

Lies my teacher told me about density functional theory: Can machine learning find the truth?

Kieron Burke and friends
UC Irvine
Physics & Chemistry

<http://dft.uci.edu>

Motivation

- I assume you are mostly interested in strongly correlated materials
- DFT does not work for such systems.
- But DFT increasingly used for inputs to model Hamiltonians
- You'd better understand the starting point.

Contents

- A. DFT in general
- B. Hubbard dimer
- C. Electronic structure in 1d
- D. Finding the strongly-correlated functional with machine learning

A DFT in general

Molecules versus materials

- Molecular calculations are easier
- Have many small cases where very accurate results (e.g., 10 x better than DFT) available
- Chemistry far ahead in DFT accuracy
- Use term *first-principles* instead of *ab initio*
- Please extract only total energies from ground-state DFT quantities

The electronic structure problem

- Use atomic units
- Born-Oppenheimer approximation
- All non-relativistic (but added back in)
- Wavefunctions antisymmetric and normalized
- Only discuss ground-state electronic problem here, but many variations.

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

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

where the kinetic and elec-elec repulsion energies are

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

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.

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

HK theorem (1964)

- Makes TF an approximation to an exact theory
- Can find both ground-state density and energy via Euler equation

- 1 Rewrite variational principle (Levy 79):

$$\begin{aligned} E &= \min_{\Psi} \langle \Psi | \hat{T} + \hat{V}_{ee} + \hat{V} | \Psi \rangle \\ &= \min_n \left\{ F[n] + \int d^3r v(\mathbf{r})n(\mathbf{r}) \right\} \end{aligned}$$

where

$$F[n] = \min_{\Psi \rightarrow n} \langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle$$

- ▶ The minimum is taken over all positive $n(\mathbf{r})$ such that $\int d^3r n(\mathbf{r}) = N$
 - 2 The external potential $v(\mathbf{r})$ and the hamiltonian \hat{H} are determined to within an additive constant by $n(\mathbf{r})$
- P. Hohenberg and W. Kohn, Phys. Rev. **136**, B 864 (1964).
 - M. Levy, Proc. Natl. Acad. Sci. (U.S.A.) **76**, 6062 (1979).

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

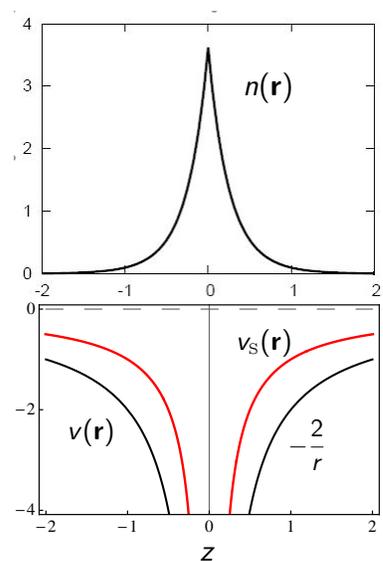
$$F = T + V_{ee} = T_S + U + E_{XC}$$

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.



Three key ingredients for a breakthrough

- A theorem (HK)
- A computational recipe (KS equations)
- An approximation (LDA)

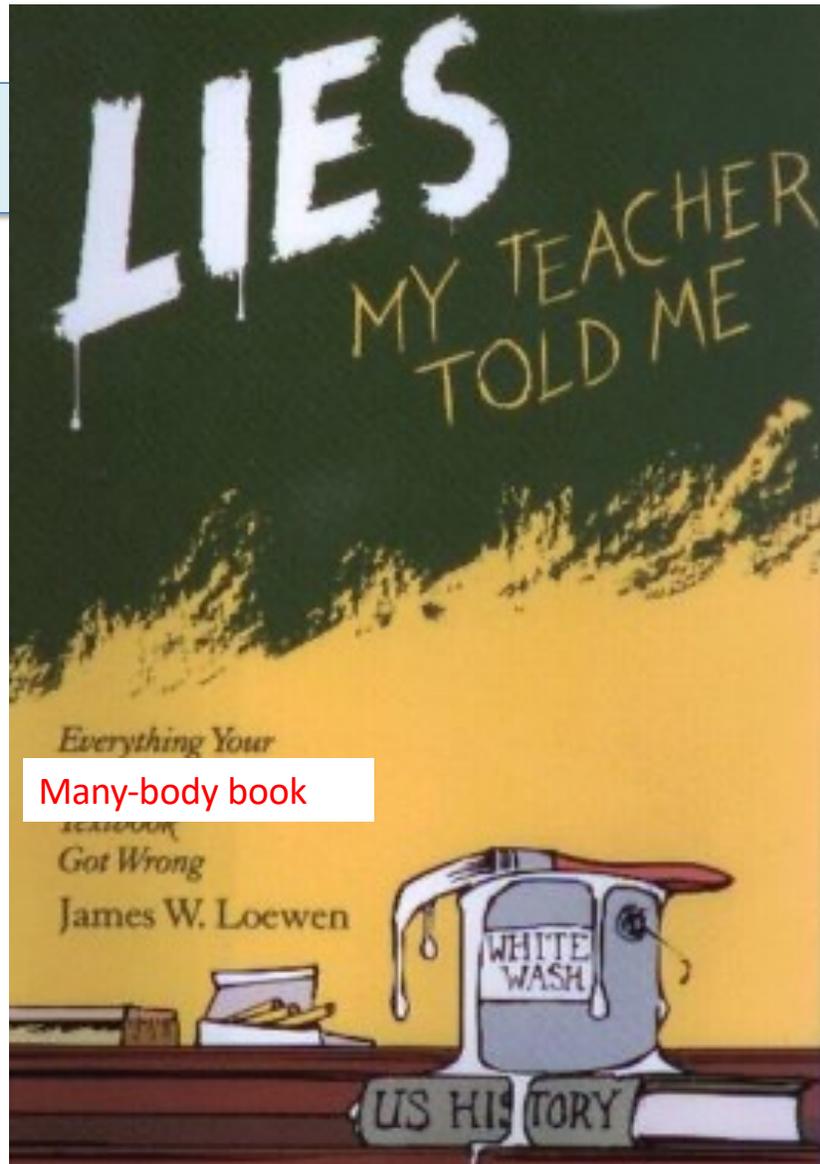
Today's commonly-used XC functionals

- Local density approximation (LDA) $E_X^{\text{LDA}}[n] = A_X \int d^3r n^{4/3}(\mathbf{r})$
 - Uses only $n(\mathbf{r})$ at a point. $A_X = -(3/4)(3/\pi)^{1/3} = -0.738$.
- Generalized gradient approx (GGA)
 - Uses both $n(\mathbf{r})$ and $|\nabla n(\mathbf{r})|$
 - Should be more accurate, corrects overbinding of LDA
 - Examples are PBE and BLYP
- Global Hybrid:
 - Mixes some fraction of HF
 - Examples are B3LYP and PBE0
- Many, many improvements on these

Applications

- 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).
- Room temperature superconductors (under pressure), predicted by DFT calculations, then made.
- Understanding Jupiter's interior using Juno probe
- Latest generation of intel chips (needed for Mac airbook) is half-size and Pb-free with help of DFT.
- Physiology of gilt-head seabream in acidic water
- Effect of water hardness on taste of coffee

B. Hubbard dimer



Many-body book

Understanding the KS idea

- Correl 21 book chapter

Lies My Teacher Told Me About Density Functional Theory: Seeing Through Them with the Hubbard Dimer
Kieron Burke and John Kozłowski arXiv:2108.11534

- Ground-state review:

The Hubbard dimer: a density functional case study of a many-body problem D J Carrascal D.J., Ferrer, J., Smith, J. and KB 2015 *J. Phys.: Condens. Matter* **27** 393001

- Linear-response TDDFT review:

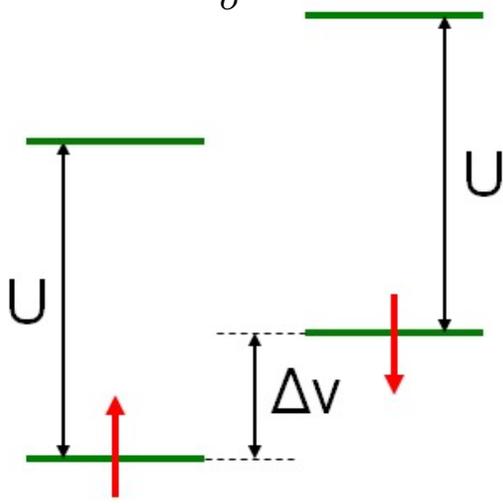
Linear response time-dependent density functional theory of the Hubbard dimer Carrascal, D.J., Ferrer, J., Maitra, N. and KB. Linear response time-dependent density functional theory of the Hubbard dimer. *Eur. Phys. J. B* **91**, 142 (2018).

