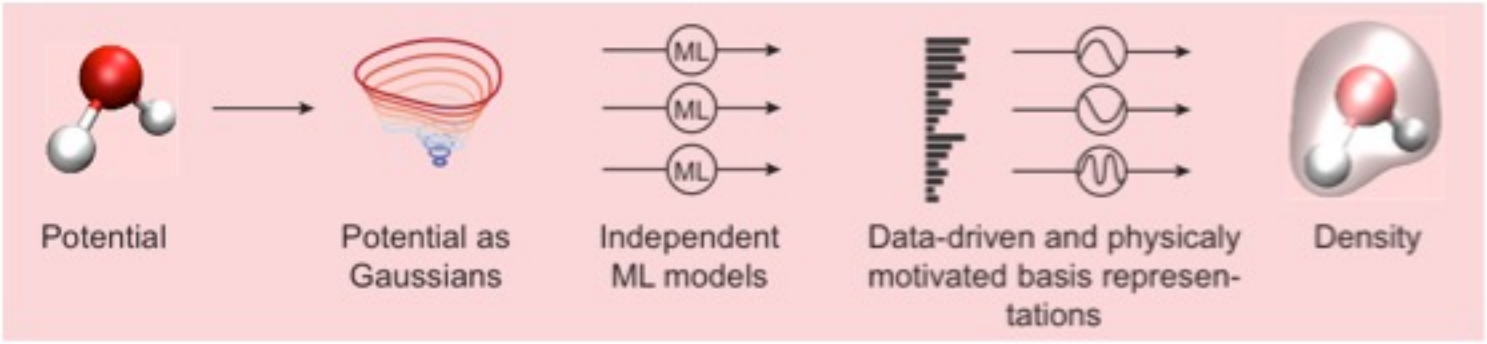
Predicting Energies From Electron Densities



Machine Learning for Reactive Molecular Dynamics

Leslie Vogt-Maranto

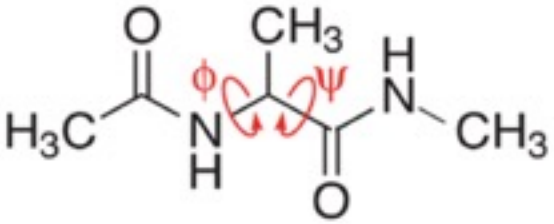
Assistant Research Scientist

Tuckerman Group
Department of Chemistry
New York University

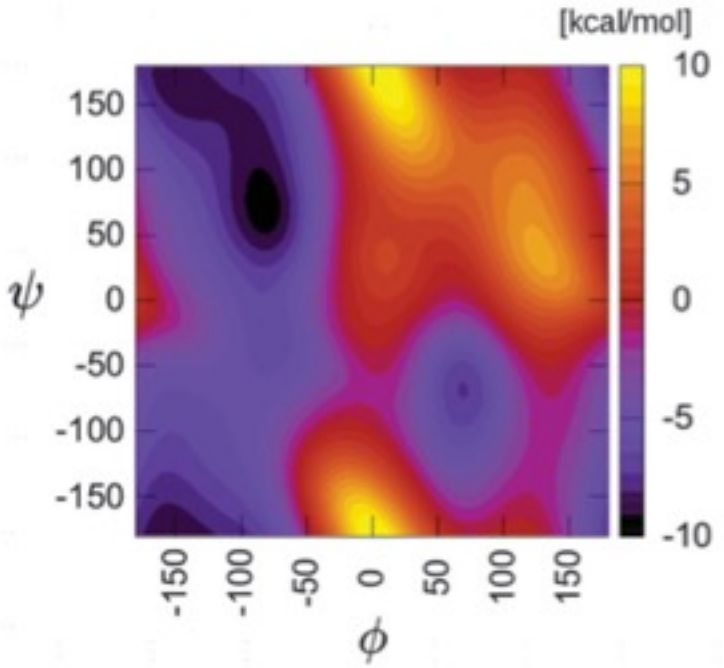
Motivation: Generating Free Energy Surfaces

Molecular systems have many degrees of freedom, but only a subset describe conformational differences

Coarse-grained variables can be used to define the FES:



