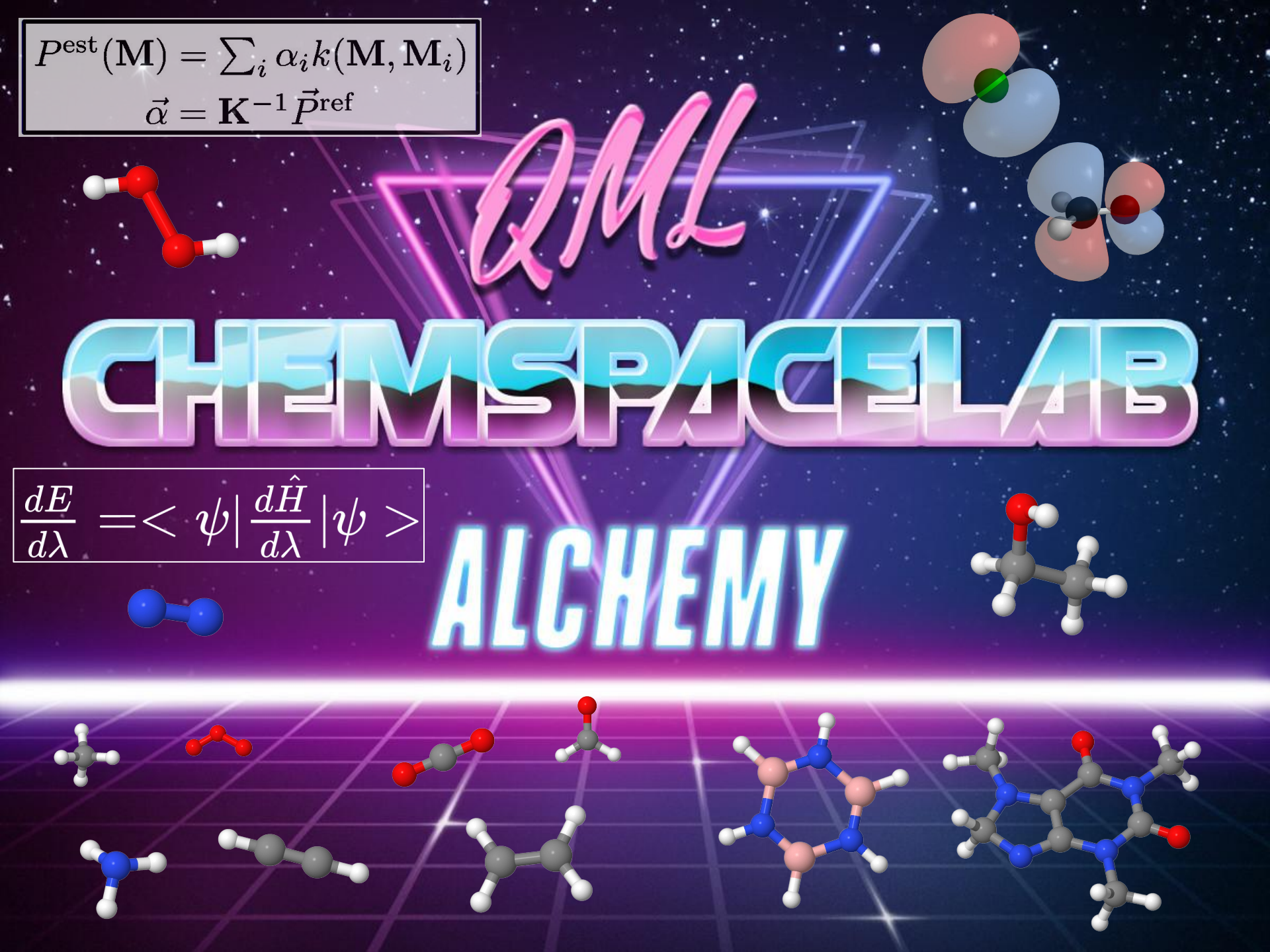
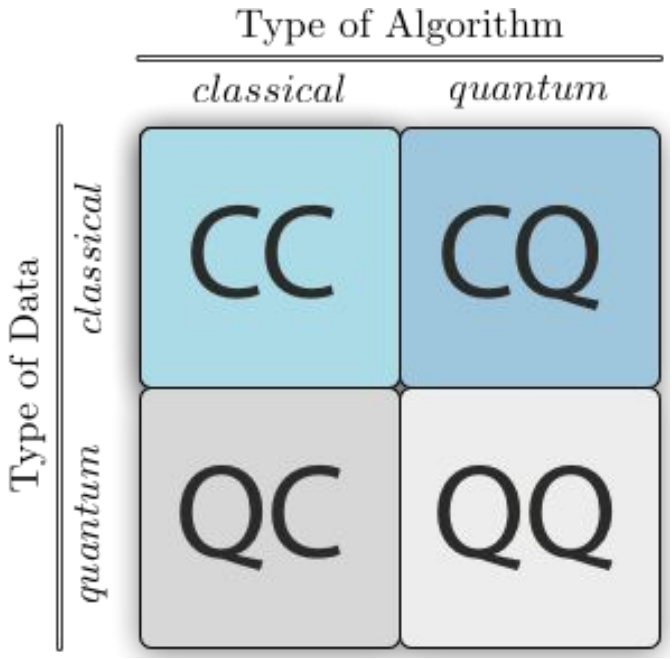

$P^{\text{est}}(\mathbf{M}) = \sum_i \alpha_i k(\mathbf{M}, \mathbf{M}_i)$
$\vec{\alpha} = \mathbf{K}^{-1}\vec{P}^{\text{ref}}$
QML
CHEMSPACELAB
$\frac{dE}{d\lambda} = < \psi | \frac{d\hat{H}}{d\lambda} | \psi >$
ALCHEMY

Quantum Machine Learning



wikipedia.org/quantum_machine_learning

QML in the QC sense and in analogy to conventional quantum simulations

- Quantum Chemistry/Mechanics Calculations
- Quantum Monte Carlo (QMC)
- Quantum Molecular Dynamics (QMD) aka AIMD/FPMD
- …
- Solving quantum problems with machine learning
  Theorists know everything!
  S. Kalinin, University of Tokyo, Feb 9 2019

AND also to clearly mark the distinction to conventional QSPR/QSAR approaches in XYZ-informatics approaches …

Machine Learning = Correlation

x

"panda"
57.7% confidence

Elephant in the room ...

Published as a conference paper at ICLR 2015

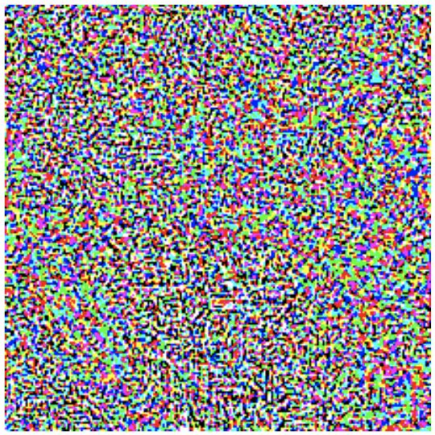
EXPLAINING AND HARNESSING ADVERSARIAL EXAMPLES

Ian J. Goodfellow, Jonathon Shlens & Christian Szegedy
Google Inc., Mountain View, CA

https://arxiv.org/abs/1412.6572

Machine Learning = Correlation

$+ .007 \times$



$=$

$\boldsymbol{x}$

"panda"
57.7% confidence

$\text{sign}(\nabla_{\boldsymbol{x}} J(\boldsymbol{\theta}, \boldsymbol{x}, y))$

"nematode"
8.2% confidence

$\boldsymbol{x} + \epsilon \text{sign}(\nabla_{\boldsymbol{x}} J(\boldsymbol{\theta}, \boldsymbol{x}, y))$

Published as a conference paper at ICLR 2015

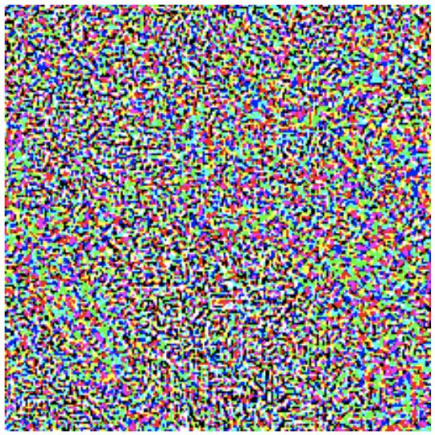
EXPLAINING AND HARNESSING ADVERSARIAL EXAMPLES

Ian J. Goodfellow, Jonathon Shlens & Christian Szegedy
Google Inc., Mountain View, CA

https://arxiv.org/abs/1412.6572

Machine Learning = Correlation

$+ .007 \times$



$=$

$\boldsymbol{x}$

"panda"
57.7% confidence

$\text{sign}(\nabla_{\boldsymbol{x}} J(\boldsymbol{\theta}, \boldsymbol{x}, y))$

"nematode"
8.2% confidence

$\boldsymbol{x} + \epsilon \text{sign}(\nabla_{\boldsymbol{x}} J(\boldsymbol{\theta}, \boldsymbol{x}, y))$

"gibbon"
99.3 % confidence

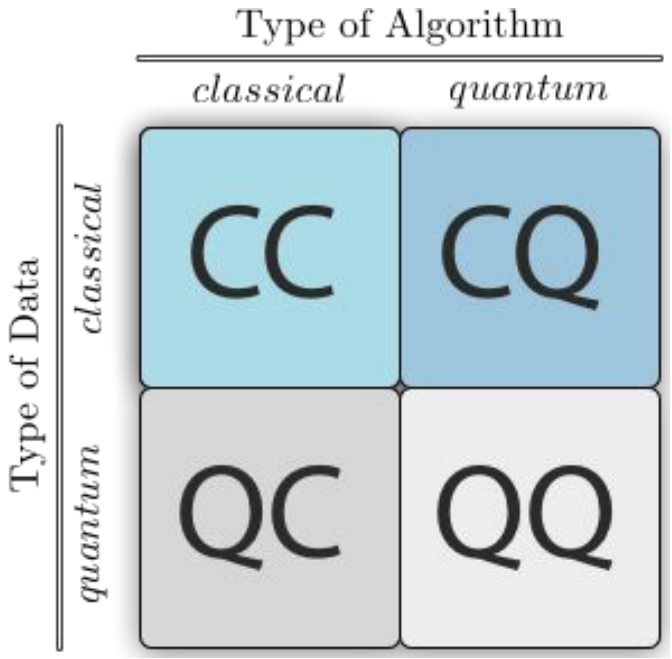
Published as a conference paper at ICLR 2015

EXPLAINING AND HARNESSING ADVERSARIAL EXAMPLES

Ian J. Goodfellow, Jonathon Shlens & Christian Szegedy
Google Inc., Mountain View, CA

https://arxiv.org/abs/1412.6572

Quantum Machine Learning



wikipedia.org/quantum_machine_learning

QML in the QC sense and in analogy to conventional quantum simulations

- Quantum Chemistry/Mechanics Calculations
- Quantum Monte Carlo (QMC)
- Quantum Molecular Dynamics (QMD) aka AIMD/FPMD
- …
- Solving quantum problems with machine learning
  Theorists know everything!
  S. Kalinin, University of Tokyo, Feb 9 2019

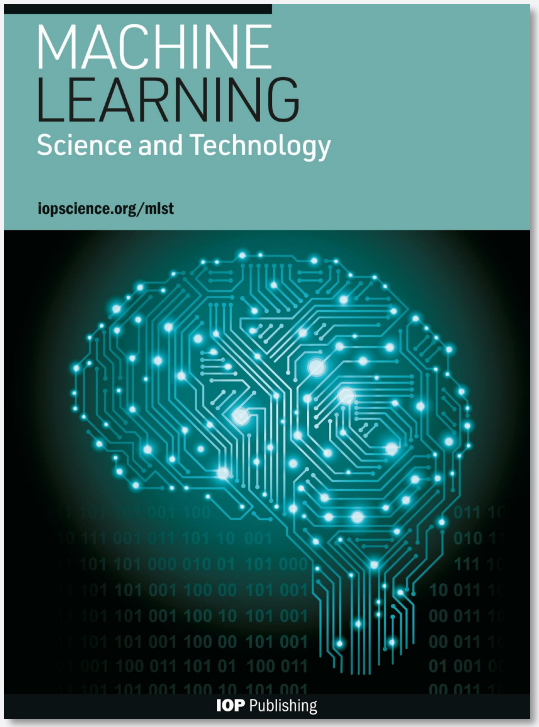
AND also to clearly mark the distinction to conventional QSPR/QSAR approaches in XYZ-informatics approaches …

Causal relationship: $\boldsymbol{M} \mapsto \boldsymbol{f}(\boldsymbol{M})$

MACHINE LEARNING
Science and Technology

NEW LAUNCH FOR 2019

OPEN ACCESS



A multidisciplinary, open access journal devoted to the application and development of machine learning for the sciences.

Why should you publish in *Machine Learning: Science and Technology*™ ?

- Multidisciplinary subject reach
- Rapid peer-review
- Include citable datasets and code
- Visibility through open access
- No APCs until July 2020
- Society-owned journal

Editor-in-Chief

Anatole von Lilienfeld
University of Basel, Switzerland

To find out more about the journal scope and how to submit your paper, please visit **iopscience.org/mlst** or e-mail **mlst@ioppublishing.org.**

Chemical space within QM?

→ 1st postulate of QM: System = wave-function

$$\hat{h}(\mathbf{r})\phi_i(\mathbf{r}) = \left(-\frac{\hbar^2}{2m}\nabla^2(\mathbf{r}) + v_{\mathrm{KS}}(\mathbf{r})\right)\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r})$$

$$v_{\mathrm{KS}}(\mathbf{r}) = v_{\mathrm{ext}}(\mathbf{r}) + \int d\mathbf{r}'\,\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + v_{xc}(\mathbf{r})$$

$$v_{\mathrm{ext}} = \sum_I \frac{Z_I}{|\mathbf{r}-\mathbf{R}_I|}$$

CCS ~ $O(4N_I + 1)$

$0 < N_I < 10^6$

Chemical space within QM?

→ 1st postulate of QM: System = wave-function

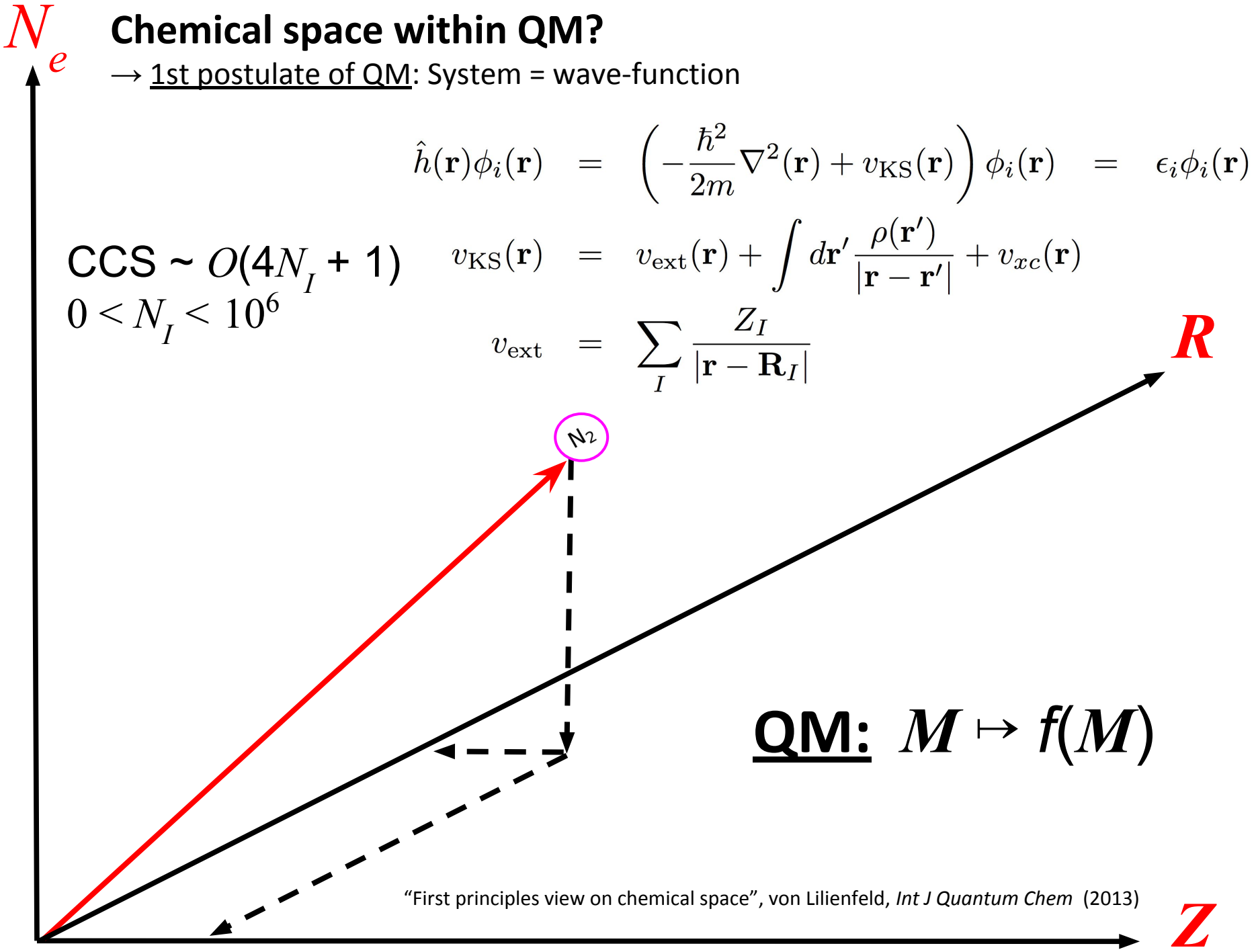
$$\hat{h}(\mathbf{r})\phi_i(\mathbf{r}) = \left(-\frac{\hbar^2}{2m}\nabla^2(\mathbf{r}) + v_{\mathrm{KS}}(\mathbf{r})\right)\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r})$$

$$v_{\mathrm{KS}}(\mathbf{r}) = v_{\mathrm{ext}}(\mathbf{r}) + \int d\mathbf{r}'\,\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + v_{xc}(\mathbf{r})$$

$$v_{\mathrm{ext}} = \sum_I \frac{Z_I}{|\mathbf{r}-\mathbf{R}_I|}$$

CCS ~ $O(4N_I + 1)$

$0 < N_I < 10^6$



QM: $M \mapsto f(M)$

"First principles view on chemical space", von Lilienfeld, *Int J Quantum Chem* (2013)

Error [Energy]

FF

SE

15 kcal/mol

HF

5 kcal/mol

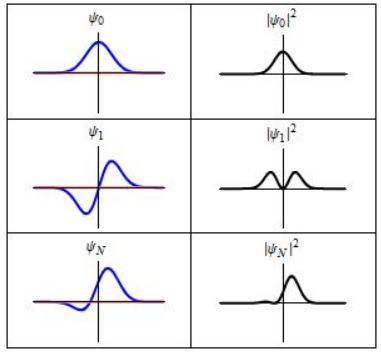
MP2

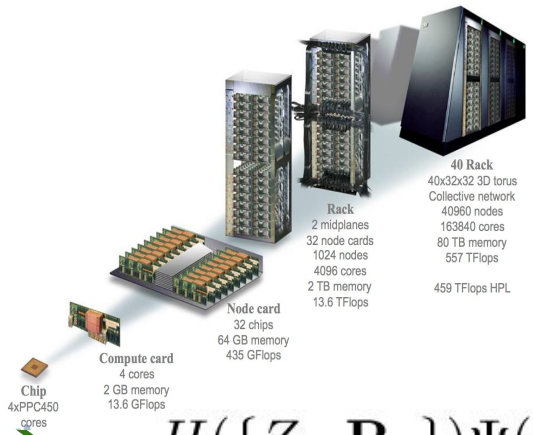
1 kcal/mol

QMC
CCSD(T)

ms min h d yr

Cost [CPU t]/compound





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

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle$$

Schrödinger
1933

Error [Energy]

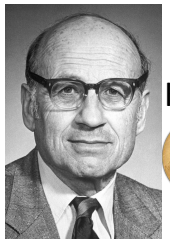
Model chemistry

FF

SE

15 kcal/mol

HF



Kohn

DFT

5 kcal/mol

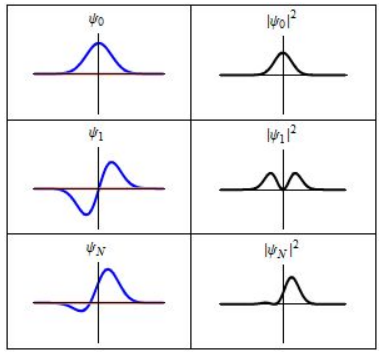
MP2

1 kcal/mol

QMC
CCSD(T)

1998

Pople



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

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle$$

Schrödinger
1933

ms min h d yr

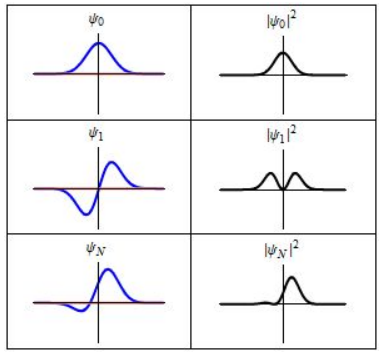
Cost [CPU t]/compound

Error [Energy]

Model chemistry

1998

Pople

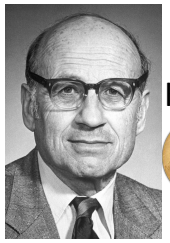


FF

SE

15 kcal/mol

HF



Kohn

DFT

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

5 kcal/mol

MP2

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle$$

Schrödinger
1933

1 kcal/mol

QMC

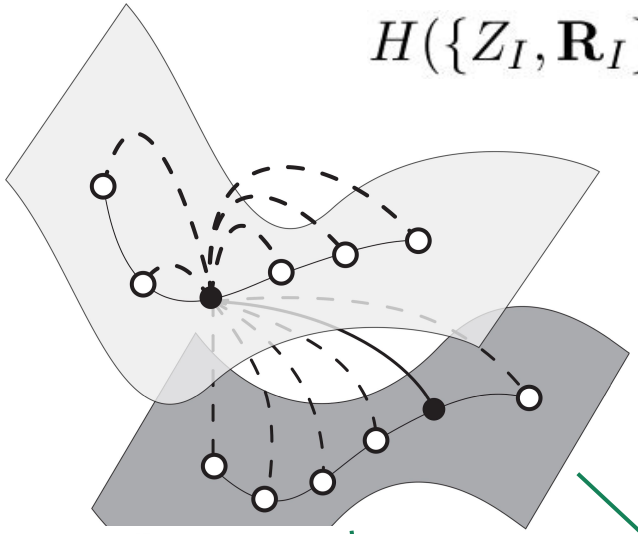
CCSD(T)

Alchemy/QML

ms min h d yr

Cost [CPU t]/compound

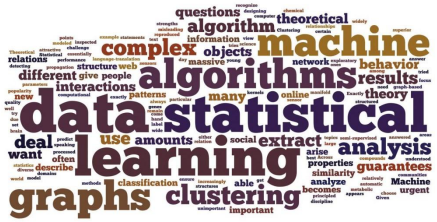
$$H(\{Z_I, \mathbf{R}_I\})\Psi(\mathbf{r}) = E\Psi(\mathbf{r})$$



Schrödinger

Alchemy

Fukui



QML

correlation (interpolate)

Vapnik

$$\{Z_I, \mathbf{R}_I\} \overset{H\Psi}{\longmapsto} E$$

supervised learning

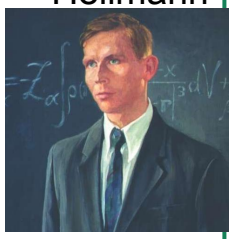
$$\{Z_I, \mathbf{R}_I\} \overset{\text{ML}}{\longmapsto} E$$

perturbation (extrapolate)

$$E(H(\lambda)) = E(H_i + \lambda(H_f - H_i))$$

$$\frac{\partial E[H]}{\partial \lambda} = \left\langle \Psi \left| \frac{\partial H(\lambda)}{\partial \lambda} \right| \Psi \right\rangle$$

Hellmann



$$\frac{\partial E[H]}{\partial R_{Ix}} = \left\langle \Psi \left| \frac{\partial H}{\partial R_{Ix}} \right| \Psi \right\rangle$$

$$\frac{\partial E[H]}{\partial Z_I} = \left\langle \Psi \left| \frac{\partial H}{\partial Z_I} \right| \Psi \right\rangle$$

Feynman

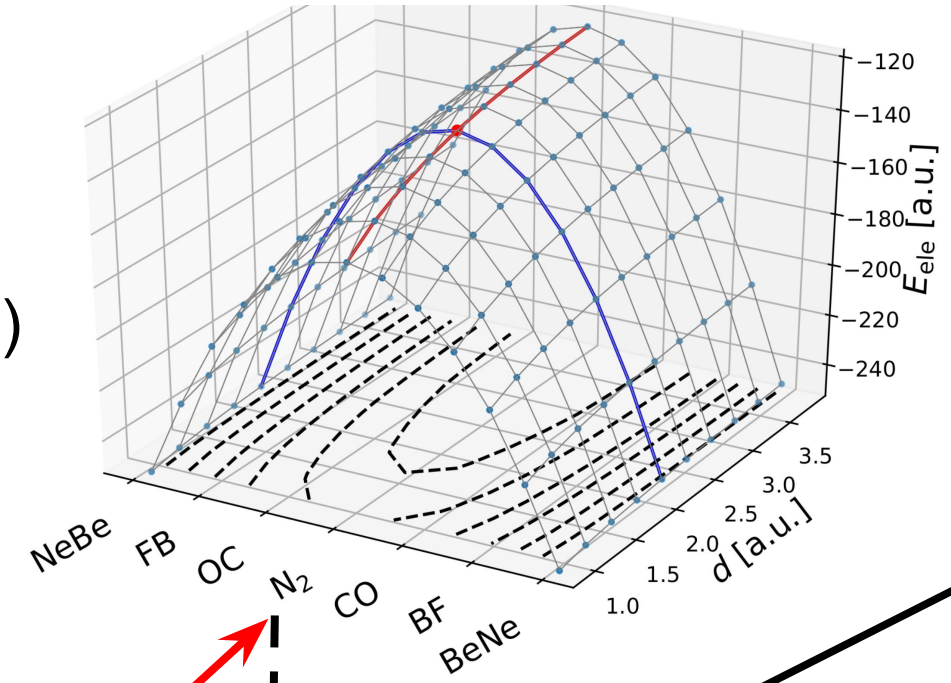
Vapnik, *The Nature of Statistical Learning Theory,* Springer (1995)

Fias, Chang, von Lilienfeld, *J Phys Chem Lett* (2019), https://arxiv.org/abs/1809.03302

N_e

CCS ~ $O(4N_I + 1)$

$0 < N_I < 10^6$



R

Stijn Fias

von Rudorff, von Lilienfeld, https://arxiv.org/abs/1809.01647

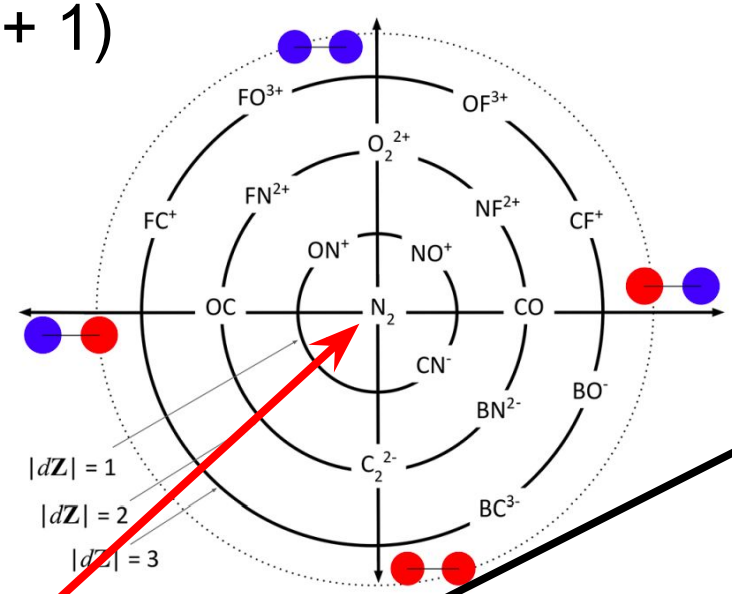
L

Fias, Chang, von Lilienfeld, *J Phys Chem Lett* (2019), https://arxiv.org/abs/1809.03302

$$\mathbf{H} = \begin{array}{c} \begin{matrix} \Delta Z_1 & \Delta Z_2 & \Delta \mathbf{R} & \Delta N \end{matrix} \\ \begin{pmatrix} -3.126 & 0.139 & 0.121 & -0.575 \\ 0.139 & -3.126 & 0.121 & -0.575 \\ 0.121 & 0.121 & -9.477 & -0.121 \\ -0.575 & -0.575 & -0.121 & 0.139 \end{pmatrix} \end{array} \begin{matrix} \Delta Z_1 \\ \Delta Z_2 \\ \Delta \mathbf{R} \\ \Delta N \end{matrix}$$

CCS ~ $O(4N_I + 1)$

$0 < N_I < 10^6$



N_e

R

Z

$$\begin{array}{c} \begin{matrix} \mathrm{ANM}_1 & \mathrm{ANM}_2 & \mathrm{ANM}_3 & \mathrm{ANM}_4 \end{matrix} \\ \begin{pmatrix} -0.018 & 0.168 & 0.707 & 0.687 \\ -0.018 & 0.168 & -0.707 & 0.687 \\ 1. & 0.016 & -0. & 0.022 \\ 0.01 & -0.971 & -0. & 0.238 \\ -9.48 & 0.34 & -3.27 & -3.18 \end{pmatrix} \end{array} \begin{matrix} \Delta Z_1 \\ \Delta Z_2 \\ \Delta \mathbf{R} \\ \Delta N \\ \epsilon \end{matrix}$$

von Rudorff, von Lilienfeld, https://arxiv.org/abs/1809.01647

$$\mathbf{x} = (Z_1, Z_2, \cdots, Z_M, \mathbf{R}_1, \mathbf{R}_2, \cdots, \mathbf{R}_M, N)$$

$$E(\mathbf{x}^t) = E(\mathbf{x}_0) + \mathbf{g}\,d\mathbf{x} + \frac{1}{2}\,d\mathbf{x}^{\mathrm{T}}\mathbf{H}\,d\mathbf{x} + \cdots$$

$$\mathbf{x}^t = \mathbf{x}_0 + d\mathbf{x}$$

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 E_0}{\partial Z_I \partial Z_J} & \frac{\partial^2 E_0}{\partial Z_I \partial \mathbf{R}_J} & \frac{\partial^2 E_0}{\partial Z_I \partial N} \\ \frac{\partial^2 E_0}{\partial \mathbf{R}_I \partial Z_J} & \frac{\partial^2 E_0}{\partial \mathbf{R}_I \partial \mathbf{R}_J} & \frac{\partial^2 E_0}{\partial \mathbf{R}_I \partial N} \\ \frac{\partial^2 E_0}{\partial N \partial Z_J} & \frac{\partial^2 E_0}{\partial N \partial \mathbf{R}_J} & \frac{\partial^2 E_0}{\partial N^2} \end{bmatrix}$$

1. Alchemical hardness
2. Alchemical force
3. Conventional Hessian
4. Alchemical Fukui function
5. Nuclear Fukui function
6. Electronic hardness

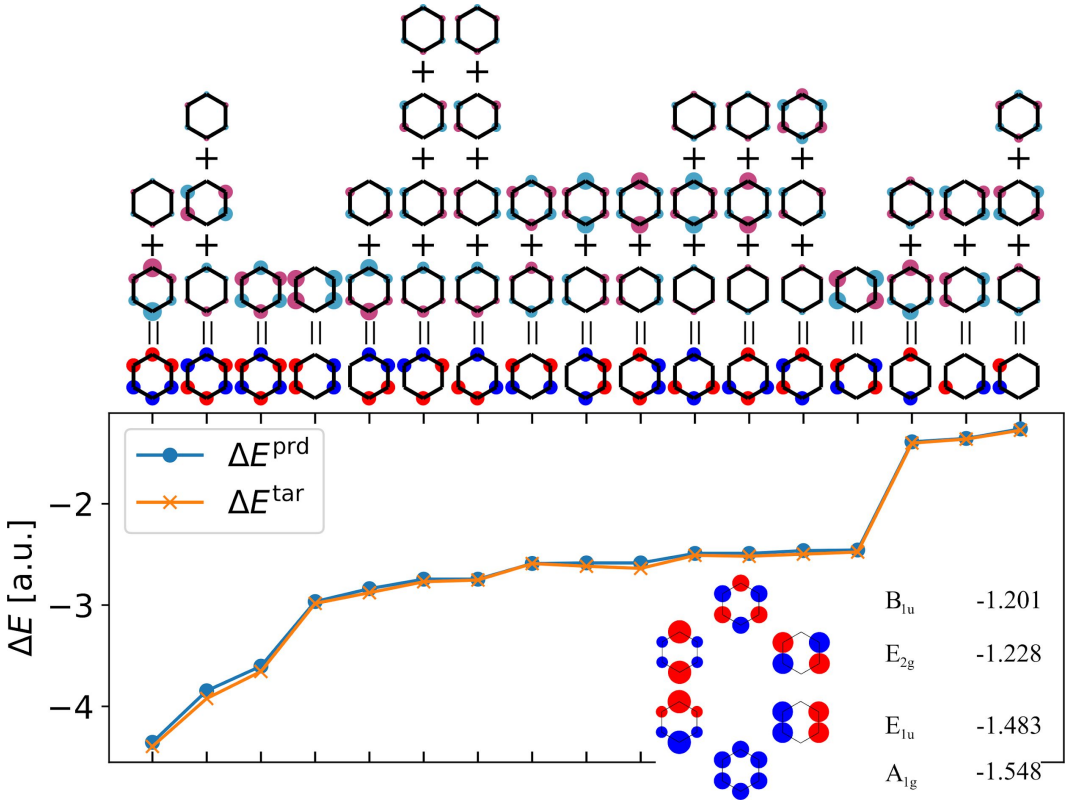
$$\mathbf{HQ} = \mathbf{Q}\boldsymbol{\Upsilon}$$

$$E(\mathbf{x}^t) \approx E(\mathbf{x}_0) + \mathbf{g}\,d\mathbf{x} + \frac{1}{2}\mathbf{c}^{\mathrm{T}}\boldsymbol{\Upsilon}\mathbf{c}$$

$$\mathbf{c} = \mathbf{Q}\,d\mathbf{x}$$

Stijn Fias

Fias, Chang, von Lilienfeld, *J Phys Chem Lett* (2019), https://arxiv.org/abs/1809.03302



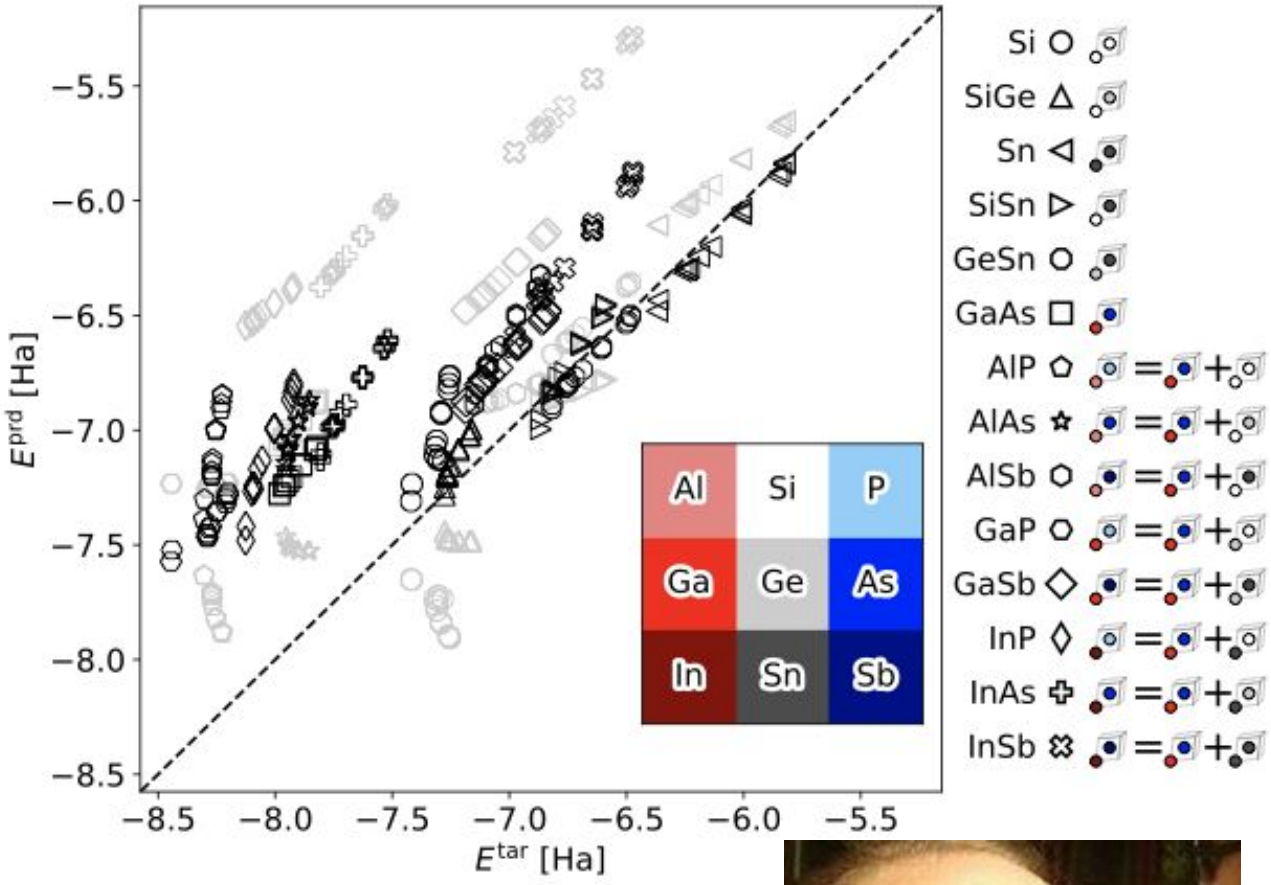
$$\mathbf{HQ} = \mathbf{Q}\boldsymbol{\Upsilon}$$

$$E(\mathbf{x}^t) \approx E(\mathbf{x}_0) + \mathbf{g}\,\mathrm{d}\mathbf{x} + \frac{1}{2}\mathbf{c}^{\mathrm{T}}\boldsymbol{\Upsilon}\mathbf{c}$$

$$\mathbf{c} = \mathbf{Q}\,\mathrm{d}\mathbf{x}$$

Stijn Fias

Fias, Chang, von Lilienfeld, *J Phys Chem Lett* (2019), https://arxiv.org/abs/1809.03302



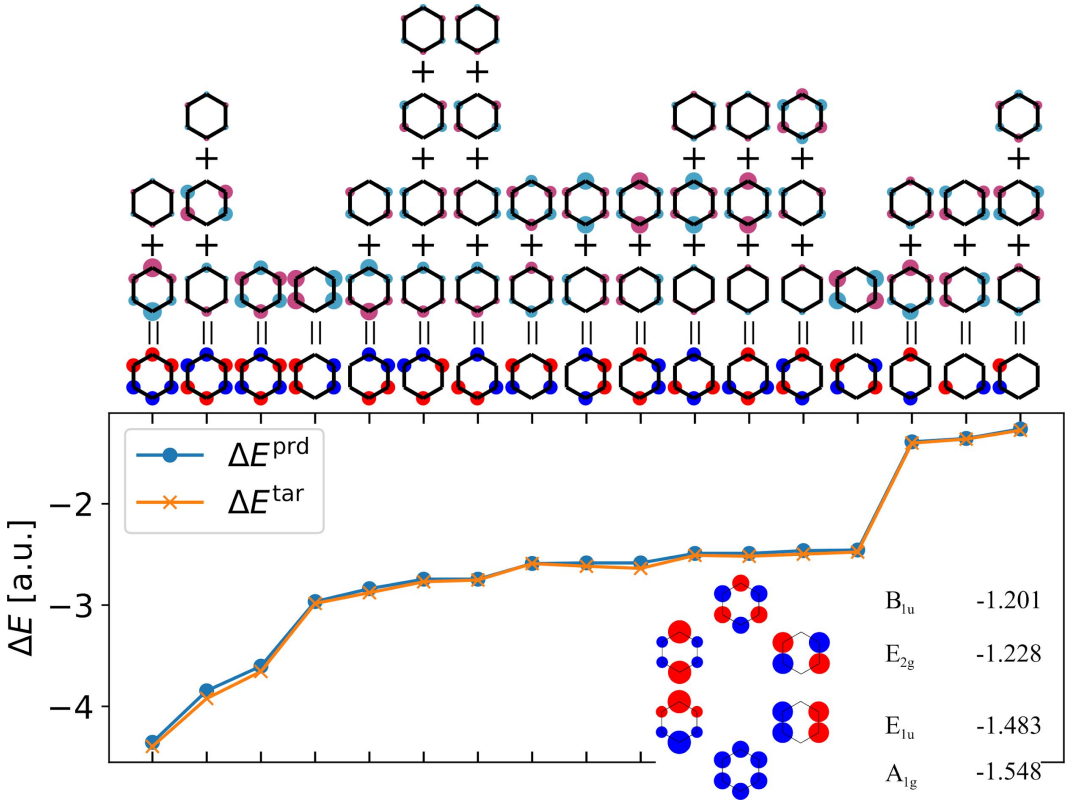
$$\mathbf{HQ} = \mathbf{Q}\boldsymbol{\Upsilon}$$

$$E(\mathbf{x}^t) \approx E(\mathbf{x}_0) + \mathbf{g}\,\mathrm{d}\mathbf{x} + \frac{1}{2}\mathbf{c}^{\mathrm{T}}\boldsymbol{\Upsilon}\mathbf{c}$$

$$\mathbf{c} = \mathbf{Q}\,\mathrm{d}\mathbf{x}$$

Stijn Fias

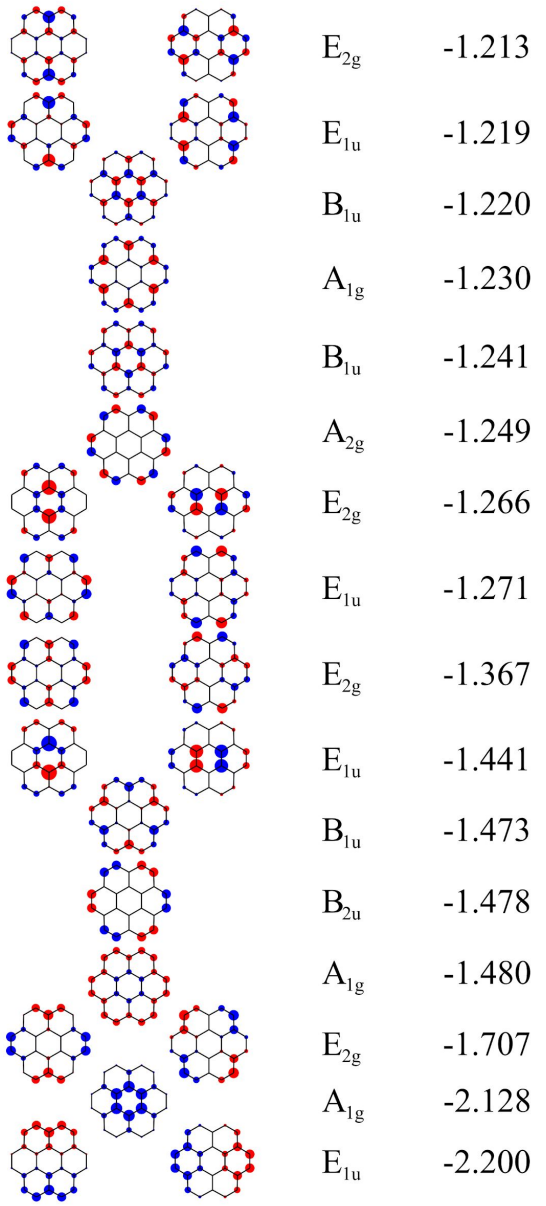
Fias, Chang, von Lilienfeld, *J Phys Chem Lett* (2019), https://arxiv.org/abs/1809.03302



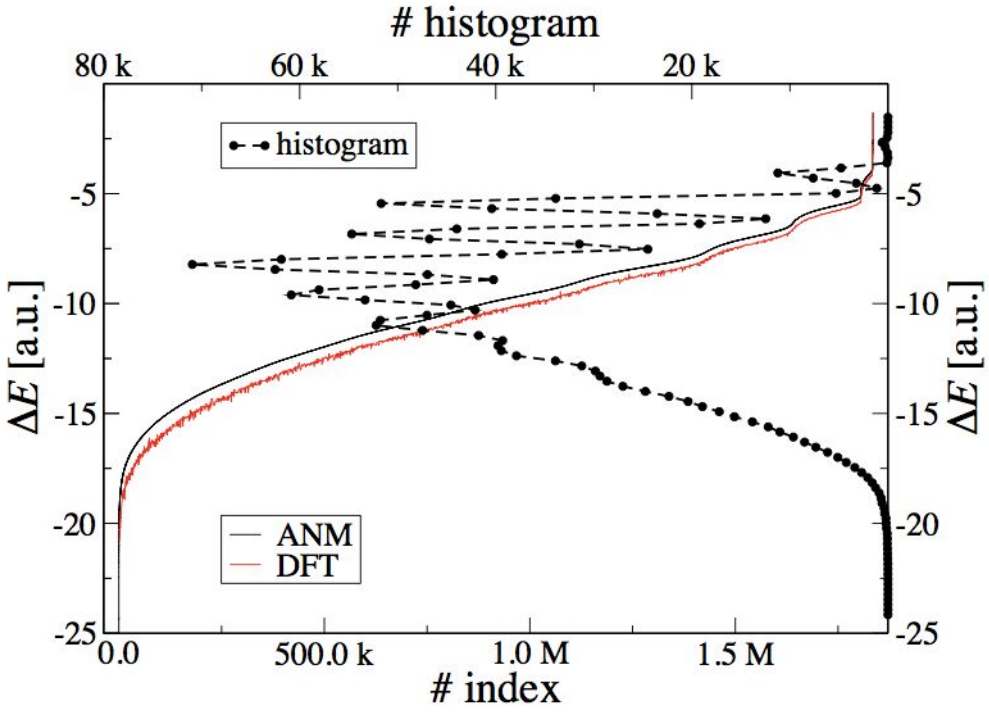
$$\mathbf{HQ} = \mathbf{Q}\boldsymbol{\Upsilon}$$

$$E(\mathbf{x}^t) \approx E(\mathbf{x}_0) + \mathbf{g}\,\mathbf{dx} + \frac{1}{2}\mathbf{c}^{\mathrm{T}}\boldsymbol{\Upsilon}\mathbf{c}$$

$$\mathbf{c} = \mathbf{Q}\,\mathbf{dx}$$



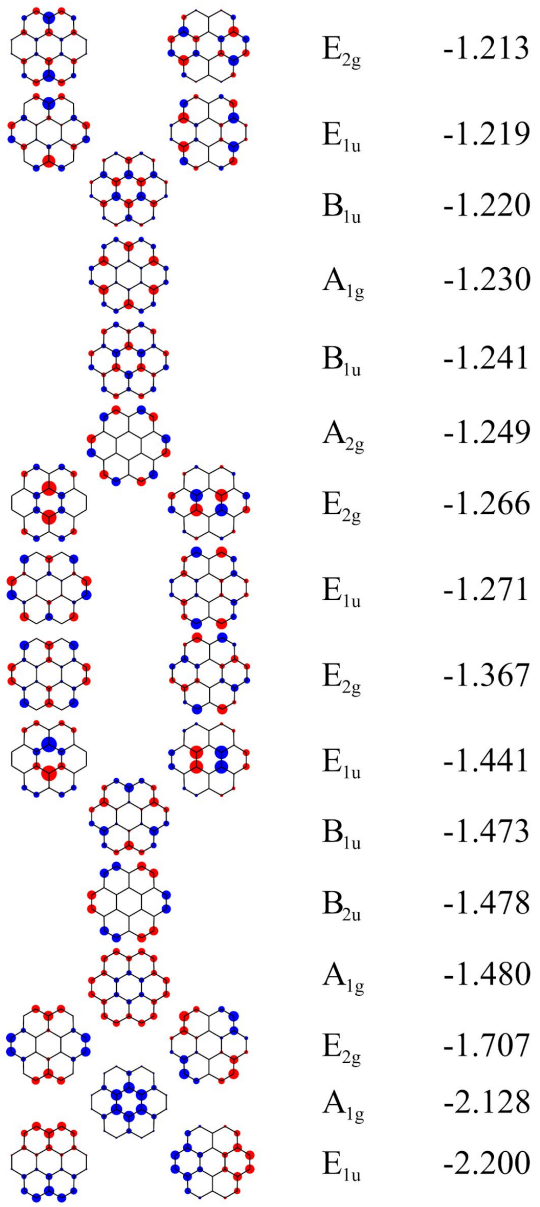
Fias, Chang, von Lilienfeld, *J Phys Chem Lett* (2019), https://arxiv.org/abs/1809.03302



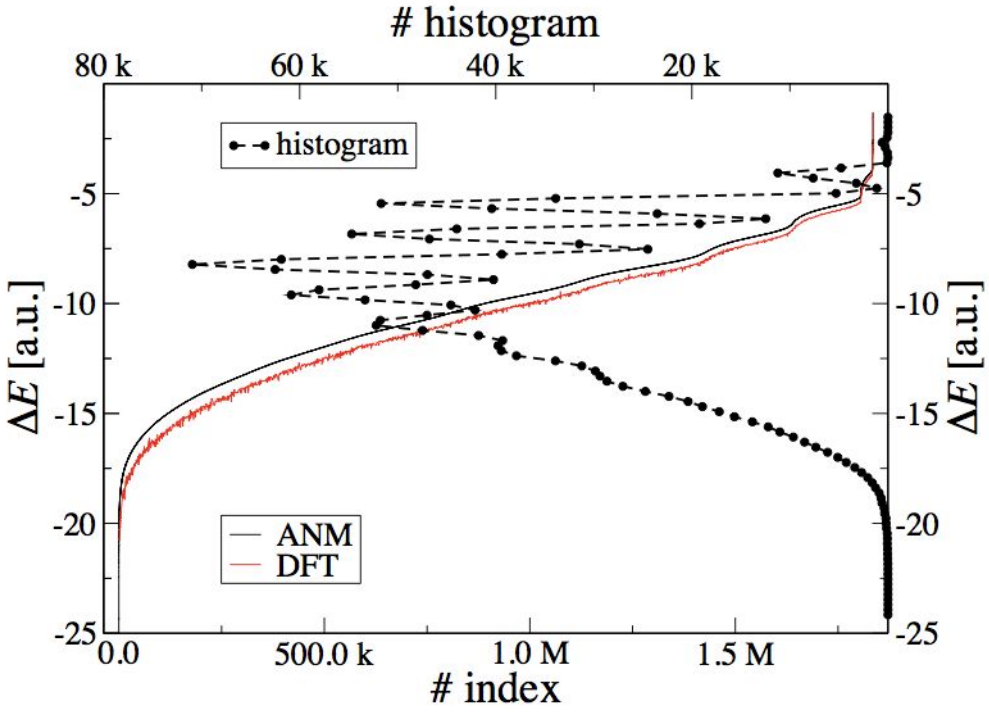
$$\mathbf{HQ} = \mathbf{Q}\boldsymbol{\Upsilon}$$

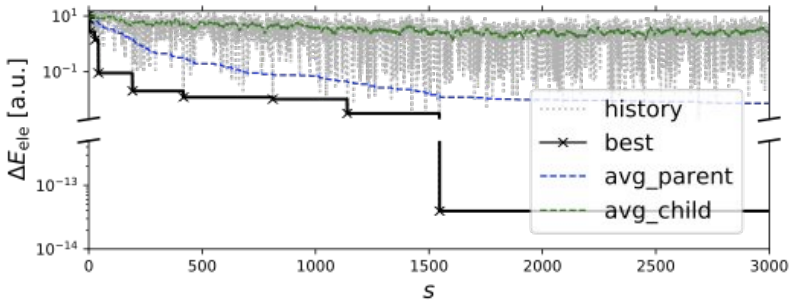
$$E(\mathbf{x}^t) \approx E(\mathbf{x}_0) + \mathbf{g}\,\mathrm{d}\mathbf{x} + \frac{1}{2}\mathbf{c}^{\mathrm{T}}\boldsymbol{\Upsilon}\mathbf{c}$$

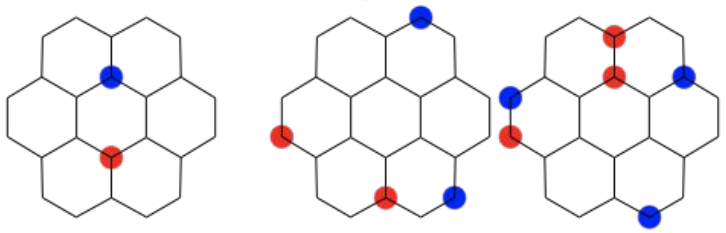
$$\mathbf{c} = \mathbf{Q}\,\mathrm{d}\mathbf{x}$$

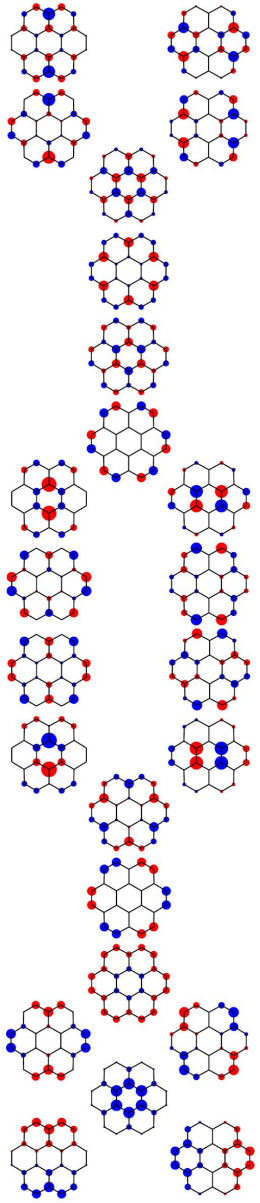


Fias, Chang, von Lilienfeld, *J Phys Chem Lett* (2019), https://arxiv.org/abs/1809.03302









E_{2g}	-1.213
E_{1u}	-1.219
B_{1u}	-1.220
A_{1g}	-1.230
B_{1u}	-1.241
A_{2g}	-1.249
E_{2g}	-1.266
E_{1u}	-1.271
E_{2g}	-1.367
E_{1u}	-1.441
B_{1u}	-1.473
B_{2u}	-1.478
A_{1g}	-1.480
E_{2g}	-1.707
A_{1g}	-2.128
E_{1u}	-2.200

Fias, Chang, von Lilienfeld, *J Phys Chem Lett* (2019), https://arxiv.org/abs/1809.03302

APDFT

APDFT is a software to allow quantum-chemistry calculations of many isoelectronic molecules at once rather than evaluating them one-by-one. This is achieved through *Alchemical Perturbation Density Functional Theory* [APDFT] where the change in external potential between molecules is treated as perturbation. This concept works just as fine for post-HF methods.

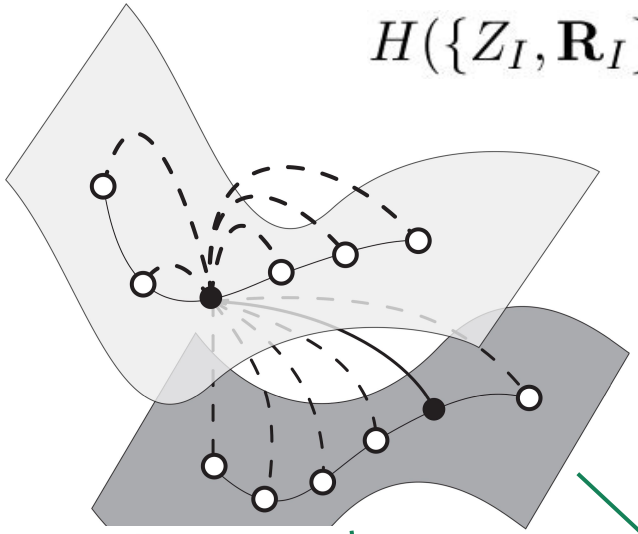
All gaussian basis sets from the EMSL Basis Set Exchange and a variety of methods (HF, LDA, PBE, PBE0, CCSD) are supported. APDFT does not reinvent the wheel but leverages other QM software in the background. Currently, we support Gaussian as only backend. For more details, please see Features and supported software.

Note

APDFT is under development. While the software is ready to use, the API may be subject to change.

[APDFT] https://arxiv.org/abs/1809.01647

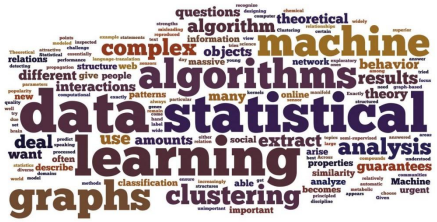
$$H(\{Z_I, \mathbf{R}_I\})\Psi(\mathbf{r}) = E\Psi(\mathbf{r})$$



Schrödinger

Alchemy

Fukui



QML

correlation (interpolate)

Vapnik

$$\{Z_I, \mathbf{R}_I\} \overset{H\Psi}{\longmapsto} E$$

supervised learning

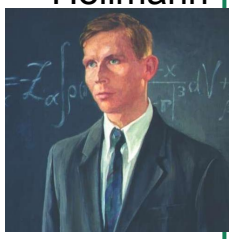
$$\{Z_I, \mathbf{R}_I\} \overset{\text{ML}}{\longmapsto} E$$

perturbation (extrapolate)

$$E(H(\lambda)) = E(H_i + \lambda(H_f - H_i))$$

$$\frac{\partial E[H]}{\partial \lambda} = \left\langle \Psi \left| \frac{\partial H(\lambda)}{\partial \lambda} \right| \Psi \right\rangle$$

Hellmann



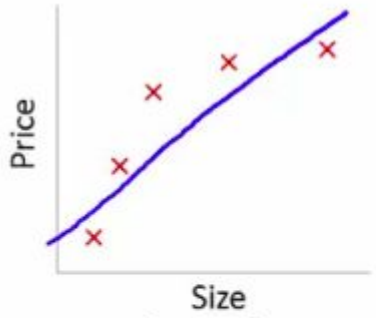
$$\frac{\partial E[H]}{\partial R_{Ix}} = \left\langle \Psi \left| \frac{\partial H}{\partial R_{Ix}} \right| \Psi \right\rangle$$

$$\frac{\partial E[H]}{\partial Z_I} = \left\langle \Psi \left| \frac{\partial H}{\partial Z_I} \right| \Psi \right\rangle$$

Feynman

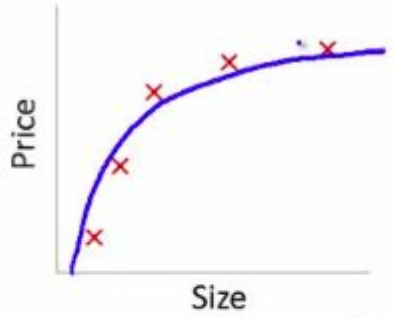
Vapnik, *The Nature of Statistical Learning Theory,* Springer (1995)

Overfitting?



