Quantum Machine Learning

Anatole von Lilienfeld
Institute of Physical Chemistry and National Center for Computational Design and Discovery of Novel Materials (MARVEL), Department of Chemistry, University of Basel, Switzerland

Many of the most relevant chemical properties of matter depend explicitly on atomistic and electronic details, rendering a first principles approach to chemistry mandatory. Alas, even when using high-performance computers, brute force high-throughput screening of compounds is beyond any capacity for all but the simplest systems and properties due to the combinatorial nature of chemical space, i.e. all compositional, constitutional, and conformational isomers. Consequently, efficient exploration algorithms need to exploit all implicit redundancies present in chemical space. I will discuss recently developed statistical learning approaches for interpolating quantum mechanical observables in compositional and constitutional space. Results for our models indicate remarkable performance in terms of accuracy, speed, universality, and size scalability.

"Machine Learning, Quantum Mechanics, and Chemical Compound Space"
By Ramakrishnan and von Lilienfeld
Reviews in Computational Chemistry, edited by Abby L. Parrill and Kenny B. Lipkowitz
Volume **30**, Chapter 5, pages 225-256 (2017)
www.arxiv.org/abs/1510.07512

Essay: “Quantum Machine Learning in Chemical Compound Space”
Just accepted, *Angew. Chem. Int. Ed.* (2018)
http://rdcu.be/AJNU/

UNI BASEL

SOPHIAE PHILOSOPHORUM ET PHYSICORUM

Departement Chemie

U.S. DEPARTMENT OF ENERGY
INCITE
LEADERSHIP COMPUTING

FNSNF
FONDS NATIONAL SUISSE
SCHWEIZERISCHER NATIONALFONDS
FONDO NAZIONALE SVIZZERO
SWISS NATIONAL SCIENCE FOUNDATION

erc

Institute for Pure & Applied Mathematics
IPAM
University of California, Los Angeles

Research at Google

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
UNITED STATES AIR FORCE

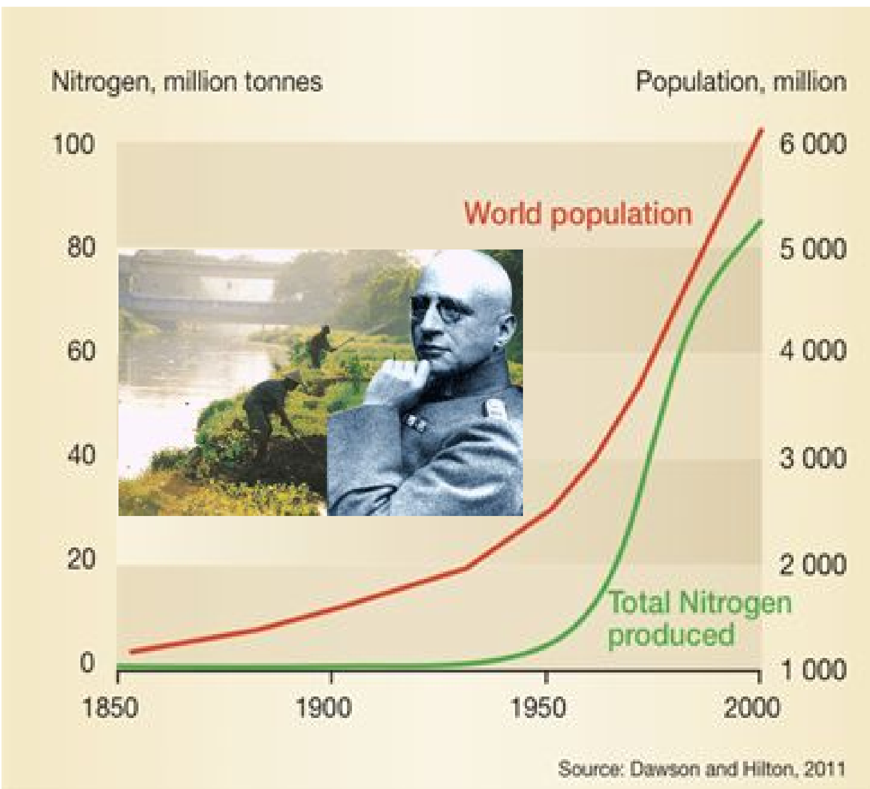
MARVEL
NATIONAL CENTRE OF COMPETENCE IN RESEARCH

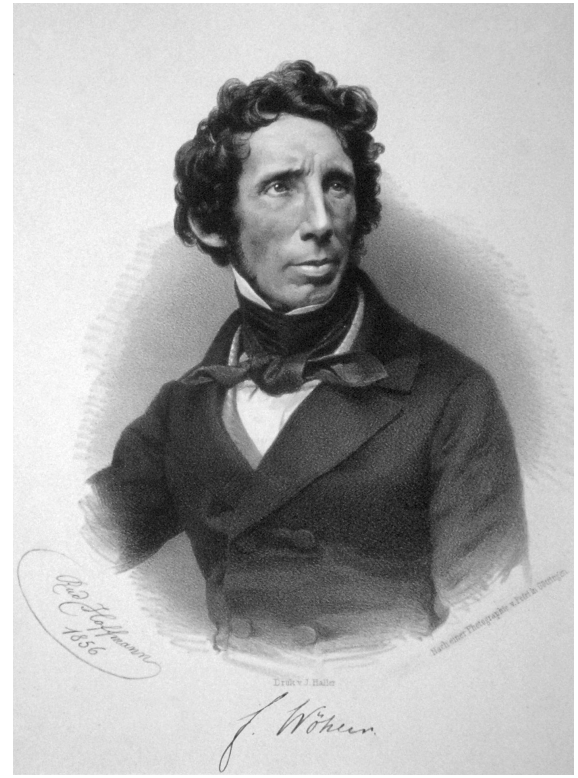
Chemistry?

Millennium Essay by V. Smil

Nature **400**, 415 (29 July 1999) | doi:10.1038/22672

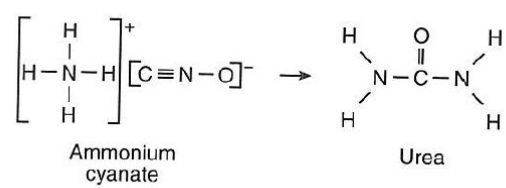
``What is the most important invention of the twentieth century? Aeroplanes, nuclear energy, space flight, television and computers will be the most common answers … Yet none of these can match the synthesis of ammonia from its elements. ''



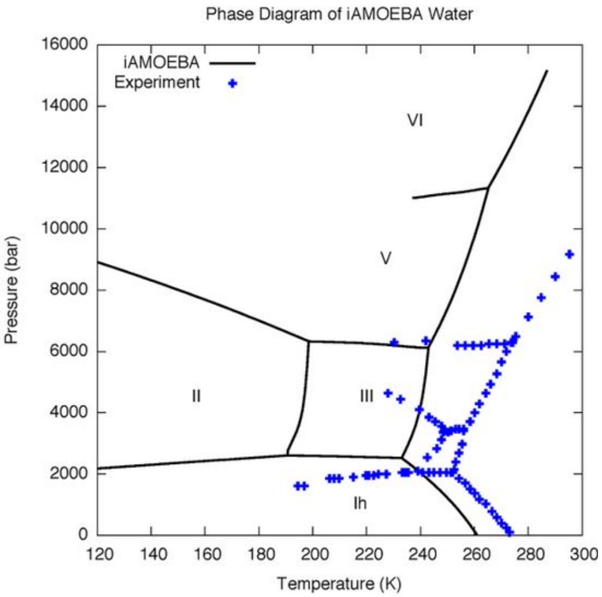


Friedrich Wöhler (1800-1882)

1828: Urea from ammonia and cyanate (in solution+heat)



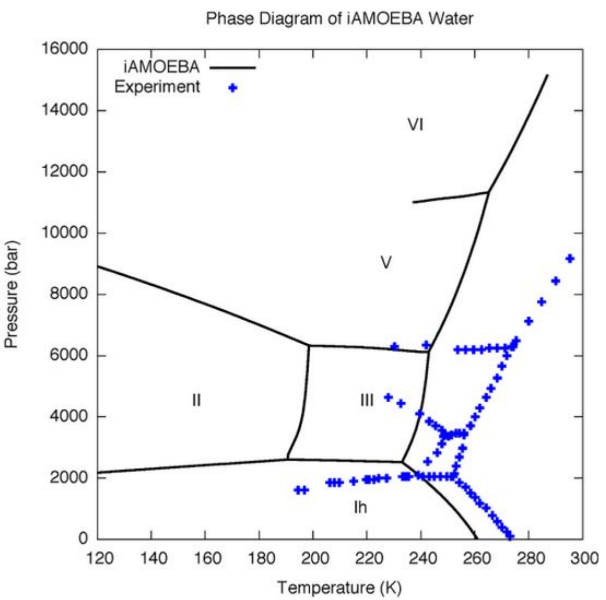
Chemistry on a computer? Predict outcomes!



Pande et al*, J. Phys. Chem B* (2013)

Structure

Chemistry on a computer? Predict outcomes!

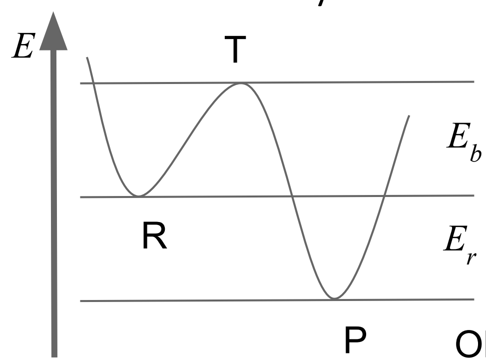


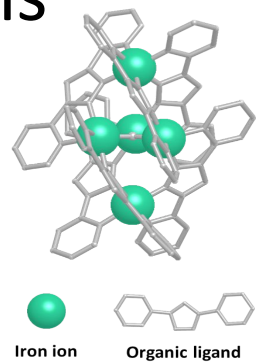
Pande et al, *J. Phys. Chem B* (2013)

Structure

Reactions

Fe based catalyst for water oxidation

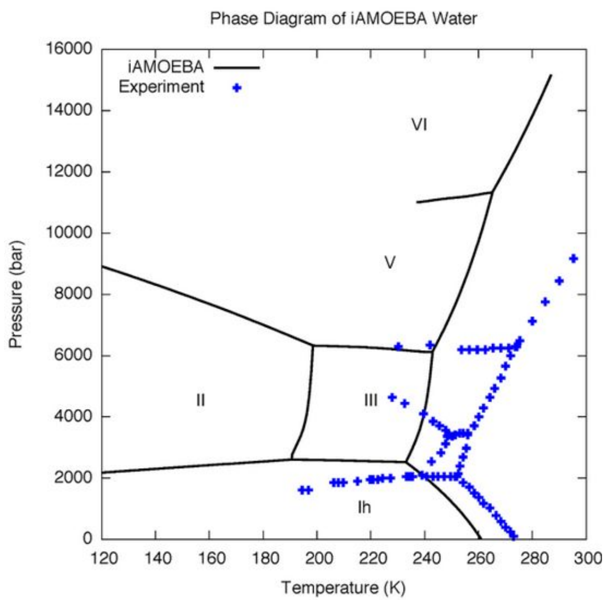




Okamura et al, *Nature* (2016)

Chemistry on a computer? Predict outcomes!

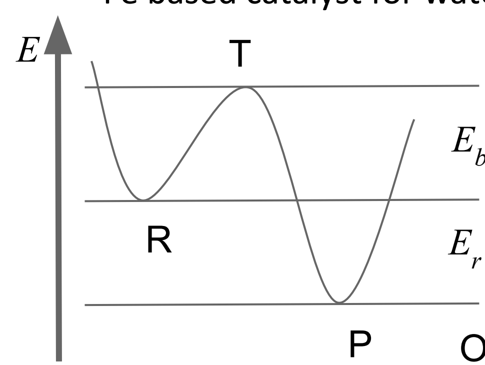
Structure

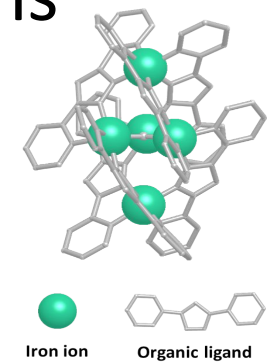


Pande et al, *J. Phys. Chem B* (2013)

Reactions

Fe based catalyst for water oxidation

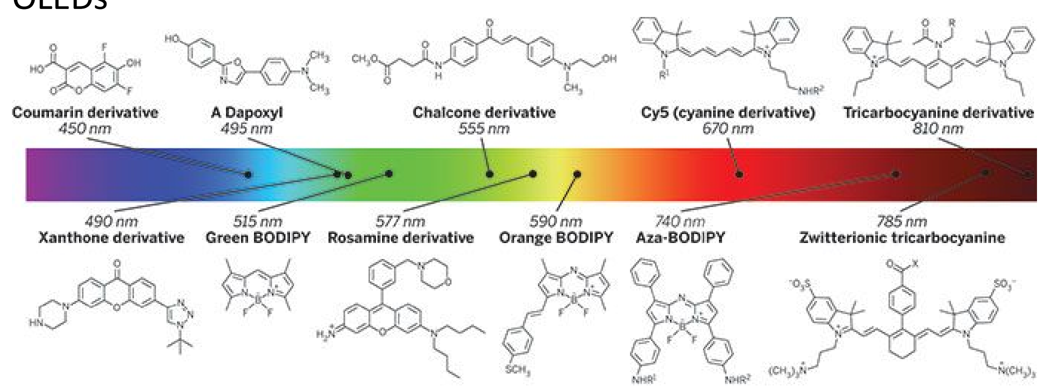




Okamura et al, *Nature* (2016)

Properties

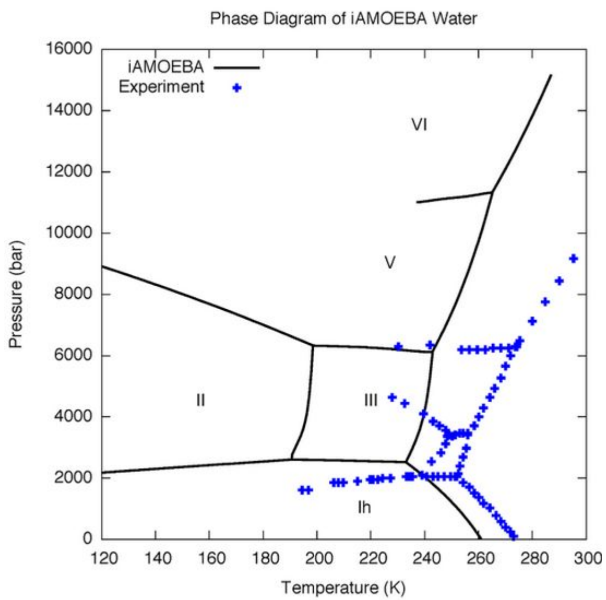
OLEDs



Chemistry on a computer? Predict outcomes!

We suck!

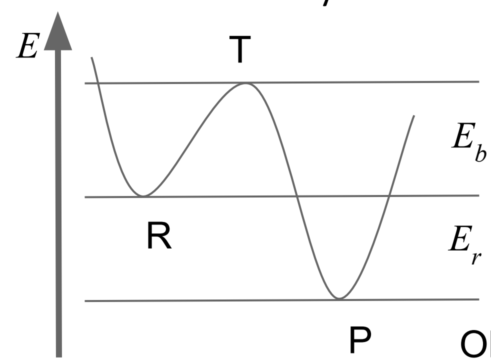
Structure

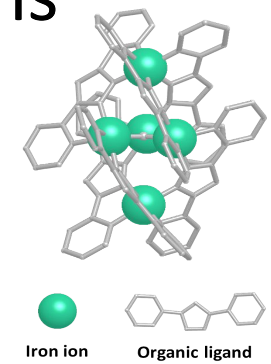


Pande et al, *J. Phys. Chem B* (2013)

Reactions

Fe based catalyst for water oxidation

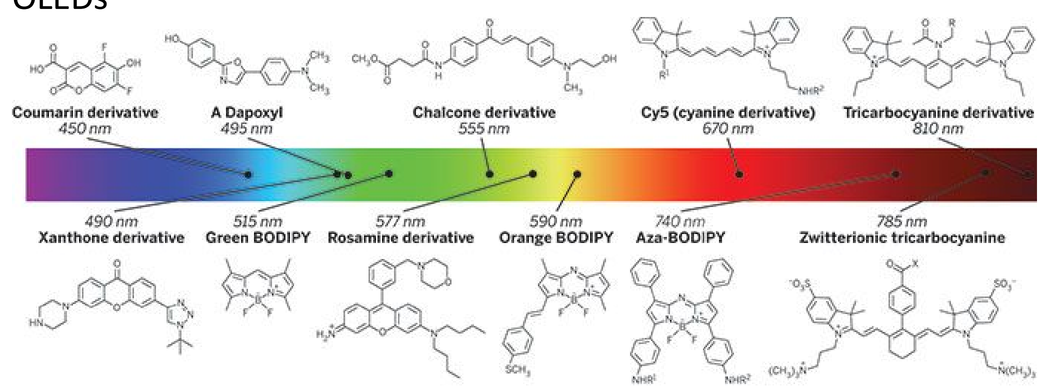




Okamura et al, *Nature* (2016)

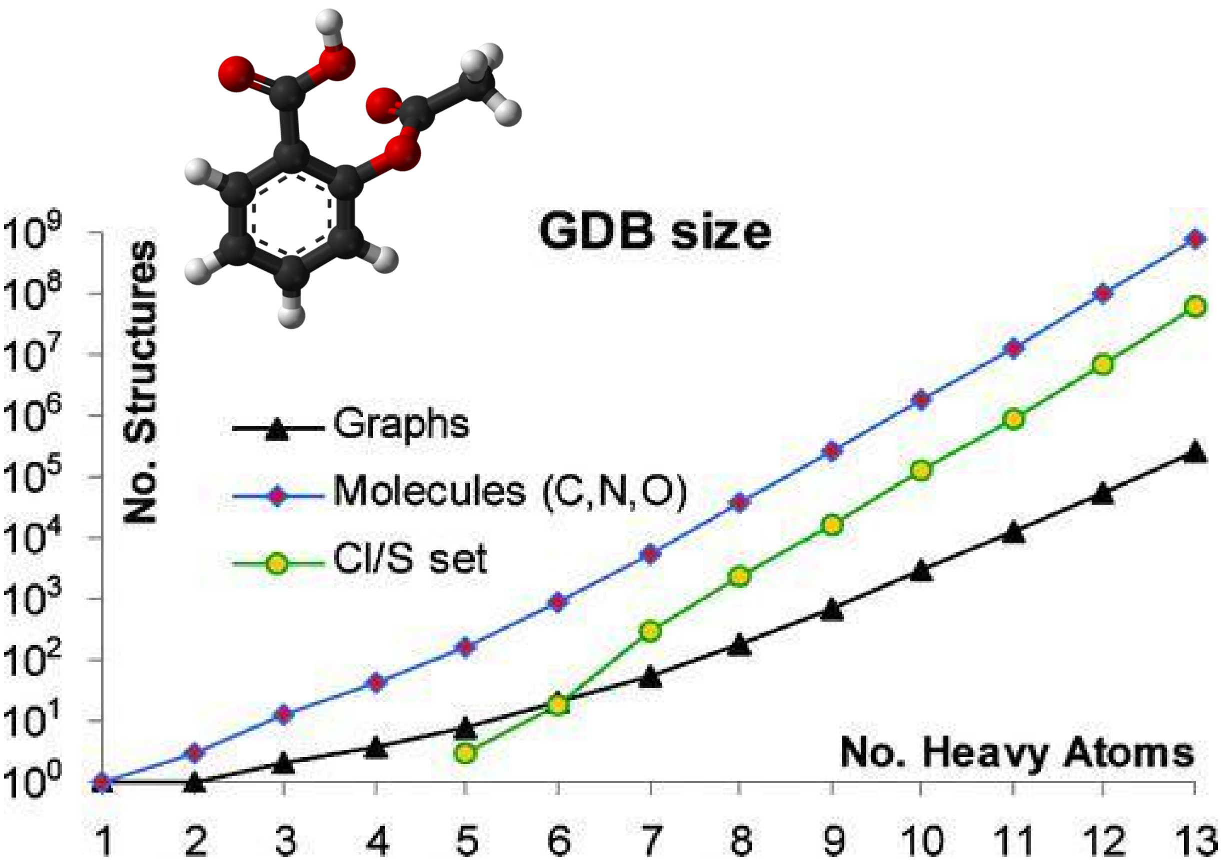
Properties

OLEDs

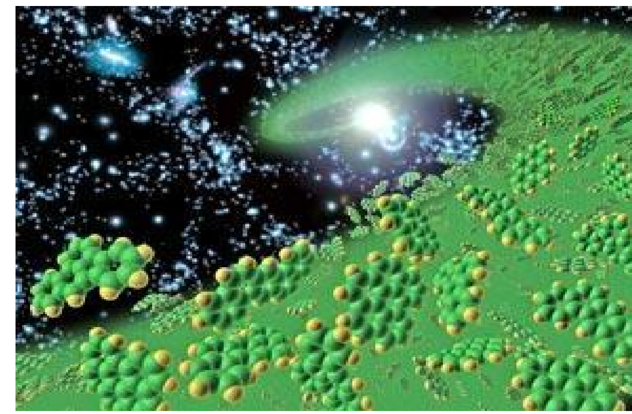


``The greatest shortcoming of the human race is our inability to understand the exponential function''

Al Bartlett, U of Colorado Boulder



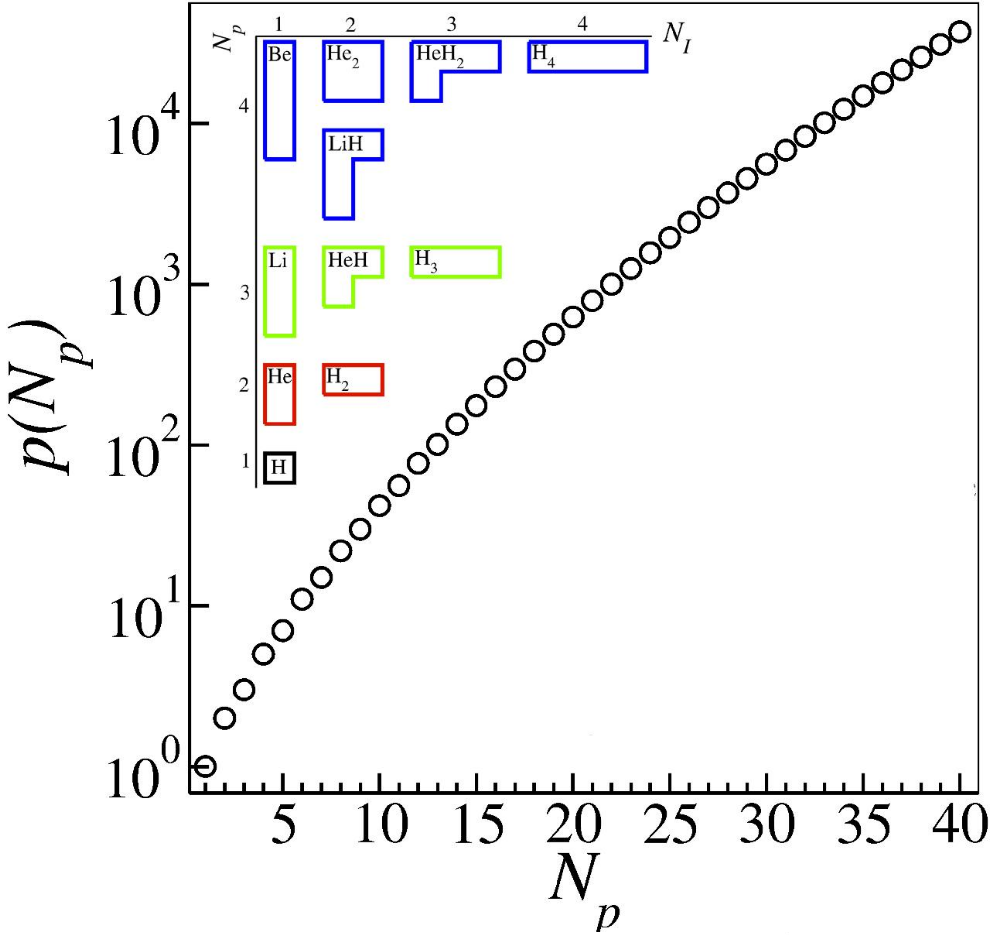
J.-L. Reymond and coworkers, *J Am Chem Soc* (2009) and ff