Alanine dipeptide



$$A(s_1,\ldots,s_n) = -\beta^{-1}\ln P(s_1,\ldots,s_n)$$

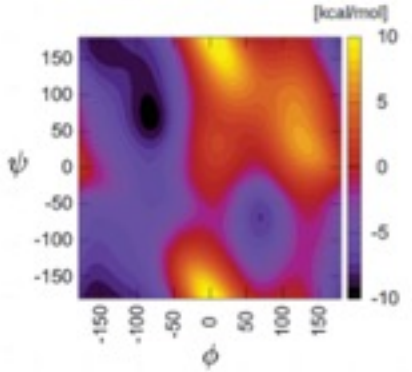
$$P(s) = \mathcal{N}^{-1}\int d^N\mathbf{r}e^{-\beta U(\mathbf{r})}\prod_{\alpha=1}^{n}\delta(q_\alpha(\mathbf{r}) - s_\alpha)$$

Schneider, Dai, Topper, Drechsel-Grau, Tuckerman (**2017**) *Phys. Rev. Lett.* 119, 150601.

Motivation: Generating Free Energy Surfaces

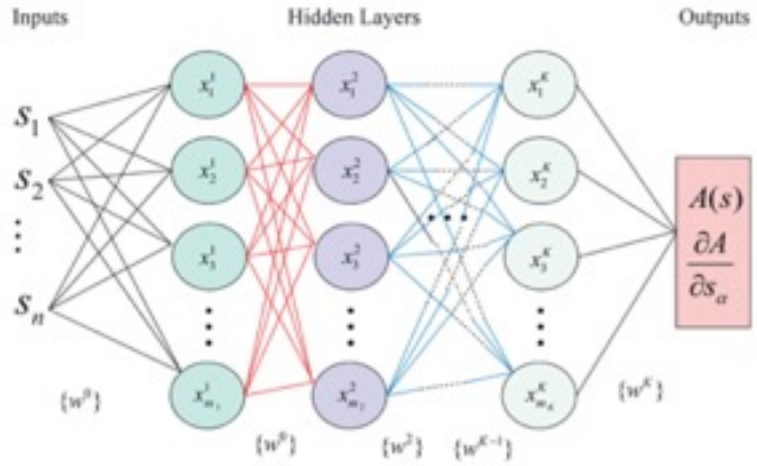
We can sample the FES for different values of *s* using molecular dynamics:

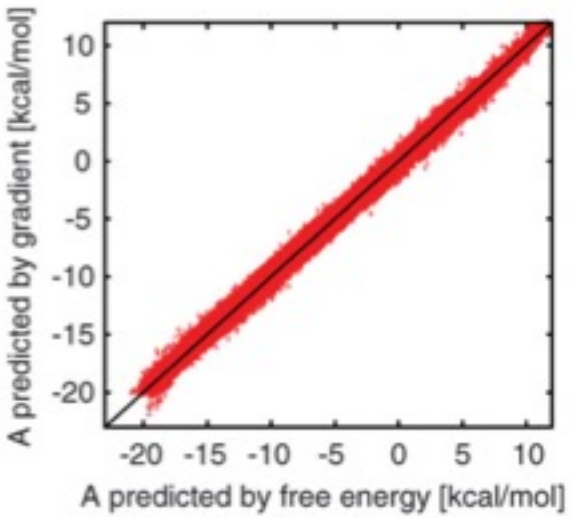
- Restrained simulations for a grid of *s* values
- Adiabatic Free Energy Dynamics (AFED)
- Stochastic activation–relaxation technique (START)



Artificial Neural Networks trained to generate full FES from sampled points

$$A(s_1,\ldots,s_n) = -\beta^{-1}\ln P(s_1,\ldots,s_n)$$





Points sampled using Metropolis MC

Schneider, Dai, Topper, Drechsel-Grau, Tuckerman (**2017**) *Phys. Rev. Lett.* 119, 150601.

