ML in electronic structure (gsDFT)


Bhupalee Kalita

Kieron Burke and friends
UC Irvine
Physics & Chemistry

http://dft.uci.edu


Ryan Pederson

1

Outline

- General ML
- ML in electronic structure
- ML to make density functionals
  - Orbital-free DFT
  - XC energy for strong correlation
- The future

2

A. What is machine learning?

- Unholy alliance of computer science, math, and statistics.
- Initial excitement in 1990's with first neural networks (NN's), inspired by brain architecture
- Computing power and data acquisition lead to revolution about 2010
- Last 5 years, increasingly dominated by neural networks (deep learning)
- ML=applied statistics + non-linearity + GPUs

3

Challenges for physical scientists

- ML designed for generic real-world data, much involving human caprice.
- Physical science governed by underlying physical laws (eg laws of thermodynamics)
- Fits that disobey such laws are obviously junk.
- Scientists have lots of prior knowledge and intuition that is very difficult to categorize.
- General purpose ML algorithms can easily give nonsense

4

More challenges

- Computer science is NOT physical science.
- Publication in computer science is very different, because the standards are very different.
- Because of enthusiasm for ML, lots of papers appearing in chemistry and materials.
- More than ½ do not meet basic quality standards for reproducibility, test selection, etc.
- 90% are likely to prove worthless (true of all good research).
- But 10% are first papers ever doing *something*

5

My qualifications

- Pioneered using ML for finding functionals (2012), with Klaus Robert-Mueller
- I teach ML for physical sciences for last 4 years
- Just gave 4-hour lecture series to all Korea
- ML is a paradigm shift in modelling
- If you're a current graduate student, you'd be crazy not to learn/do a little ML
- My ML graduate students intern at Google, etc. and take jobs at startups
- I'm not much good at ML

6

B.ML applications in electronic structure

- ML-designed force fields
  - Configuration space: Behler-Parinello, Csanyi,..
  - Compound space: s-GDML, ANI,…
- Data repositories of DFT calculations
  - NOMAD
  - Materials Genome project
- Accelerating MD,…

Editorial: Special Topic on Data-enabled Theoretical Chemistry Matthias Rupp, O. Anatole von Lilienfeld, Kieron Burke, Journal of Chemical Physics **148**, 241401 (2018)

Retrospective on a decade of machine learning for chemical discovery von Lilienfeld, O. Anatole; Burke, Kieron, *Nature Communications* **11**, 4895 (2020).

7

ML application: classical forces

- Run DFT (or better) calculations to make training set
- Train a deep neural network to create a force field

8

Data Descriptor | Open Access | Published: 01 May 2020

The ANI-1ccx and ANI-1x data sets, coupled-cluster and density functional theory properties for molecules

Justin S. Smith, Roman Zubatyuk, Benjamin Nebgen, Nicholas Lubbers, Kipton Barros, Adrian E. Roitberg, Olexandr Isayev & Sergei Tretiak

Scientific Data **7**, Article number: 134 (2020) | Cite this article
1703 Accesses | **3** Citations | **12** Altmetric | Metrics

Abstract

Maximum diversification of data is a central theme in building generalized and accurate machine learning (ML) models. In chemistry, ML has been used to develop models for predicting molecular properties, for example quantum mechanics (QM) calculated potential energy surfaces and atomic charge models. The ANI-1x and ANI-1ccx ML-based general-purpose potentials for organic molecules were developed through active learning; an automated data diversification process. Here, we describe the ANI-1x and ANI-1ccx data sets. To demonstrate data diversity, we visualize it with a dimensionality reduction scheme, and contrast against existing data sets. The ANI-1x data set contains multiple QM properties from 5 M density functional theory calculations, while the ANI-1ccx data set contains 500 k data points obtained with an accurate CCSD(T)/CBS extrapolation. Approximately 14 million CPU core-hours were expended to generate this data. Multiple QM calculated properties for the chemical elements C, H, N, and O are provided: energies, atomic forces, multipole moments, atomic charges, etc. We provide this data to the community to aid research and development of ML models for chemistry.

