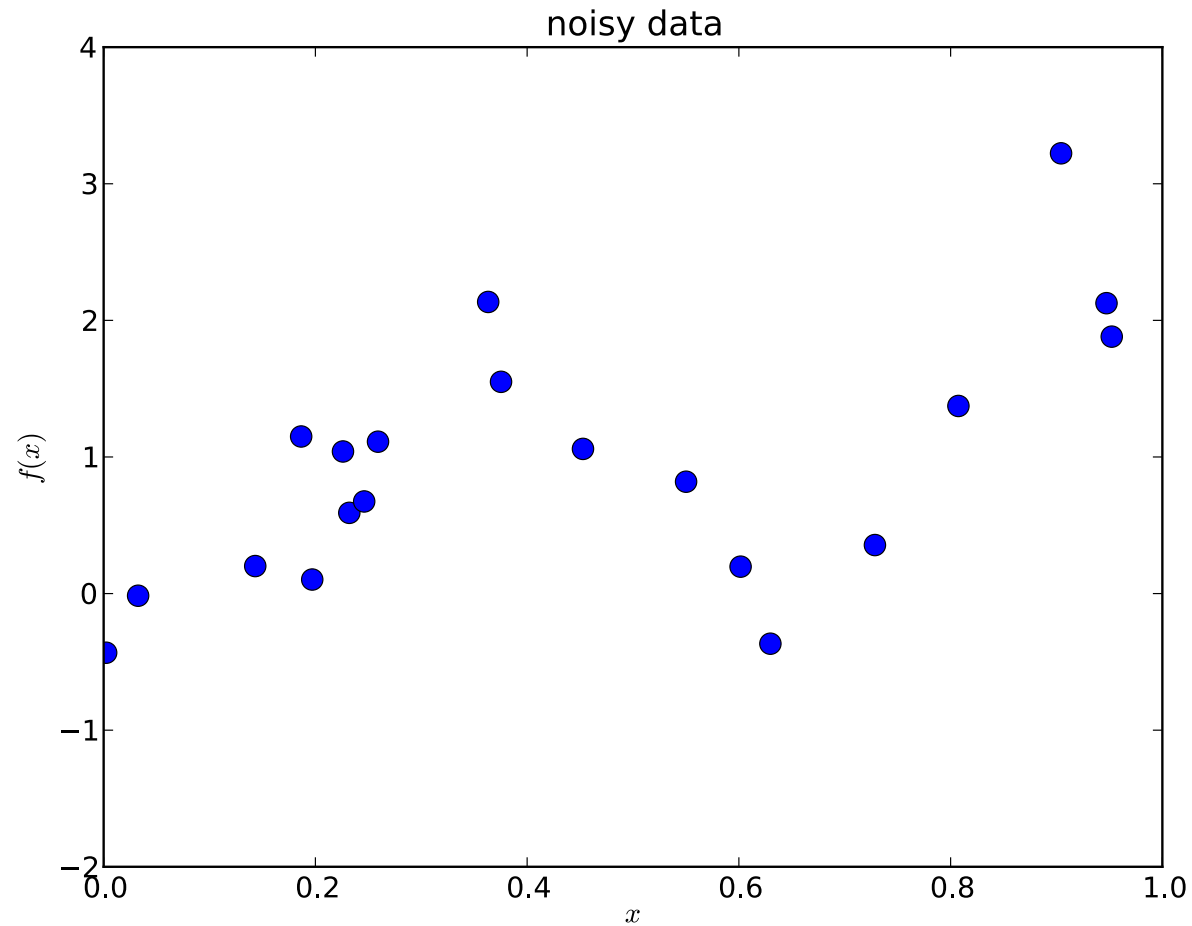
- Minimize:

$$\mathcal{C}(\boldsymbol{\alpha}) = \sum_{j=1}^M (\hat{f}(\mathbf{x}_j) - f_j)^2 + \lambda^2 \|\boldsymbol{\alpha}\|^2$$