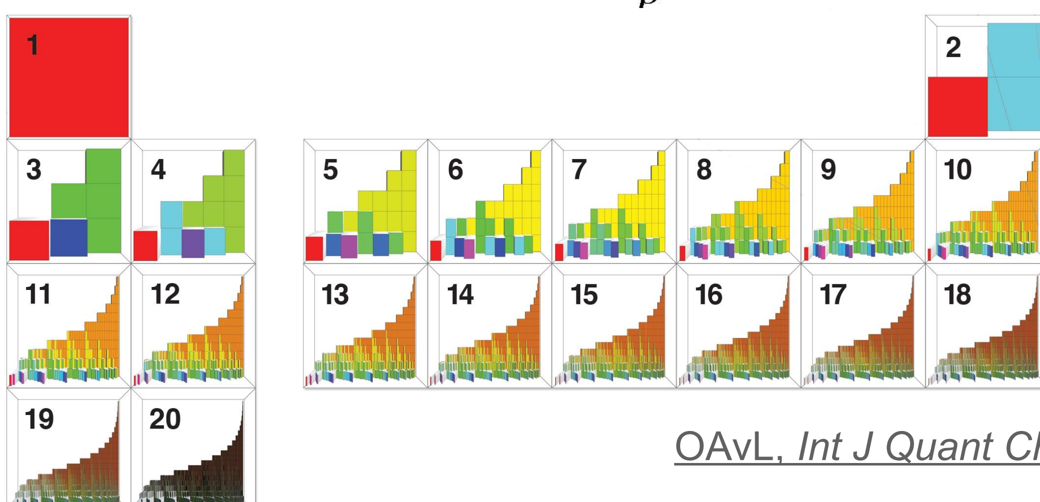


© by NASA

Composition

Young-Ferrers

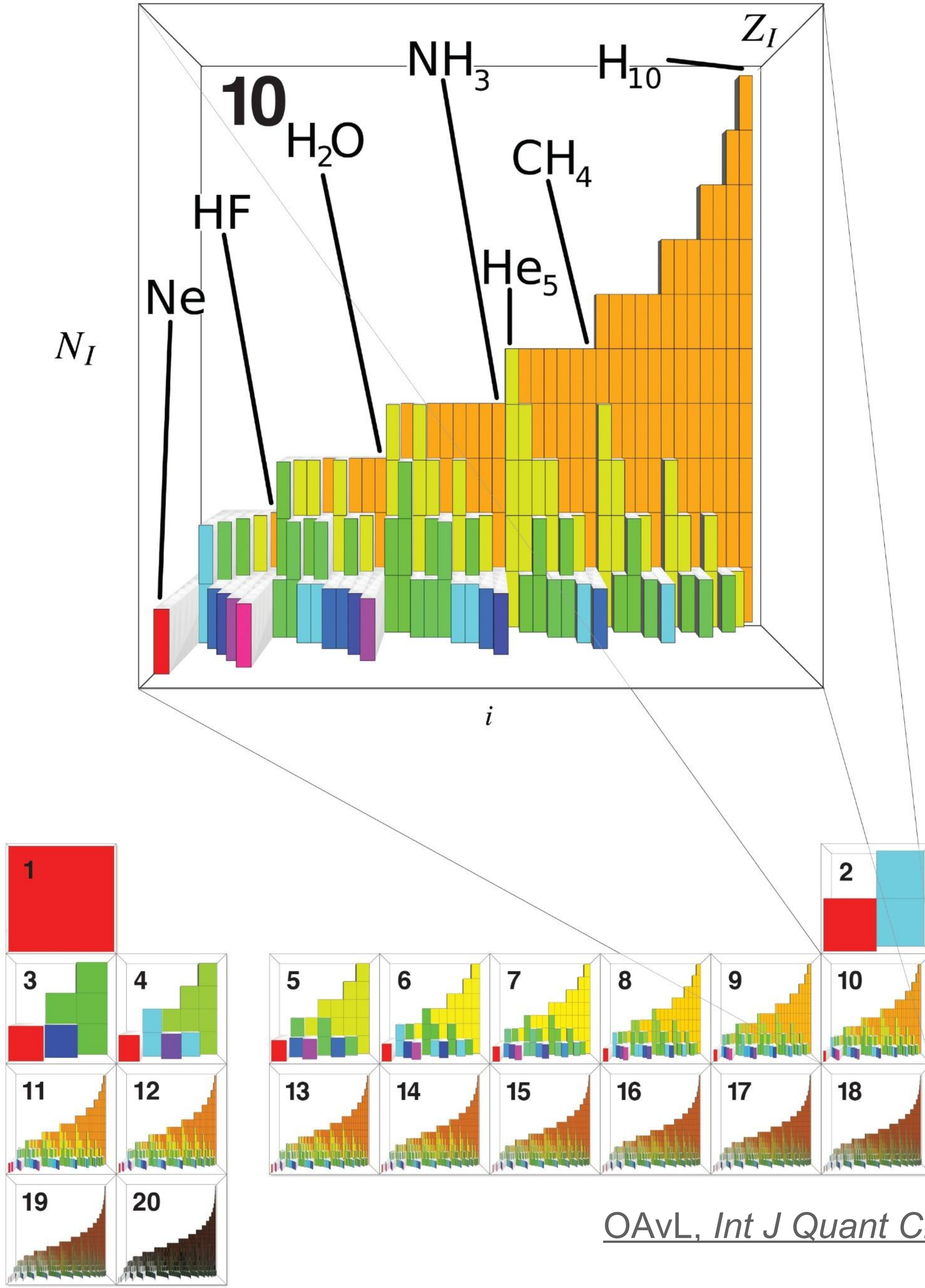




OAvL, *Int J Quant Chem*, (2013)

Composition

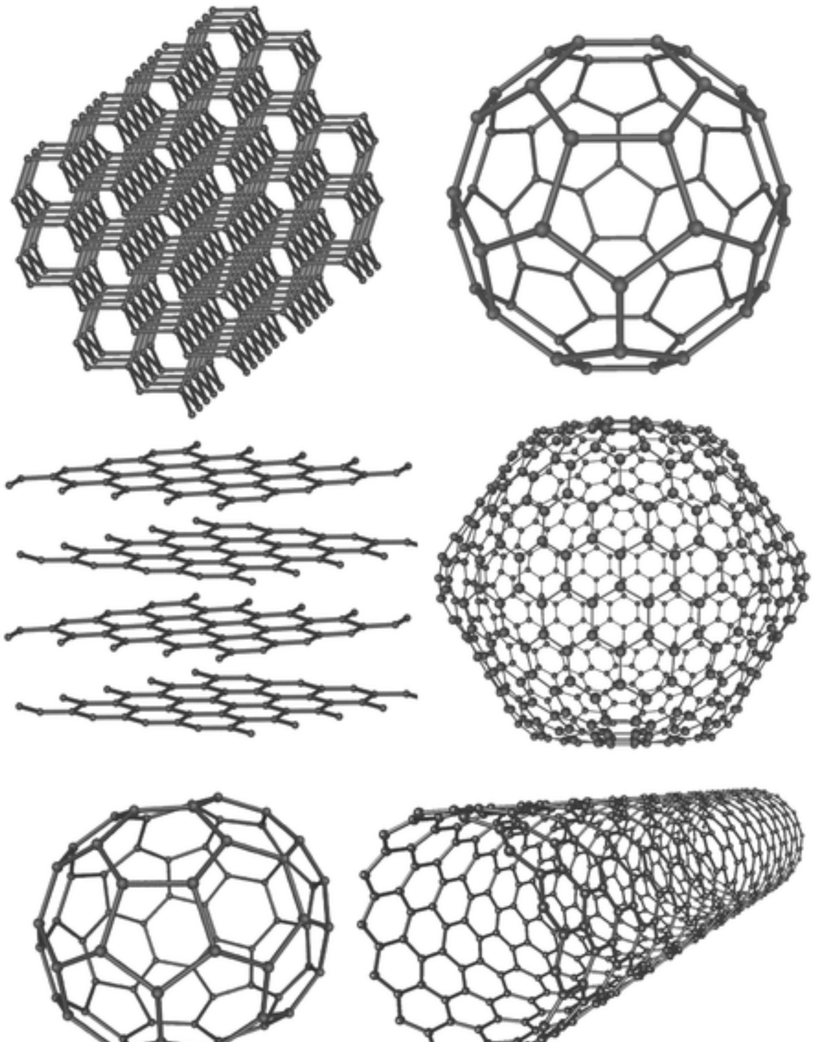
10 protons

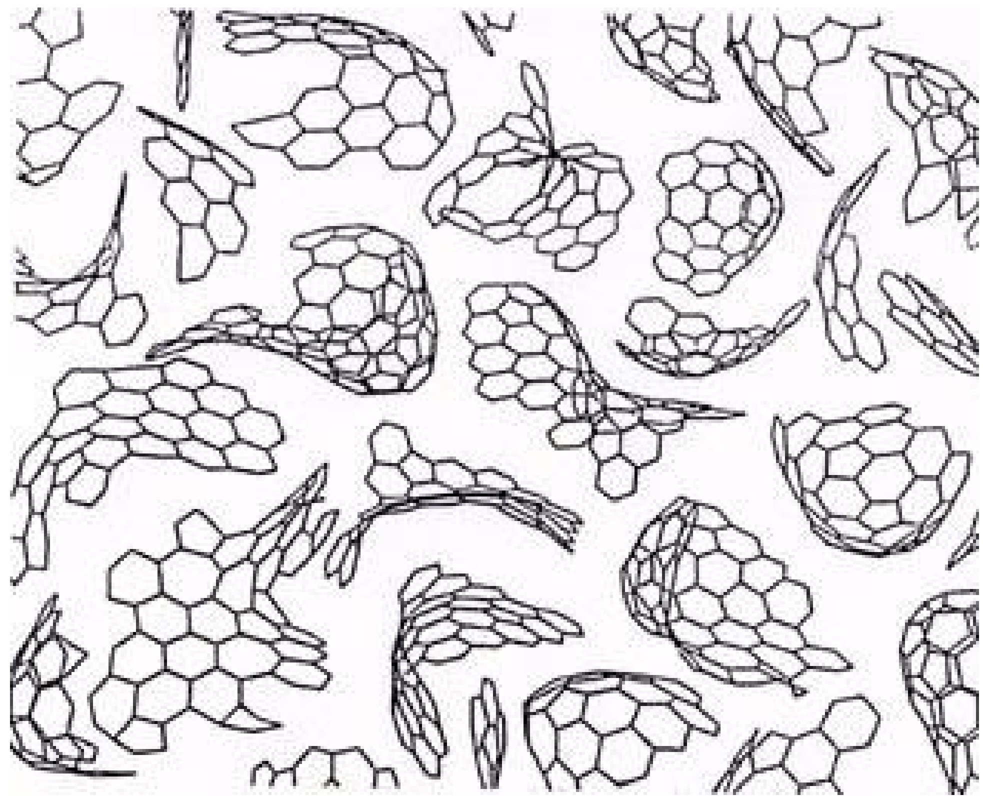


OAvL, *Int J Quant Chem*, (2013)

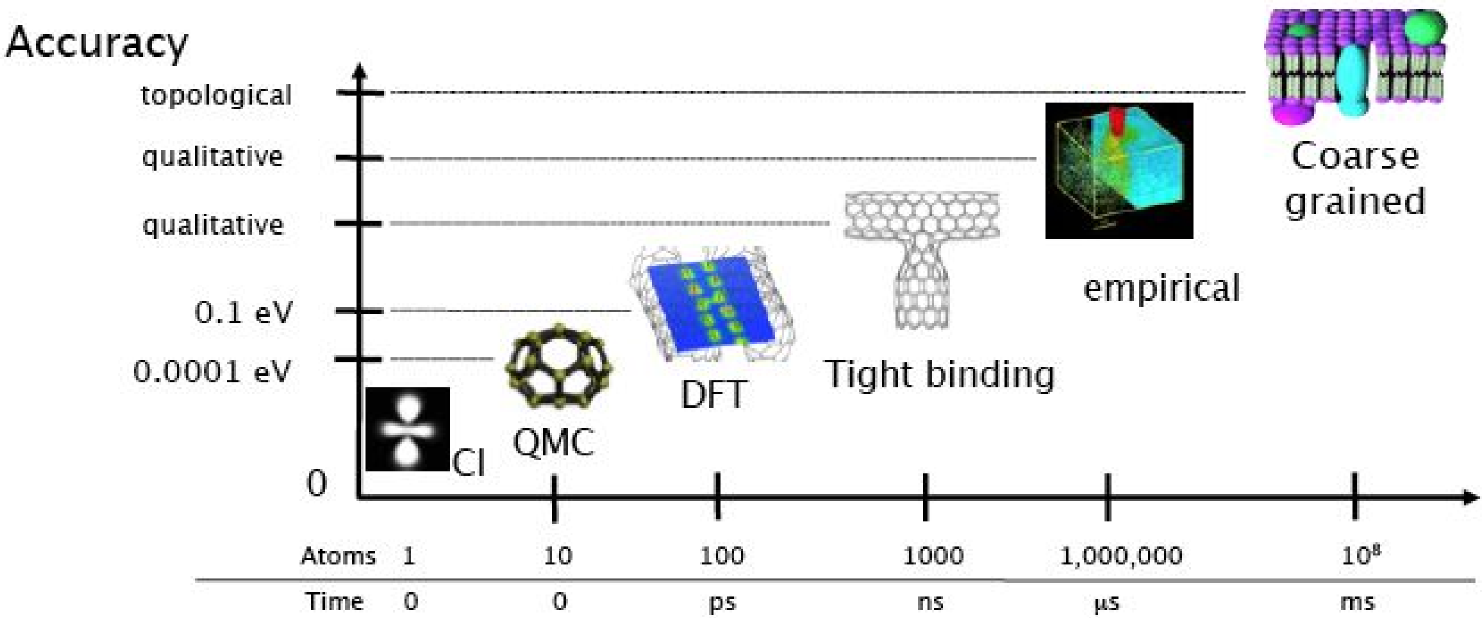
Spatial configuration

Carbon allotropes

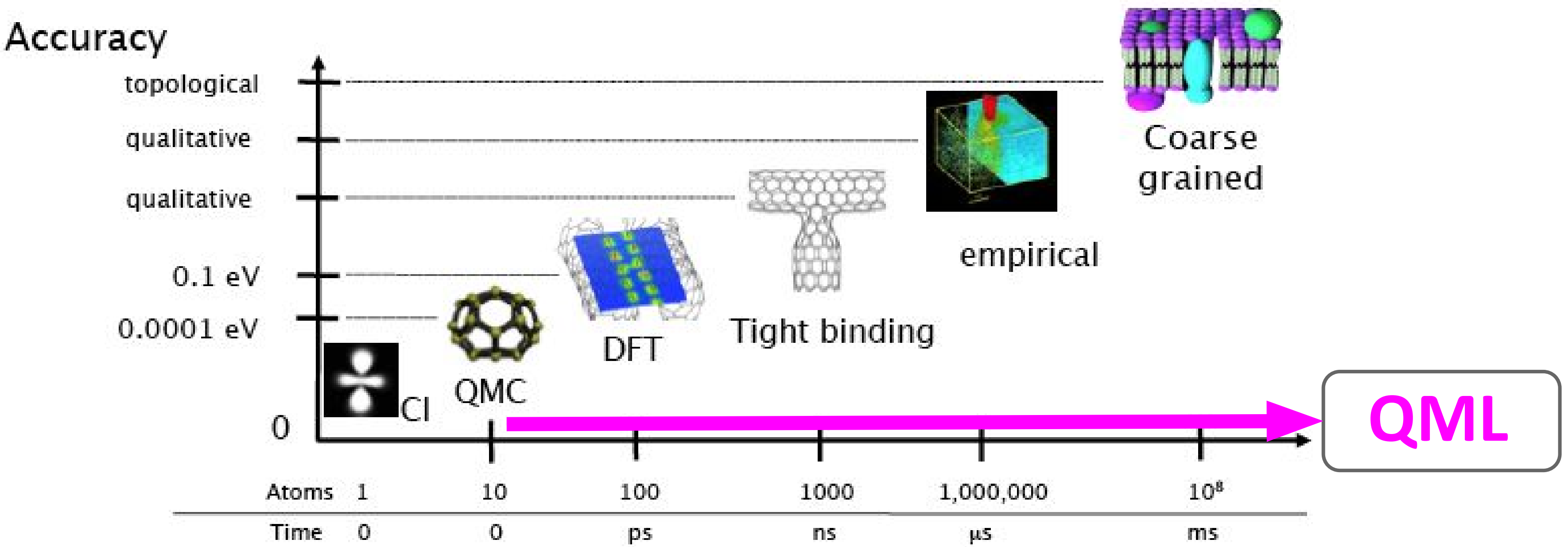




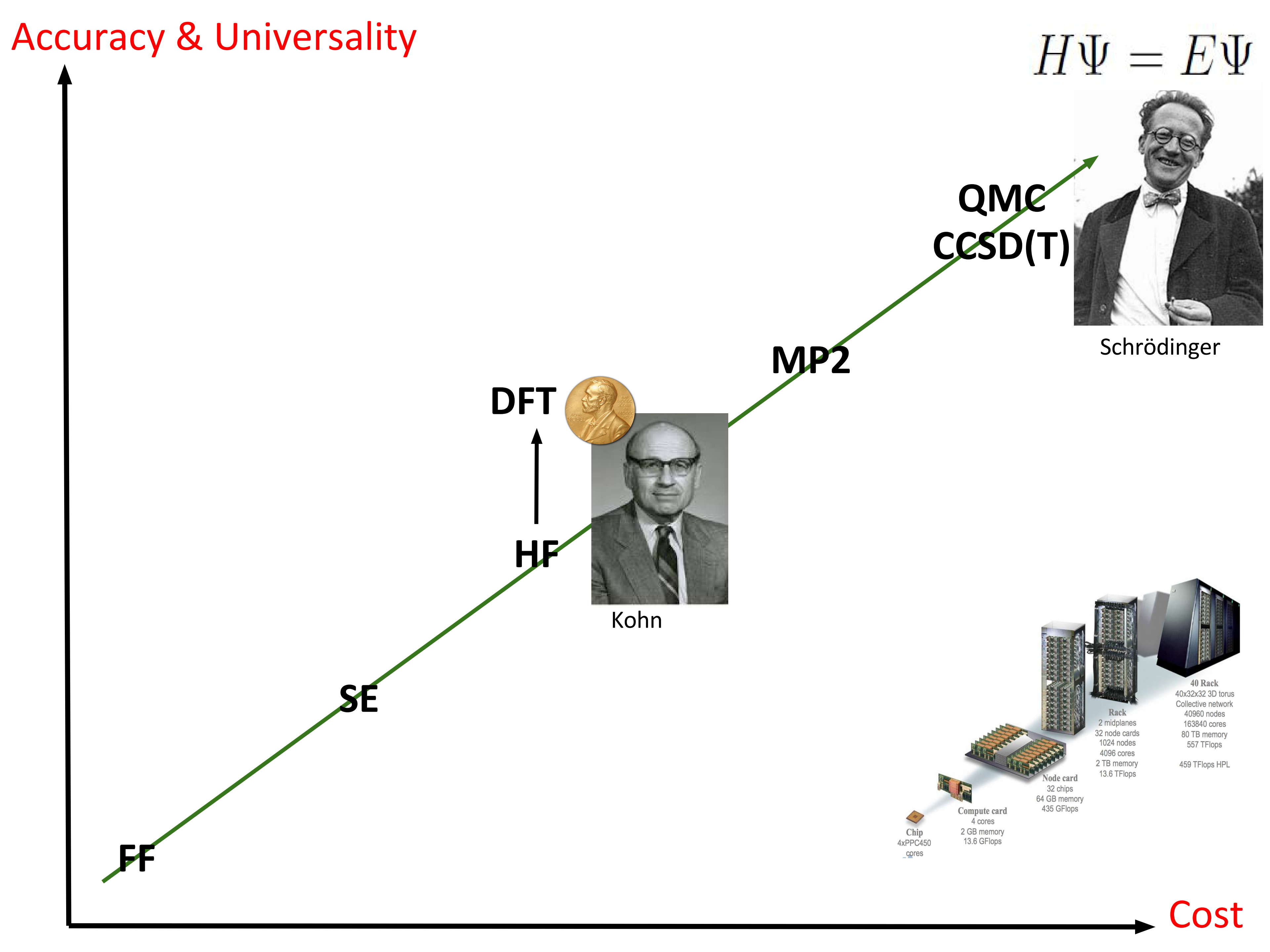
CCS >> 10^{60}
"Chemical Space", by Kirkpatrick and Ellis, *Nature* (2004)

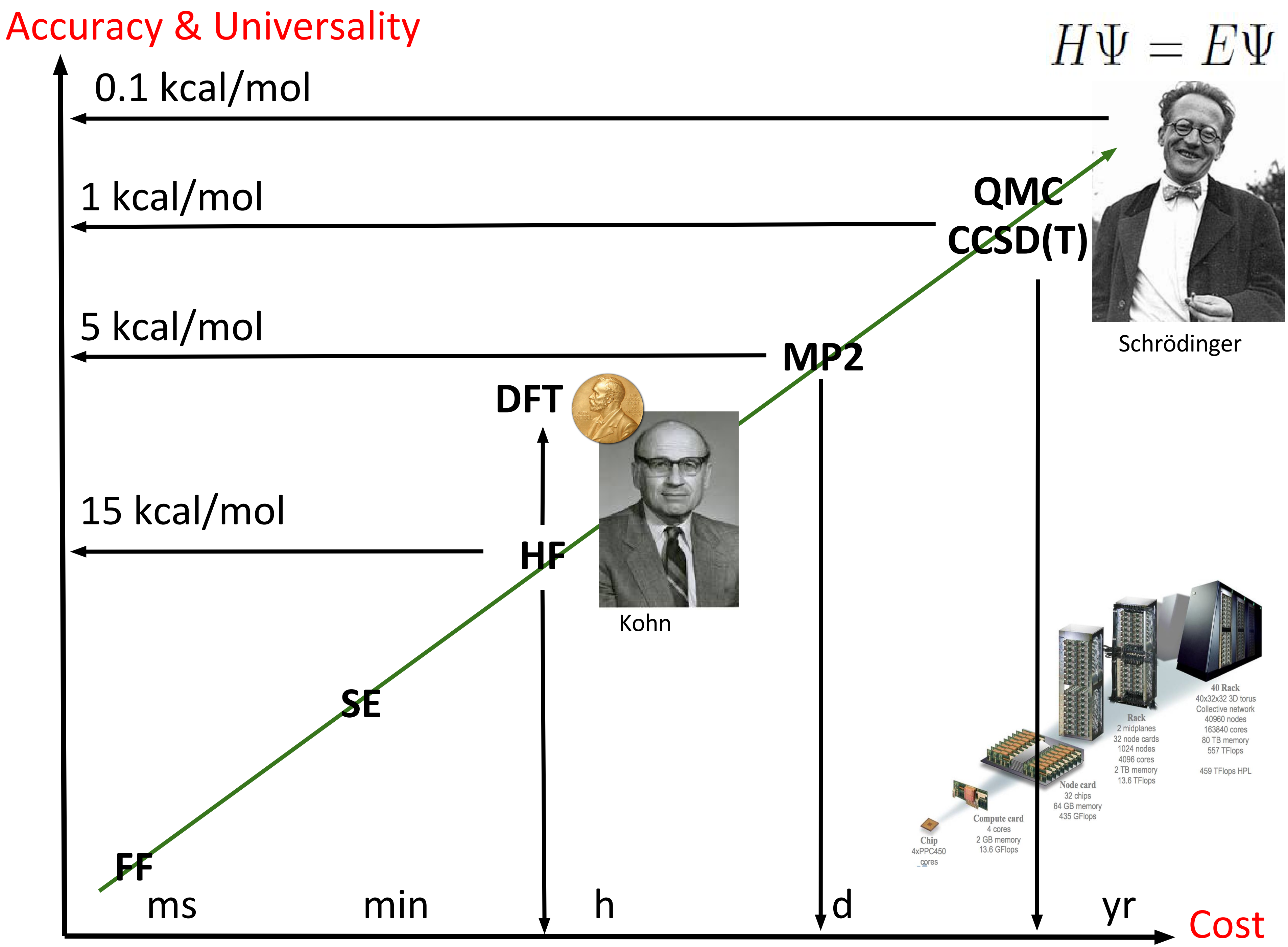


Picture from
Gabor Csanyi

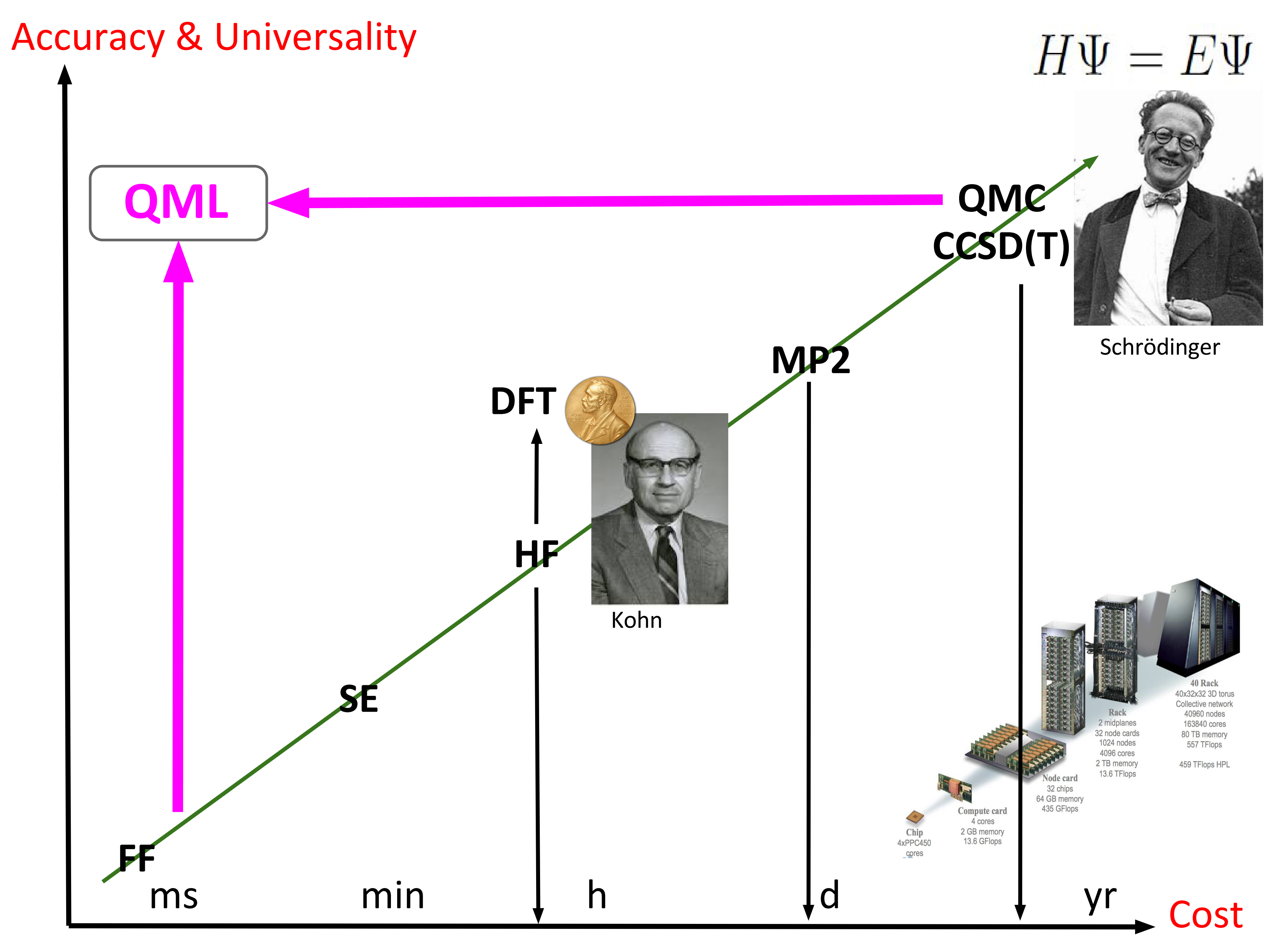


Picture from
Gabor Csanyi

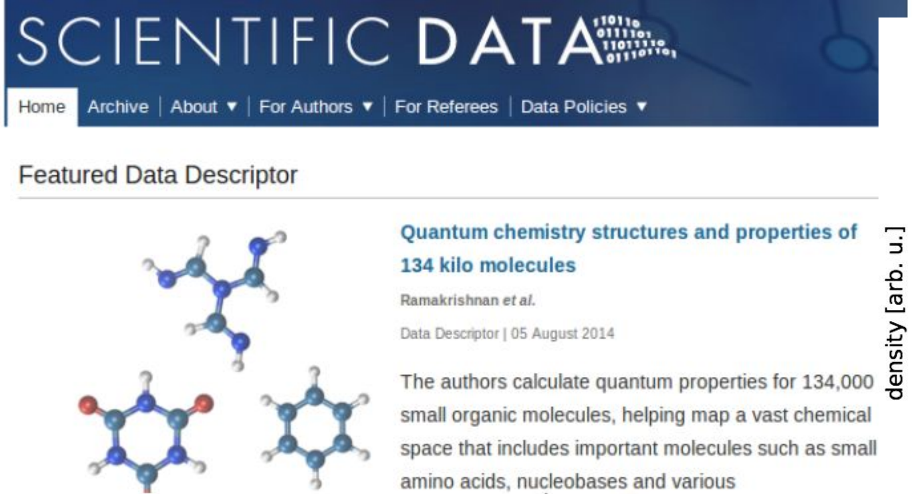

Accuracy & Universality
$H\Psi = E\Psi$
QMC
CCSD(T)
Schrödinger
MP2
DFT
HF
Kohn
SE
FF
Cost
40 Rack
40x32x32 3D torus
Collective network
40960 nodes
163840 cores
80 TB memory
557 TFlops
459 TFlops HPL
Rack
2 midplanes
32 node cards
1024 nodes
4096 cores
2 TB memory
13.6 TFlops
Node card
32 chips
64 GB memory
435 GFlops
Compute card
4 cores
2 GB memory
13.6 GFlops
Chip
4xPPC450
cores