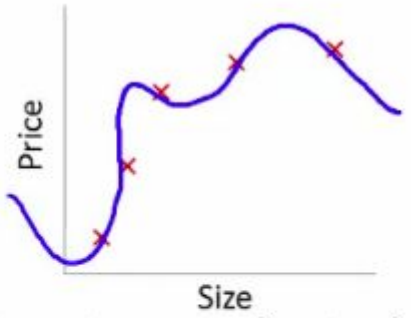
$\theta_0 + \theta_1 x$

High bias
(underfit)



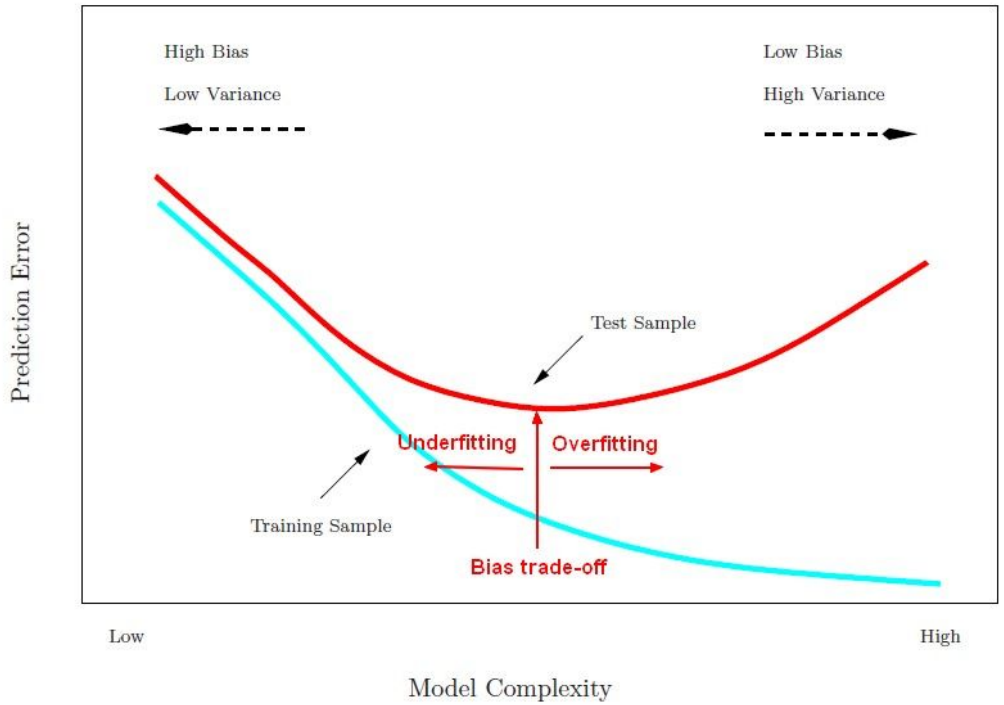
$\theta_0 + \theta_1 x + \theta_2 x^2$

"Just right"



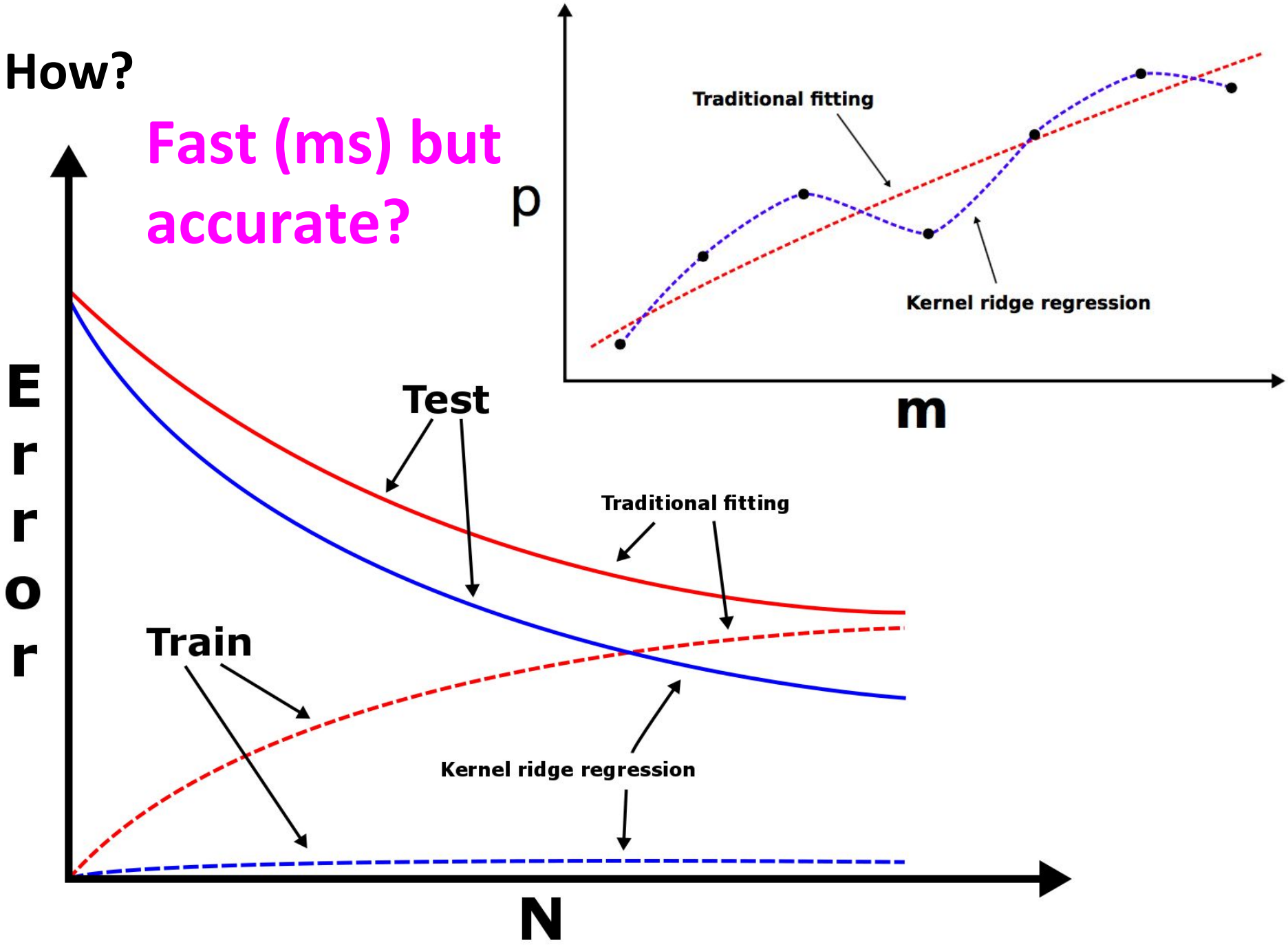
$\theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 + \theta_4 x^4$

High variance
(overfit)



How?

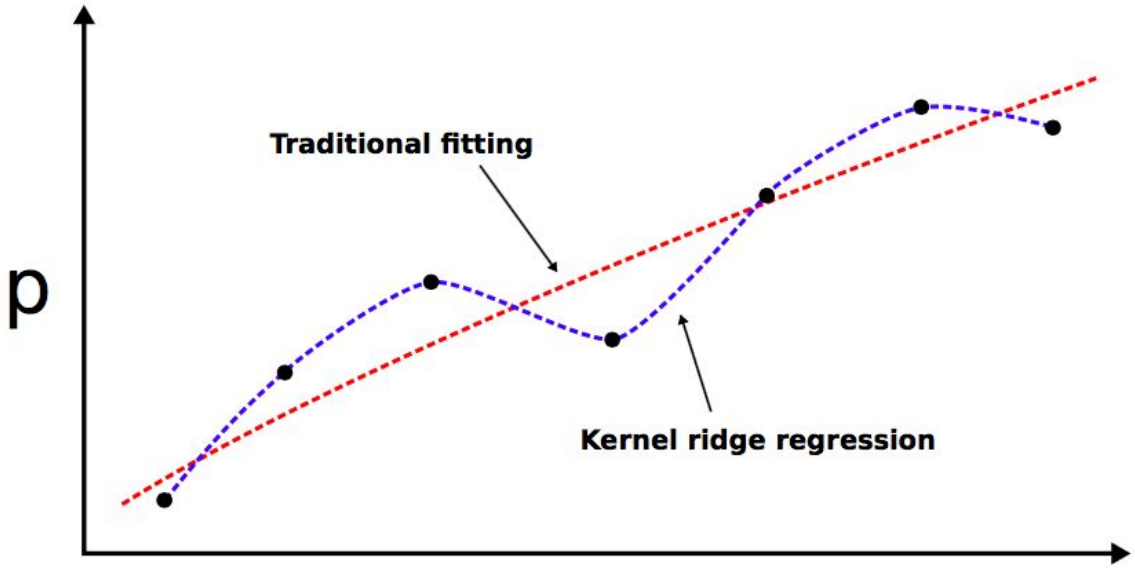
Fast (ms) but accurate?



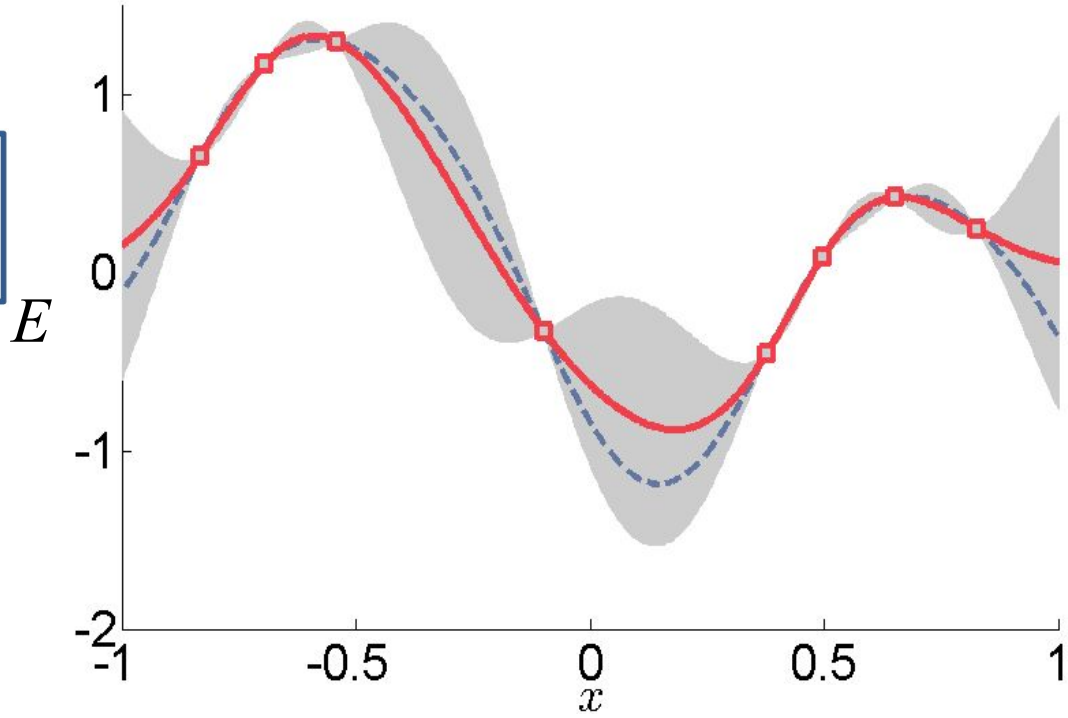
Vapnik, *The Nature of Statistical Learning Theory,* Springer (1995)

How?

Fast (ms) but accurate?

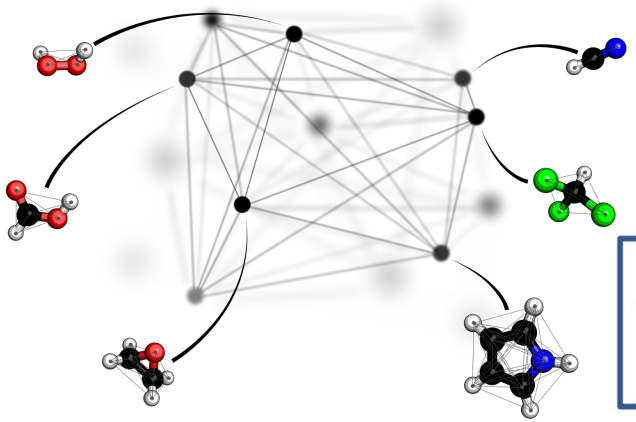


$$P^{\text{est}}(\mathbf{M}) = \sum_i \alpha_i k(\mathbf{M}, \mathbf{M}_i)$$
$$\vec{\alpha} = \mathbf{K}^{-1}\vec{P}^{\text{ref}}$$



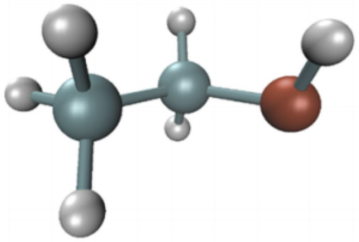
Vapnik, *The Nature of Statistical Learning Theory,* Springer (1995)

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

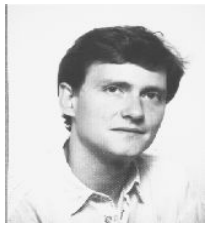


$$\mathrm{CM}_{IJ}^{(n)} = \frac{Z_I Z_J}{R_{IJ}^n}$$

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

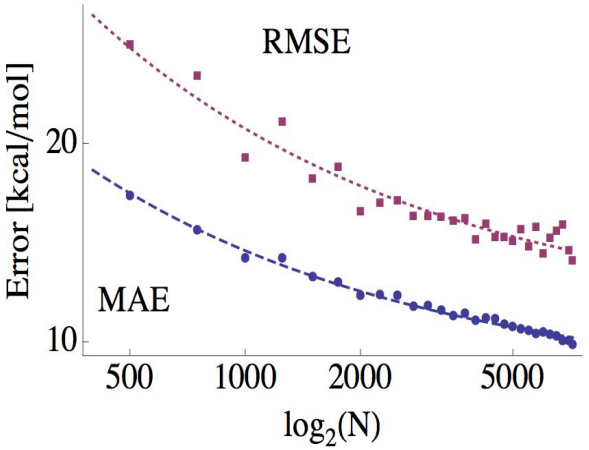


	O	C	C	H	H	H	H	H	H
O	O	OC	OC	OH	OH	OH	OH	OH	OH
C	OC	C	CC	CH	CH	CH	CH	CH	CH
C	OC	CC	C	CH	CH	CH	CH	CH	CH
H	OH	CH	CH	H	HH	HH	HH	HH	HH
H	OH	CH	CH	HH	H	HH	HH	HH	HH
H	OH	CH	CH	HH	HH	H	HH	HH	HH
H	OH	CH	CH	HH	HH	HH	H	HH	HH
H	OH	CH	CH	HH	HH	HH	HH	H	HH
H	OH	CH	CH	HH	HH	HH	HH	HH	H


Mueller


Tkatchenko

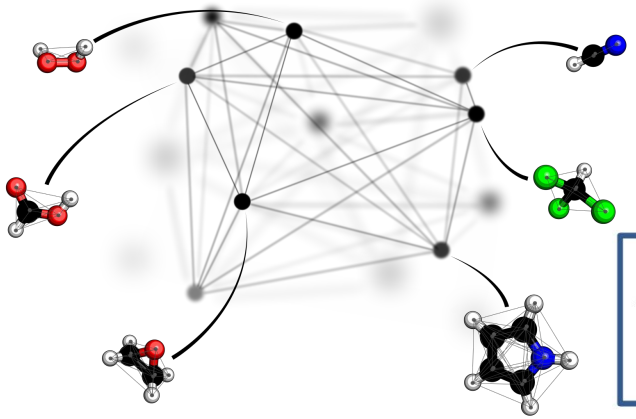

Rupp




Vapnik

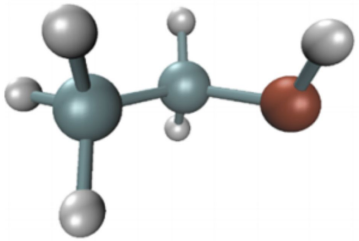
``Fast and accurate modeling of molecular atomization energies with machine learning'', Rupp et al, *Phys Rev Lett* (2012)

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



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

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

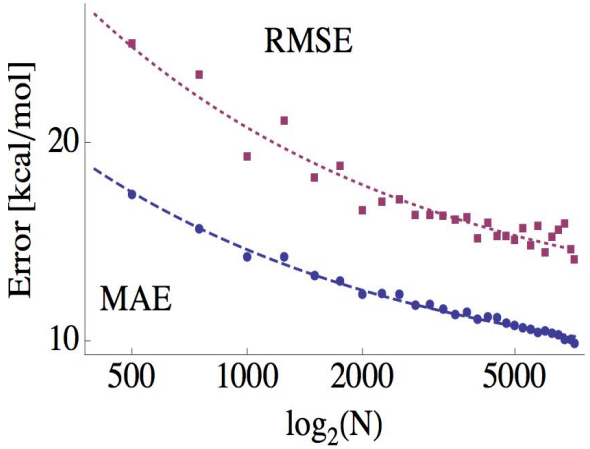


	O	C	C	H	H	H	H	H	H
O	O	OC	OC	OH	OH	OH	OH	OH	OH
C	OC	C	CC	CH	CH	CH	CH	CH	CH
C	OC	CC	C	CH	CH	CH	CH	CH	CH
H	OH	CH	CH	H	HH	HH	HH	HH	HH
H	OH	CH	CH	HH	H	HH	HH	HH	HH
H	OH	CH	CH	HH	HH	H	HH	HH	HH
H	OH	CH	CH	HH	HH	HH	H	HH	HH
H	OH	CH	CH	HH	HH	HH	HH	H	HH
H	OH	CH	CH	HH	HH	HH	HH	HH	H

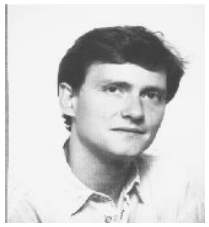
Error ~ a/N^b

→ log(Error) = log(a) - $b$log(N)

Cortes, Vapnik et al, Adv in NIPS (1992),
Muller et al, Neural Comp (1996) …



Vapnik



Mueller

Tkatchenko

Rupp

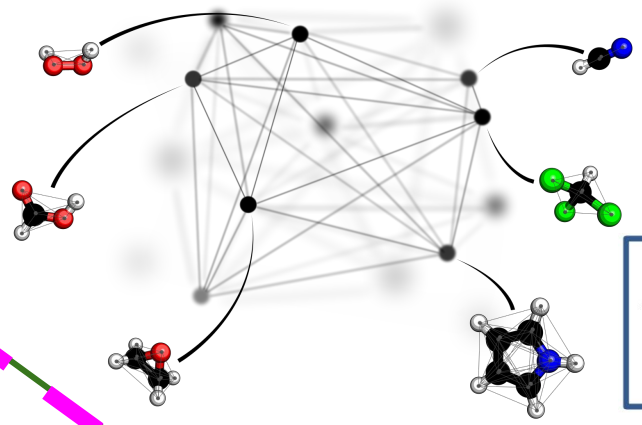
``Fast and accurate modeling of molecular atomization energies with machine learning'', Rupp et al, *Phys Rev Lett* (2012)

Error [Energy]

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

FF

SE



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

15 kcal/mol

HF

Error ~ a/N^b

→ log(Error) = log(a) - $b$log(N)

QML

DFT

5 kcal/mol

MP2

1 kcal/mol

QMC

CCSD(T)

Vapnik

10 100 1’000 10’000 100’000

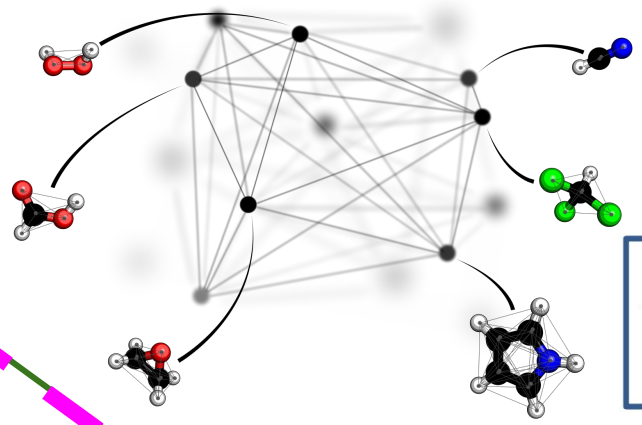
Training examples

Error [Energy]

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

FF

SE



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

15 kcal/mol

HF

Error ~ a/N^b

→ log(Error) = log(a) - $b$log(N)

DFT

5 kcal/mol

MP2

Vapnik

1 kcal/mol

QM: > 1000 seconds

QML: ~milli seconds

QMC

CCSD(T)

QML

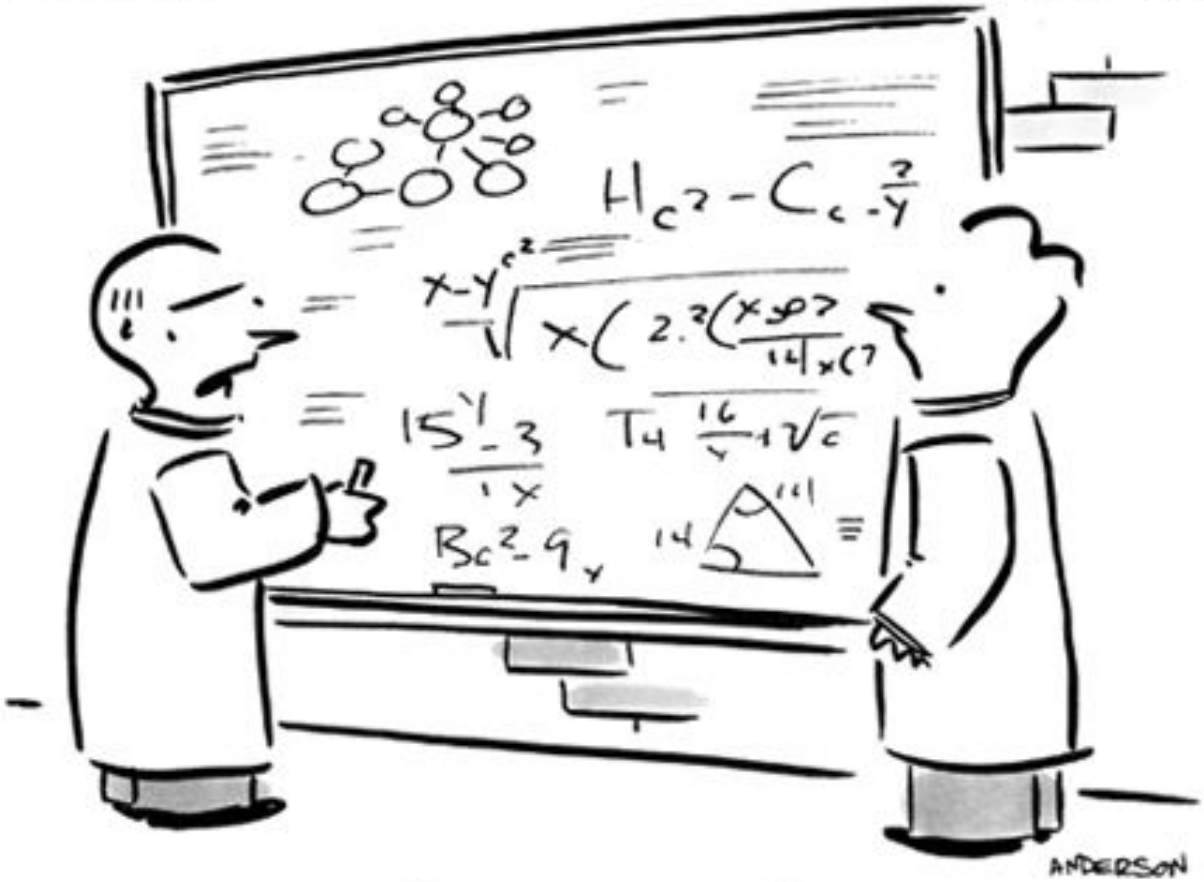
10 100 1’000 10’000 100’000

Training examples

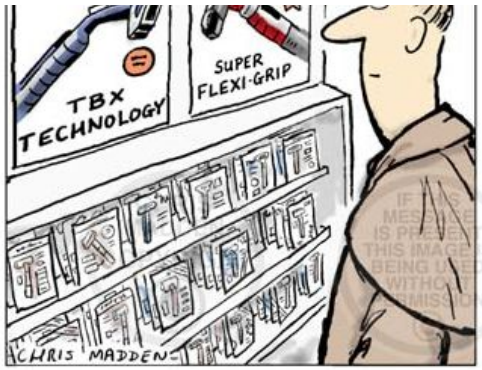
``Simplicity has its value!''
Ralf Drautz, IPAM reunion 2005, Lake Arrowhead, CA

``As simple as possible but not simpler''
``It can scarcely be denied that the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible without having to surrender the adequate representation of a single datum of experience.''
Einstein

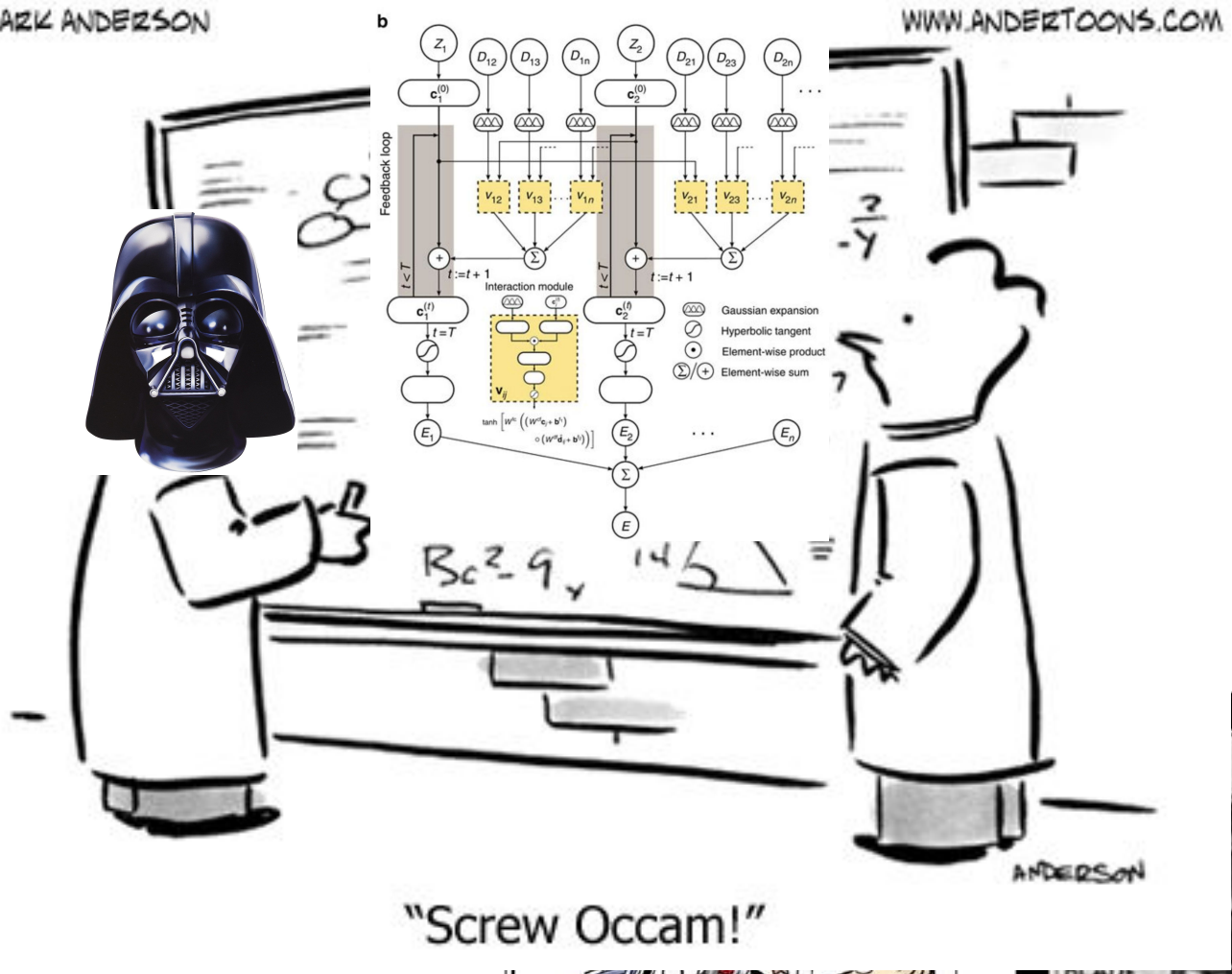
Ockham chooses a razor



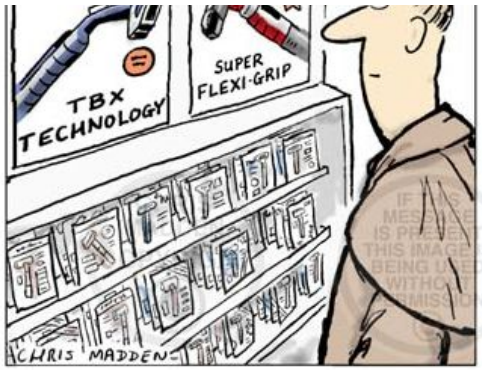
"Screw Occam!"



Ockham chooses a razor



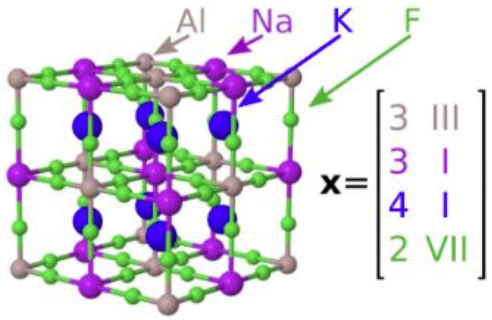
"Screw Occam!"



Ockham chooses a razor

Crystals

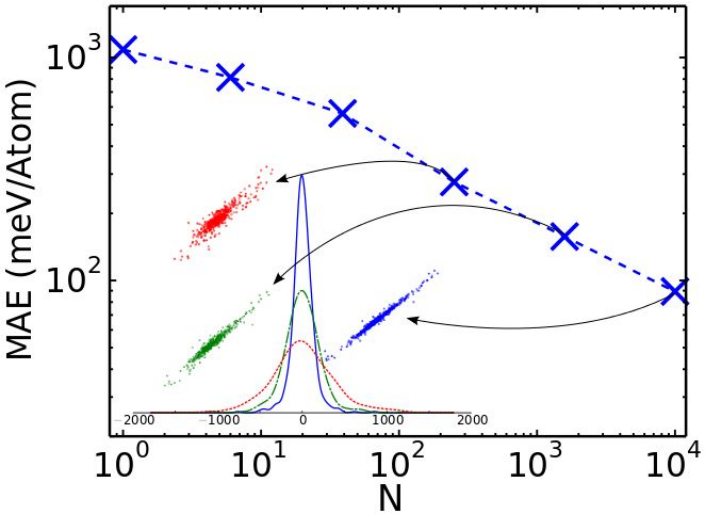
Elpasolite (K_2NaAlF_6-symmetry) is a vitreous, transparent, luster, colorless and soft **quaternary** crystal in the Fm3m space group which can be found in the Rocky Mountains, Virginia, or the Apennines. It is the most abundant quaternary crystal present in the Inorganic Crystal Structure Database. Some Elpasolites emit light when exposed to ionic radiation. This makes them interesting material candidates for scintillator devices.

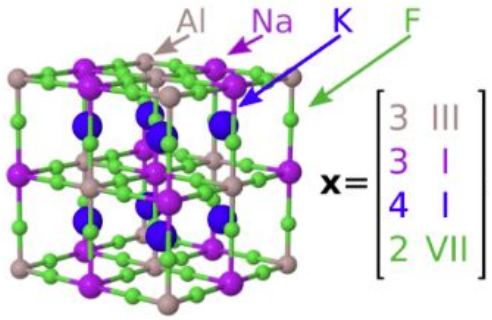


$$E_t(R_t) \approx E_b(R_b) + \Delta_b^t(R_b)$$

$$E^{est}(\mathbf{M}) = \sum_i^N \alpha_i k(\mathbf{M}, \mathbf{M}_i)$$

$$\alpha = (\mathbf{K} + \lambda \mathbf{I})^{-1} \mathbf{E}^{ref}$$





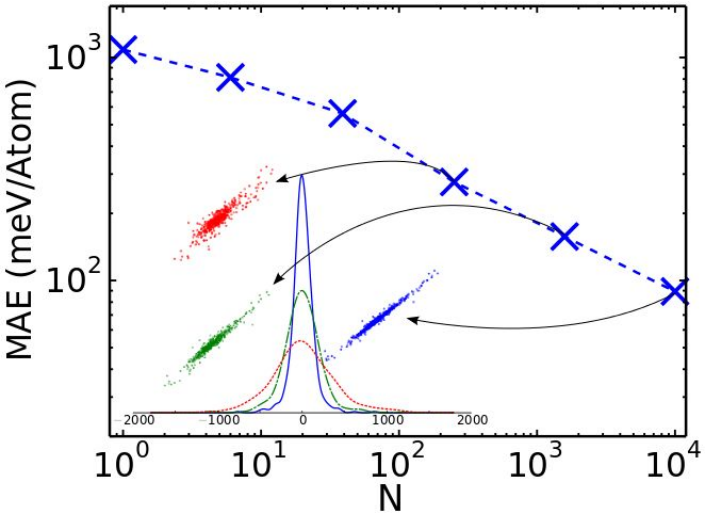
$$E_t(R_t) \approx E_b(R_b) + \Delta_b^t(R_b)$$

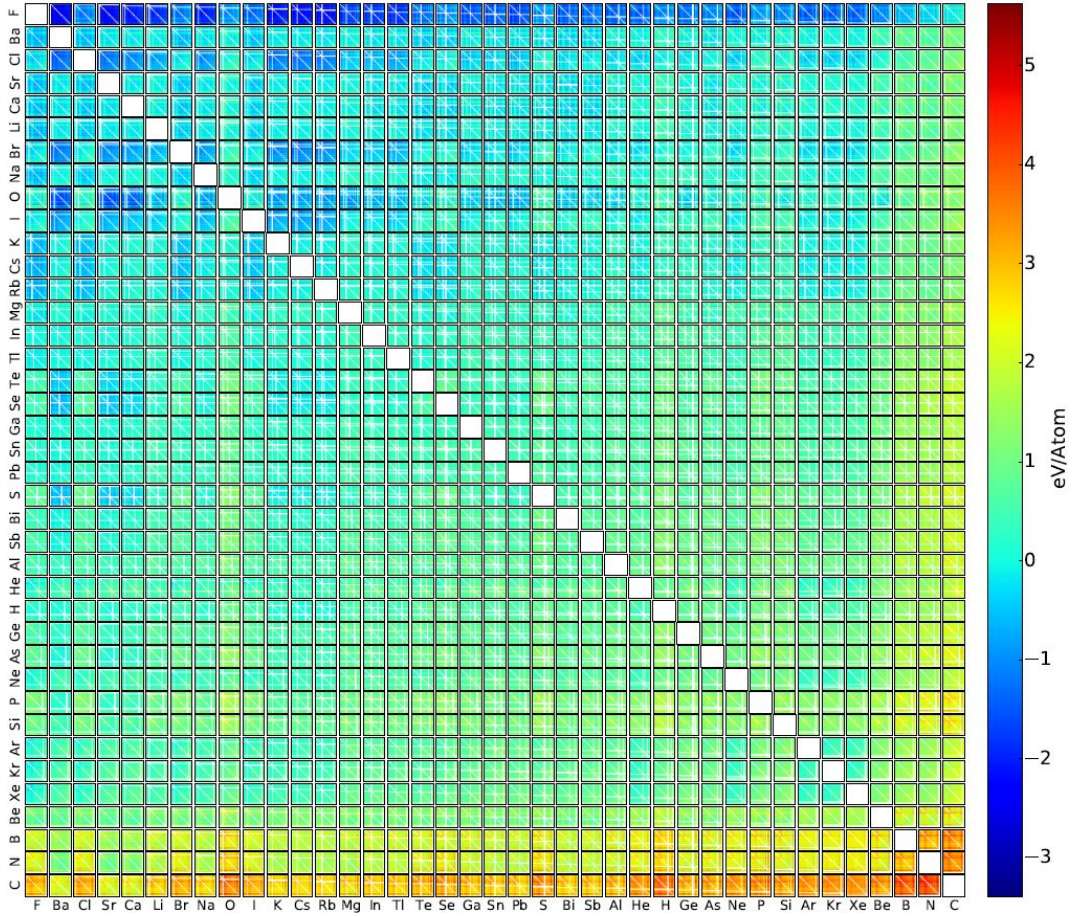
$$E^{est}(\mathbf{M}) = \sum_i^N \alpha_i k(\mathbf{M}, \mathbf{M}_i)$$

$$\alpha = (\mathbf{K} + \lambda \mathbf{I})^{-1} \mathbf{E}^{ref}$$

Machine Learning Energies of 2 Million Elpasolite (ABC_2D_6) Crystals

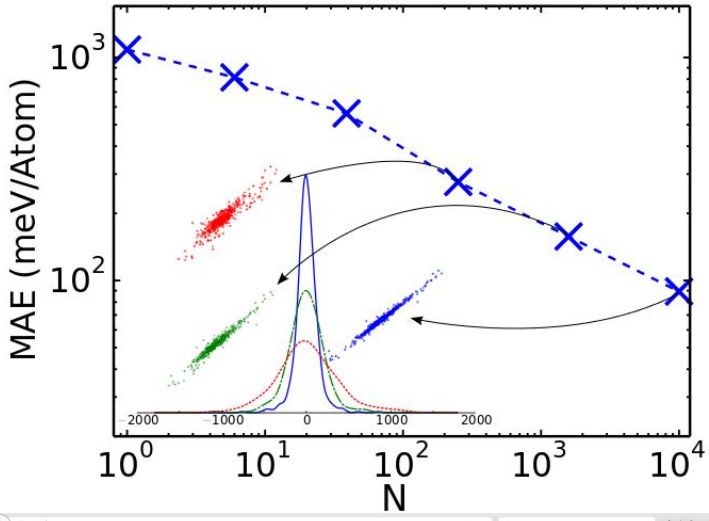
Felix A. Faber,[1] Alexander Lindmaa,[2] O. Anatole von Lilienfeld,[1,3,*] and Rickard Armiento[2,†]

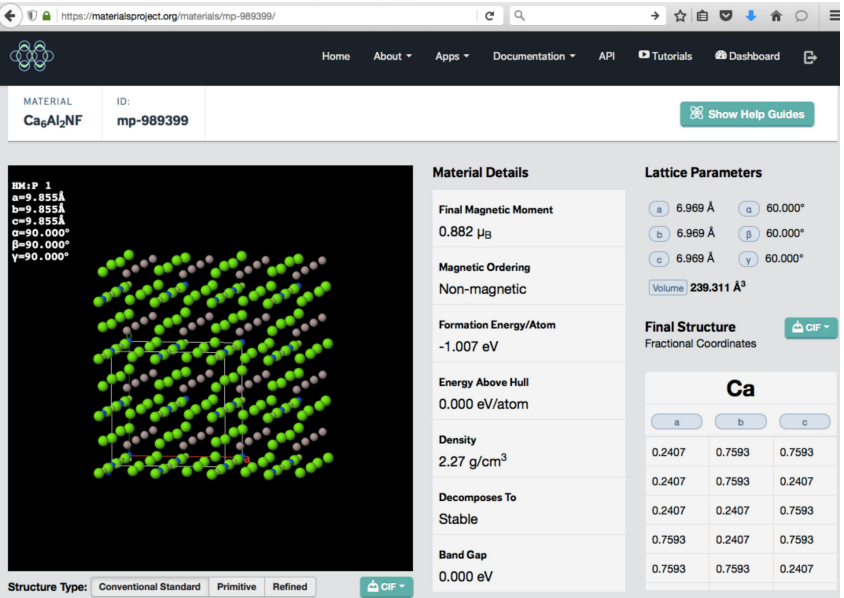


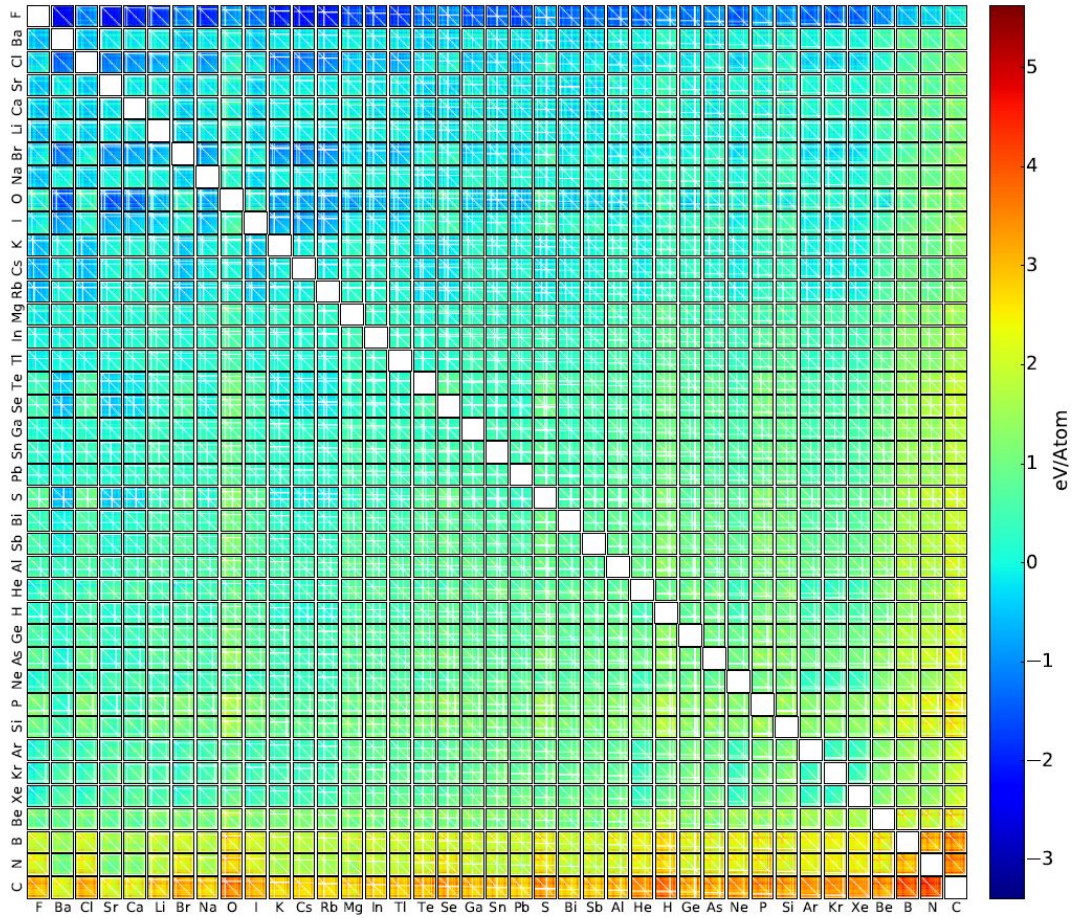


Machine Learning Energies of 2 Million Elpasolite (ABC_2D_6) Crystals

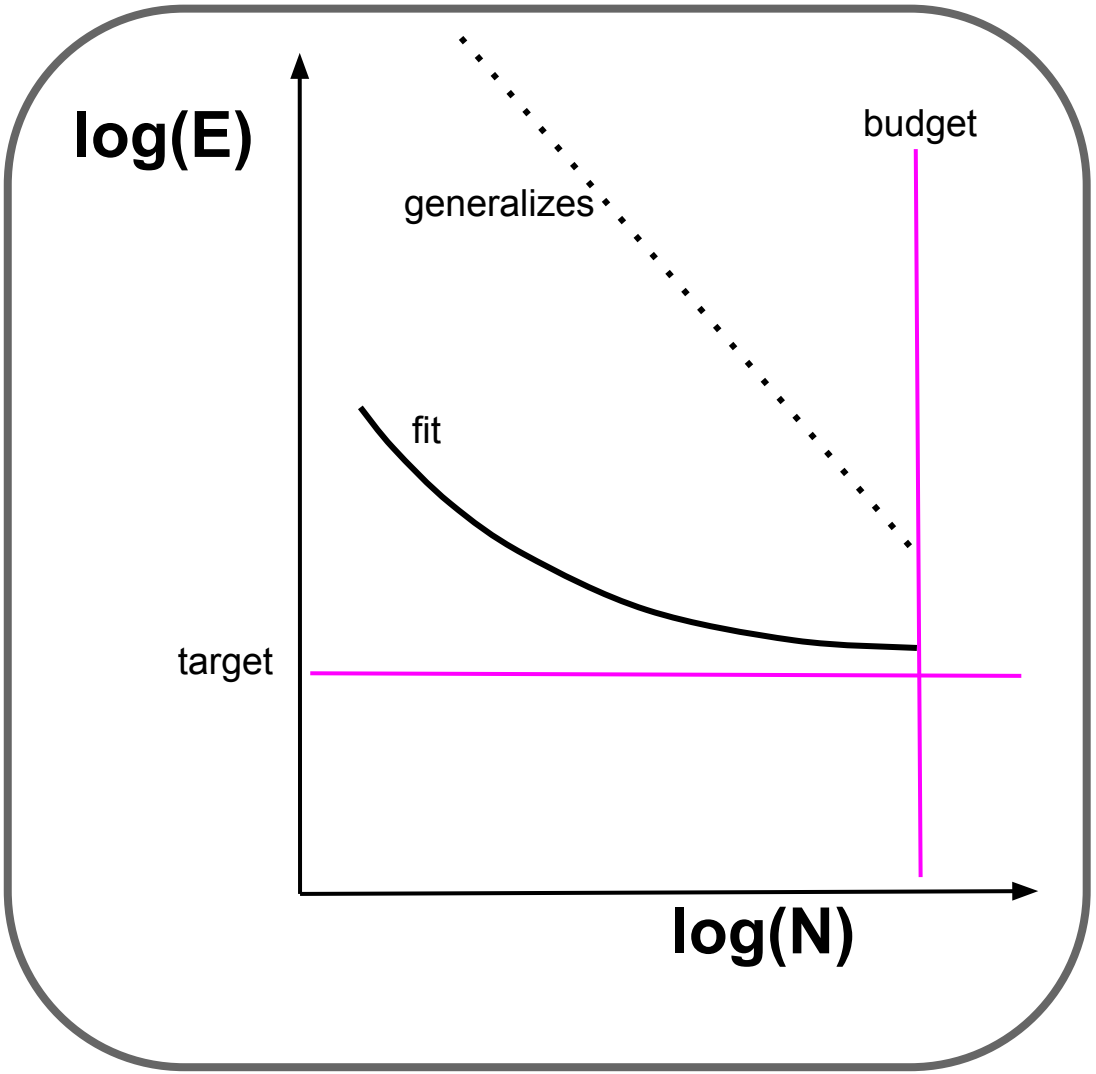
Felix A. Faber,[1] Alexander Lindmaa,[2] O. Anatole von Lilienfeld,[1,3,*] and Rickard Armiento[2,†]







Learning curves



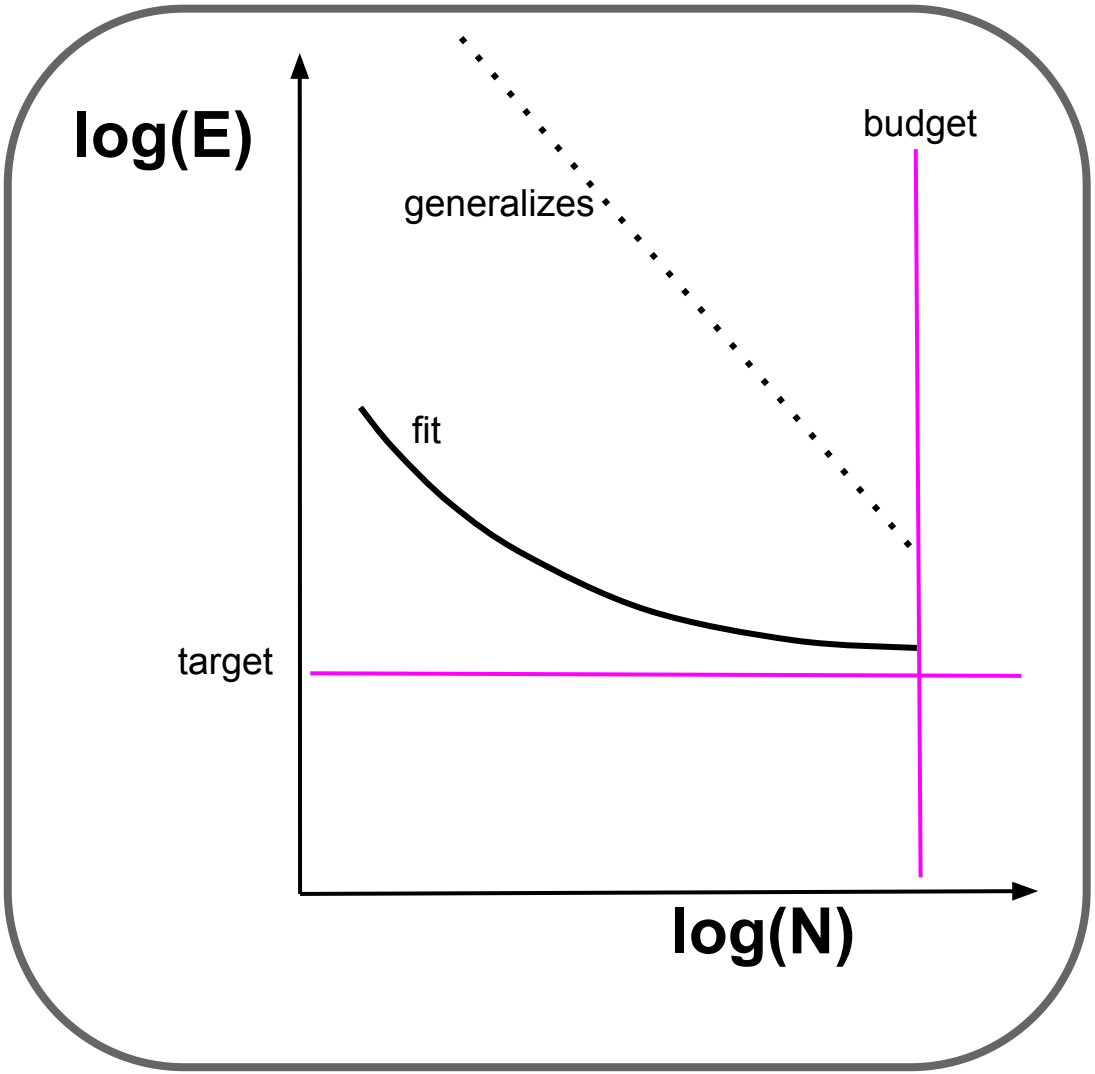
Reminder:
Machine ``learning'' iff the error decays w data!

Evolution vs. Revolution?

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

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

Learning curves



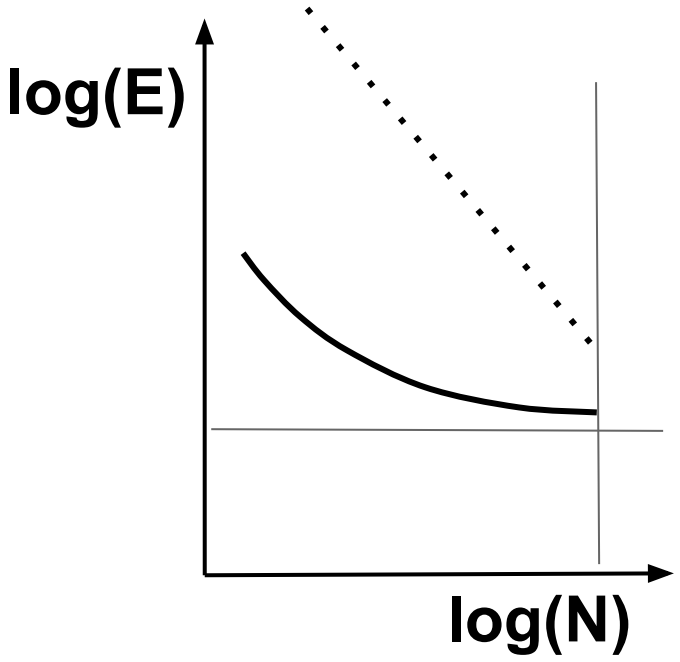
Reminder:
Machine ``learning'' iff the error decays w data!

~~**Evolution vs.**~~
Revolution!!!!
models joke

Twofold:
1. **Control accuracy**
2. **Speed:**
   QM ~ CPUh
   QML ~ CPUms

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

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



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

κ ~ Ψ

α ~ Ô

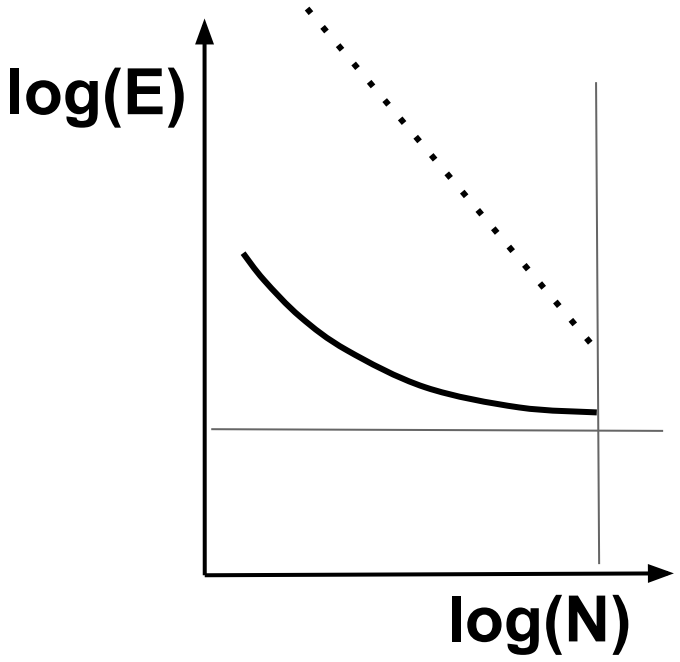
Ramakrishnan, OAvL, *CHIMIA* (2015)

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

Error ~ a/N^b

→ log(Error) = log(a) - $b$log(N)

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



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

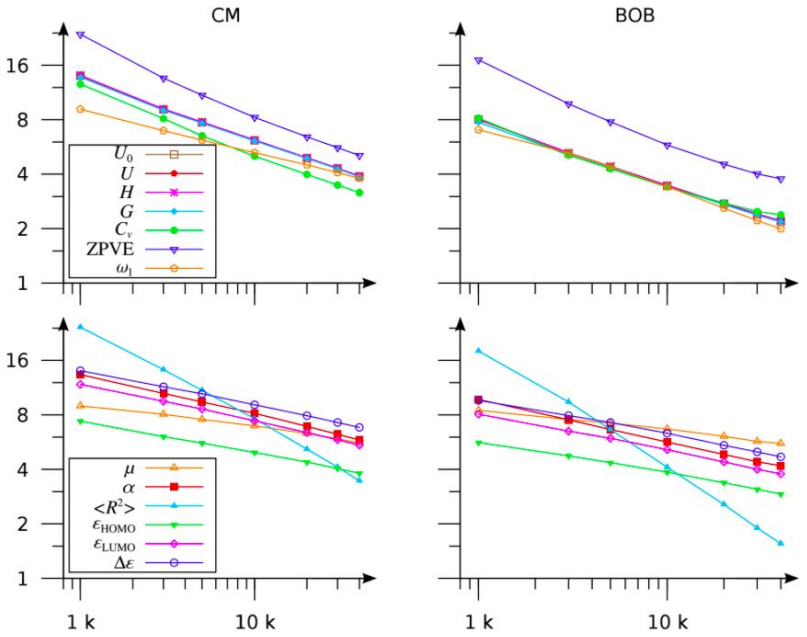
κ ~ Ψ

α ~ Ô

Ramakrishnan, OAvL, *CHIMIA* (2015)

QML vs QSPR

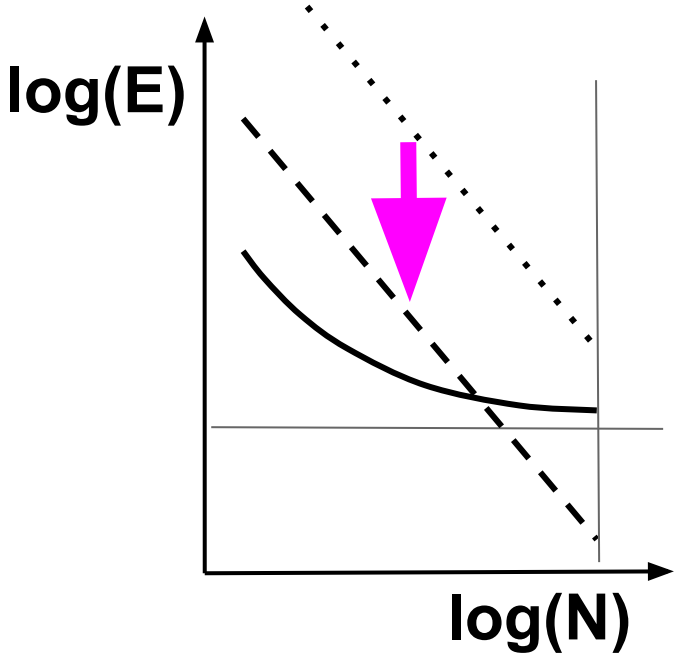
$$P^{\text{est}}(\mathbf{M}) = \sum_i \alpha_i k(\mathbf{M}, \mathbf{M}_i)$$
$$\vec{\alpha} = \mathbf{K}^{-1} \vec{P}^{\text{ref}}$$



Error ~ a/N^b

→ log(Error) = log(a) - $b$log(N)

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



How???

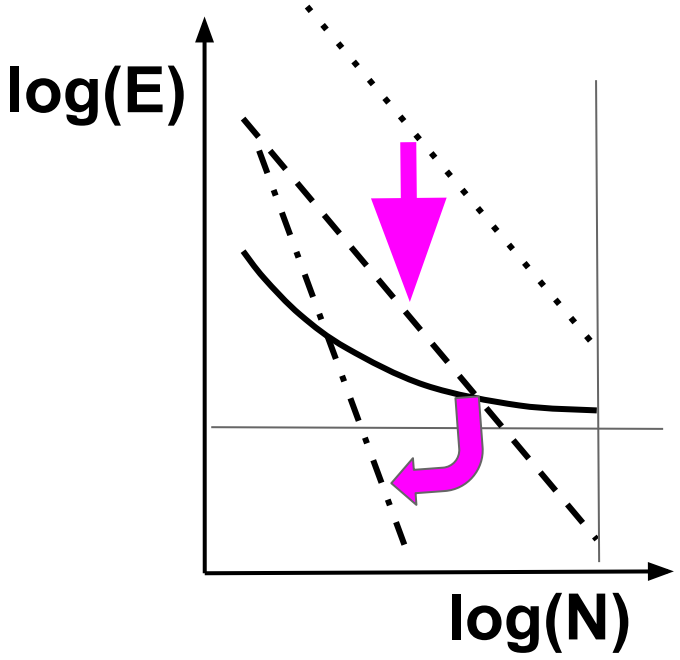
$$P^{\text{est}}(\mathbf{M}) = \sum_i \alpha_i k(\mathbf{M}, \mathbf{M}_i)$$
$$\vec{\alpha} = \mathbf{K}^{-1} \vec{P}^{\text{ref}}$$

Error ~ a/N^b

→ log(Error) = log(a) - $b$log(N)

`Quantum Machine Learning', von Lilienfeld, *Angew. Chem. Int. Ed.* (2018)

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



How???