9

DF 1

Pushing the limit of molecular dynamics with ~~*ab initio*~~ accuracy to 100 million atoms with machine learning

Weile Jia*, Han Wang†, Mohan Chen‡, Denghui Lu‡, Lin Lin¶‖, Roberto Car‖, Weinan E‖, Linfeng Zhang‖ §
*University of California, Berkeley, USA
Email: jiaweile@berkeley.edu, linlin@math.berkeley.edu
† Laboratory of Computational Physics, Institute of Applied Physics and Computational Mathematics, Beijing, China
Email: wang_han@iapcm.ac.cn
‡CAPT, HEDPS,College of Engineering, Peking University, Beijing, China
Email: mohanchen@pku.edu.cn, denghuilu@pku.edu.cn
¶Lawrence Berkeley National Laboratory, Berkeley, USA
‖ Princeton University, Princeton, USA

14 Sep 2020

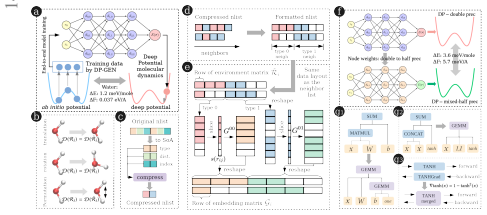


Fig. 2: Key steps in the optimized DeePMD-kit, taking water as an example.

10

C. Finding density functionals

11

Crucial theme

- Almost all ML for functionals in literature takes existing approximate forms and uses data to 'improve' them
- These are still approximations (usually semilocal) that fail in difficult cases!
- Almost all my work uses the entire density to find the exact functional to a given level of accuracy
- Creates functionals that solve problems where existing approximations fail.

12

First 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:

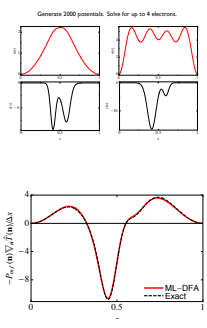
$$\hat{T}(\mathbf{n}) = \bar{T}\sum_{j=1}^{M} \alpha_j k(\mathbf{n}_j, \mathbf{n})$$

k[n,n']=exp (-∫ dx (n(x)-n'(x))²/(2σ²))

Kieron Burke ML for DFT 13

13

Test case: KS electrons in a box



kcal/mol

N	M	λ	σ	$\overline{\lvert\Delta T\rvert}$	$\lvert\Delta T\rvert^{\text{std}}$	$\lvert\Delta T\rvert^{\max}$
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)

Machine Learning Approaches toward Orbital-free Density Functional Theory: Simultaneous Training on the Kinetic Energy Density Functional and Its Functional Derivative Ralf Meyer, Manuel Weichselbaum, and Andreas W. Hauser *Journal of Chemical Theory and Computation* **2020** 16 (9), 5685-5694

Kieron Burke ML for DFT 14

14

Application: MD with ML functional

Malonaldehyde Proton Transfer

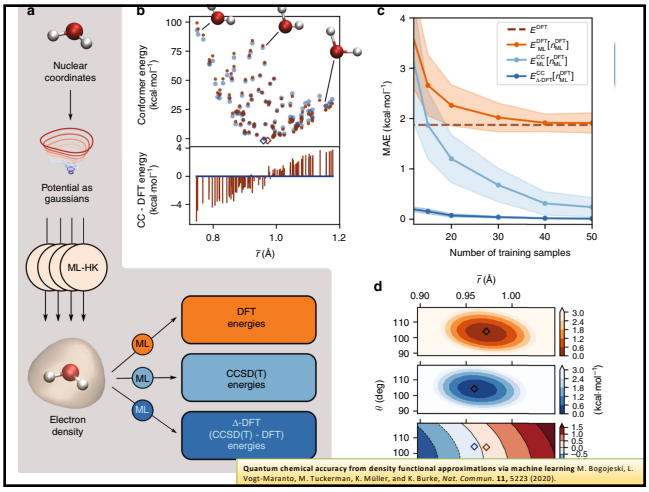
Dr. Felix Brockherde
Senior Data Scientist