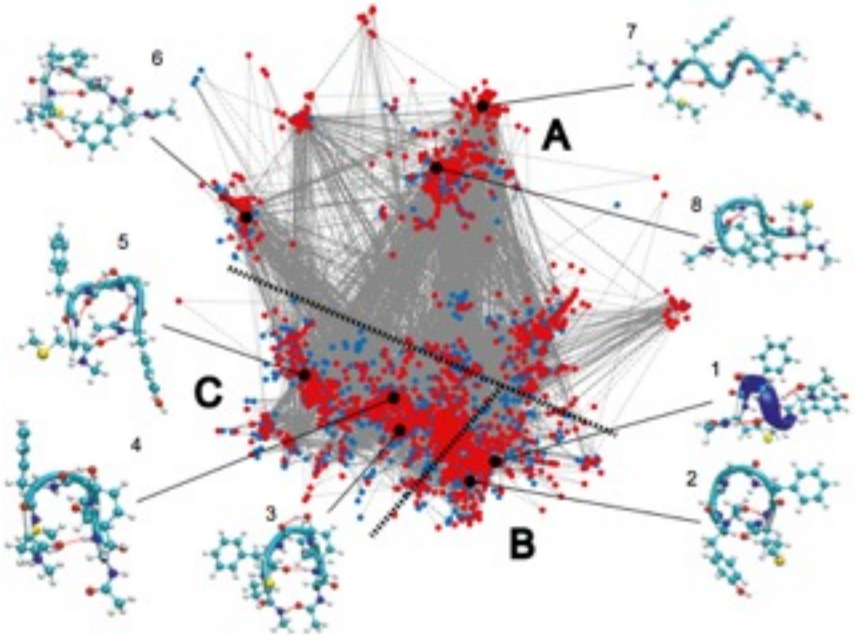
Motivation: Calculating Observables

met-enkephalin: small peptide that can bind to opioid receptors

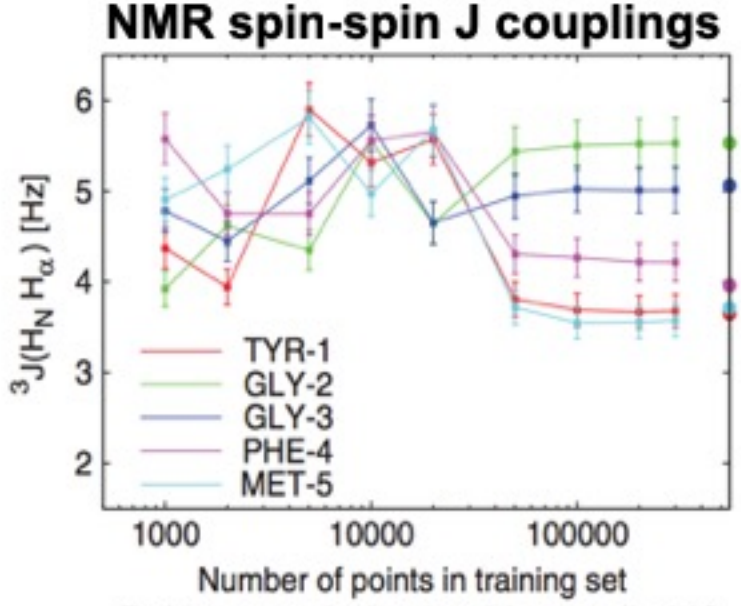
→ 10 coarse-grained variables!

$$\langle O \rangle = \frac{\int d^n s \langle O \rangle_{\mathbf{r}}(s) e^{-\beta A_{\mathrm{ANN}}(s;w_{\mathrm{opt}})}}{\int d^n s e^{-\beta A_{\mathrm{ANN}}(s;w_{\mathrm{opt}})}}$$

$$\langle O \rangle_{\mathbf{r}}(s) = \frac{\mathcal{N}^{-1}}{P(s)} \int d^N \mathbf{r} O(\mathbf{r}) e^{-\beta U(\mathbf{r})} \prod_{\alpha=1}^{n} \delta(q_\alpha(\mathbf{r}) - s_\alpha)$$



Clustered populations from START simulations



Points sampled using Metropolis MC

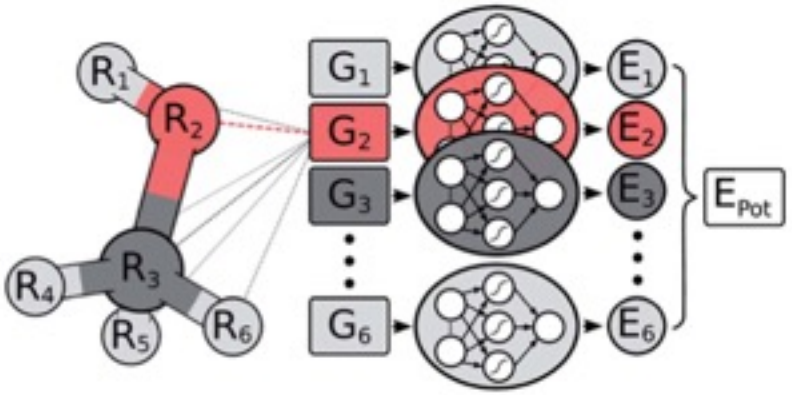
Chen, Yu, Tuckerman (2015) *PNAS* 112, 3235.