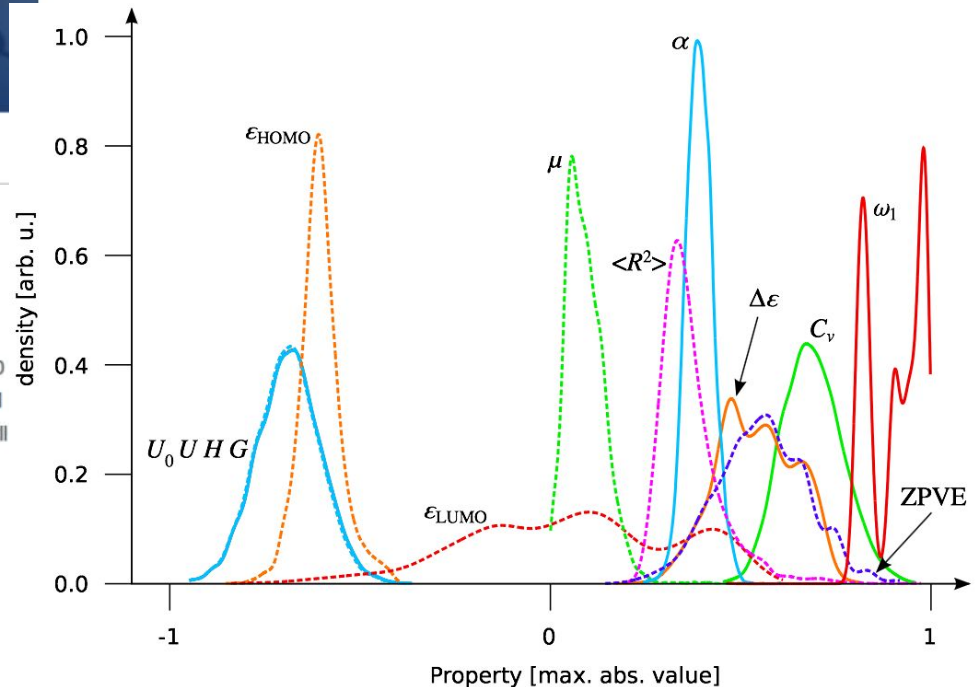


Accuracy & Universality
$H\Psi = E\Psi$
0.1 kcal/mol
1 kcal/mol
5 kcal/mol
15 kcal/mol
QMC
CCSD(T)
Schrödinger
MP2
DFT
Kohn
HF
SE
FF
ms
min
h
d
yr
Cost
Chip
4xPPC450 cores
Compute card
4 cores
2 GB memory
13.6 GFlops
Node card
32 chips
64 GB memory
435 GFlops
Rack
2 midplanes
32 node cards
1024 nodes
4096 cores
2 TB memory
13.6 TFlops
40 Rack
40x32x32 3D torus
Collective network
40960 nodes
163840 cores
80 TB memory
557 TFlops
459 TFlops HPL

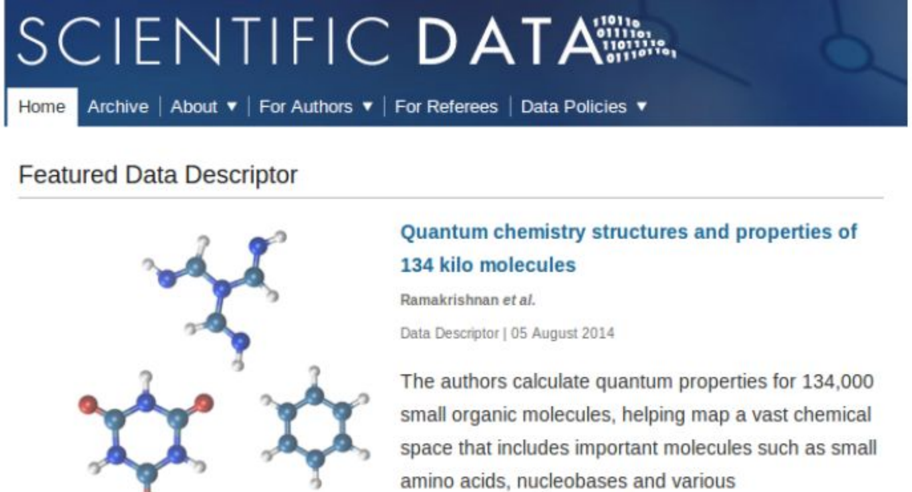


Accuracy & Universality
$H\Psi = E\Psi$
QML
QMC
CCSD(T)
Schrödinger
MP2
DFT
HF
Kohn
SE
FF
ms
min
h
d
yr
Cost
Chip
4xPPC450
cores
Compute card
4 cores
2 GB memory
13.6 GFlops
Node card
32 chips
64 GB memory
435 GFlops
Rack
2 midplanes
32 node cards
1024 nodes
4096 cores
2 TB memory
13.6 TFlops
40 Rack
40x32x32 3D torus
Collective network
40960 nodes
163840 cores
80 TB memory
557 TFlops
459 TFlops HPL

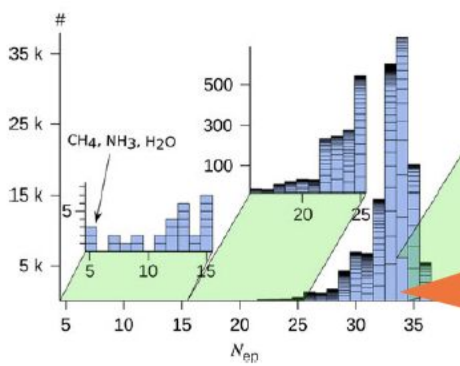


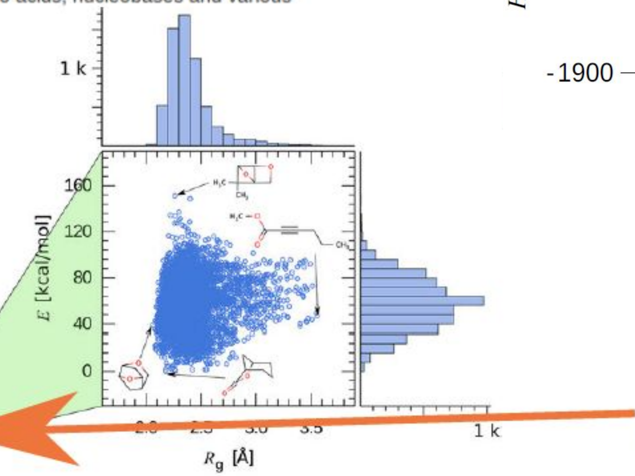


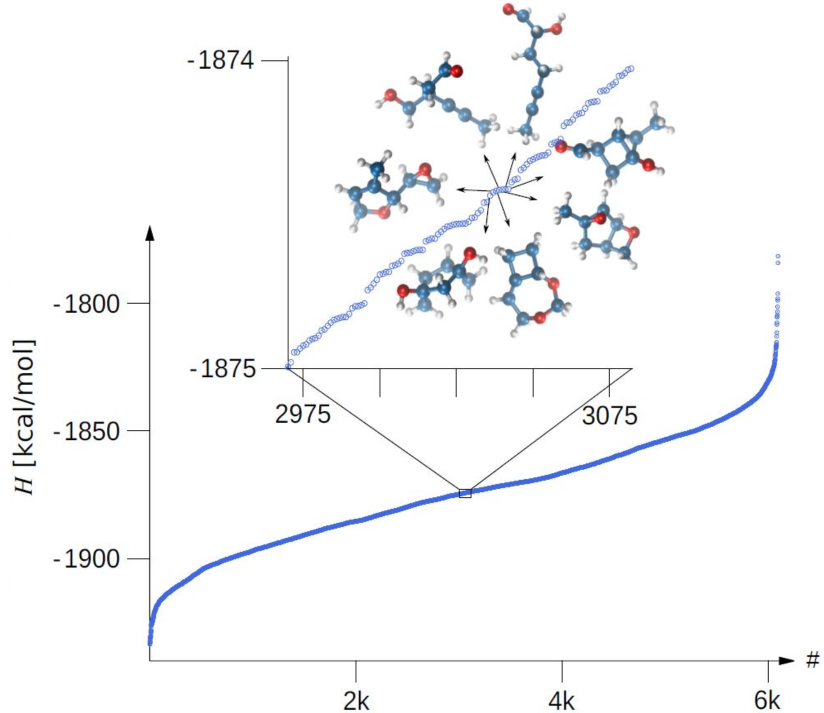
Ramakrishnan et al, *Scientific Data* (2014)

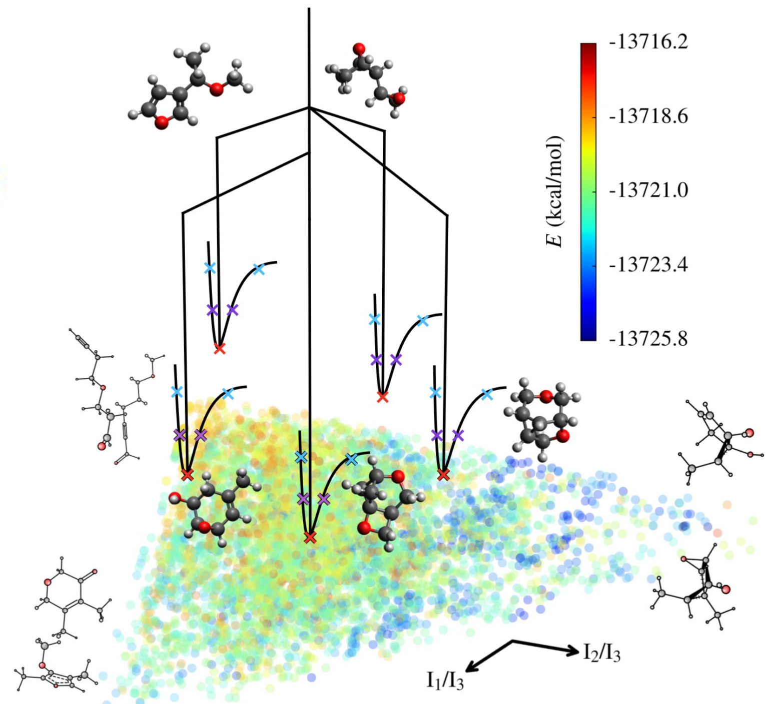


Organic molecules
with up to 9 CONF atoms





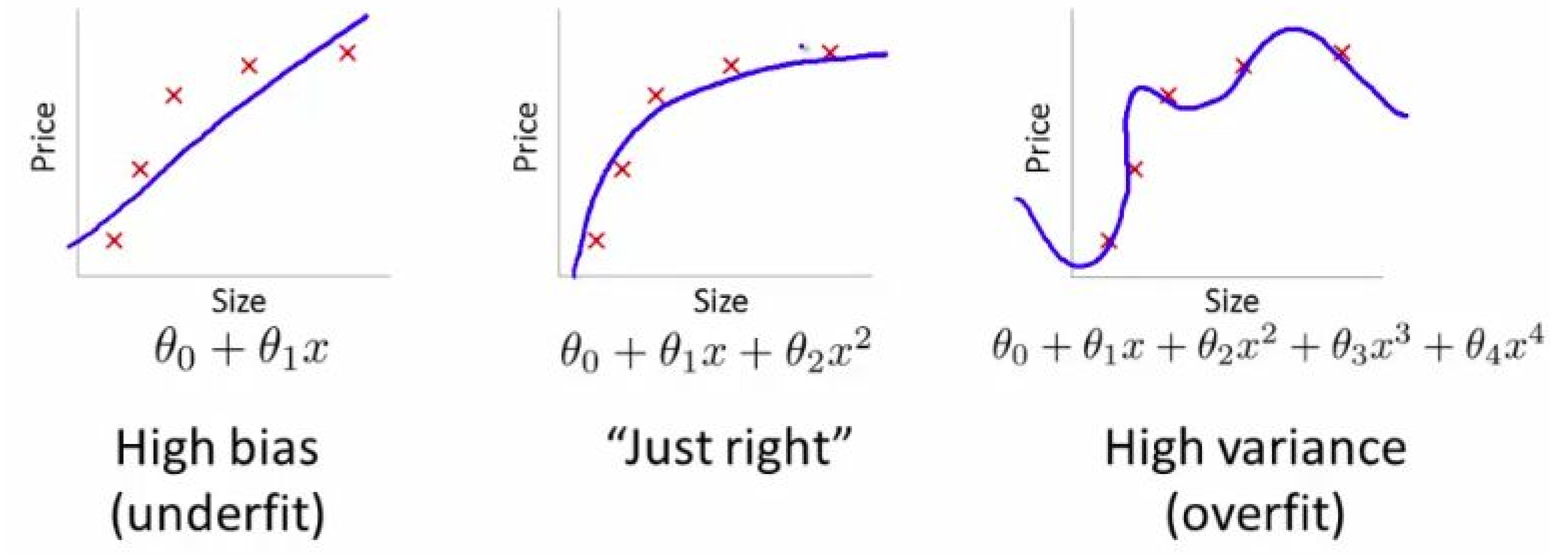


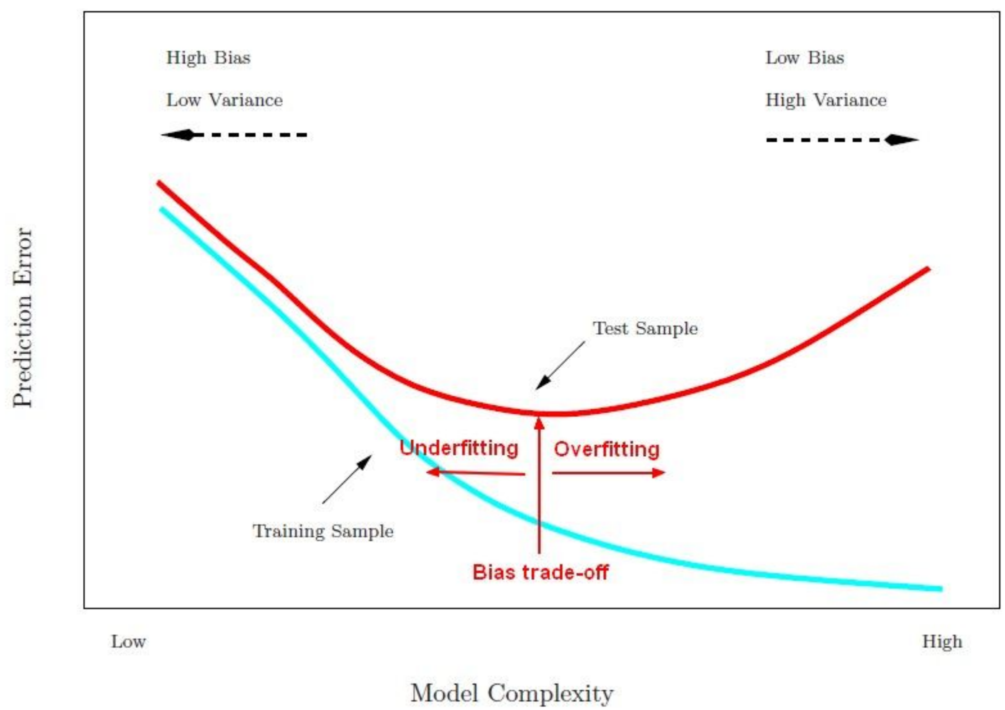


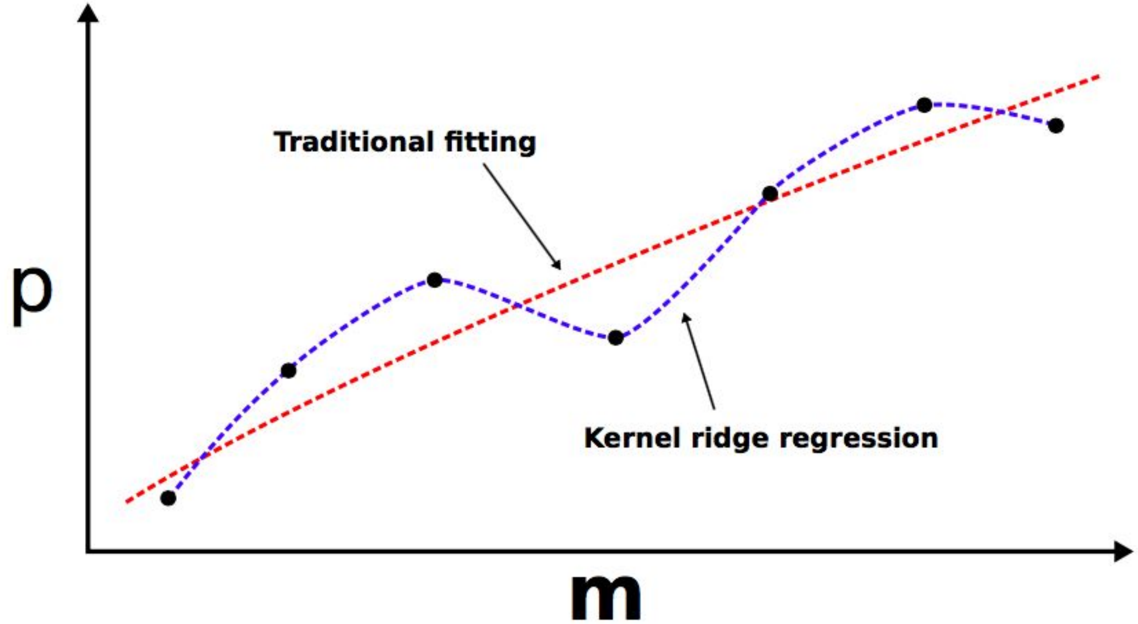
``Enumeration surpasses human imagination''
J.-L. Reymond

Ramakrishnan et al, *Scientific Data* (2014)

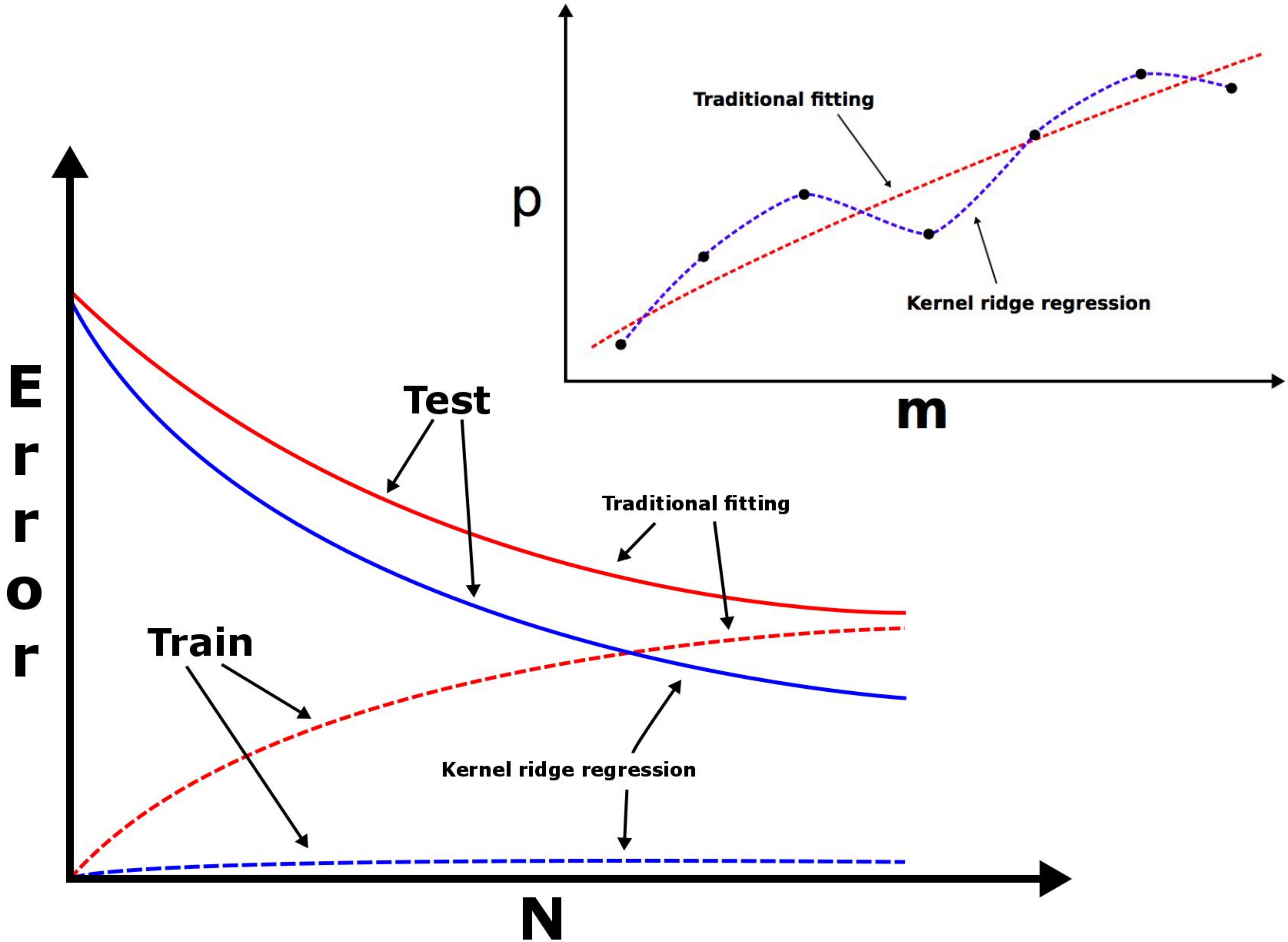
Overfitting?



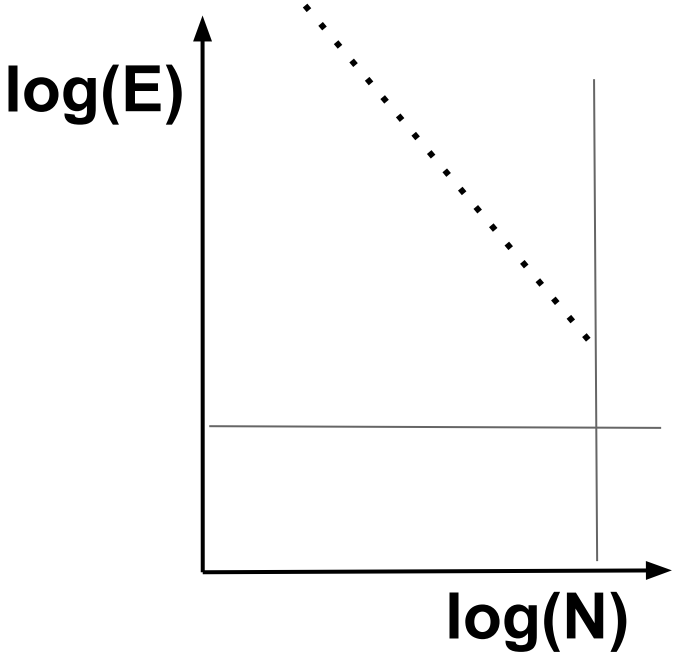




Vapnik, *The Nature of Statistical Learning Theory,* Springer (1995)



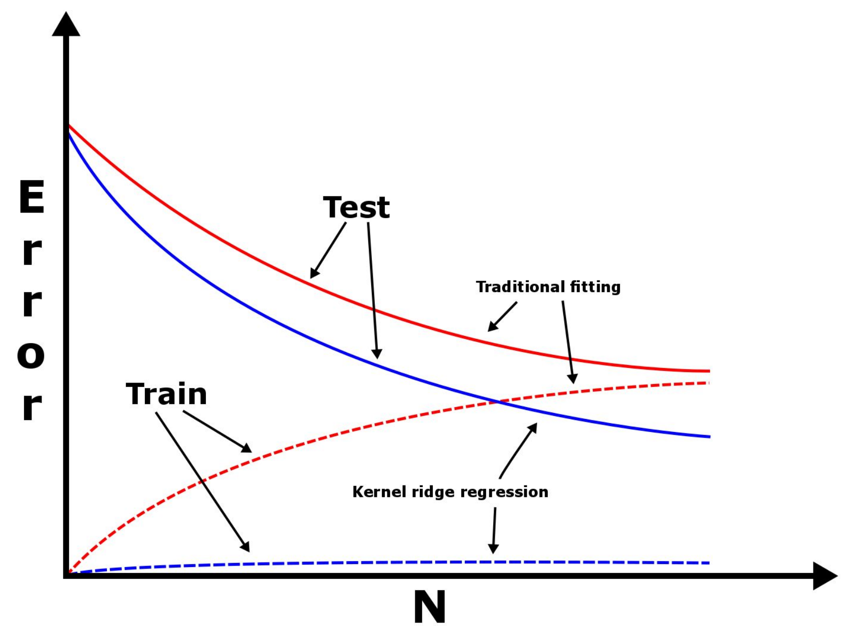
Vapnik, *The Nature of Statistical Learning Theory,* Springer (1995)



The bigger the data the better …

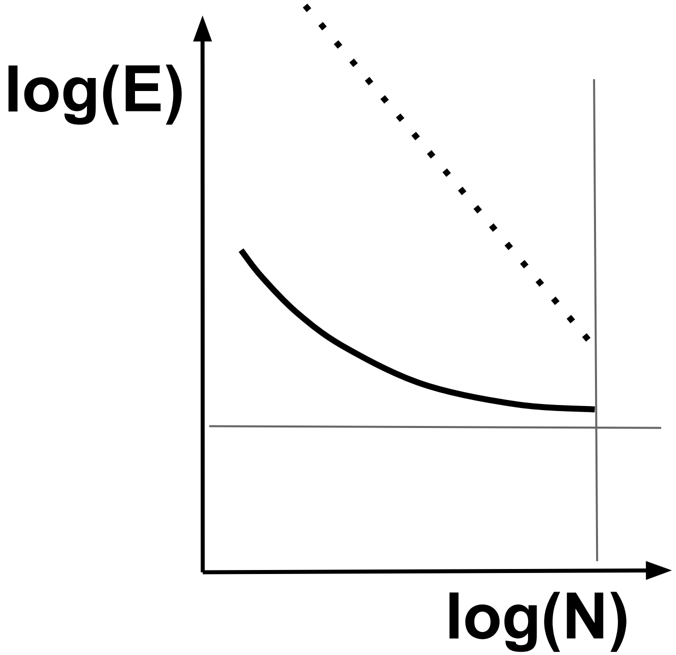
Vapnik, V., *The Nature of Statistical Learning Theory,* Springer (1995)

$$\boxed{\begin{gathered} P^{\text{est}}(\mathbf{M}) = \sum_i \alpha_i k(\mathbf{M}, \mathbf{M}_i) \\ \vec{\alpha} = \mathbf{K}^{-1} \vec{P}^{\text{ref}} \end{gathered}}$$