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

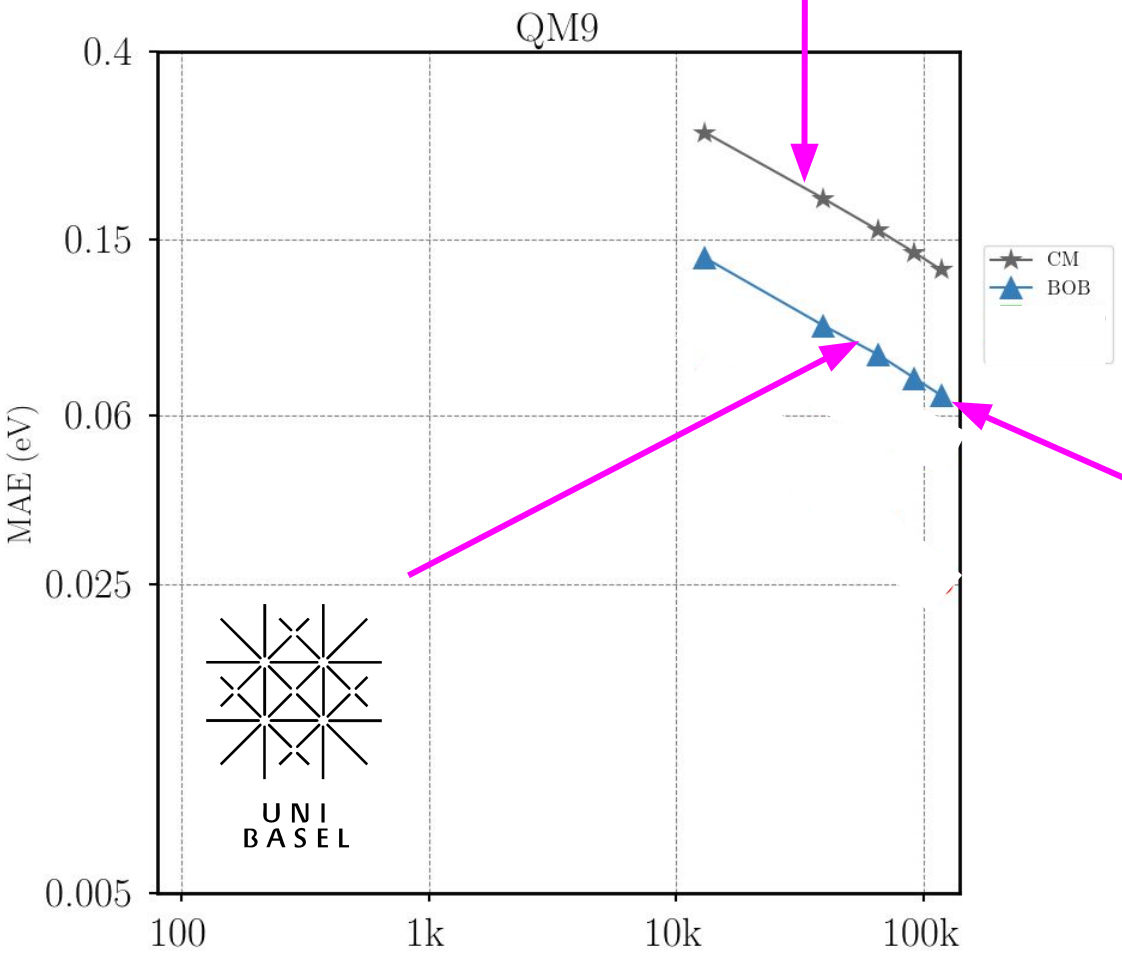
Error ~ a/N^b

→ log(Error) = log(a) - $b$log(N)

`Quantum Machine Learning', von Lilienfeld, *Angew. Chem. Int. Ed.* (2018)

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

Learning curves for QM9 energies (pre 2017)



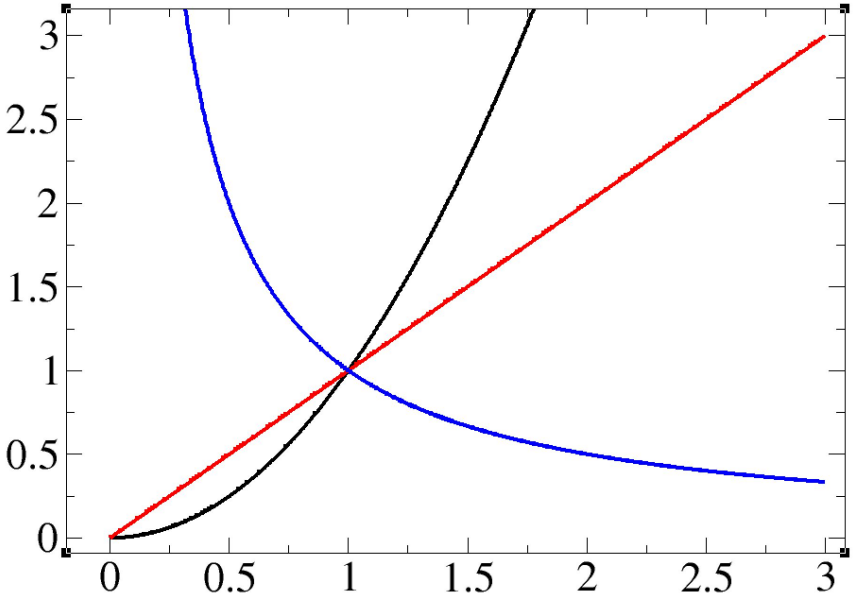
2015 BoB: Hansen, Tkatchenko (FHI)/Muller (TU Berlin)/Lilienfeld (Basel)

2014 QM9: Ramakrishnan, …, Lilienfeld (Basel). B3LYP for GDB-9 subset of GDB-17 from Reymond and co-workers (Berne)

2012 CM: Rupp, Tkatchenko, Muller, Lilienfeld (IPAM) (~10kcal/mol)

Representation

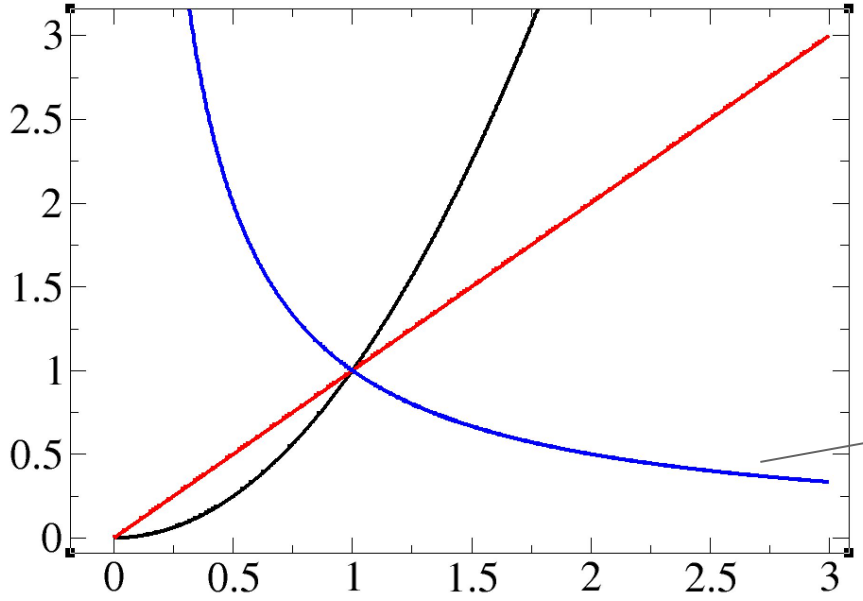
$$\mathrm{CM}_{IJ}^{(n)} = \frac{Z_I Z_J}{R_{IJ}^n}$$

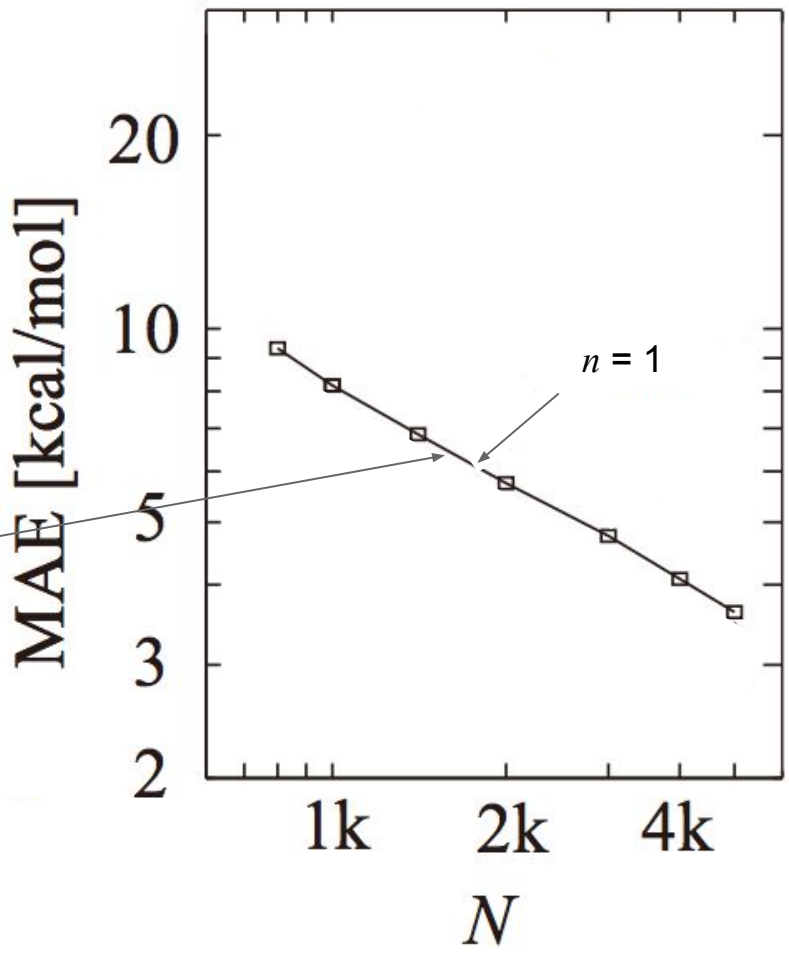


OAvL et al, *Int J Quantum Chem* (2015), Huang, OAvL, *J Chem Phys* (2016)

Representation

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

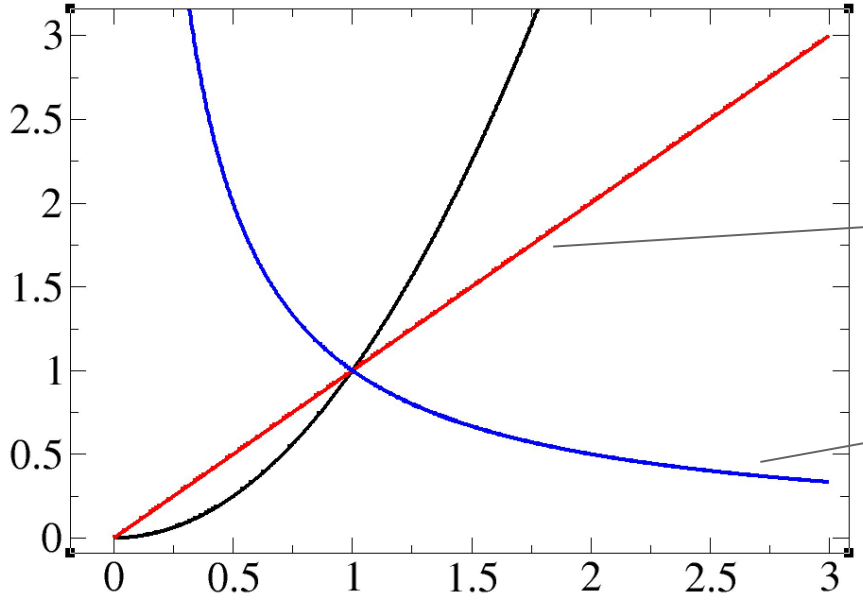


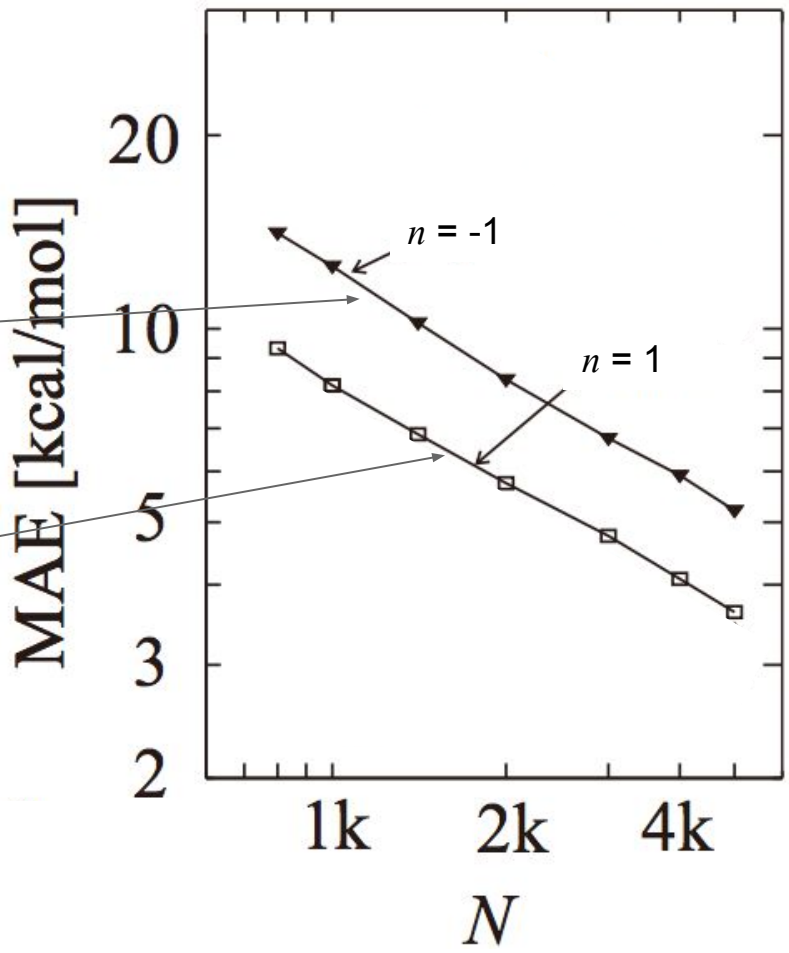


OAvL et al, *Int J Quantum Chem* (2015), Huang, OAvL, *J Chem Phys* (2016)

Representation

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

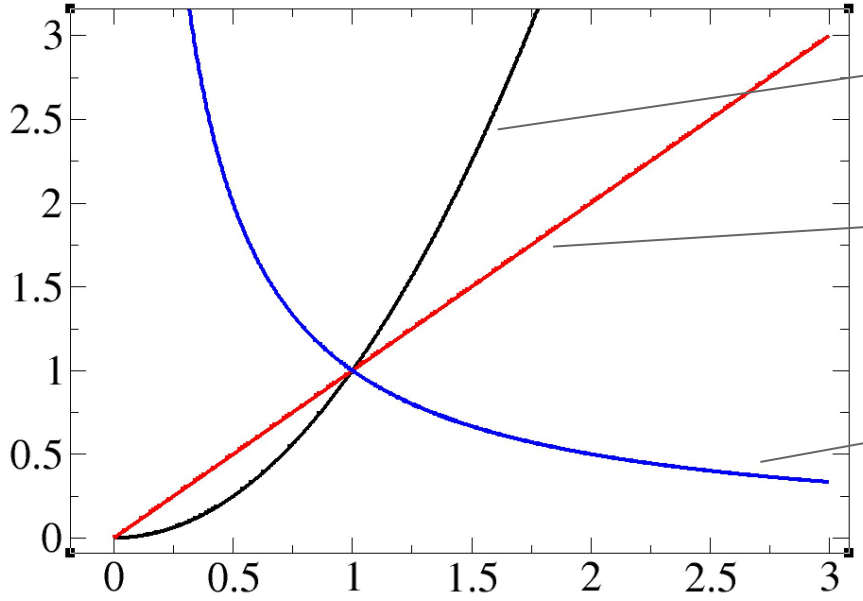


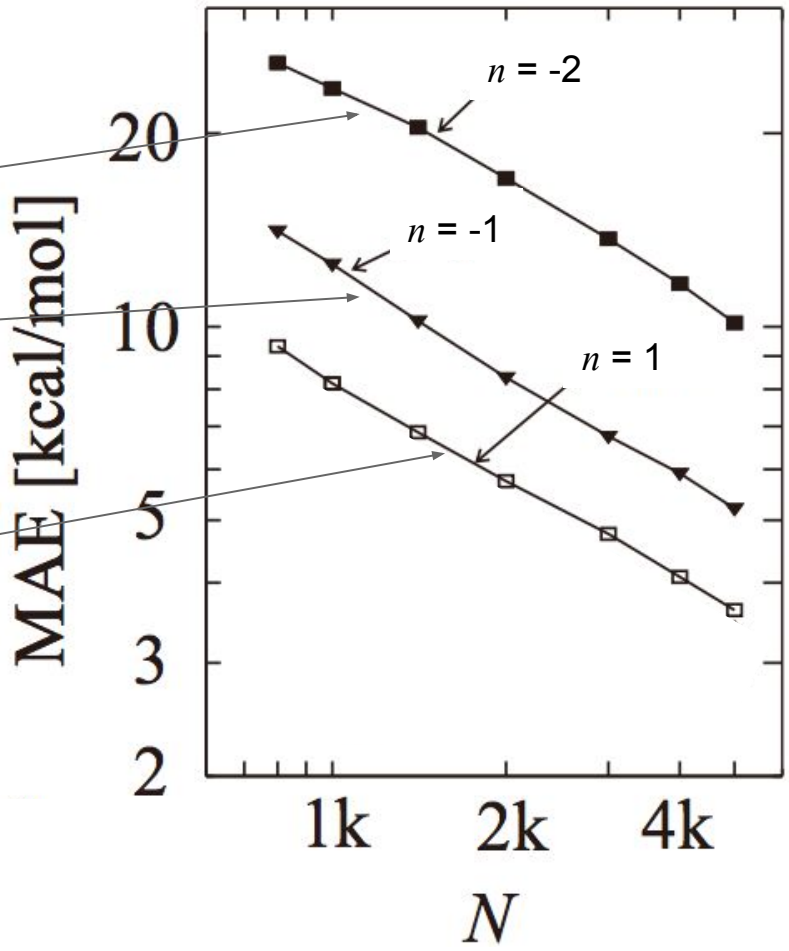


OAvL et al, *Int J Quantum Chem* (2015), Huang, OAvL, *J Chem Phys* (2016)

Representation

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

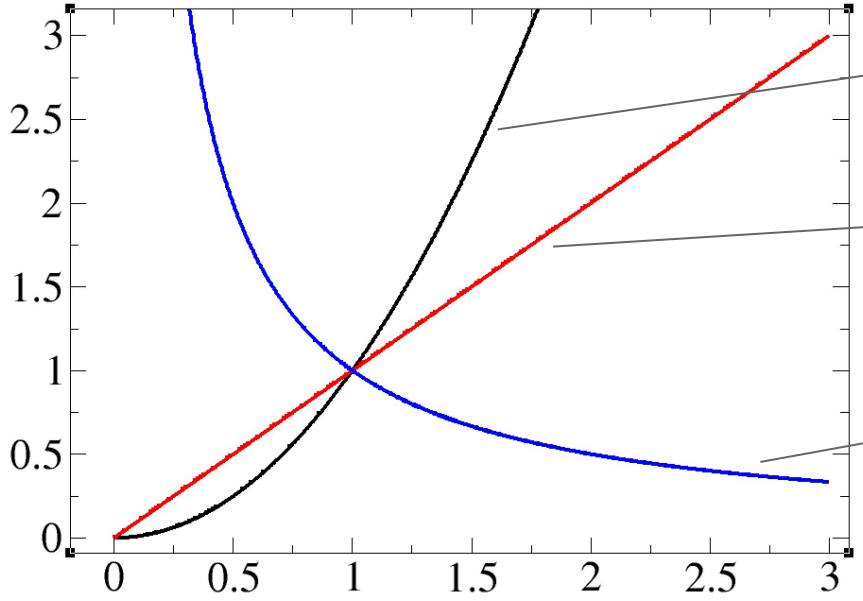


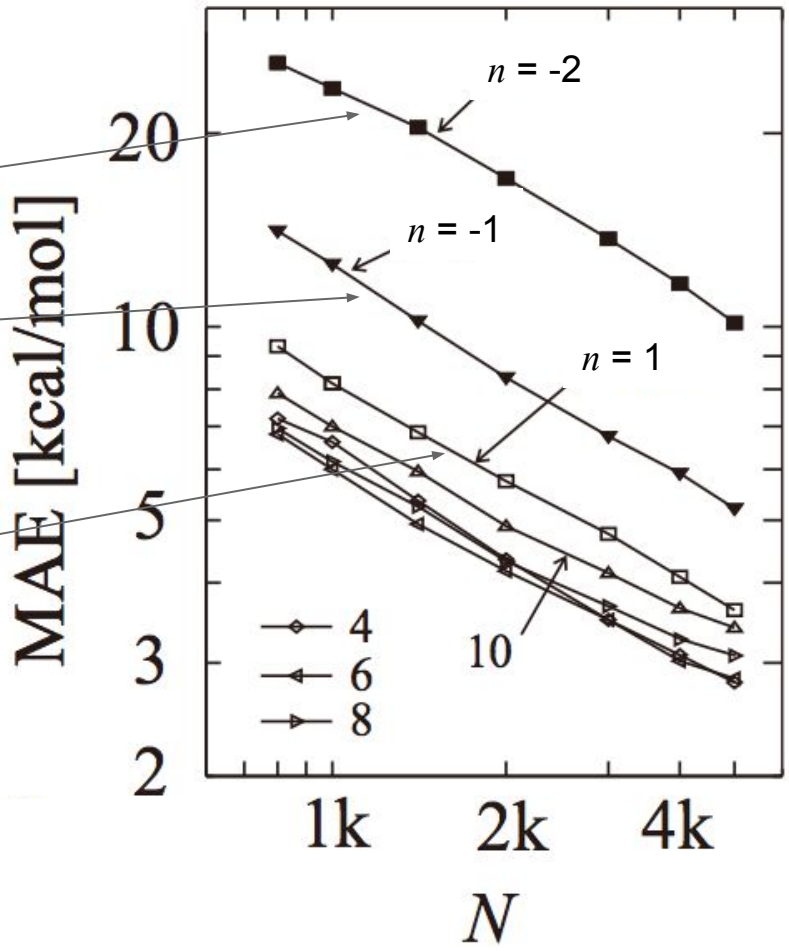


OAvL et al, *Int J Quantum Chem* (2015), Huang, OAvL, *J Chem Phys* (2016)

Representation

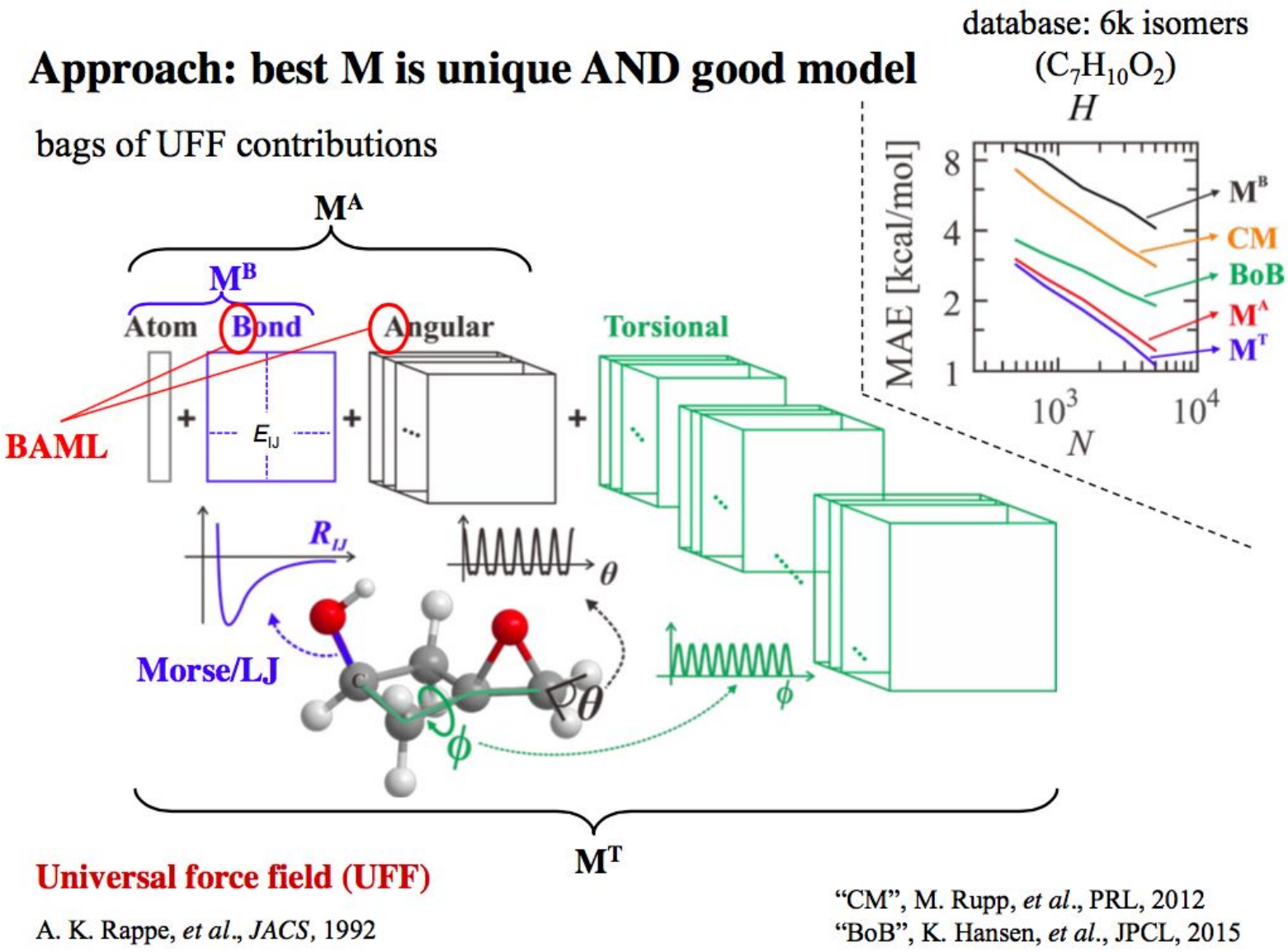
$$\mathrm{CM}_{IJ}^{(n)} = \frac{Z_I Z_J}{R_{IJ}^n}$$





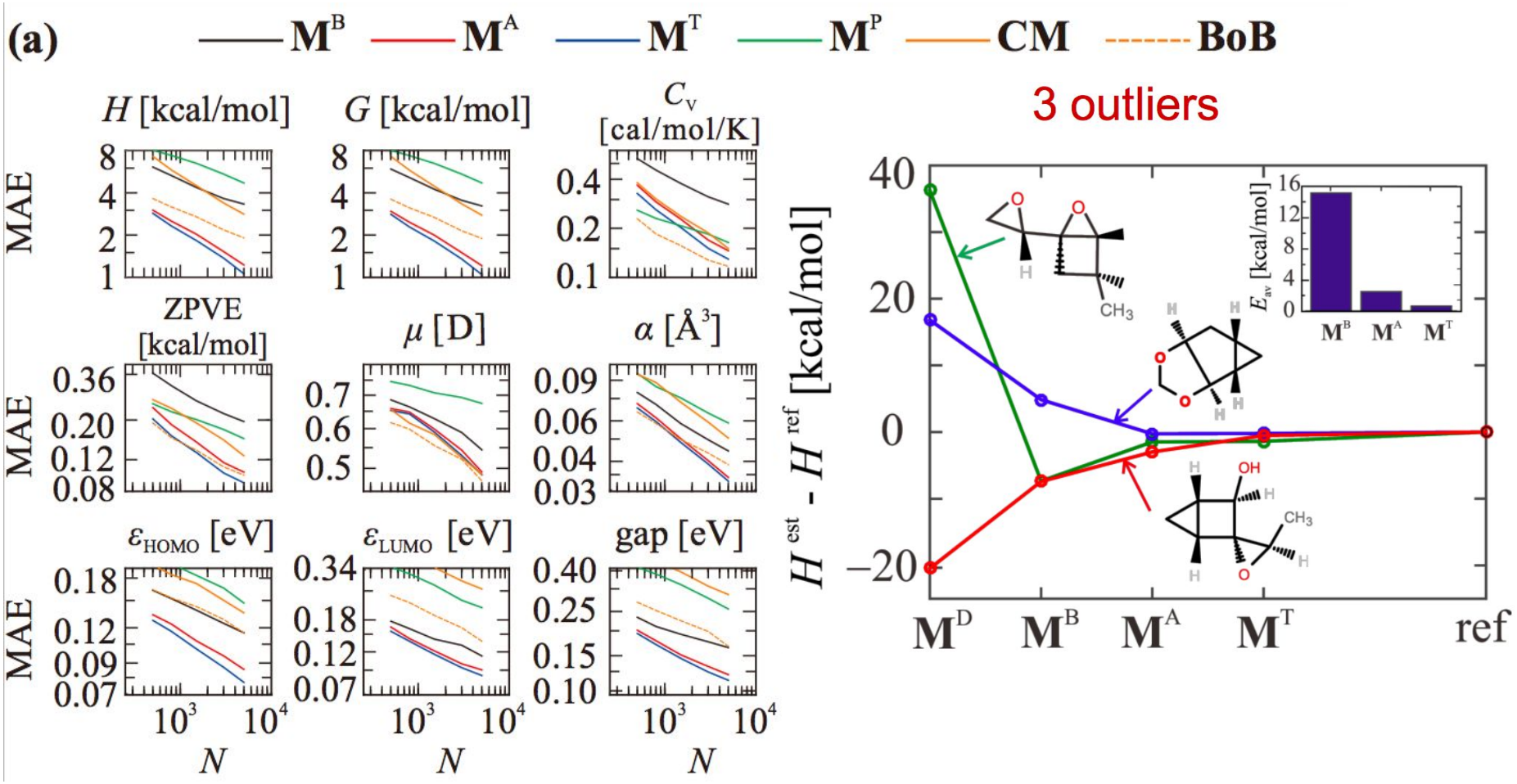
OAvL et al, *Int J Quantum Chem* (2015), Huang, OAvL, *J Chem Phys* (2016)

BAML



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

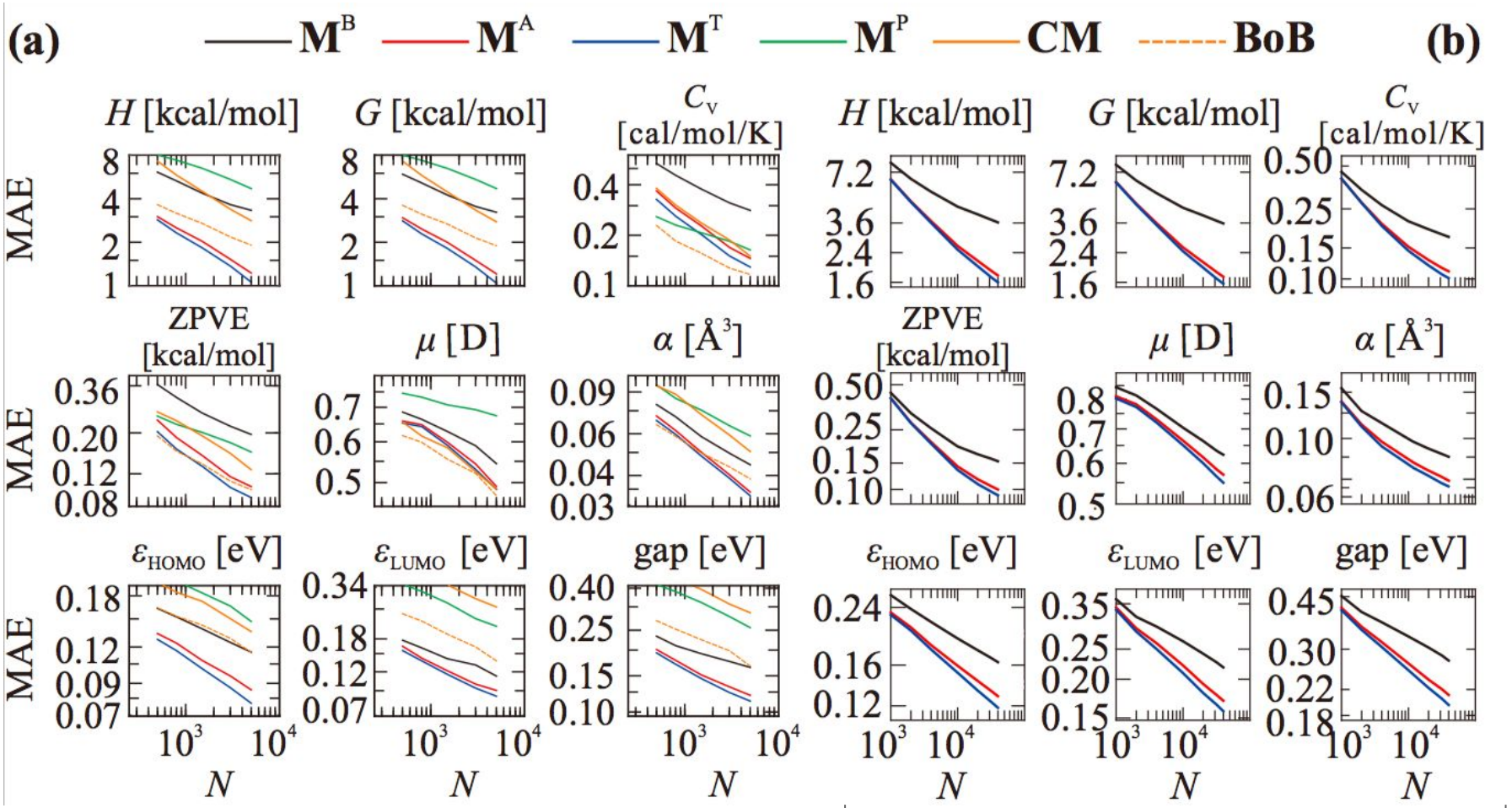
BAML



6k constitutional isomers of $C_7O_2H_{10}$

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

BAML

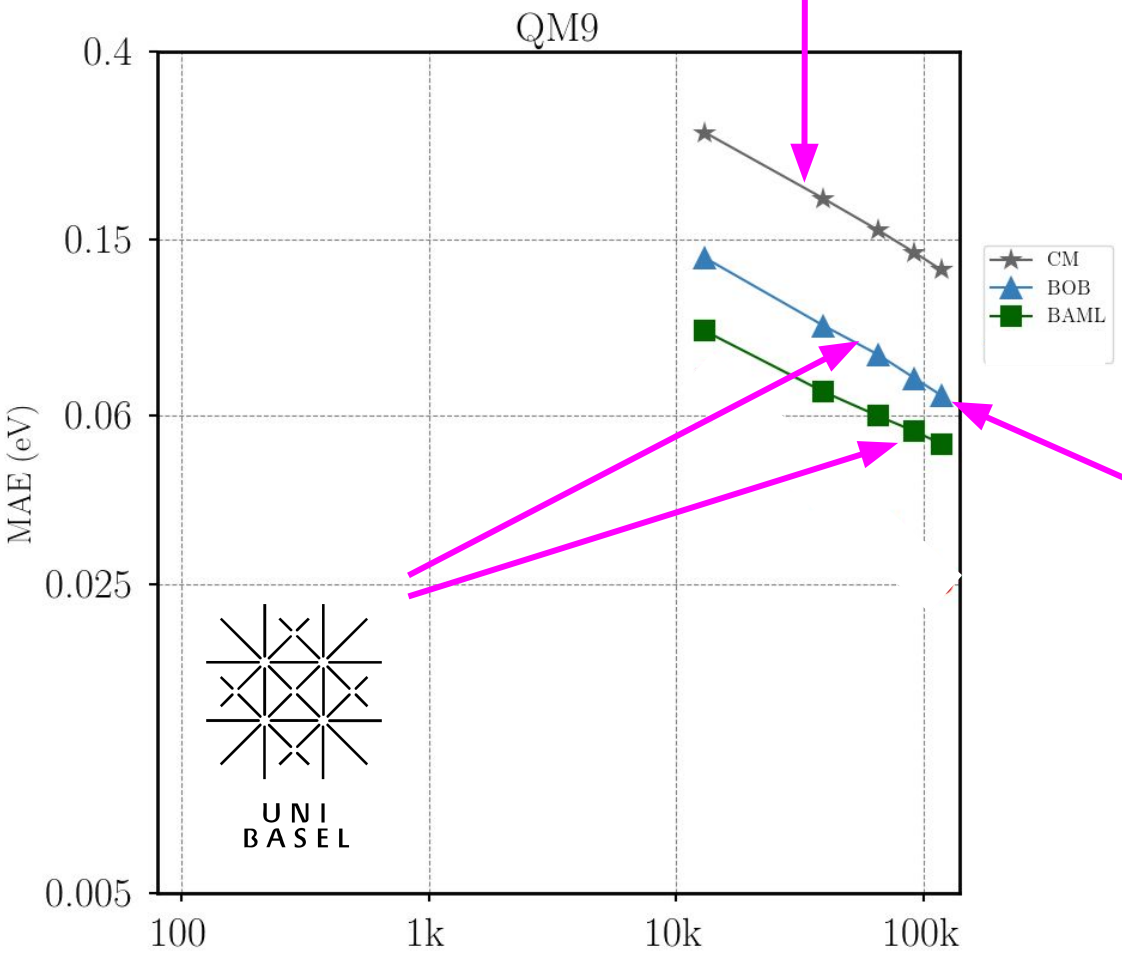


6k constitutional isomers of $C_7O_2H_{10}$

QM9 (134k molecules)

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

Learning curves for QM9 energies (pre 2017)



2016 BAML: Huang, Lilienfeld (Basel)

2015 BoB: Hansen, Tkatchenko (FHI)/Muller (TU Berlin)/Lilienfeld (Basel)

2014 QM9: Ramakrishnan, …, Lilienfeld (Basel). B3LYP for GDB-9 subset of GDB-17 from Reymond and co-workers (Berne)

2012 CM: Rupp, Tkatchenko, Muller, Lilienfeld (IPAM) (~10kcal/mol)

JCTC Journal of Chemical Theory and Computation

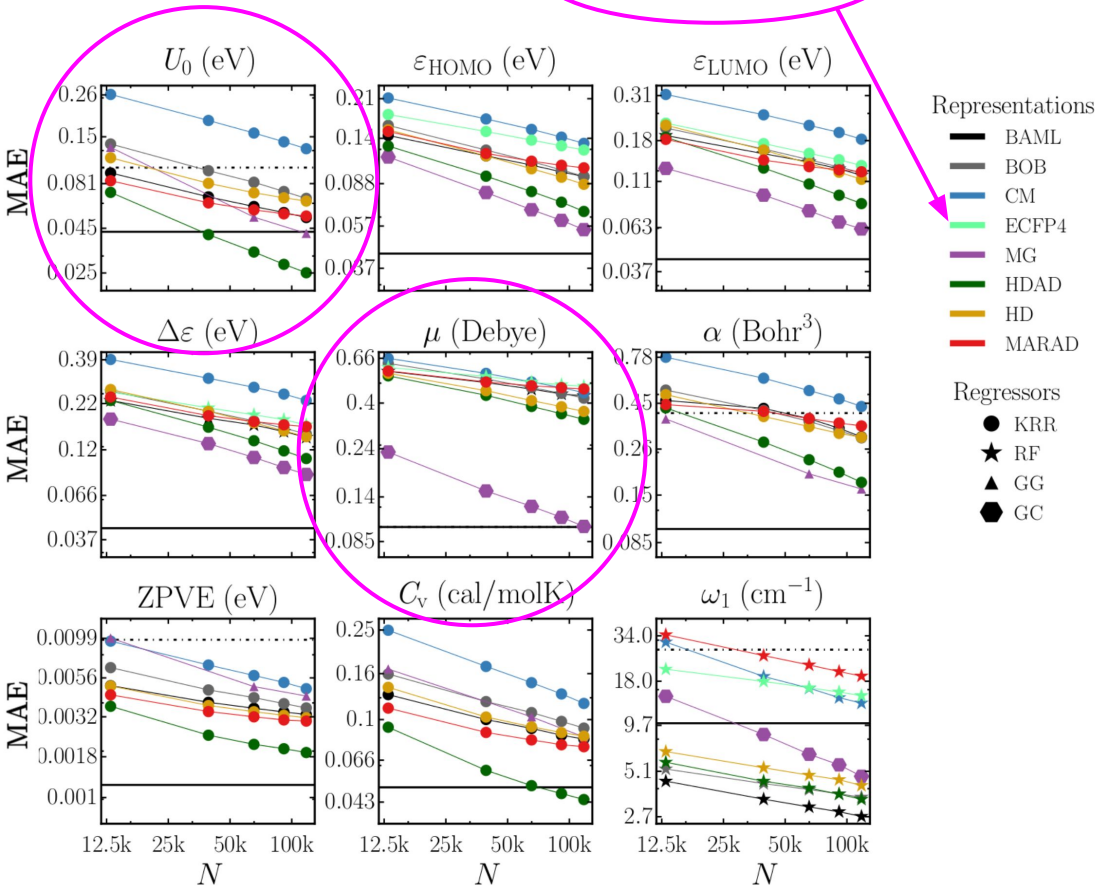
Article

Cite This: J. Chem. Theory Comput. 2017, 13, 5255-5264

pubs.acs.org/JCTC

Prediction Errors of Molecular Machine Learning Models Lower than Hybrid DFT Error

Felix A. Faber,[†] Luke Hutchison,[‡] Bing Huang,[†] Justin Gilmer,[‡] Samuel S. Schoenholz,[‡] George E. Dahl,[‡] Oriol Vinyals,[¶] Steven Kearnes,[‡] Patrick F. Riley,[‡] and O. Anatole von Lilienfeld*,[†]



Research at Google

JCTC Journal of Chemical Theory and Computation

Article

Cite This: J. Chem. Theory Comput. 2017, 13, 5255-5264

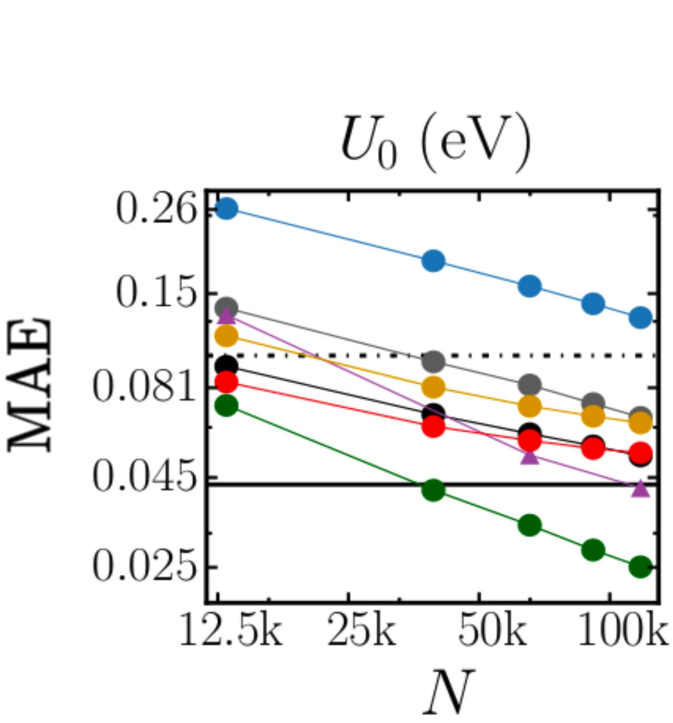
pubs.acs.org/JCTC

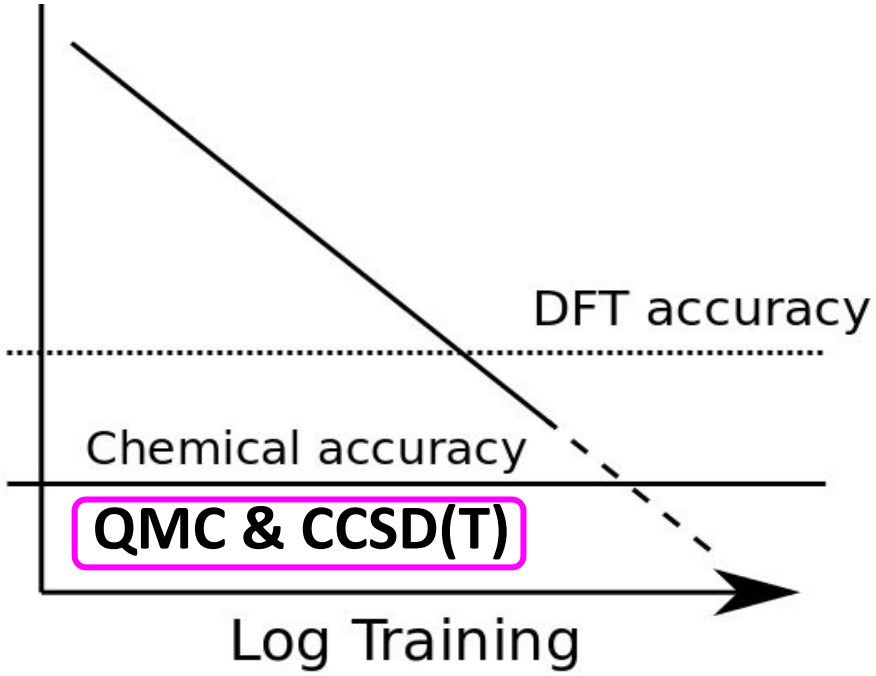
Prediction Errors of Molecular Machine Learning Models Lower than Hybrid DFT Error

Felix A. Faber,[†] Luke Hutchison,[‡] Bing Huang,[†] J
Oriol Vinyals,[¶] Steven Kearnes,[‡] Patrick F. Riley,

Log Error

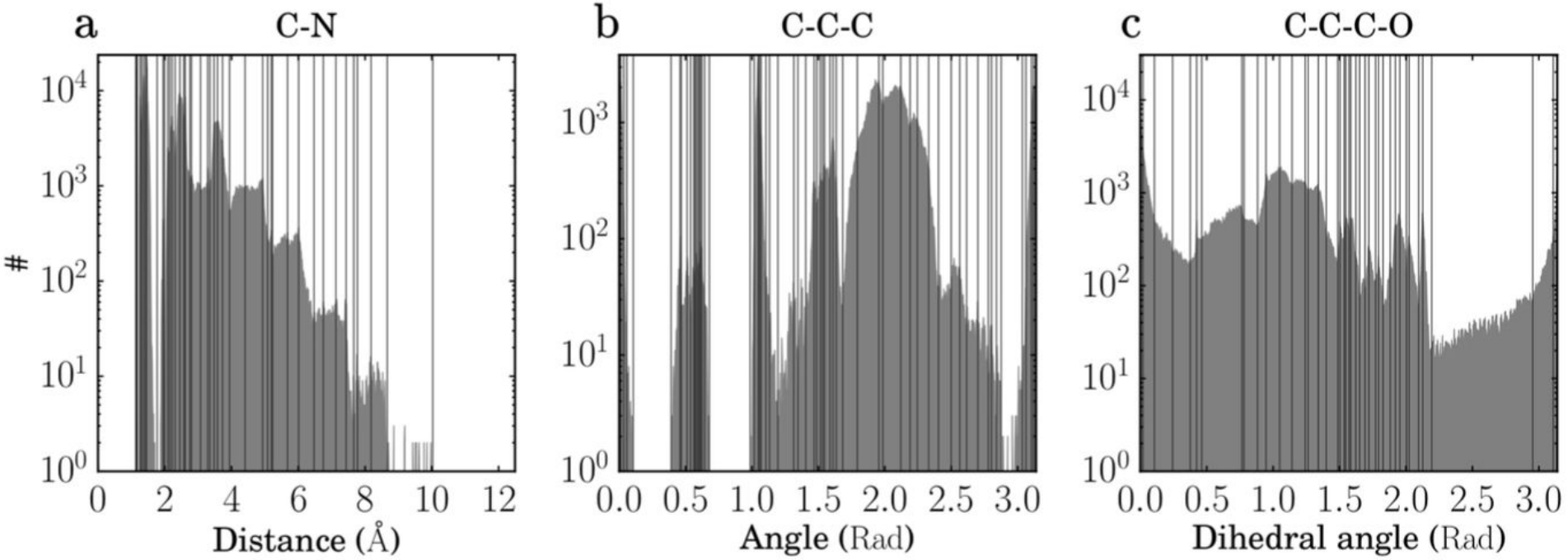
Kernel Ridge Regression





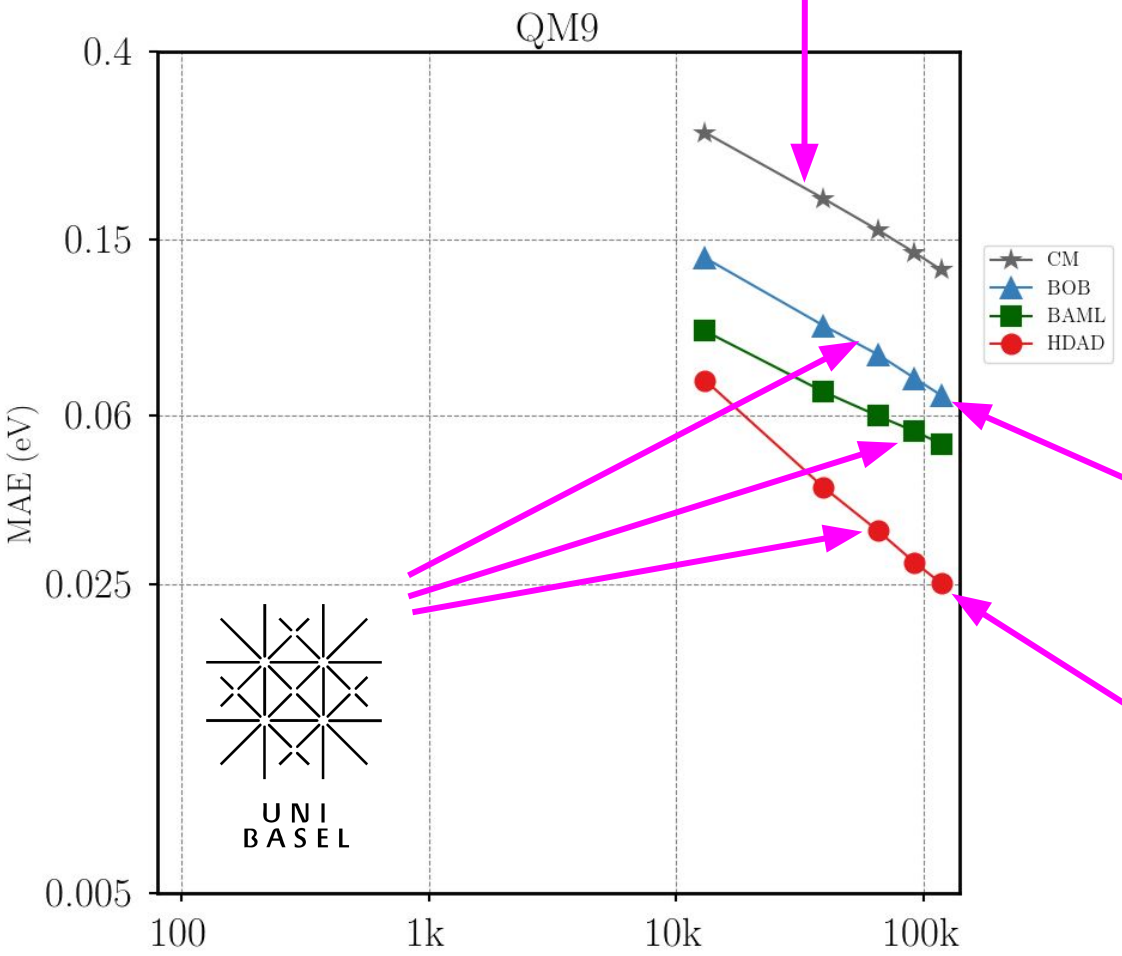
JCTC Journal of Chemical Theory and Computation
Article
Cite This: J. Chem. Theory Comput. 2017, 13, 5255-5264
pubs.acs.org/JCTC

Prediction Errors of Molecular Machine Learning Models Lower than Hybrid DFT Error



Histogram Distances Angles Dihedrals (HDAD) representation

Learning curves for QM9 energies (2017)



2017 HDAD: Faber et al (Basel)/Riley et al (Google Accelerated Science)

2016 BAML: Huang, Lilienfeld (Basel)

2015 BoB: Hansen, Tkatchenko (FHI)/Muller (TU Berlin)/Lilienfeld (Basel)

2014 QM9: Ramakrishnan, …, Lilienfeld (Basel). B3LYP for GDB-9 subset of GDB-17 from Reymond and co-workers (Berne)

2012 CM: Rupp, Tkatchenko, Muller, Lilienfeld (IPAM) (~10kcal/mol)

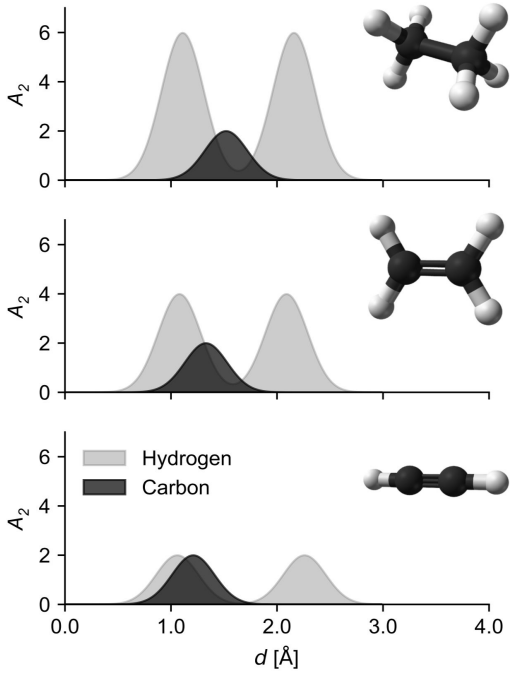
FCHL representation

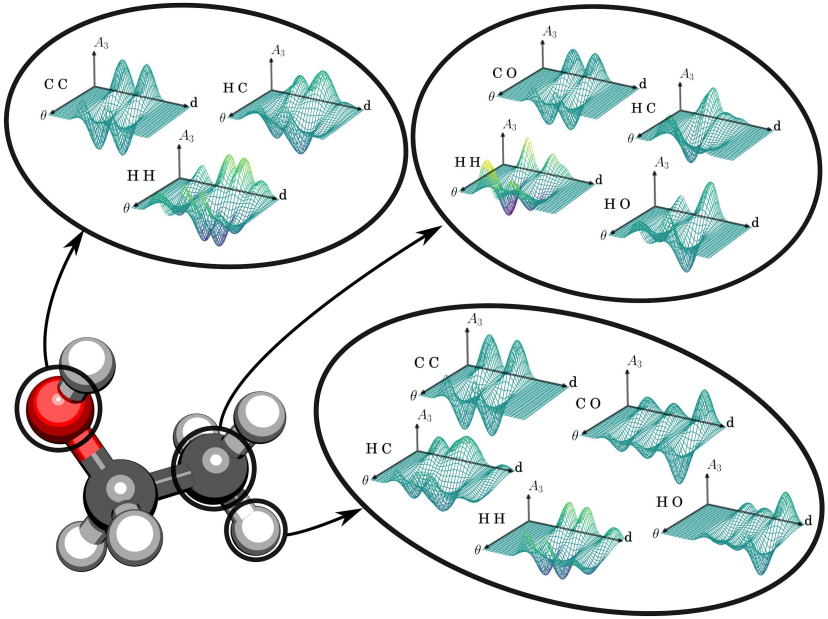
$$\mathcal{A}_M(I) \;=\; \{A_1(I), A_2(I), A_3(I), \ldots, A_M(I)\}$$

$$A_1(I) = \mathcal{N}(\mathbf{x}_I^{(1)}) \;=\; e^{-\frac{(P_I-\chi_1)^2}{2\sigma_P^2} - \frac{(G_I-\chi_2)^2}{2\sigma_G^2}}$$

$$A_2(I) \quad = \quad \mathcal{N}(\mathbf{x}_I^{(1)}) \sum_{i\neq I} \mathcal{N}(\mathbf{x}_{iI}^{(2)}) \xi_2(d_{iI})$$

$$\Delta(\mathcal{A}_M(I), \mathcal{A}_M(J))^2 \equiv \sum_{m=0}^{M} \beta_m \frac{1}{\varsigma_m^2} \int_{\mathbb{R}^{3m+2}} (A_m(I) - A_m(J))^2 d\chi_1 \cdots d\chi_{3m+2}$$





"Alchemical and structural distribution based representation for universal QML", Faber et al, *J Chem Phys* (2018), *https://arxiv.org/abs/1712.08417*

FCHL representation

$$\mathcal{A}_M(I) \;=\; \{A_1(I), A_2(I), A_3(I), \ldots, A_M(I)\}$$

$$A_1(I) = \mathcal{N}(\mathbf{x}_I^{(1)}) \;=\; e^{-\frac{(P_I-\chi_1)^2}{2\sigma_P^2} - \frac{(G_I-\chi_2)^2}{2\sigma_G^2}}$$

$$A_2(I) \quad = \quad \mathcal{N}(\mathbf{x}_I^{(1)}) \sum_{i\neq I} \mathcal{N}(\mathbf{x}_{iI}^{(2)}) \xi_2(d_{iI})$$

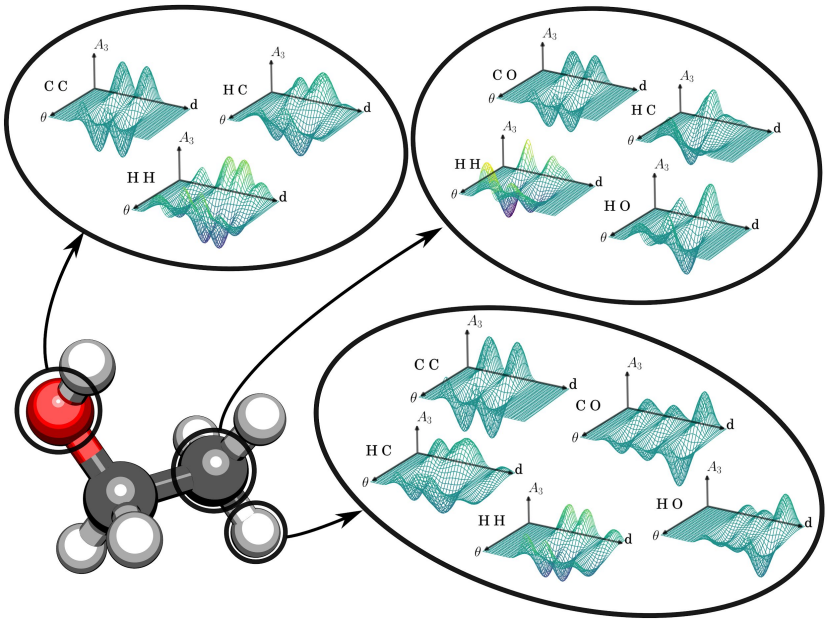
$$\Delta(\mathcal{A}_M(I), \mathcal{A}_M(J))^2 \equiv \sum_{m=0}^{M} \beta_m \frac{1}{\varsigma_m^2} \int_{\mathbb{R}^{3m+2}} (A_m(I) - A_m(J))^2 d\chi_1 \cdots d\chi_{3m+2}$$

$$\Delta(A_m(I), A_m(J))^2 = \frac{1}{\varsigma_m^2} \int_{\mathbb{R}^{3m-1}} d\chi_1 \cdots d\chi_{3m-1} (A_m(I) - A_m(J))^2$$

$$\frac{1}{\varsigma_1^2} \int_{\mathbb{R}^2} d\chi_1 d\chi_2 A_1(I) A_1(J) = \frac{1}{2} \exp\left(-\frac{(P_I - P_J)^2}{4\sigma_P^2} - \frac{(G_I - G_J)^2}{4\sigma_G^2}\right)$$