Schneider, Dai, Topper, Drechsel-Grau, Tuckerman (**2017**) *Phys. Rev. Lett* 119, 150601.

Machine Learning Molecular Energies

Can we use ML to generate molecular energies? Yes!

Local atom environments can be represented with radial and angular symmetry functions, with a separate NN for each element



Gastegger, Behler, Marquetand (**2017**) *Chem. Sci.* 8, 6924.

Exploding in popularity, with contributions from:

Barros, Behler, Ceriotti, Csanyi, Isayev, Müller, Parkhill, Roitberg, Rupp, von Lilienfeld, …

However, most of these potentials work best for isolated molecules and must be re-trained to get any properties other than energies…

1) Can we make a reactive potential where bonds can be broken?
2) How can we use what we know from physics to improve ML algorithms?

Energies via Electron Densities: DFT

Density functional theory allows us to utilize the one-to-one mapping between a potential and the corresponding electron density:

$$v(\mathbf{r}) \longleftrightarrow n(\mathbf{r})$$

And, the system energy is a functional of the density:

$$E[n] = T_{\mathrm{s}}[n] + U[n] + \int d^3r\, n(\mathbf{r})\, v(\mathbf{r}) + E_{\mathrm{XC}}[n]$$

The plan: learn energies from representations of the electron density

- The atomic environment is reflected in the ground state electron density
- Can be used for any known or trial functional
- Allows us to run Born-Oppenheimer molecular dynamics on the DFT potential energy surface

Data can be expensive to generate, so problem is a good fit for kernel methods

ML-DFT

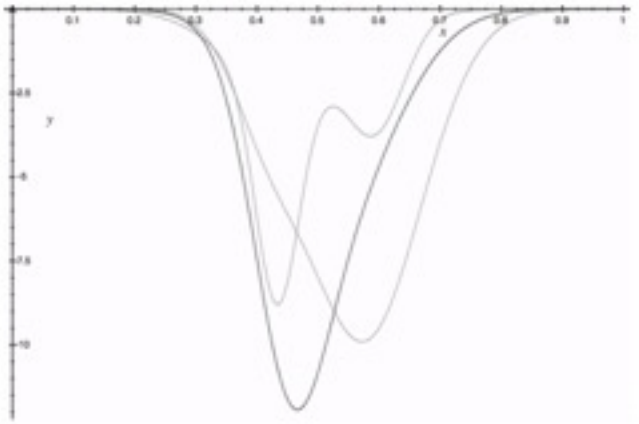
For 1D test case…

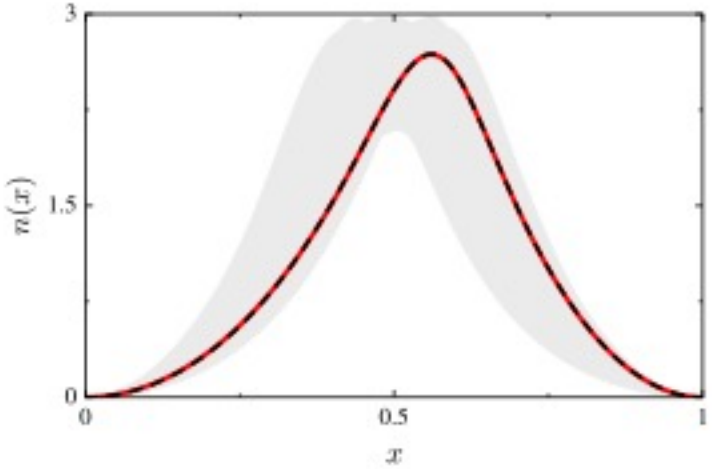
Kernel Ridge Regression is successful in generating the electron density with reasonable error using a Gaussian kernel to quantify similarity in densities