Error ~ a/N^b

$\rightarrow \log(\text{Error}) = \log(a) - b\log(N)$

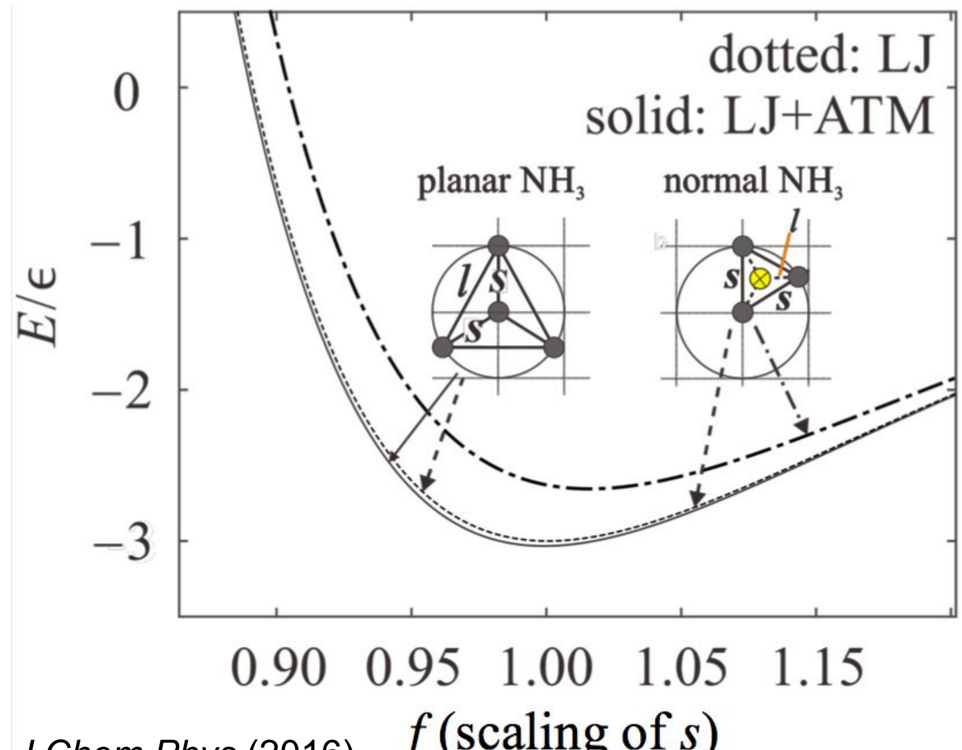


The bigger the data the better …

Vapnik, V., *The Nature of Statistical Learning Theory,* Springer (1995)

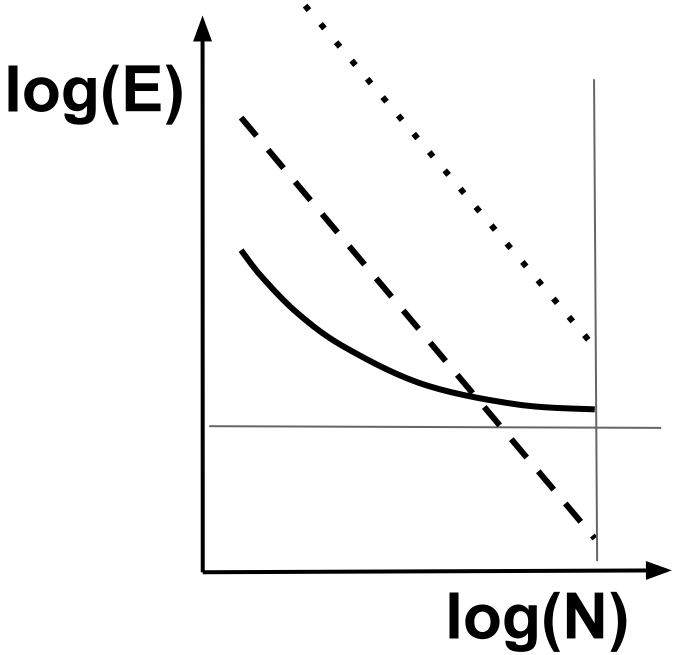
$$P^{\text{est}}(\mathbf{M}) = \sum_i \alpha_i k(\mathbf{M}, \mathbf{M}_i)$$
$$\vec{\alpha} = \mathbf{K}^{-1} \vec{P}^{\text{ref}}$$

Error ~ a/N^b



Huang, OAvL, *J Chem Phys* (2016)

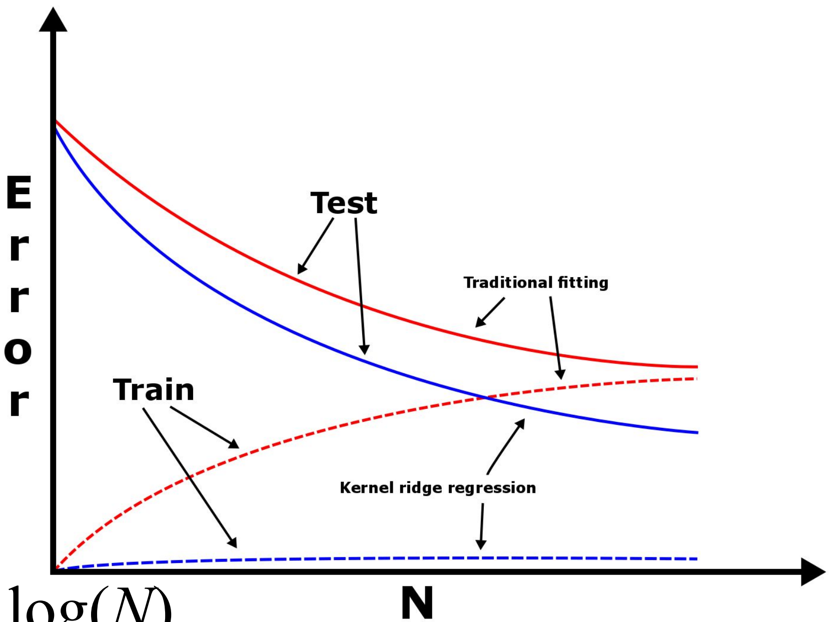
LJ: Lennard-Jones 2-body vdW potential
ATM: Axilrod-Teller-Muto 3-body vdW potential



The bigger the data the better …

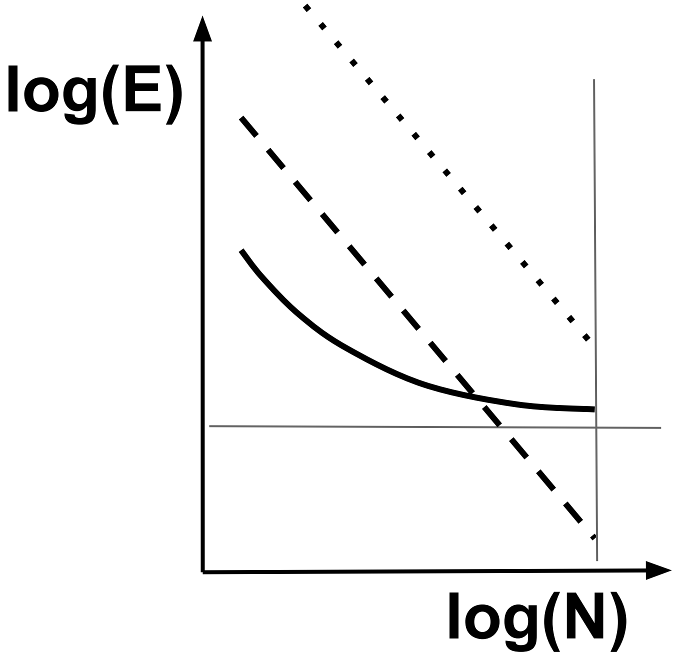
Vapnik, V., *The Nature of Statistical Learning Theory,* Springer (1995)

$$P^{\text{est}}(\mathbf{M}) = \sum_i \alpha_i k(\mathbf{M}, \mathbf{M}_i)$$
$$\vec{\alpha} = \mathbf{K}^{-1}\vec{P}^{\text{ref}}$$



Error ~ $a/(N')^b$, e.g. $N' = N\,c$

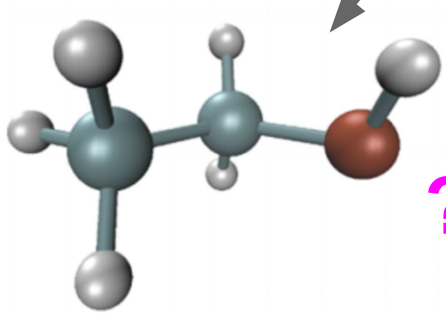
$\rightarrow$ log(Error) = log(a) - b log(c) - b log(N)



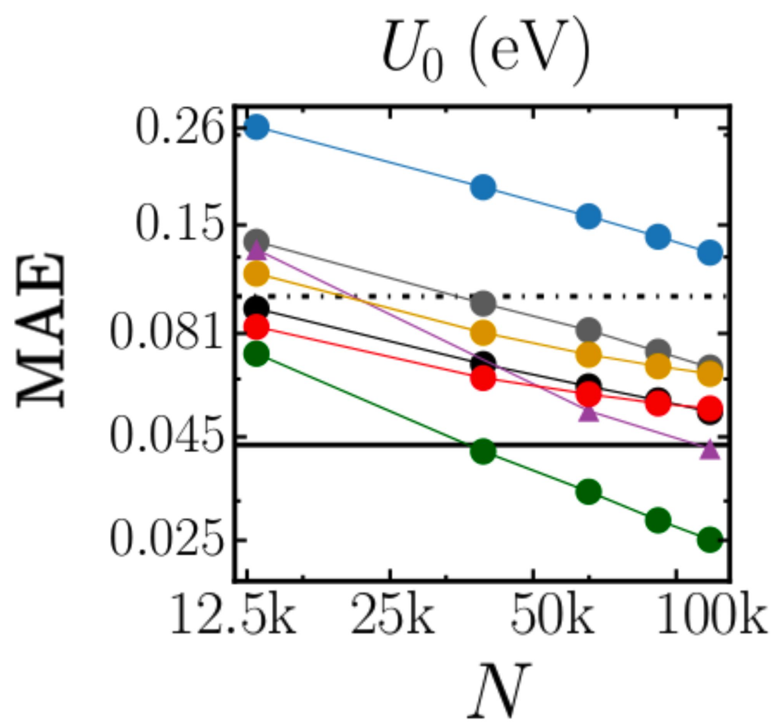
The bigger the data the better …

Vapnik, V., *The Nature of Statistical Learning Theory*, Springer (1995)

$$P^{\mathrm{est}}(\mathbf{M}) = \sum_i \alpha_i k(\mathbf{M}, \mathbf{M}_i)$$
$$\vec{\alpha} = \mathbf{K}^{-1} \vec{P}^{\mathrm{ref}}$$



???



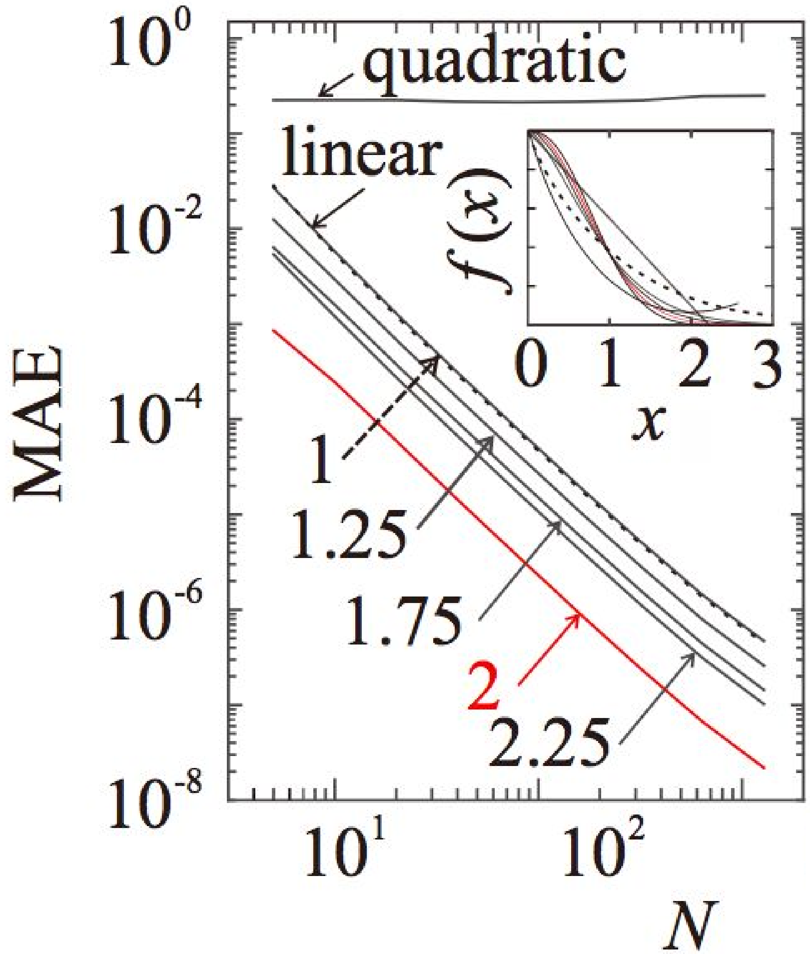
"Prediction errors of molecular machine learning models lower than hybrid DFT errors", F. A. Faber, L. Hutchison, B. Huang, J. Gilmer, S. S. Schoenholz, G. E. Dahl, O. Vinyals, S. Kearnes, P. F. Riley, OAvL *J Chem Theory Comput* (2017) arxiv.org/abs/1702.05532

Representation

$$\log(\text{Error}) = a - b\log(N)$$

target
similarity

$$f^{\text{est}}(x) = \sum \alpha_i k(\underbrace{ax_i + b}_{\text{M}_i}, \underbrace{ax + b}_{\text{M}})$$



Huang, OAvL, *J Chem Phys* (2016)

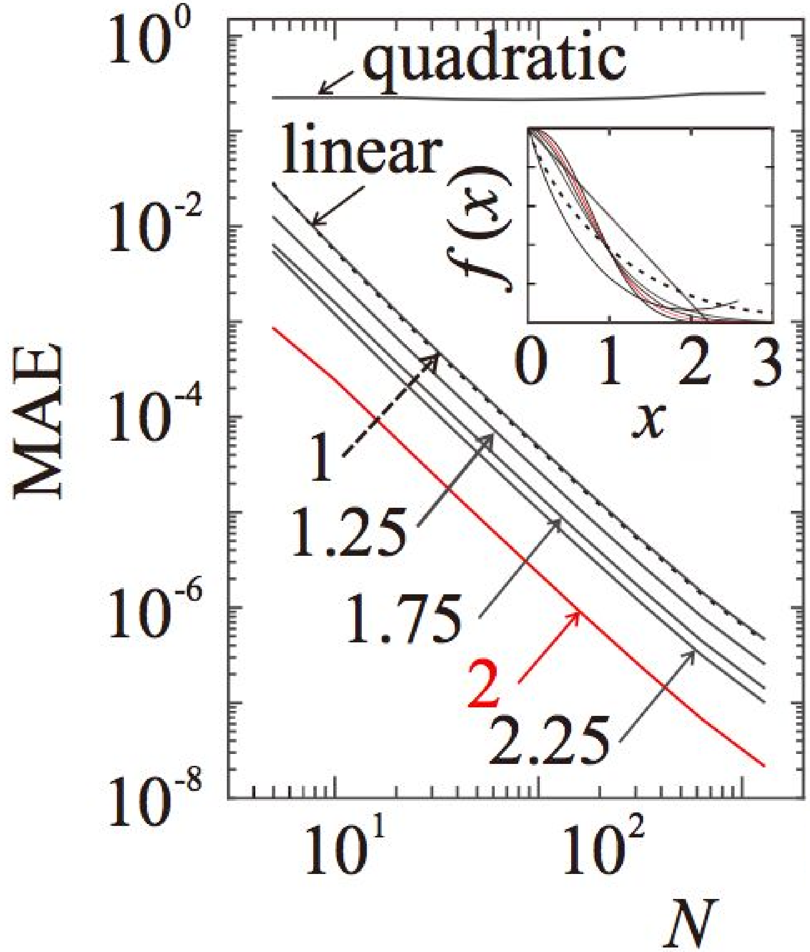
Representation

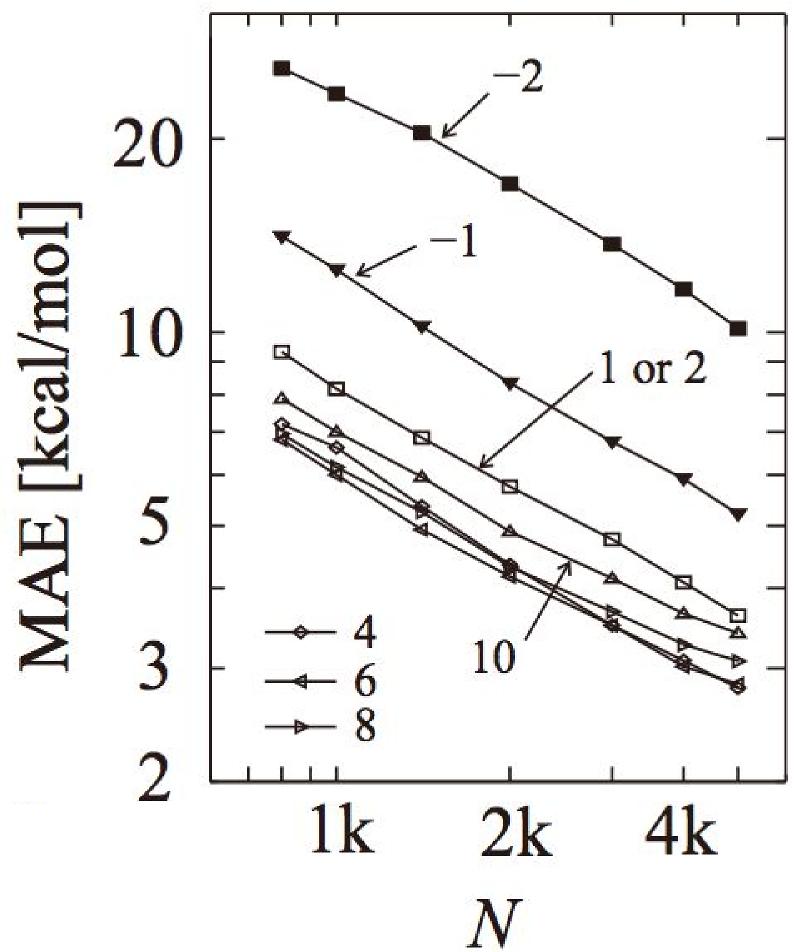
$$f^{\mathrm{est}}(x) = \sum \alpha_i k(\underbrace{ax_i + b}_{\mathrm{M}_i}, \underbrace{ax + b}_{\mathrm{M}})$$

$$\log(\mathrm{Error}) = a - b\log(N)$$

target similarity

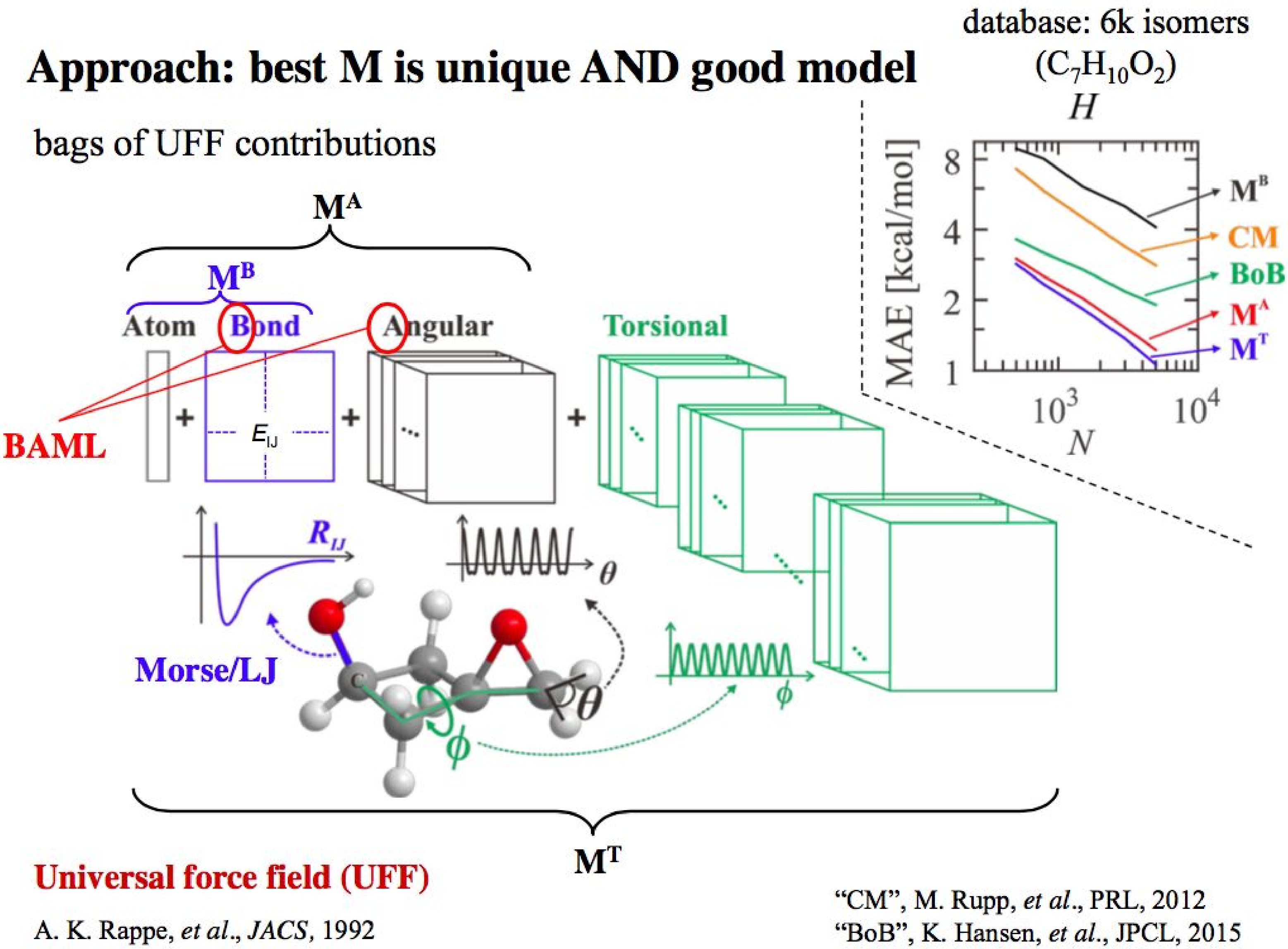
$$\mathrm{CM}_{IJ}^{(n)} = \frac{Z_I Z_J}{R_{IJ}^n}$$





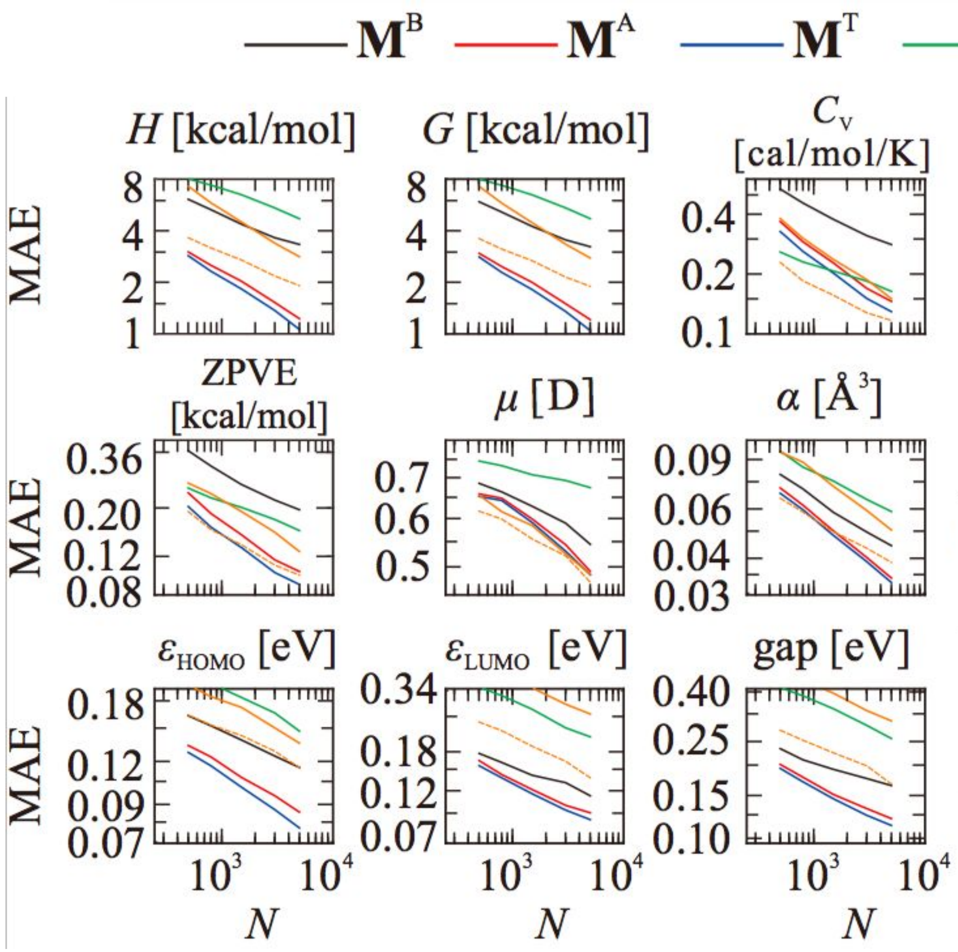
Huang, OAvL, *J Chem Phys* (2016)

BAML

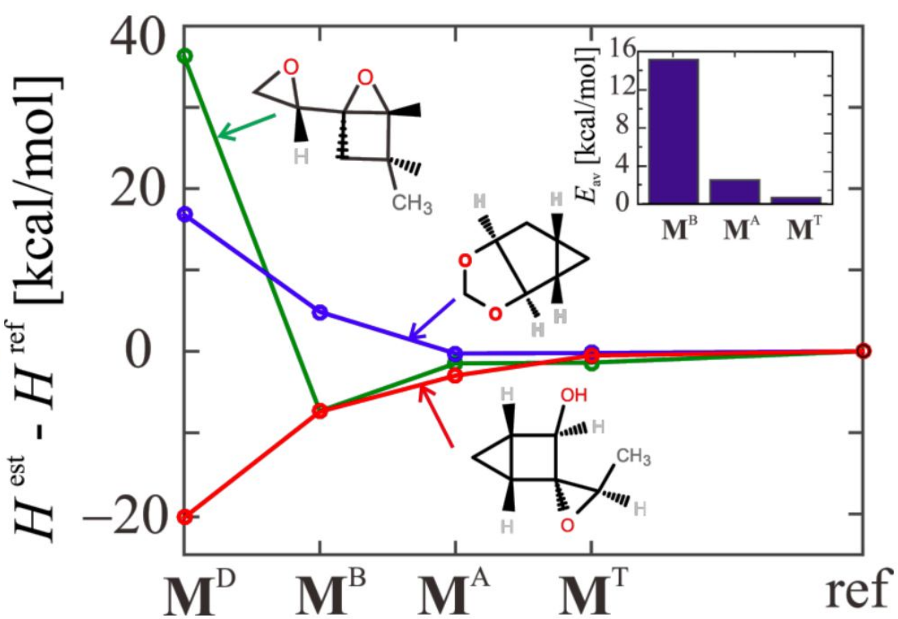


Huang, OAvL, *J Chem Phys* (2016)