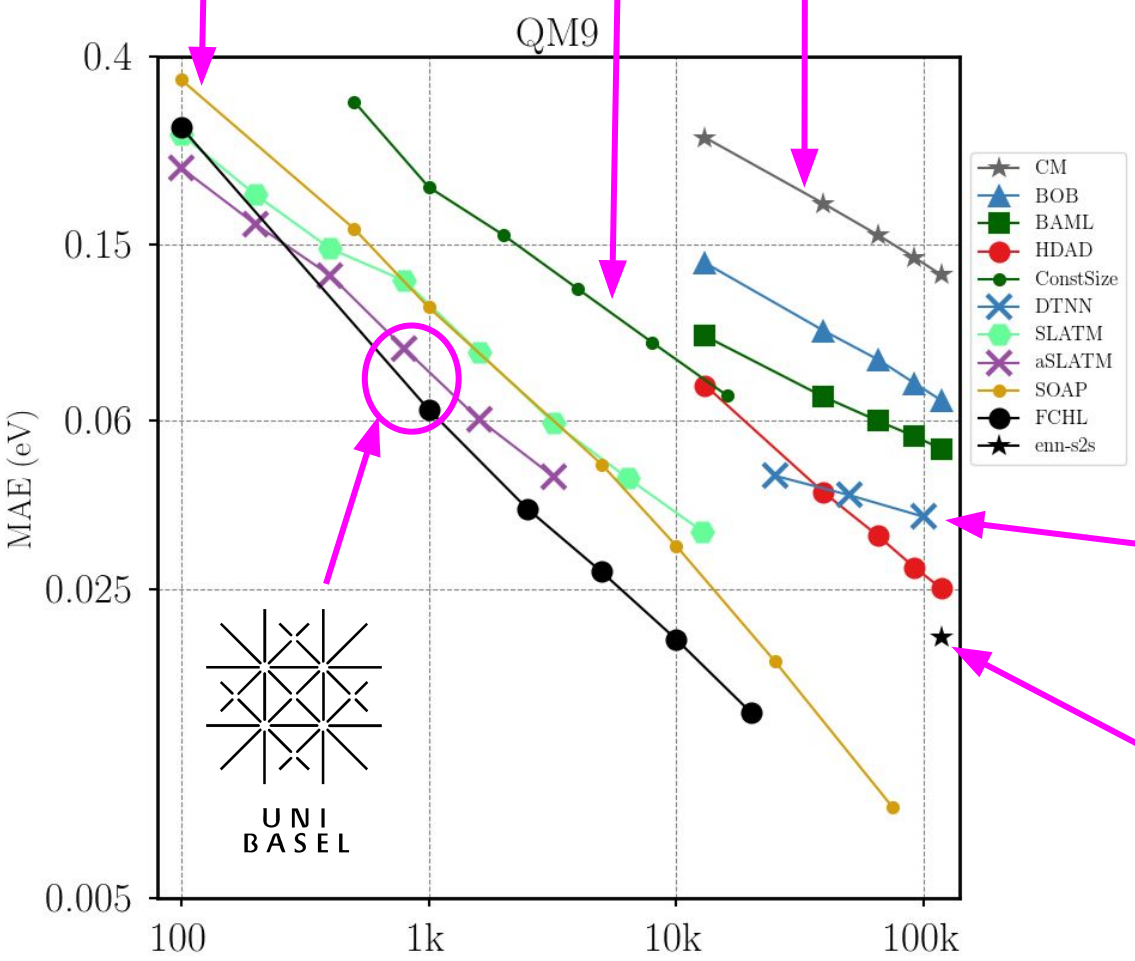
$$\frac{1}{\varsigma_2^2} \int_{\mathbb{R}^5} d\chi_1 \cdots d\chi_5 A_2(I) A_2(J) = \frac{1}{2\sqrt{2}} \exp\left(-\frac{(P_I - P_J)^2}{4\sigma_P^2} - \frac{(G_I - G_J)^2}{4\sigma_G^2}\right)$$
$$\sum_{i\neq I}^{n_I} \xi_2(d_{iI}) \sum_{j\neq J}^{n_J} \exp\left(-\frac{(d_{jJ} - d_{iI})^2}{4\sigma_d^2} - \frac{(P_i - P_j)^2}{4\sigma_P^2} - \frac{(G_i - G_j)^2}{4\sigma_G^2}\right) \xi_2(d_{jJ})$$

$$\frac{1}{\varsigma_3^2} \int_{\mathbb{R}^8} d\chi_1 \cdots d\chi_8 A_3(I) A_3(J) = \frac{1}{16} \exp\left(-\frac{(P_I - P_J)^2}{4\sigma_P^2} - \frac{(G_I - G_J)^2}{4\sigma_G^2}\right)$$
$$\sum_{i\neq I}^{n_I} \sum_{j\neq J}^{n_J} \exp\left(-\frac{(d_{jJ} - d_{iI})^2}{4\sigma_d^2} - \frac{(P_i - P_j)^2}{4\sigma_P^2} - \frac{(G_i - G_j)^2}{4\sigma_G^2}\right)$$
$$\sum_{k\neq i,I}^{n_I} \xi_2(d_{iI}, d_{kI}, \theta_{ik}^I) \sum_{l\neq j,J}^{n_J} \exp\left(-\frac{(\theta_{ik}^I - \theta_{jl}^J)^2}{4\sigma_\theta^2} - \frac{(P_k - P_l)^2}{4\sigma_P^2} - \frac{(G_k - G_l)^2}{4\sigma_G^2}\right) \xi_3(d_{jJ}, d_{lJ}, \theta_{jk}^J)$$



"Alchemical and structural distribution based representation for universal QML", Faber et al, *J Chem Phys* (2018), https://arxiv.org/abs/1712.08417

Learning curves for QM9 energies (pre 2018)



2017 FCHL: Faber, Christensen, Huang, Lilienfeld (Basel)
2017 aSLATM: Huang, Lilienfeld (Basel)
2017 SOAP multi-kernel active sampling: Bartok, Csanyi (Cambridge)/Ceriotti (EPFL)
2017 Const. Size: Yaron (Carnegie Mellon)
2017 enn-s2s: Gilmer et al (Google)
2017 DTNN: Schutt, Chmiela, Muller (TU Berlin)/Tkatchenko (Luxemburg)
2017 HDAD: Faber et al (Basel)/Riley et al (Google Accelerated Science)

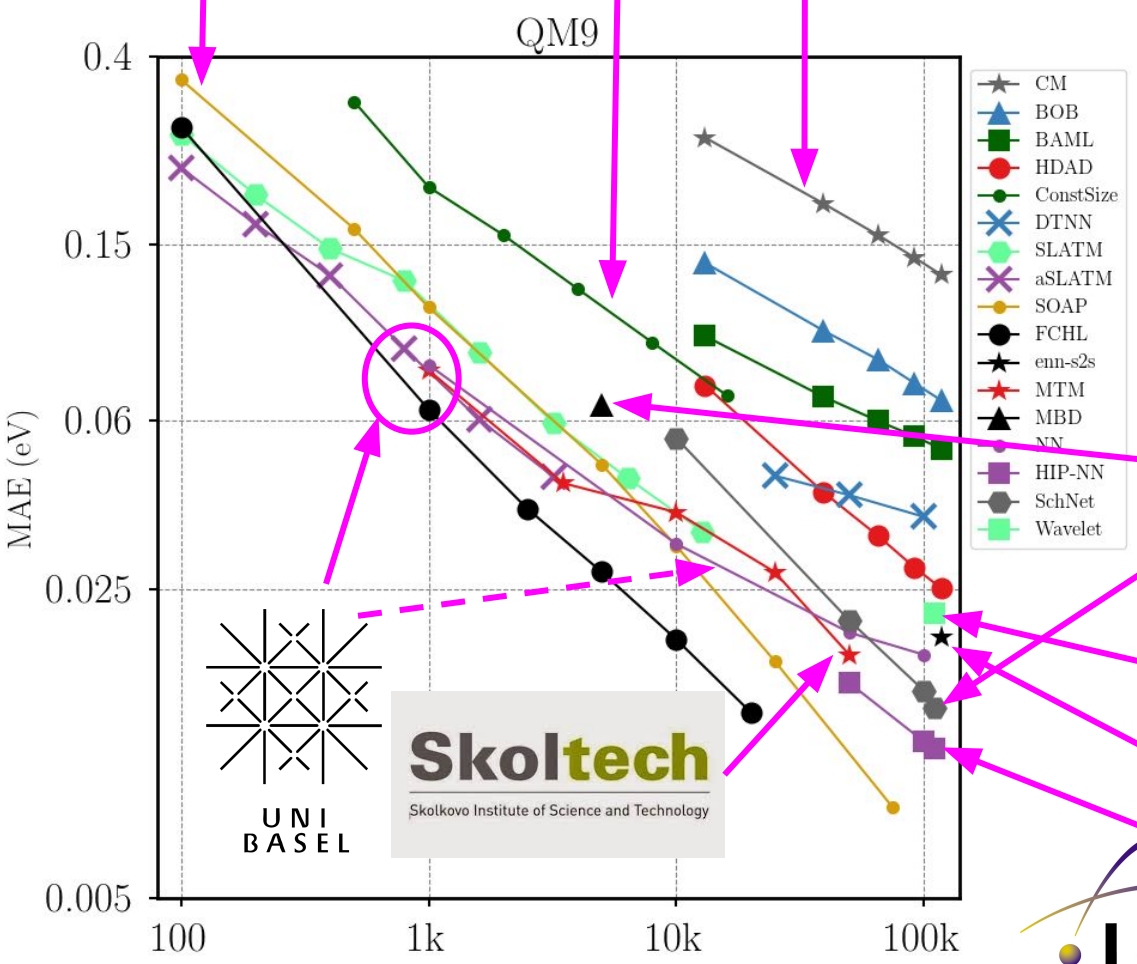
2016 BAML: Huang, Lilienfeld (Basel)

2015 BoB: Hansen, Tkatchenko (FHI)/Muller (TU Berlin)/Lilienfeld (Basel)

2014 QM9: Ramakrishnan, …, Lilienfeld (Basel). B3LYP for GDB-9 subset of GDB-17 from Reymond and co-workers (Berne)

2012 CM: Rupp, Tkatchenko, Muller, Lilienfeld (IPAM) (~10kcal/mol)

Learning curves for QM9 energies (current)



2018 MTM: Shapeev (Skolkovo)
2018 Wavelet: Eickenberg, Exarchakis, Thiry, Mallat (ENS), Hirn (Michigan)
2018 HIP-NN: Lubbers, Smith, Barros (Los Alamos National Lab)
2018 NN: Unke, Meuwly (Basel)
2018 SchNet: Schutt, Muller (TU Berlin)/Tkatchenko (Luxemburg)
2018 MBD: Pronobis, Muller (TU Berlin)/Tkatchenko (Luxemburg)

2017 FCHL: Faber, Christensen, Huang, Lilienfeld (Basel)
2017 aSLATM: Huang, Lilienfeld (Basel)
2017 SOAP multi-kernel active sampling: Bartok, Csanyi (Cambridge)/Ceriotti (EPFL)
2017 Const. Size: Yaron (Carnegie Mellon)
2017 enn-s2s: Gilmer et al (Google)
2017 DTNN: Schutt, Chmiela, Muller (TU Berlin)/Tkatchenko (Luxemburg)
2017 HDAD: Faber et al (Basel)/Riley et al (Google Accelerated Science)

2016 BAML: Huang, Lilienfeld (Basel)

2015 BoB: Hansen, Tkatchenko (FHI)/Muller (TU Berlin)/Lilienfeld (Basel)

2014 QM9: Ramakrishnan, …, Lilienfeld (Basel). B3LYP for GDB-9 subset of GDB-17 from Reymond and co-workers (Berne)

2012 CM: Rupp, Tkatchenko, Muller, Lilienfeld (IPAM) (~10kcal/mol)

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

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

CARNEGIE MELLON UNIVERSITY
PITTSBURGH PENNSYLVANIA 1900

UNI BASEL

Skoltech
Skolkovo Institute of Science and Technology

berlin

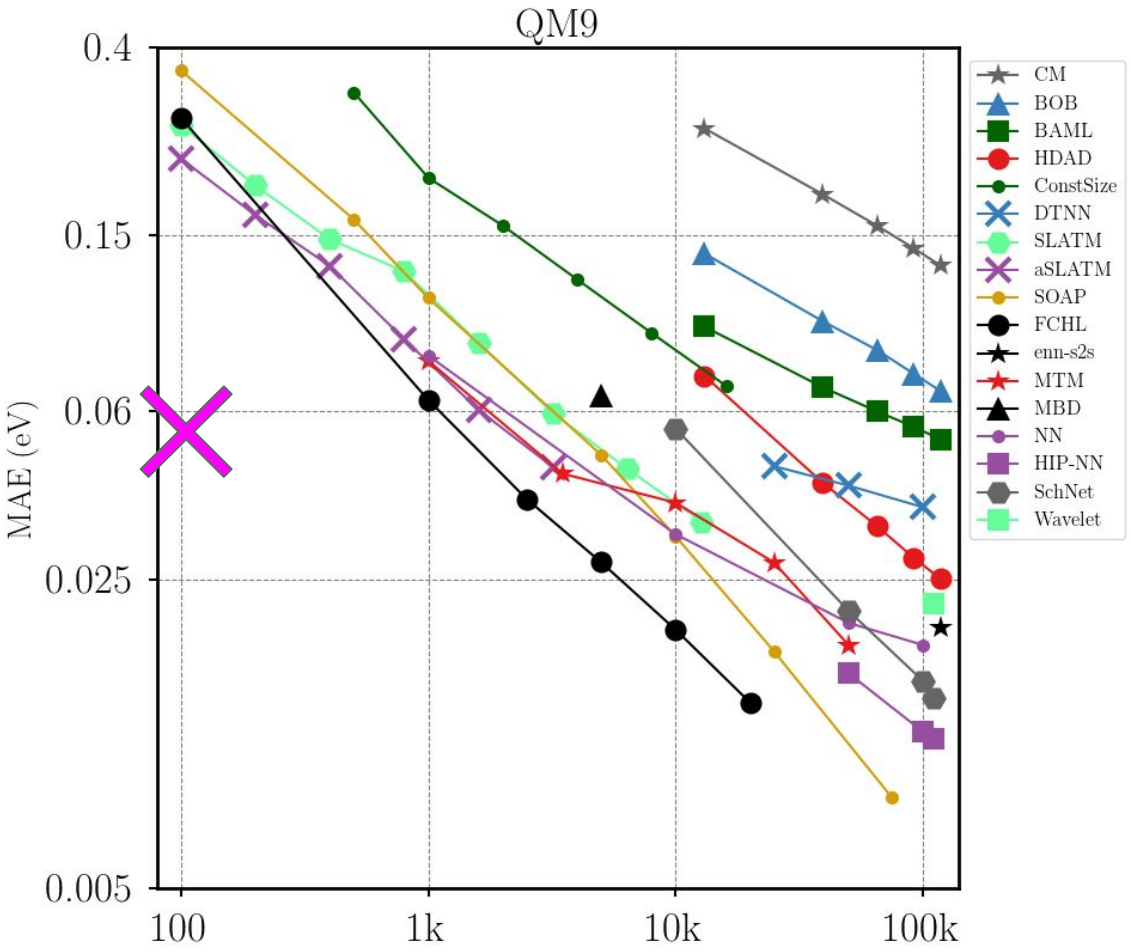
1794
ENS
ÉCOLE NORMALE SUPÉRIEURE

Research at Google

Los Alamos
NATIONAL LABORATORY

QM9 IPAM challenge

- Direct QML model (no fitting on other data-sets, milli-second execution time)
- Active learning/optimal training molecule selection & use of QM is desirable
- Reach chemical accuracy (MAE~0.05 eV) for U_0 of QM9 @ N = 100, or prove that it's impossible
- Present paper at 2nd reunion



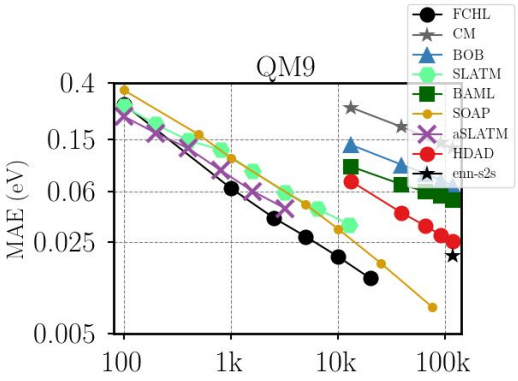
Decision

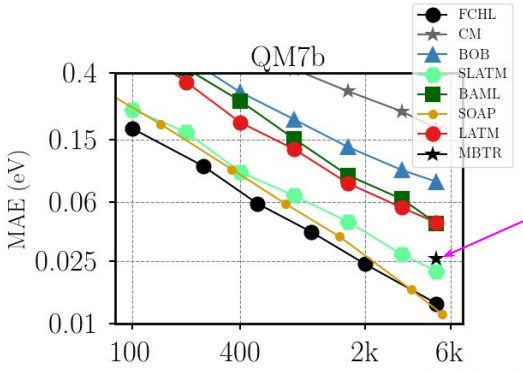
- Cash award ($100/Prof)
- Reproducible (no cheating)
- Unanimously agreed
- Shared by all winners

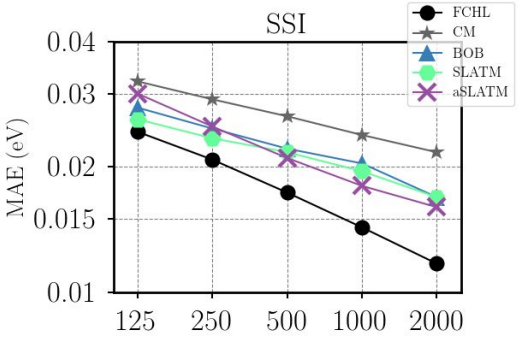
Panel: Anatole von Lilienfeld and

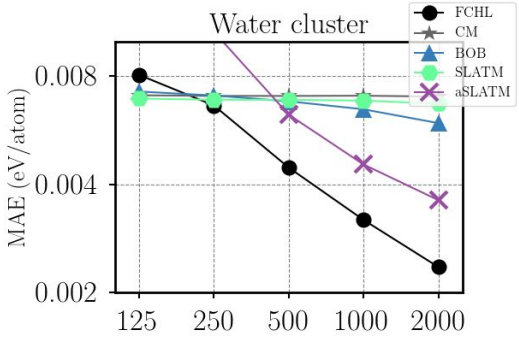
- Alexandre Tkatchenko
- Stefan Tautz
- Klaus Mueller
- Reinhard Maurer
- Frank Noe
- Cecilia Clementi
- Olexander Isayev
- Risi Kondor
- Alan Aspuru-Guzik
- Adrian Roitberg
- Jean-Philip Piquemal
- Clemence Corminboeuf
- …

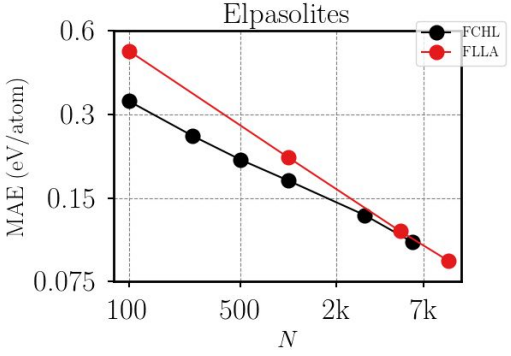
FCHL

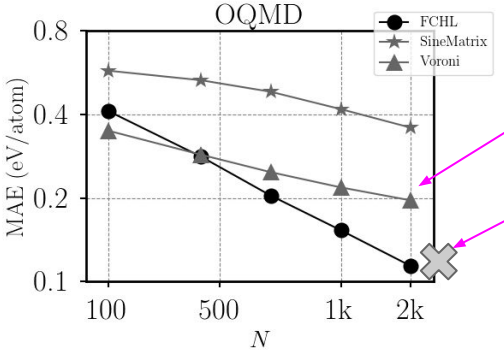












"Machine Learning Energies of 2 M Elpasolite (ABC2D6) Crystals", Faber et al, *Phys Rev Lett* (2016)

"Alchemical and structural distribution based representation for universal QML", Faber et al, *J Chem Phys* (2018)

"Unified representation for machine learning of molecules and crystals", Huo, Rupp, *arXiv preprint* arXiv:1704.06439

Fritz-Haber-Institut

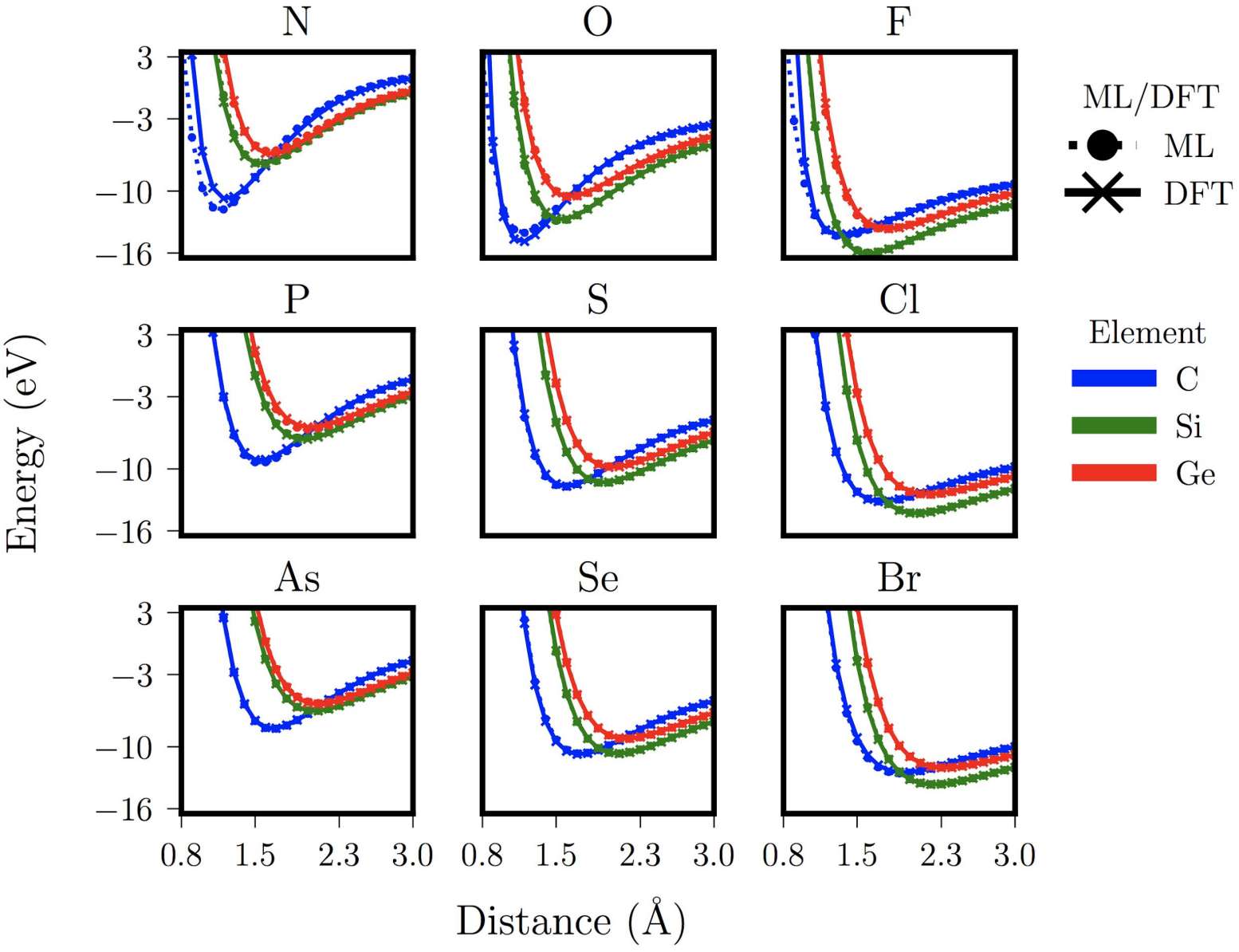
"Including crystal structure attributes in machine learning models of formation energies via Voronoi tessellations", Logan, Wolverton, *Phys Rev B* (2017)

Northwestern University

"SchNet - A deep learning architecture for molecules and materials", Schuett et al, *J Chem Phys* (2018)

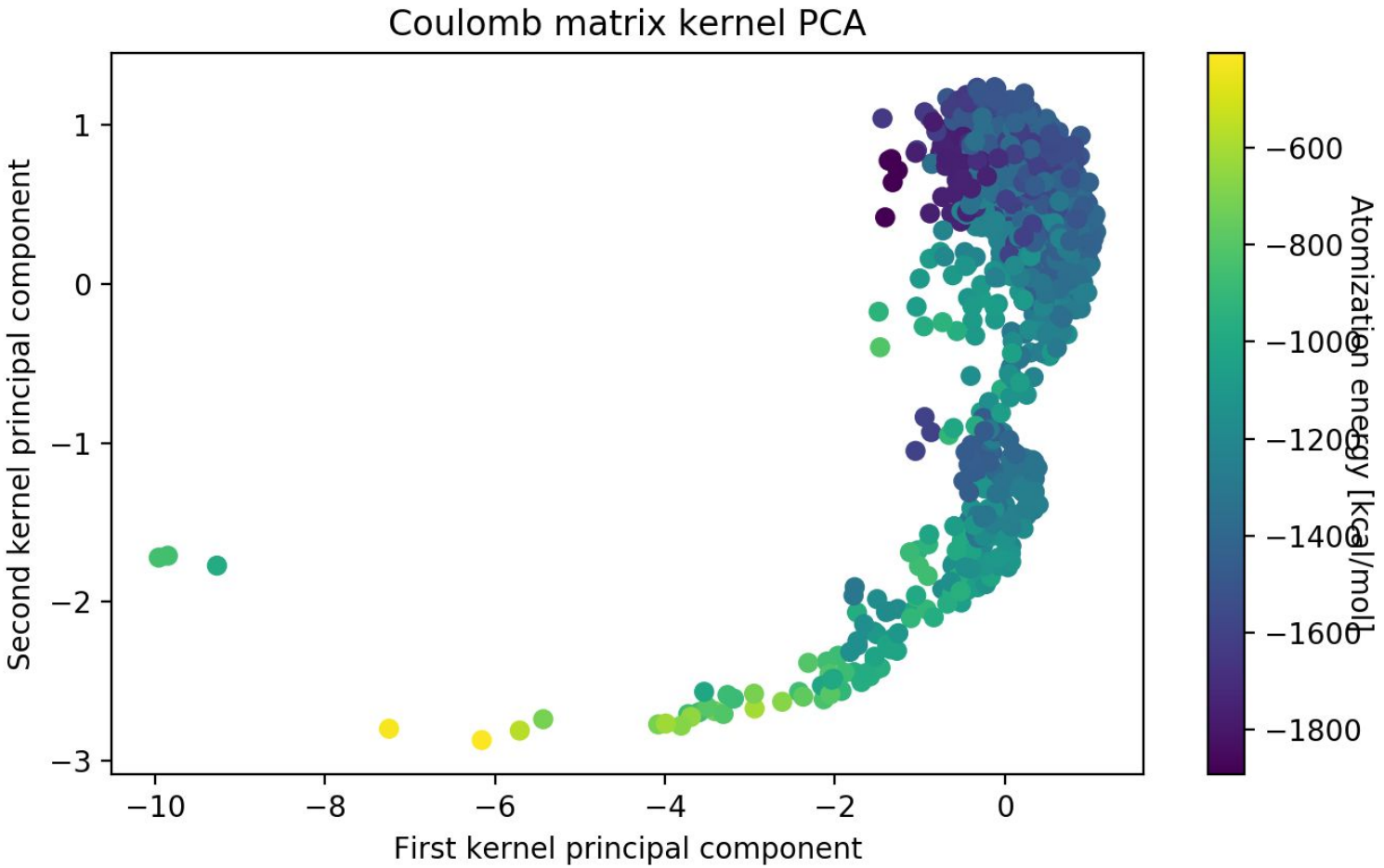
TU berlin

FCHL representation



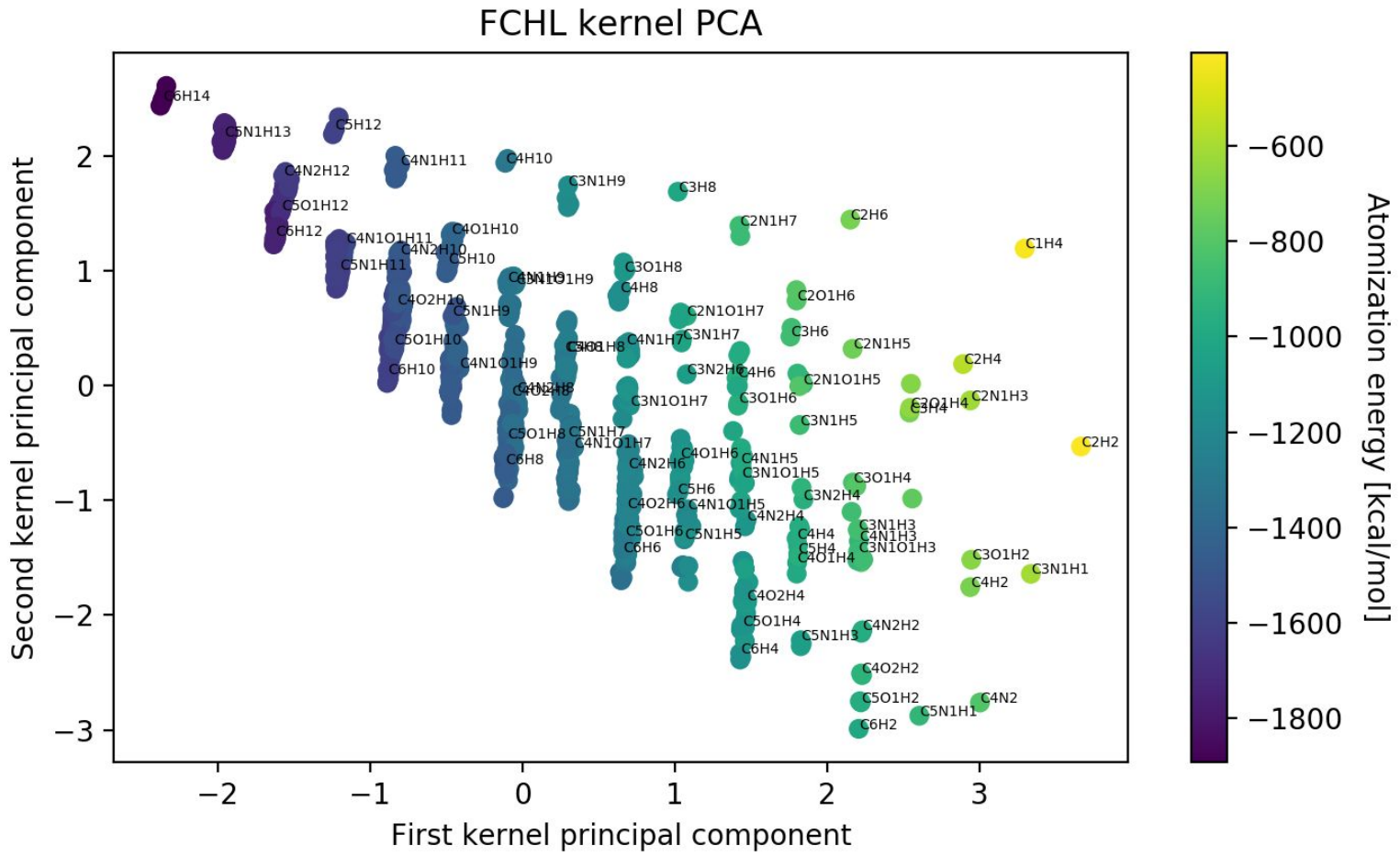
"Alchemical and structural distribution based representation for universal QML", Faber et al, *J Chem Phys* (2018), *https://arxiv.org/abs/1712.08417*

FCHL representation



"Alchemical and structural distribution based representation for universal QML", Faber et al, *J Chem Phys* (2018), *https://arxiv.org/abs/1712.08417*

FCHL representation



"Alchemical and structural distribution based representation for universal QML", Faber et al, *J Chem Phys* (2018), *https://arxiv.org/abs/1712.08417*

JCTC Journal of Chemical Theory and Computation

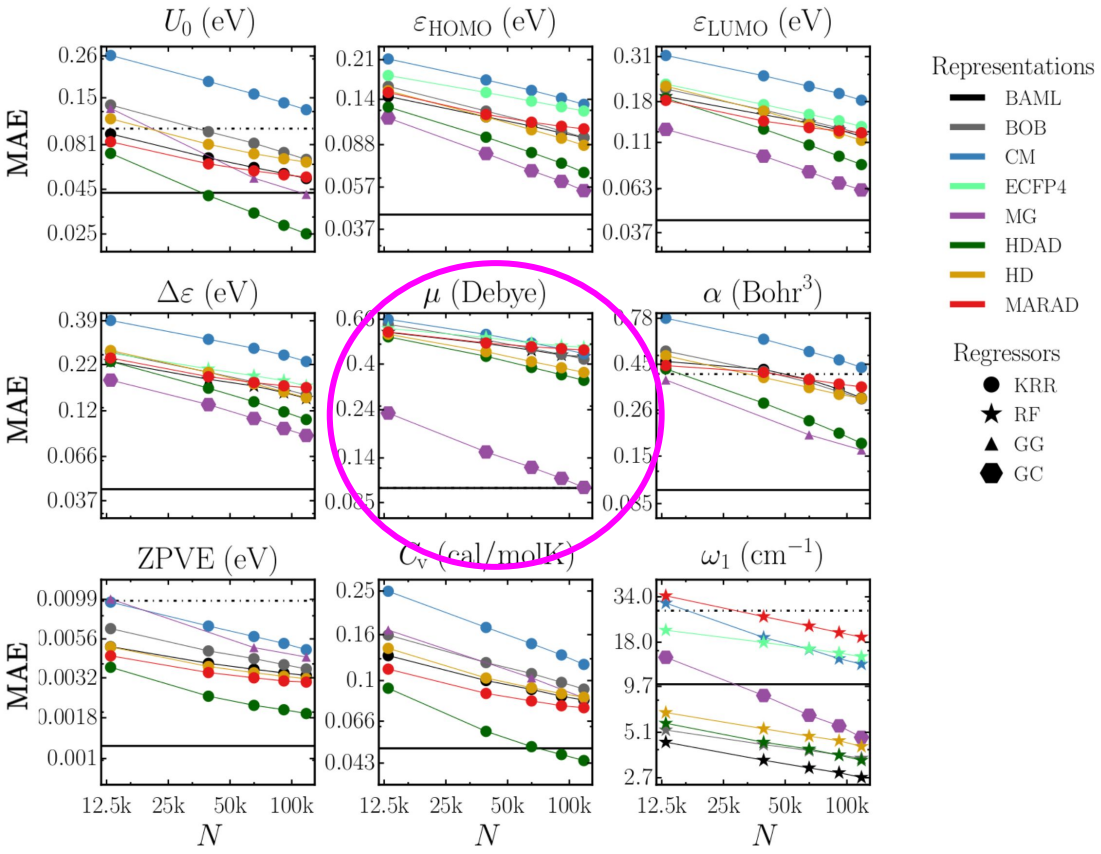
Article

Cite This: J. Chem. Theory Comput. 2017, 13, 5255-5264

pubs.acs.org/JCTC

Prediction Errors of Molecular Machine Learning Models Lower than Hybrid DFT Error

Felix A. Faber,[†] Luke Hutchison,[‡] Bing Huang,[†] Justin Gilmer,[‡] Samuel S. Schoenholz,[‡] George E. Dahl,[‡] Oriol Vinyals,[¶] Steven Kearnes,[‡] Patrick F. Riley,[‡] and O. Anatole von Lilienfeld*[,†]



Research at Google

OQML: response properties

$$\mathbf{U} = \mathbf{K}\boldsymbol{\alpha} \qquad\qquad \omega = \mathcal{O}[\mathbf{U}] = \mathcal{O}[\mathbf{K}]\boldsymbol{\alpha}$$

$$J(\boldsymbol{\alpha}) = \sum_{\gamma} \beta_{\gamma} \| \mathcal{O}_{\gamma}[\mathbf{U}^{\mathrm{ref}}] - \mathcal{O}_{\gamma}[\mathbf{K}\boldsymbol{\alpha}] \|^2_{L_2(\Omega_{\gamma})}$$

$$\equiv \sum_{\gamma} \beta_{\gamma} \int_{\Omega_{\gamma}} \Big[\mathcal{O}_{\gamma}[\mathbf{U}^{\mathrm{ref}}] - \mathcal{O}_{\gamma}[\mathbf{K}\boldsymbol{\alpha}] \Big]^T \Big[\mathcal{O}_{\gamma}[\mathbf{U}^{\mathrm{ref}}] - \mathcal{O}_{\gamma}[\mathbf{K}\boldsymbol{\alpha}] \Big]$$

"Operators in Machine Learning: Response Properties in Chemical Space", Christensen et al, *J Chem Phys* (2019), arxiv.org/abs/1807.08811

OQML: response properties

$$\mathbf{U} = \mathbf{K}\boldsymbol{\alpha} \qquad\qquad \omega = \mathcal{O}[\mathbf{U}] = \mathcal{O}[\mathbf{K}]\boldsymbol{\alpha}$$

$$J(\boldsymbol{\alpha}) = \sum_{\gamma} \beta_{\gamma} \| \mathcal{O}_{\gamma}[\mathbf{U}^{\mathrm{ref}}] - \mathcal{O}_{\gamma}[\mathbf{K}\boldsymbol{\alpha}] \|^2_{L_2(\Omega_{\gamma})}$$

$$\equiv \sum_{\gamma} \beta_{\gamma} \int_{\Omega_{\gamma}} \Big[\mathcal{O}_{\gamma}[\mathbf{U}^{\mathrm{ref}}] - \mathcal{O}_{\gamma}[\mathbf{K}\boldsymbol{\alpha}] \Big]^T \Big[\mathcal{O}_{\gamma}[\mathbf{U}^{\mathrm{ref}}] - \mathcal{O}_{\gamma}[\mathbf{K}\boldsymbol{\alpha}] \Big]$$

$$\boldsymbol{\alpha} = \Big[\sum_{\gamma} \beta_{\gamma} \int_{\Omega_{\gamma}} \mathcal{O}_{\gamma}[\mathbf{K}]^T \mathcal{O}_{\gamma}[\mathbf{K}] \Big]^{-1} \Big[\sum_{\gamma} \beta_{\gamma} \int_{\Omega_{\gamma}} \mathcal{O}_{\gamma}[\mathbf{U}^{\mathrm{ref}}]^T \mathcal{O}_{\gamma}[\mathbf{K}] \Big]$$

"Operators in Machine Learning: Response Properties in Chemical Space", Christensen et al, *J Chem Phys* (2019), arxiv.org/abs/1807.08811

OQML: response properties

$$\int_{\Omega_{\delta\vec{\eta}}} \mathcal{O}_{\delta\vec{\eta}}[\mathbf{K}]^T \mathcal{O}_{\delta\vec{\eta}}[\mathbf{K}] = \frac{1}{3} \sum_{\nu \in x,y,z} \left(\frac{\partial}{\partial \eta_k}\mathbf{K}\right)^T \left(\frac{\partial}{\partial \eta_\nu}\mathbf{K}\right)$$

$$\int_{\Omega_{\delta\vec{\eta}}} \mathcal{O}_{\delta\vec{\eta}}[\mathbf{U}^{\mathrm{ref}}]^T \mathcal{O}_{\delta\vec{\eta}}[\mathbf{K}] = \frac{1}{3} \sum_{\nu \in x,y,z} \left(\frac{\partial}{\partial \eta_\nu}\mathbf{K}\right)^T \left(\frac{\partial}{\partial \eta_\nu}\mathbf{U}^{\mathrm{ref}}\right)$$

$$\boldsymbol{\alpha} = \left[\sum_\gamma \beta_\gamma \int_{\Omega_\gamma} \mathcal{O}_\gamma[\mathbf{K}]^T \mathcal{O}_\gamma[\mathbf{K}]\right]^{-1} \left[\sum_\gamma \beta_\gamma \int_{\Omega_\gamma} \mathcal{O}_\gamma[\mathbf{U}^{\mathrm{ref}}]^T \mathcal{O}_\gamma[\mathbf{K}]\right]$$

"Operators in Machine Learning: Response Properties in Chemical Space", Christensen et al, *J Chem Phys* (2019), arxiv.org/abs/1807.08811

OQML: response properties

$$\int_{\Omega_{\delta\vec{\eta}}} \mathcal{O}_{\delta\vec{\eta}}[\mathbf{K}]^T \mathcal{O}_{\delta\vec{\eta}}[\mathbf{K}] = \frac{1}{3} \sum_{\nu \in x,y,z} \left(\frac{\partial}{\partial \eta_k}\mathbf{K}\right)^T \left(\frac{\partial}{\partial \eta_\nu}\mathbf{K}\right)$$

$$\int_{\Omega_{\delta\vec{\eta}\delta\vec{\eta'}}} \mathcal{O}_{\delta\vec{\eta}\delta\vec{\eta'}}[\mathbf{K}]^T \mathcal{O}_{\delta\vec{\eta}\delta\vec{\eta'}}[\mathbf{K}]$$

$$= \frac{1}{9} \sum_{\nu,\nu' \in x,y,z} \left(\frac{\partial^2}{\partial \eta_\nu \partial \eta'_{\nu'}}\mathbf{K}\right)^T \left(\frac{\partial^2}{\partial \eta_\nu \partial \eta'_{\nu'}}\mathbf{K}\right)$$

$$\int_{\Omega_{\delta\vec{\eta}}} \mathcal{O}_{\delta\vec{\eta}}[\mathbf{U}^{\mathrm{ref}}]^T \mathcal{O}_{\delta\vec{\eta}}[\mathbf{K}] = \frac{1}{3} \sum_{\nu \in x,y,z} \left(\frac{\partial}{\partial \eta_\nu}\mathbf{K}\right)^T \left(\frac{\partial}{\partial \eta_\nu}\mathbf{U}^{\mathrm{ref}}\right)$$

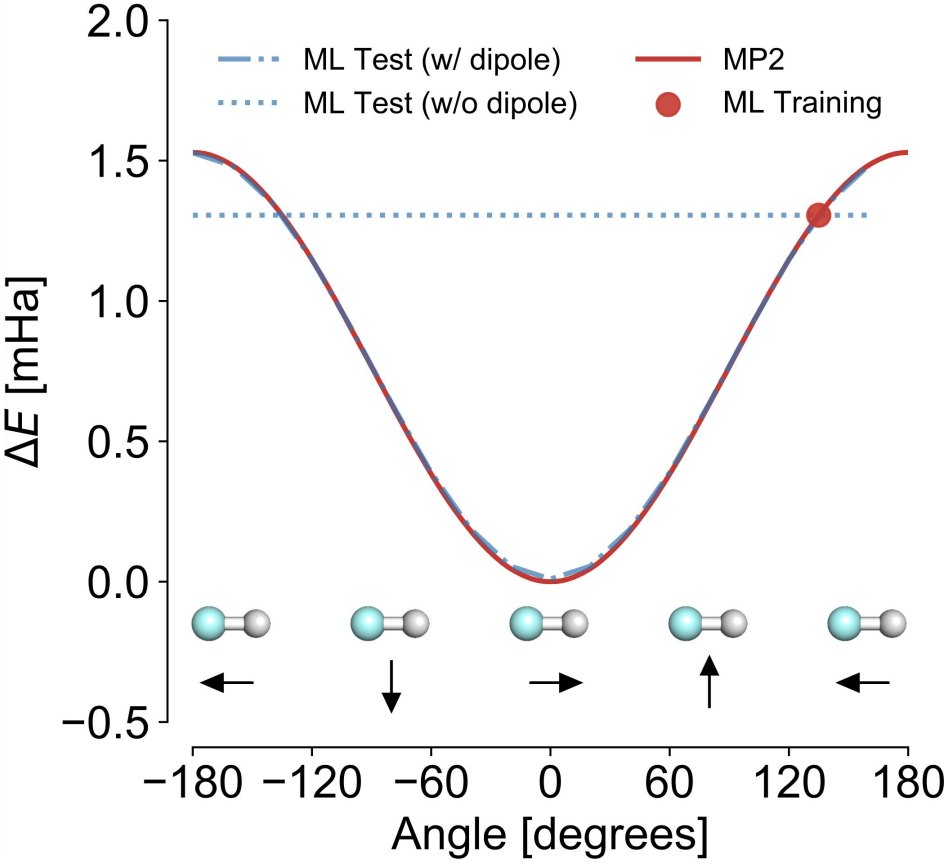
$$\int_{\Omega_{\delta\vec{\eta}\delta\vec{\eta'}}} \mathcal{O}_{\delta\vec{\eta}\delta\vec{\eta'}}[\mathbf{U}^{\mathrm{ref}}]^T \mathcal{O}_{\delta\vec{\eta}\delta\vec{\eta'}}[\mathbf{K}]$$

$$= \frac{1}{9} \sum_{\nu,\nu' \in x,y,z} \left(\frac{\partial^2}{\partial \eta_\nu \partial \eta'_{\nu'}}\mathbf{U}^{\mathrm{ref}}\right)^T \left(\frac{\partial^2}{\partial \eta_\nu \partial \eta'_{\nu'}}\mathbf{K}\right)$$

$$\boldsymbol{\alpha} = \left[\sum_\gamma \beta_\gamma \int_{\Omega_\gamma} \mathcal{O}_\gamma[\mathbf{K}]^T \mathcal{O}_\gamma[\mathbf{K}]\right]^{-1} \left[\sum_\gamma \beta_\gamma \int_{\Omega_\gamma} \mathcal{O}_\gamma[\mathbf{U}^{\mathrm{ref}}]^T \mathcal{O}_\gamma[\mathbf{K}]\right]$$

"Operators in Machine Learning: Response Properties in Chemical Space", Christensen et al, *J Chem Phys* (2019), arxiv.org/abs/1807.08811

OQML: Dipole moment = electric field derivative

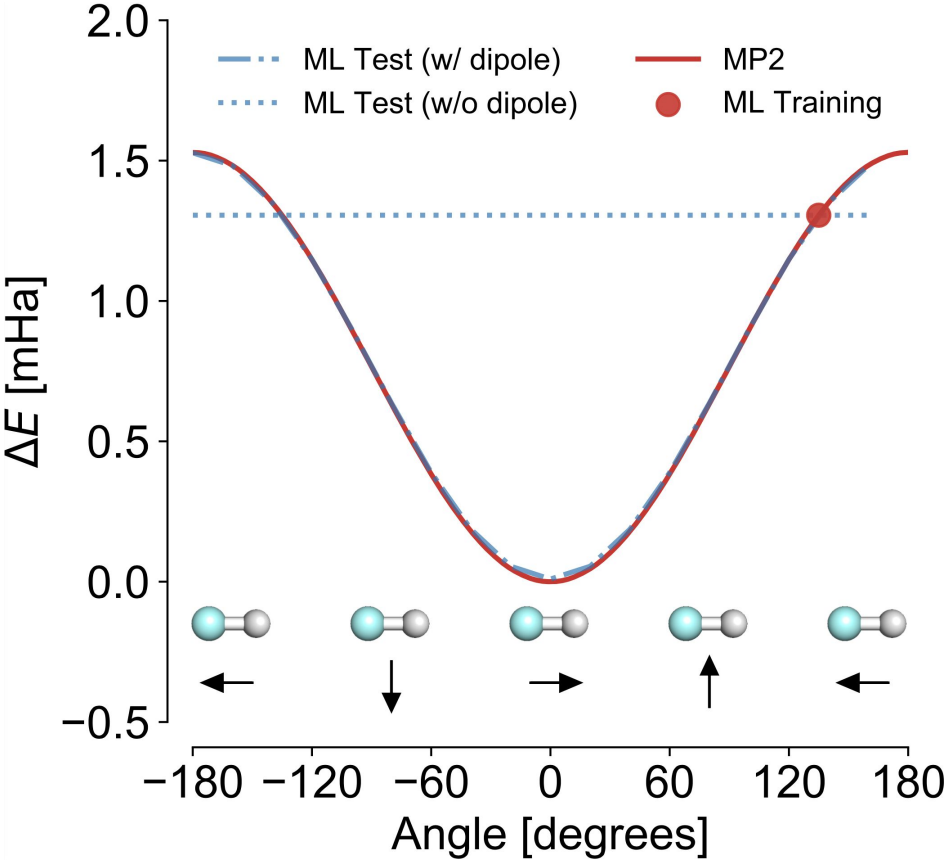


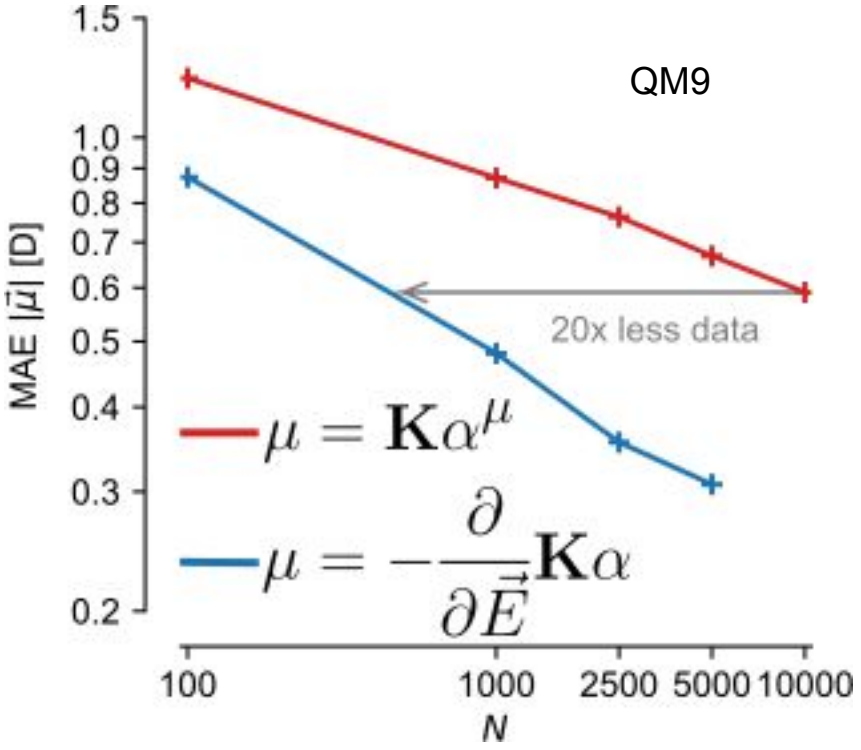
Must use extended FCHL* to account for electric field derivative: Update \xi-functions

$$\xi_2^{*IJ} = \xi_2^{IJ} - \epsilon(\vec{\mu}_{IJ} \cdot \vec{E}) \qquad \xi_3^{*IJK} = \xi_3^{IJK} - \epsilon(\vec{\mu}_{IJK} \cdot \vec{E})$$

"Operators in Machine Learning: Response Properties in Chemical Space", Christensen et al, *J Chem Phys* (2019), arxiv.org/abs/1807.08811

OQML: Dipole moment = electric field derivative



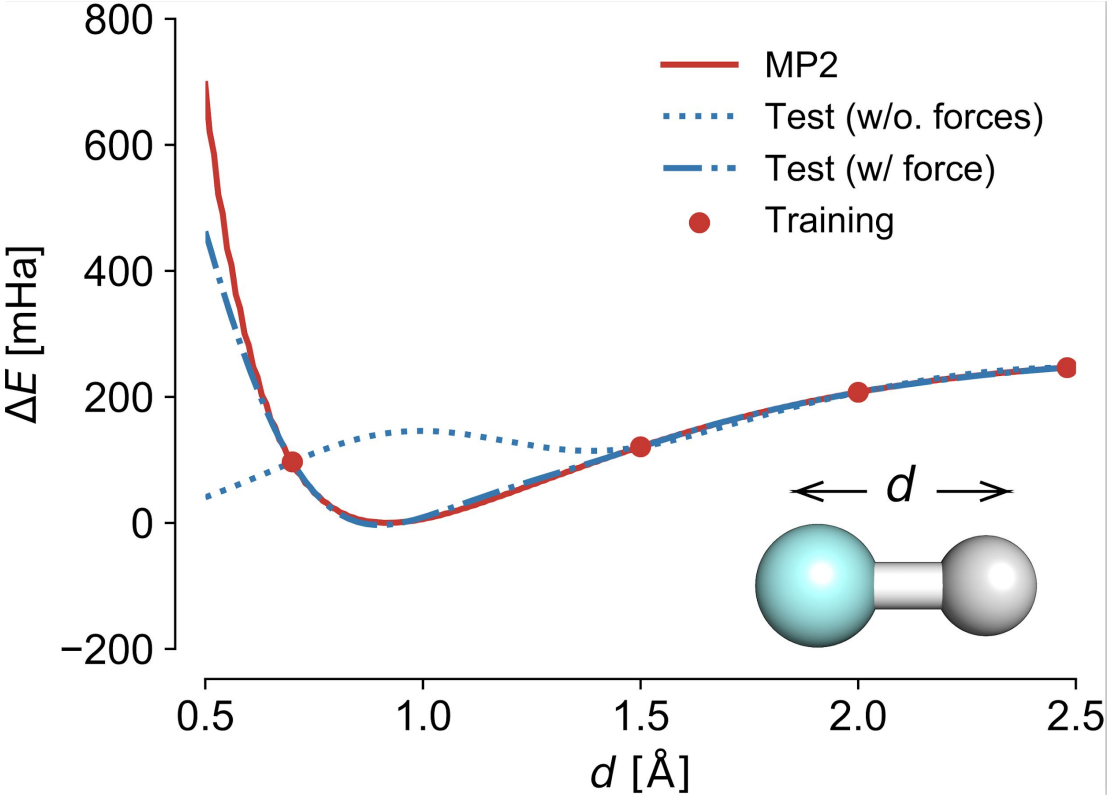


Must use extended FCHL* to account for electric field derivative: Update \xi-functions

$$\xi_2^{*IJ} = \xi_2^{IJ} - \epsilon(\vec{\mu}_{IJ} \cdot \vec{E}) \qquad \xi_3^{*IJK} = \xi_3^{IJK} - \epsilon(\vec{\mu}_{IJK} \cdot \vec{E})$$

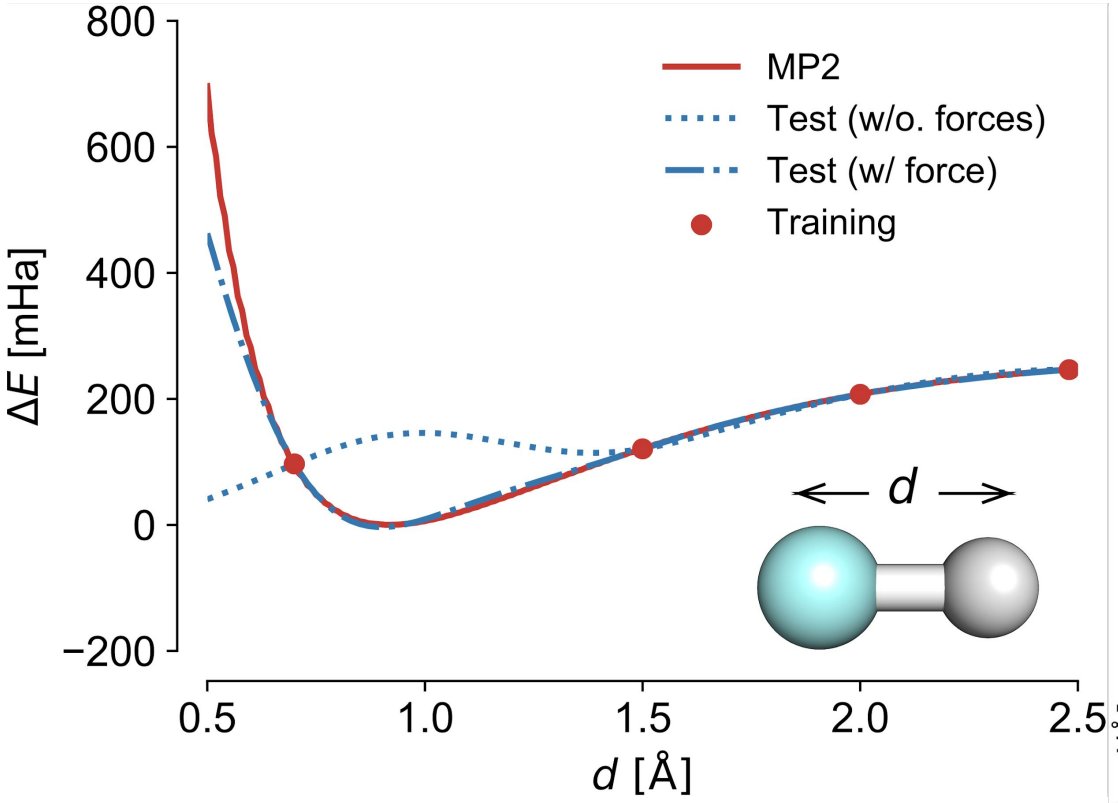
"Operators in Machine Learning: Response Properties in Chemical Space", Christensen et al, *J Chem Phys* (2019), arxiv.org/abs/1807.08811

OQML: Forces = Atomic displacement derivatives

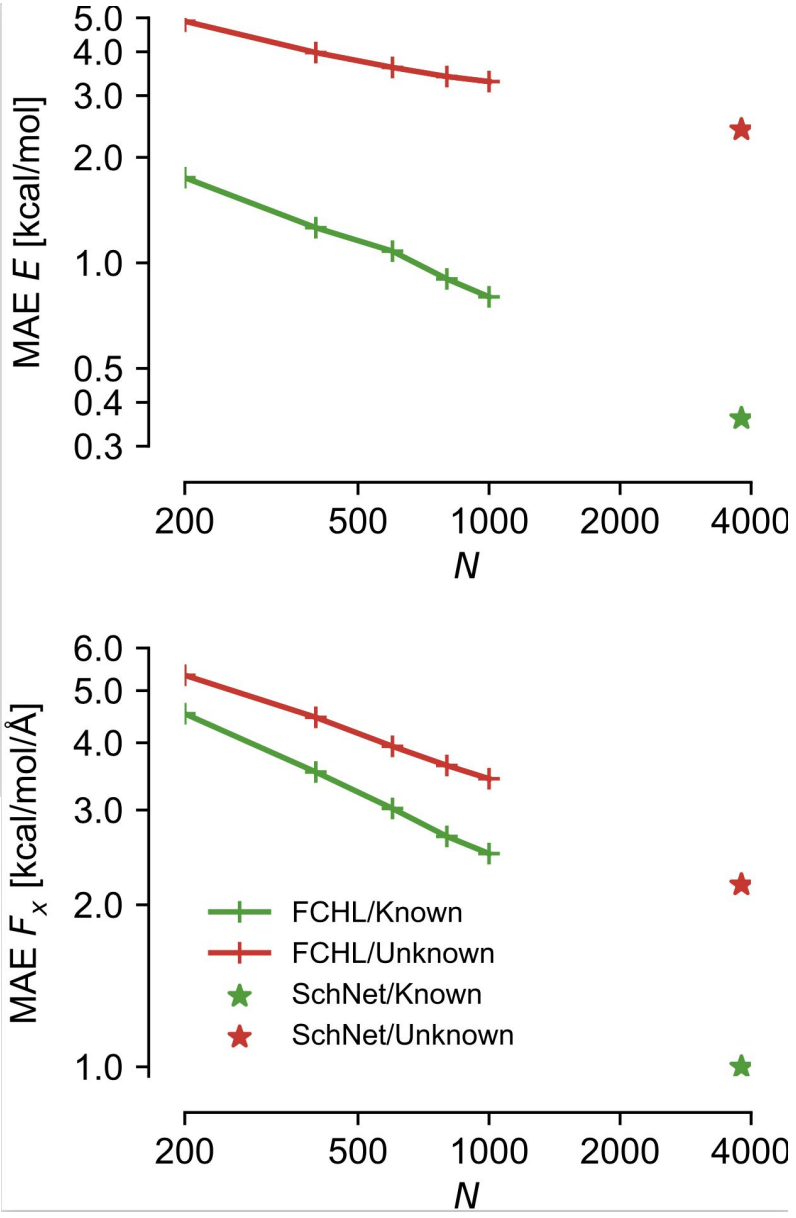


"Operators in Machine Learning: Response Properties in Chemical Space", Christensen et al, *J Chem Phys* (2019), arxiv.org/abs/1807.08811