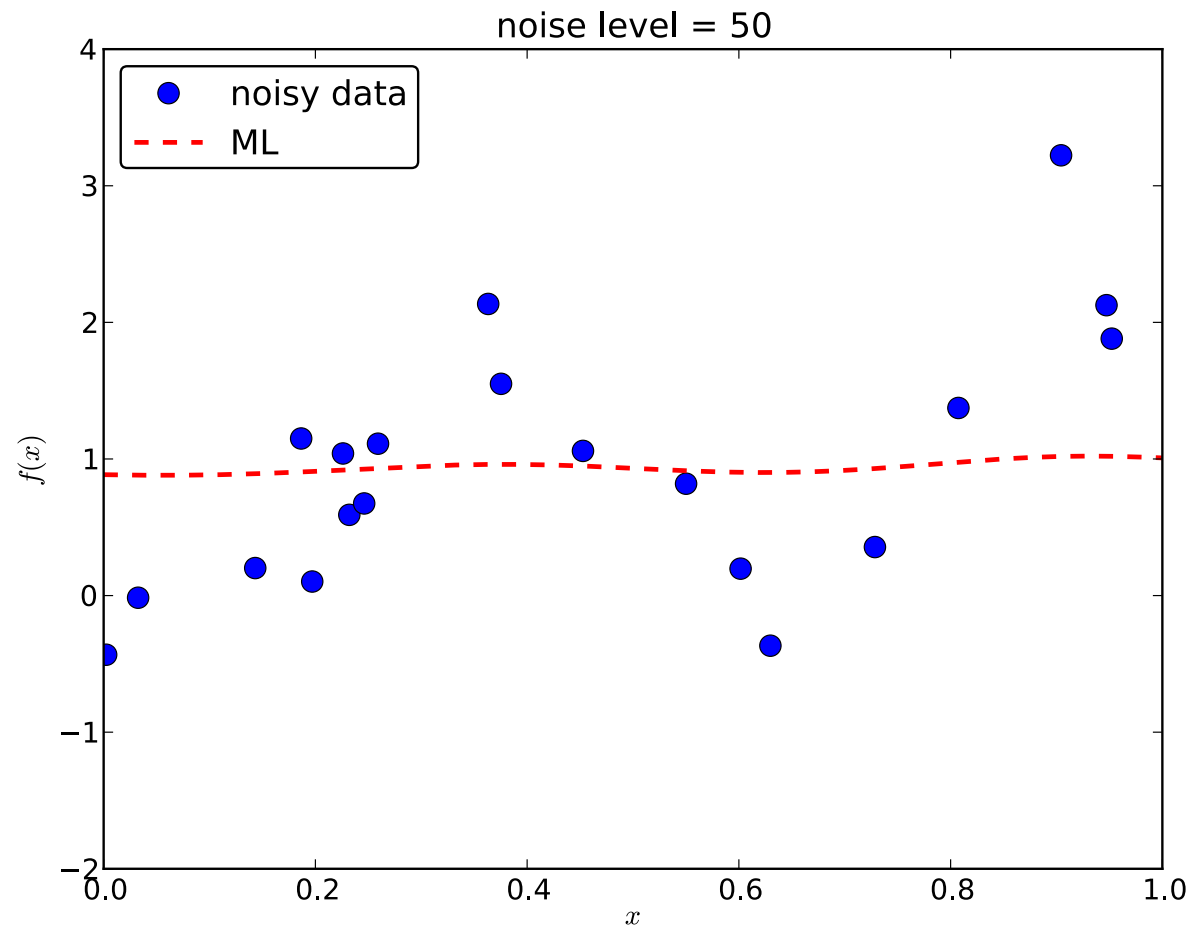
noise level

$$\boldsymbol{\alpha} = (K + \lambda^2 I)^{-1} \mathbf{f}$$

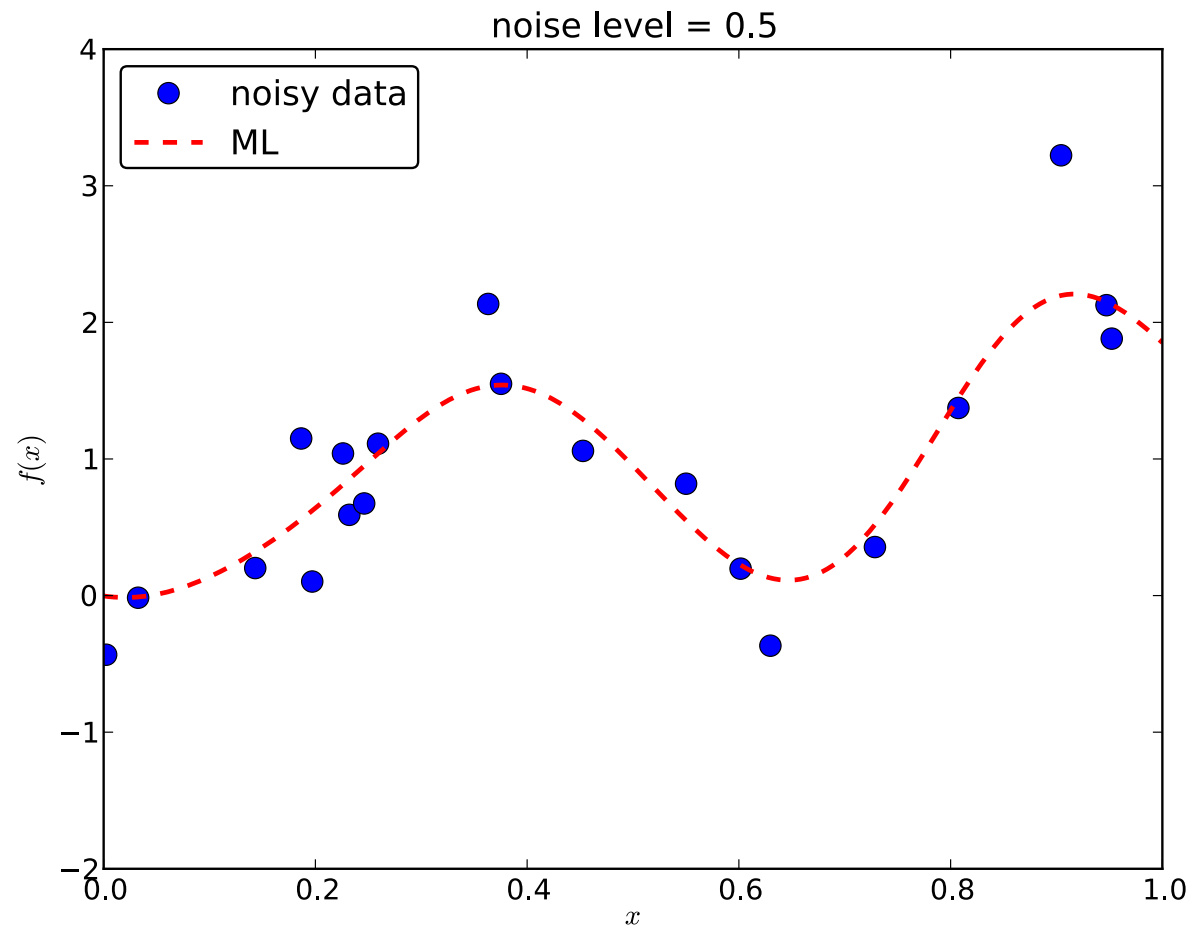
Fitting a simple function



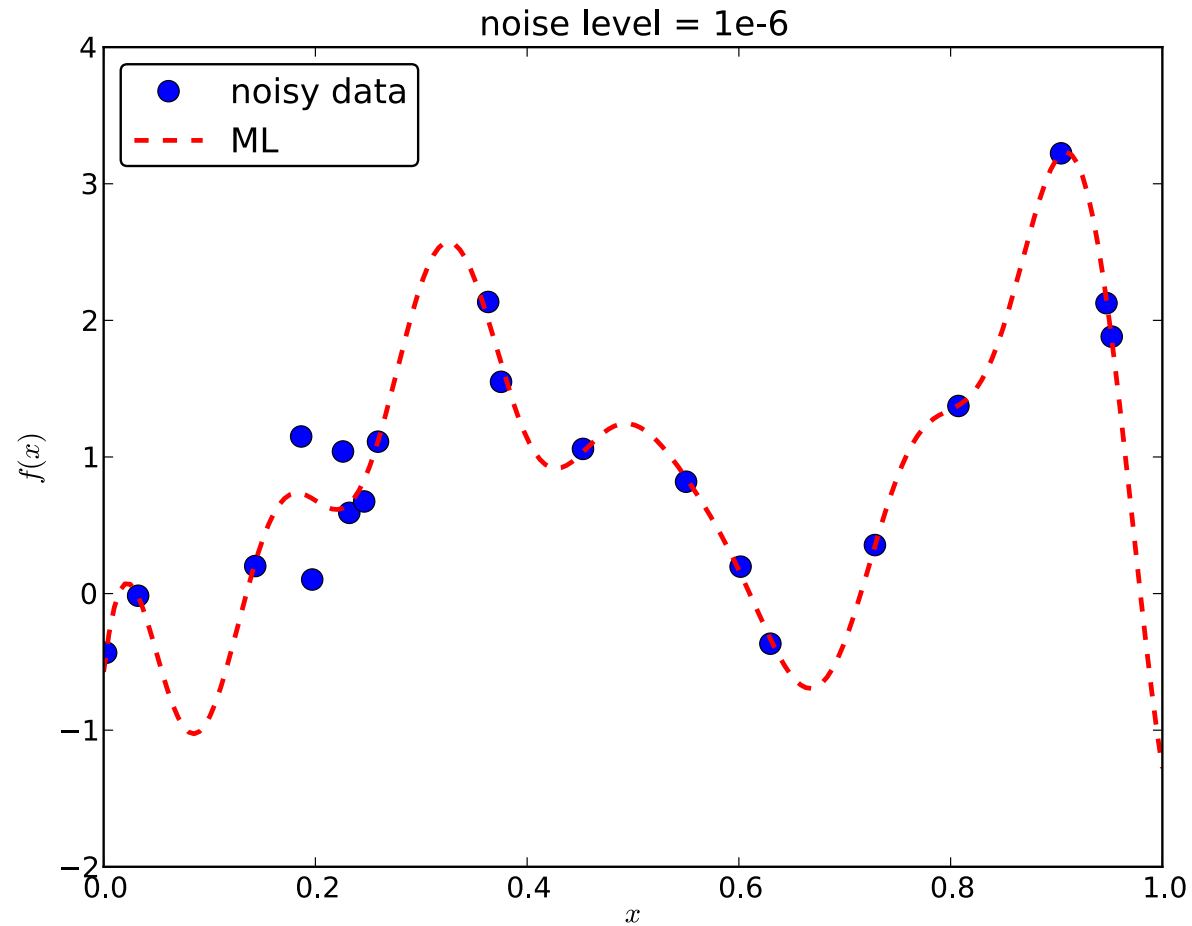
Too high noise level: underfit



Medium noise level