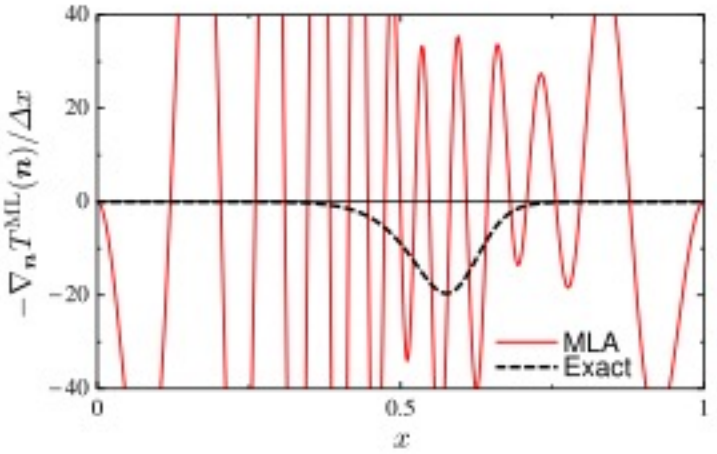
$$k(\boldsymbol{n}, \boldsymbol{n}') = \exp[-\|\boldsymbol{n} - \boldsymbol{n}'\|^2/(2\sigma^2)],$$

$$T^{\mathrm{ML}}(\boldsymbol{n}) = \bar{T} \sum_{j=1}^{M} \alpha_j k(\boldsymbol{n}_j, \boldsymbol{n}),$$

$$\mathcal{C}(\boldsymbol{\alpha}) = \sum_{j=1}^{M} \Delta T_j^2 + \lambda \|\boldsymbol{\alpha}\|^2,$$







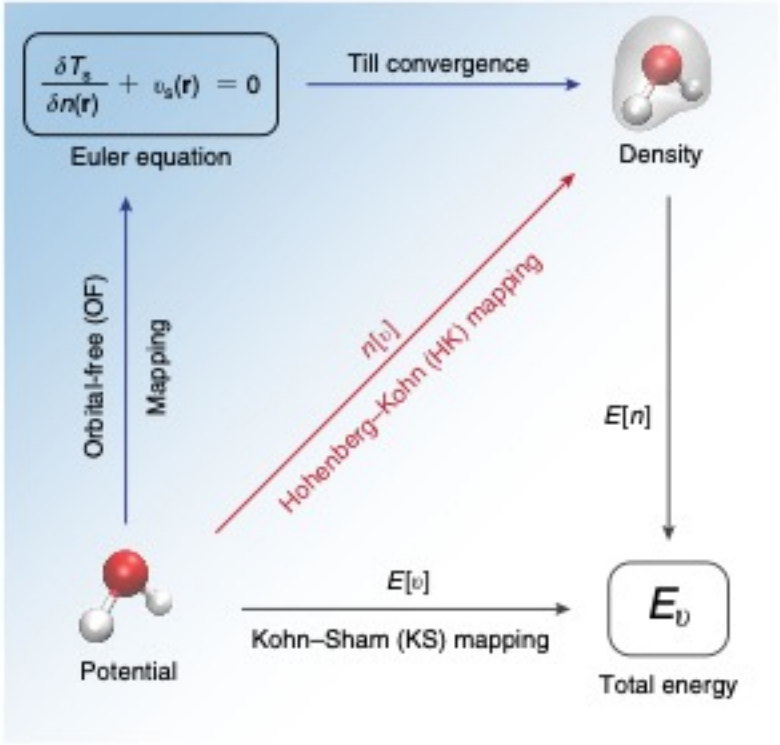
…but the functional derivatives leave room for improvement

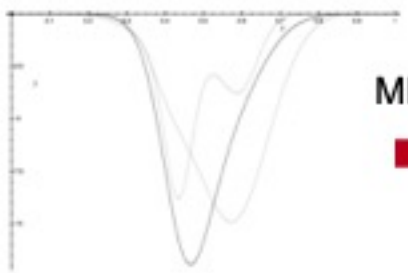
Snyder, Rupp, Hansen, Müller, Burke (**2012**) *Phys. Rev. Lett.* 108, 253002.