Felix is a Senior Data Scientist at Fresenius Medical Care Data Solutions. He excels on challenging data-driven-projects. Before starting in the medical technology sector, he founded a consulting company working on the explainability of deep learning decisions with clients in the automotive industry and natural language-processing of technical publications in the aviation industry.
Felix taught and researched in the Machine Learning group at TU Berlin, the Institute of Pure and Applied Mathematics at UCLA, and the Max Planck Institute of Microstructure Physics. He holds a PhD in data science, a MSc in computer science and a BSc in pure mathematics.

"Bypassing the Kohn-Sham equations with machine learning", Felix Brockherde, Leslie Vogt, Li Li, Mark E. Tuckerman, Kieron Burke, Klaus-Robert Müller, *Nature Communications*, 2017

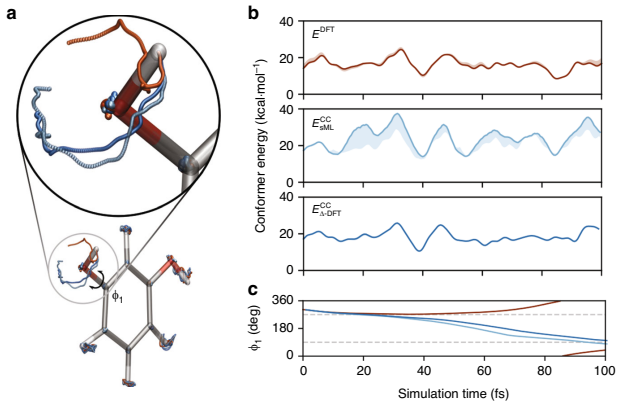
Kieron Burke ML for DFT

15



Quantum chemical accuracy from density functional approximations via machine learning M. Bogojeski, L. Vogt-Maranto, M. Tuckerman, K. Müller, and K. Burke, *Nat. Commun.* **11**, 5223 (2020).

16



17

ML for XC

- Aim: To find $E_{xc}[n]$ for static (strong) correlation
- Background
  - All standard approximations to $E_{xc}[n]$ fail as bonds are stretched.
  - Called static correlation in chemistry
  - Also effects accuracy at equilibrium for multiple bonds
  - Origin of mixing fraction of E_x in global hybrids
  - Gets worse as length of a chain grows

18

First attempts ML for $E_{xc}[n]$

Exchange-correlation potentials

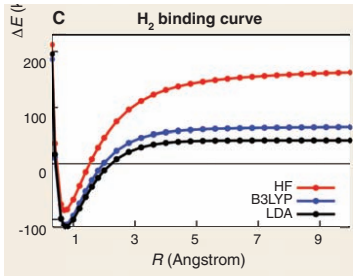
David J. Tozer, Victoria E. Ingamells, and Nicholas C. Handy
Department of Chemistry, University of Cambridge, Cambridge, CB2 1EW, United Kingdom

(Received 14 May 1996; accepted 22 August 1996)

We describe our implementation of the Zhao, Morrison, and Parr method [Phys. Rev. A **50**, 2138 (1994)] for the calculation of molecular exchange-correlation potentials from high-level *ab initio* densities. The use of conventional Gaussian basis sets demands careful consideration of the value of the Lagrange multiplier associated with the constraint that reproduces the input density. Although formally infinite, we demonstrate that a finite value should be used in finite basis set calculations. The potential has been determined for Ne, HF, N_2, H_2O, and $N_2(1.5r_e)$, and compared with popular analytic potentials. We have then examined how well the Zhao, Morrison, Parr potential can be represented using a computational neural network. Assuming $v_{xc}=v_{xc}(\rho)$, we incorporate the neural network into a regular Kohn–Sham procedure [Phys. Rev. A **140**, 1133 (1965)] with encouraging results. The extension of this method to include density derivatives is briefly outlined. © *1996 American Institute of Physics*. [S0021-9606(96)01444-4]