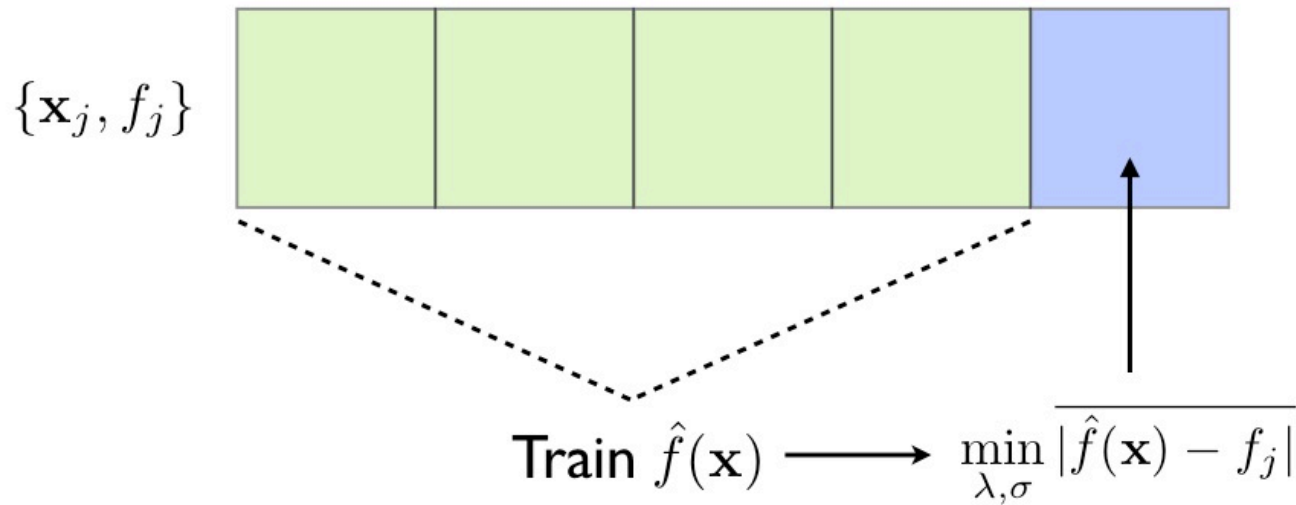


Small noise level: overfit



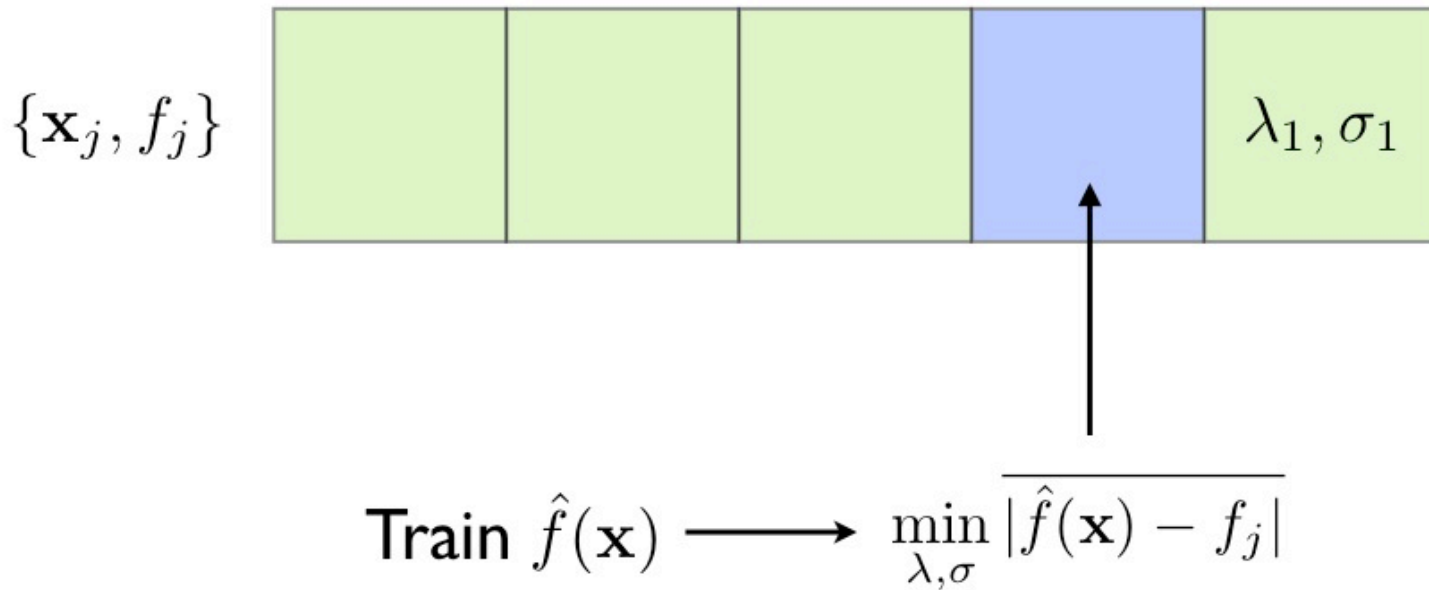
Cross validation

- Always optimize on samples *not* in the training set



Cross validation

- Always optimize on samples *not* in the training set



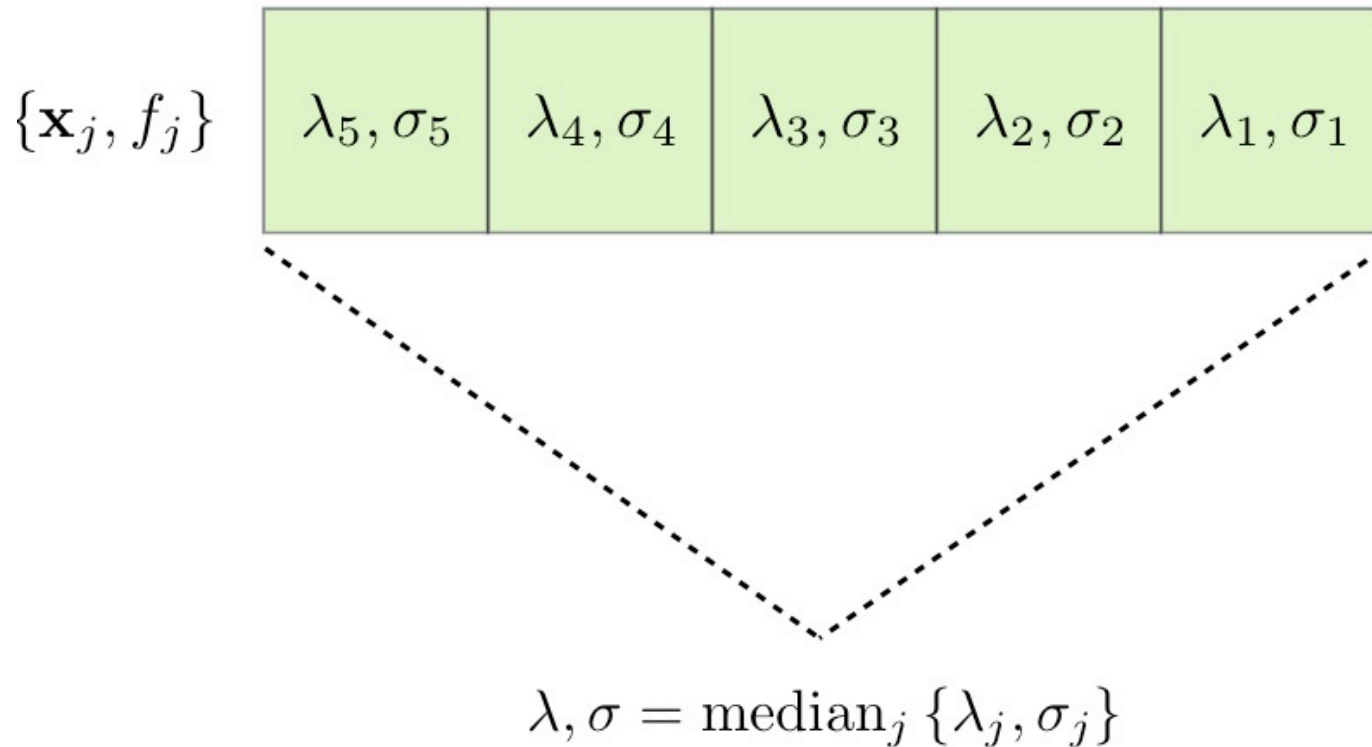