- General intro to DFT in real-space:

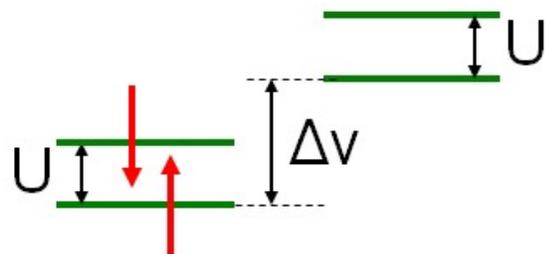
DFT in a nutshell, Kieron Burke, Lucas O. Wagner, *Int. J. Quant. Chem.* **113**, 96-101 (2013).

Asymmetric 2-site Hubbard

$$\hat{H} = -t \sum_{\sigma} (\hat{c}_{1\sigma}^{\dagger} \hat{c}_{2\sigma} + h.c.) + U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow} + \sum_i v_i \hat{n}_i$$



Mott-Hubbard



Charge-Transfer

E as function of Δv

- $E(\Delta v)$ is analytic
- Symmetric

$$E = -\sqrt{1 + (U/2)^2} + U/2, \quad \Delta n = 0$$

- Tight-binding

$$E = -\sqrt{1 + \Delta v^2}, \quad \Delta n = -2 \frac{\Delta v}{\sqrt{1 + \Delta v^2}} \quad (U=0) \quad \Delta v = \frac{\Delta n}{\sqrt{4 - \Delta n^2}}$$

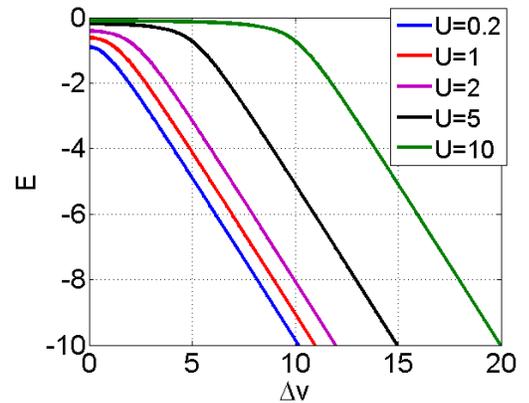
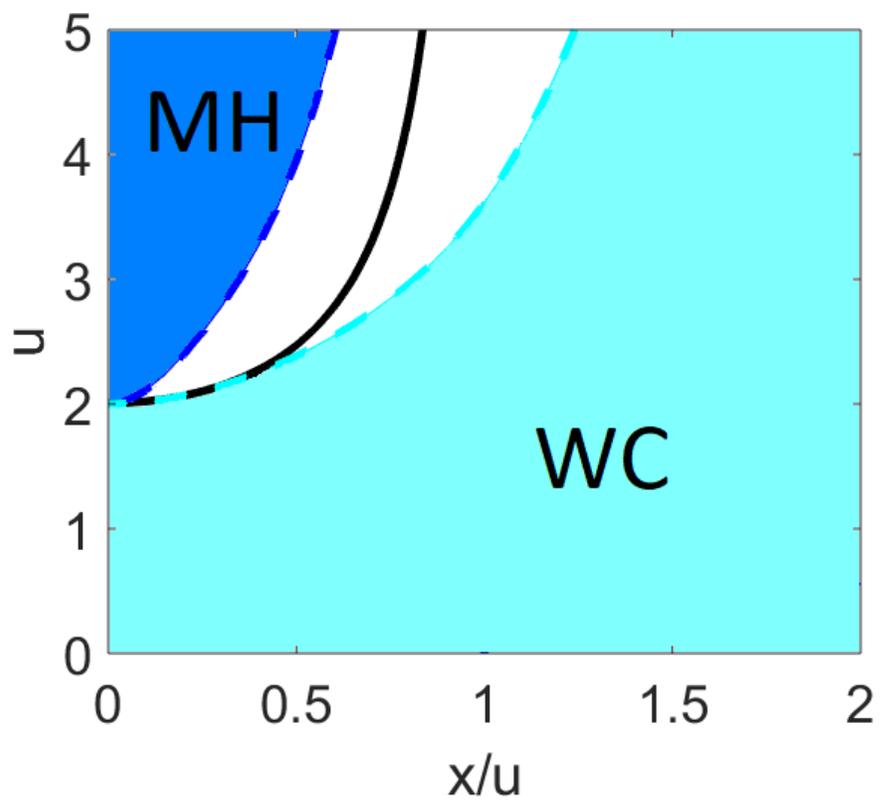


FIG. 2. Exact ground-state energy of the Hubbard dimer as a function of Δv for several values of U . The qualitative behavior changes as Δv passes through U .

Weak and strong correlation



HK I

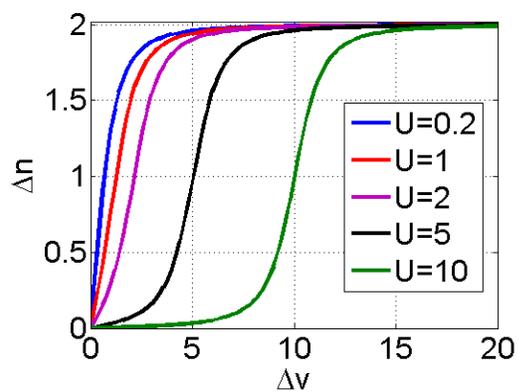


FIG. 3. Ground-state occupation of the Hubbard dimer as function of Δv for several values of U .

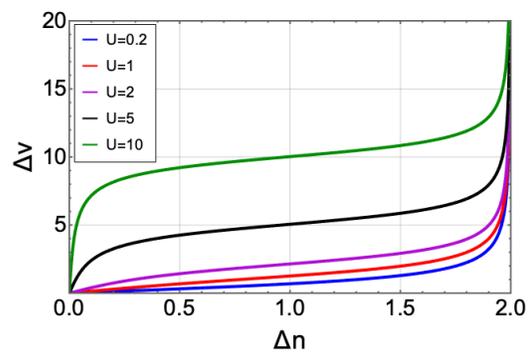


FIG. 4. Ground-state potential difference as a function of Δn for several values of U .

HK II: Part of E is universal

$$F_U(n_1) = \min_{\Psi \rightarrow n_1} \langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle = \max_{\Delta v} \left\{ E(\Delta v) - \Delta v \Delta n / 2 \right\}.$$

- Simply means F independent of Δv , i.e., is the same for every different Δv

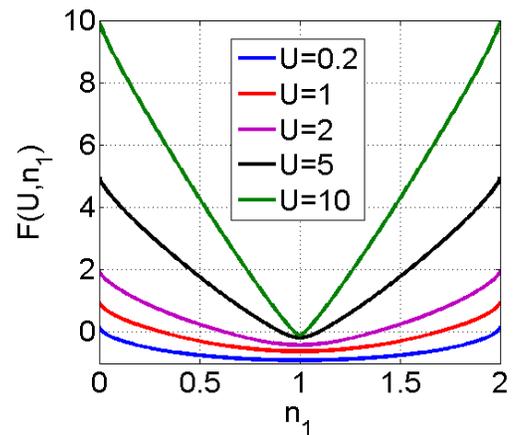


FIG. 5. Universal part of the energy function(al) of a Hubbard dimer as a function of n_1 for several values of U . As U increases, F tends to $U|1-n_1|$.