19

CMY Science 2008



Insights into Current Limitations of Density Functional Theory

Aron J. Cohen, Paula Mori-Sánchez, Weitao Yang*

20

$E_{xc}[n]$ from a few molecules

npj | Computational Materials

www.nature.com/npjcompumats

ARTICLE OPEN

Check for updates

Completing density functional theory by machine learning hidden messages from molecules

Ryo Nagai[1,2✉], Ryosuke Akashi[1] and Osamu Sugino[1,2]

Kohn–Sham density functional theory (DFT) is the basis of modern computational approaches to electronic structures. Their accuracy heavily relies on the exchange-correlation energy functional, which encapsulates electron–electron interaction beyond the classical model. As its universal form remains undiscovered, approximated functionals constructed with heuristic approaches are used for practical studies. However, there are problems in their accuracy and transferability, while any systematic approach to improve them is yet obscure. In this study, we demonstrate that the functional can be systematically constructed using accurate density distributions and energies in reference molecules via machine learning. Surprisingly, a trial functional machine learned from only a few molecules is already applicable to hundreds of molecules comprising various first- and second-row elements with the same accuracy as the standard functionals. This is achieved by relating density and energy using a flexible feed-forward neural network, which allows us to take a functional derivative via the back-propagation algorithm. In addition, simply by introducing a nonlocal density descriptor, the nonlocal effect is included to improve accuracy, which has hitherto been impractical. Our approach thus will help enrich the DFT framework by utilizing the rapidly advancing machine-learning technique.

npj Computational Materials (2020)6:43 ; https://doi.org/10.1038/s41524-020-0310-0

Machine-learning-based exchange correlation functional with physical asymptotic constraints
Ryo Nagai, Ryosuke Akashi, and Osamu Sugino Phys. Rev. Research 4, 013106 (2022)

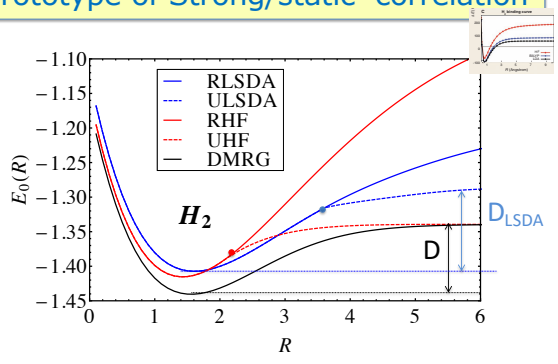
21

1D electronic structure lab

- Very difficult to get realistic benchmark results for strong correlation for bulk materials.
- Have very efficient solver (DMRG) for 1D problems
- Previously applied only to model Hamiltonians (with great effect)
- Creator of DMRG is Steve White, UCI physics

22

Prototype of Strong/static correlation



Kohn-Sham calculations with the exact functional, Lucas O. Wagner, Thomas E. Baker, E. M. Stoudenmire, Kieron Burke, Steven R. White, *Phys. Rev.* **B90**, 045109 (2014).

23

Kohn-Sham regularizer

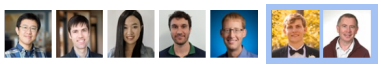
PHYSICAL REVIEW LETTERS **126,** 036401 (2021)

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

Li Li (李力)[1,*] Stephan Hoyer[1] Ryan Pederson[2] Ruoxi Sun (孙若溪)[1] Ekin D. Cubuk[1] Patrick Riley[1] and Kieron Burke[2,3]

[1]*Google Research, Mountain View, California 94043, USA*
[2]*Department of Physics and Astronomy, University of California, Irvine, California 92697, USA*
[3]*Department of Chemistry, University of California, Irvine, California 92697, USA*