More cross validation

- Always optimize on samples *not* in the training set

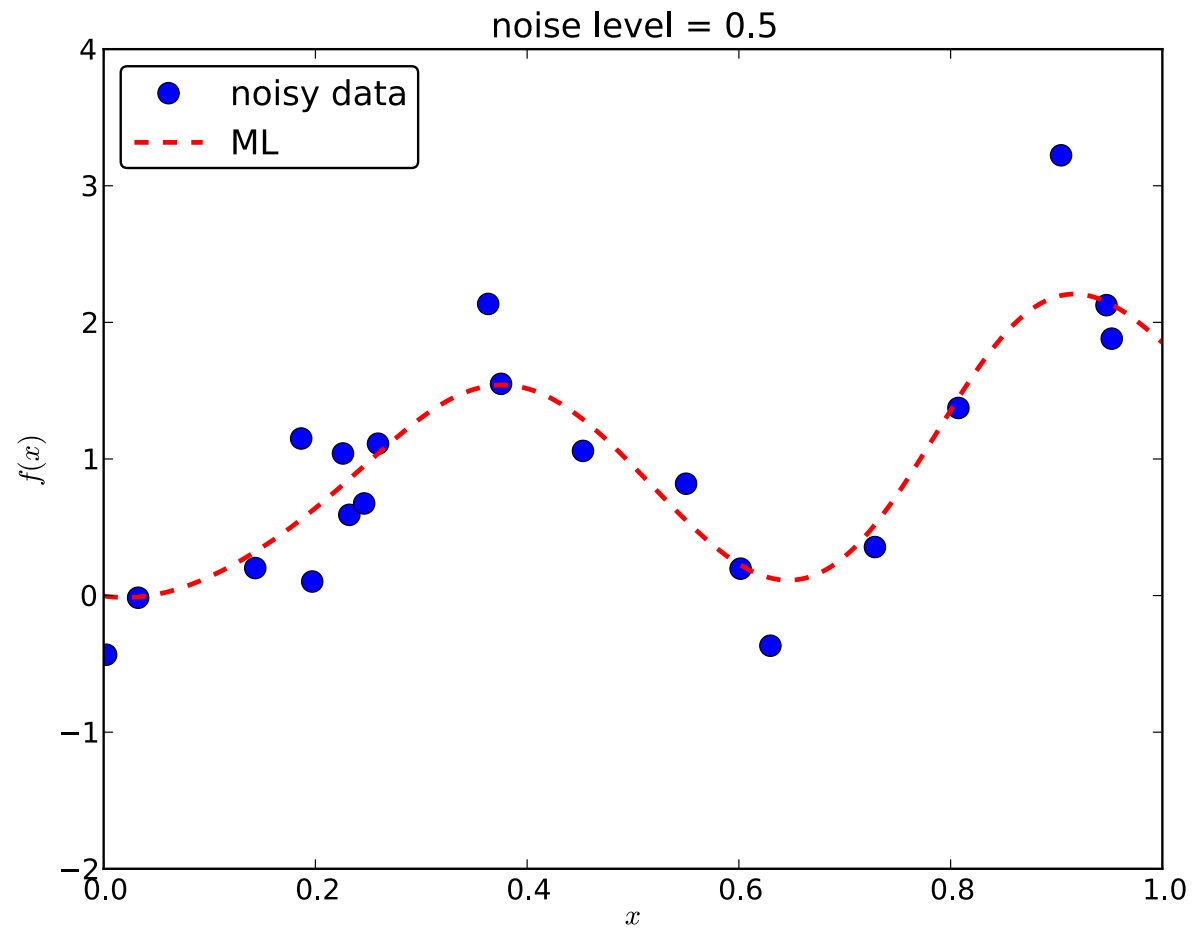
$\{\mathbf{x}_j, f_j\}$	λ_5, σ_5	λ_4, σ_4	λ_3, σ_3	λ_2, σ_2	λ_1, σ_1
-------------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------