HK III: Density variational principle

$$E(\Delta v) = \min_{n_1} \left\{ F_U(n_1) + \Delta v \Delta n / 2 \right\}.$$

- Euler equation

$$\frac{dF_U(n_1)}{dn_1} - \frac{\Delta v}{2} = 0;$$

KS DFT

- Consider fake unique non-interacting electrons with same density

$$F_U(n_1) = T_S(n_1) + U_H(n_1) + E_{XC}(n_1). \quad U_H = \frac{U}{2}(n_1^2 + n_2^2),$$

- Euler equation implies

$$v_{S,i} = v_i + Un_i + \frac{\partial E_{XC}}{\partial n_i}$$

KS potential components

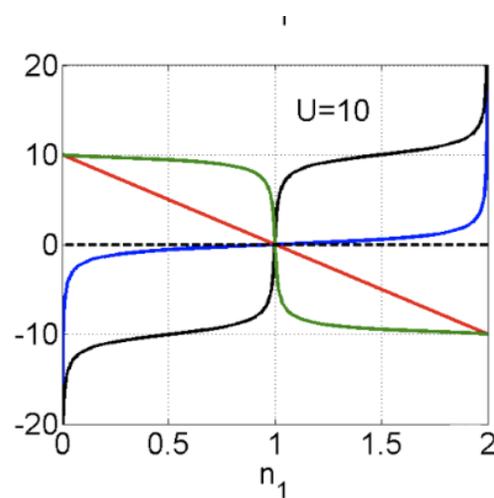
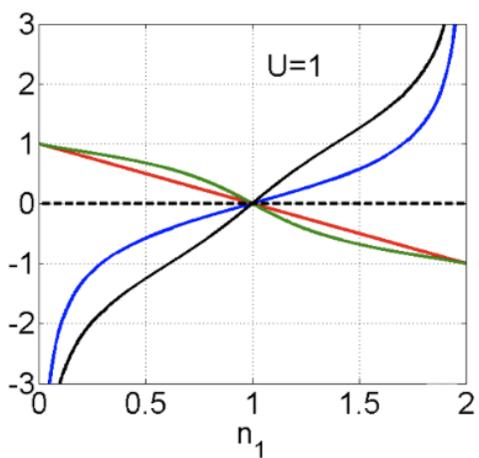
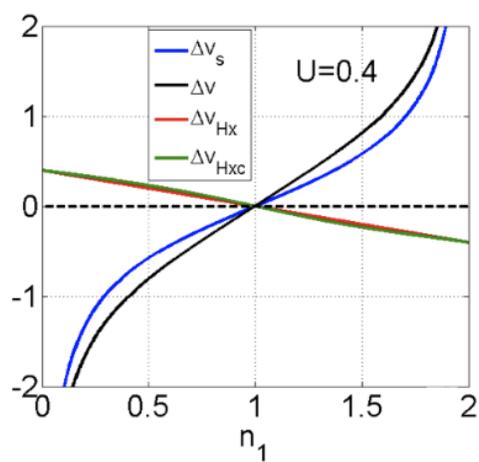


FIG. 7. Plots of Δv_s (blue) and its components, the one-body potential Δv (black), the Hartree plus exchange potentials, $U\Delta n/2$ (red), and the same with correlation added, $U\Delta n/2 + \Delta v_c$ (green) plotted against n_1 for various values of U .

Comments

- Ground-state DFT is a machinery for extracting ground-state electronic energies
- Almost all calculations use KS scheme plus approximate XC
- 99% of all applications of gs DFT are to find E as function of nuclear coordinates
- Even the density can be extracted from E
- Uses about 30% of NSERC's supercomputers



Fundamental gap

$$I = E(N - 1) - E(N)$$

$$A = E(N) - E(N + 1).$$

$$E_g = I - A,$$

$$E_g = I - A,$$

$$\epsilon^{\text{HOMO}} = -I,$$

$$E_g = E_{gs} + \Delta_{\text{XC}}$$

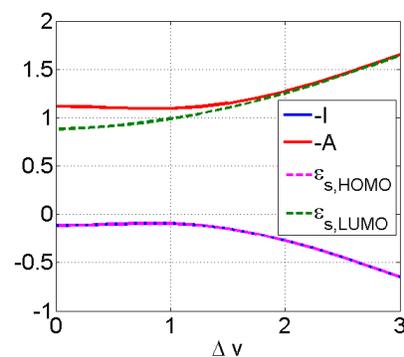


FIG. 11. Plot of $-A$, $-I$, ϵ^{HOMO} , and ϵ^{LUMO} as a function of Δv with $U = 1$ and $2t = 1$.

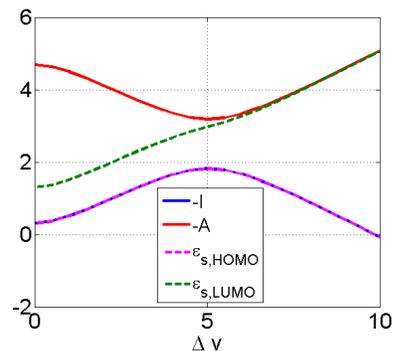


FIG. 12. Plot of $-A$, $-I$, ϵ^{HOMO} , and ϵ^{LUMO} as a function of Δv with $U = 5$ and $2t = 1$.

Spectral functions, real and KS

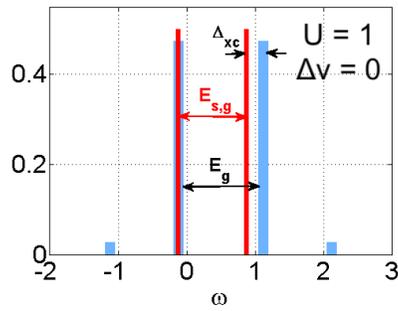


FIG. 9. Spectral function of the symmetric dimer for $U = 1$ and $\Delta v = 0$. The physical MB peaks are plotted in blue, the KS in red. Here $I = 0.1$, $A = -1.1$, and $\varepsilon^{LU} = 0.9$.

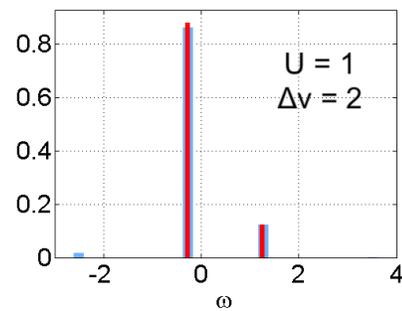


FIG. 11. Same as Fig. 9, but now $U = 1$, $\Delta v = 2$. Here $I = 0.27$, $A = -1.27$, and $\varepsilon^{LU} = 1.25$.

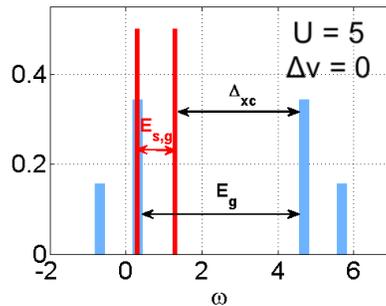


FIG. 10. Same as Fig. 9, but now $U = 5$. Here $I = -0.3$, $A = -4.7$, and $\varepsilon^{LU} = 1.3$. Note that the KS gap remains unchanged by the alteration of U because $\Delta n = 0$ in both cases.

Most important point

- When system is strongly correlated, KS spectral function not close to physical Green's function
- But all XC needs do is yield accurate energy differences
- So when KS spectral function 'fails', DFT may still be working fine.
- What has failed is the naïve picture of the KS excitations as quasiparticle excitations

Important points

- KS-DFT is NOT mean-field theory
- The KS-band gap does not match charge gap, and is not supposed to
- KS transitions are NOT quasi-particle excitations
- TD-DFT shows how to correct KS transitions to OPTICAL excitations
- For weakly-correlated systems, KS transitions often interpreted as excitations, but that's NOT based on DFT

About other stuff

- DFT+U is outside first-principles DFT but, when used carefully and appropriately, is a good fix
- DMFT is not a first-principles scheme for real-space Hamiltonians at the same level as DFT
- DMFT is an excellent approach to capturing moderately strong fluctuations
- The KS Hamiltonian is usually well-defined and is that unique single-particle Hamiltonian which has the exact ground-state density

C. Electronic structure in 1d

Importance of real-space H

- DFT approximations rely on H being real space
- Theorems can apply to lattice systems, but approximations do not transfer
- LDA is universal semiclassical limit of ALL electronic systems
- Expansion around limit is asymptotic and not captured by single GGA.
- Actually, in general HK fails on a lattice! See Penz and van Leeuwen [arXiv:2106.15370](https://arxiv.org/abs/2106.15370)