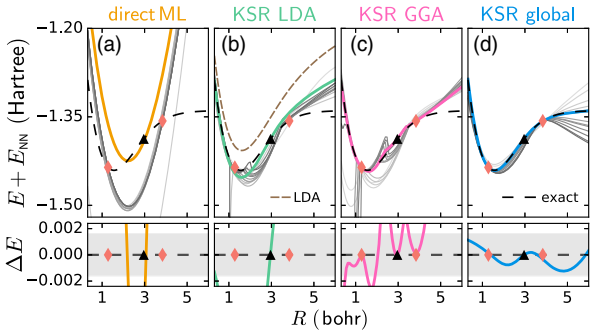
(Received 18 September 2020; accepted 3 December 2020; published 20 January 2021)



Google UC Irvine

24

1D H_2 binding curves



25

Would it work for weak correlation?

- Appeared on arXiv:2110.14846, **Generalizability of density functionals learned from differentiable programming on weakly correlated spin-polarized systems** Bhupalee Kalita, Ryan Pederson, Li Li, Kieron Burke
- Was presented Neurips workshop on differentiable programming Dec 13, 2021
- Latest version scientific article
- Tried it for (1D) molecules at equilibrium, i.e., weakly correlated system

How Well Does Kohn–Sham Regularizer Work for Weakly Correlated Systems? B. Kalita, R. Pederson, J. Chen, L. Li, and K. Burke, *J. Phys. Chem. Lett.* **13**, *11*, 2540-2547 (2022)

26

Training and testing

- Total energies and densities are generated from accurate 1D DMRG calculations with exponential approximation, solved in real space on a grid of 513 points.
- Differentiable DFT built using JAX library
- Number of KS iterations fixed based on the training or test examples.
- Trained using L-BFGS.
- Calculations are repeated for 30 random seeds and optimal parameters are chosen based on validation loss.
- Training and testing can be performed on GPU or CPU.

Training	Validation	Testing
H, He, Li Be, Be^{++}	Be^{+}	H_2, H_3, H_4, H_2^+, H_3^+, LiH, BeH_2, HeH^+, H-He-He-H^{2+}, He-H-H-He^{2+}

27

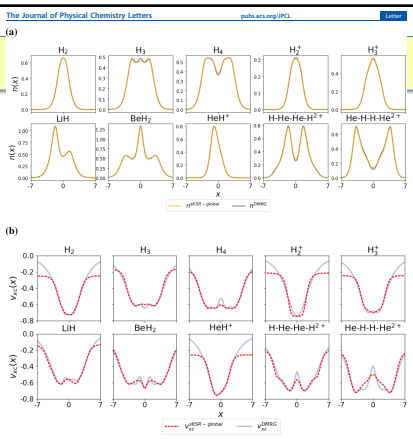
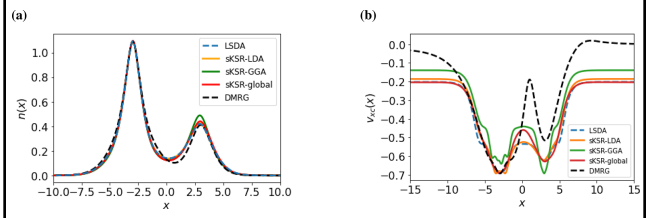


Figure 3. (a) The densities obtained using sKSR-global (orange dashes) and the exact ground-state densities (gray), (b) average XC potentials calculated from sKSR-global approximation (red dashes) to v_{xc} and their exact counterparts calculated with DMRG (light blue) for the test molecules in Table 1 at equilibrium separations. The sKSR potentials are shifted by a constant for a better comparison with the exact XC potentials. sKSR-global was trained on H, He, Li, Be, and Be^{2+} and validated on Be^{+}. Note that, in general, these 1D densities and XC potentials can differ even qualitatively from their 3D analogues.