Average over samples

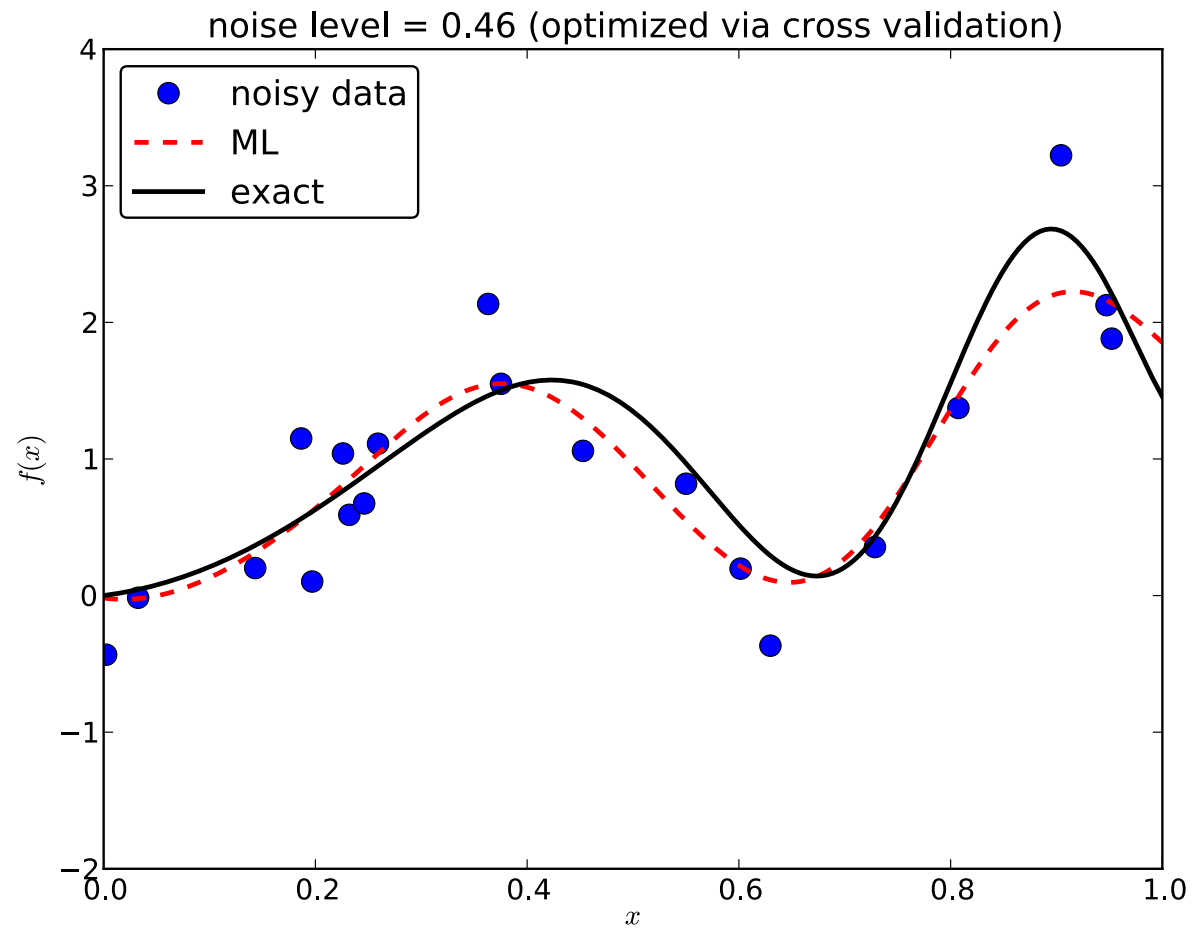
- Always optimize on samples *not* in the training set



Medium noise level



Exact function and best fit



Demo problem in DFT

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

- Define class of potential:

$$v(x) = - \sum_{i=1}^3 a_i \exp(-(x - b_i)^2 / (2c_i^2))$$

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

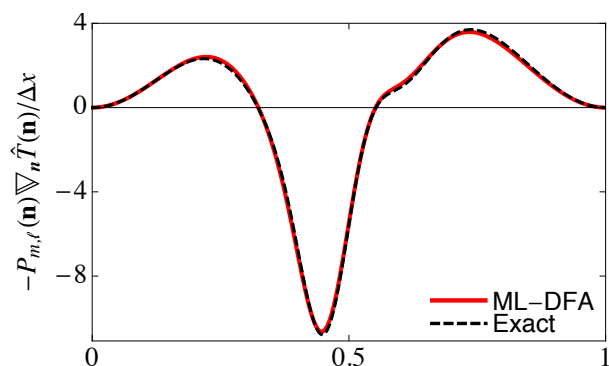
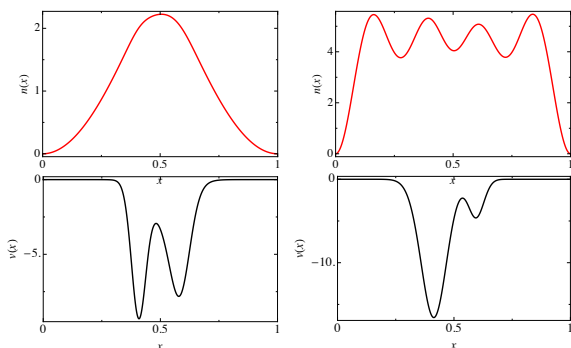
- ML-DFA for KE:

$$\hat{T}(\mathbf{n}) = \bar{T} \sum_{j=1}^M \alpha_j k(\mathbf{n}_j, \mathbf{n})$$

$$k[\mathbf{n}, \mathbf{n}'] = \exp \left(- \int dx (n(x) - n'(x))^2 / (2\sigma^2) \right)$$

Test case: KS electrons in a box

Generate 2000 potentials. Solve for up to 4 electrons.