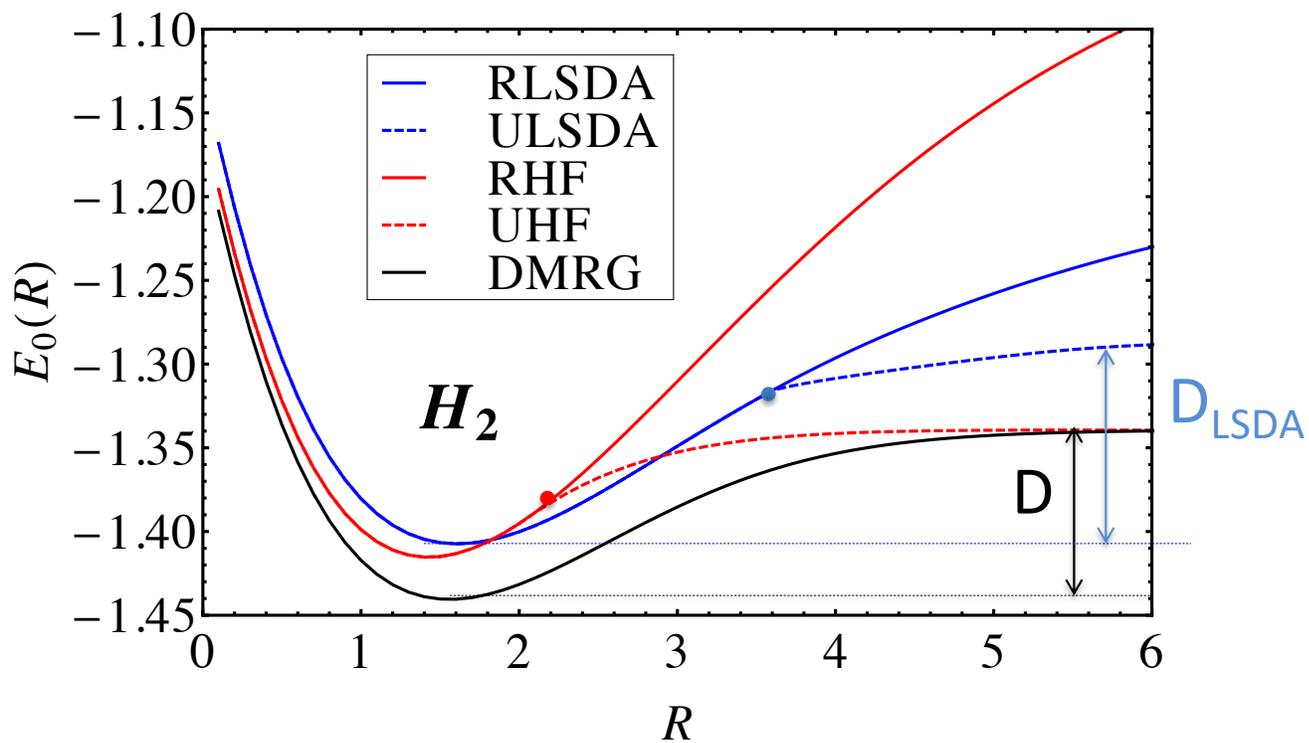
Electronic structure in 1d

- Much easier and faster in 1d
- Use DMRG to get essentially exact solution
- Use grid of about 20 points per atom
- Choose single exponential (mimicking soft-Coulomb repulsion) for both e-e repulsion and e-nuc attraction
- Do DMRG on 1d uniform gas to find LDA

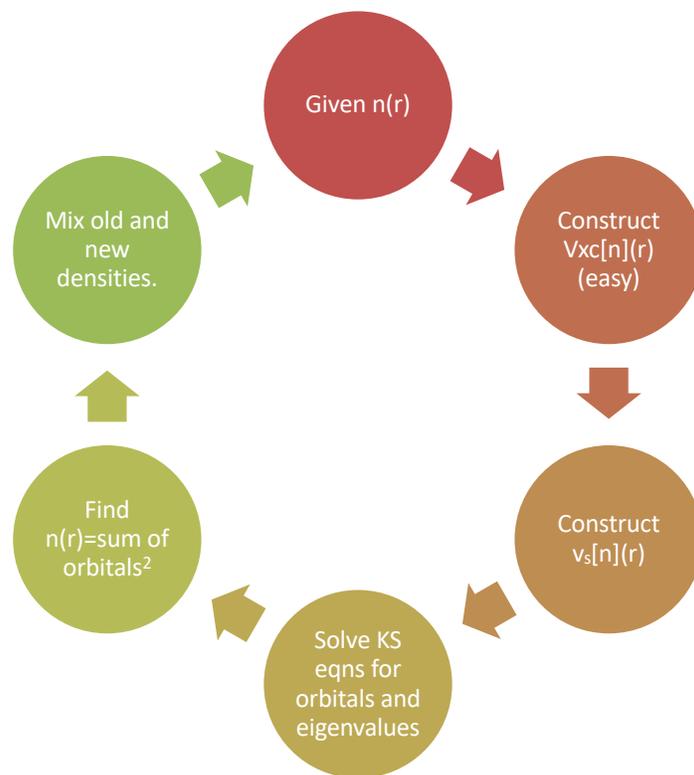
Generic failures of DFT approximations

- As a bond is stretched, eventually break spin symmetry but yield correct dissociated limit
- Exact wavefunction and exact KS wavefunctions remain as singlets
- Happens with almost all other practical DFT approximations as they contain local components
- Problem worsens with H_4 as now 4 broken symmetry solutions

Prototype of Strong/static correlation

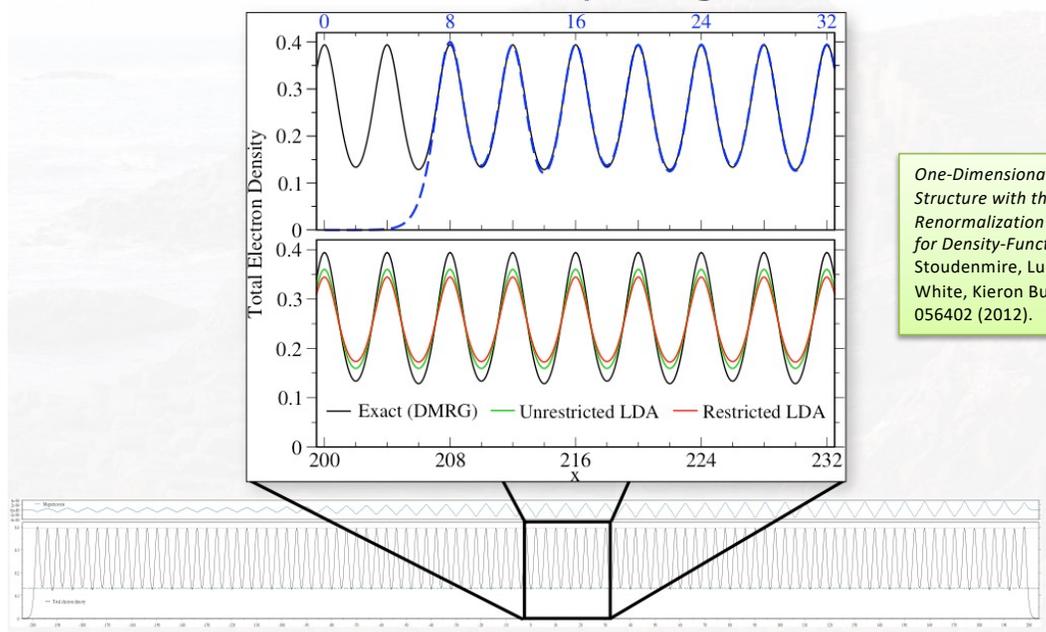


KS algorithm (standard)



Example of long chains

DMRG powerful enough to solve a chain of 100 stretched soft Hydrogen atoms



Mott-Hubbard gap

- Classic prototype of condensed matter
- Infinite chain of H atoms
- When lattice spacing is large, must be an insulator
- But with one electron per site, always a band metal