OQML: Forces = Atomic displacement derivatives



ISO17 & SchNet: Schutt et al, *J Chem Phys* (2018)



"Operators in Machine Learning: Response Properties in Chemical Space", Christensen et al, *J Chem Phys* (2019), arxiv.org/abs/1807.08811

OQML forces vs. GPR forces vs. GDML forces

$$\begin{bmatrix} \mathbf{U} \\ \mathbf{F} \end{bmatrix} = \begin{bmatrix} \mathbf{K}^{\mathrm{OQML}} \\ -\frac{\partial}{\partial \vec{r}^{*}} \mathbf{K}^{\mathrm{OQML}} \end{bmatrix} \boldsymbol{\alpha}^{\mathrm{OQML}}$$

$$\mathbf{K}^{\mathrm{OQML}}_{Ji} = \sum_{I \in i} k(\mathbf{q}_J, \mathbf{q}^{*}_I)$$

$$\left(-\frac{\partial}{\partial \vec{r}^{*}} \mathbf{K}^{\mathrm{OQML}} \right)_{JiK} = -\sum_{I \in i} \frac{\partial k(\mathbf{q}_J, \mathbf{q}^{*}_I)}{\partial \vec{r}^{*}_{iK}}$$

"Operators in Machine Learning: Response Properties in Chemical Space", Christensen et al, *J Chem Phys* (2019), arxiv.org/abs/1807.08811

OQML forces vs. GPR forces vs. GDML forces

$$\begin{bmatrix} \mathbf{U} \\ \mathbf{F} \end{bmatrix} = \begin{bmatrix} \mathbf{K}^{\mathrm{OQML}} \\ -\frac{\partial}{\partial \vec{r}^*}\mathbf{K}^{\mathrm{OQML}} \end{bmatrix} \boldsymbol{\alpha}^{\mathrm{OQML}}$$

$$\mathbf{K}^{\mathrm{OQML}}_{Ji} = \sum_{I \in i} k(\mathbf{q}_J, \mathbf{q}^*_I)$$

$$\left(-\frac{\partial}{\partial \vec{r}^*}\mathbf{K}^{\mathrm{OQML}}\right)_{JiK} = -\sum_{I \in i} \frac{\partial k(\mathbf{q}_J, \mathbf{q}^*_I)}{\partial \vec{r}^*_{iK}}$$

$$\begin{bmatrix} \mathbf{U} \\ \mathbf{F} \end{bmatrix} = \begin{bmatrix} \mathbf{K}^{\mathrm{KRR}} & -\frac{\partial}{\partial \vec{r}^*}\mathbf{K}^{\mathrm{KRR}} \\ -\frac{\partial}{\partial \vec{r}}\mathbf{K}^{\mathrm{KRR}} & \frac{\partial^2}{\partial \vec{r}\partial \vec{r}^*}\mathbf{K}^{\mathrm{KRR}} \end{bmatrix} \boldsymbol{\alpha}^{\mathrm{GPR}}$$

$$\mathbf{K}^{\mathrm{KRR}}_{ij} = \sum_{I \in i}\sum_{J \in j} k(\mathbf{q}_J, \mathbf{q}^*_I)$$

$$\left(-\frac{\partial}{\partial \vec{r}^*}\mathbf{K}^{\mathrm{KRR}}\right)_{jiK} = -\sum_{I \in i}\sum_{J \in j} \frac{\partial k(\mathbf{q}_J, \mathbf{q}^*_I)}{\partial \vec{r}^*_{iK}}$$

$$\left(\frac{\partial^2}{\partial \vec{r}\partial \vec{r}^*}\mathbf{K}^{\mathrm{KRR}}\right)_{jLiK} = \sum_{I \in i}\sum_{J \in j} \frac{\partial k(\mathbf{q}_J, \mathbf{q}^*_I)}{\partial \vec{r}_{jL}\partial \vec{r}^*_{iK}}$$

S. Mathias, *A Kernel-Based Learning Method for an efficient Approxi- mation of the high-dimensional Born-Oppenheimer Potential Energy Sur- face*, Master's thesis, Mathematisch-Naturwissenschaftliche Fakultät der- Rheinischen Friedrich-Wilhelms-Universität Bonn, Germany (2015)

A. Bartok, G. Csanyi, *Gaussian Approximation Potentials: A Brief Tutorial Introduction,* Int. J. Quantum Chem. (2015)

"Operators in Machine Learning: Response Properties in Chemical Space", Christensen et al, *J Chem Phys* (2019), arxiv.org/abs/1807.08811

OQML forces vs. GPR forces vs. GDML forces

$$\begin{bmatrix} \mathbf{U} \\ \mathbf{F} \end{bmatrix} = \begin{bmatrix} \mathbf{K}^{\mathrm{OQML}} \\ -\frac{\partial}{\partial \vec{r}^*}\mathbf{K}^{\mathrm{OQML}} \end{bmatrix} \boldsymbol{\alpha}^{\mathrm{OQML}}$$

$$\mathbf{K}^{\mathrm{OQML}}_{Ji} = \sum_{I \in i} k(\mathbf{q}_J, \mathbf{q}^*_I)$$

$$\left(-\frac{\partial}{\partial \vec{r}^*}\mathbf{K}^{\mathrm{OQML}}\right)_{JiK} = -\sum_{I \in i} \frac{\partial k(\mathbf{q}_J, \mathbf{q}^*_I)}{\partial \vec{r}^*_{iK}}$$

$$\begin{bmatrix} \mathbf{U} \\ \mathbf{F} \end{bmatrix} = \begin{bmatrix} \mathbf{K}^{\mathrm{KRR}} & -\frac{\partial}{\partial \vec{r}^*}\mathbf{K}^{\mathrm{KRR}} \\ -\frac{\partial}{\partial \vec{r}}\mathbf{K}^{\mathrm{KRR}} & \frac{\partial^2}{\partial \vec{r}\partial \vec{r}^*}\mathbf{K}^{\mathrm{KRR}} \end{bmatrix} \boldsymbol{\alpha}^{\mathrm{GPR}}$$

$$\mathbf{K}^{\mathrm{KRR}}_{ij} = \sum_{I \in i}\sum_{J \in j} k(\mathbf{q}_J, \mathbf{q}^*_I)$$

$$\mathbf{F} = \frac{\partial^2}{\partial \vec{r}\partial \vec{r}^*}\mathbf{K}^{\mathrm{KRR}}\boldsymbol{\alpha}^{\mathrm{GDML}}$$

$$\left(-\frac{\partial}{\partial \vec{r}^*}\mathbf{K}^{\mathrm{KRR}}\right)_{jiK} = -\sum_{I \in i}\sum_{J \in j} \frac{\partial k(\mathbf{q}_J, \mathbf{q}^*_I)}{\partial \vec{r}^*_{iK}}$$

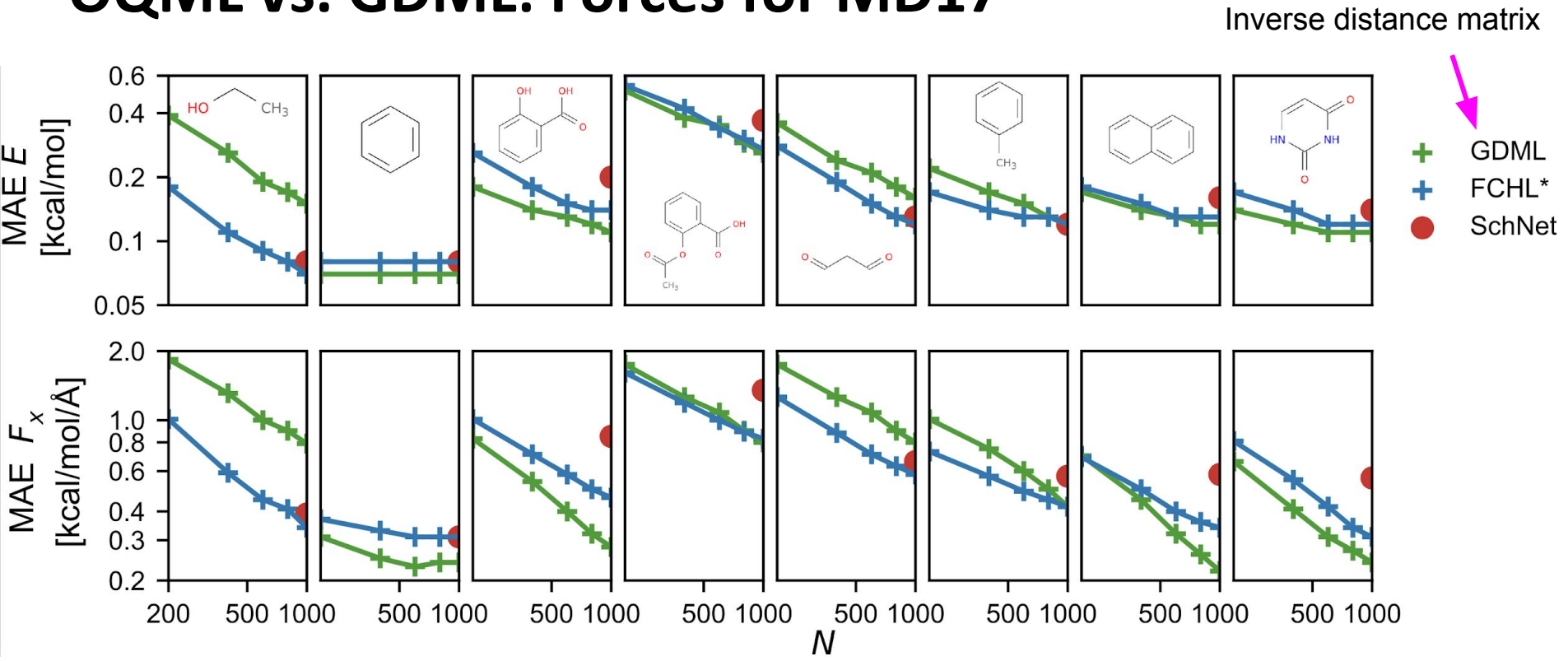
$$\mathbf{U} = -\frac{\partial}{\partial \vec{r}}\mathbf{K}^{\mathrm{KRR}}\boldsymbol{\alpha}^{\mathrm{GDML}}$$

$$\left(\frac{\partial^2}{\partial \vec{r}\partial \vec{r}^*}\mathbf{K}^{\mathrm{KRR}}\right)_{jLiK} = \sum_{I \in i}\sum_{J \in j} \frac{\partial k(\mathbf{q}_J, \mathbf{q}^*_I)}{\partial \vec{r}_{jL}\partial \vec{r}^*_{iK}}$$

MD17 & GDML: Chmiela et al, *Sci. Adv.* (2017)

"Operators in Machine Learning: Response Properties in Chemical Space", Christensen et al, *J Chem Phys* (2019), arxiv.org/abs/1807.08811

OQML vs. GDML: Forces for MD17



MD17 & GDML: Chmiela et al, *Sci. Adv.* (2017)

"Operators in Machine Learning: Response Properties in Chemical Space", Christensen et al, *J Chem Phys* (2019), arxiv.org/abs/1807.08811

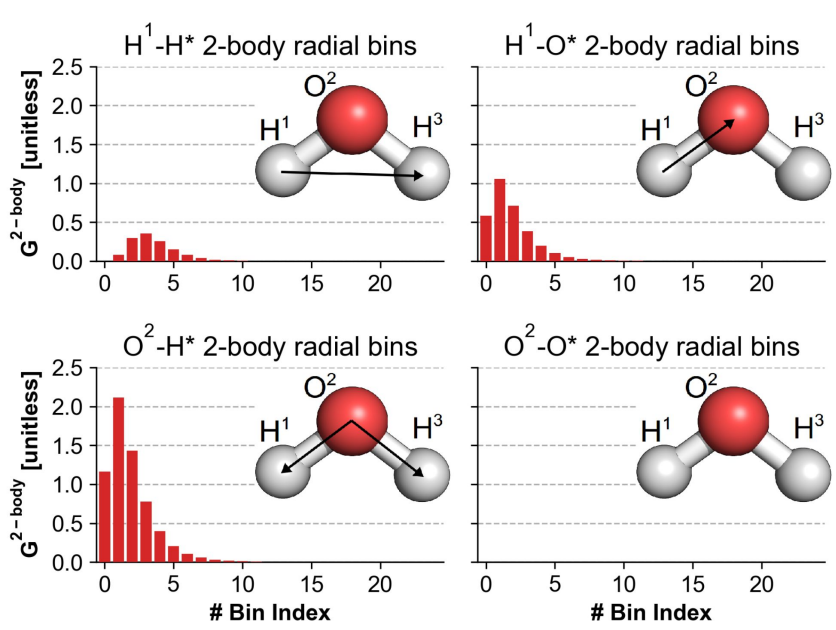
FCHL19 revision

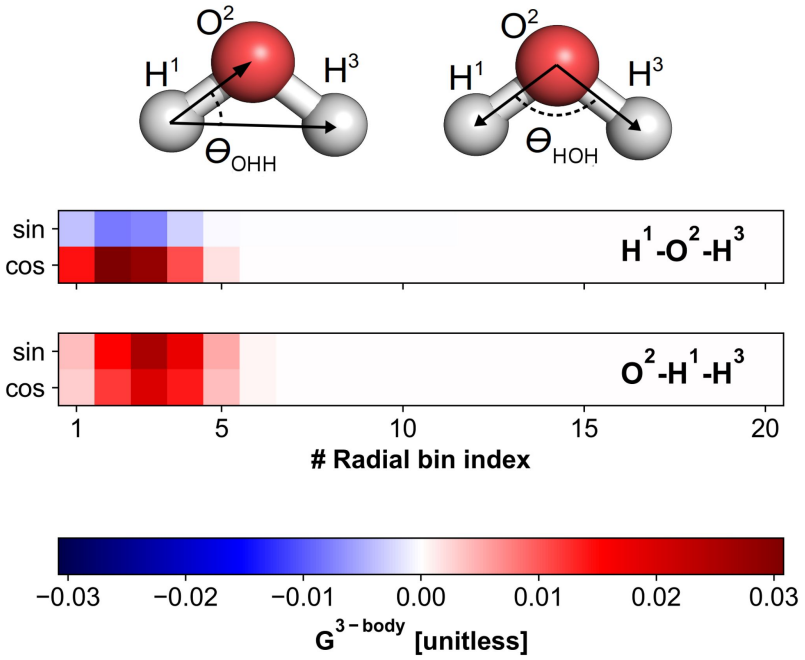
$$\mathcal{A}_M(I) \;=\; \{A_1(I), A_2(I), A_3(I), \ldots, A_M(I)\}$$

$$A_1(I) = \mathcal{N}(\mathbf{x}_I^{(1)}) \;=\; e^{-\frac{(P_I-\chi_1)^2}{2\sigma_P^2} - \frac{(G_I-\chi_2)^2}{2\sigma_G^2}}$$

$$A_2(I) \quad = \quad \mathcal{N}(\mathbf{x}_I^{(1)}) \sum_{i\neq I} \mathcal{N}(\mathbf{x}_{iI}^{(2)}) \xi_2(d_{iI})$$

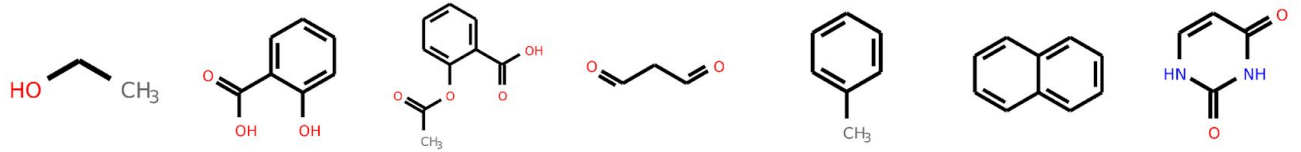
$$\Delta(\mathcal{A}_M(I), \mathcal{A}_M(J))^2 \equiv \sum_{m=0}^{M} \beta_m \frac{1}{\varsigma_m^2} \int_{\mathbb{R}^{3m+2}} (A_m(I) - A_m(J))^2 d\chi_1 \cdots d\chi_{3m+2}$$

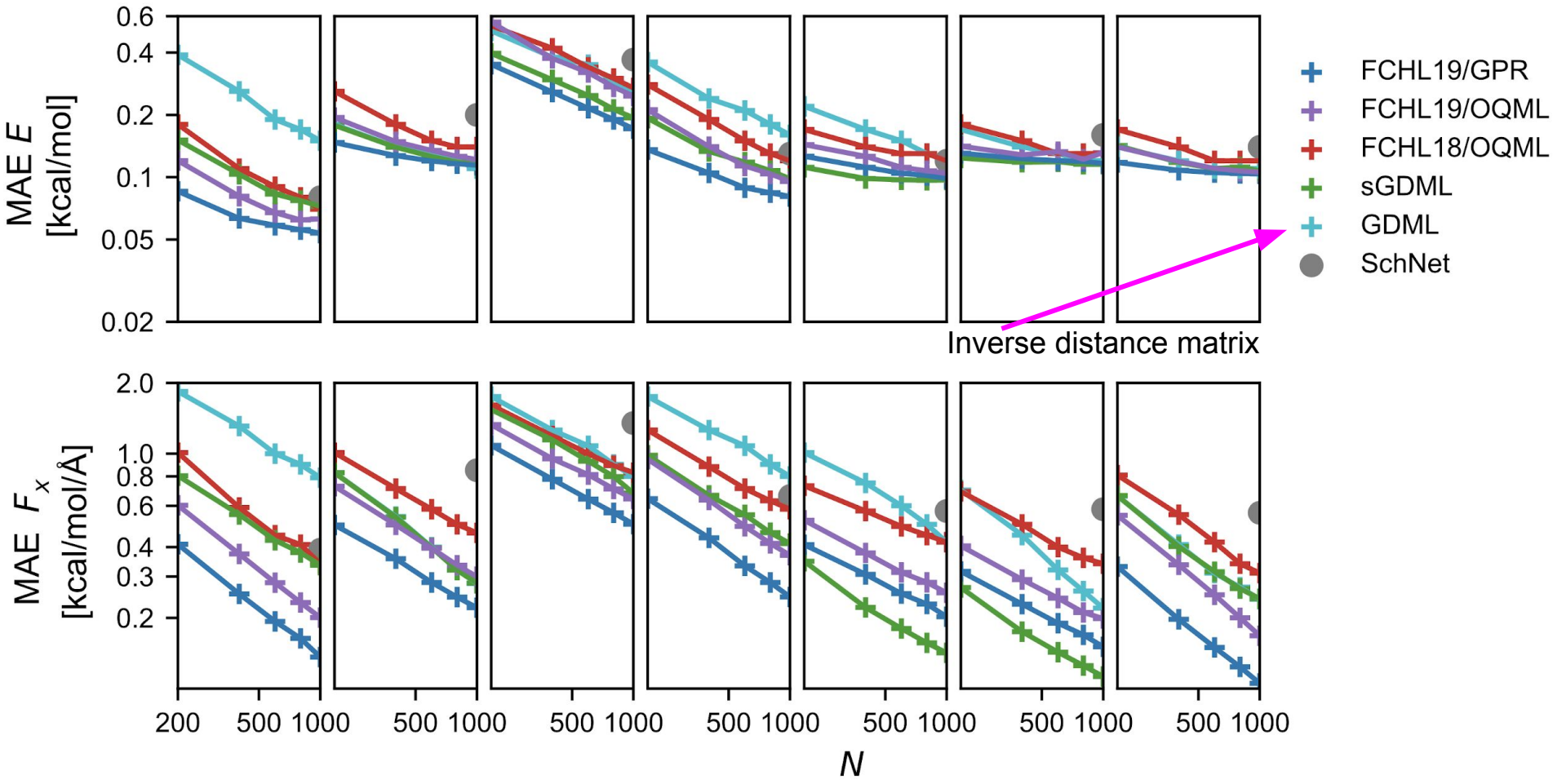




FCHL19: "FCHL revisited: Faster and more accurate quantum machine learning", Christensen et al, submitted to *J Chem Phys*

OQML revisted vs. GDML revisited: Forces for MD17

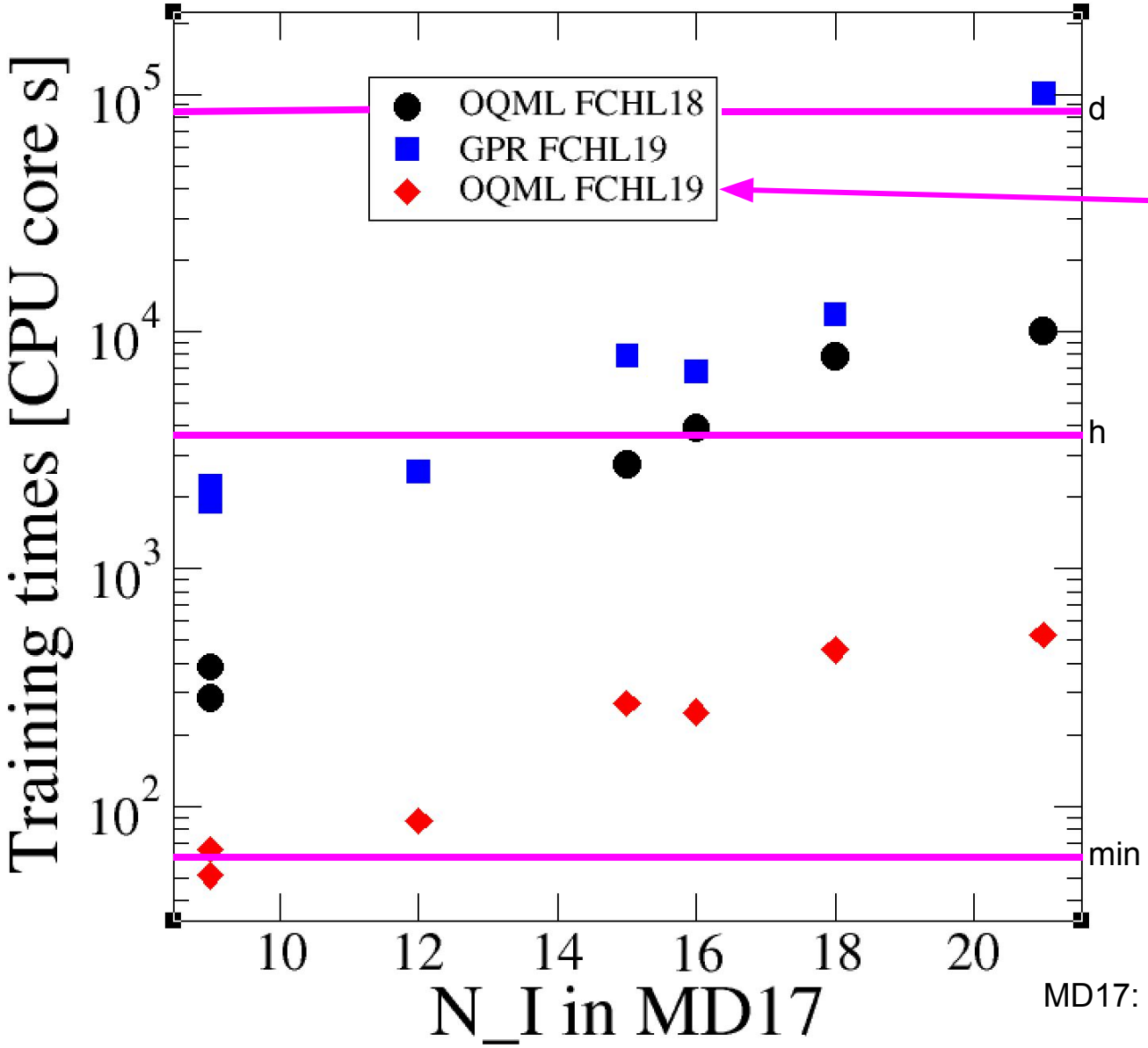




MD17 & sGDML: Chmiela et al, *Sci. Adv.* (2017); *Nature Comm.* (2018); Comput. Phys. Comm. (2019)

"Operators in Machine Learning: Response Properties in Chemical Space", Christensen et al, *J Chem Phys* (2019), arxiv.org/abs/1807.08811
FCHL19: "FCHL revisited: Faster and more accurate quantum machine learning", Christensen et al, submitted to *J Chem Phys*

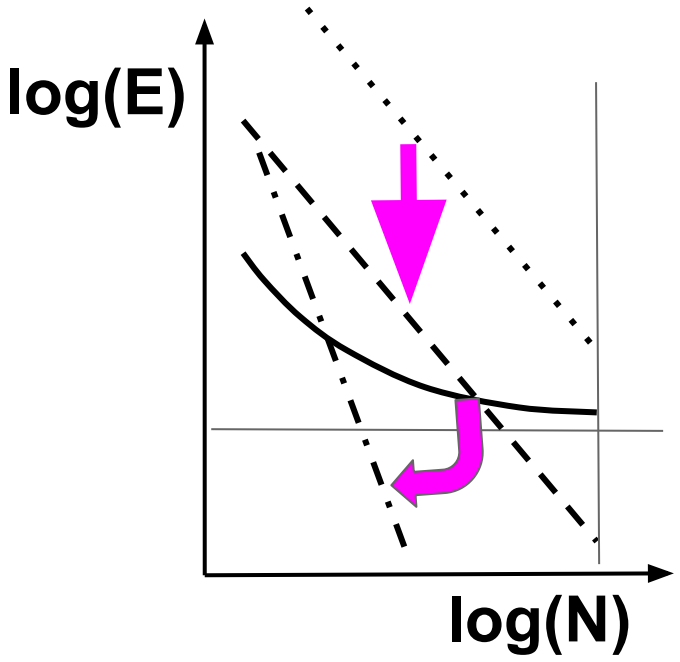
Timings to train 1 k configurations/MD17-molecule



Prediction *t* ~ ms/atom

MD17: Chmiela et al, *Sci. Adv.* (2017)

"Operators in Machine Learning: Response Properties in Chemical Space", Christensen et al, *J Chem Phys* (2019), arxiv.org/abs/1807.08811
FCHL19: "FCHL revisited: Faster and more accurate quantum machine learning", Christensen et al, submitted to *J Chem Phys*



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

Error $\sim a/N^b$

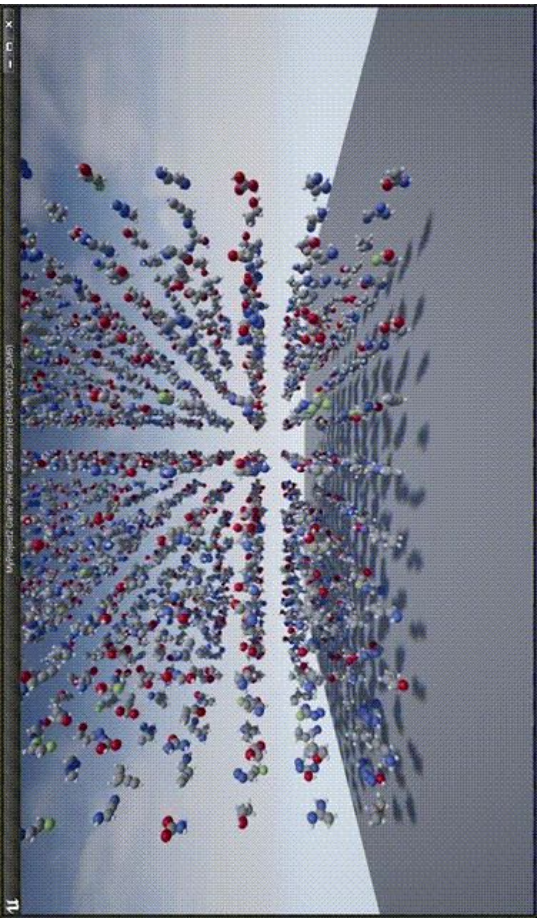
$\rightarrow$ log(Error) = log(a) - $b$log(N)

`Quantum Machine Learning', von Lilienfeld, *Angew. Chem. Int. Ed.* (2018)

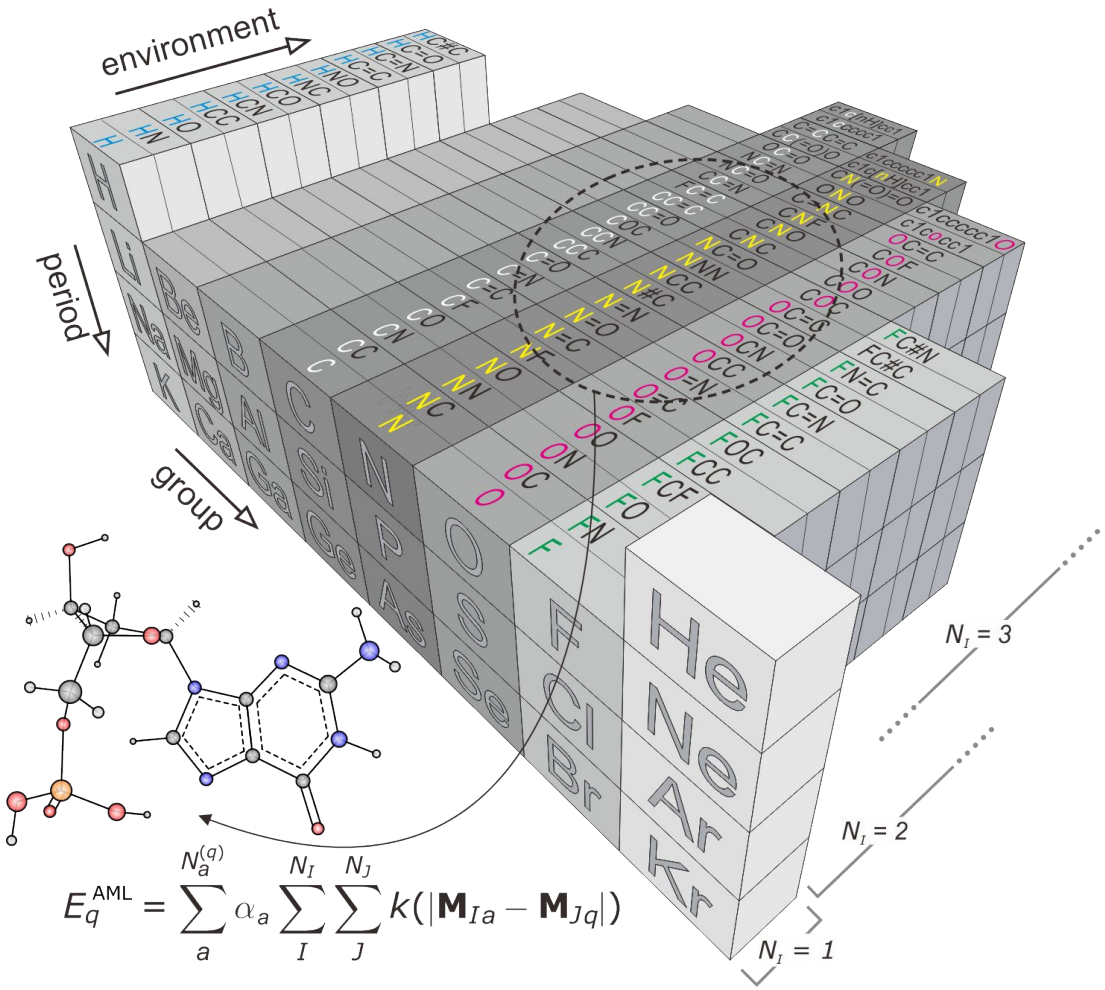
Vapnik, *The Nature of Statistical Learning Theory,* Springer (1995)

Atom in a Molecule: “AM-on”

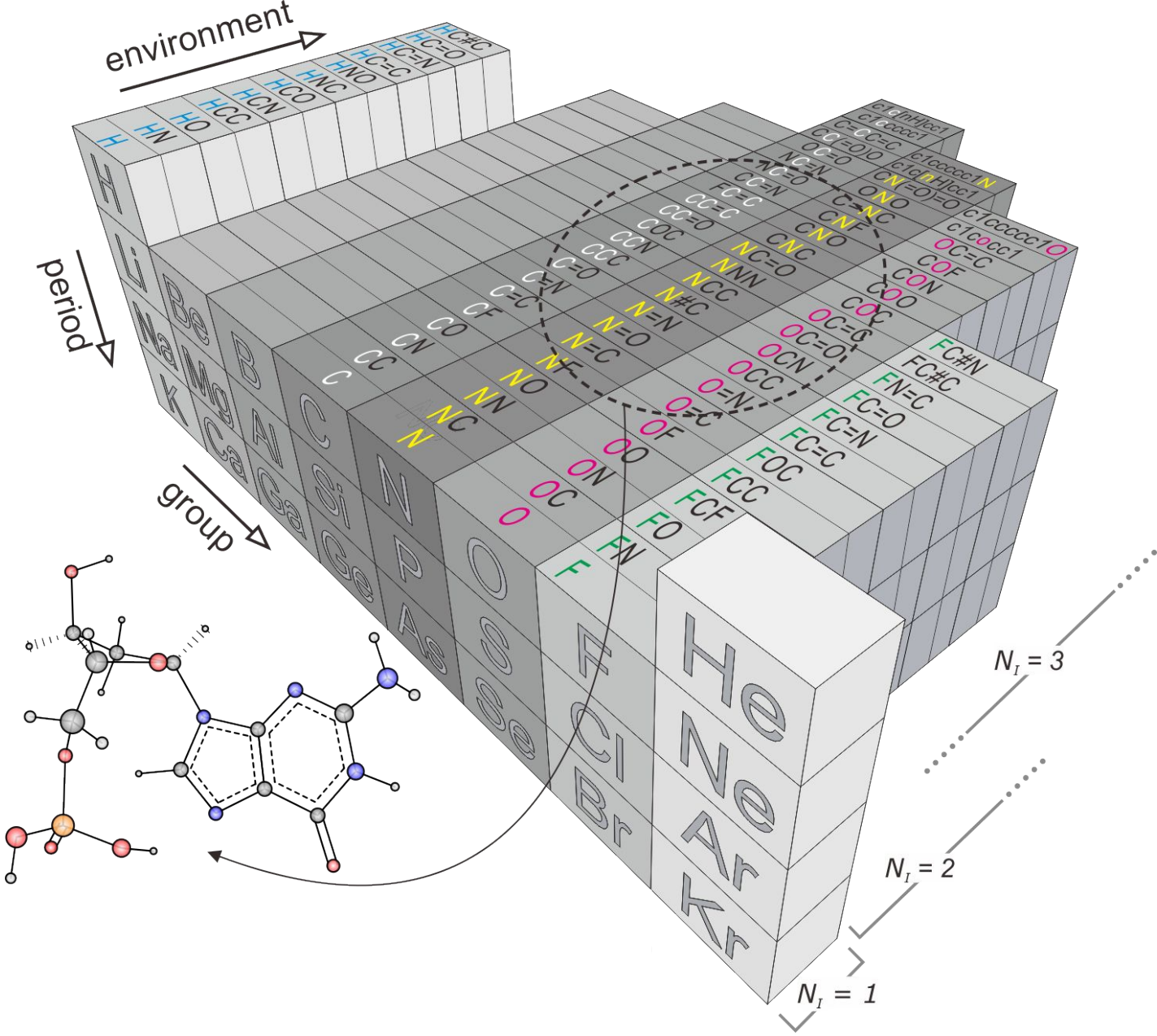
elementary	→	building blocks	→	structure	→	effect
letters	→	words	→	sentence	→	meaning
atoms	→	DNA	→	gene	→	function
atoms	→	**AMONS**	→	molecule	→	property



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

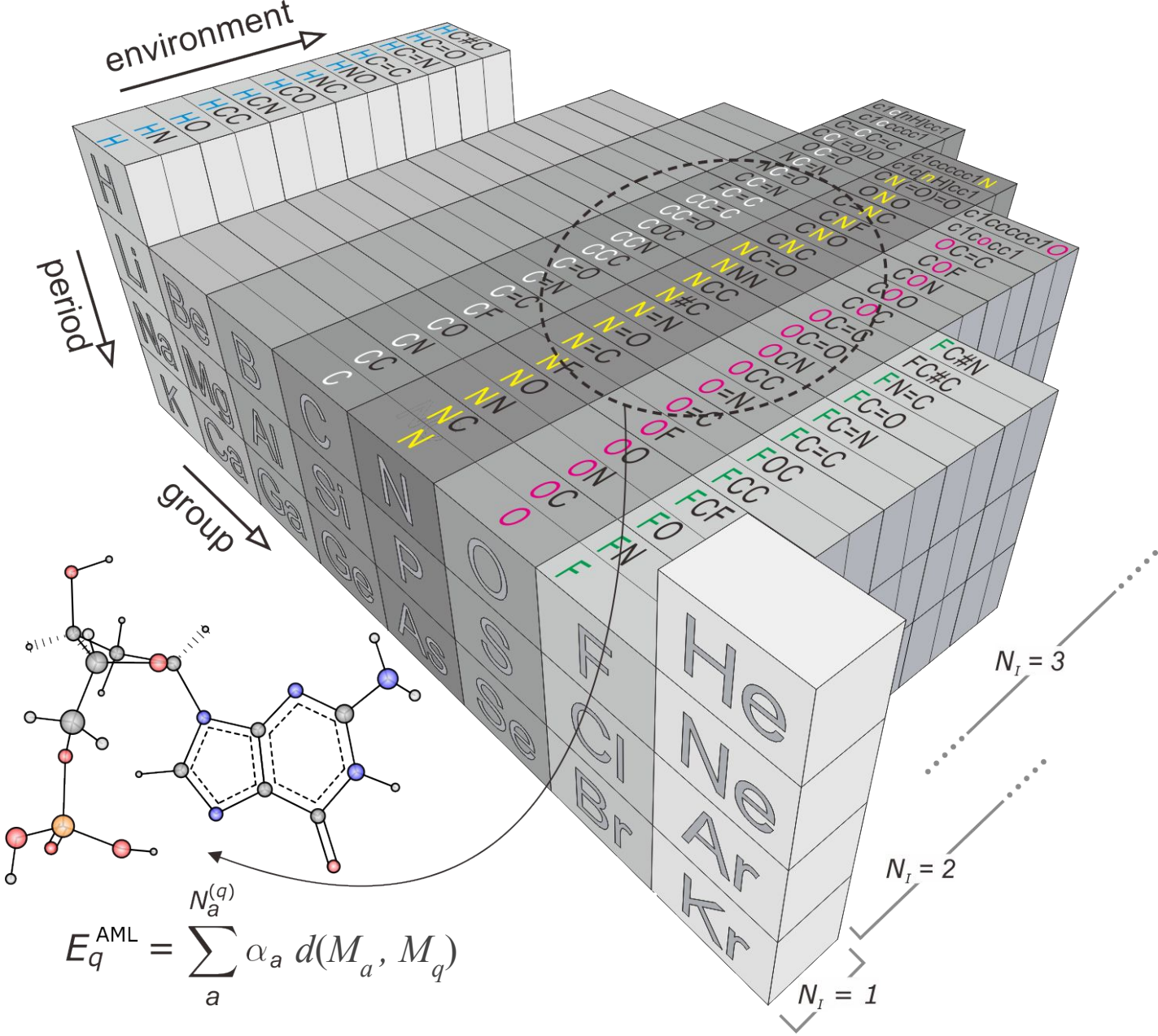


AMON QML



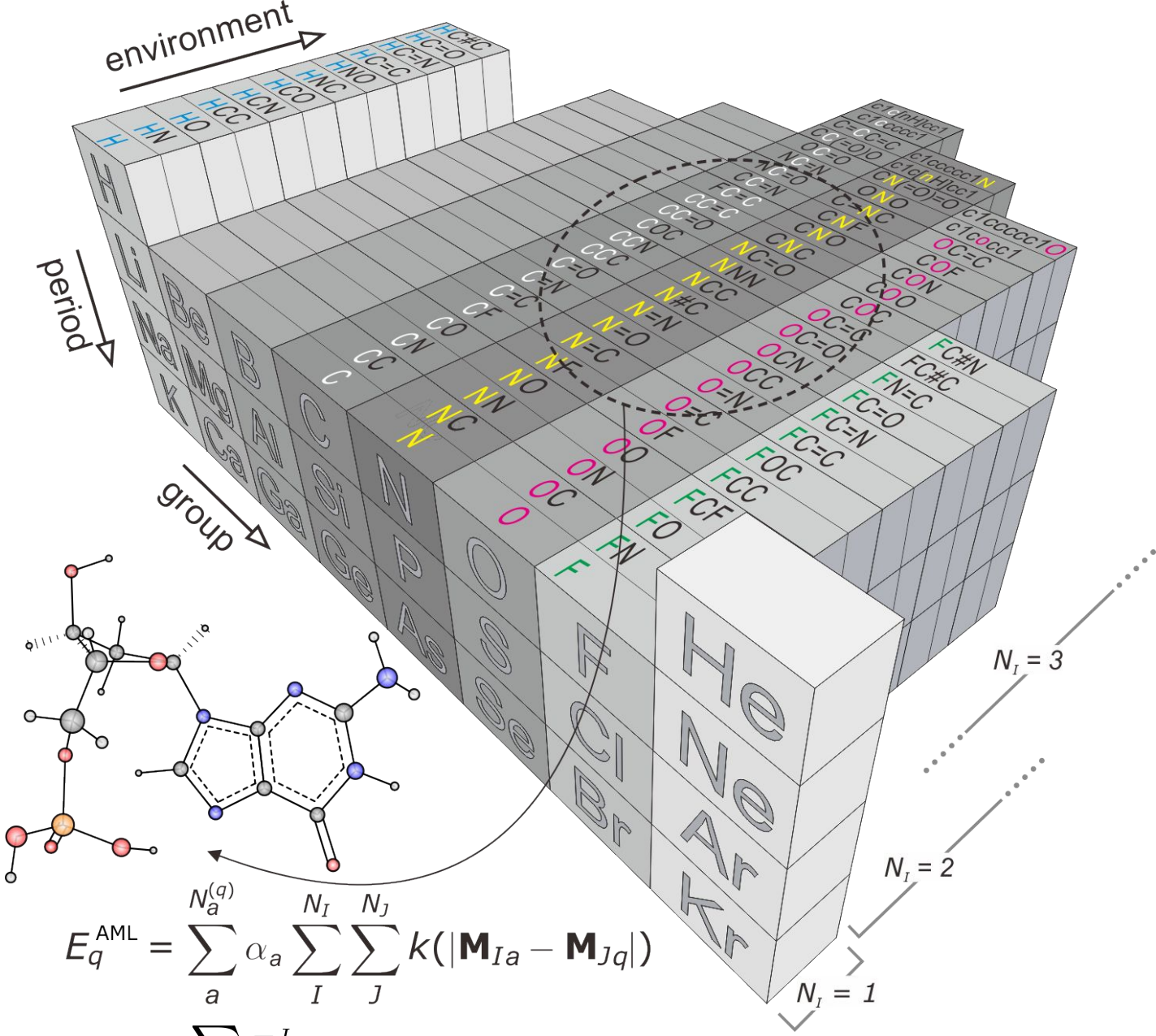
Huang, von Lilienfeld, https://arxiv.org/abs/1707.04146

AMON QML



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

AMON QML



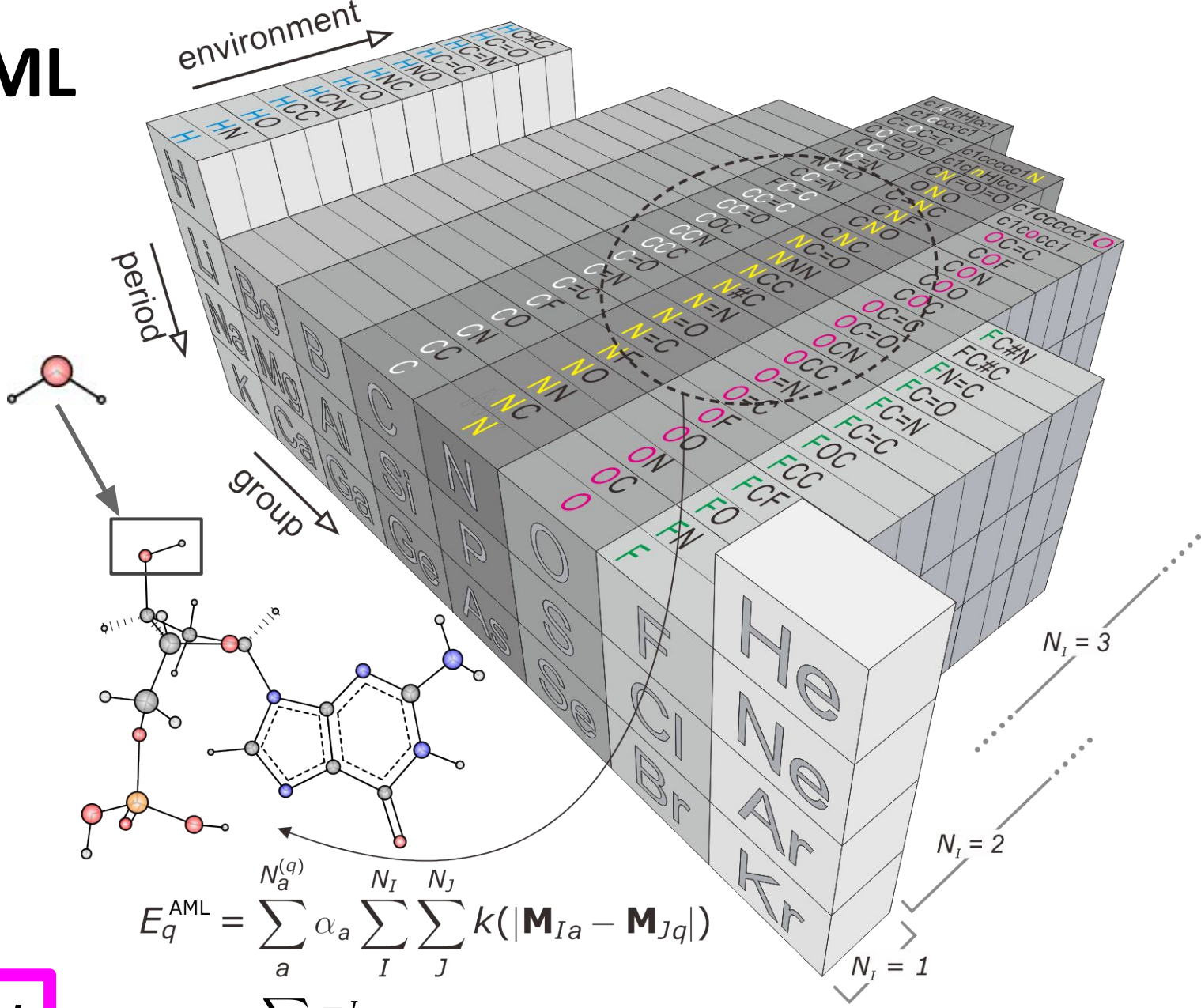
$$E_q^{\text{AML}} = \sum_a^{N_a^{(q)}} \alpha_a \sum_I^{N_I} \sum_J^{N_J} k(|\mathbf{M}_{Ia} - \mathbf{M}_{Jq}|)$$

$$= \sum_J E_q^J$$

$$N_I < N_J$$

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

AMON QML



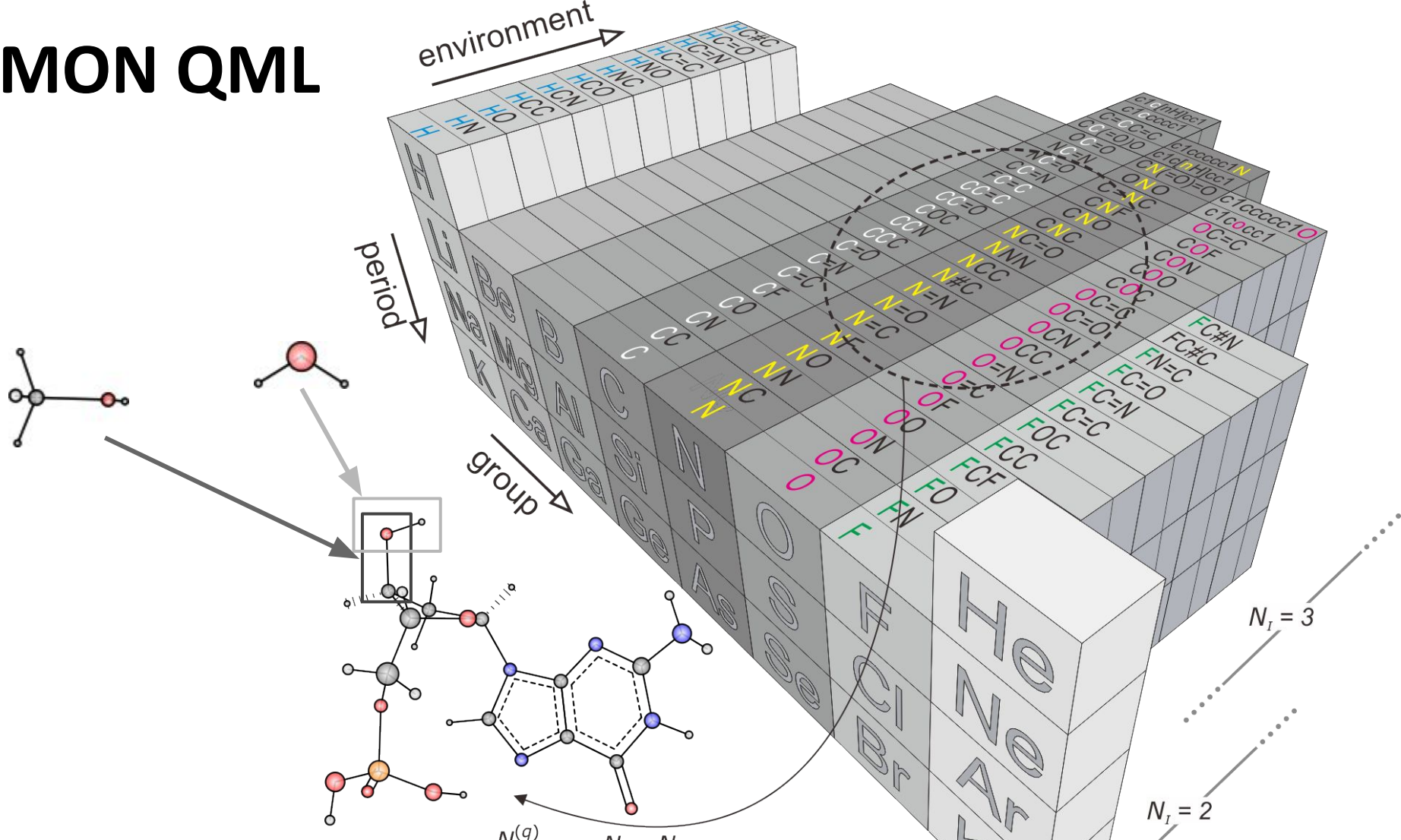
$$E_q^{\text{AML}} = \sum_a^{N_a^{(q)}} \alpha_a \sum_I^{N_I} \sum_J^{N_J} k(|\mathbf{M}_{Ia} - \mathbf{M}_{Jq}|)$$

$$= \sum_J E_q^J$$

$$N_I < N_J$$

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

AMON QML



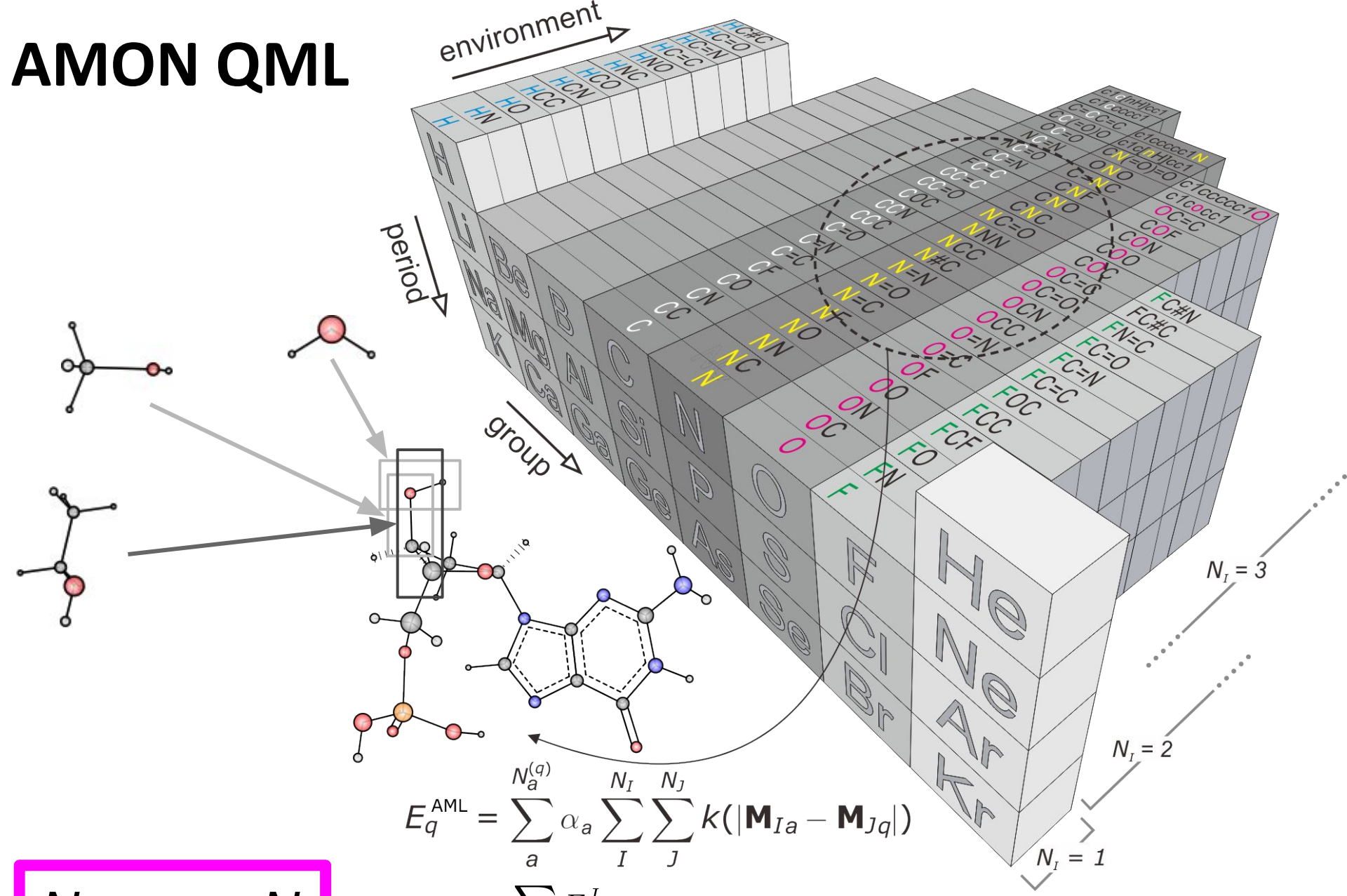
$$E_q^{\text{AML}} = \sum_a^{N_a^{(q)}} \alpha_a \sum_I^{N_I} \sum_J^{N_J} k(|\mathbf{M}_{Ia} - \mathbf{M}_{Jq}|)$$

$$= \sum_J E_q^J$$

$$N_I < N_J$$

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

AMON QML



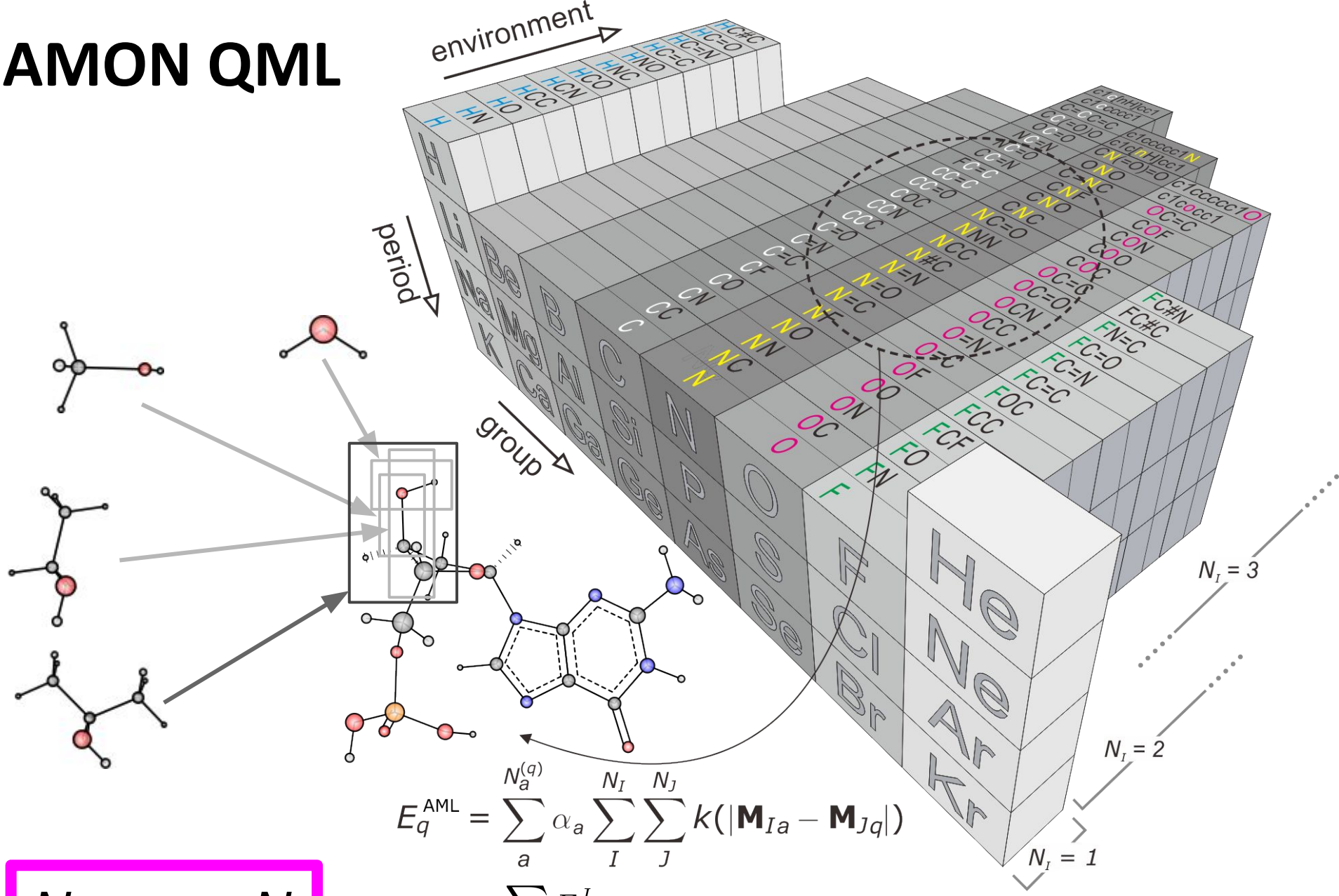
$$E_q^{\text{AML}} = \sum_a^{N_a^{(q)}} \alpha_a \sum_I^{N_I} \sum_J^{N_J} k(|\mathbf{M}_{Ia} - \mathbf{M}_{Jq}|)$$

$$= \sum_J E_q^J$$

$$N_I < N_J$$

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

AMON QML



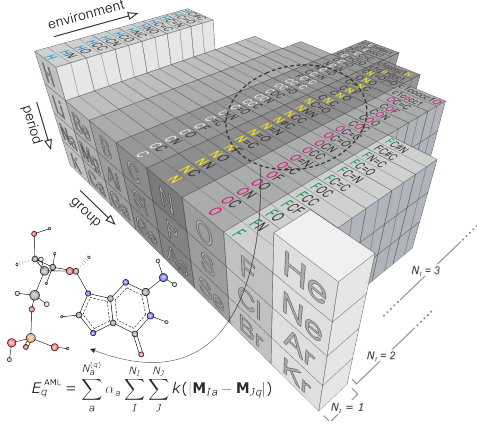
$$E_q^{\text{AML}} = \sum_a^{N_a^{(q)}} \alpha_a \sum_I^{N_I} \sum_J^{N_J} k(|\mathbf{M}_{Ia} - \mathbf{M}_{Jq}|)$$