BAML



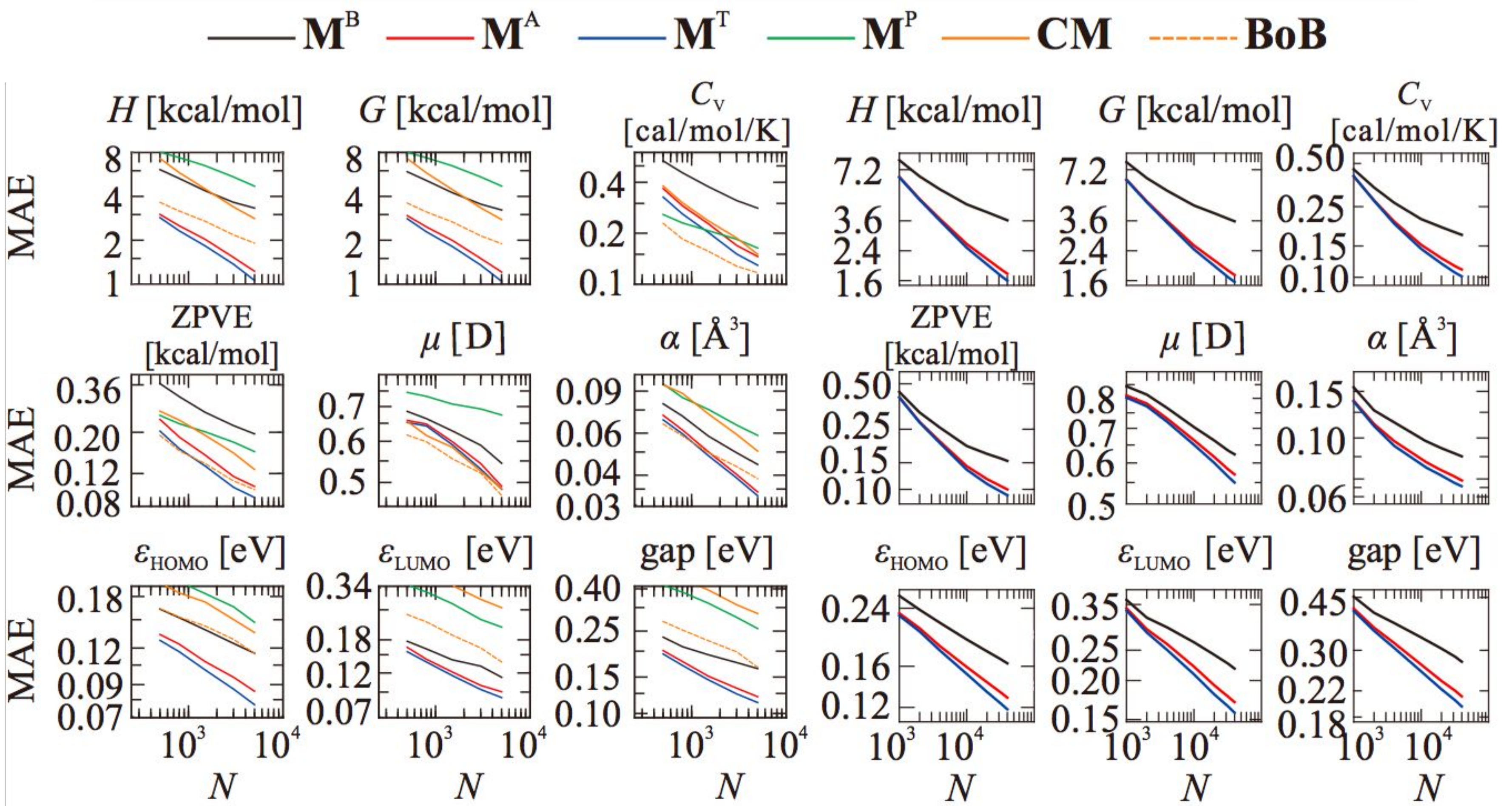
3 outliers



6k constitutional isomers of $C_7O_2H_{10}$

Huang, OAvL, *J Chem Phys* (2016)

BAML

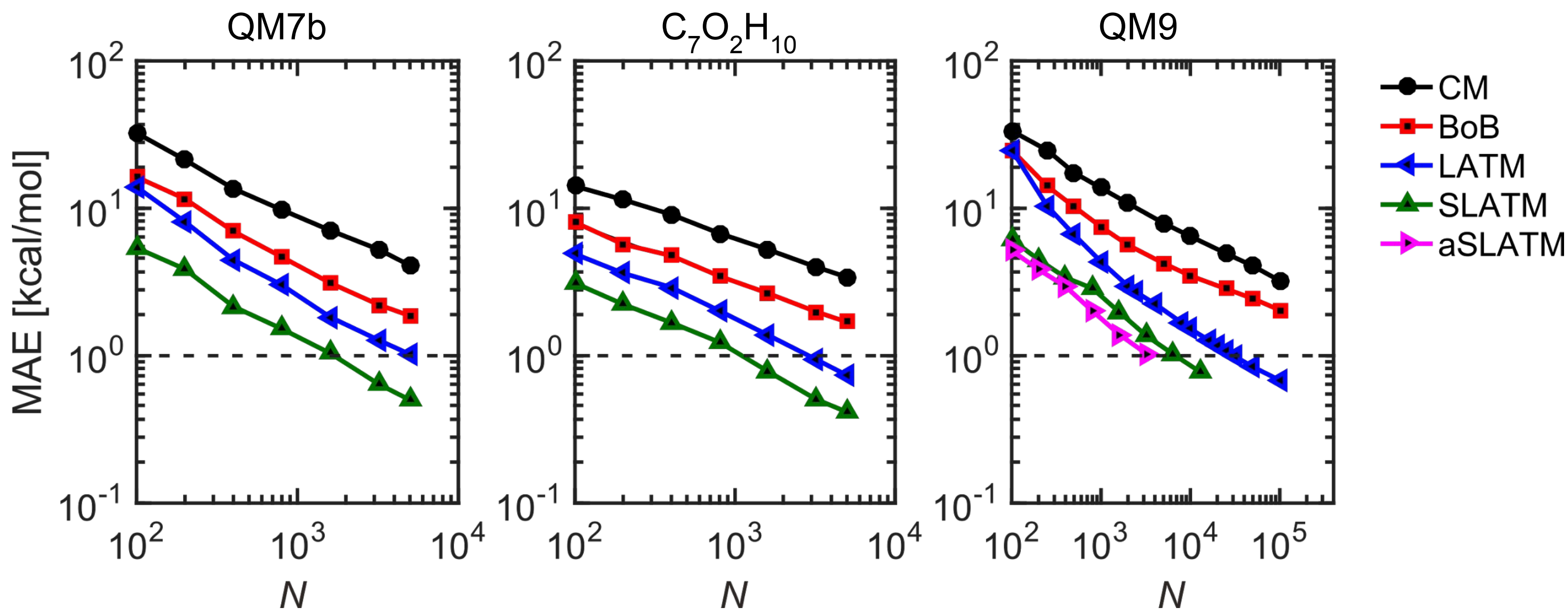


6k constitutional isomers of $C_7O_2H_{10}$

QM9 (134k molecules)

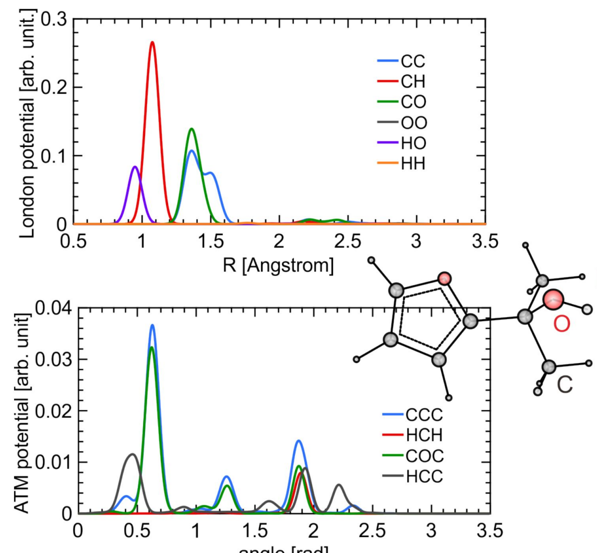
Huang, OAvL, *J Chem Phys* (2016)

Atoms + **L**ondon + **A**xilrod-**T**eller-**M**uto **(LATM)**

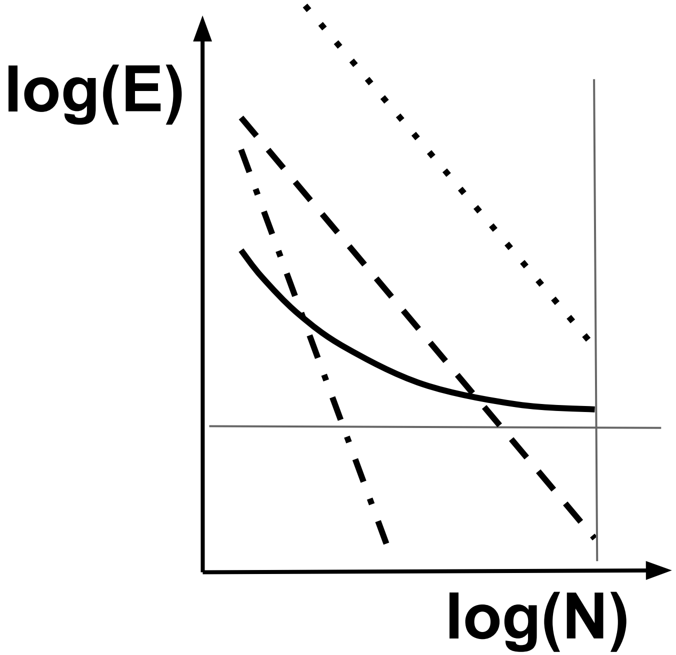


$$E^{(2)}\left(\mathbf{R_I},\mathbf{R_J}\right)=-\frac{C_{6_{IJ}}}{R_{IJ}^{6}}$$

$$E^{(3)}\left(\mathbf{R_I},\mathbf{R_J},\mathbf{R_K}\right)=C_{9_{IJK}}\frac{3\cos\left[\phi_I\right]\cos\left[\phi_J\right]\cos\left[\phi_K\right]+1}{R_{IJ}^{3}R_{IK}^{3}R_{JK}^{3}}$$



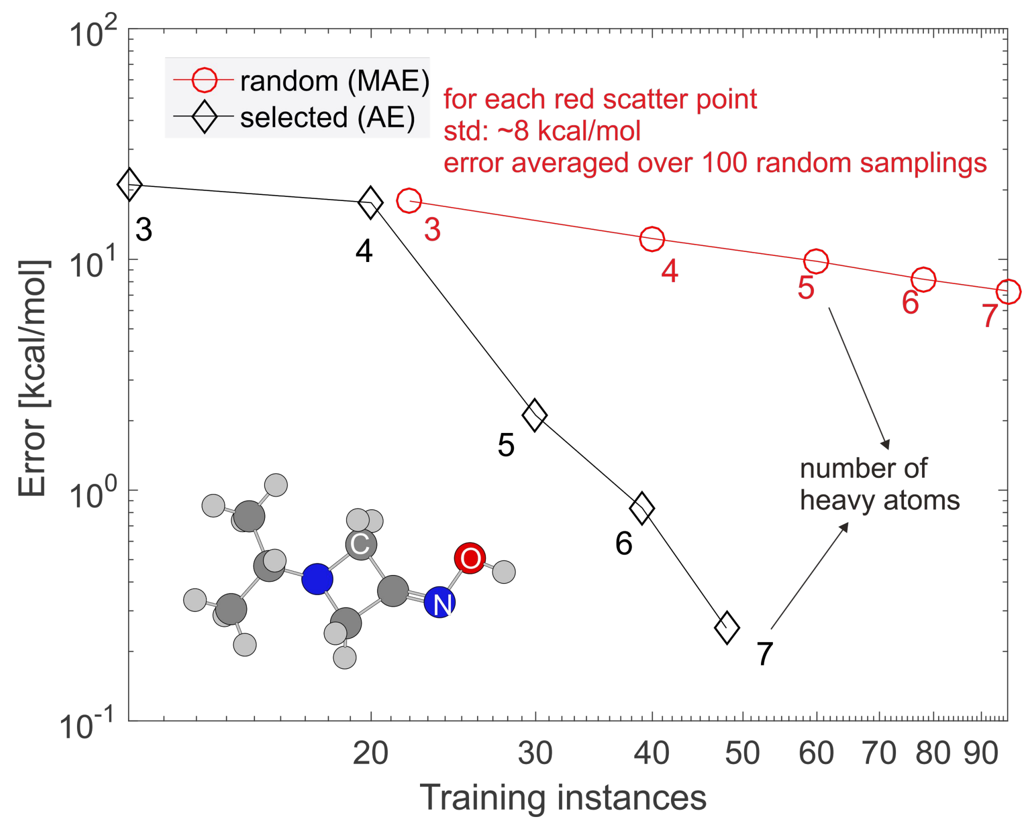
Huang, von Lilienfeld, https://arxiv.org/abs/1707.04146



The bigger the data the better …

Vapnik, V., *The Nature of Statistical Learning Theory,* Springer (1995)

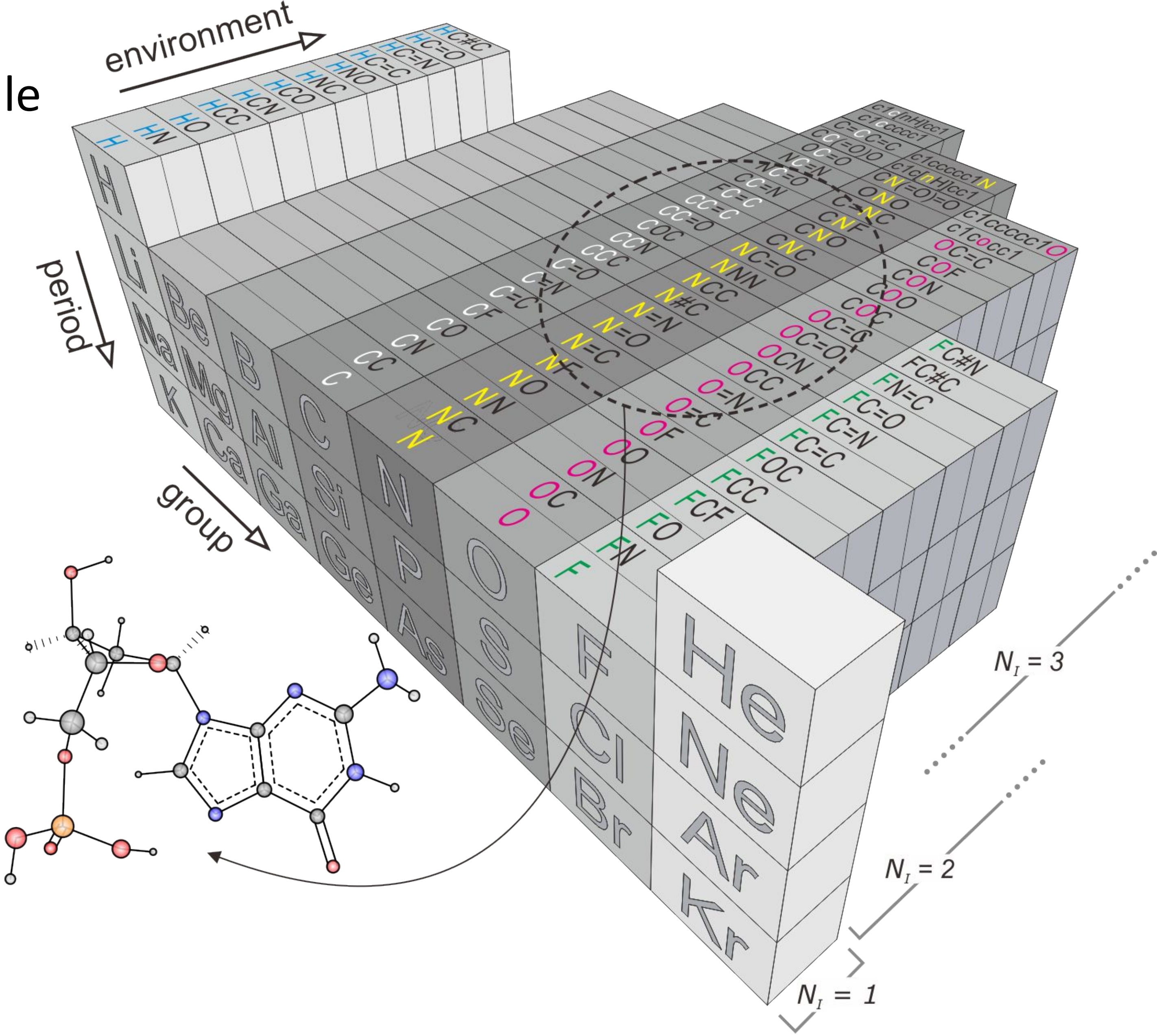
$$P^{\text{est}}(\mathbf{M}) = \sum_i \alpha_i k(\mathbf{M}, \mathbf{M}_i)$$
$$\vec{\alpha} = \mathbf{K}^{-1} \vec{P}^{\text{ref}}$$



Huang, OAvL, *J Chem Phys* (2016)

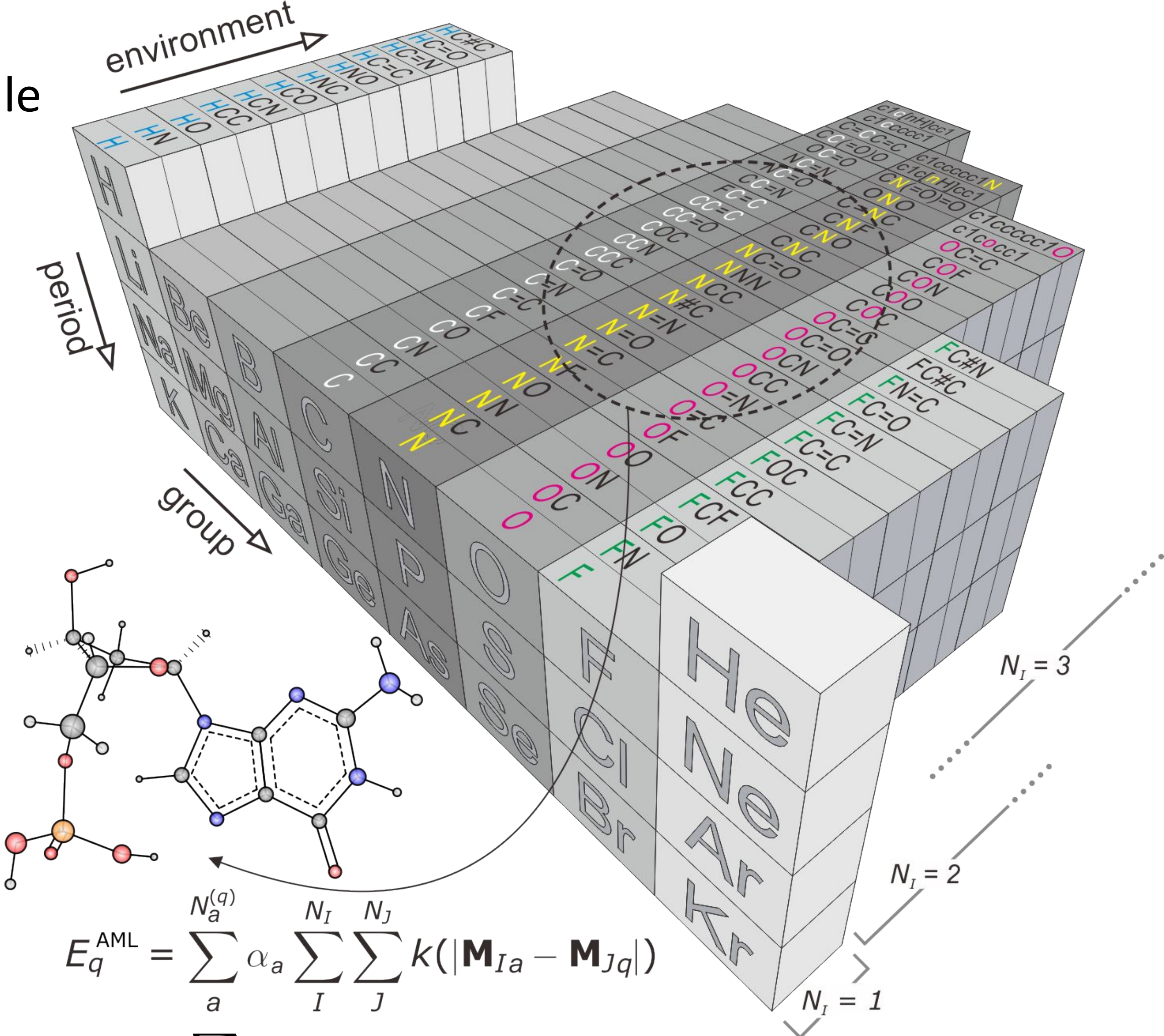
Huang, von Lilienfeld, https://arxiv.org/abs/1707.04146

<u>A</u>tom in a <u>M</u>olecule
"AM-on"



Huang, von Lilienfeld, https://arxiv.org/abs/1707.04146

Atom in a Molecule
"AM-on"



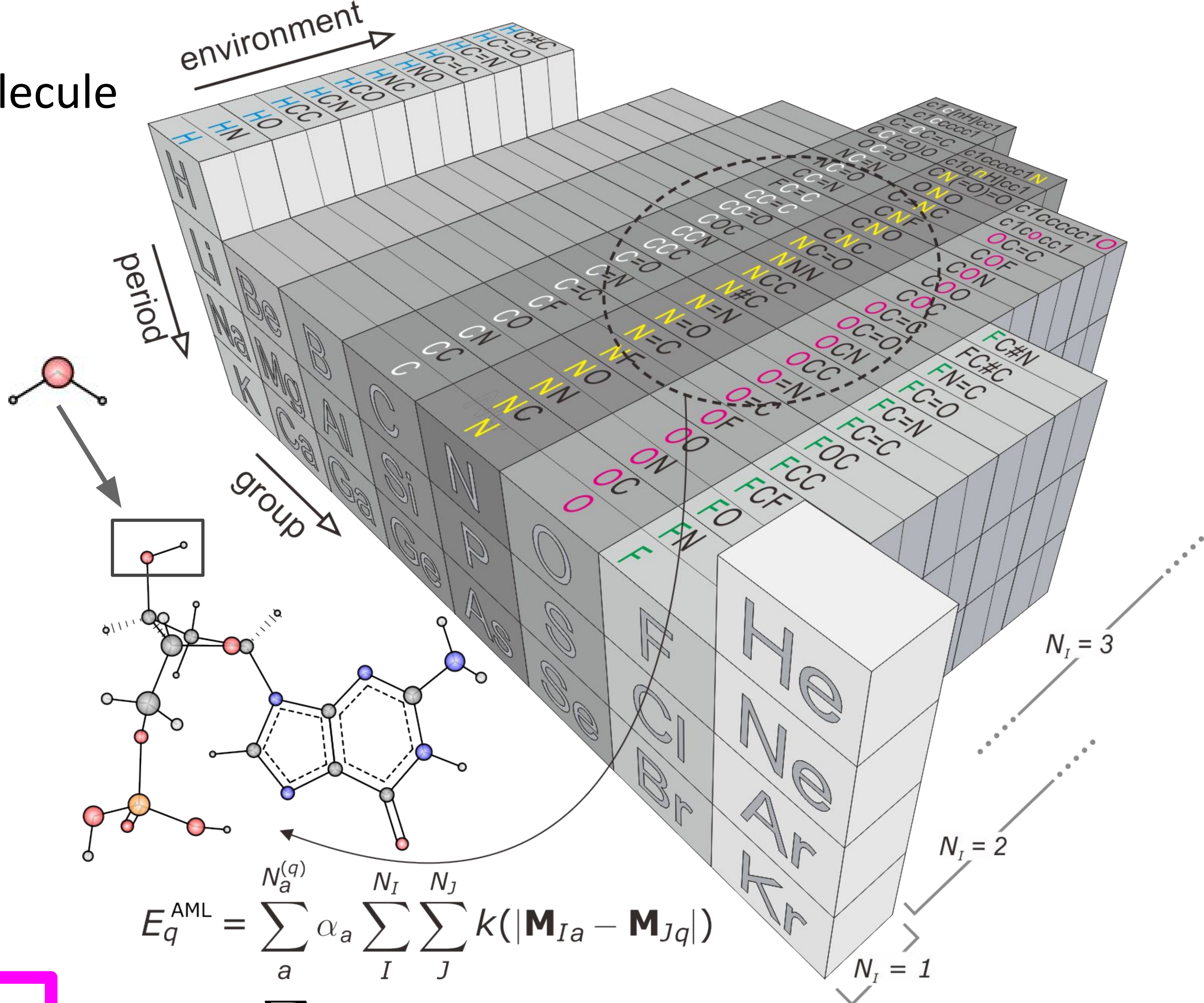
$$E_q^{\text{AML}} = \sum_a^{N_a^{(q)}} \alpha_a \sum_I^{N_I} \sum_J^{N_J} k(|\mathbf{M}_{Ia} - \mathbf{M}_{Jq}|)$$

$$= \sum_J E_q^J$$

$$\boxed{N_I < N_J}$$

Huang, von Lilienfeld, https://arxiv.org/abs/1707.04146

<u>A</u>tom in a <u>M</u>olecule
“AM-on”