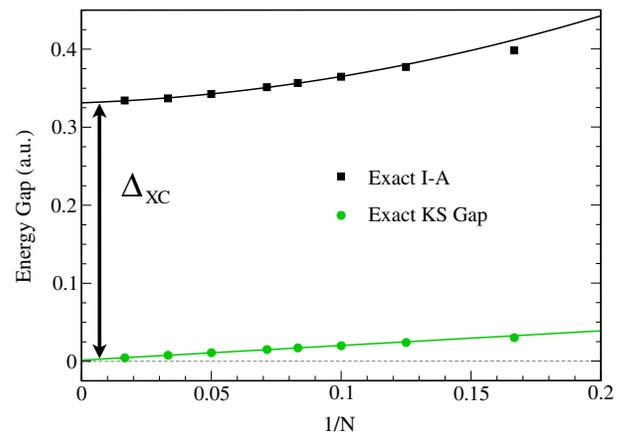


FIG. 3 (color online). Exact gaps for chains of N soft hydrogen atoms with atomic separation $b = 4$ (error bars are less than symbol sizes). The upper curve is a quadratic fit of exact gaps of the largest six systems and extrapolates to a finite value $E_g \approx 0.33$. The exact Kohn-Sham gaps, in contrast, extrapolate to zero showing that for $N \rightarrow \infty$ the true KS system is metallic (lower curve is a linear fit of exact KS gaps of the largest six systems).

Possible problems with KS calculations for strongly correlated systems

- Uniqueness of the KS potential
 - No problem, because of HK
- Existence of KS potential
 - Uniqueness does not guarantee existence, and neither HK nor Levy/Lieb prove this.
 - Always find it exists for all systems we've looked at.
- Convergence
 - Even if a unique KS exists, what if you can never find it?
 - What if you get stuck in an endless limit cycle?

Interacting inversion (insanity)

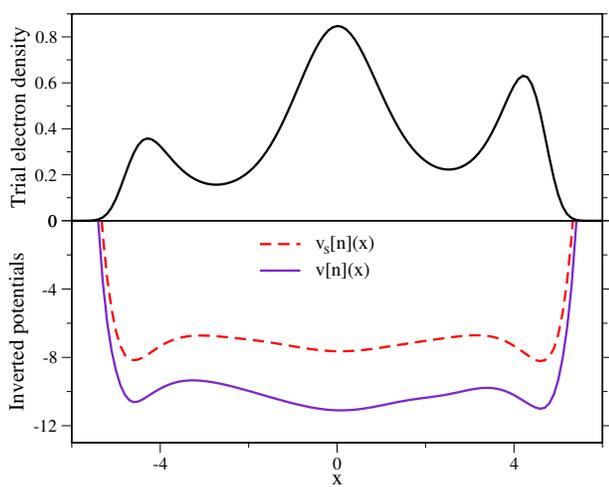
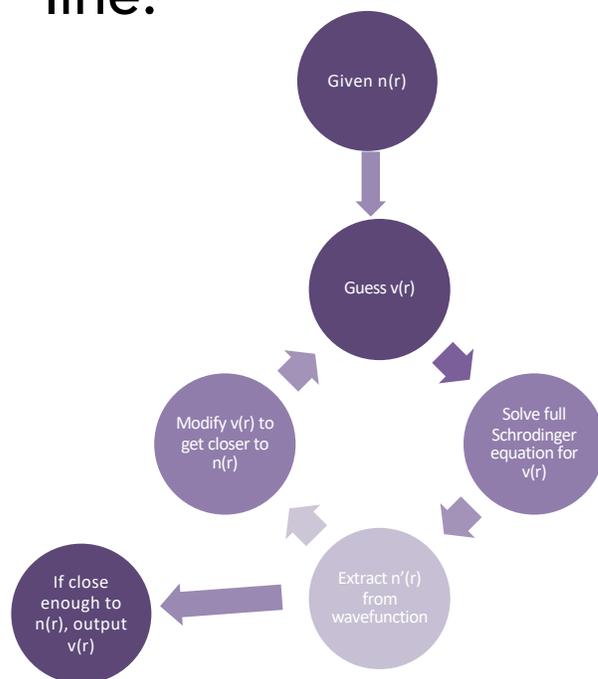
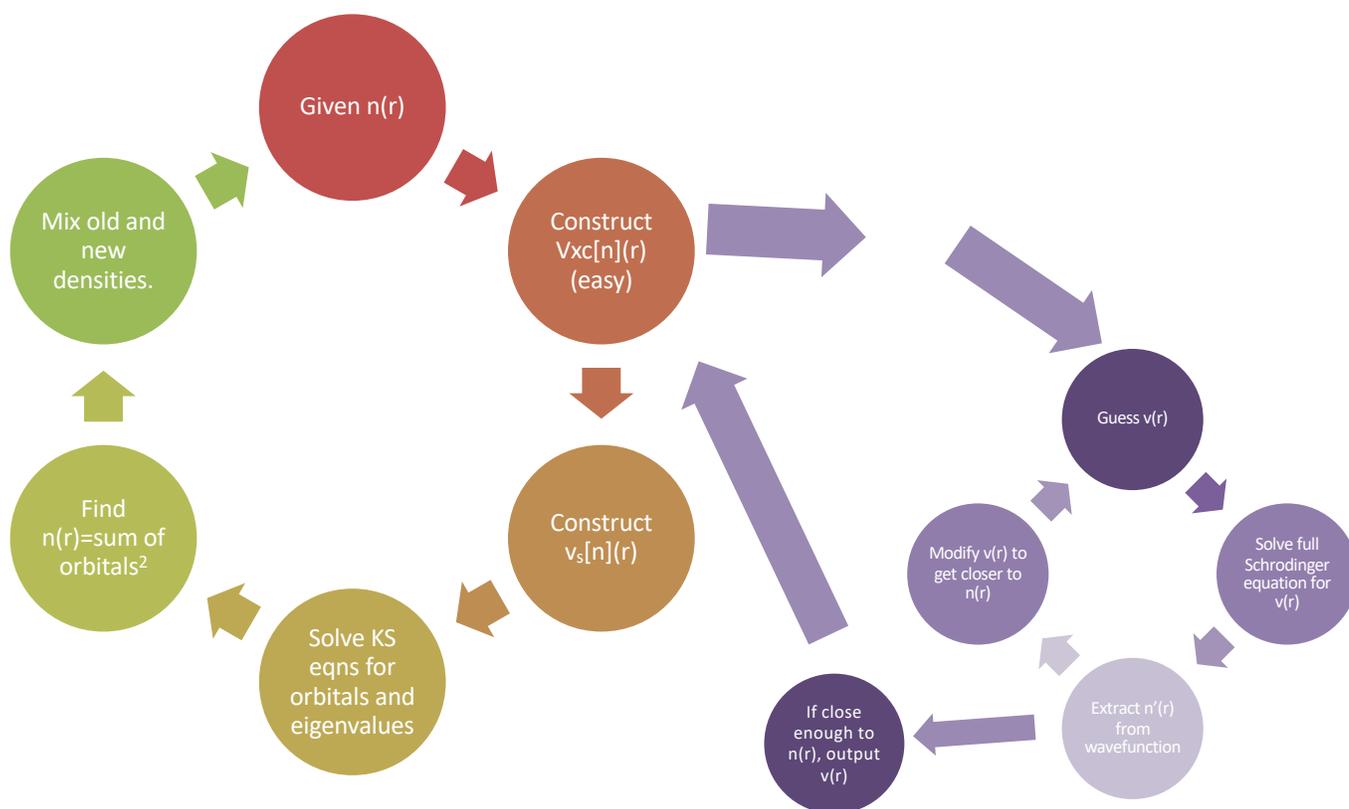


FIG. 2. Arbitrary density inversion for non-interacting and interacting potentials.

- To find the purple line:



KS algorithm (with exact XC)



Proof of convergence

- Lemma: $\int d^3r (v[n'](\mathbf{r}) - v[n](\mathbf{r}))(n'(\mathbf{r}) - n(\mathbf{r})) < 0$.
- Oops: Already proven by Gritsenko and Baerends (2005) – see our erratum.
- Consequence: take one step for some
 - $dE/d\lambda$ always < 0 at ends of curve
 - Guarantees a minimum
 - Can prove always converges for $\lambda < \lambda_c$
 - Assume Hilbert space finite

Guaranteed Convergence of the Kohn-Sham Equations Lucas O. Wagner, E. M. Stoudenmire, Kieron Burke, Steven R. White, Phys. Rev. Lett. **111**, 093003 (2013).