28

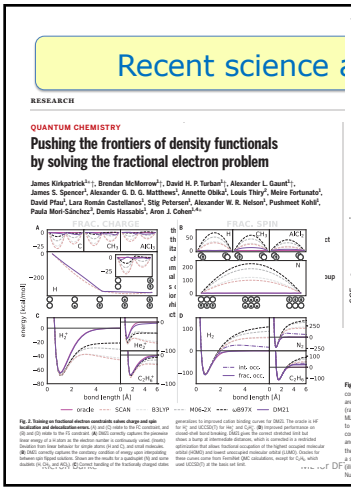
Stretched LiH



re 6. (a) The total density and (b) the average XC potentials of LiH at a bond distance of 5.92 bohr calculated with the three neura tionals as well as uniform-gas LSDA. The exact (DMRG) average XC potentials are included for comparison.

29

Recent science a

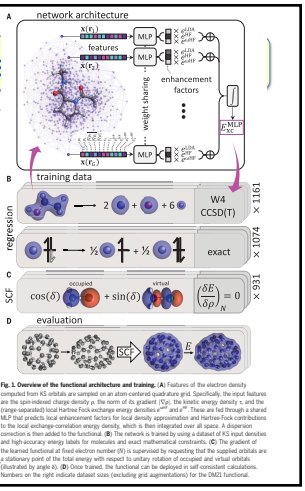


RESEARCH

QUANTUM CHEMISTRY

Pushing the frontiers of density functionals by solving the fractional electron problem

James Kirkpatrick[1,2]†, Brendan McMorrow[2]†, David H. P. Turban[2]†, Alexander L. Gaunt[2]†, James S. Spencer[2], Alexander G. D. G. Matthews[2], Annette Obika[2], Louis Thiry[2], Meire Fortunato[2], David Pfau[2], Lara Román Castellanos[2], Stig Petersen[2], Alexander W. R. Nelson[2], Pushmeet Kohli[2], Paula Mori-Sánchez[3], Dennis Hassabis[1], Aron J. Cohen[1,4]*

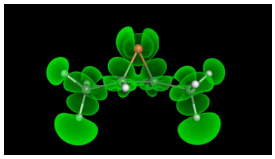


Layman's view

QUANTUM PHYSICS

Quantum Complexity Tamed by Machine Learning

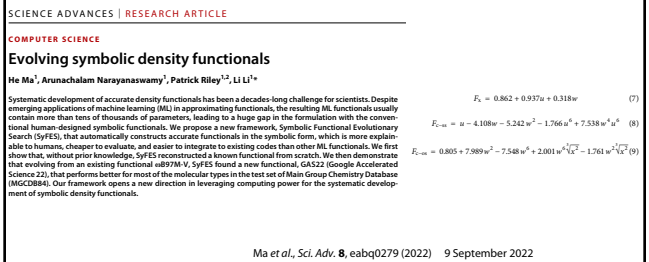
If scientists understood exactly how electrons act in molecules, they'd be able to predict the behavior of everything from experimental drugs to high-temperature superconductors. Following decades of physics-based insights, artificial intelligence systems are taking the next leap.



Machine learning and density functional theory R. Pederson, B. Kalita, and K. Burke, *Nat. Rev. Phys.* (2022). DOI: 10.1038/s42254-022-00470-2

31

Last work from Li Li



SCIENCE ADVANCES | RESEARCH ARTICLE

COMPUTER SCIENCE

Evolving symbolic density functionals

He Ma[1], Arunachalam Narayanaswamy[1], Patrick Riley[1,2], Li Li[1]*

Systematic development of accurate density functionals has been a decades-long challenge for scientists. Despite emerging applications of machine learning (ML) in approximating functionals, the resulting ML functionals usually contain more than tens of thousands of parameters, leading to a huge gap in the formulation with the conventional human-designed symbolic functionals. We propose a new framework, Symbolic Functional Evolutionary Search (SyFES), that automatically constructs accurate functionals in the symbolic form, which is more explainable to humans, cheaper to evaluate, and easier to integrate to existing codes than other ML functionals. We first show that, without prior knowledge, SyFES reconstructed a known functional from scratch. We then demonstrate that evolving from an existing functional ωB97M-V, SyFES found a new functional, GAS22 (Google Accelerated Science 22), that performs better for most of the molecular types in the test set of Main Group Chemistry Database (MGCDB84). Our framework opens a new direction in leveraging computing power for the systematic development of symbolic density functionals.