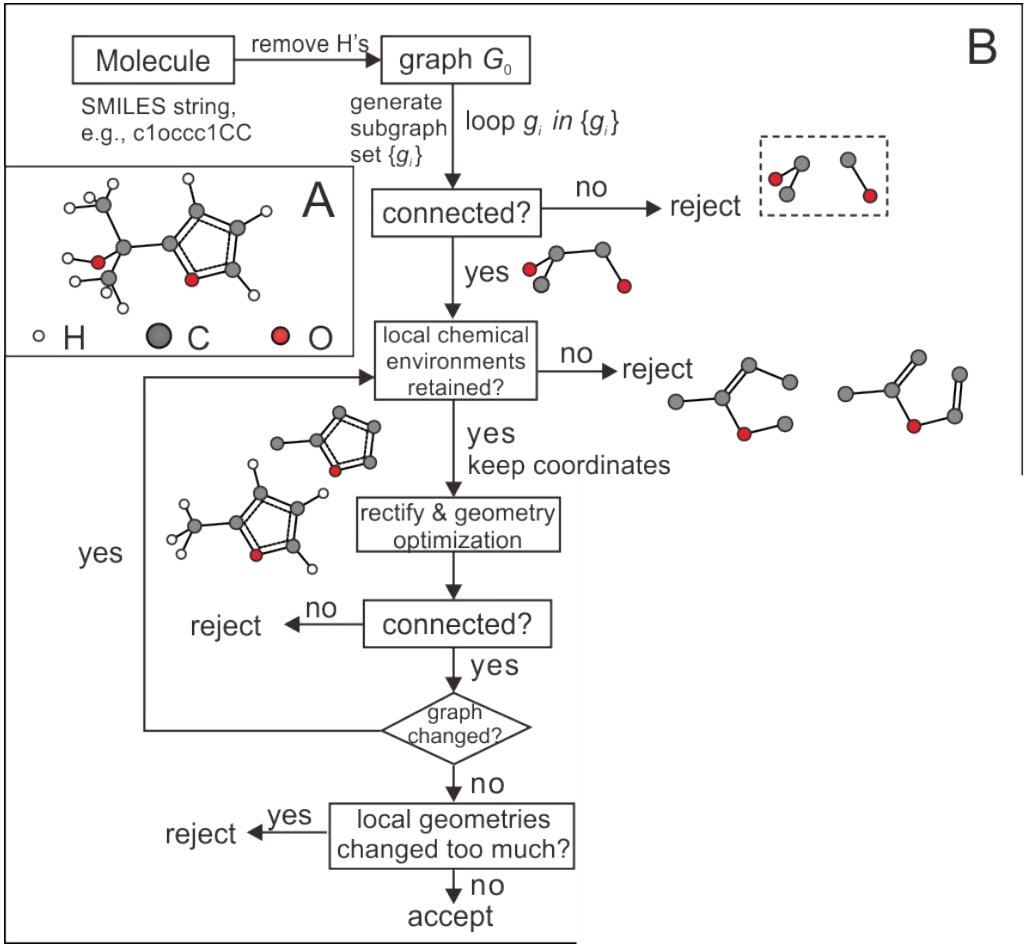
$$= \sum_J E_q^J$$

$$N_I < N_J$$

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

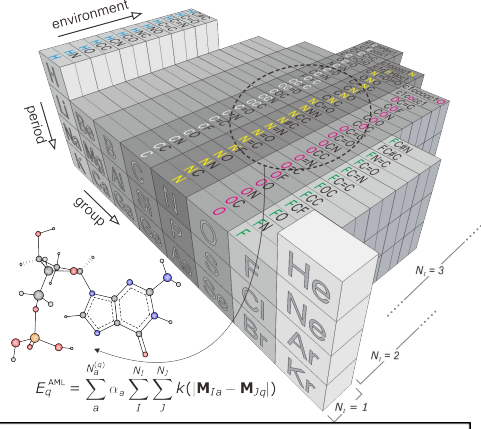
AMON QML

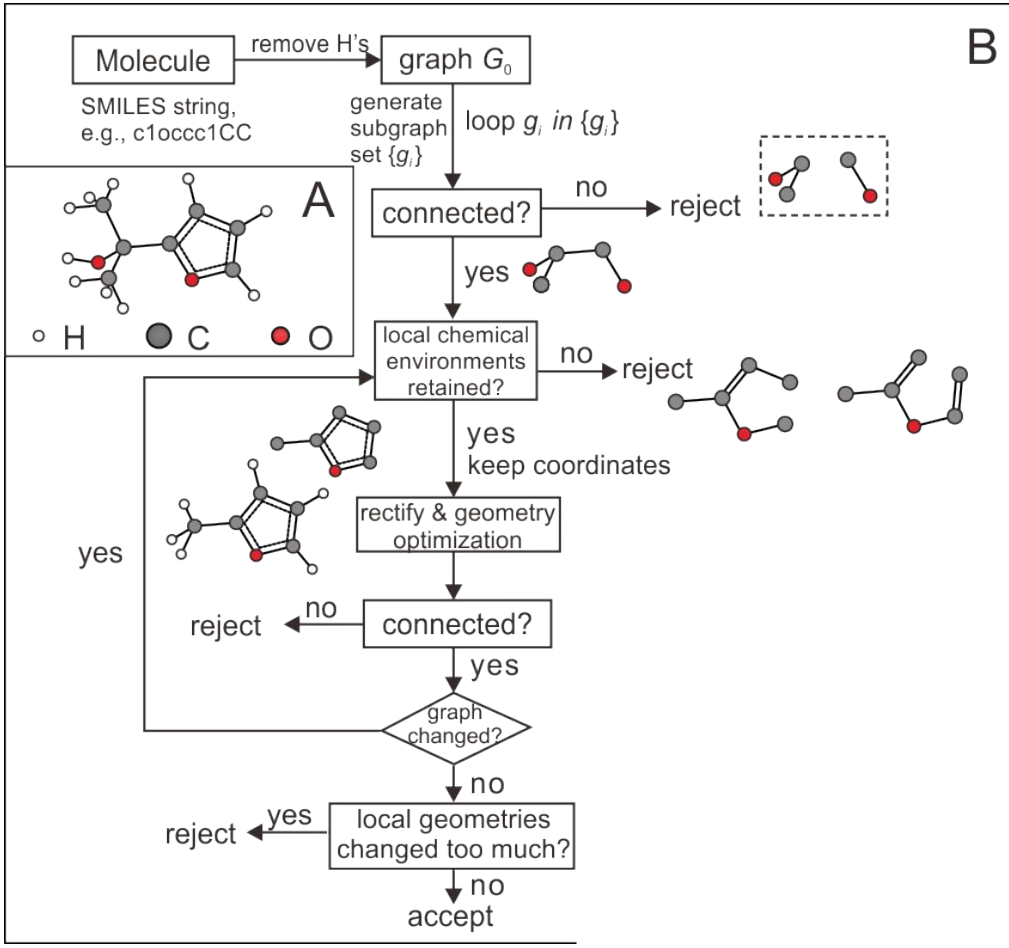


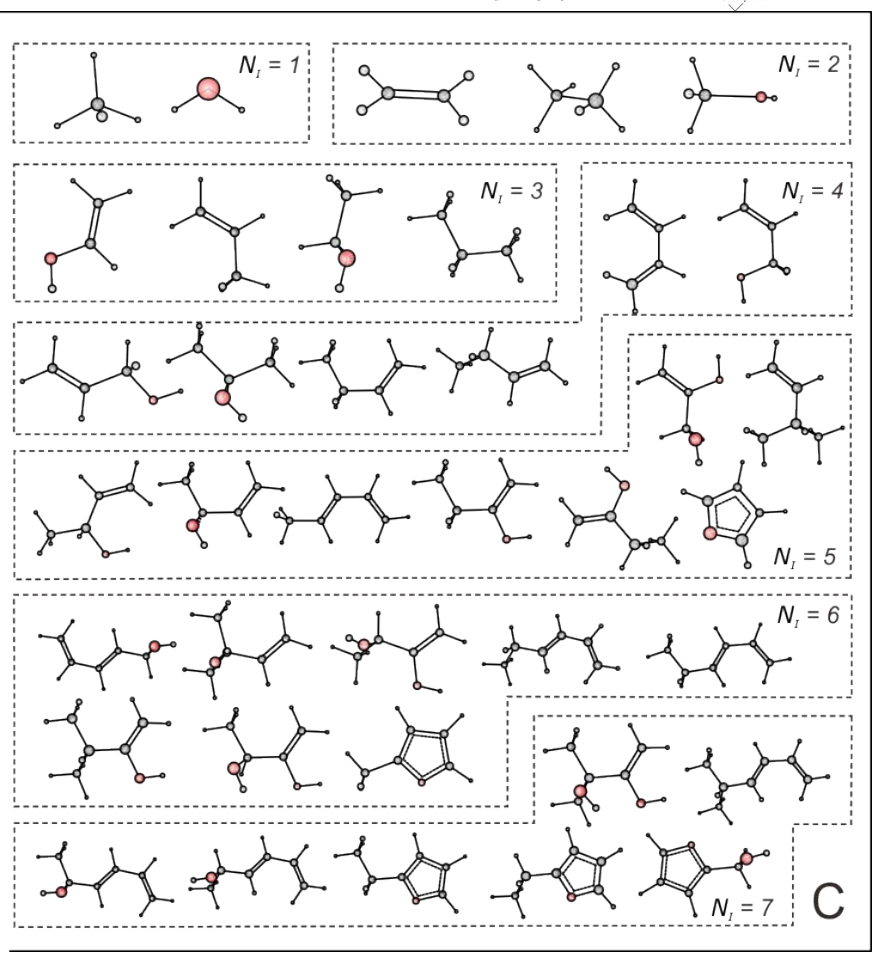


Huang, von Lilienfeld, https://arxiv.org/abs/1707.04146

AMON QML

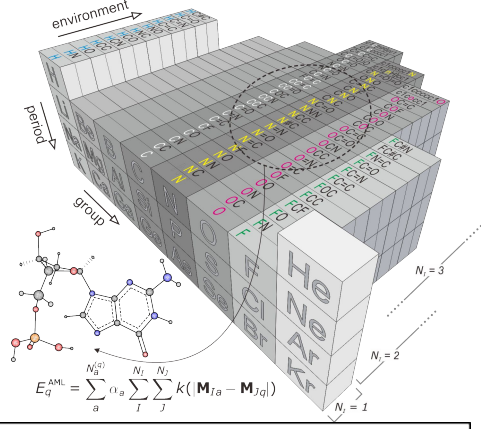


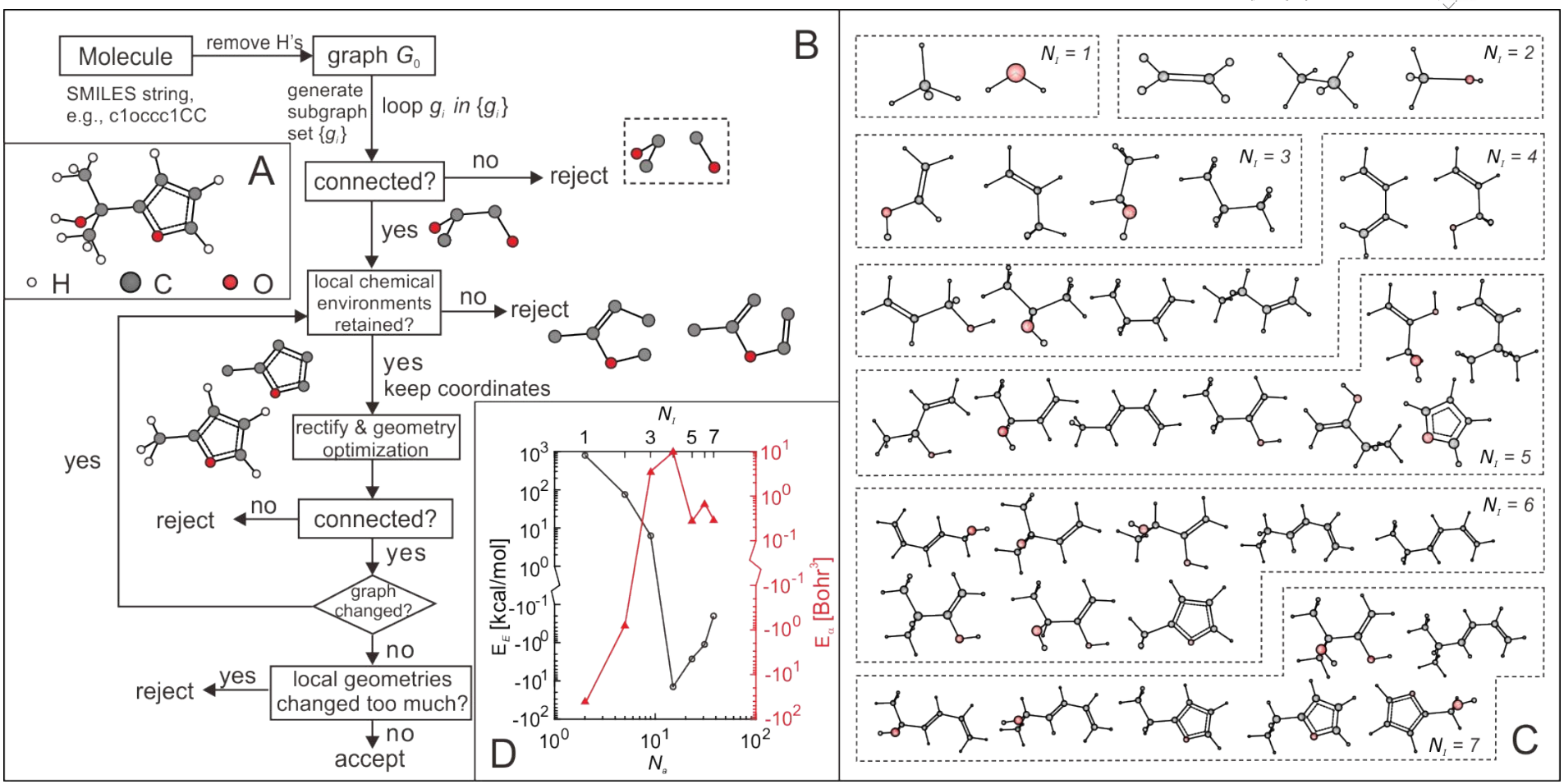




Huang, von Lilienfeld, https://arxiv.org/abs/1707.04146

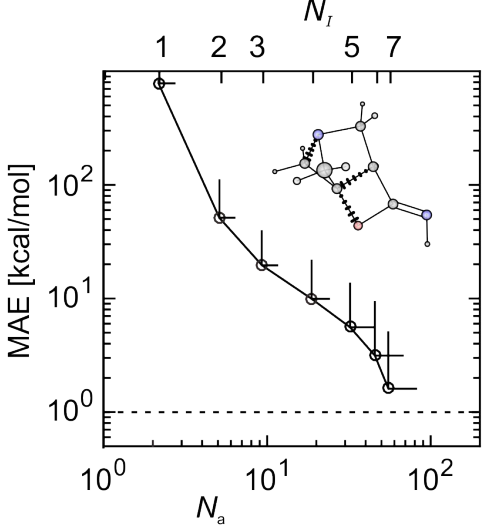
AMON QML

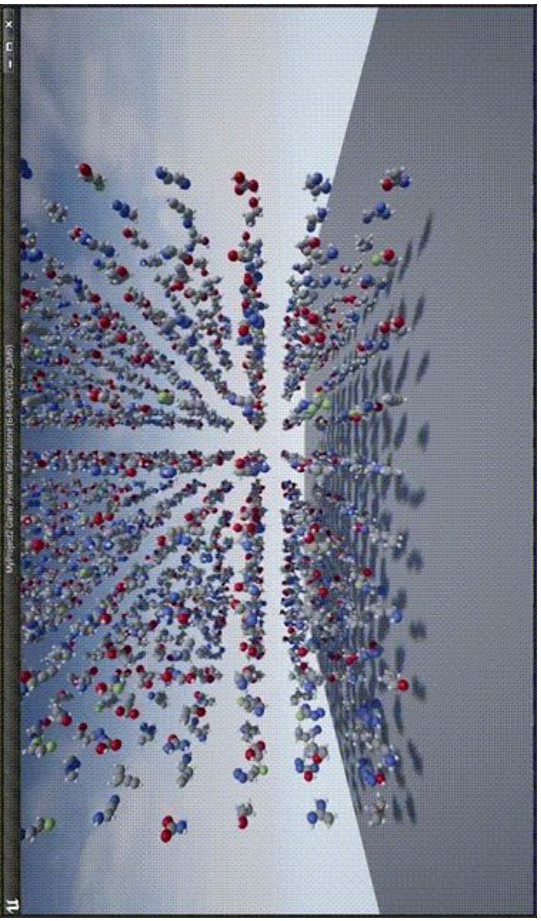




Huang, von Lilienfeld, https://arxiv.org/abs/1707.04146

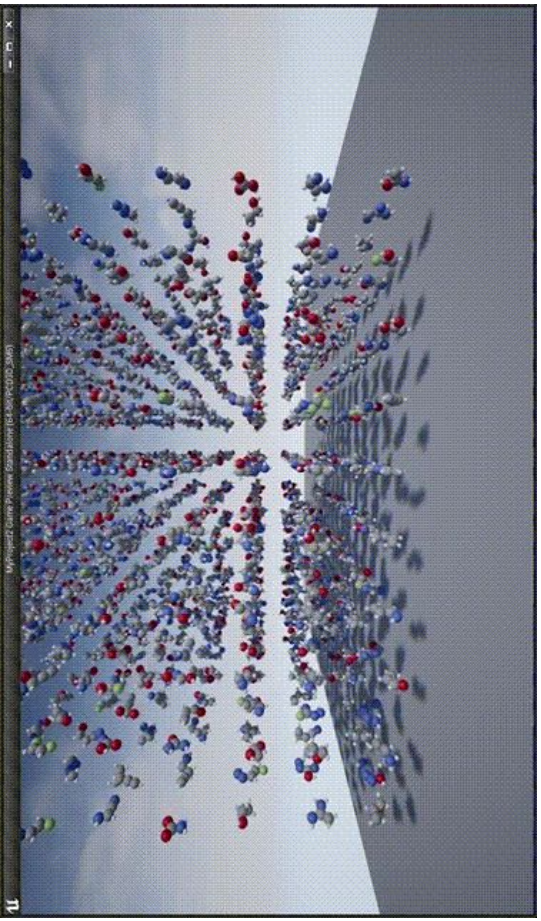
AMON QML

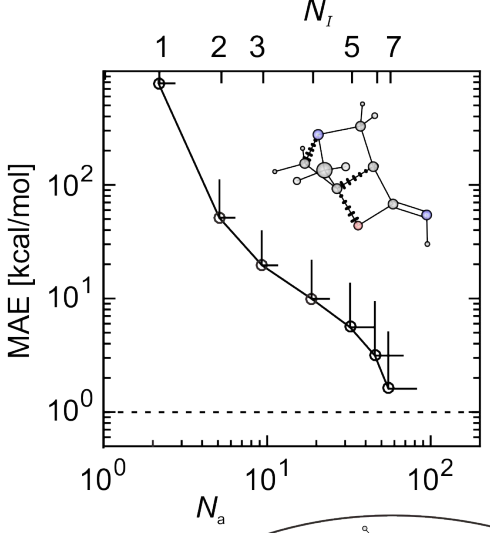


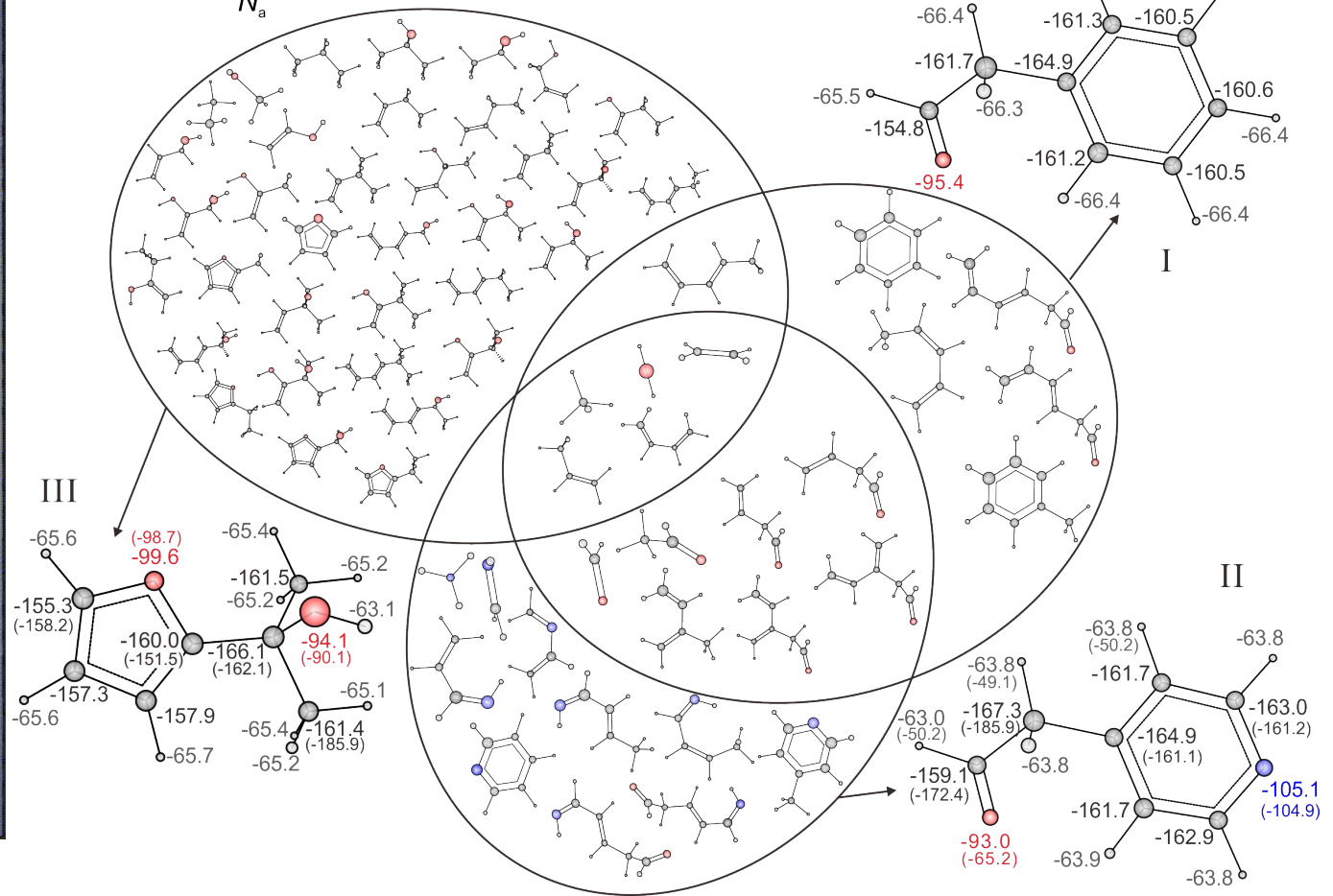


Huang, von Lilienfeld, https://arxiv.org/abs/1707.04146

AMON QML

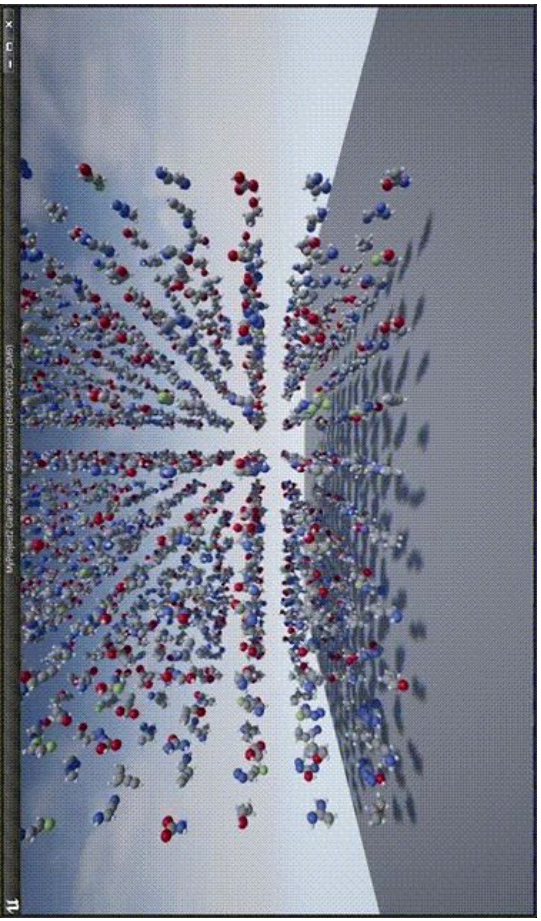




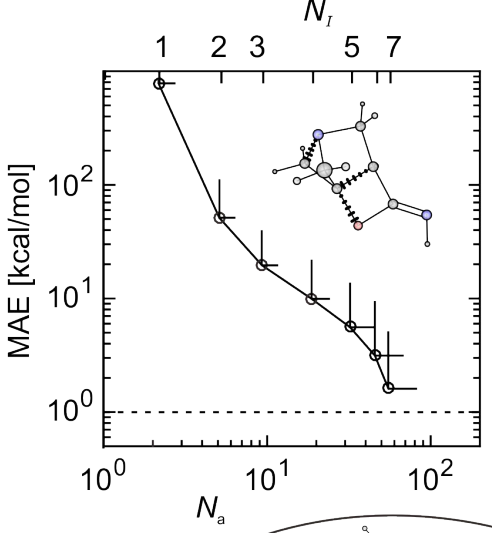


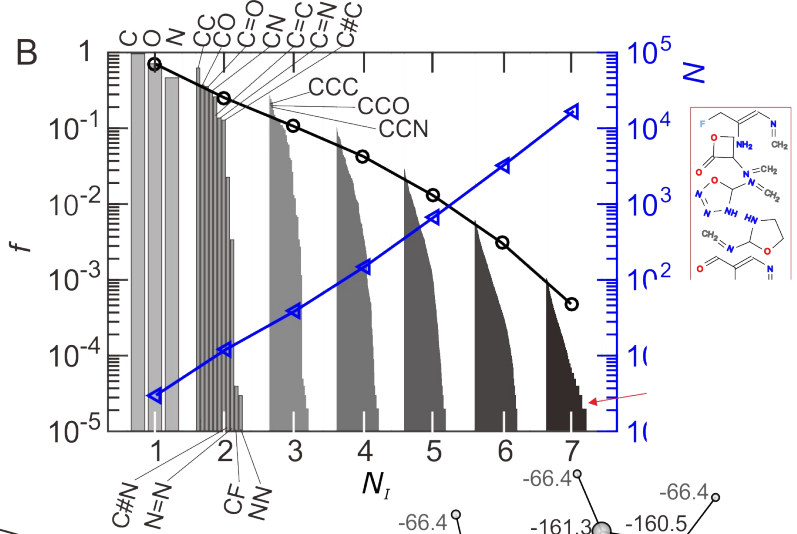
Huang, von Lilienfeld, https://arxiv.org/abs/1707.04146

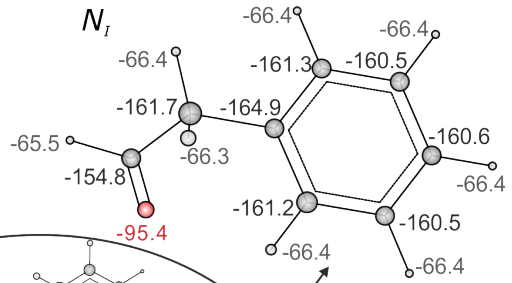
AMON QML



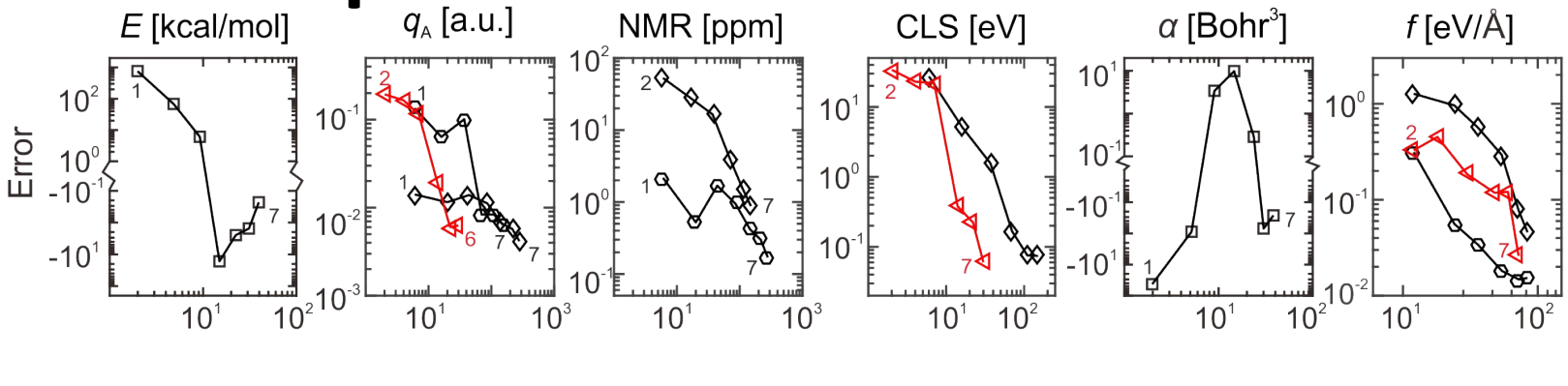
Huang, von Lilienfeld, https://arxiv.org/abs/1707.04146





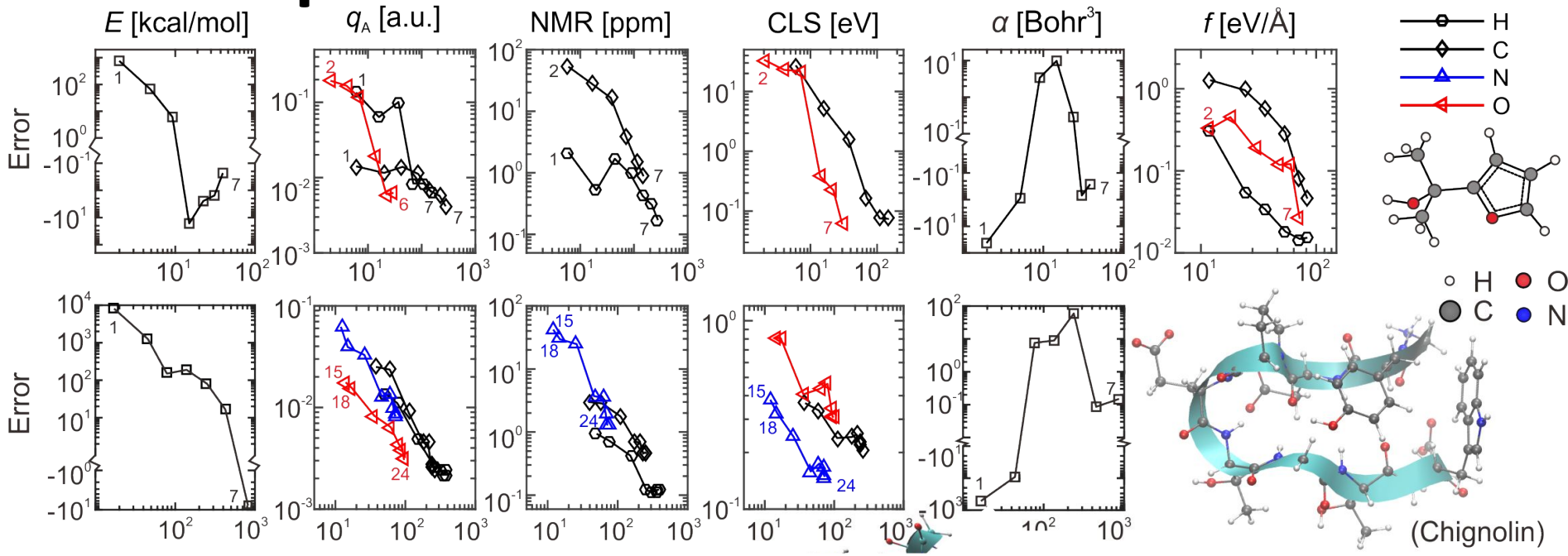


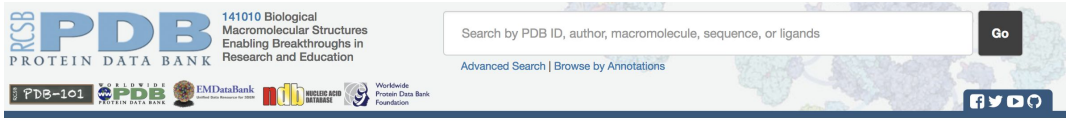
3 Examples

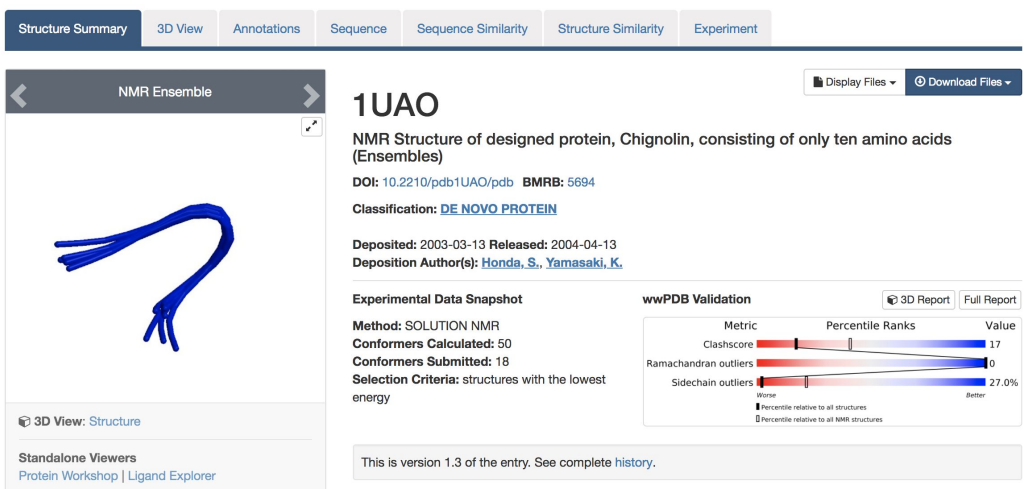


9 heavy atoms

3 Examples

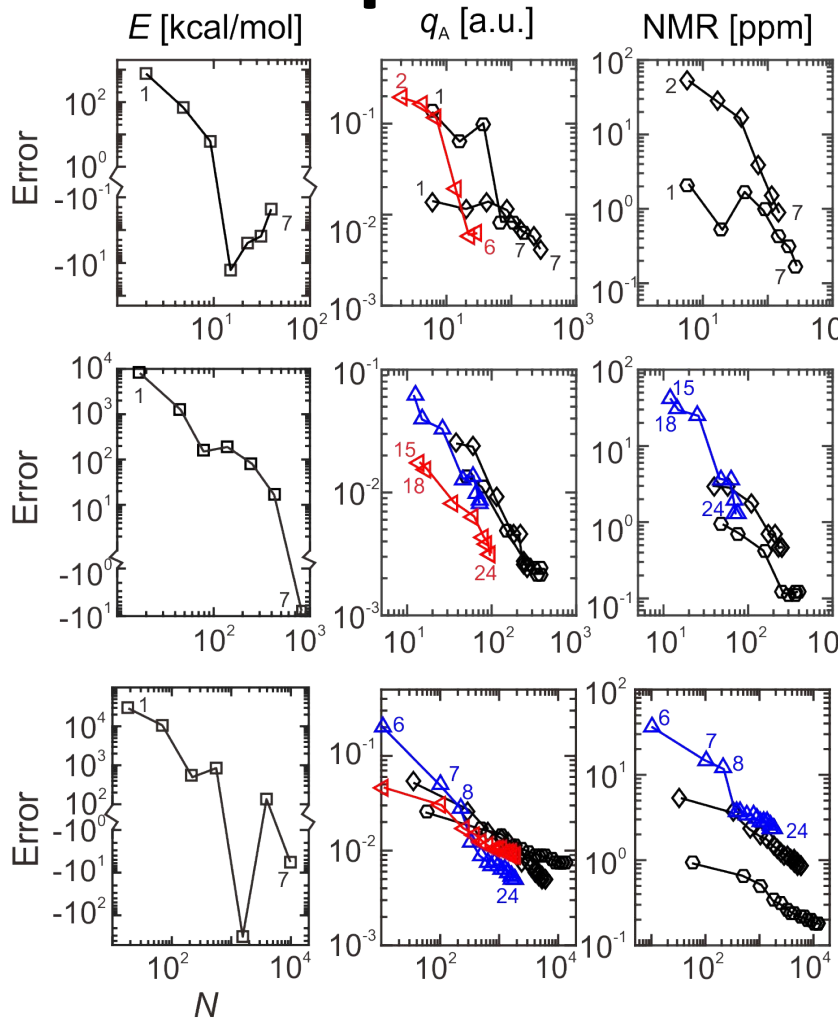


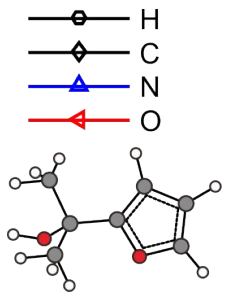


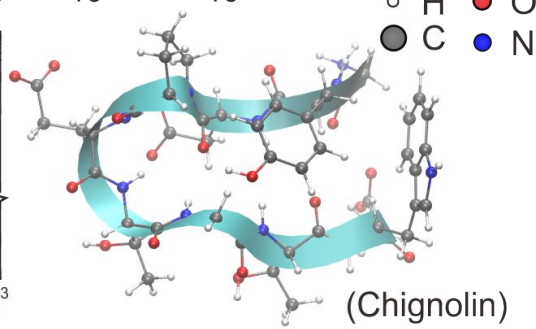


77 heavy atoms

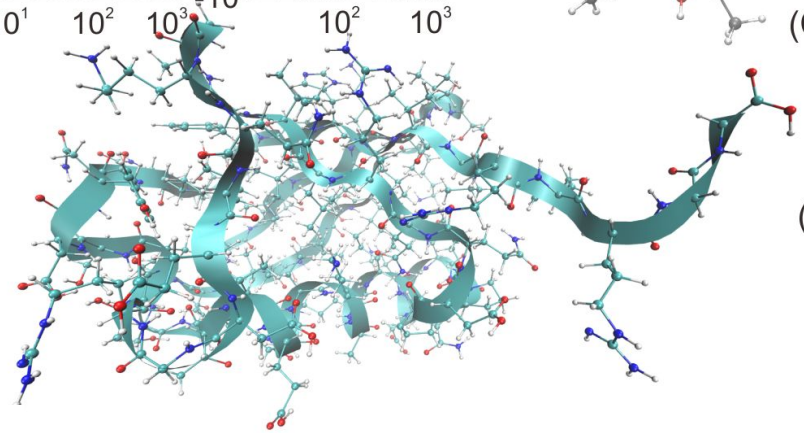
3 Examples







(Chignolin)



(Ubiquitin)

QM: > 10 hours

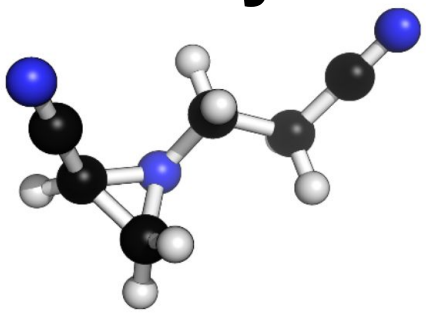
QML: ~milli seconds

602 heavy atoms

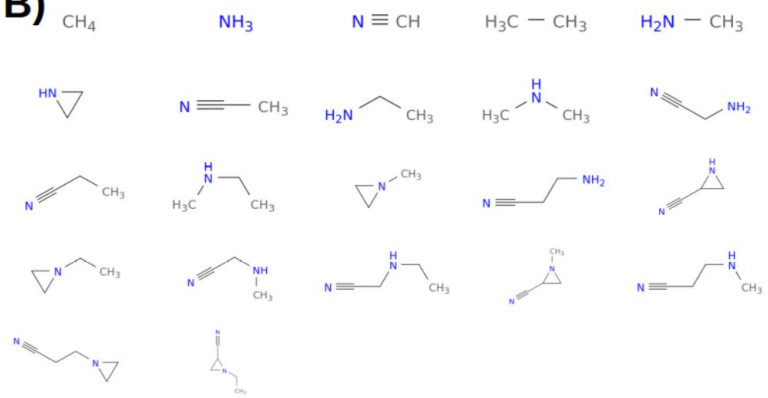
Forces w AQML

9 heavy atoms

A)



B)



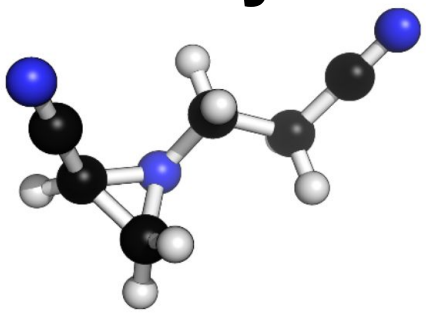
< 8 heavy atoms

"Operators in Machine Learning: Response Properties in Chemical Space", Christensen et al, *J Chem Phys* (2019), arxiv.org/abs/1807.08811

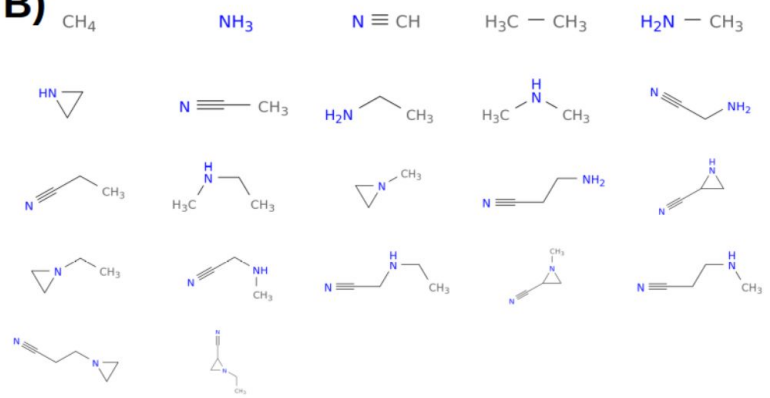
Forces w AQML

9 heavy atoms

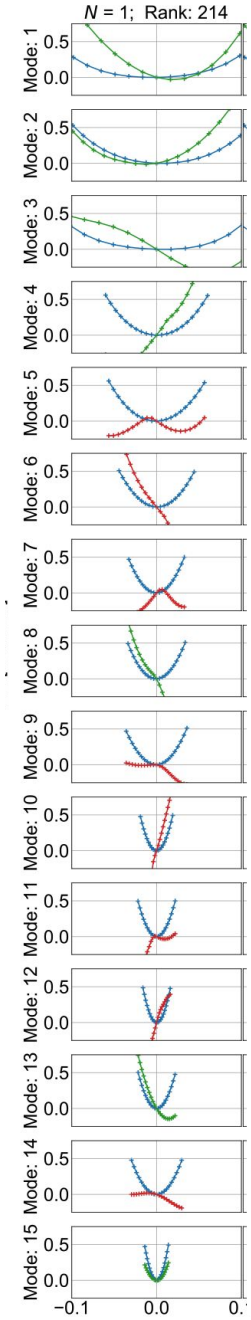
A)



B)



< 8 heavy atoms

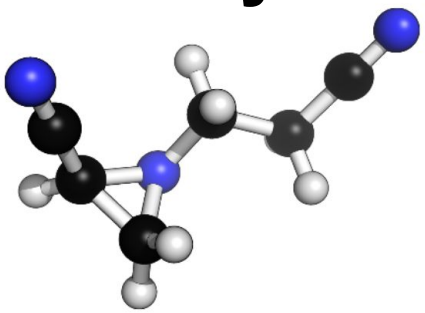


"Operators in Machine Learning: Response Properties in Chemical Space", Christensen et al, *J Chem Phys* (2019), arxiv.org/abs/1807.08811

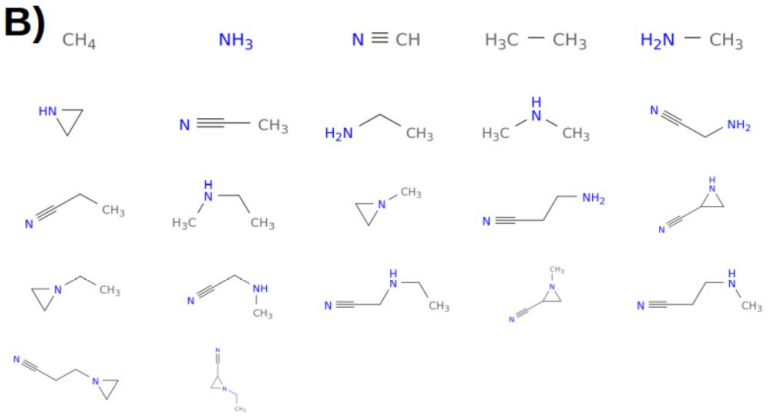
Forces w AQML

9 heavy atoms

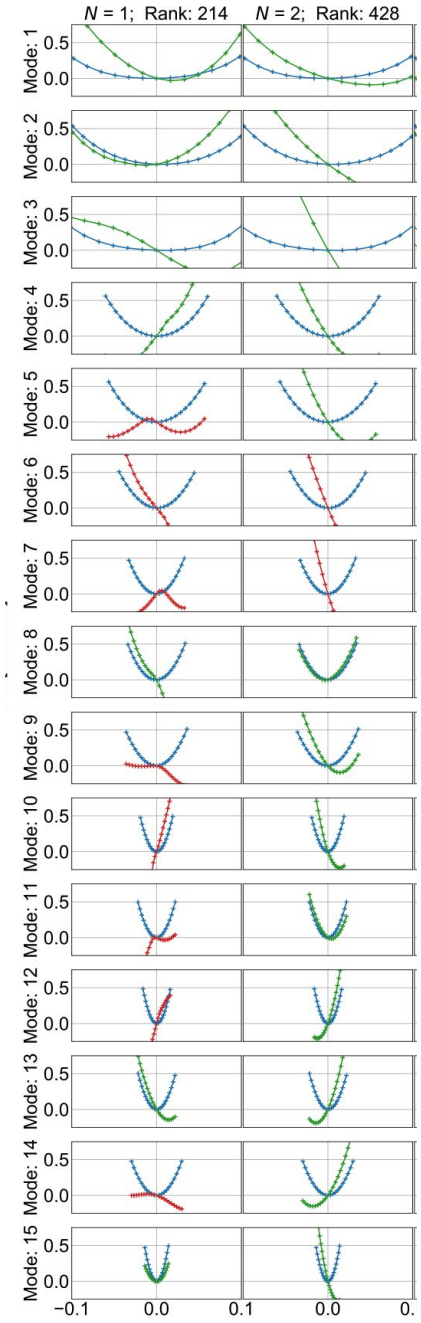
A)



B)



< 8 heavy atoms

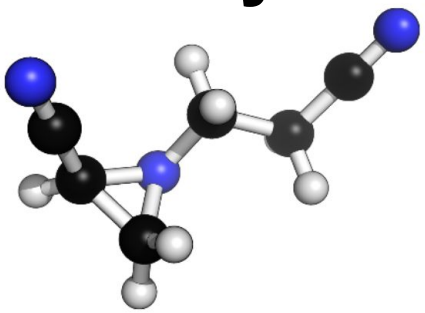


"Operators in Machine Learning: Response Properties in Chemical Space", Christensen et al, *J Chem Phys* (2019), arxiv.org/abs/1807.08811

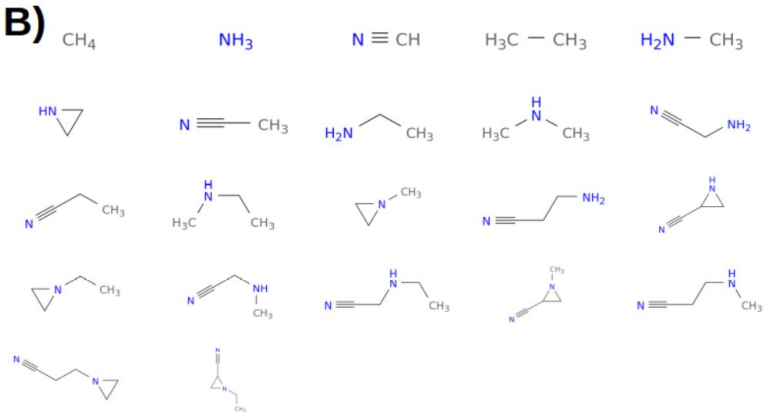
Forces w AQML

9 heavy atoms

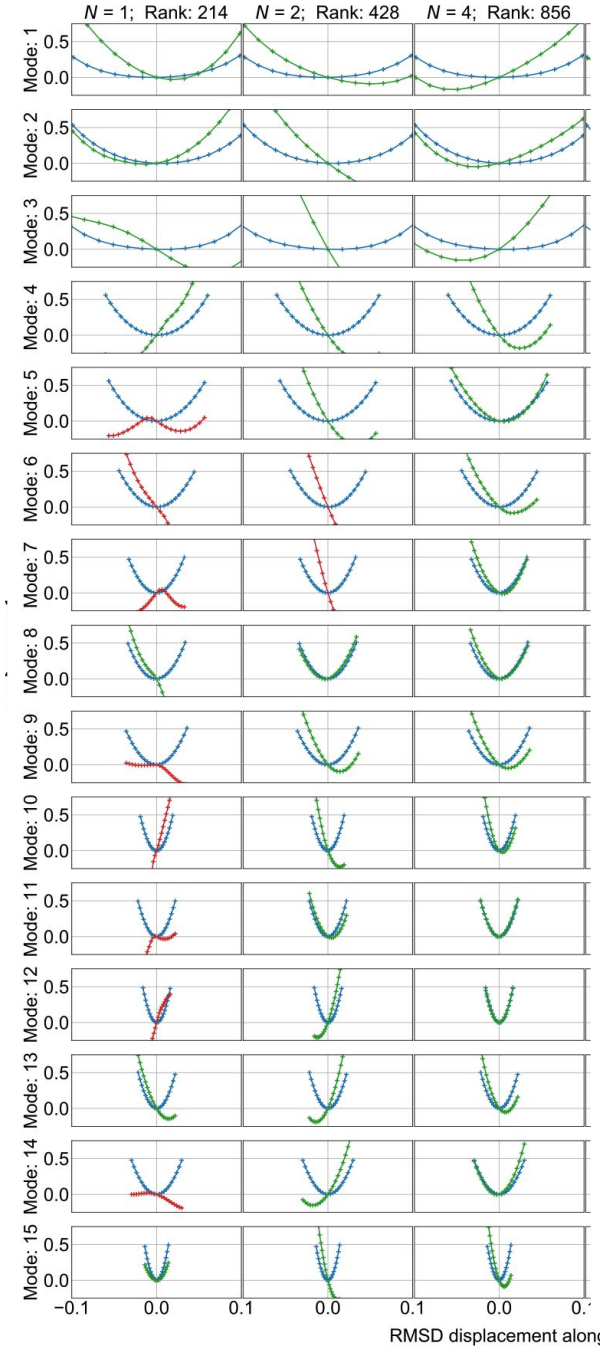
A)



B)



< 8 heavy atoms

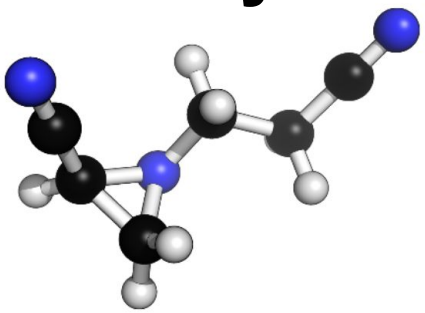


"Operators in Machine Learning: Response Properties in Chemical Space", Christensen et al, *J Chem Phys* (2019), arxiv.org/abs/1807.08811

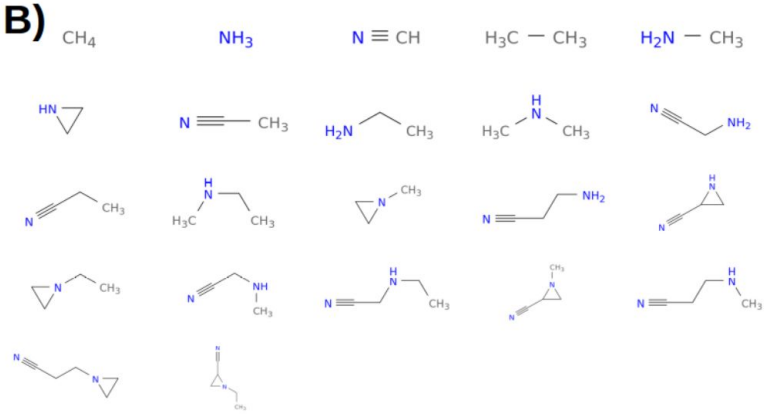
Forces w AQML

9 heavy atoms

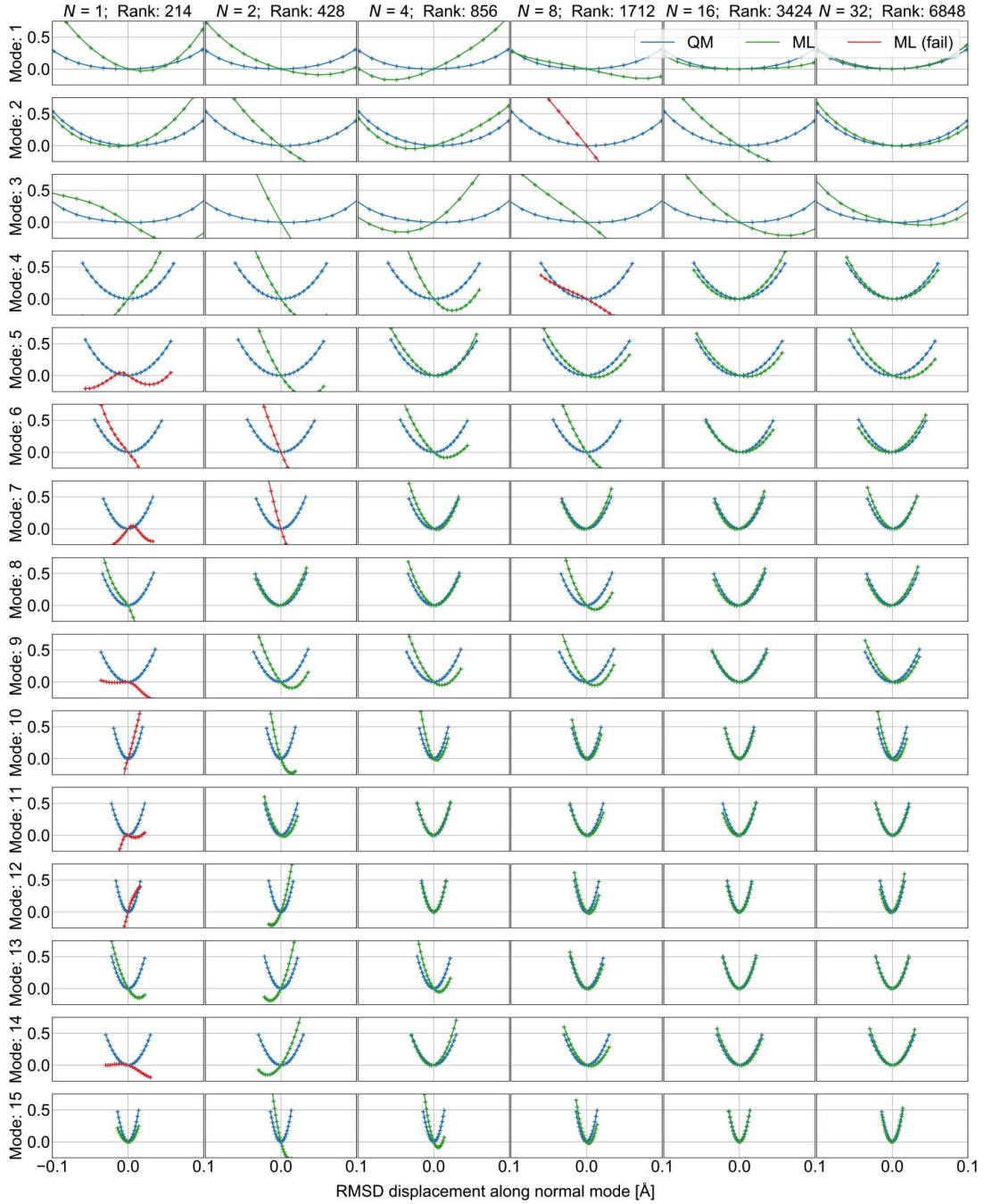
A)



B)

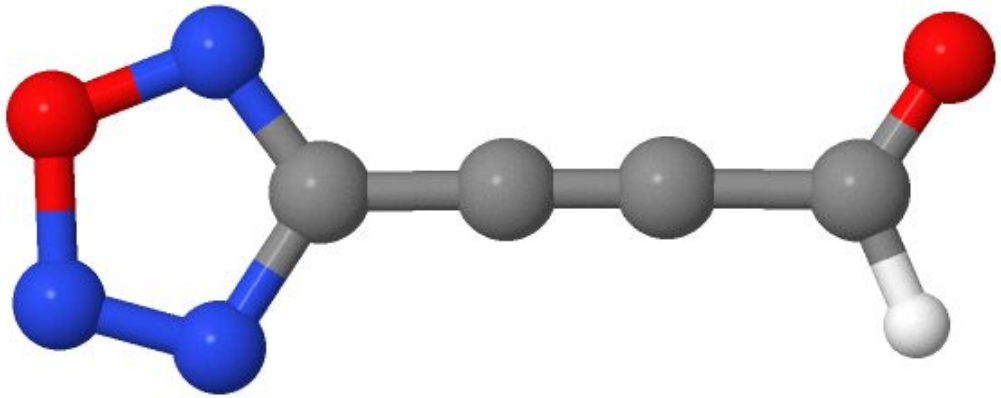


< 8 heavy atoms

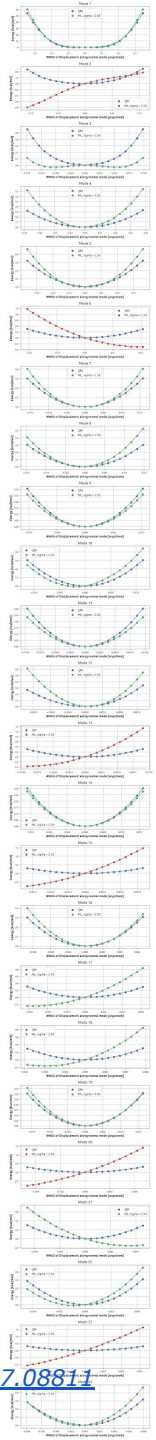


"Operators in Machine Learning: Response Properties in Chemical Space", Christensen et al, *J Chem Phys* (2019), arxiv.org/abs/1807.08811