Conditional Probability DFT

Bypassing the Energy Functional in Density Functional Theory: Direct Calculation of Electronic Energies from Conditional Probability Densities

Ryan J. McCarty¹, Dennis Perchak¹, Ryan Pederson², Robert Evans³, Yiheng Qiu,²
Steven R. White² and Kieron Burke^{1,2,*}

¹Department of Chemistry, University of California, Irvine, California 92697, USA

²Department of Physics and Astronomy, University of California, Irvine, California 92697, USA

³H H Wills Physics Laboratory, University of Bristol, Bristol BS8 1TL, United Kingdom

(Received 3 July 2020; accepted 16 November 2020; published 23 December 2020)

Density functional calculations can fail for want of an accurate exchange-correlation approximation. The energy can instead be extracted from a sequence of density functional calculations of conditional probabilities (*CP* DFT). Simple *CP* approximations yield usefully accurate results for two-electron ions, the hydrogen dimer, and the uniform gas at all temperatures. *CP* DFT has no self-interaction error for one electron, and correctly dissociates H_2 , both major challenges. For warm dense matter, classical *CP* DFT calculations can overcome the convergence problems of Kohn-Sham DFT.

DOI: 10.1103/PhysRevLett.125.266401

PHYSICAL REVIEW LET

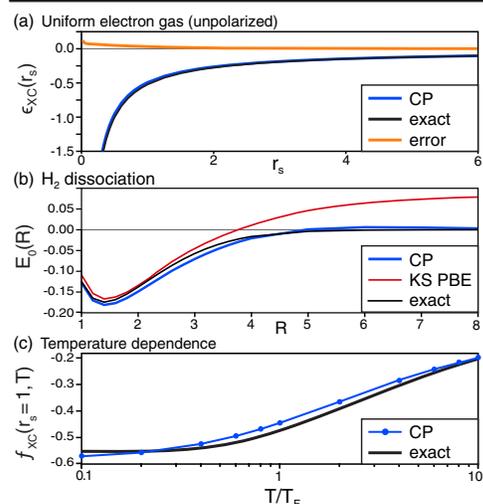
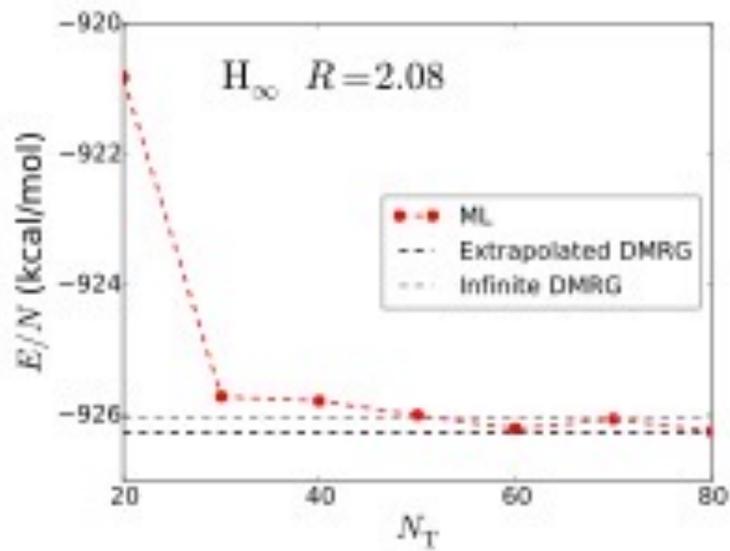


FIG. 1. CP (blue curve) and exact (black curve). (a) XC energy per particle in uniform gas at increasing Wigner-Seitz radii (r_s) and $T = 0$, (b) binding energy curve for H_2 (red is KS DFT using PBE [14]), and (c) XC free energy per particle at $r_s = 1$ as a function of reduced temperature (T_F is the Fermi temperature). Exact from Ref. [17] in (a), Ref. [18] in (c). Hartree atomic units used throughout. See supplemental materials for numerical values.

D. Finding the functional with machine learning

Convergence for infinite chain

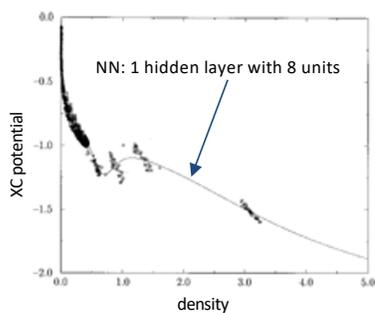


Kieron Burke

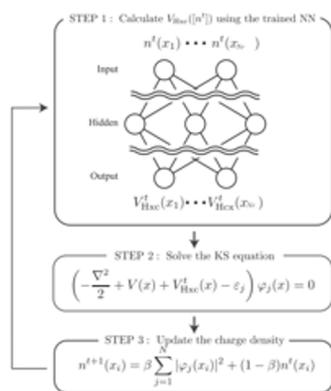
Theory

Pure density functional for strong correlations and the thermodynamic limit from machine learning Li Li, Thomas E. Baker, Steven R. White, Kieron Burke, *Phys. Rev. B* **94**, 245129 (2016).

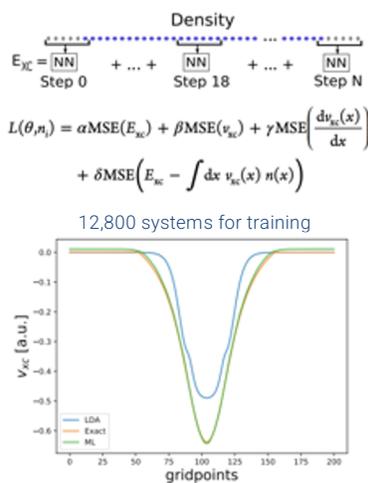
Learning to approximate XC ~~functionals~~ potentials



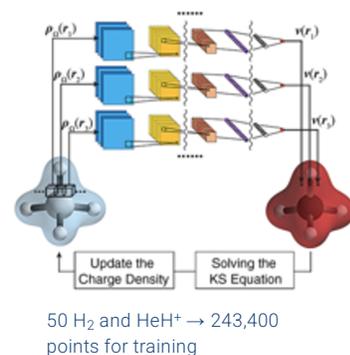
J. Chem. Phys. 105, 9200 (1996)



J. Chem. Phys. 148, 241737 (2018)



J. Phys. Chem. Lett. 10, 6425 (2019)



J. Phys. Chem. Lett. 10, 7264 (2019)

Supreme importance of energy

- Almost all applications of ground-state KS-DFT are designed to produce E (not $v(\mathbf{r})$, not $n(\mathbf{r})$).
- Most energy errors are dominated by the functional error, not the density error (and therefore not the potential error).
- Map from $n(\mathbf{r})$ to $v(\mathbf{r})$ notoriously tricky.

Understanding and reducing errors in density functional calculations Min-Cheol Kim, Eunji Sim, Kieron Burke, Phys. Rev. Lett. **111**, 073003 (2013).

Why the density?

- Density is all possible derivatives of energy with respect to the one-body potential

$$n(\mathbf{r}) = \delta E / \delta v(\mathbf{r})$$

- Eg density contains the force, i.e., dE/dR , at a data point
- Bypasses all need for analysis of relation between density and potential.