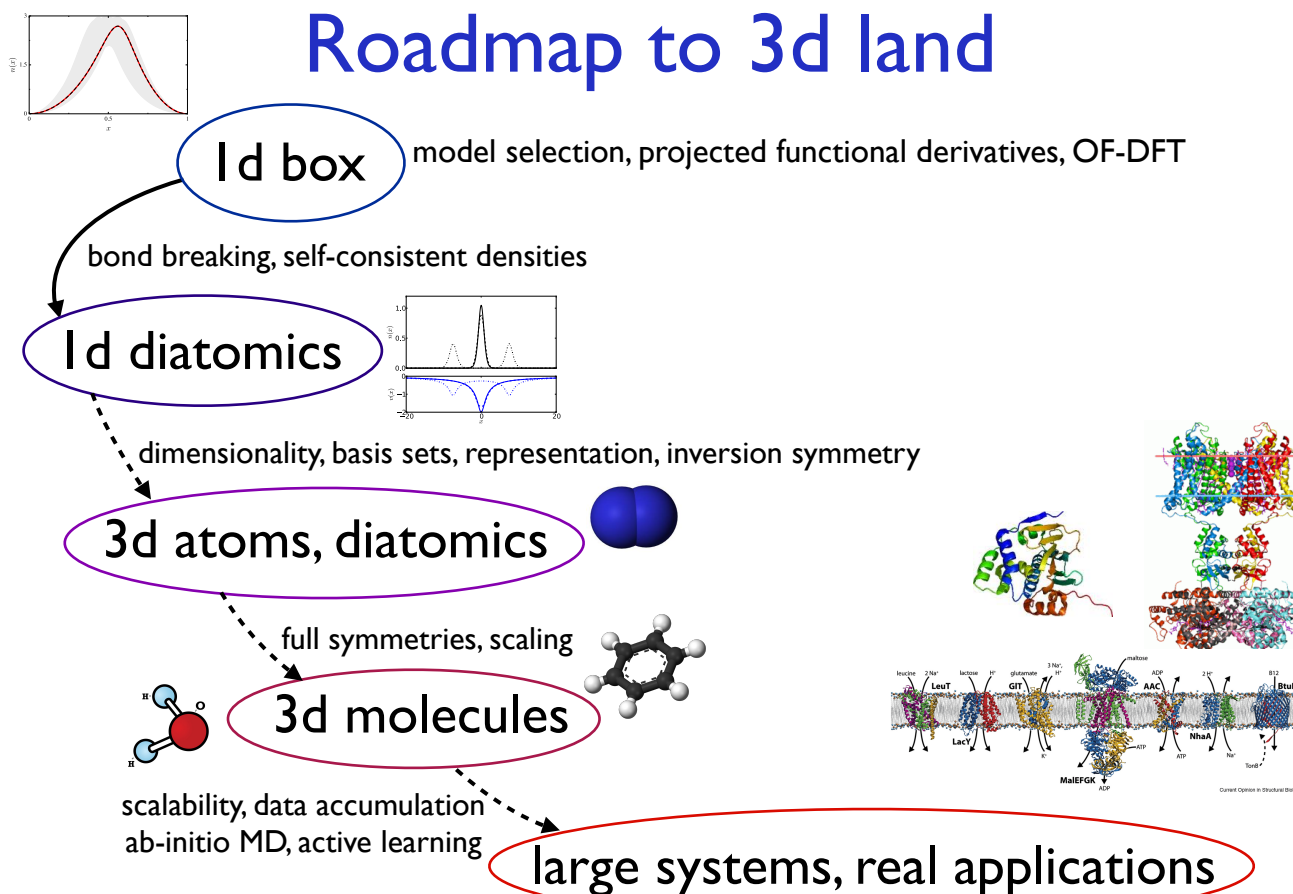


kcal/mol						
N	M	λ	σ	$\overline{ \Delta T }$	$ \Delta T ^{\text{std}}$	$ \Delta T ^{\text{max}}$
1	40	2.4×10^{-5}	238	3.3	3.0	23.
	60	1.0×10^{-5}	95	1.2	1.2	10.
	80	6.7×10^{-6}	48	0.43	0.54	7.1
	100	3.4×10^{-7}	43	0.15	0.24	3.2
	150	2.5×10^{-7}	33	0.060	0.10	1.3
	200	1.7×10^{-7}	28	0.031	0.053	0.65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

LDA ~ 223 kcal/mol, Gradient correction ~ 159 kcal/mol

Finding Density Functionals with Machine Learning John C. Snyder, Matthias Rupp, Katja Hansen, Klaus-Robert Müller, Kieron Burke, Phys. Rev. Lett. **108**, 253002 (2012)

Road map back to reality



Molecules and movies

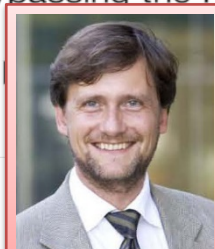
Malonaldehyde Proton Transfer



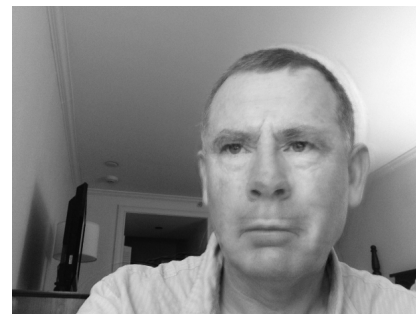
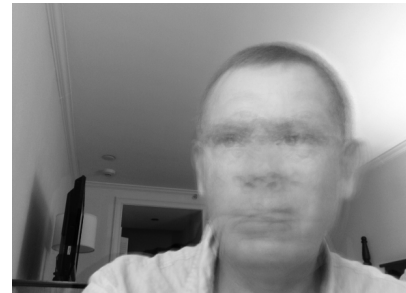
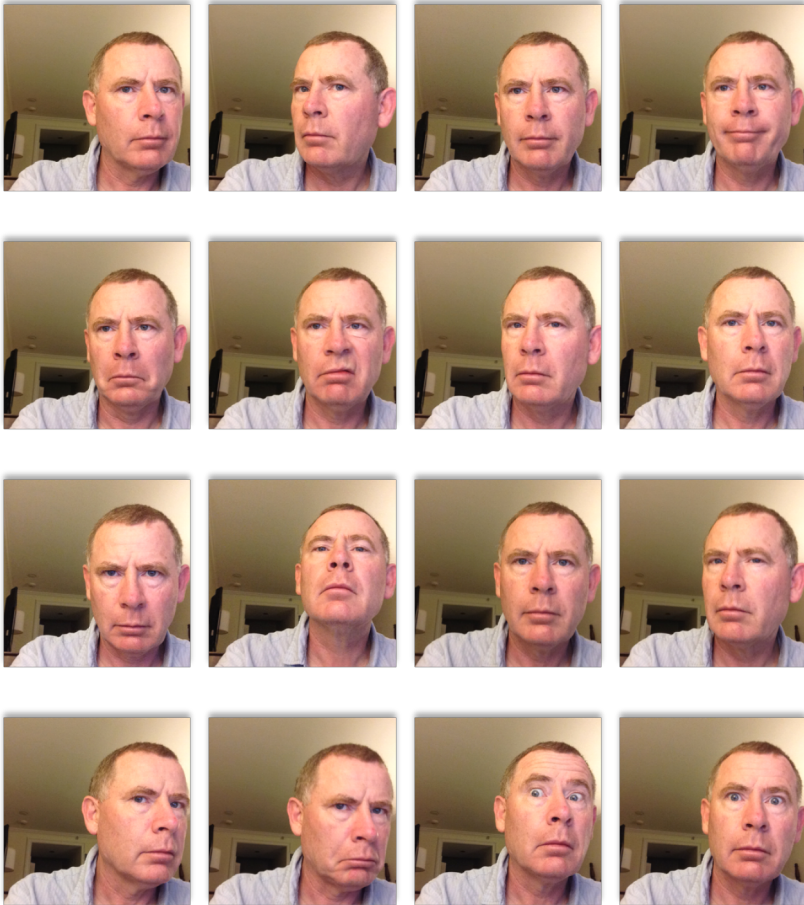
MAX-PLANCK-GESELLSCHAFT