Relaxation

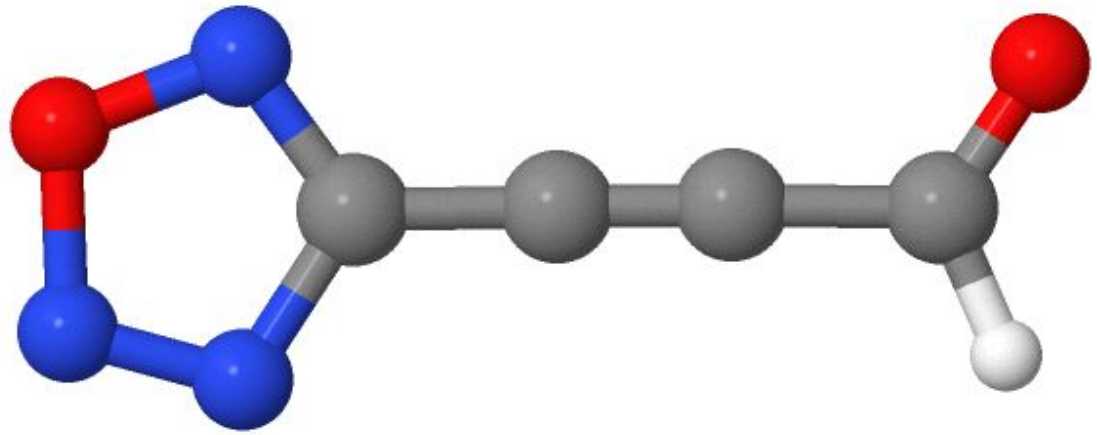


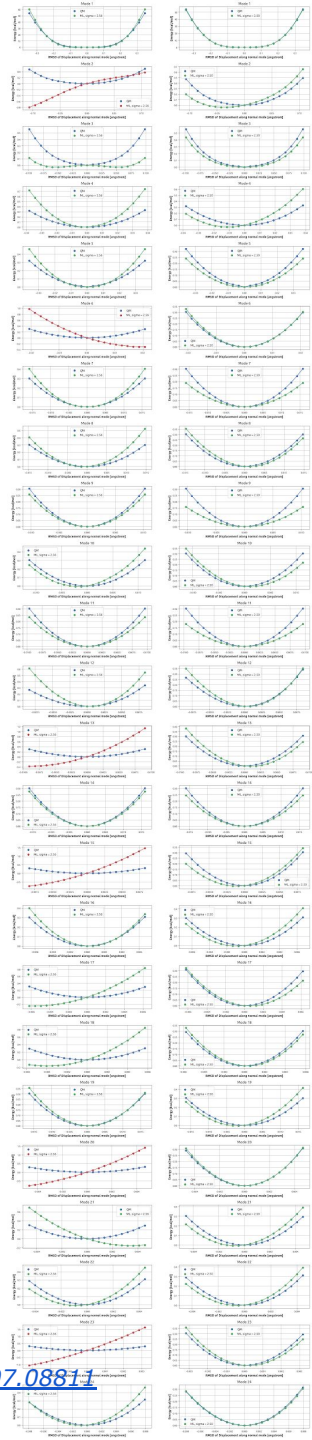
Jmol



"Operators in Machine Learning: Response Properties in Chemical Space", Christensen et al, *J Chem Phys* (2019), arxiv.org/abs/1807.08811

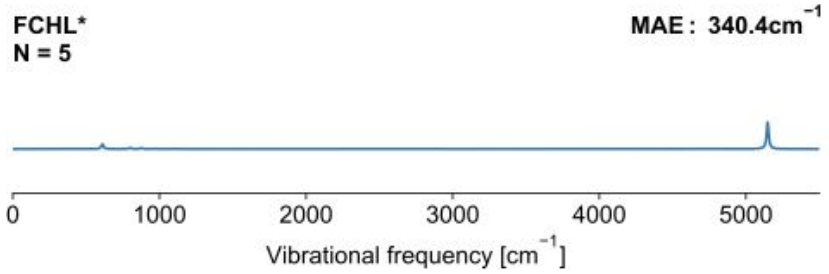
Relaxation





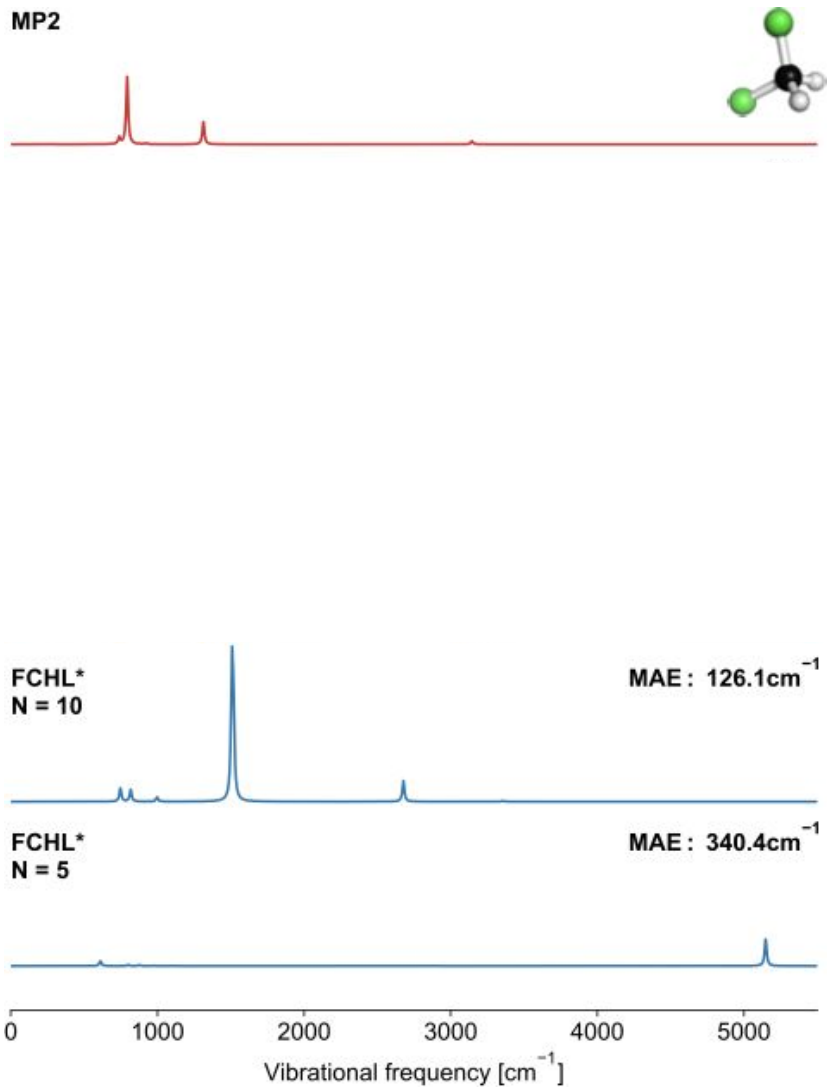
"Operators in Machine Learning: Response Properties in Chemical Space", Christensen et al, *J Chem Phys* (2019), arxiv.org/abs/1807.08811

→ IR: normal modes & dipole



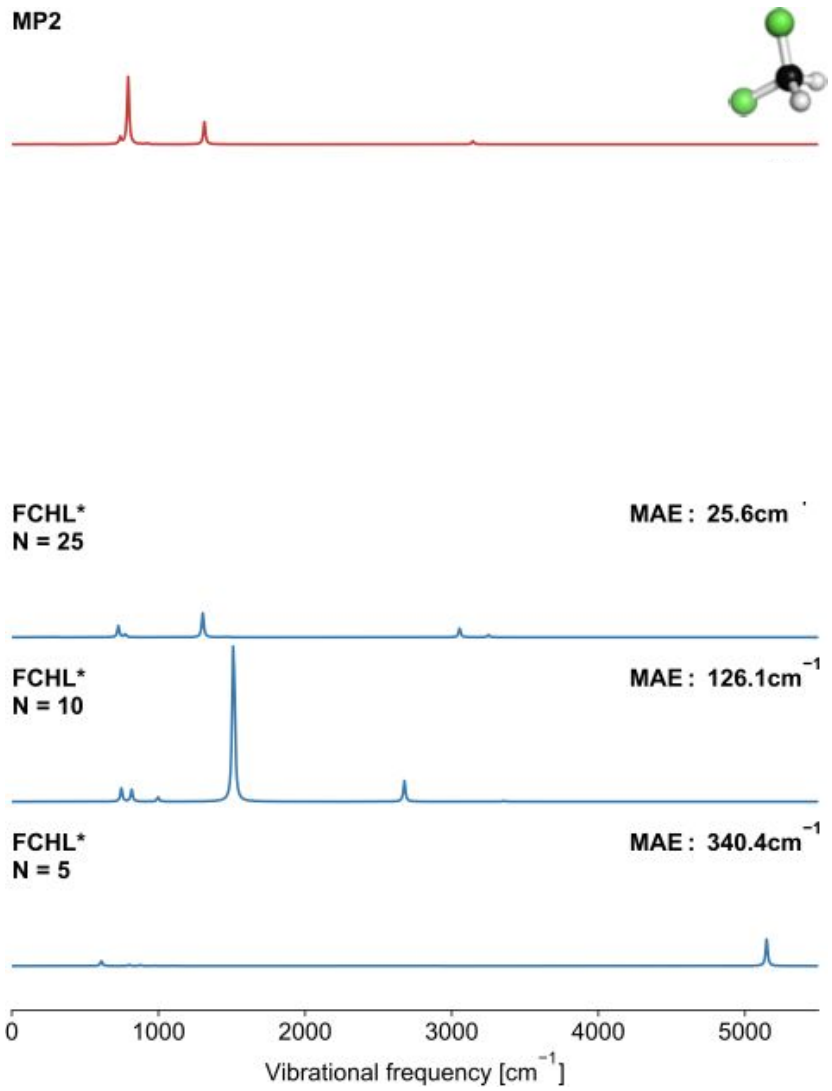
"Operators in Machine Learning: Response Properties in Chemical Space", Christensen et al, *J Chem Phys* (2019), arxiv.org/abs/1807.08811

→ IR: normal modes & dipole



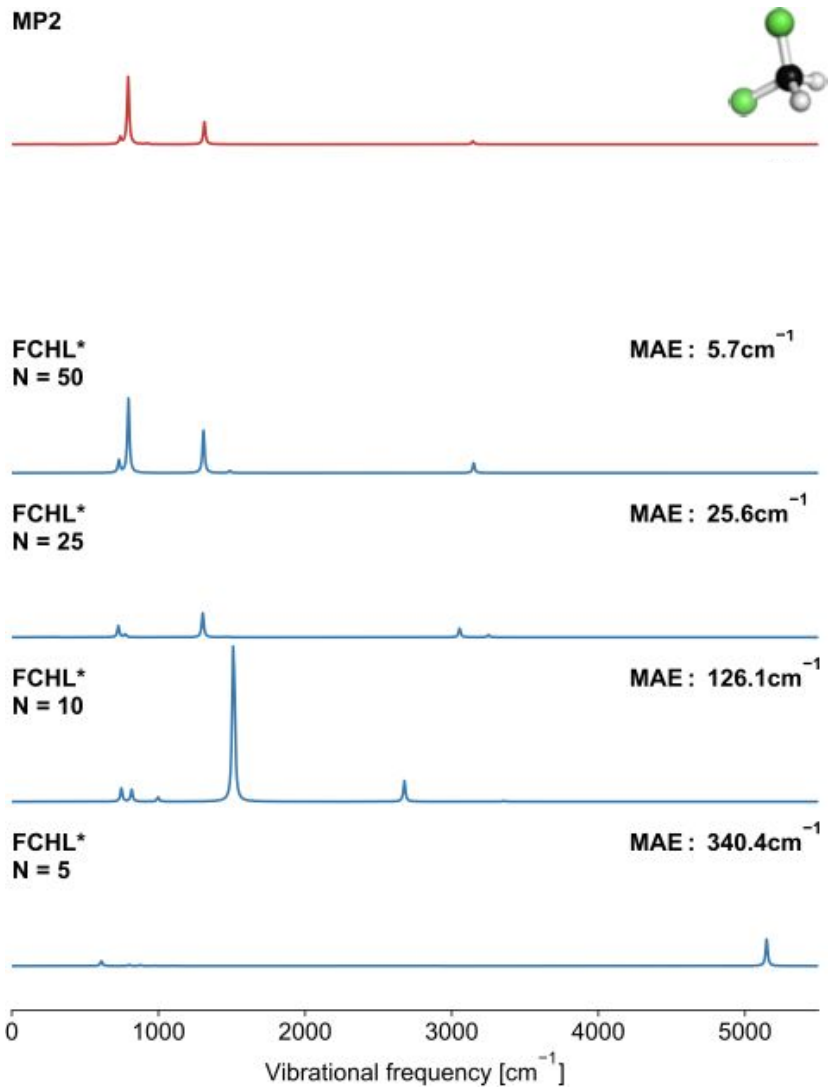
"Operators in Machine Learning: Response Properties in Chemical Space", Christensen et al, *J Chem Phys* (2019), arxiv.org/abs/1807.08811

→ IR: normal modes & dipole



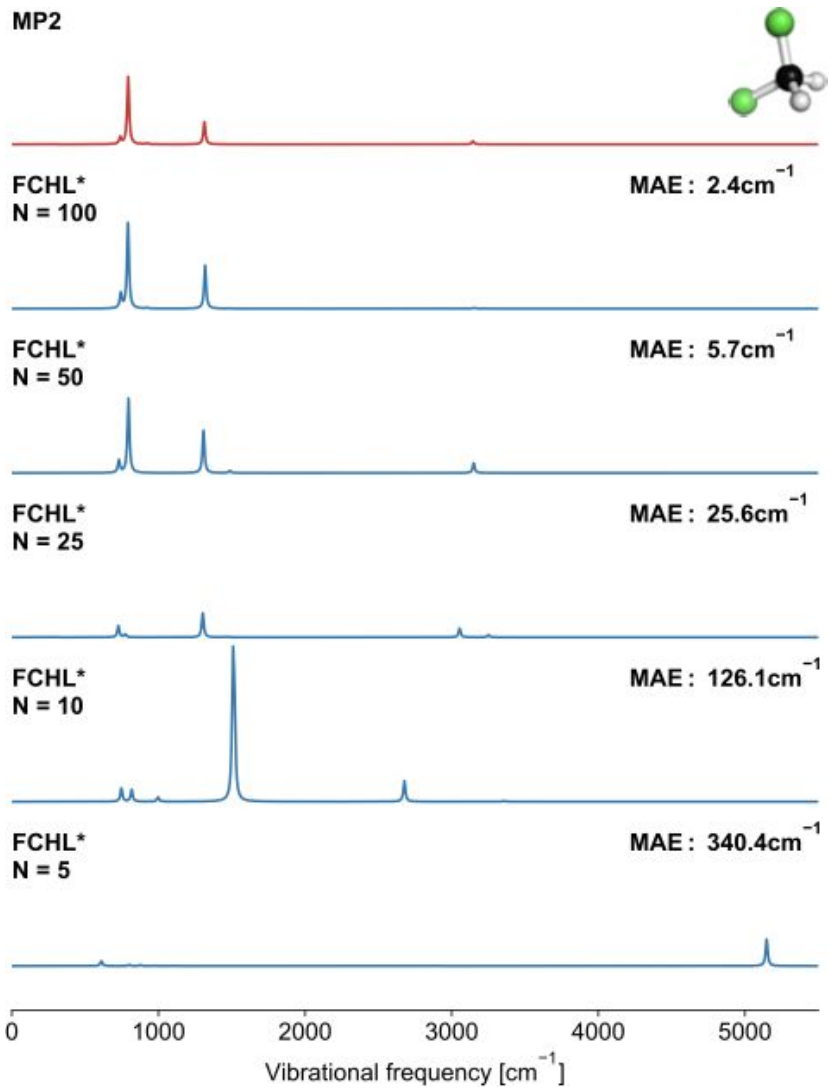
"Operators in Machine Learning: Response Properties in Chemical Space", Christensen et al, *J Chem Phys* (2019), arxiv.org/abs/1807.08811

→ IR: normal modes & dipole

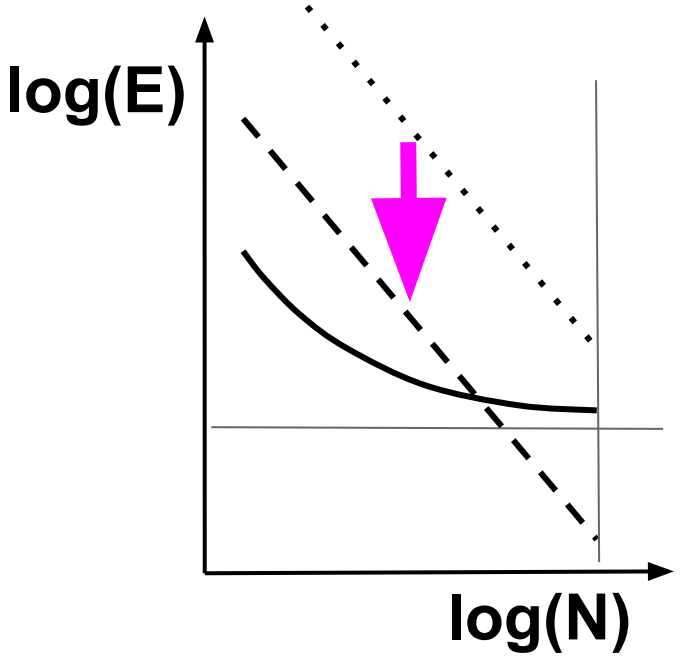


"Operators in Machine Learning: Response Properties in Chemical Space", Christensen et al, *J Chem Phys* (2019), arxiv.org/abs/1807.08811

→ IR: normal modes & dipole



"Operators in Machine Learning: Response Properties in Chemical Space", Christensen et al, *J Chem Phys* (2019), arxiv.org/abs/1807.08811



How???

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

Error ~ a/N^b

→ log(Error) = log(a) - $b$log(N)

`Quantum Machine Learning', von Lilienfeld, *Angew. Chem. Int. Ed.* (2018)

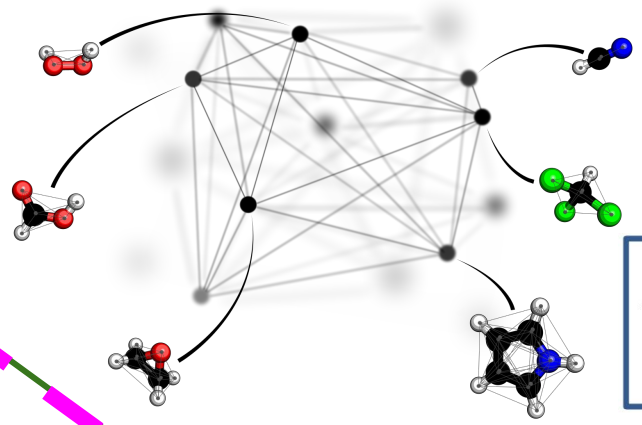
Vapnik, *The Nature of Statistical Learning Theory,* Springer (1995)

Error [Energy]

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

FF

SE



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

15 kcal/mol

HF

Error ~ a/N^b

→ log(Error) = log(a) - $b$log(N)

DFT

5 kcal/mol

MP2

Vapnik

1 kcal/mol

QM: > 1000 seconds

QML: ~milli seconds

QMC

CCSD(T)

QML

10 100 1’000 10’000 100’000

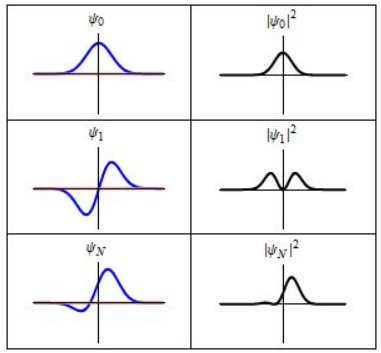
Training examples

Error [Energy]

Model chemistry

1998

Pople

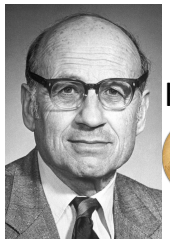


FF

SE

15 kcal/mol

HF



Kohn

DFT

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

5 kcal/mol

MP2

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle$$

Schrödinger
1933

1 kcal/mol

QMC
CCSD(T)

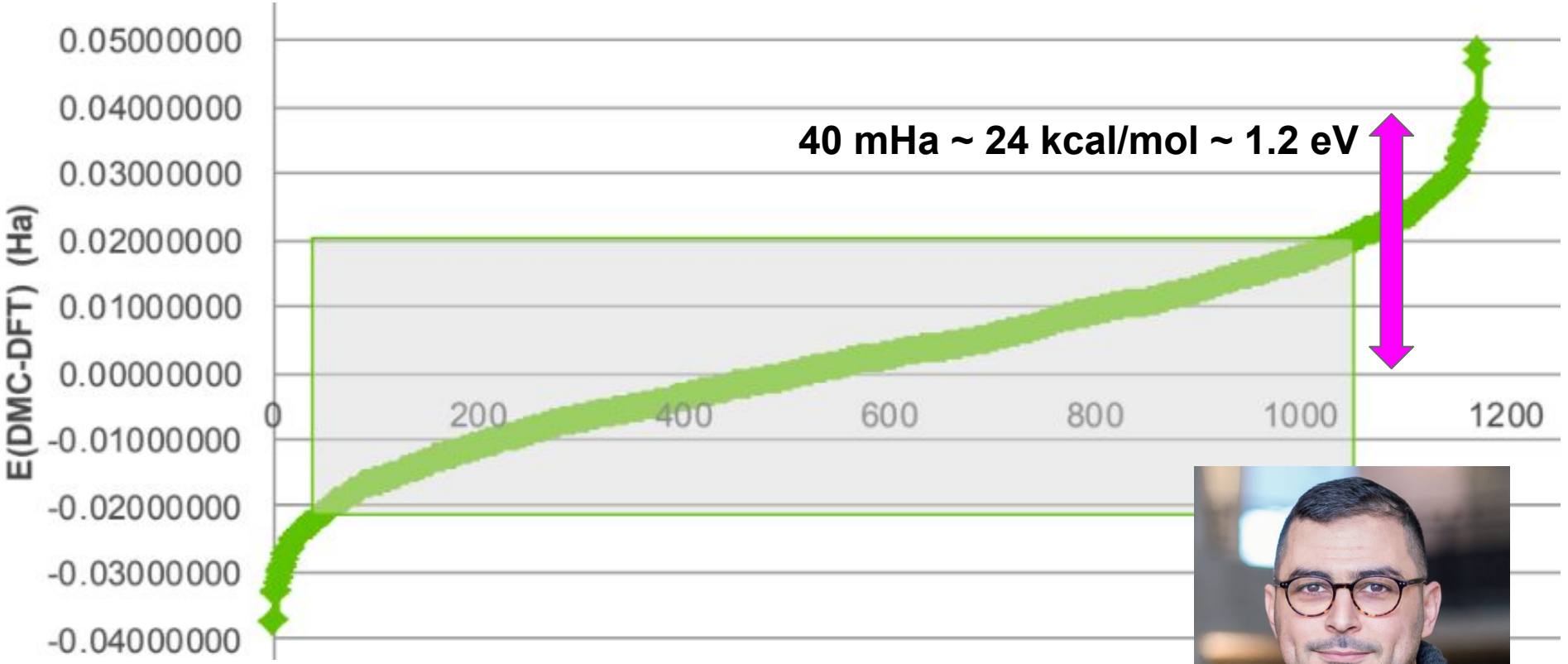
QML

ms min h d yr

Cost [CPU t]/compound

Is DFT great?

Deviation of hybrid DFT (B3LYP) from QMC atomization energy for 1.2k small neutral relaxed organic molecules with < 6 heavy atoms (not counting H)

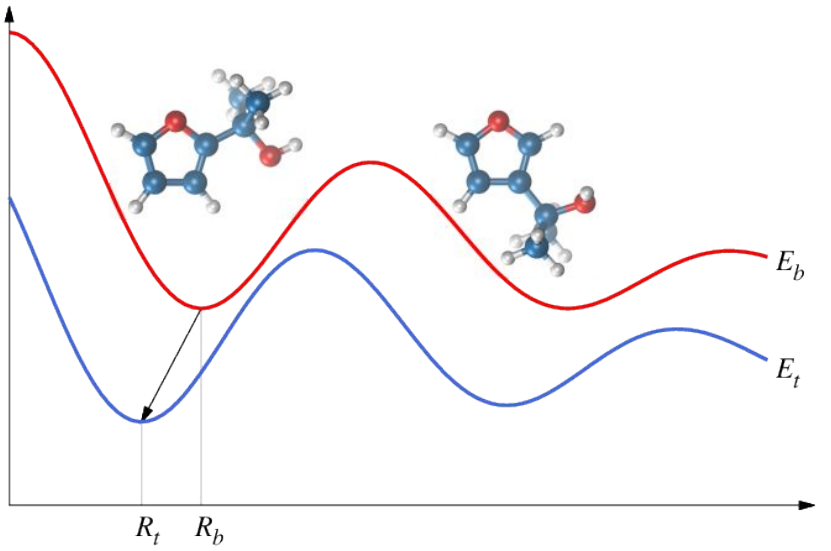


See also:
``Big Data meets Quantum Chemistry Approximations: The **Δ-Machine Learning** Approach'' Ramakrishnan et al, *JCTC* (2015) [over 170 citations, Google Scholar]

Anouar Benali

Δ-ML

$$E_t(R_t) \approx E_b(R_b) + \Delta_b^t(R_b)$$



JCTC Journal of Chemical Theory and Computation — Article
pubs.acs.org/JCTC

Big Data Meets Quantum Chemistry Approximations: The Δ-Machine Learning Approach

Raghunathan Ramakrishnan,[†] Pavlo O. Dral,[¶,‡] Matthias Rupp,[†] and O. Anatole von Lilienfeld*[,†,§]

[†]Institute of Physical Chemistry and National Center for Computational Design and Discovery of Novel Materials, Department of Chemistry, University of Basel, Klingelbergstraße 80, CH-4056 Basel, Switzerland

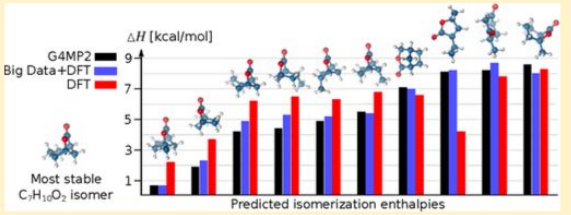
[¶]Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, 45470 Mülheim an der Ruhr, Germany

[‡]Computer-Chemie-Centrum and Interdisciplinary Center for Molecular Materials, Department Chemie und Pharmazie, Friedrich-Alexander-Universität Erlangen-Nürnberg, Nägelsbachstraße 25, 91052 Erlangen, Germany

[§]Argonne Leadership Computing Facility, Argonne National Laboratory, 9700 S. Cass Avenue, Lemont, Illinois 60439, United States

Supporting Information

ABSTRACT: Chemically accurate and comprehensive studies of the virtual space of all possible molecules are severely limited by the computational cost of quantum chemistry. We introduce a composite strategy that adds machine learning corrections to computationally inexpensive approximate legacy quantum methods. After training, highly accurate predictions of enthalpies, free energies, entropies, and electron correlation energies are possible, for significantly larger molecular sets than used for training. For thermochemical properties of up to 16k isomers of $C_7H_{10}O_2$ we present numerical evidence that chemical accuracy can be reached. We also predict electron correlation energy in post Hartree−Fock methods, at the computational cost of Hartree−Fock, and we establish a qualitative relationship between molecular entropy and electron correlation. The transferability of our approach is demonstrated, using

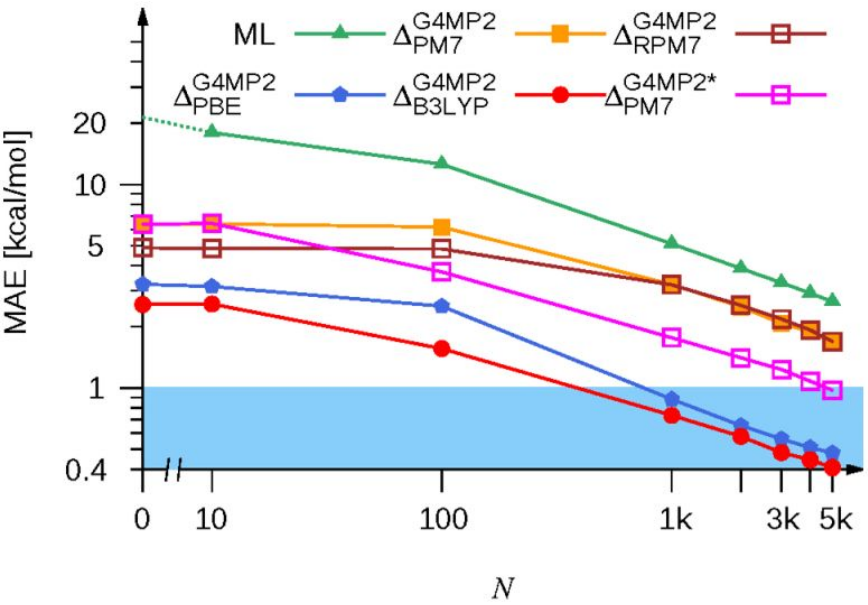


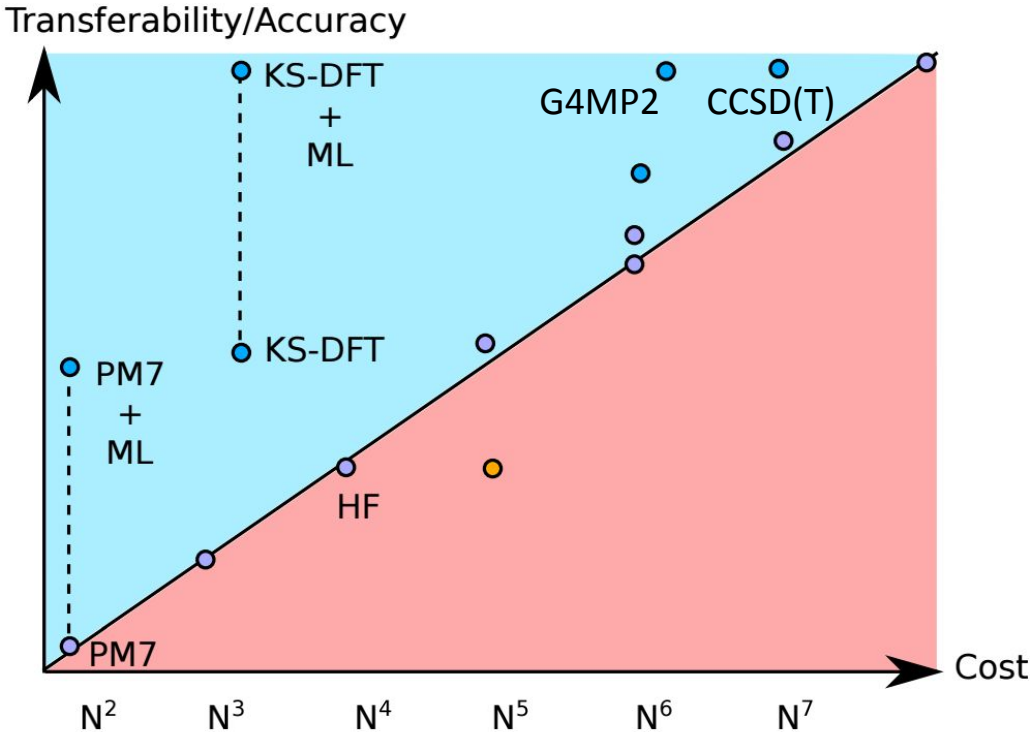
Ramakrishnan et al, *J Chem Theory Comput* (2015)

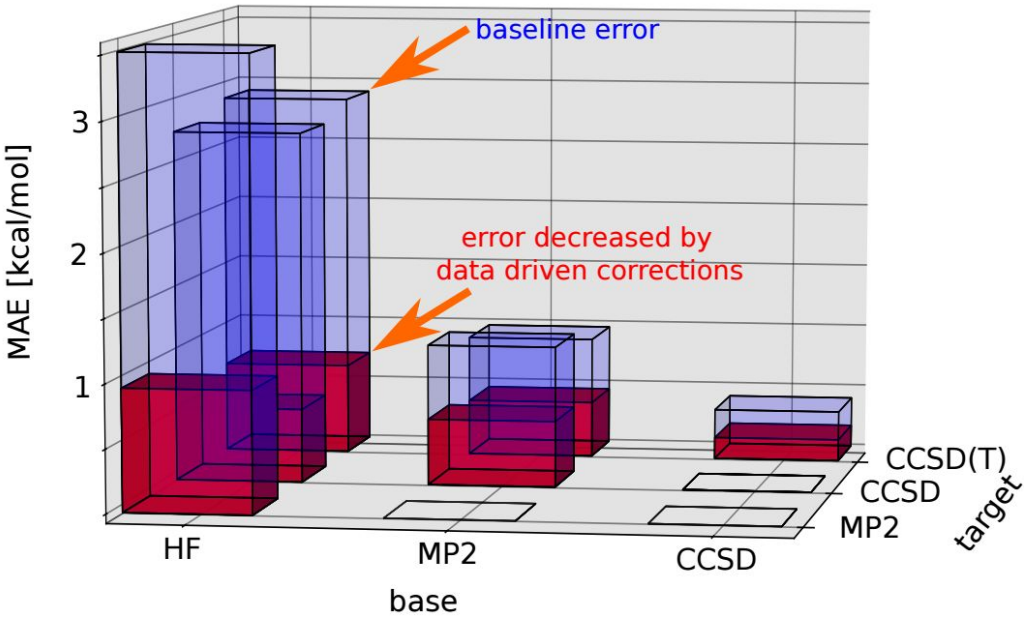
Δ-ML

$$E_t(R_t) \approx E_b(R_b) + \Delta_b^t(R_b)$$

6k constitutional isomers of $C_7O_2H_{10}$





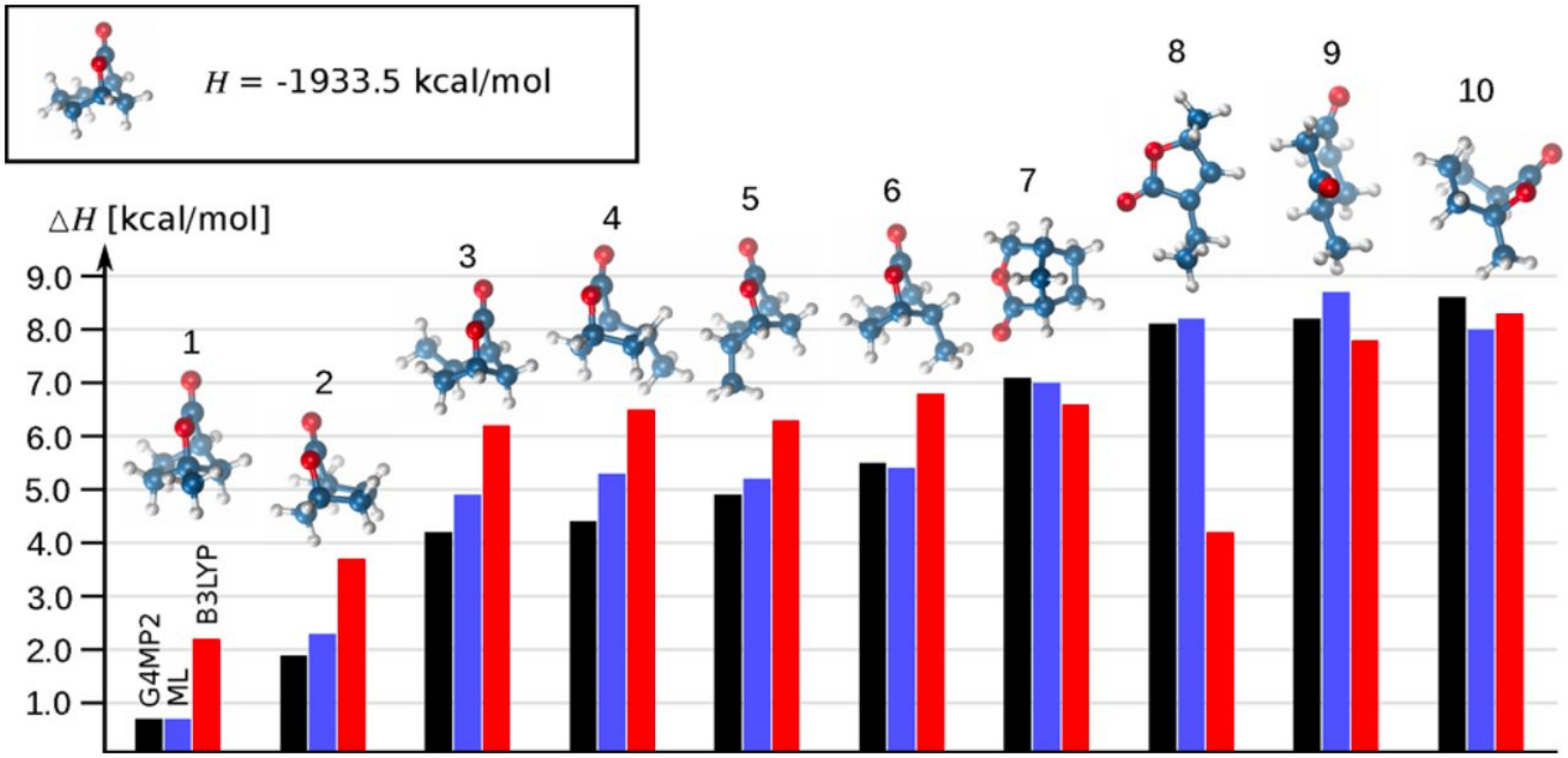


Ramakrishnan et al, *J Chem Theory Comput* (2015)

Δ-ML

$$E_t(R_t) \approx E_b(R_b) + \Delta_b^t(R_b)$$

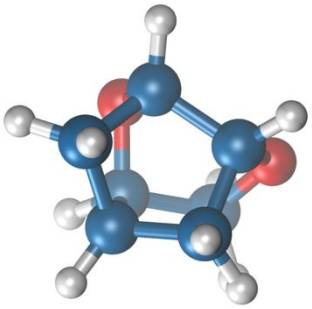
Ranking 10k diastereomers derived from 6k constitutional isomers of $C_7O_2H_{10}$
→ Global minimum, and its 10 closest isomers …



Ramakrishnan et al, *J Chem Theory Comput* (2015)

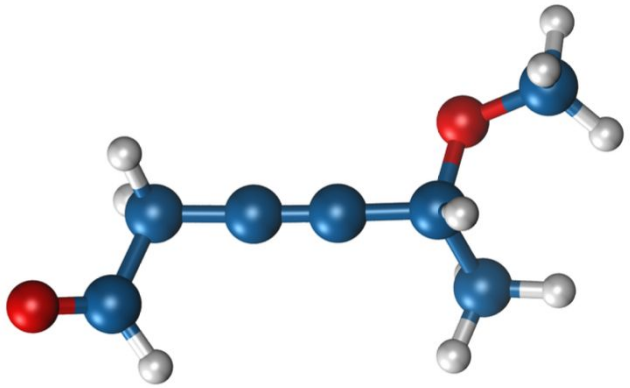
Δ-ML

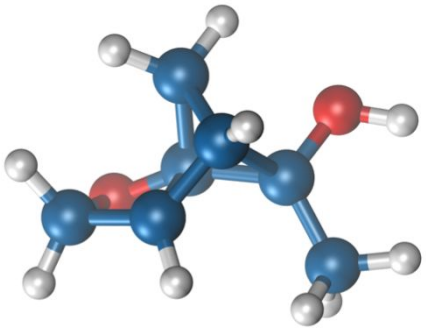
$$\left(\Delta_E^H - \Delta_E^G\right) / T = S$$



$\triangle S$ = 32 cal/mol/K

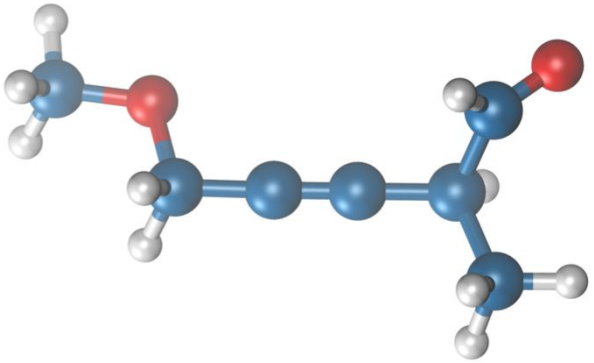
(31 cal/mol/K)



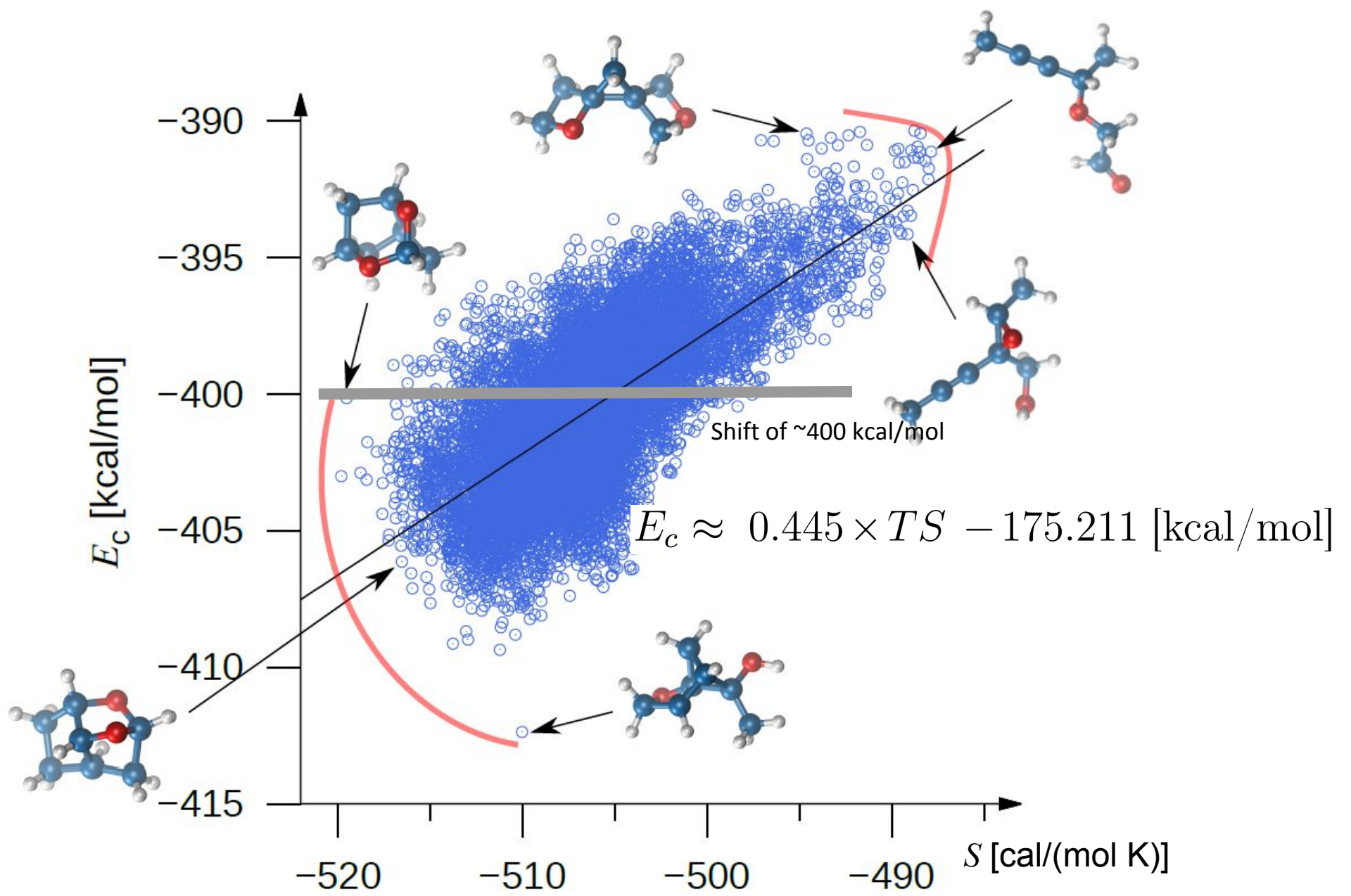


$\triangle E_c$ = 24 kcal/mol

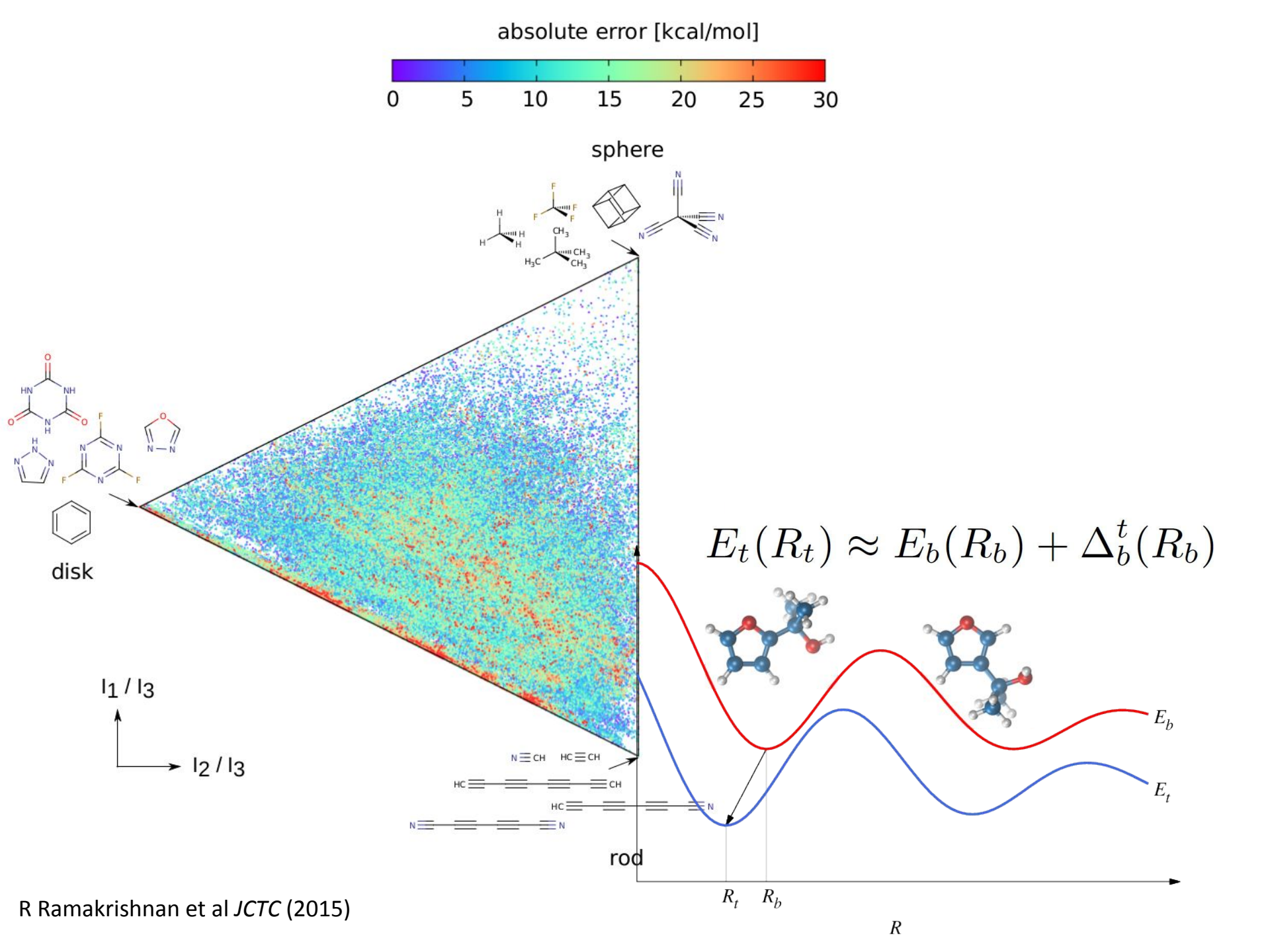
(23 kcal/mol)



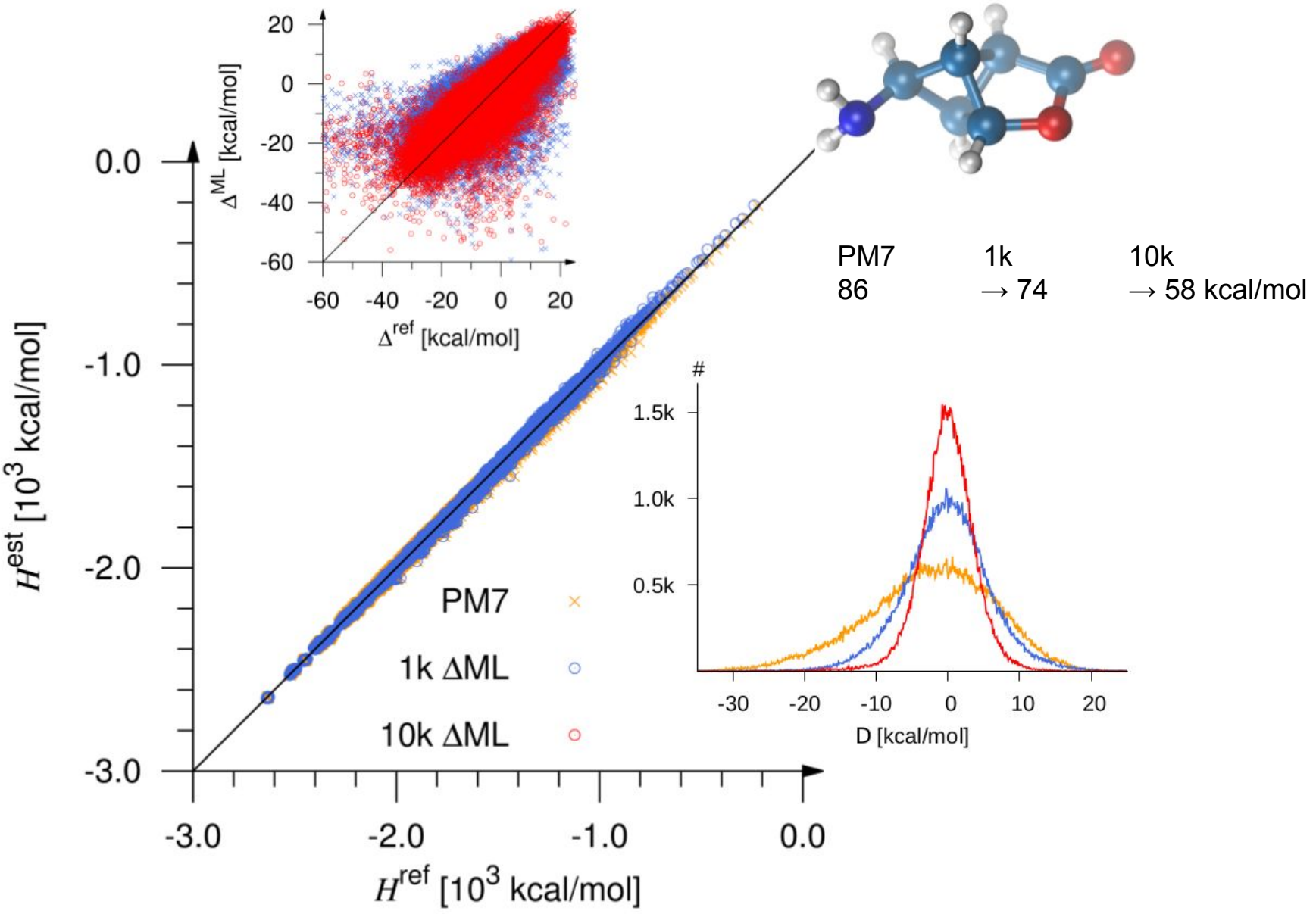
Ramakrishnan et al, *J Chem Theory Comput* (2015)



Ramakrishnan et al, *J Chem Theory Comput* (2015)

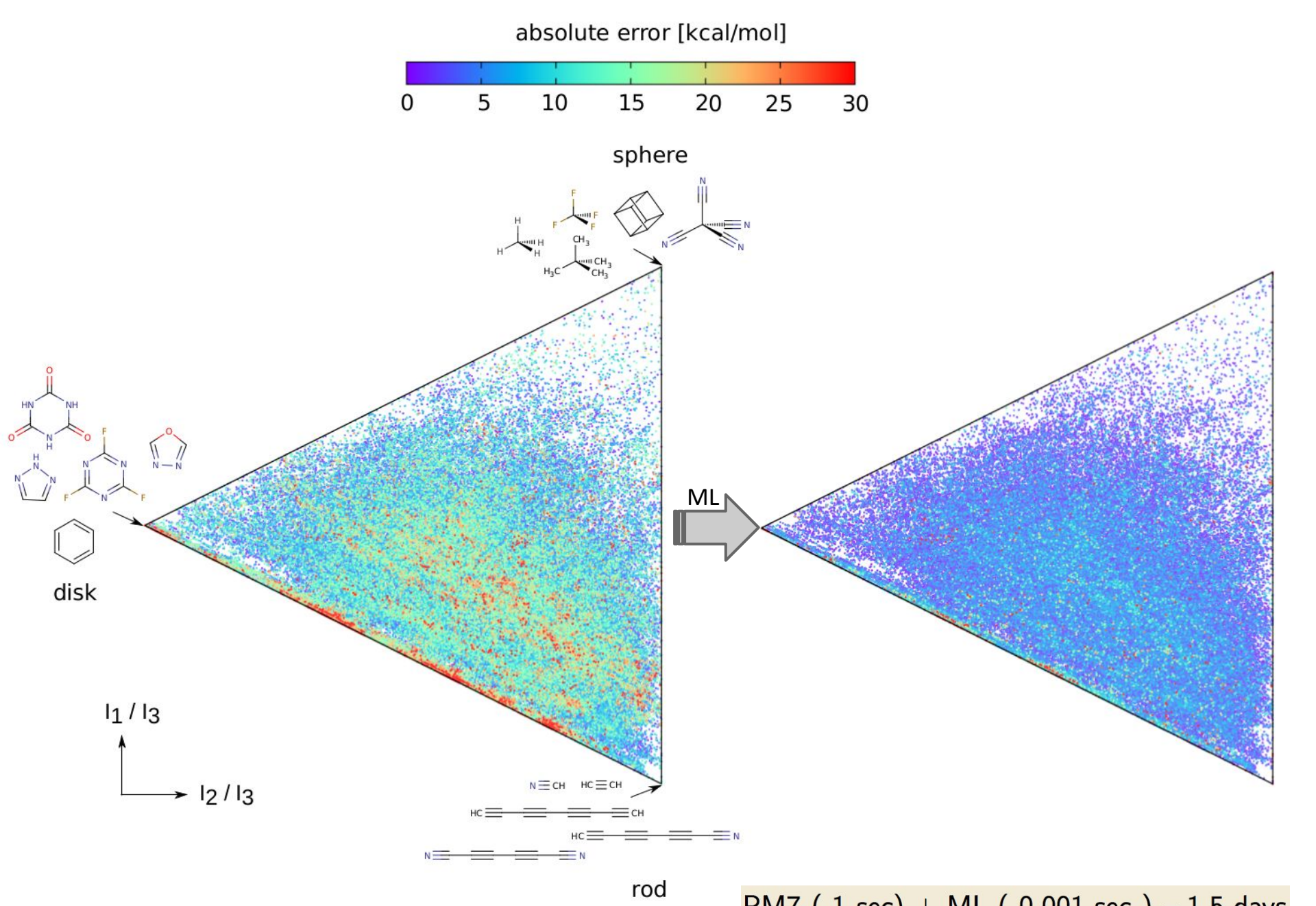


R Ramakrishnan et al *JCTC* (2015)



PM7 86 | 1k → 74 | 10k → 58 kcal/mol

R Ramakrishnan et al *JCTC* (2015)



R Ramakrishnan et al *JCTC* (2015)

PM7 (1 sec) + ML (0.001 sec) – 1.5 days

DFT (30 min) – 8 yrs !

FCHL based NCI-SCAN

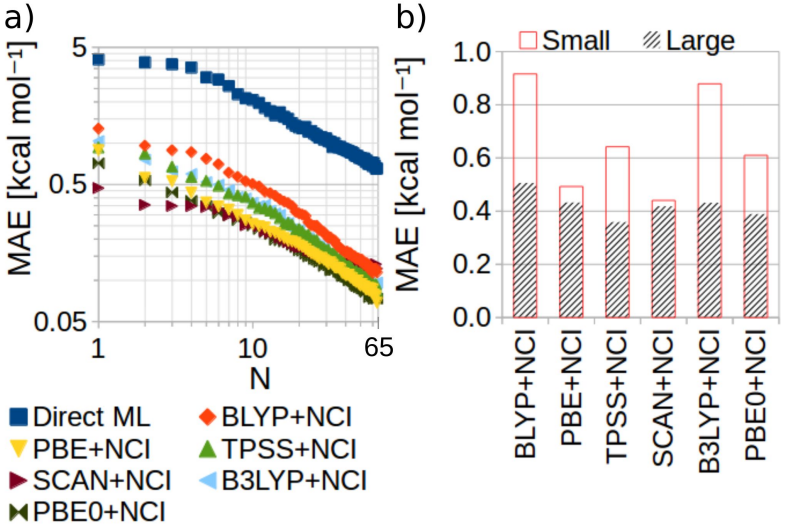
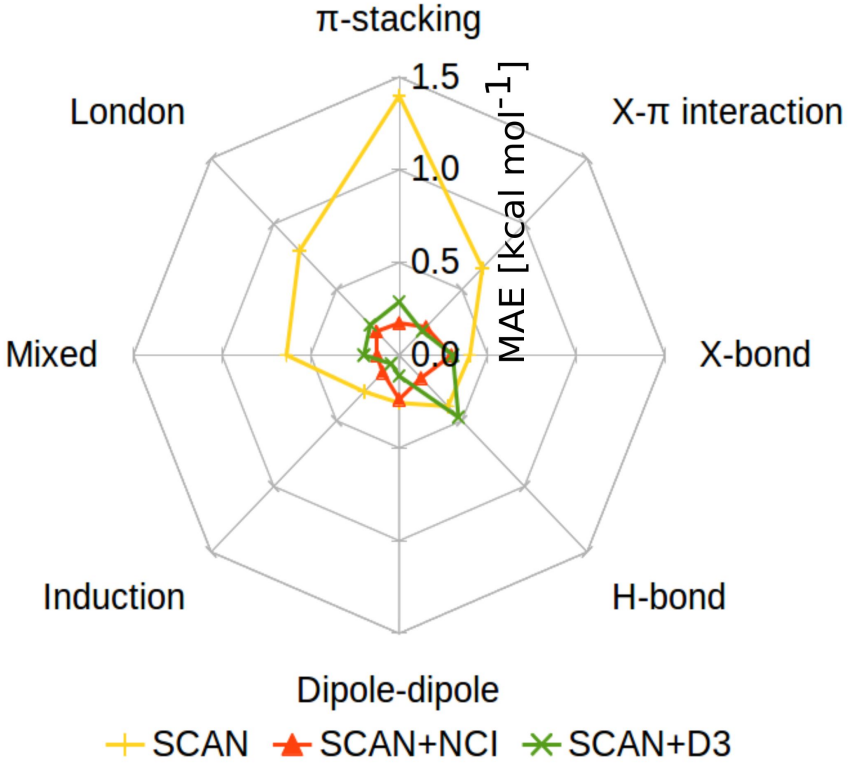


TABLE I. Mean absolute errors (kcal mol^{-1}) in the interaction energies for the different test sets with our most robust NCI approach or D3 corrections.