Kohn-Sham regularizer

PHYSICAL REVIEW LETTERS **126**, 036401 (2021)

Kohn-Sham Equations as Regularizer: Building Prior Knowledge into Machine-Learned Physics

Li Li (李力)^{1,*} Stephan Hoyer¹ Ryan Pederson² Ruoxi Sun (孙若溪)¹ Ekin D. Cubuk¹
Patrick Riley¹ and Kieron Burke^{2,3}

¹Google Research, Mountain View, California 94043, USA

²Department of Physics and Astronomy, University of California, Irvine, California 92697, USA

³Department of Chemistry, University of California, Irvine, California 92697, USA



(Received 18 September 2020; accepted 3 December 2020; published 20 January 2021)



Li Li



Stephan
Hoyer



Ruoxi
Sun



Ekin
Dogus
Cubuk



Patrick
Riley



Ryan
Pederson

Kieron
Burke

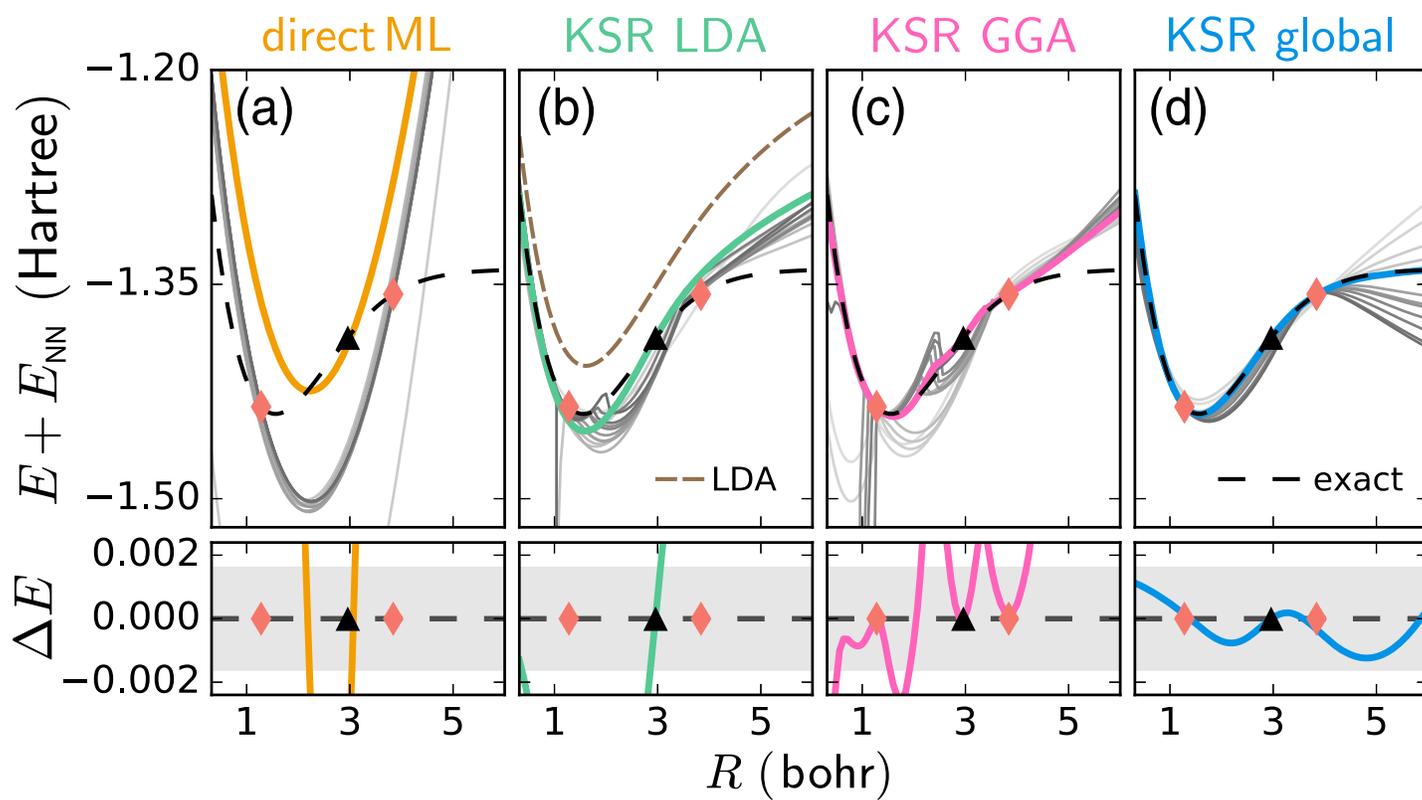
Kieron Burke

Theory Cond Matt
Google

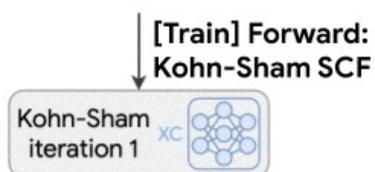
UC Irvine

48

1D H₂ binding curves



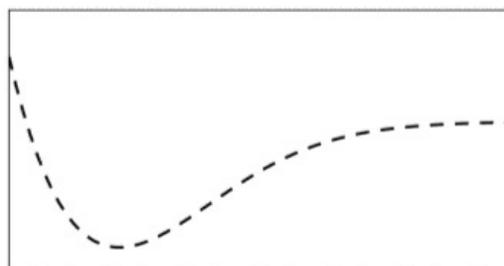
Scheme



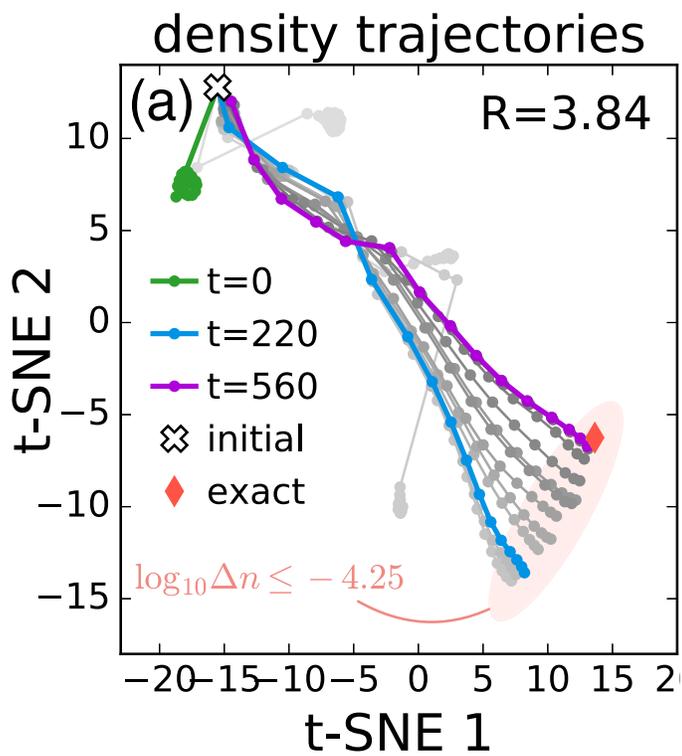
Training on two molecules



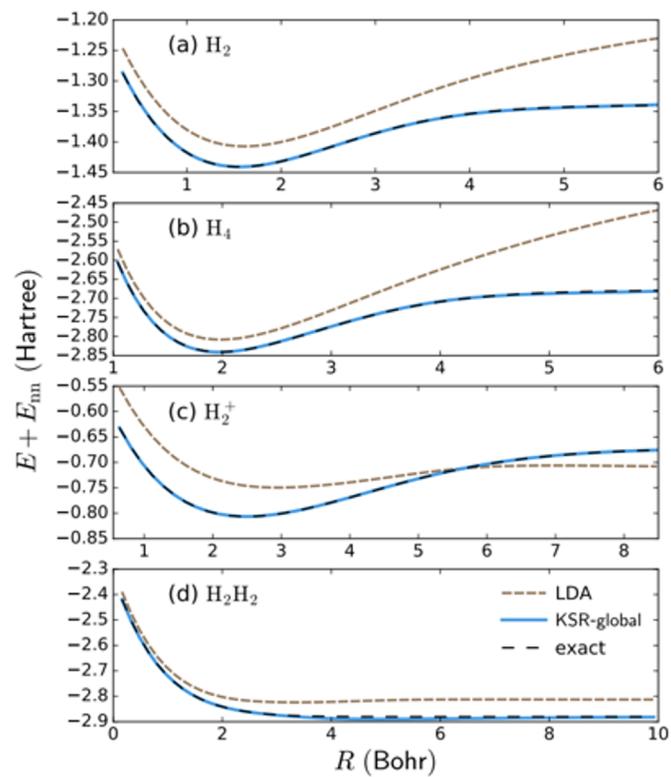
Test on H_2 dissociation curve



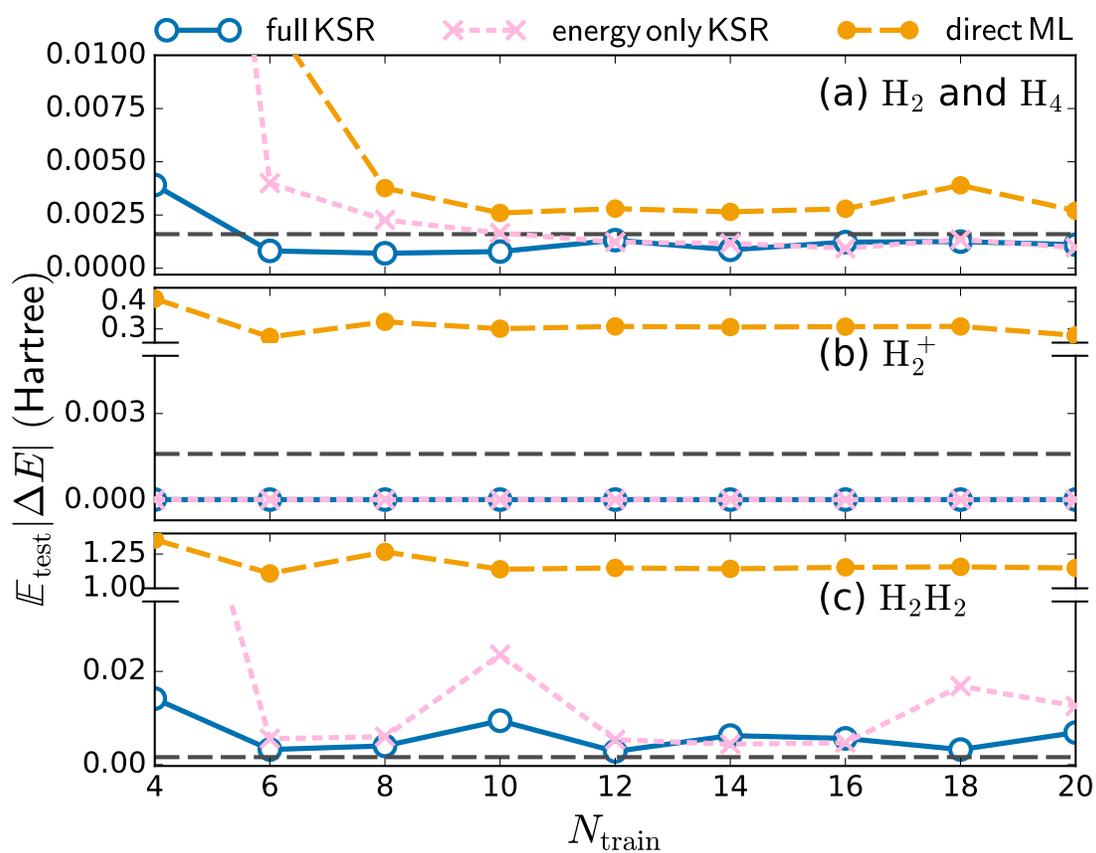
Regularization and validation



Generalization



Learning curves



Better densities via training

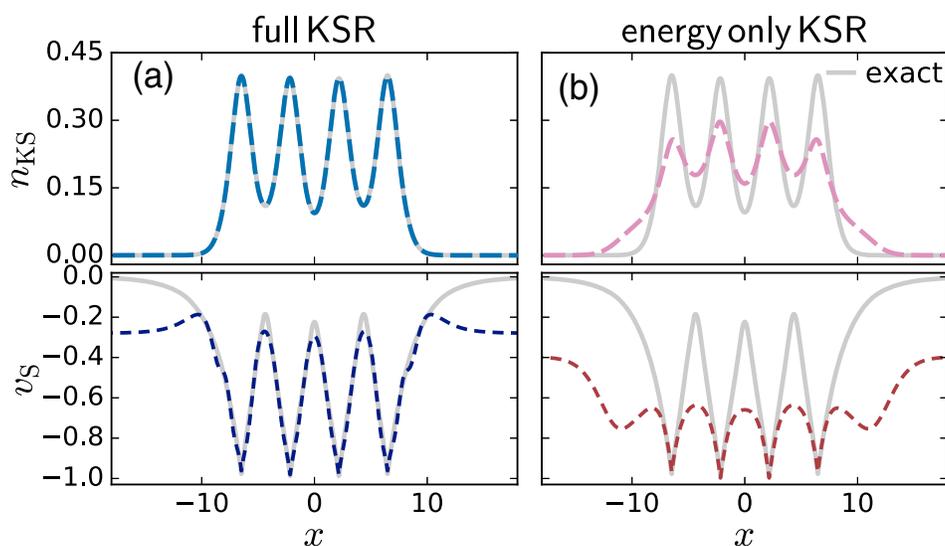


FIG. 5. Density and KS potential of H_4 with $R = 4.32$ from neural XC functionals trained with (a) full KSR (blue) and (b) energy only KSR (pink) on training set of size $N_{\text{train}} = 20$. Exact curves are in gray. v_s are shifted by a constant for better comparison.

Future directions

- Generalize to spin DFT
- Test many different variations and flavors
- Does it work for weak correlation?
- Does it work for asymmetric cases?
- Should all work for 3D just as well

Recent review of our work

ACCOUNTS

of chemical research

pubs.acs.org/accounts

Article

Learning to Approximate Density Functionals
Bhupalee Kalita, Li Li, Ryan J. McCarty, and Kieron Burke
Acc. Chem. Res. 2021, 54, 4, 818–826

Learning to Approximate Density Functionals

Published as part of the Accounts of Chemical Research special issue "Data Science Meets Chemistry".

Bhupalee Kalita, Li Li, Ryan J. McCarty, and Kieron Burke*

Cite This: *Acc. Chem. Res.* 2021, 54, 818–826

Read Online

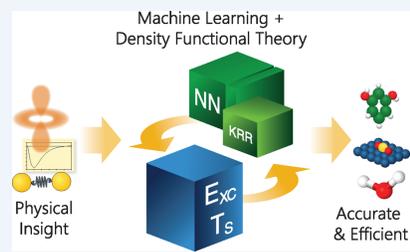
ACCESS |

Metrics & More

Article Recommendations