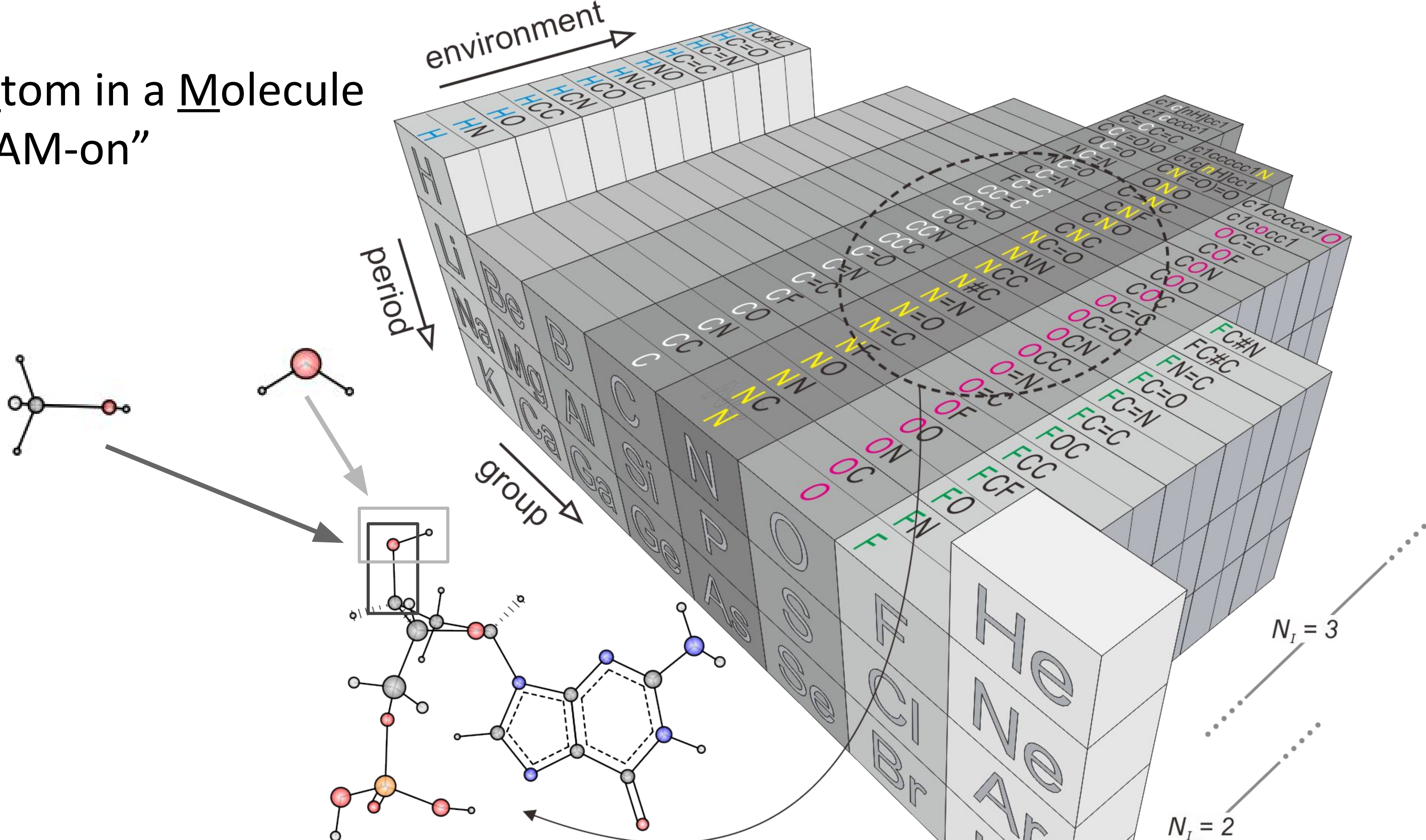
$$E_q^{\mathrm{AML}} = \sum_{a}^{N_a^{(q)}} \alpha_a \sum_{I}^{N_I} \sum_{J}^{N_J} k(|\mathbf{M}_{Ia} - \mathbf{M}_{Jq}|)$$

$$= \sum_J E_q^J$$

$$\boxed{N_I \quad < \quad N_J}$$

Huang, von Lilienfeld, https://arxiv.org/abs/1707.04146

Atom in a Molecule
“AM-on”

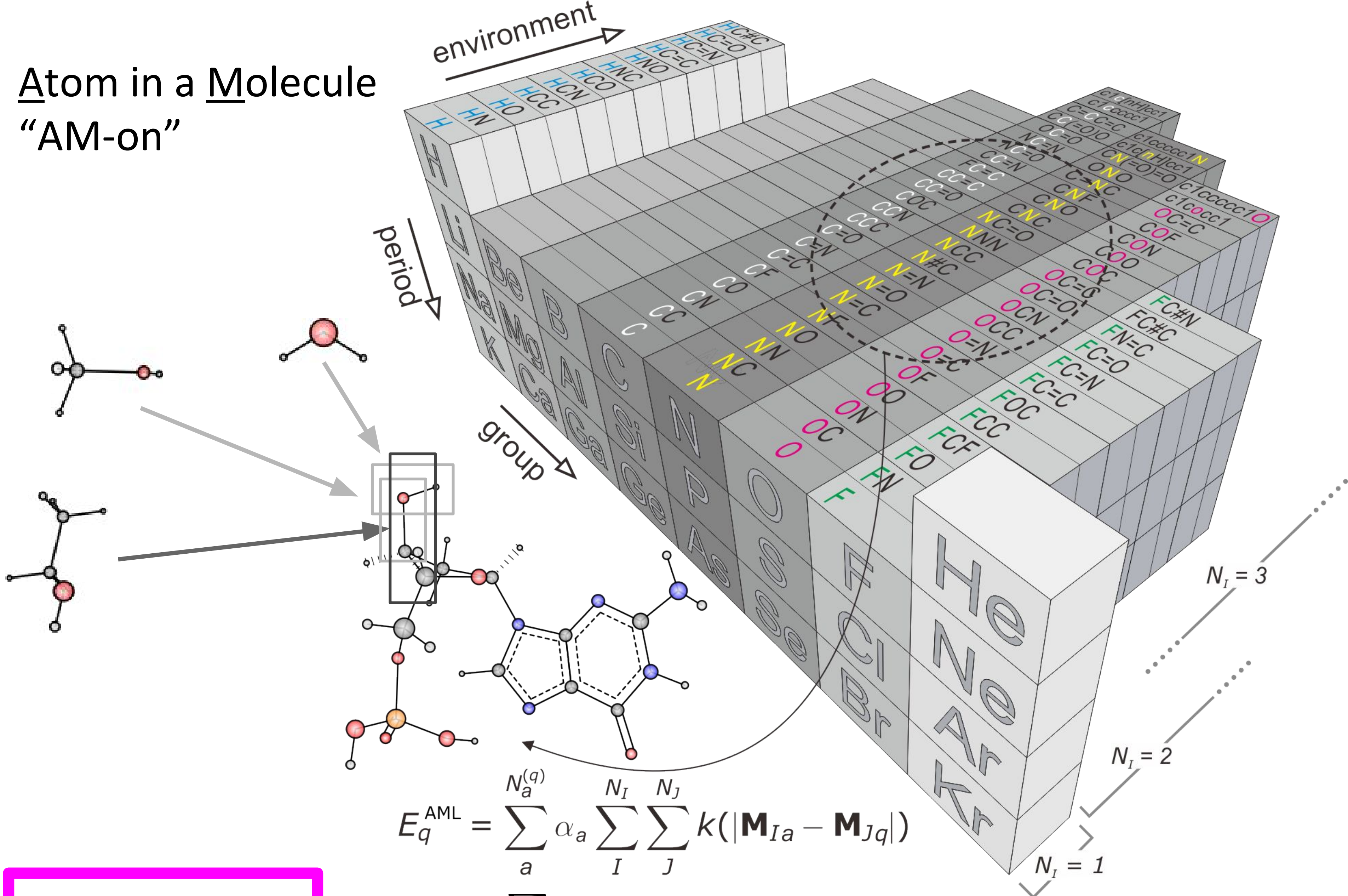


$$E_q^{\text{AML}} = \sum_{a}^{N_a^{(q)}} \alpha_a \sum_{I}^{N_I} \sum_{J}^{N_J} k(|\mathbf{M}_{Ia} - \mathbf{M}_{Jq}|)$$
$$= \sum_J E_q^J$$

$$\boxed{N_I \quad < \quad N_J}$$

Huang, von Lilienfeld, https://arxiv.org/abs/1707.04146

Atom in a Molecule
"AM-on"



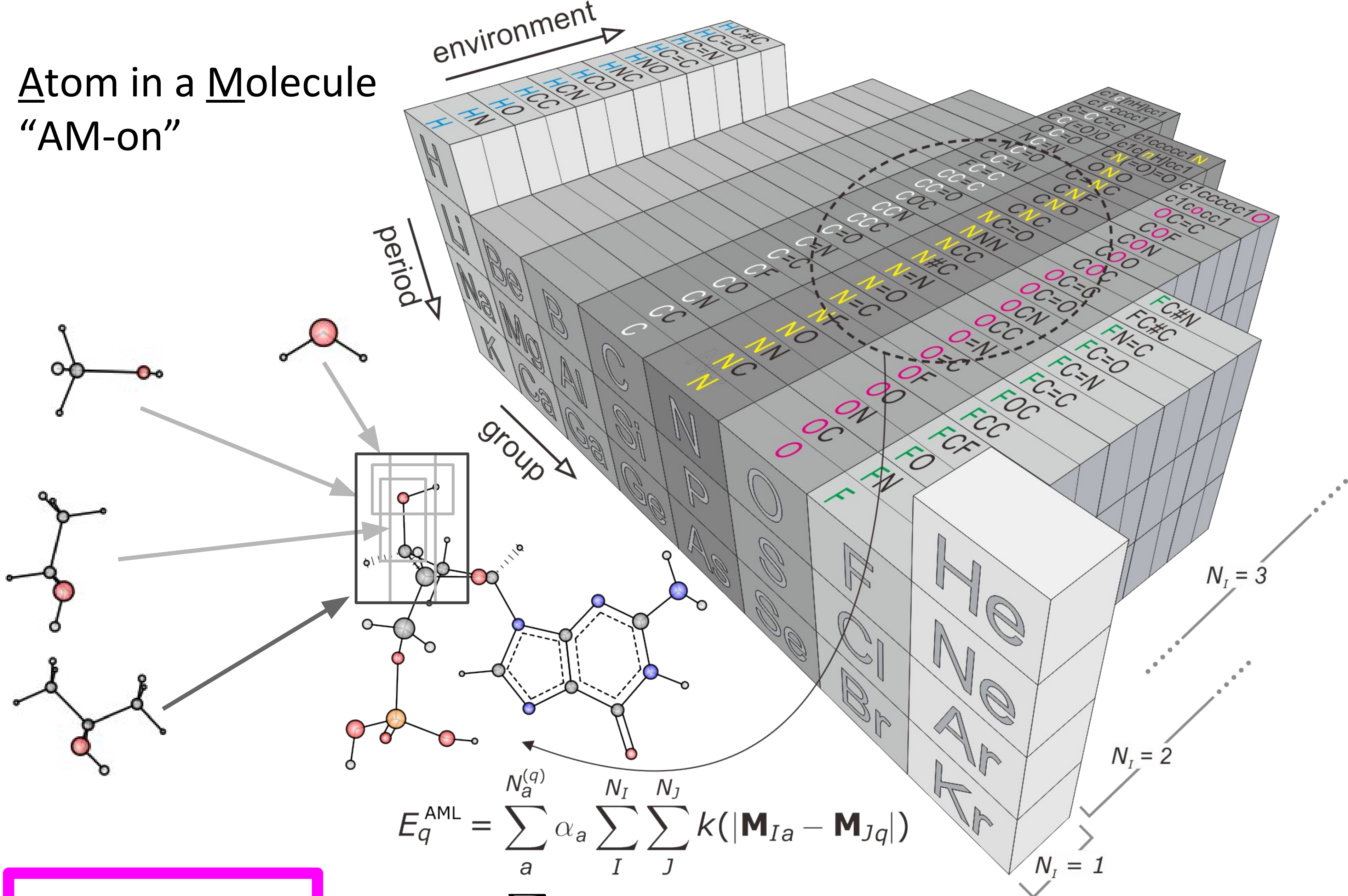
$$E_q^{\mathrm{AML}} = \sum_{a}^{N_a^{(q)}} \alpha_a \sum_{I}^{N_I} \sum_{J}^{N_J} k(|\mathbf{M}_{Ia} - \mathbf{M}_{Jq}|)$$

$$= \sum_J E_q^J$$

$$\boxed{N_I < N_J}$$

Huang, von Lilienfeld, https://arxiv.org/abs/1707.04146

Atom in a Molecule
"AM-on"



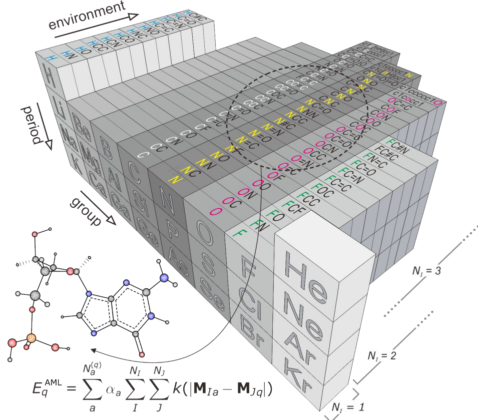
$$E_q^{\mathrm{AML}} = \sum_a^{N_a^{(q)}} \alpha_a \sum_I^{N_I} \sum_J^{N_J} k(|\mathbf{M}_{Ia} - \mathbf{M}_{Jq}|)$$

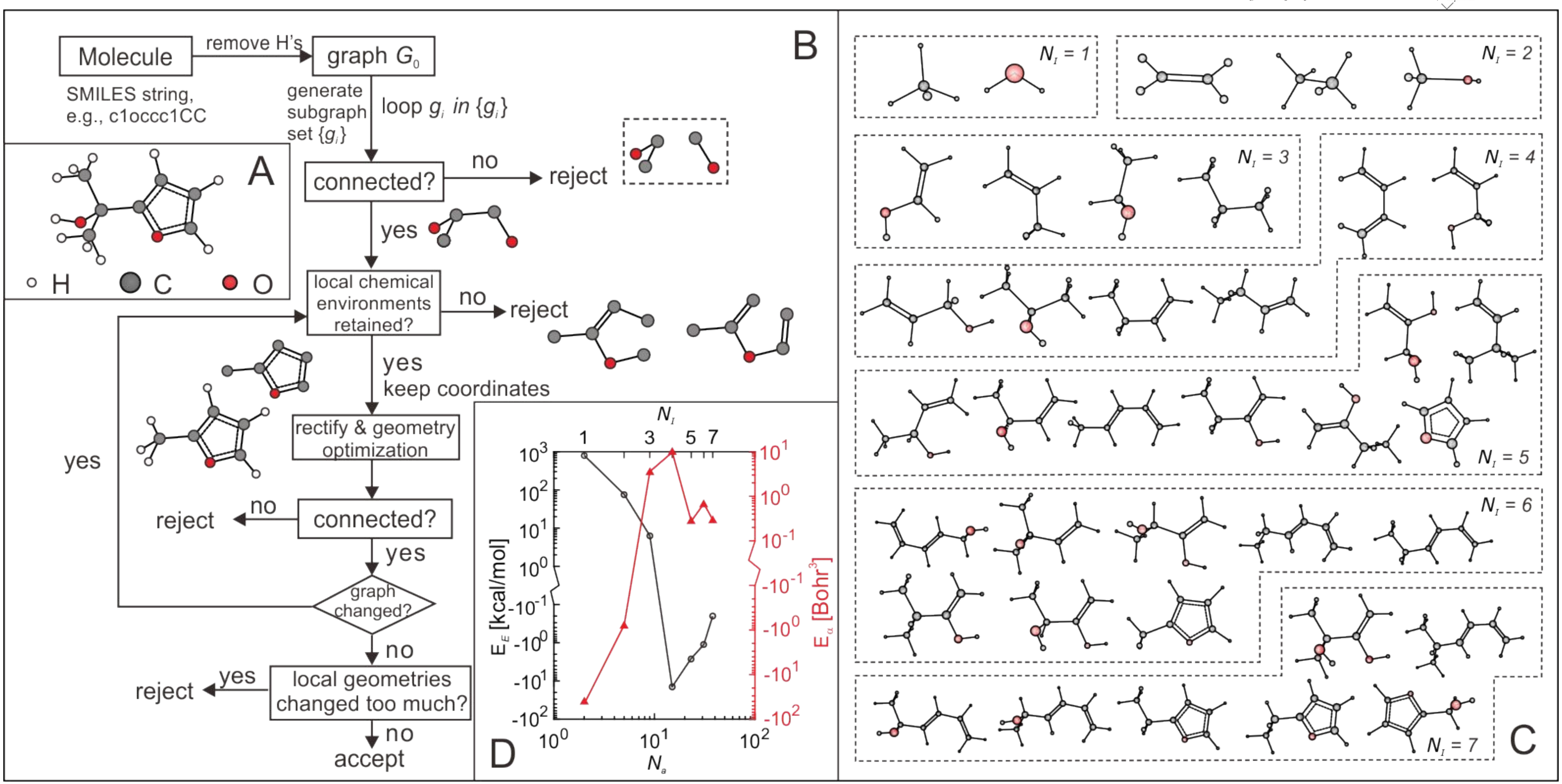
$$= \sum_J E_q^J$$

$$\boxed{N_I \quad < \quad N_J}$$

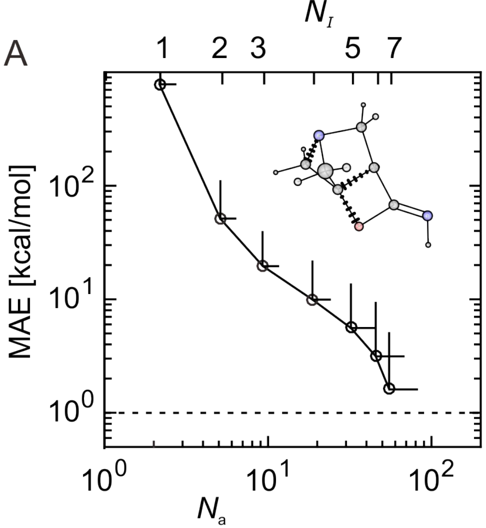
Huang, von Lilienfeld, https://arxiv.org/abs/1707.04146

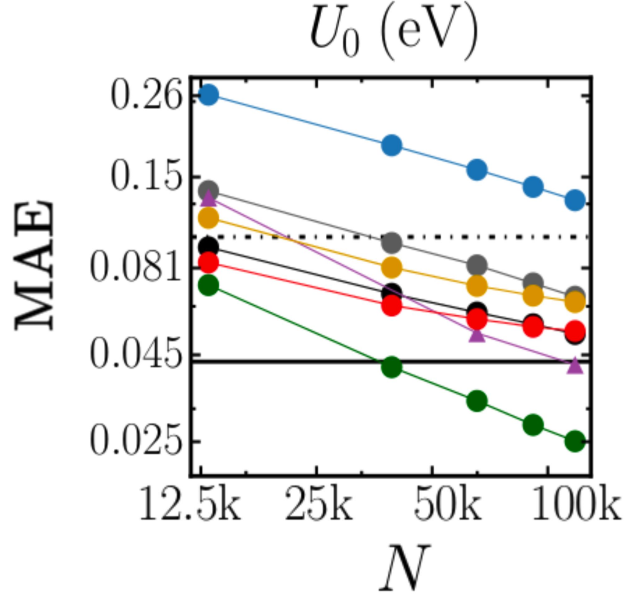
an atom in a molecule: “AM-on”

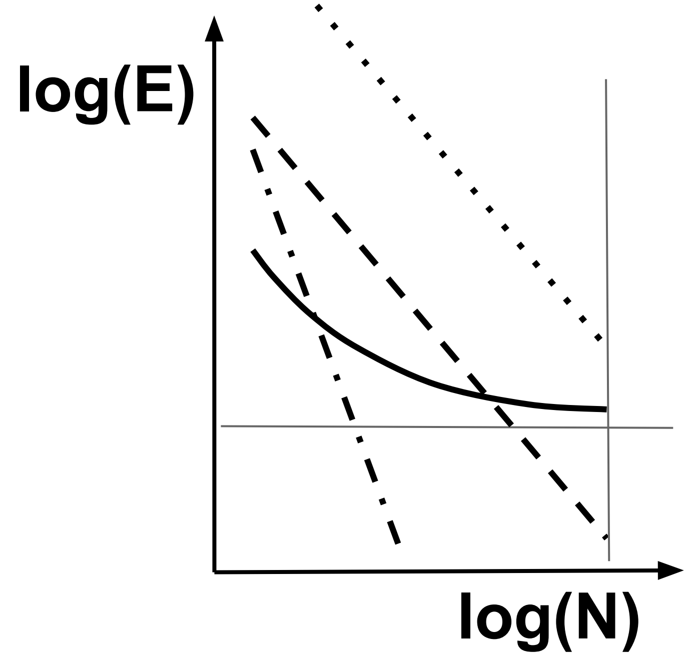


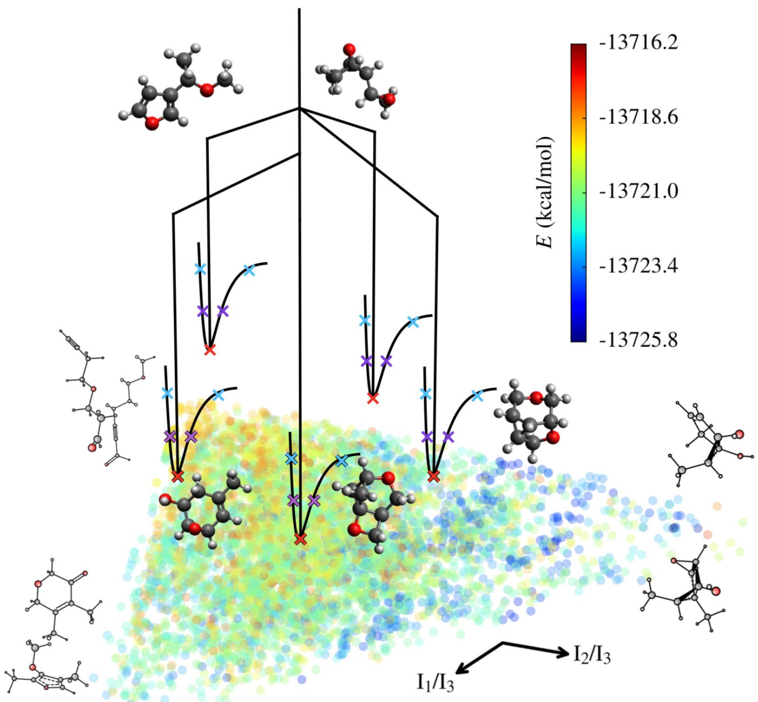


Huang, von Lilienfeld, https://arxiv.org/abs/1707.04146

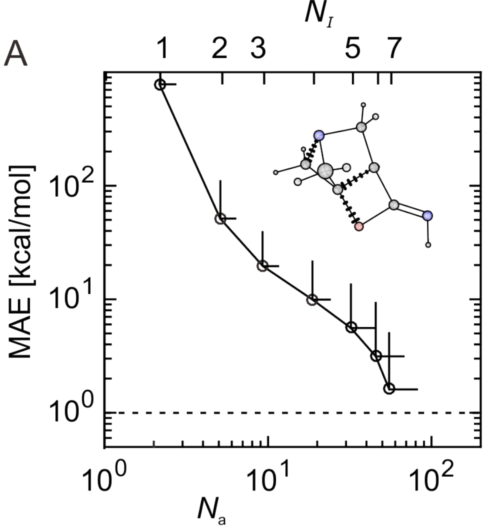


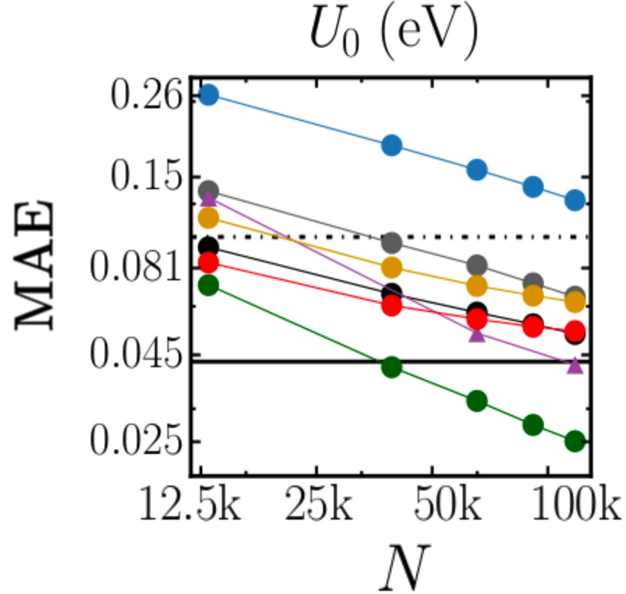


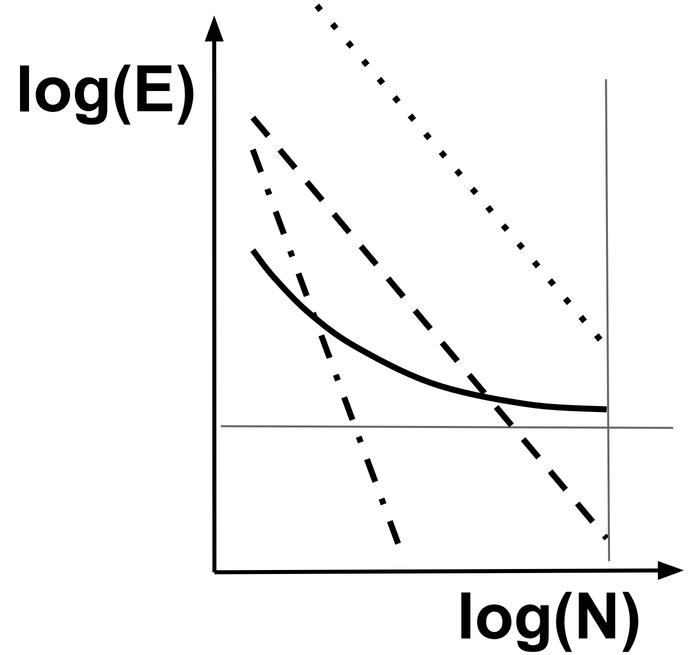


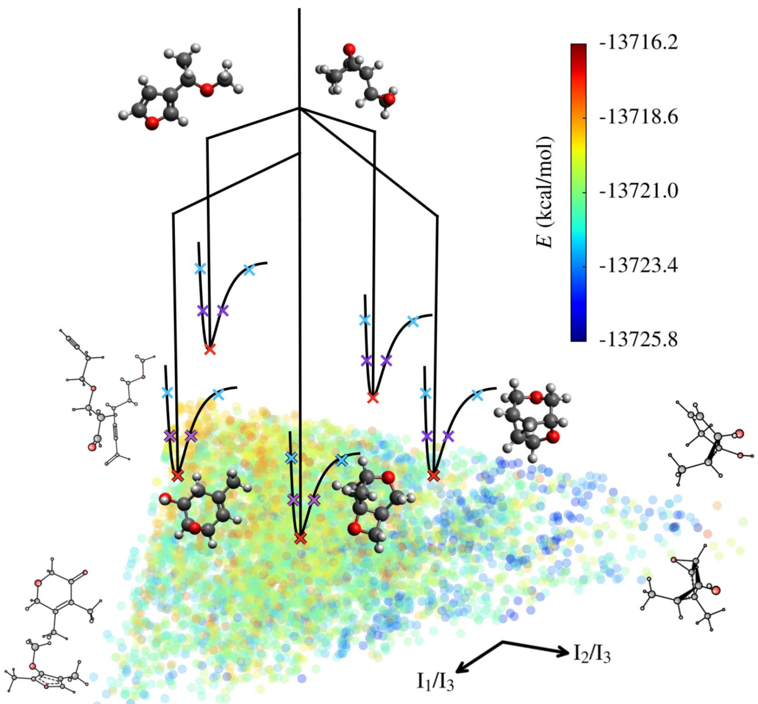


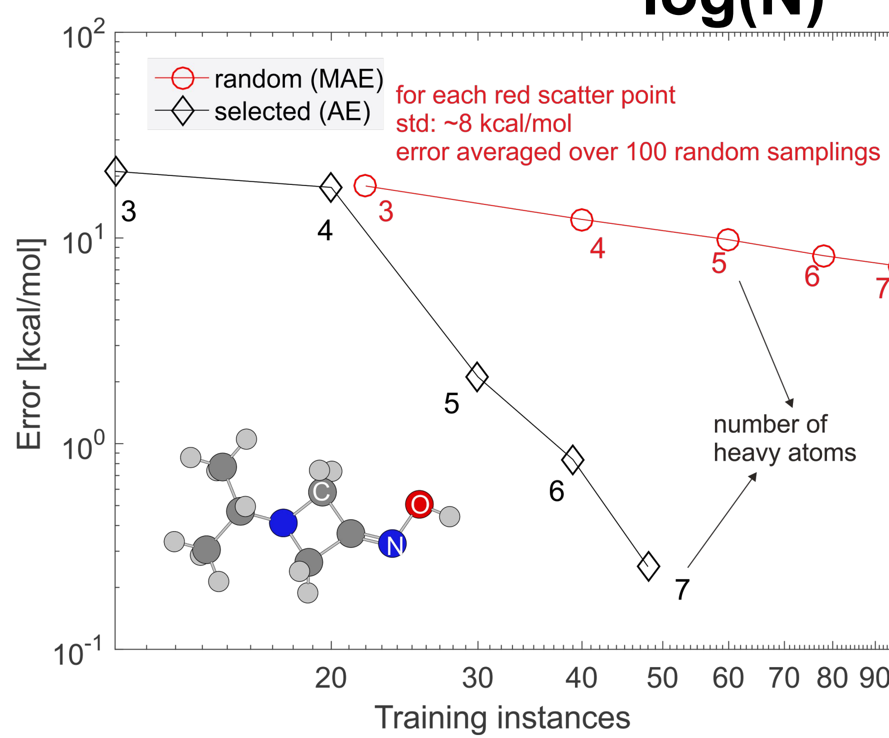
Huang, von Lilienfeld, https://arxiv.org/abs/1707.04146