“Bypassing the Kohn-Sham equations with machine learning”,
Kieron Burke, Leslie Vogt, Li Li,
Klaus-Robert Müller, Kieron Burke, Klaus-Robert Müller,
Nature Communications, 2017



Facial recognition via PCA



Kieron Burke

At the crossroads.

PCA basis for atomic densities

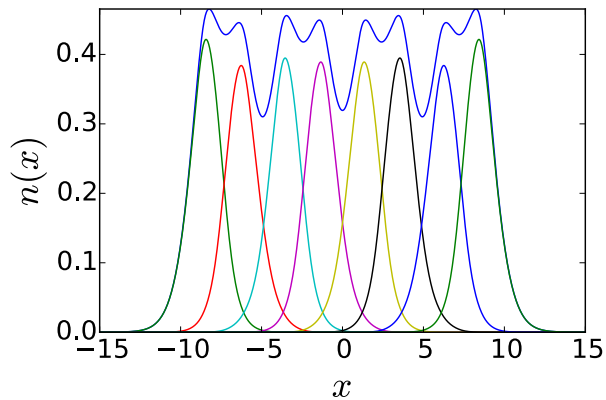


FIG. 5. Partition density of each H atom in H_8 .

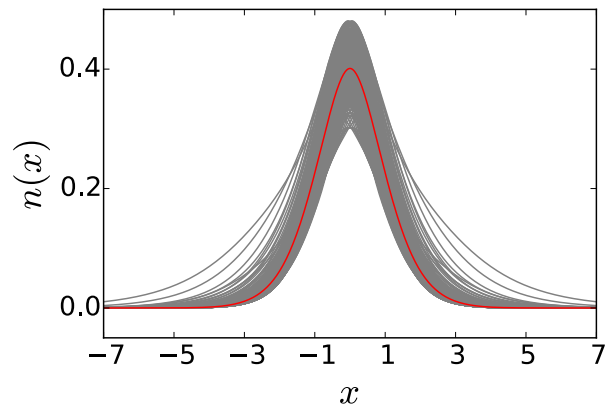


FIG. 6. Single H atom densities for H atoms in different chains and atomic distance (gray). The average density is plotted in red.

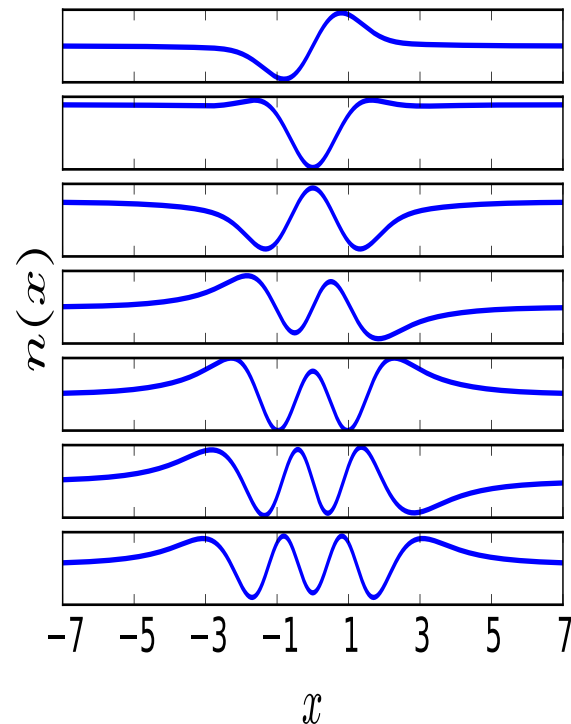
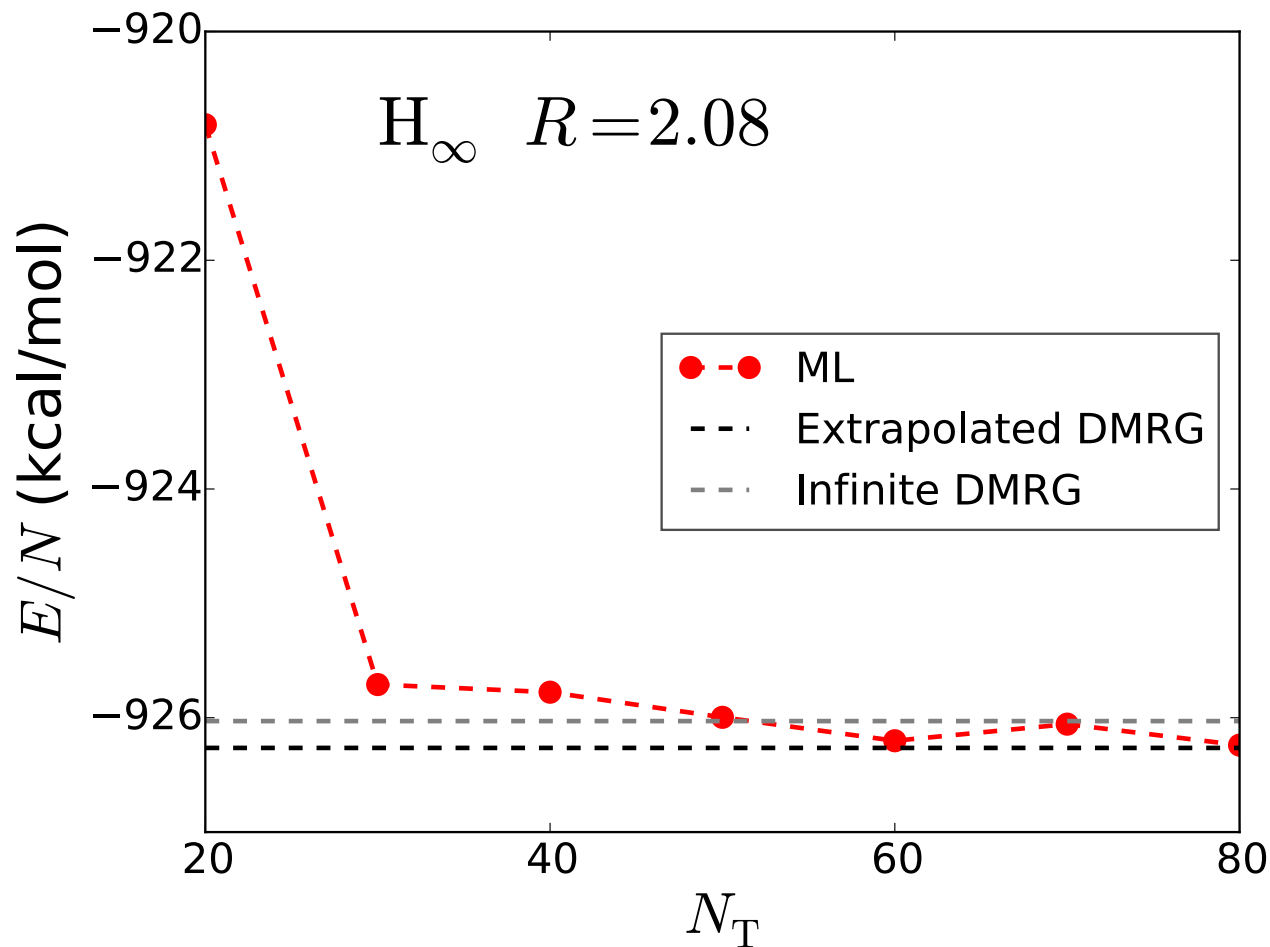