Machine learning for DFT

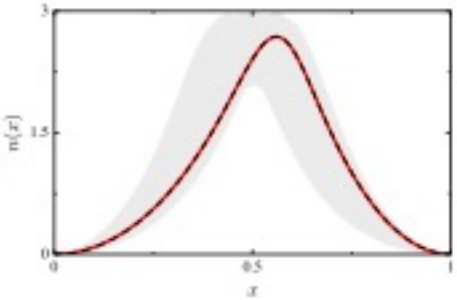
Bypassing the derivative…

We can use a "Hohenberg-Kohn" mapping to learn the converged densities directly from a nuclear potential!

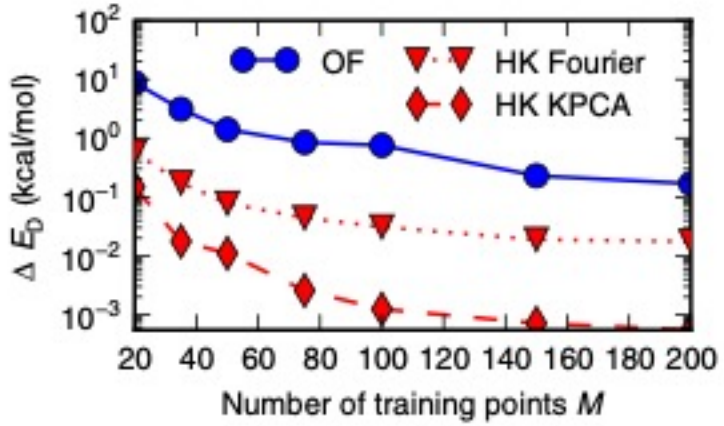




ML-HK

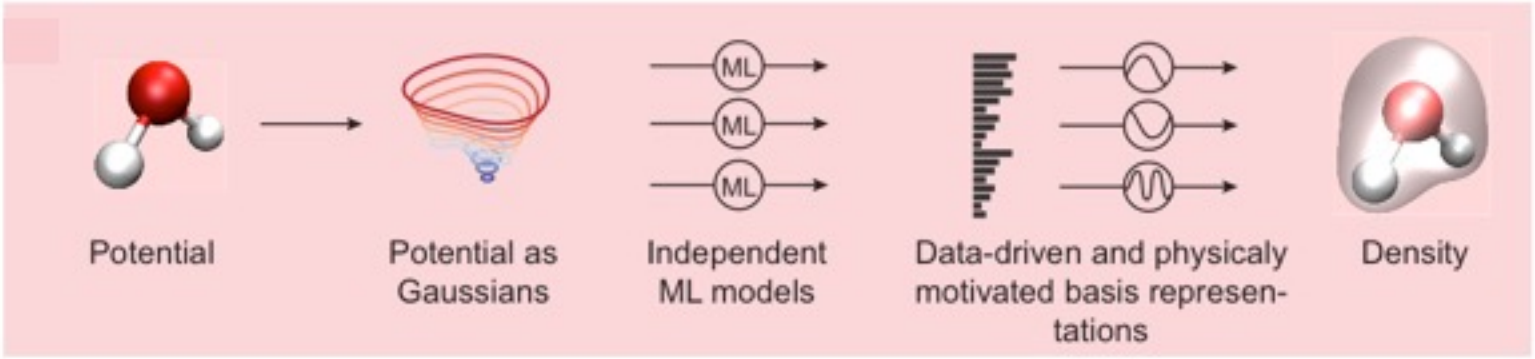


The density-driven error is small for ML models, leading to improved energy prediction performance



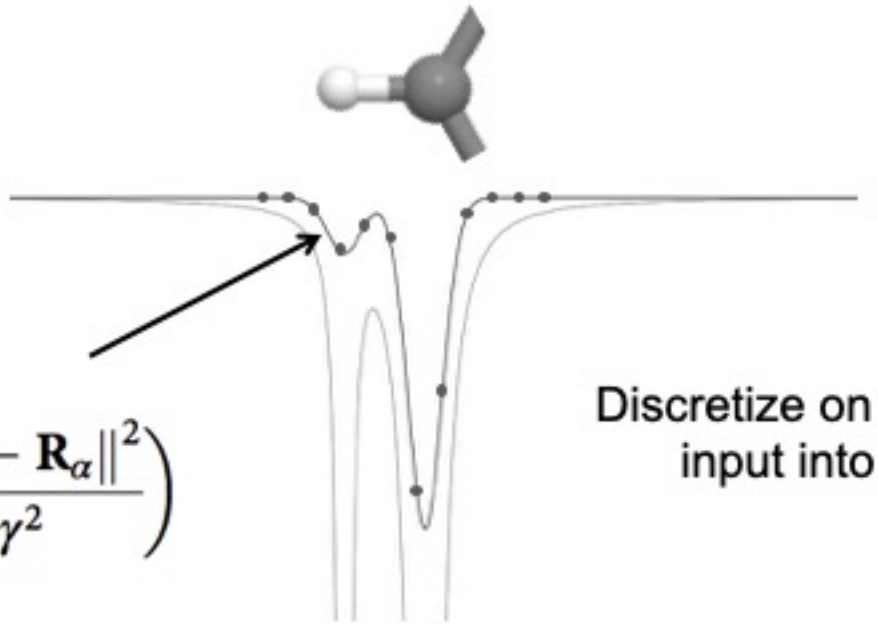
Brockherde, Vogt, Li, Tuckerman, Burke, Müller (**2017**) *Nat. Commun.* 8, 872.