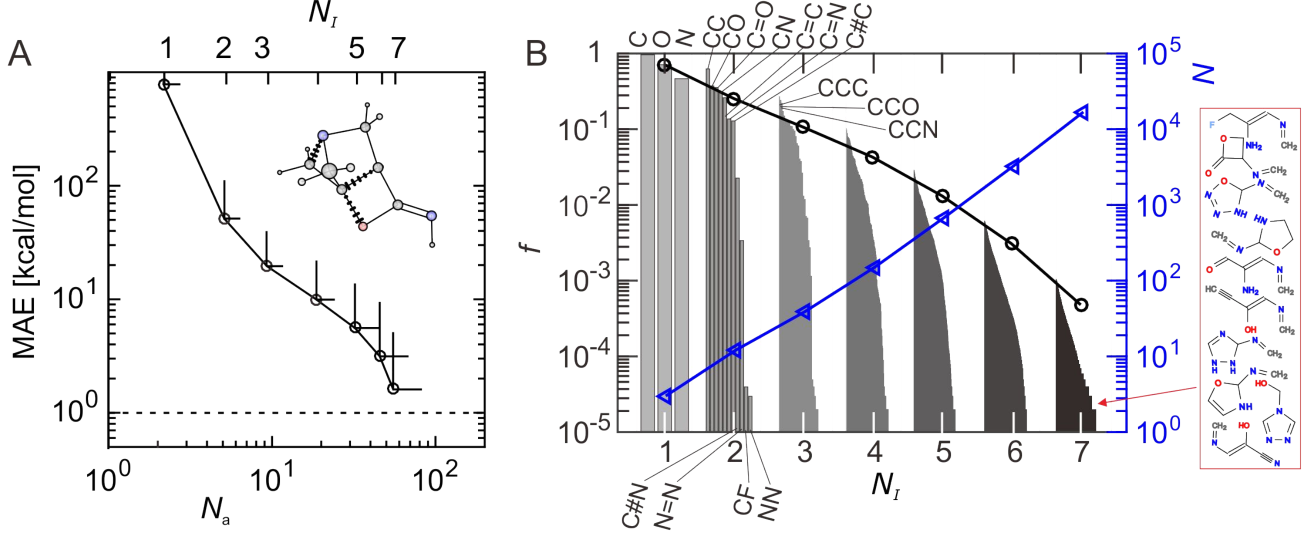


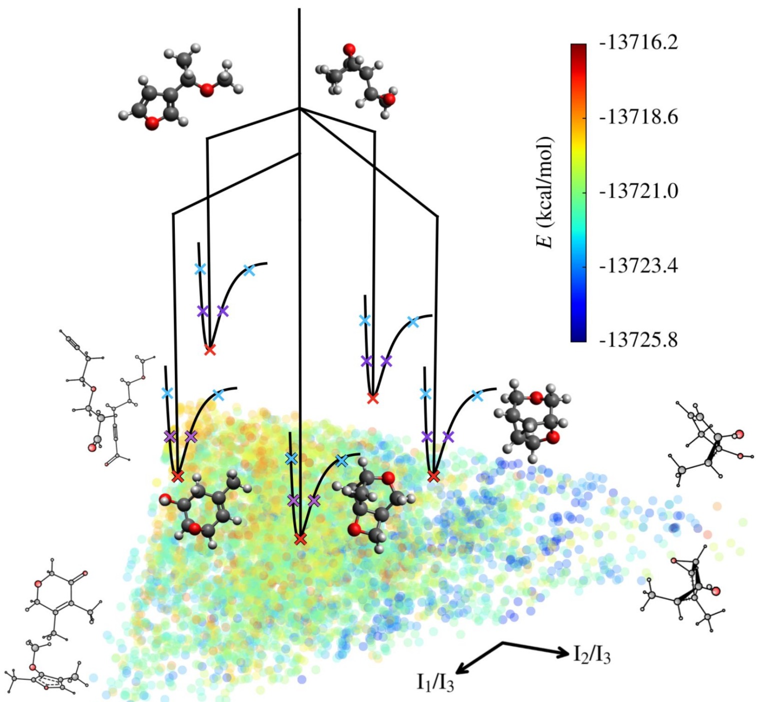


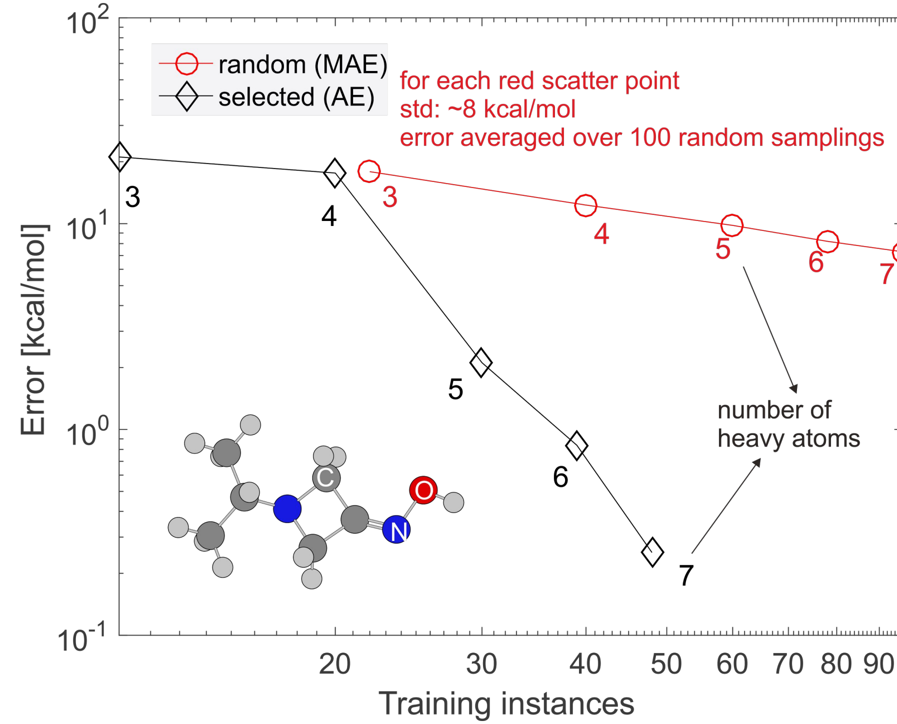




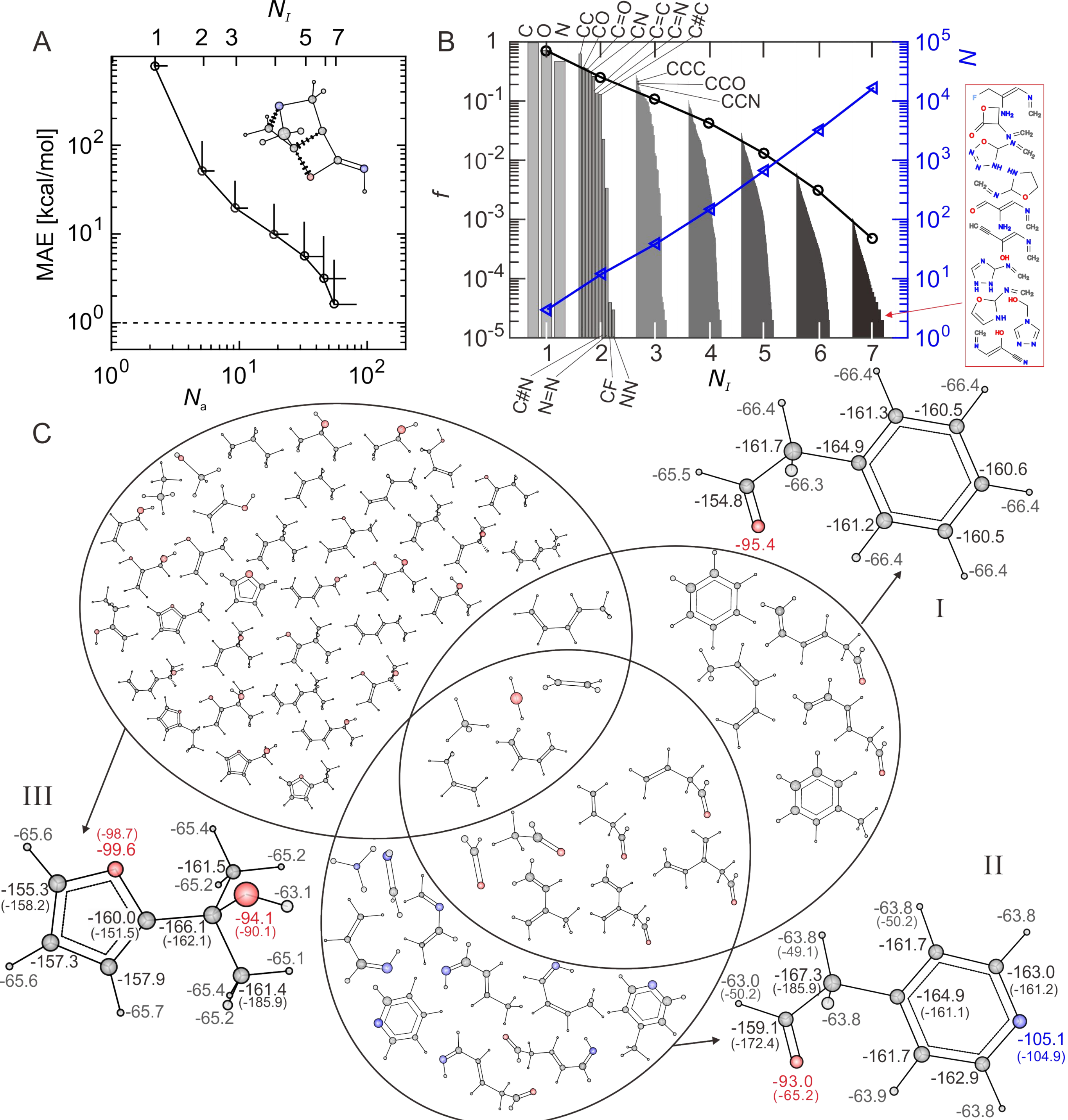
Huang, von Lilienfeld, https://arxiv.org/abs/1707.04146



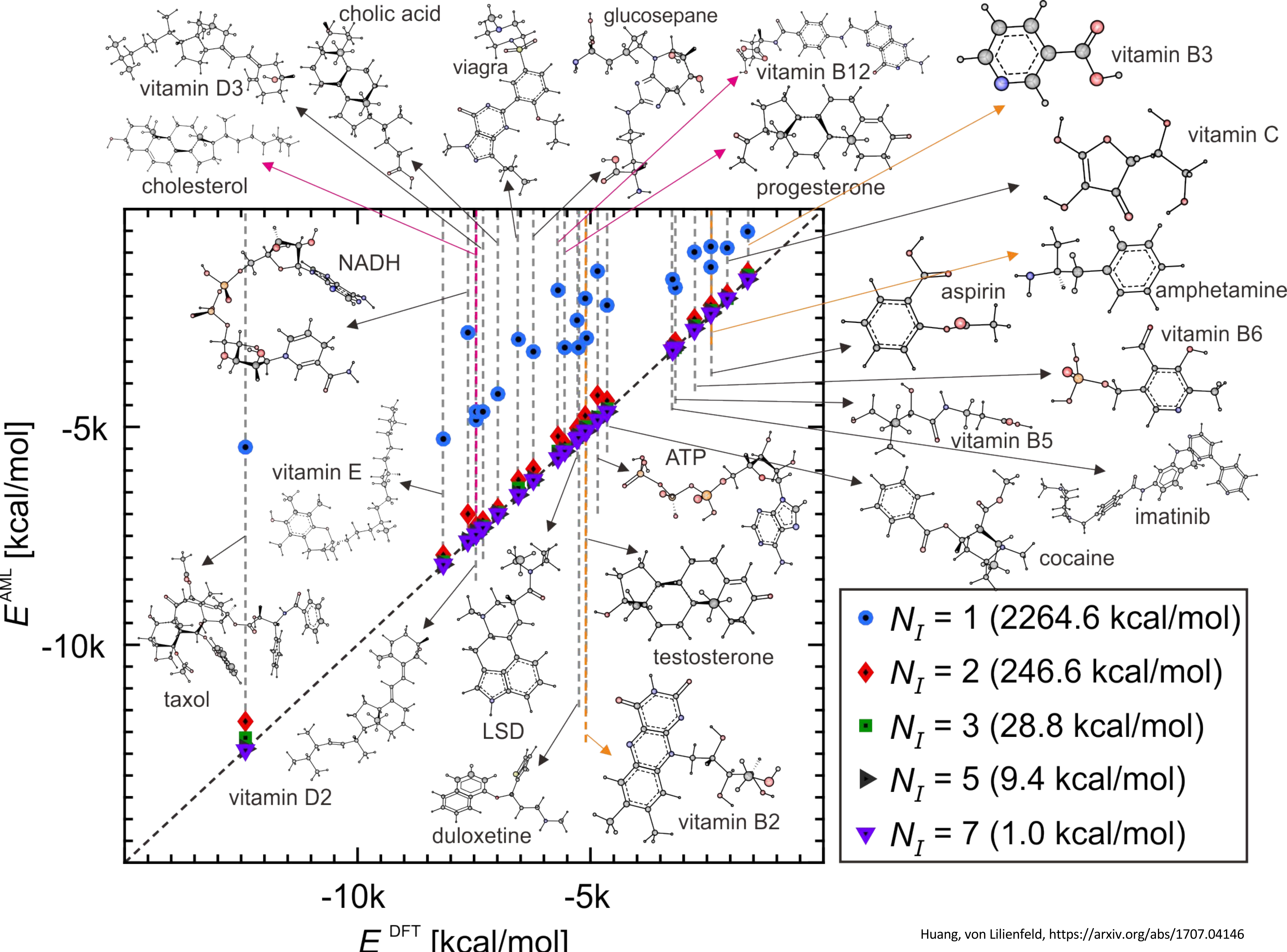




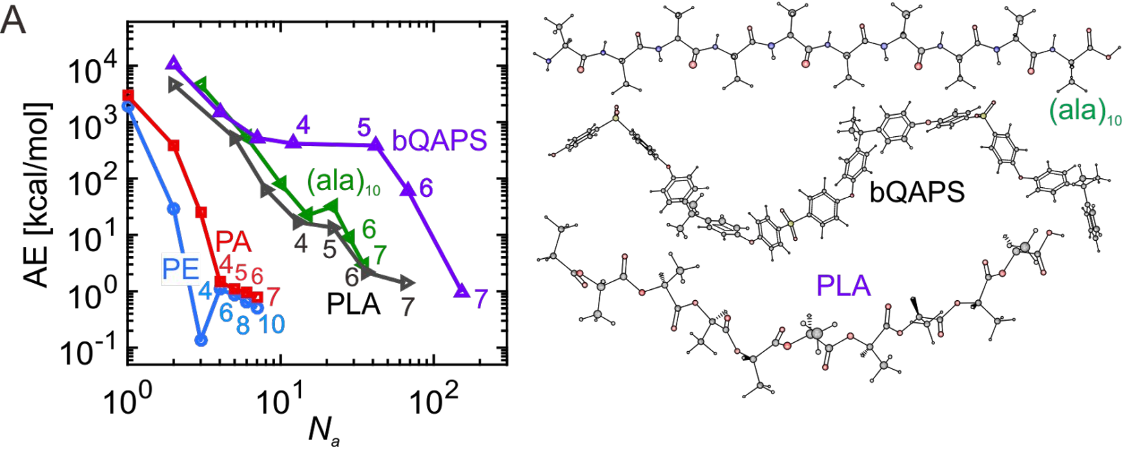
Huang, von Lilienfeld, https://arxiv.org/abs/1707.04146



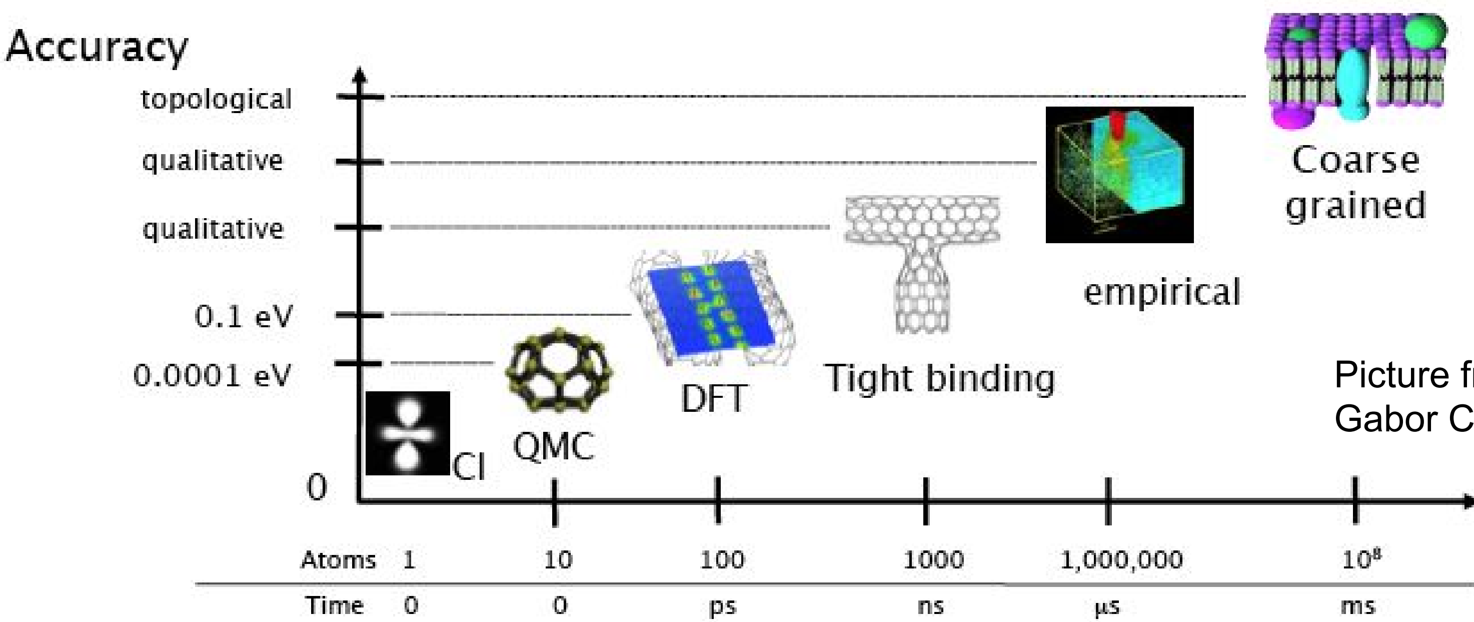
Huang, von Lilienfeld, https://arxiv.org/abs/1707.04146



Huang, von Lilienfeld, https://arxiv.org/abs/1707.04146

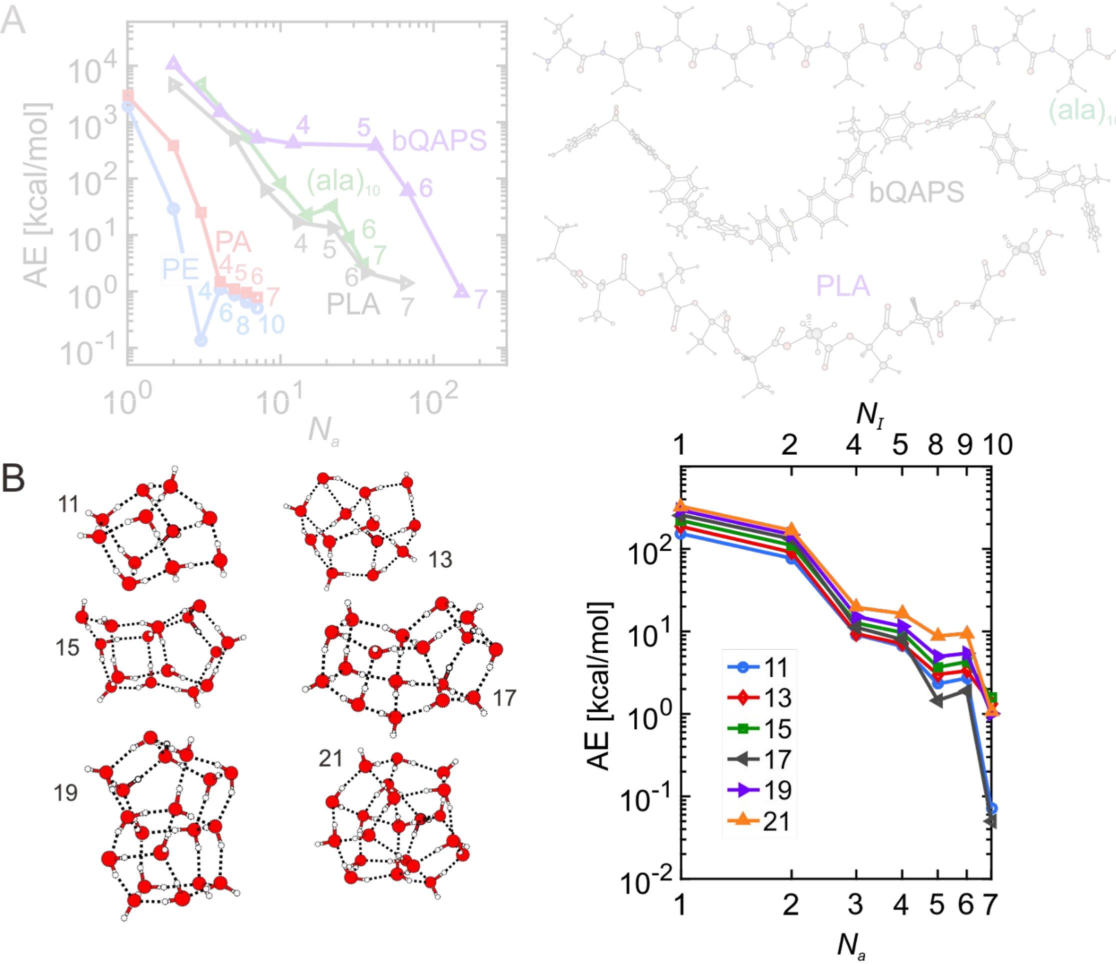


Polymer	# monomers
Polyethylene (PE)	28
Polyacetylene (PA)	15
Polyalanine	10
Polylactic acid (PLA)	10
Polysulphone (bQAPS)	3