FIG. 7. First 7 principal components of the densities shown in Fig. 6, from top to bottom.

Learning $E_{xc}[n]$



20 years ago

VOLUME 82, NUMBER 2

PHYSICAL REVIEW LETTERS

11 JANUARY 1999

Several Theorems in Time-Dependent Density Functional Theory

Paul Hessler,* Jang Park,[†] and Kieron Burke

Department of Chemistry, Rutgers University, 315 Penn Street, Camden, New Jersey 08102

(Received 25 September 1998)

The time dependence of the exchange-correlation energy in density functional theory is given in terms of the exchange-correlation potential. The virial theorem for the exchange-correlation potential is shown to hold for *time-dependent* electronic systems and is illustrated by an exactly solved model: Hooke's atom with a time-dependent force constant. A relation between the coupling constant and functionals evaluated on scaled densities is derived. [S0031-9007(98)08169-1]

PACS numbers: 71.15.Mb, 31.15.Ew, 71.45.Gm

My experience with teachers



A unique grant renewal



**Research
Corporation**

101 North Wilmot Road, Suite 250
Tucson, Arizona 85711
Telephone 520/571-1111
FAX 520/571-1119
<http://www.rescorp.org>
e-mail: awards@rescorp.org

Michael P. Doyle, Ph.D.
Vice President
mdoyle@rescorp.org

April 25, 2000

Dr. Kieron Burke
Department of Chemistry

We are pleased to support the continuation of your collaboration with high school teacher Paul Hessler. The two of you break all of the rules. Paul was not educated as a chemist or physicist, the persons who we generally support. The kind of research that Paul does with you is not considered to be possible by most science educators. You are in a graduate institution that, by some elements of the media, is not attentive to teachers in high schools. All-in-all you defy popular misconceptions in a unique way.

..... The expenses that are contingent costs for research will be borne by your research accounts. In addition, Rutgers University has pledged additional funds, in the amount of \$9,000 over the same time period, for travel and additional salary for Paul Hessler. The net outcome is a unique partnership that serves the best interests of all concerned, as well as provides for the advancement of science.

Please be aware that Partners in Science conferences will continue in 2001 and beyond in much the same way as when they were organized by Research Corporation. The M. J. Murdock Charitable Trust is organizing the 2001 meeting to be held in San Diego January 12-14, 2001. If Paul intends to attend that meeting, Dr. John Van Zytveld of Murdock should be contacted.

The enclosed Acceptance of Conditions form should be completed and returned to schedule payment. Any restrictions or conditions that have been placed on this award are spelled out in this letter and on the form.

Yours truly,

A handwritten signature in black ink that reads "Mike".

Michael P. Doyle

MPD:mb
Enclosure

Lessons

- Machine learning is a huge huge thing, impacting all aspects of human society
- Just beginning to see how it benefits theoretical physics
- Extremely democratic – just need access to web and a netbook
- You can help create a more egalitarian world

TDDFT for WDM

- Many processes require thermal TDDFT
- E.g. stopping of a fast nucleus in matter
- But any description must involve coupling of electrons and nuclei, beyond Born-Oppenheimer and beyond Mermin functional.
- Many, many illegal calculations nonetheless.
- Aim: Less lofty goal
 - just linear response to weak time-dependent perturbation
 - Joule heating is 2nd order in perturbation, so temperature stays fixed
 - Yields corrections to conductivities in MKS calculations



Aurora Pribram-Jones

Thermal Density Functional Theory: Time-Dependent Linear Response and Approximate Functionals from the Fluctuation-Dissipation Theorem

Aurora Pribram-Jones, Paul E. Grabowski, Kieron Burke, *Phys. Rev. Lett.* **116**, 233001 (2016)

Summary

- I believe machine-learning represents a paradigm shift for STEM
- Generating and storing vast amounts of data and analyzing it
- Data is generated by both experiment and computation
- Fake data is now used to train networks
- Thanks to
 - Students: Tom Baker, Li Li, John Snyder, Kevin Vu, Isabelle Pelaschier
 - Collaborators: Klaus Mueller, Matthias Rupp, Katia Hansen, Felix Brockherde, Leslie Vogt, Mark Tuckerman
 - Institute of Pure and Applied Math, UCLA
 - Funders: NSF from chem, DMR, math