$$F_x = 0.862 + 0.937u + 0.318w \quad (7)$$

$$F_{c-\omega} = u - 4.108w - 5.242\,w^2 - 1.766\,u^2 + 7.538\,w^3 u^2 \quad (8)$$

$$F_{c-\alpha} = 0.805 + 7.989\,w^2 - 7.548\,w^3 + 2.001\,w^4\sqrt{\chi_\sigma^2} - 1.761\,w^2\sqrt{\chi_\sigma^2} \quad (9)$$

Ma *et al.*, *Sci. Adv.* **8**, eabq0279 (2022) 9 September 2022

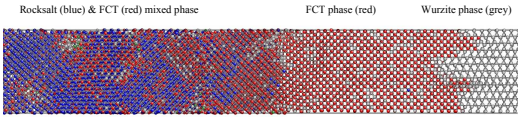
32

D Some closing thoughts

33

LAMMPS simulation

- CdSe being shocked using LAMMPS



- Recently, saw movie from Aidan Thompson of C being shocked with 18 billion atoms in ML potential

34

BRIEF COMMUNICATION
https://doi.org/10.1038/s41587-019-0224-x
nature biotechnology

Deep learning enables rapid identification of potent DDR1 kinase inhibitors

Alex Zhavoronkov[1*], Yan A. Ivanenkov[1], Alex Aliper[1], Mark S. Veselov[1], Vladimir A. Aladinskiy[1], Anastasiya V. Aladinskaya[1], Victor A. Terentiev[1], Daniil A. Polykovskiy[1], Maksim D. Kuznetsov[1], Arip Asadulaev[1], Yury Volkov[1], Artem Zholus[1], Rim R. Shayakhmetov[1], Alexander Zhebrak[1], Lidiya I. Minaeva[1], Bogdan A. Zagribelnyy[1], Lennart H. Lee[2], Richard Soll[2], David Madge[2], Li Xing[2], Tao Guo[2] and Alán Aspuru-Guzik[3,4,5,6]

We have developed a deep generative model, generative tensorial reinforcement learning (GENTRL), for de novo small-molecule design. GENTRL optimizes synthetic feasibility, novelty, and biological activity. We used GENTRL to discover potent inhibitors of discoidin domain receptor 1 (DDR1), a kinase target implicated in fibrosis and other diseases, in 21 days. Four compounds were active in biochemical assays, and two were validated in cell-based assays. One lead candidate was tested and demonstrated favorable pharmacokinetics in mice.

experimentally tested in 46 days, which demonstrates the potential of this approach to provide rapid and effective molecular design (Fig. 1a).

To create GENTRL, we combined reinforcement learning, variational inference, and tensor decompositions into a generative two-step machine learning algorithm (Supplementary Fig. 1)[19]. First, we learned a mapping of chemical space, a set of discrete molecular graphs, to a continuous space of 50 dimensions. We parameterized the structure of the learned manifold in the tensor train format to use partially known properties. Our auto-encoder-based model compresses

35

A feedback loop?

- Serious ML funding is 100 x bigger than electronic structure funding
- ML is eager for new applications in new domains
- Electronic structure is eager for resources
- Do a little ML in DFT, get 10 x usual money
- With 10 x usual money, triple the size of your elec struc group
- With triple ML output, ask for more ML money

36

Paola Gori-Giorgi

- Highly respected cond matter and quantum chemistry theorist
- DFT developer
- Postdoc with Perdew and Andeas Savin
- Full Prof at VU Amsterdam
- Quit for Microsoft AI for Science in Nov 2022
- Says working conditions much better there
- Has hired several former group members

37

Summary

- Kohn-Sham regularizer, using both energy and density losses and full differential programming, is very efficient way to learn chemical accuracy for strong correlation with minimal data.
- Also works to generate good functional for weakly-correlated systems.
- Challenge: Avoid using every point in the system as input.
- Thanks to NSF and DOE for funding.

38