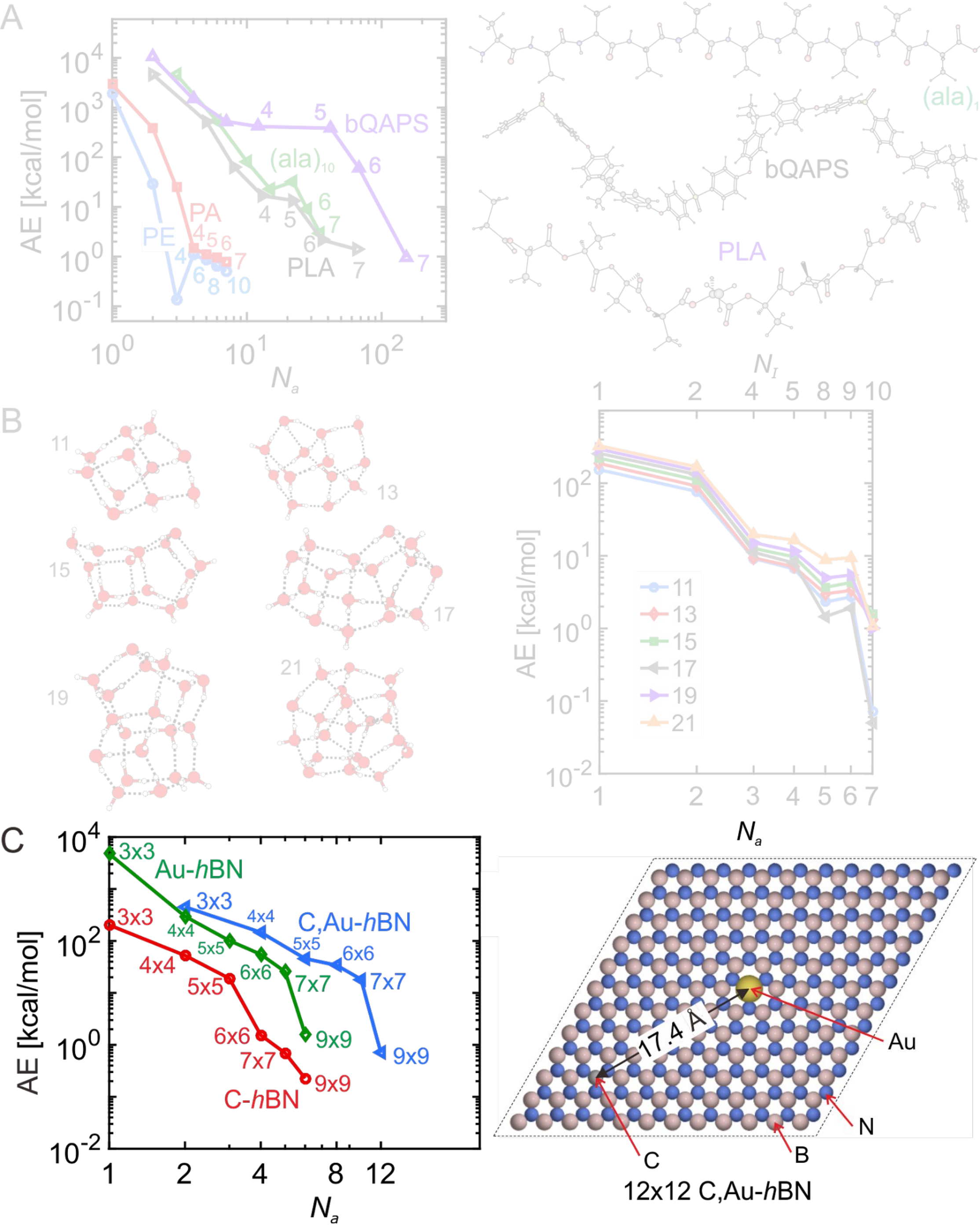


Picture from Gabor Csanyi

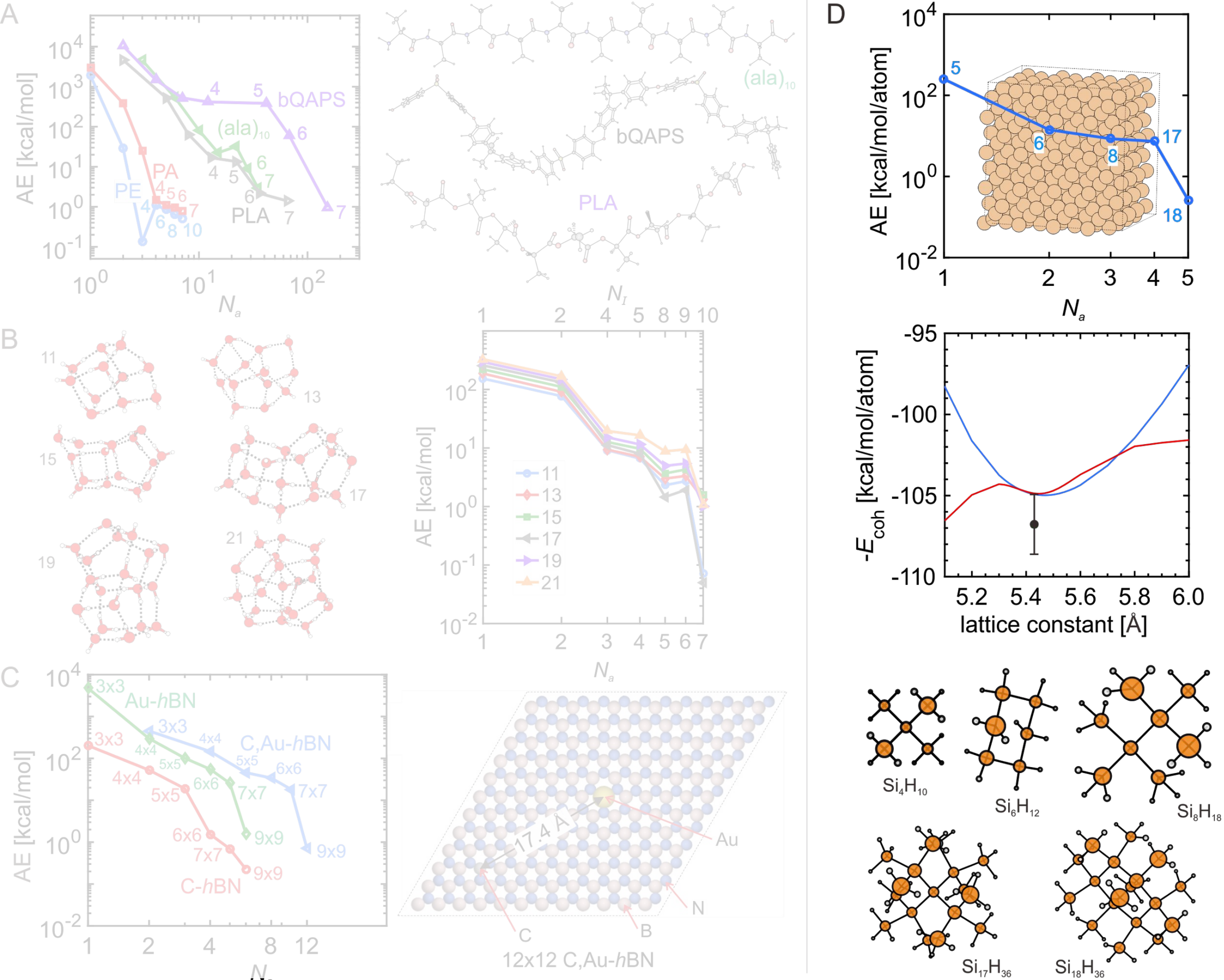
Huang, von Lilienfeld, https://arxiv.org/abs/1707.04146



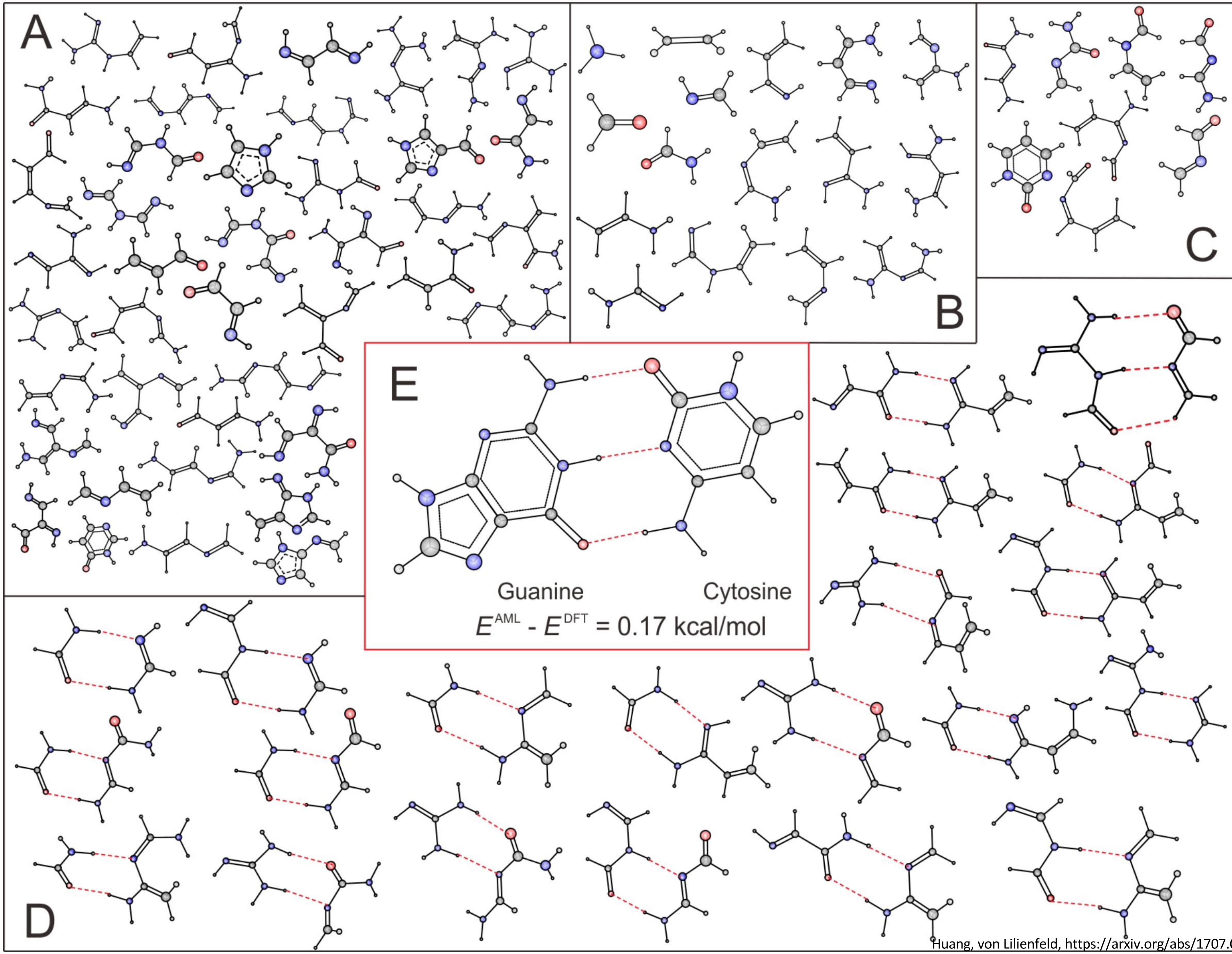
Huang, von Lilienfeld, https://arxiv.org/abs/1707.04146



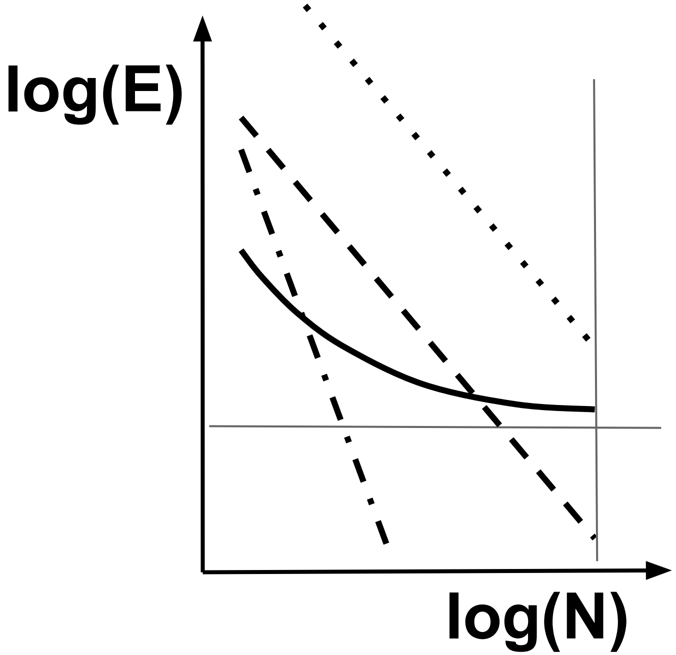
Huang, von Lilienfeld, https://arxiv.org/abs/1707.04146



Huang, von Lilienfeld, https://arxiv.org/abs/1707.04146



Huang, von Lilienfeld, https://arxiv.org/abs/1707.04146



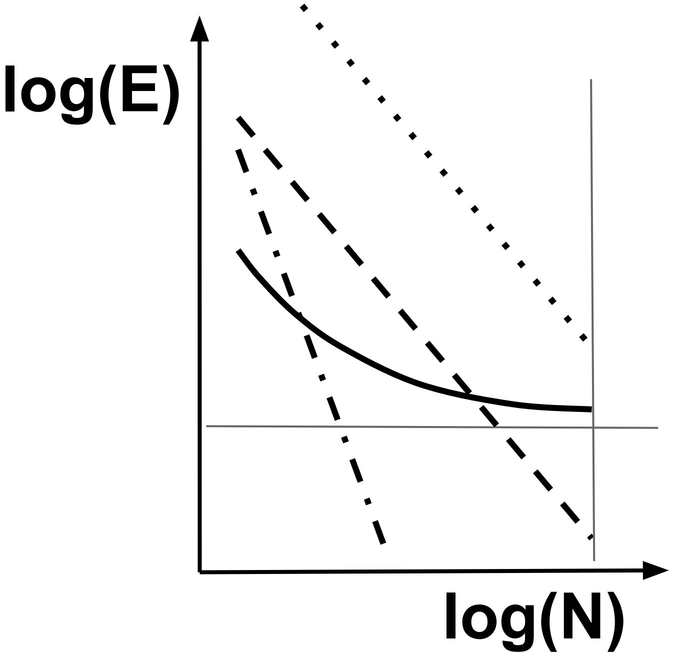
$$E_q = \langle \Psi_q | \hat{H} | \Psi_q \rangle$$
$$O_q = \langle \Psi_q | \hat{O} | \Psi_q \rangle$$

K ~ Ψ

α ~ Ô

Ramakrishnan, OAvL, *CHIMIA* (2015)

$$P^{\text{est}}(\mathbf{M}) = \sum_i \alpha_i k(\mathbf{M}, \mathbf{M}_i)$$
$$\vec{\alpha} = \mathbf{K}^{-1} \vec{P}^{\text{ref}}$$



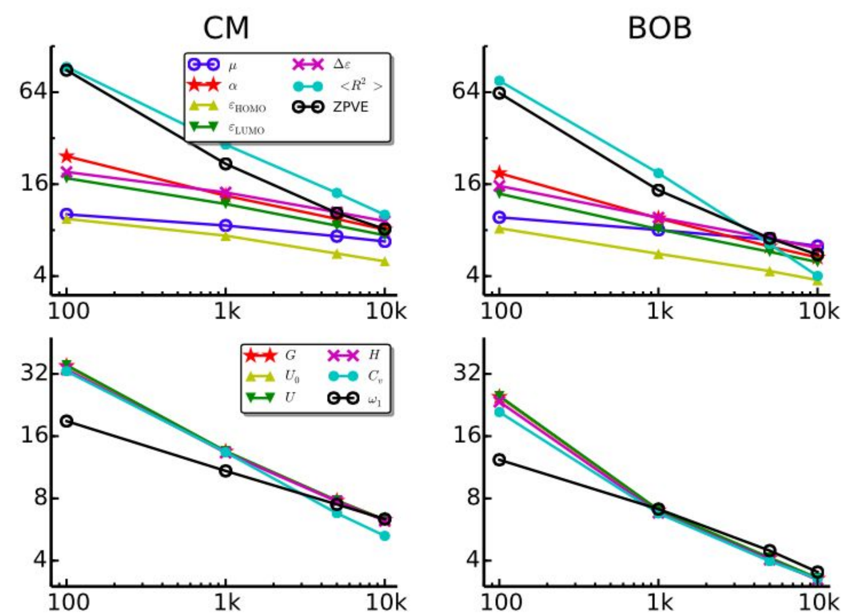
$$E_q = \langle \Psi_q | \hat{H} | \Psi_q \rangle$$
$$O_q = \langle \Psi_q | \hat{O} | \Psi_q \rangle$$

K ~ Ψ

α ~ Ô

Ramakrishnan, OAvL, *CHIMIA* (2015)

$$P^{\mathrm{est}}(\mathbf{M}) = \sum_i \alpha_i k(\mathbf{M}, \mathbf{M}_i)$$
$$\vec{\alpha} = \mathbf{K}^{-1} \vec{P}^{\mathrm{ref}}$$



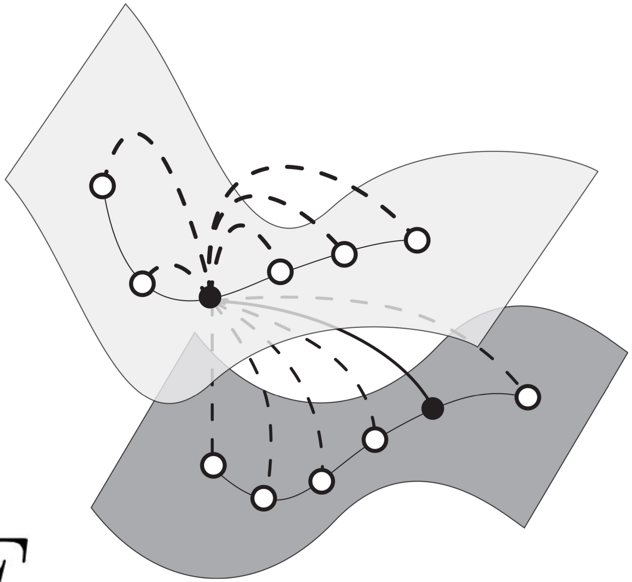
Ramakrishnan, OAvL, *CHIMIA* (2015)

Final remarks

correlations (inductive) vs. laws (deductive)

$$H(\{Z_I, \mathbf{R}_I\})\Psi(\mathbf{r}) = E\Psi(\mathbf{r})$$

Erwin



Chang, von Lilienfeld, *CHIMIA* (2014)

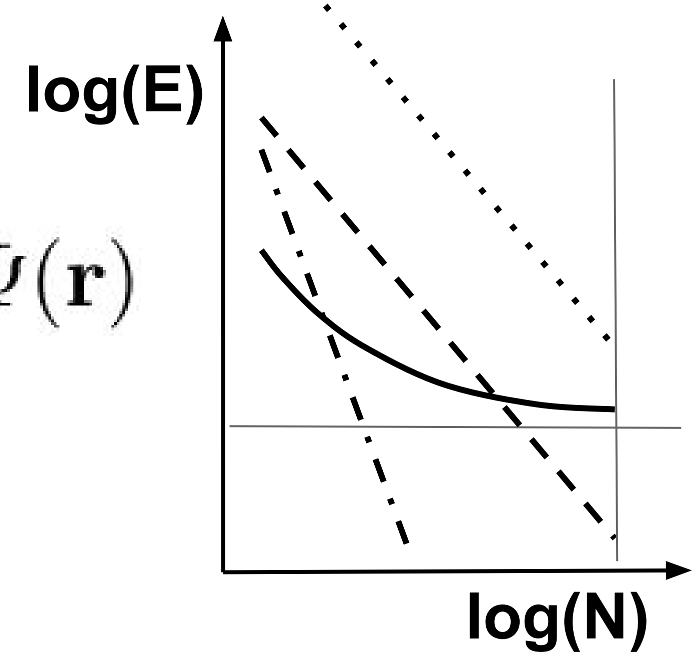
$$H(\{Z_I, \mathbf{R}_I\}) \xmapsto{\Psi} E$$

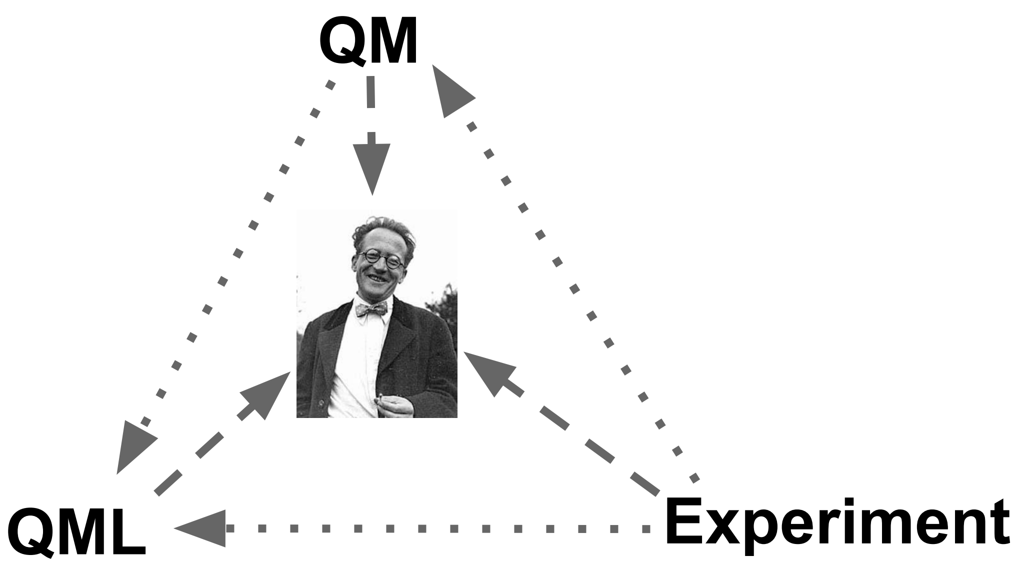
$$\{Z_I, \mathbf{R}_I\} \xmapsto{\mathrm{ML}} E$$

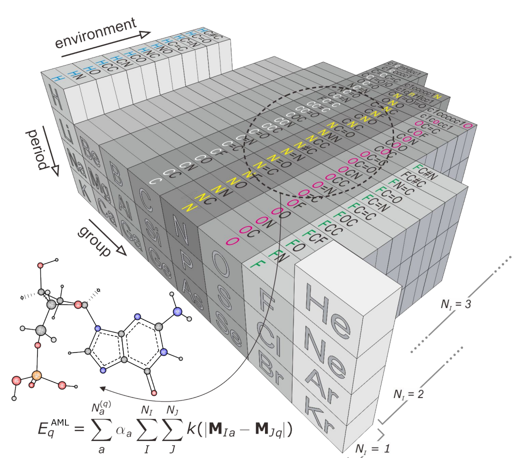
Rupp, Tkatchenko, Muller, von Lilienfeld, *Phys Rev Lett* (2012)

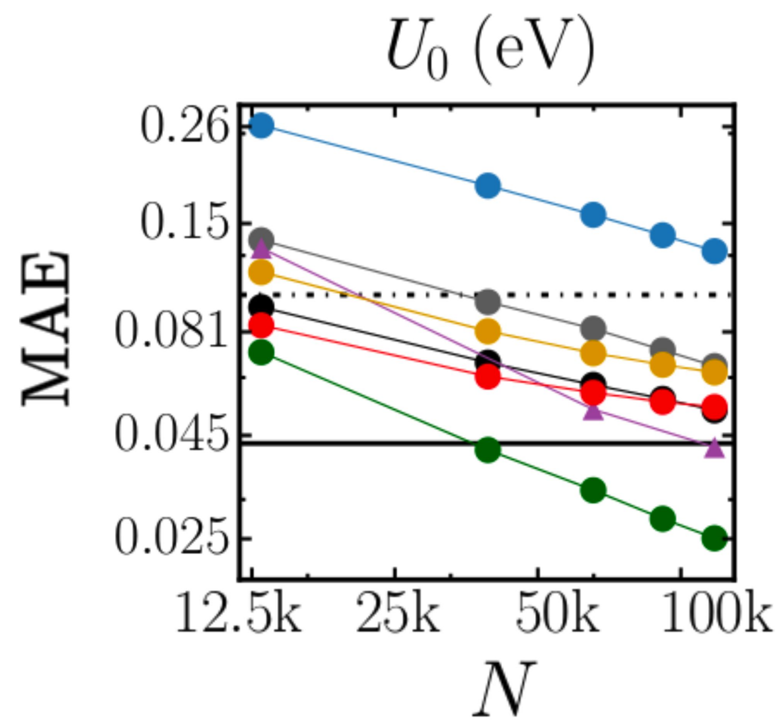
Final remarks

$$H(\{Z_I, \mathbf{R}_I\})\Psi(\mathbf{r}) \quad = \quad E\Psi(\mathbf{r})$$

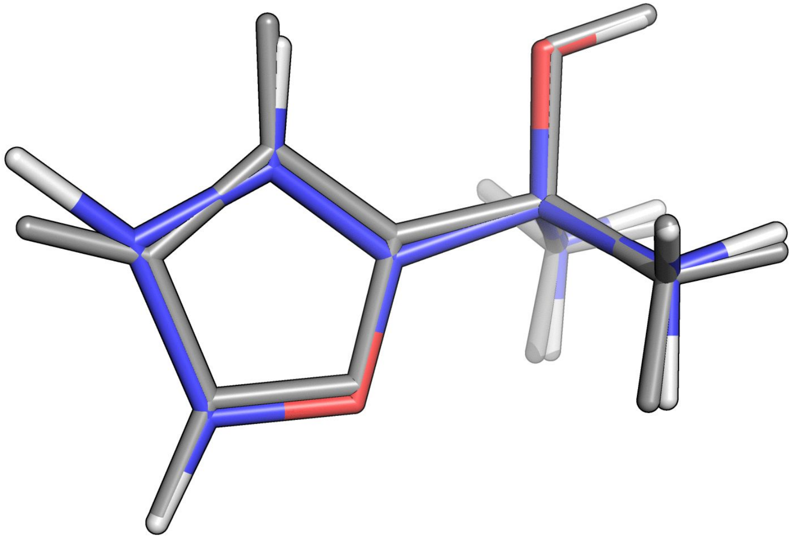


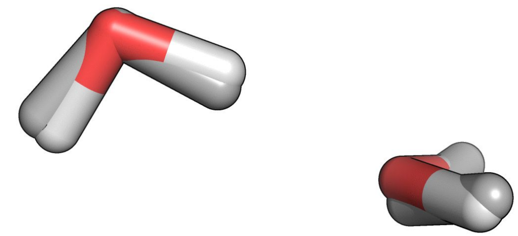




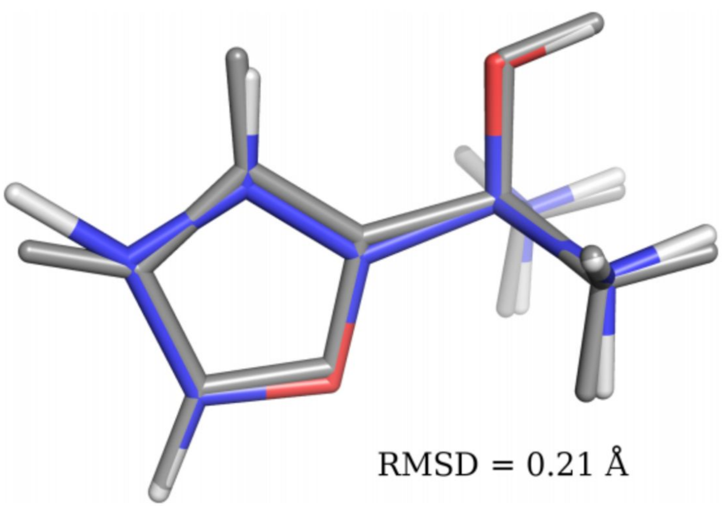


Outlook: Forces

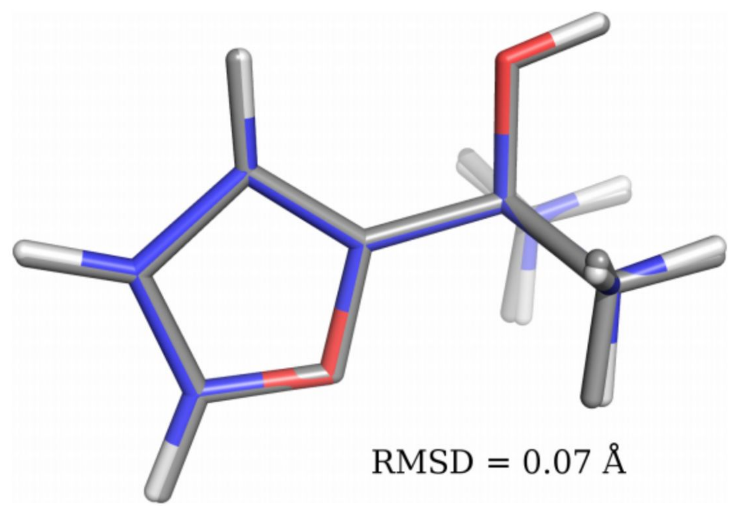




Initial



Converged



Christensen et al, in prep (2017)

Background of group members
- Computer Science
- Mathematics
- Physics (atomistic/molecular/solid)
- Chemistry (physical/computational)

WE WANT YOU!

Multiple openings for postdocs and PhD students

www.chemie.unibas.ch/~anatole