CONSPECTUS: Density functional theory (DFT) calculations are used in over 40,000 scientific papers each year, in chemistry, materials science, and far beyond. DFT is extremely useful because it is computationally much less expensive than *ab initio* electronic structure methods and allows systems of considerably larger size to be treated. However, the accuracy of any Kohn–Sham DFT calculation is limited by the approximation chosen for the exchange–correlation (XC) energy. For more than half a century, humans have developed the art of such approximations, using general principles, empirical data, or a combination of both, typically yielding useful results, but with errors well above the chemical accuracy limit (1 kcal/mol). Over the last 15 years, machine learning (ML) has made major breakthroughs in many applications and is now being applied to electronic structure calculations. This recent rise of ML begs the question: Can ML propose or improve density functional approximations? Success could greatly enhance the accuracy and usefulness of DFT calculations without increasing the cost.

In this work, we detail efforts in this direction, beginning with an elementary proof of principle from 2012, namely, finding the kinetic energy of several Fermions in a box using kernel ridge regression. This is an example of orbital-free DFT, for which a successful general-purpose scheme could make even DFT calculations run much faster. We trace the development of that work to state-of-the-art molecular dynamics simulations of resorcinol with chemical accuracy. By training on *ab initio* examples, one bypasses the need to find the XC functional explicitly. We also discuss how the exchange–correlation energy itself can be modeled with such methods, especially for strongly correlated materials. Finally, we show how deep neural networks with differentiable programming can be used to construct accurate density functionals from very few data points by using the Kohn–Sham equations themselves as a regularizer. All these cases show that ML can create approximations of greater accuracy than humans, and is capable of finding approximations that can deal with difficult cases such as strong correlation. However, such ML-designed functionals have not been implemented in standard codes because of one last great challenge: generalization. We discuss how effortlessly human-designed functionals can be applied to a wide range of situations, and how difficult that is for ML.



Summary

- DFT is an enormously successful approach to finding ground-state energies in electronic structure of everything.
- `Failures' of DFT for strong correlation are failures of approximations (bias toward local approximations)
- KS spectral function not close to true function is not a failure of DFT.
- Can show
 - exact gap does not match KS gap
 - KS equations with exact $E_{xc}[n]$ converge
 - ML can produce fully non-local functionals which capture strong correlation.
- Thanks to US DOE for funding.