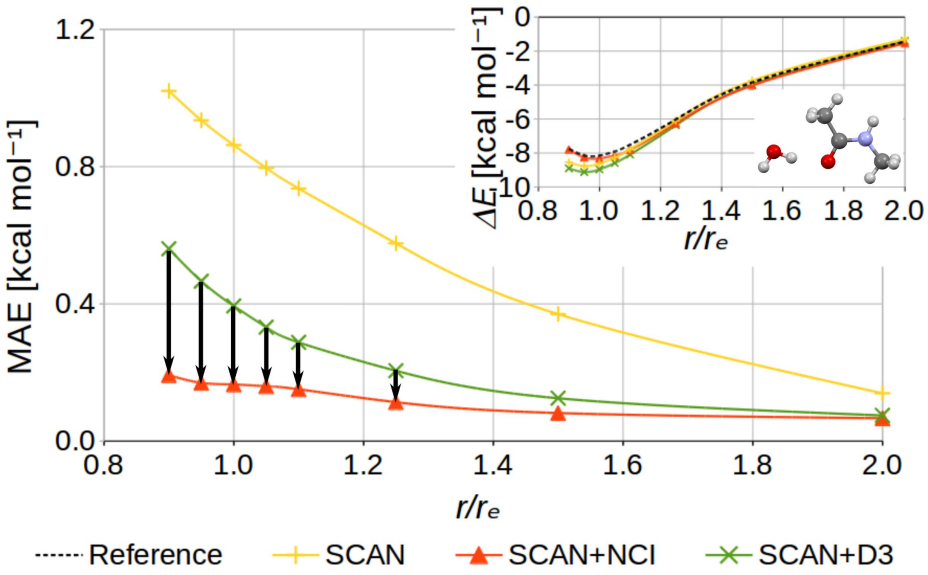
Dataset	SCAN	+NCI	+D3
Biochemically relevant [46, 47]	0.70	0.14	0.25
Non-covalent blind test [48, 49]	0.40	0.25	0.16
Halogen-containing [50, 51]	0.49	0.24	0.36
Water clusters [52–55]	4.85	0.94	6.89
Molecular crystals[a] [56–60]	4.69	2.38	2.20
Host-guest complexes[a] [61, 62]	9.98	3.66	1.35

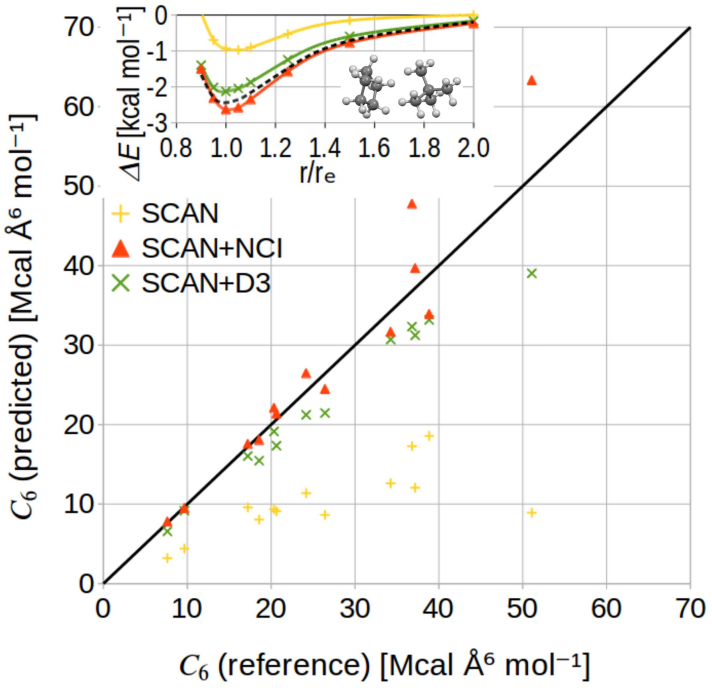
[a] Not represented in the training set.

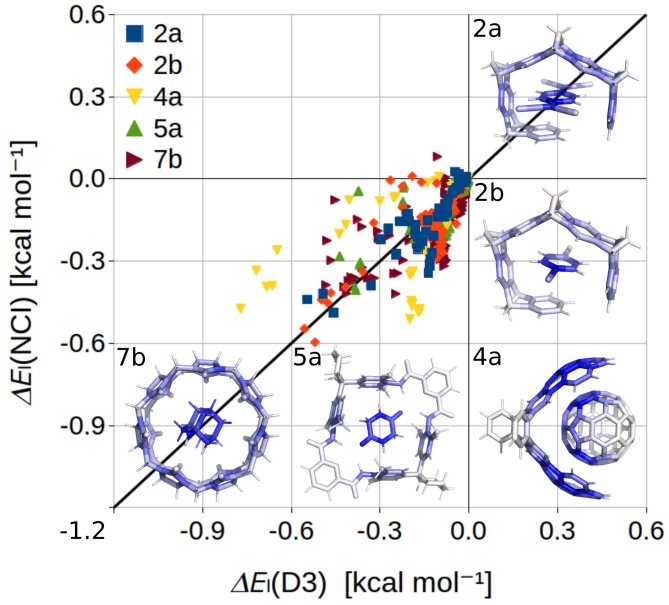


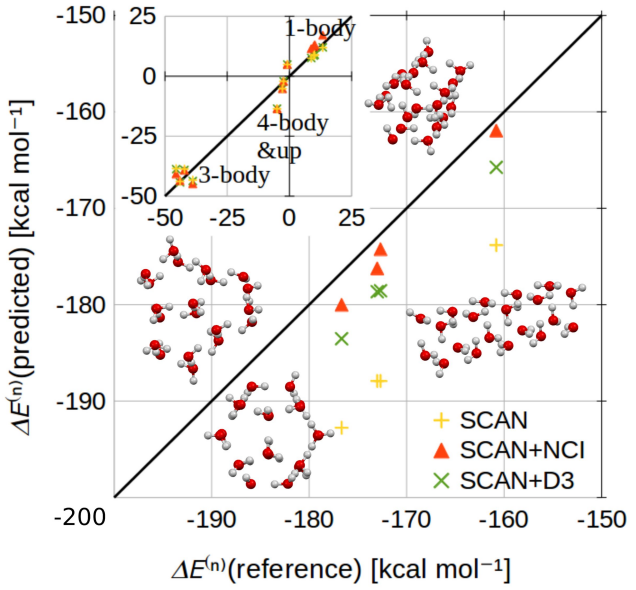
" Non-covalent quantum machine learning corrections to density functionals", P. Mezei, OAvL, *submitted* (2019), https://arxiv.org/abs/1903.09010

FCHL based NCI-SCAN









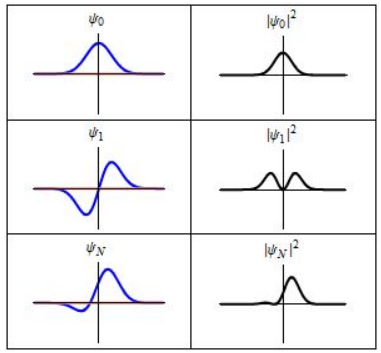
" Non-covalent quantum machine learning corrections to density functionals", P. Mezei, OAvL, *submitted* (2019), https://arxiv.org/abs/1903.09010

Error [Energy]

Model chemistry

1998

Pople



FF

SE

15 kcal/mol

HF

John

DFT

5 kcal/mol

MP2

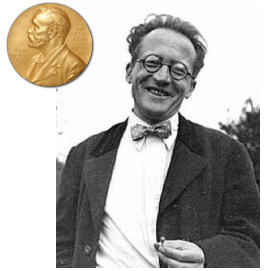
1 kcal/mol

QML

QMC
CCSD(T)

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

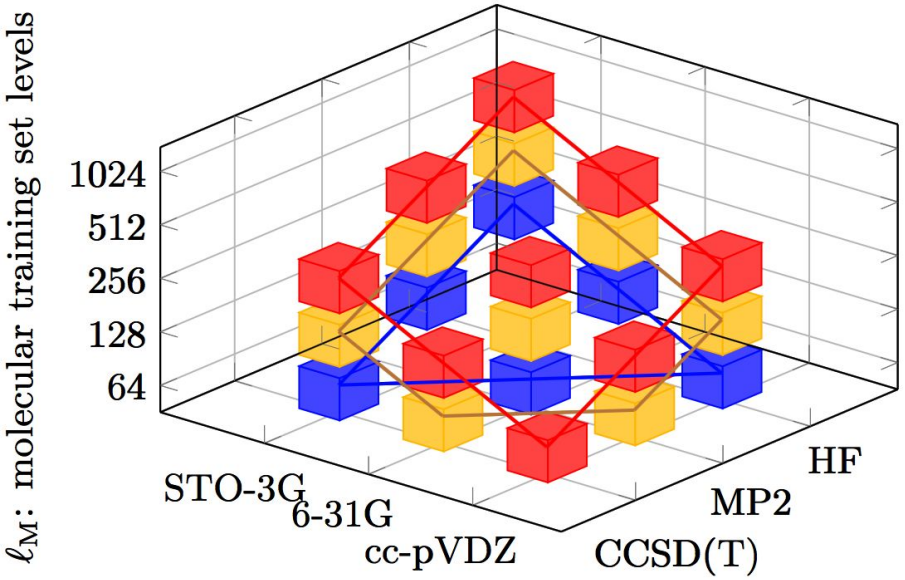
$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle$$



Schrödinger
1933

ms min h d yr

Cost [CPU t]/compound



ℓ_B: basis set levels

ℓ_C: correlation levels

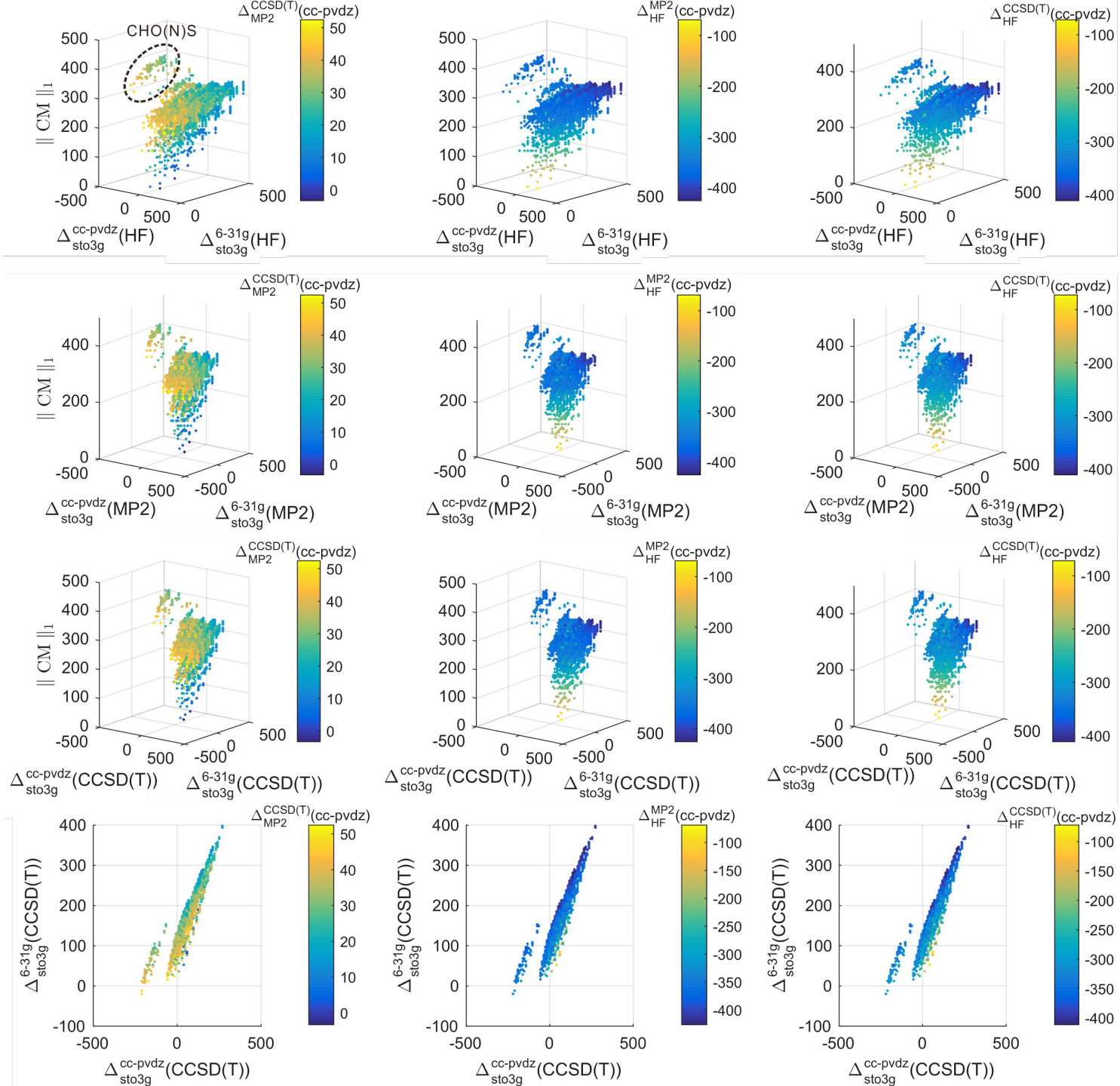
s	$r_{\ell_M=L-1}$	$r_{\ell_M=L-2}$	$r_{\ell_M=L-3}$
1	1	2	4
2	1	4	16

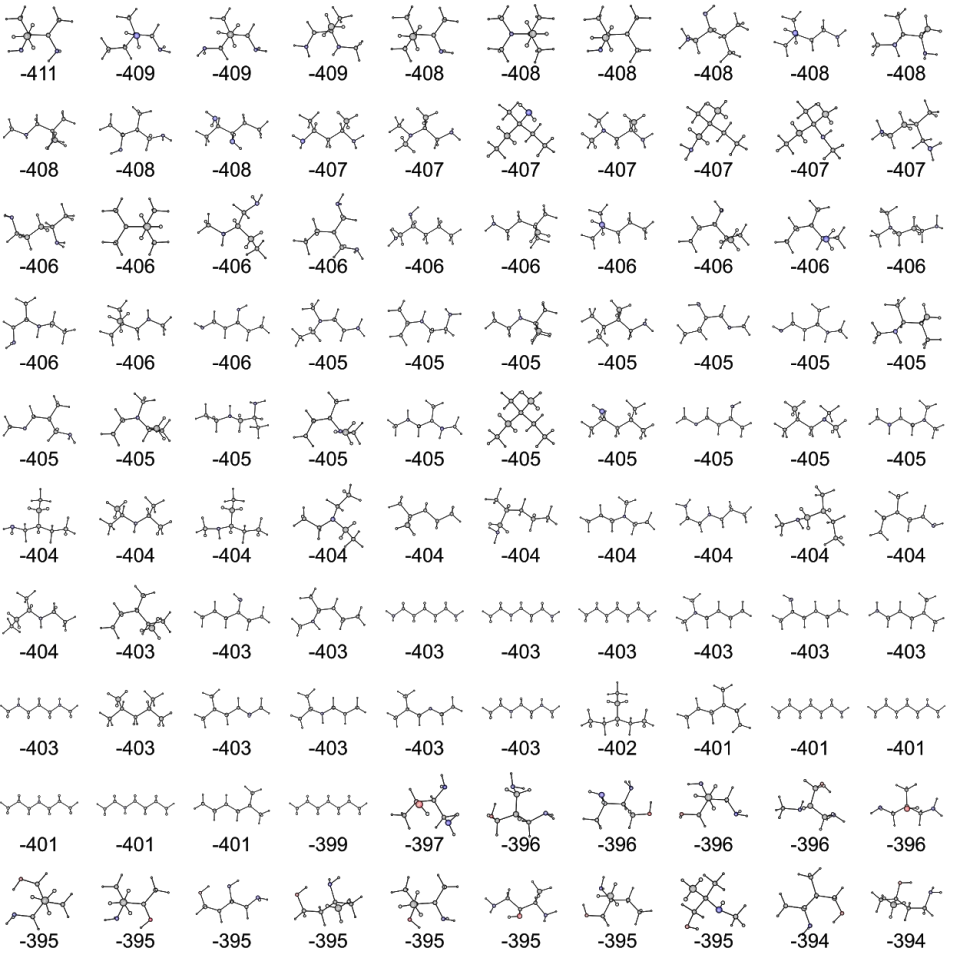
level	0	1	2
ℓ_C	HF	MP2	CCSD(T)
ℓ_B	sto-3g	6-31g	cc-pvdz
ℓ_M	N_0	N_1	N_2

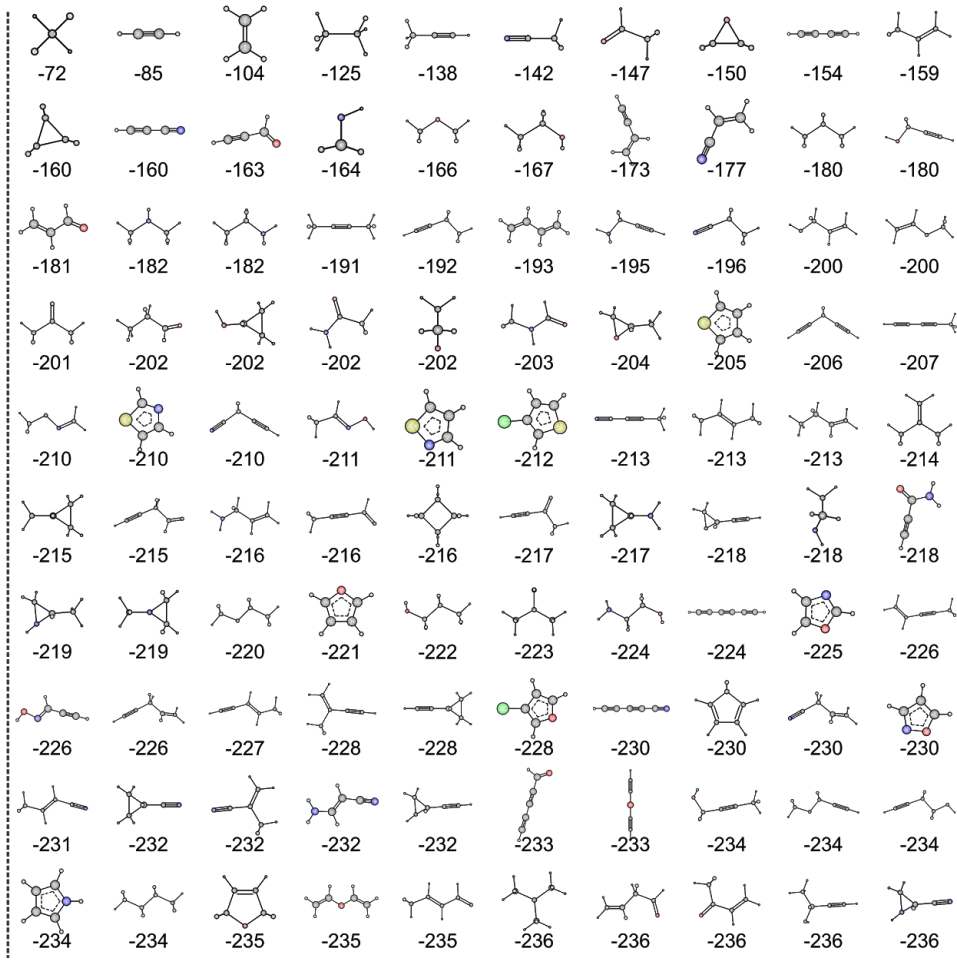
Multi-level sparse grid **C**ombination technique (CQML)

$$\mathcal{M}^{\ell}_{\ell_1}(\boldsymbol{R}_q) := \mathcal{M}^{\ell-1}_{\ell_1}(\boldsymbol{R}_q) + \sum_{i=1}^{N_\ell} \alpha_i^{(\ell)} k\left(\boldsymbol{R}_q, \boldsymbol{R}_i\right)$$

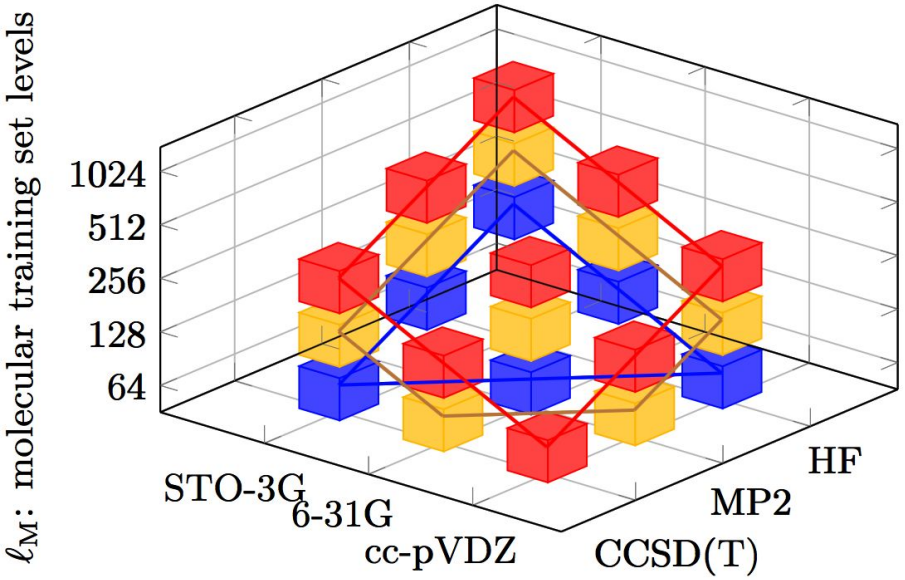
P Zaspel, H Harbrecht, B Huang, OAvL *JCTC* (2018)


CHO(N)S
$\| \mathrm{CM} \|_1$
$\Delta^{\mathrm{CCSD(T)}}_{\mathrm{MP2}}$(cc-pvdz)
$\Delta^{\mathrm{MP2}}_{\mathrm{HF}}$(cc-pvdz)
$\Delta^{\mathrm{CCSD(T)}}_{\mathrm{HF}}$(cc-pvdz)
$\Delta^{\mathrm{cc\text{-}pvdz}}_{\mathrm{sto3g}}$(HF)
$\Delta^{\mathrm{6\text{-}31g}}_{\mathrm{sto3g}}$(HF)
$\Delta^{\mathrm{cc\text{-}pvdz}}_{\mathrm{sto3g}}$(MP2)
$\Delta^{\mathrm{6\text{-}31g}}_{\mathrm{sto3g}}$(MP2)
$\Delta^{\mathrm{cc\text{-}pvdz}}_{\mathrm{sto3g}}$(CCSD(T))
$\Delta^{\mathrm{6\text{-}31g}}_{\mathrm{sto3g}}$(CCSD(T))





P Zaspel, H Harbrecht, B Huang, OAvL *JCTC* (2018)



ℓ_B: basis set levels

ℓ_C: correlation levels

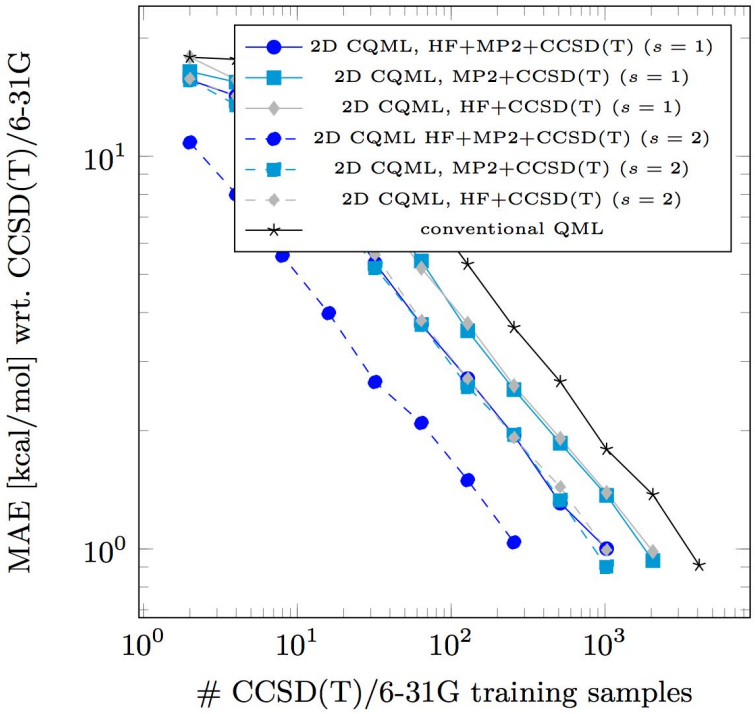
s	$r_{\ell_M=L-1}$	$r_{\ell_M=L-2}$	$r_{\ell_M=L-3}$
1	1	2	4
2	1	4	16

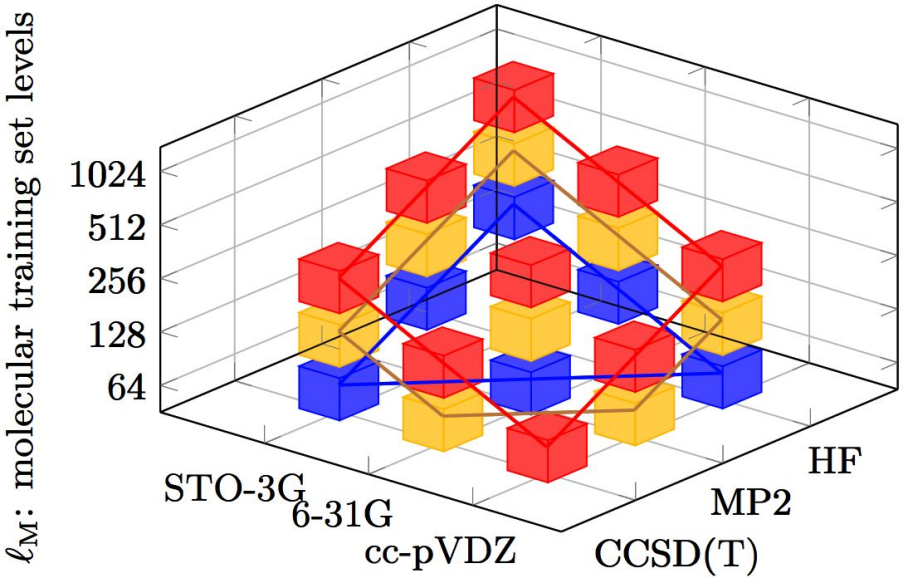
level	0	1	2
ℓ_C	HF	MP2	CCSD(T)
ℓ_B	sto-3g	6-31g	cc-pvdz
ℓ_M	N_0	N_1	N_2

Multi-level sparse grid **C**ombination technique (CQML)

$$\mathcal{M}^{\ell}_{\ell_1}(\boldsymbol{R}_q) := \mathcal{M}^{\ell-1}_{\ell_1}(\boldsymbol{R}_q) + \sum_{i=1}^{N_\ell} \alpha_i^{(\ell)} k\left(\boldsymbol{R}_q, \boldsymbol{R}_i\right)$$

P Zaspel, H Harbrecht, B Huang, OAvL *JCTC* (2018)





level	0	1	2
ℓ_C	HF	MP2	CCSD(T)
ℓ_B	sto-3g	6-31g	cc-pvdz
ℓ_M	N_0	N_1	N_2

ℓ_B: basis set levels

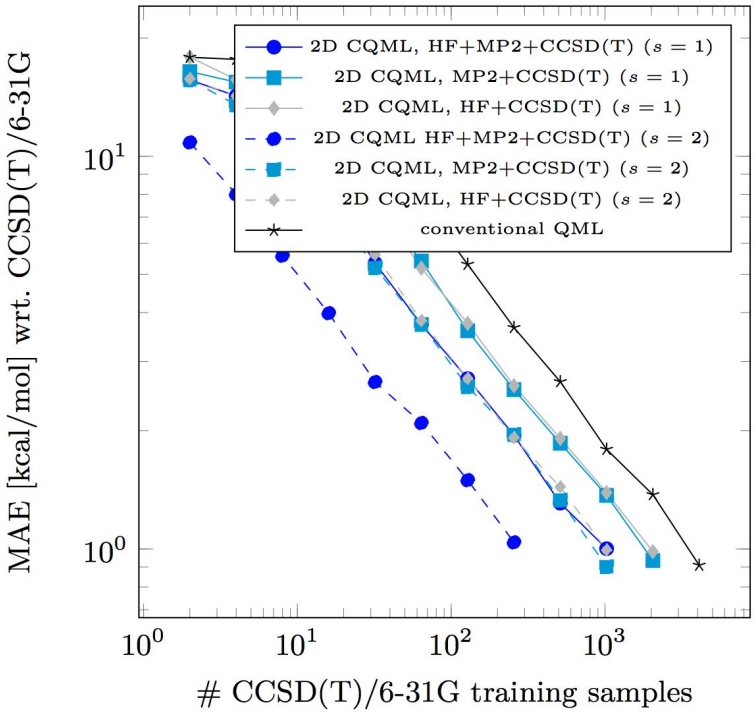
ℓ_C: correlation levels

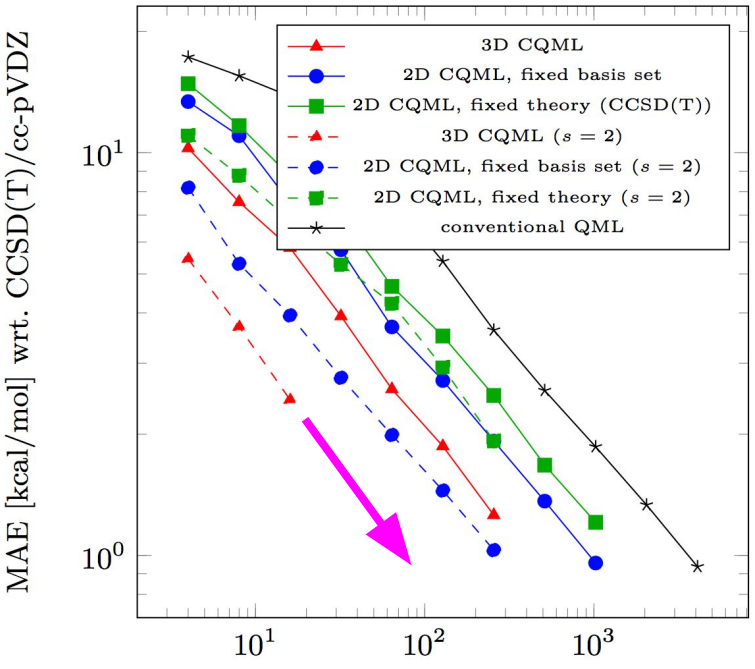
s	$r_{\ell_M=L-1}$	$r_{\ell_M=L-2}$	$r_{\ell_M=L-3}$
1	1	2	4
2	1	4	16

Multi-level sparse grid **C**ombination technique (CQML)

$$\mathcal{M}^{\ell}_{\ell_1}(\boldsymbol{R}_q) := \mathcal{M}^{\ell-1}_{\ell_1}(\boldsymbol{R}_q) + \sum_{i=1}^{N_\ell} \alpha_i^{(\ell)} k\left(\boldsymbol{R}_q, \boldsymbol{R}_i\right)$$

P Zaspel, H Harbrecht, B Huang, OAvL *JCTC* (2018)





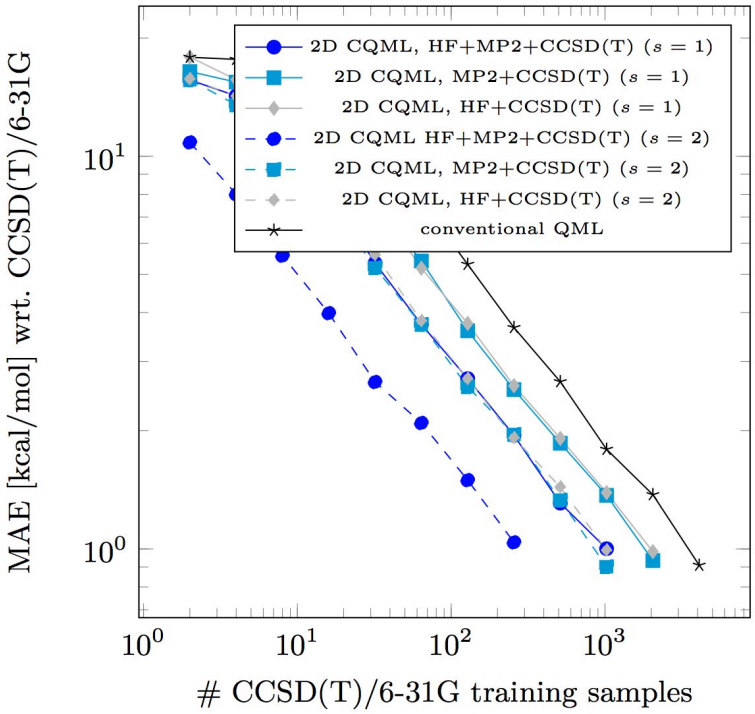
$level$	0	1	2
ℓ_{C}	HF	MP2	CCSD(T)
ℓ_{B}	sto-3g	6-31g	cc-pvdz
ℓ_{M}	N_0	N_1	N_2

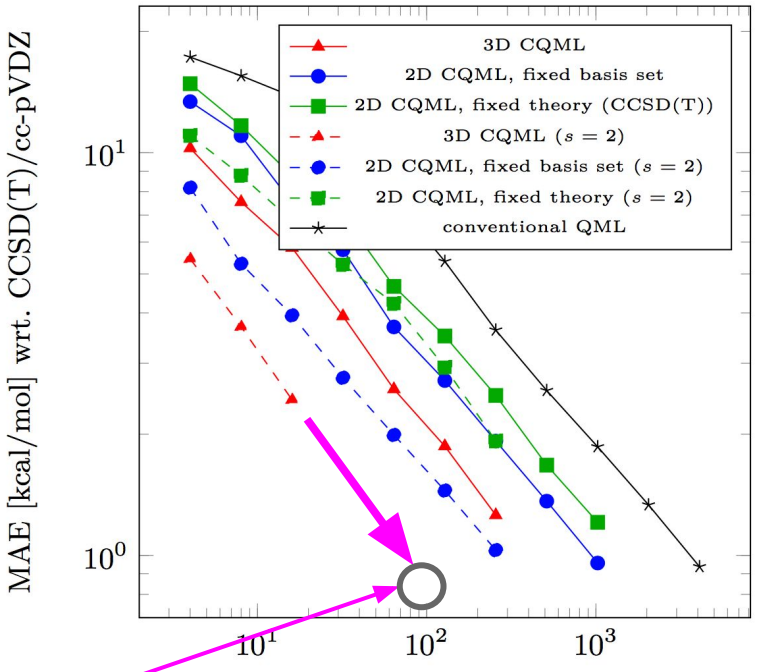
s	$r_{\ell_{\mathrm{M}}=L-1}$	$r_{\ell_{\mathrm{M}}=L-2}$	$r_{\ell_{\mathrm{M}}=L-3}$
1	1	2	4
2	1	4	16

Multi-level sparse grid **C**ombination technique (CQML)

$$\mathcal{M}^{\ell}_{\ell_1}(\boldsymbol{R}_q) := \mathcal{M}^{\ell-1}_{\ell_1}(\boldsymbol{R}_q) + \sum_{i=1}^{N_\ell} \alpha_i^{(\ell)} k\left(\boldsymbol{R}_q, \boldsymbol{R}_i\right)$$

P Zaspel, H Harbrecht, B Huang, OAvL *JCTC* (2018)

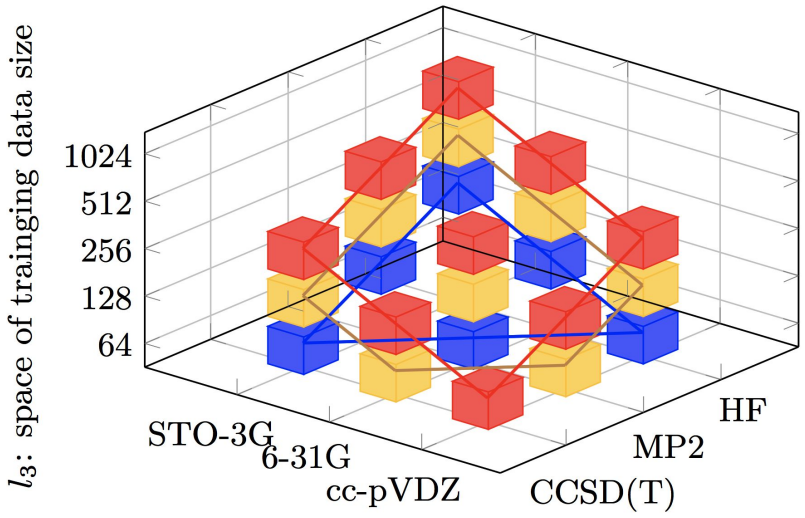




s	$r_{\ell_M=L-1}$	$r_{\ell_M=L-2}$	$r_{\ell_M=L-3}$
1	1	2	4
2	1	4	16

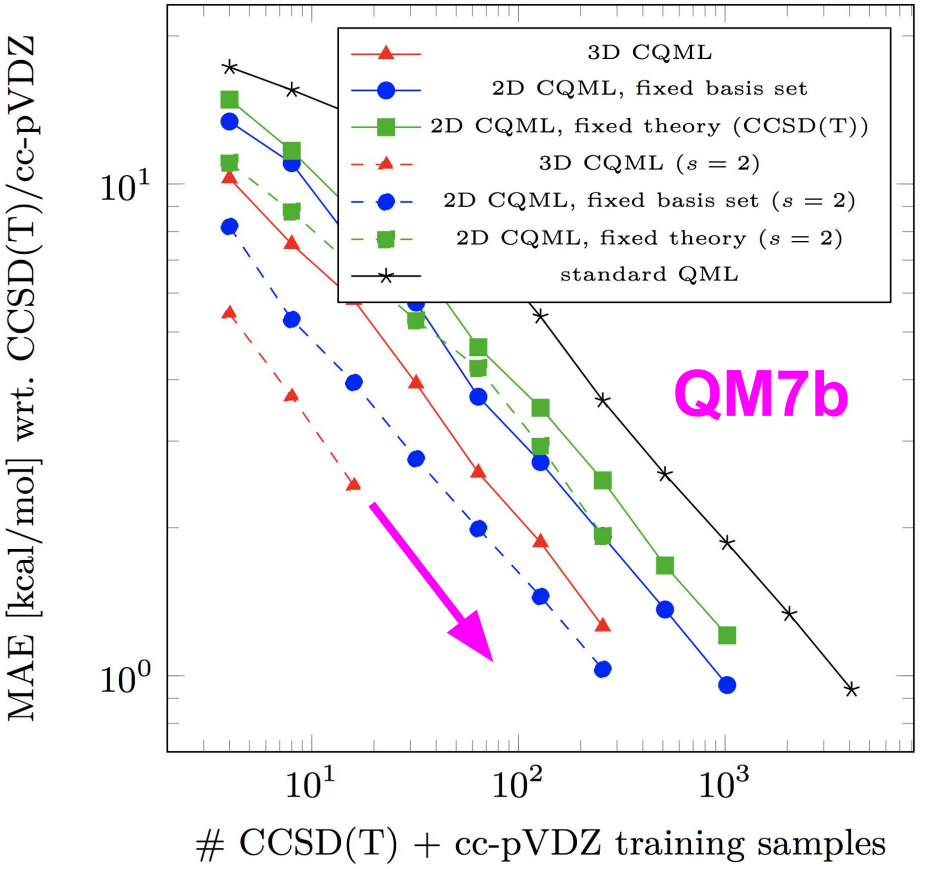
Would have required > 25k HF/STO-3G training molecules

P Zaspel, H Harbrecht, B Huang, OAvL *JCTC* (2018)



l_2: space of basis sets

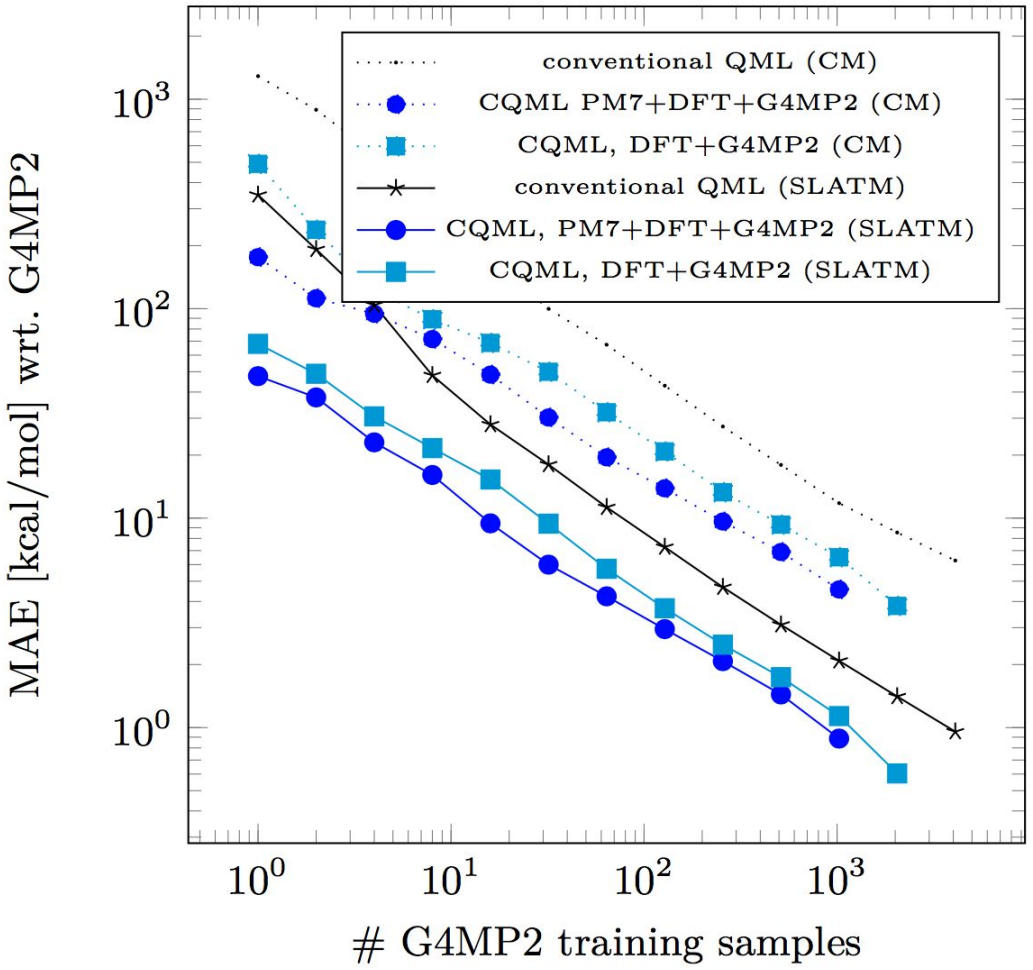
l_1: space of theories



Multi-level sparse grid **C**ombination technique (CQML)

$$\mathcal{M}^{\ell}_{\ell_1}(\boldsymbol{R}_q) := \mathcal{M}^{\ell-1}_{\ell_1}(\boldsymbol{R}_q) + \sum_{i=1}^{N_\ell} \alpha_i^{(\ell)} k\left(\boldsymbol{R}_q, \boldsymbol{R}_i\right)$$

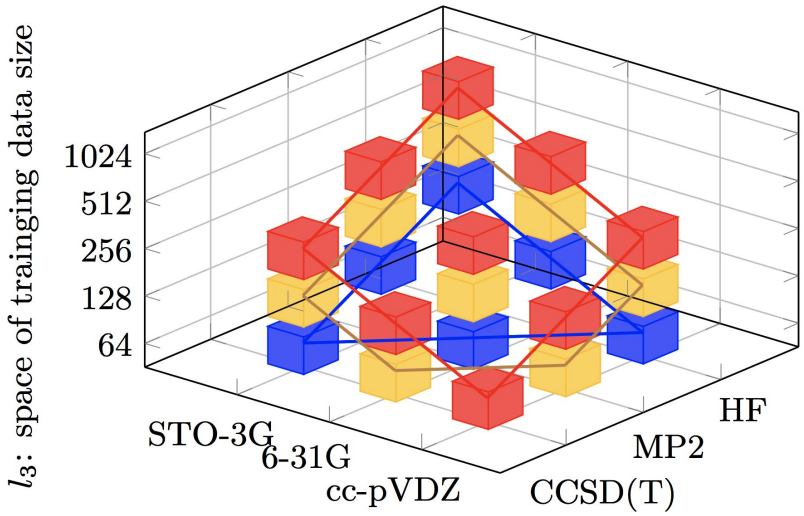
P Zaspel, H Harbrecht, B Huang, OAvL *JCTC* (2019)



Multi-level sparse grid **C**ombination technique (CQML)

$$\mathcal{M}_{\ell_1}^{\ell}(\boldsymbol{R}_q) := \mathcal{M}_{\ell_1}^{\ell-1}(\boldsymbol{R}_q) + \sum_{i=1}^{N_\ell} \alpha_i^{(\ell)} k\left(\boldsymbol{R}_q, \boldsymbol{R}_i\right)$$

P Zaspel, H Harbrecht, B Huang, OAvL *JCTC* (2018)



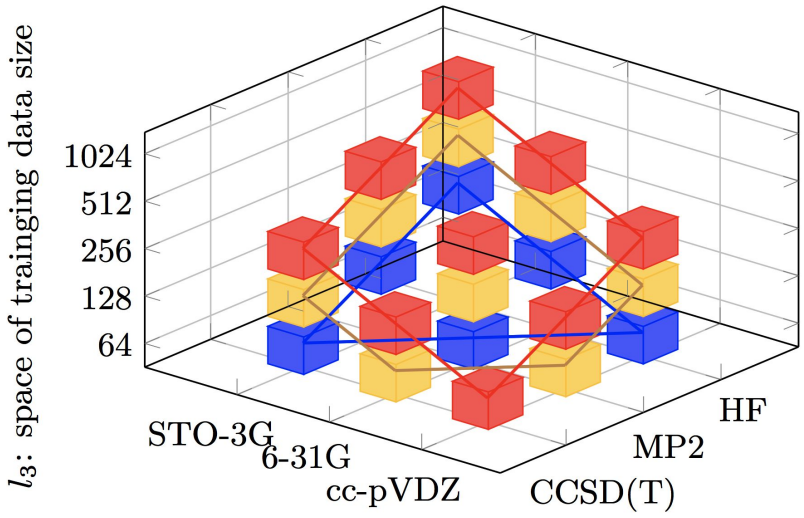
l_2: space of basis sets

l_1: space of theories

Multi-level sparse grid **C**ombination technique (CQML)

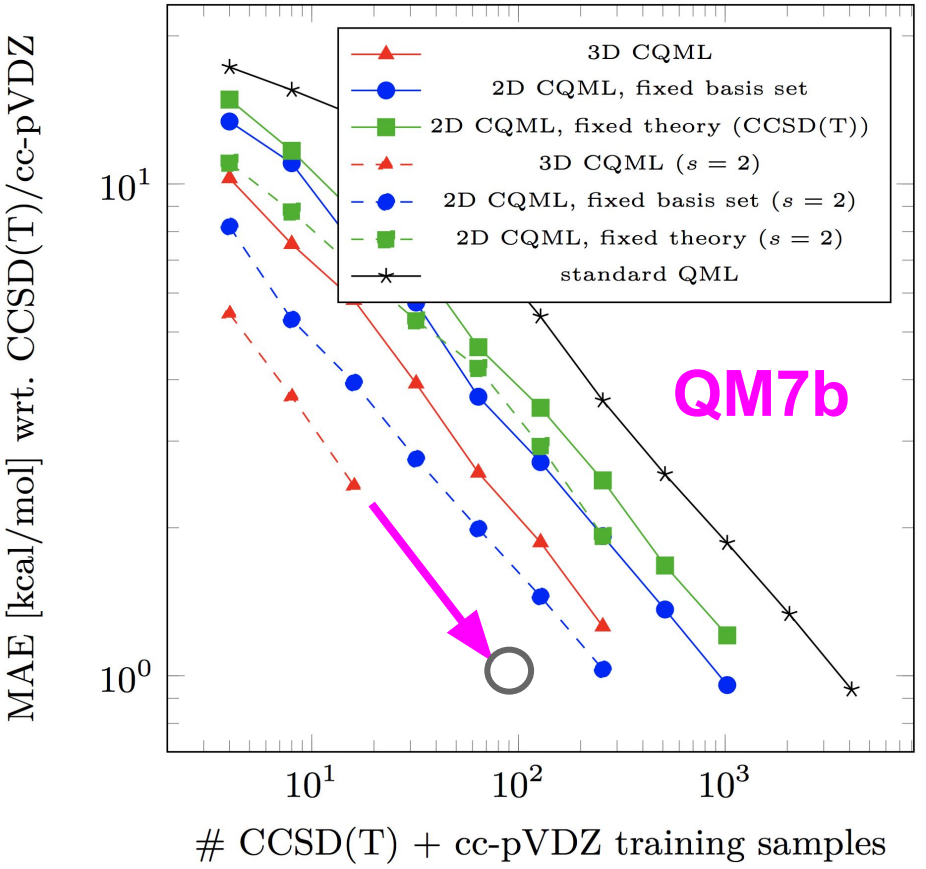
$$\mathcal{M}^{\ell}_{\ell_1}(\boldsymbol{R}_q) := \mathcal{M}^{\ell-1}_{\ell_1}(\boldsymbol{R}_q) + \sum_{i=1}^{N_\ell} \alpha_i^{(\ell)} k\left(\boldsymbol{R}_q, \boldsymbol{R}_i\right)$$

P Zaspel, H Harbrecht, B Huang, OAvL *JCTC* (2019)



l_2: space of basis sets

l_1: space of theories



Multi-level sparse grid **C**ombination technique (CQML)

$$\mathcal{M}^{\ell}_{\ell_1}(\boldsymbol{R}_q) := \mathcal{M}^{\ell-1}_{\ell_1}(\boldsymbol{R}_q) + \sum_{i=1}^{N_\ell} \alpha_i^{(\ell)} k\left(\boldsymbol{R}_q, \boldsymbol{R}_i\right)$$

P Zaspel, H Harbrecht, B Huang, OAvL *JCTC* (2019)

QML

Search docs

GETTING STARTED:

Installing QML

Citing use of QML

QML Tutorial

Examples

SOURCE DOCUMENTATION:

Python API documentation

QML: A Python Toolkit for Quantum Machine Learning

QML is a Python2/3-compatible toolkit for representation learning of properties of molecules and solids. QML is not a high-level framework where you can do `model.train()`, but supplies the building blocks to carry out efficient and accurate machine learning on chemical compounds. As such, the goal is to provide usable and efficient implementations of concepts such as representations and kernels.

Current list of contributors:

- Anders S. Christensen (University of Basel)
- Felix A. Faber (University of Basel)
- Bing Huang (University of Basel)
- Lars A. Bratholm (University of Copenhagen)
- Alexandre Tkatchenko (University of Luxembourg)
- Klaus-Robert Müller (Technische Universität Berlin/Korea University)
- O. Anatole von Lilienfeld (University of Basel)

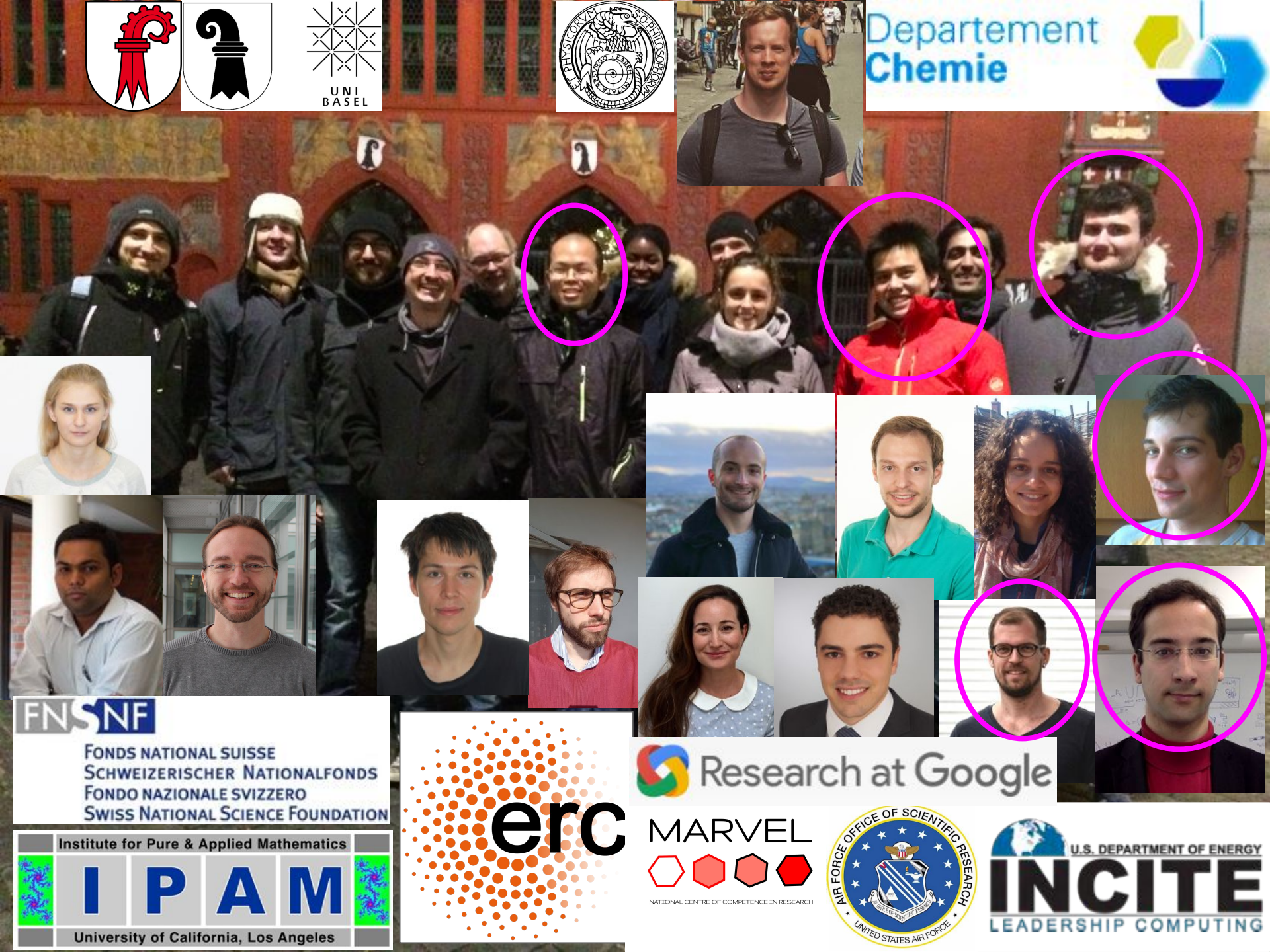
Code development

The QML code is developed through our GitHub repository:

https://github.com/qmlcode/qml

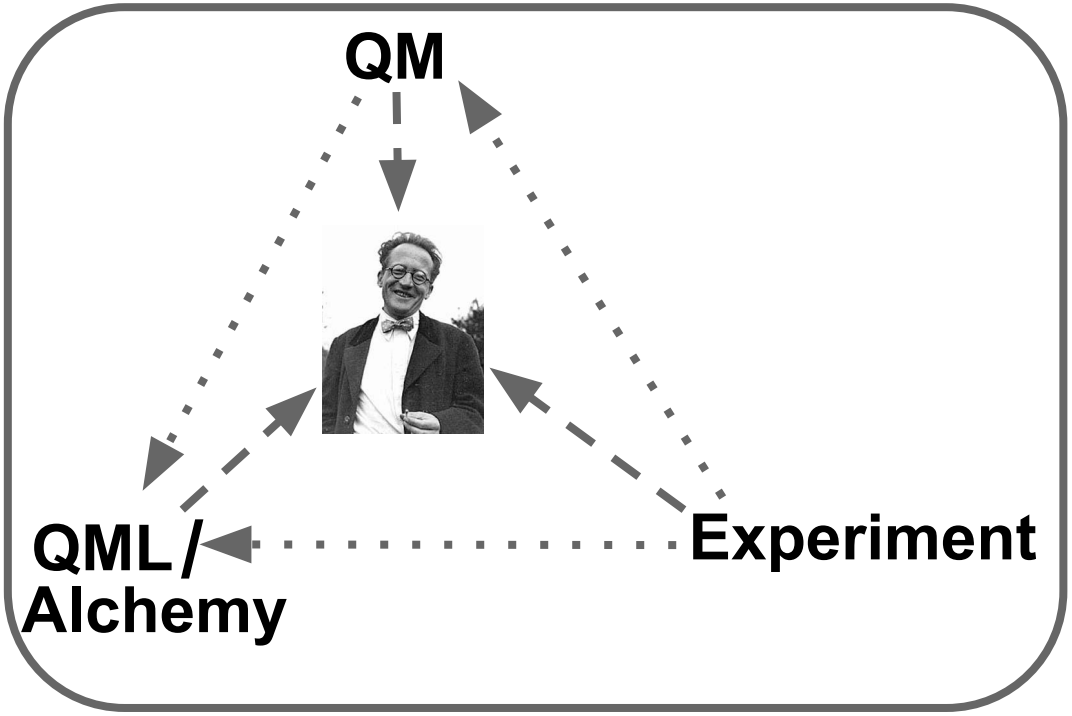
Please add you code to QML by forking and making pull-requests to the
now and then develop branch is pushed to the "master" branch and auto
where the latest stable version is hosted.

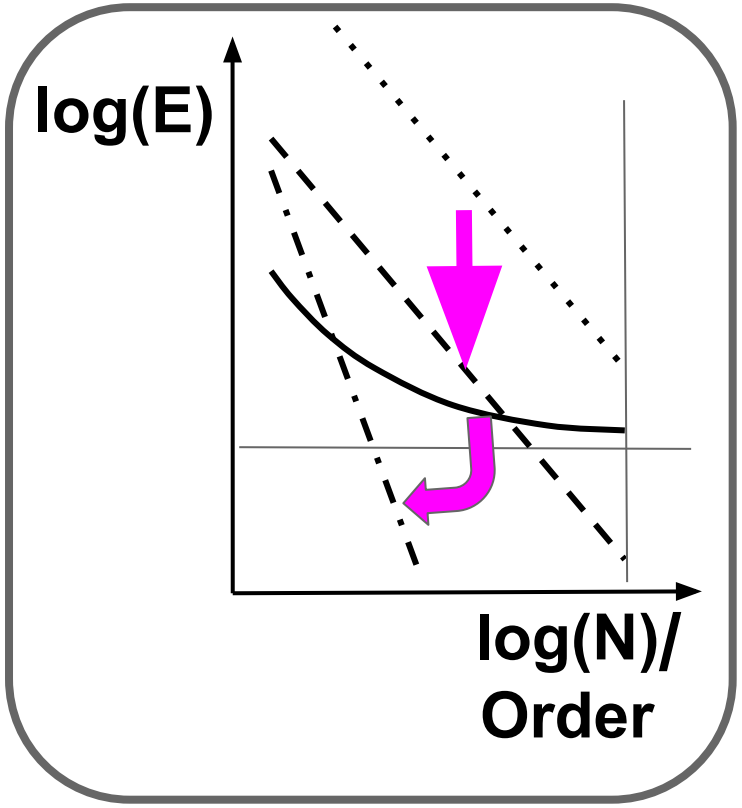
See the "Installing QML" page for up-to-date installation instructions.


UNI
BASEL
Departement
Chemie
FNSNF
FONDS NATIONAL SUISSE
SCHWEIZERISCHER NATIONALFONDS
FONDO NAZIONALE SVIZZERO
SWISS NATIONAL SCIENCE FOUNDATION
Institute for Pure & Applied Mathematics
I P A M
University of California, Los Angeles
erc
Research at Google
MARVEL
NATIONAL CENTRE OF COMPETENCE IN RESEARCH
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
UNITED STATES AIR FORCE
U.S. DEPARTMENT OF ENERGY
INCITE
LEADERSHIP COMPUTING

Summary

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





`First principles view on chemical space', *Int. J. Quantum Chem.* (2013)
`Quantum Machine Learning', *Angew. Chem. Int. Ed.* (2018)