Positions → potentials → densities



Gaussians potential

$$v(\mathbf{r}) = \sum_{\alpha=1}^{N^a} \mathbf{Z}_\alpha \exp\left(\frac{-\|\mathbf{r} - \mathbf{R}_\alpha\|^2}{2\gamma^2}\right)$$

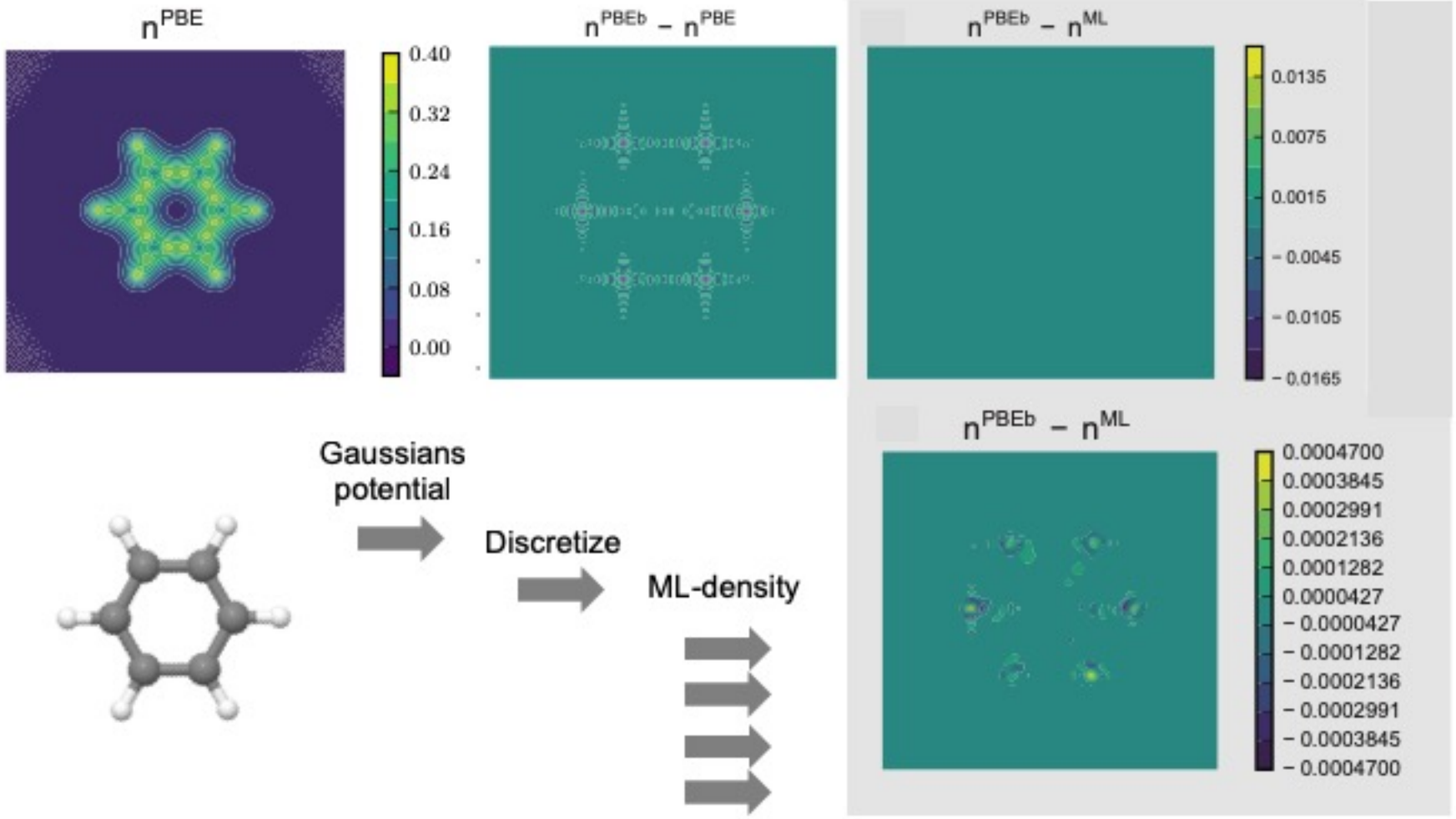


Discretize on a grid for input into KRR

Brockherde, Vogt, Li, Tuckerman, Burke, Müller (**2017**) *Nat. Commun.* 8, 872.

Machine learning electron densities

For compactness, we can represent the electron density in a Fourier basis

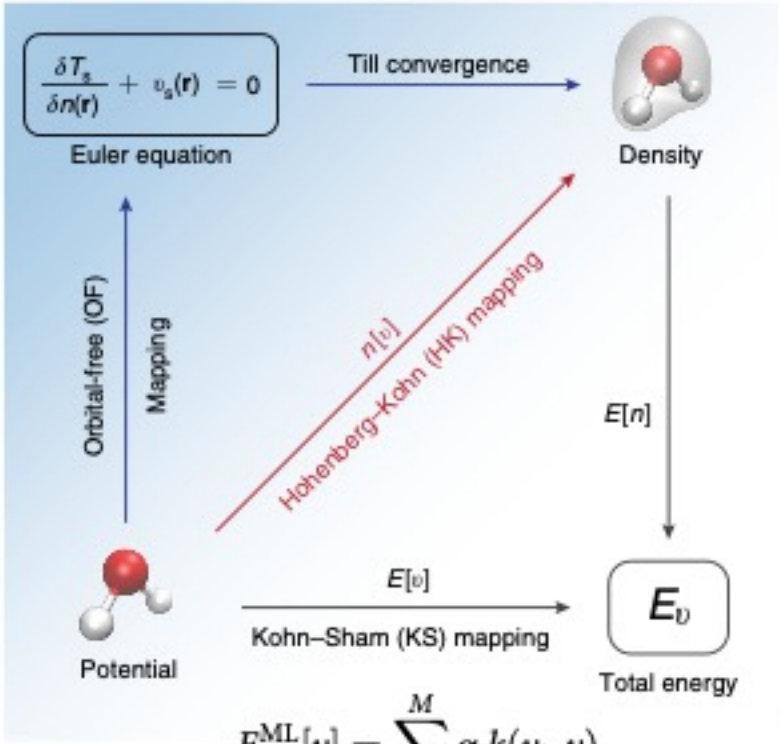


Brockherde, Vogt, Li, Tuckerman, Burke, Müller (**2017**) *Nat. Commun.* 8, 872.

Machine learning for DFT…for molecules!

$$E[n] = T_s[n] + U[n] + \int d^3r\, n(\mathbf{r})\, v(\mathbf{r}) + E_{XC}[n]$$

Several steps are required to translate the molecular geometry into an energy…



$$E^{\mathrm{ML}}[v] = \sum_{i=1}^{M} \alpha_i k(v_i, v)$$

(Just learn energies directly for each geometry)

1) Discretize a Gaussian representation of the potential

$$v(\mathbf{r}) = \sum_{\alpha=1}^{N^a} \mathbf{Z}_\alpha \exp\left(\frac{-\|\mathbf{r} - \mathbf{R}_\alpha\|^2}{2\gamma^2}\right),$$

2) Generate compact representation of density

$$n(x) = \sum_{l=1}^{L} \mathbf{u}^{(l)} \phi_l(x),$$

$$u^{(l)}[v] = \sum_{i=1}^{M} \boldsymbol{\beta}_i^{(l)} k(v, v_i),$$

3) Learn densities as a functional of v(x)

$$n^{\mathrm{ML}}[v](x) = \sum_{l=1}^{L} u^{(l)}[v] \phi_l(x),$$

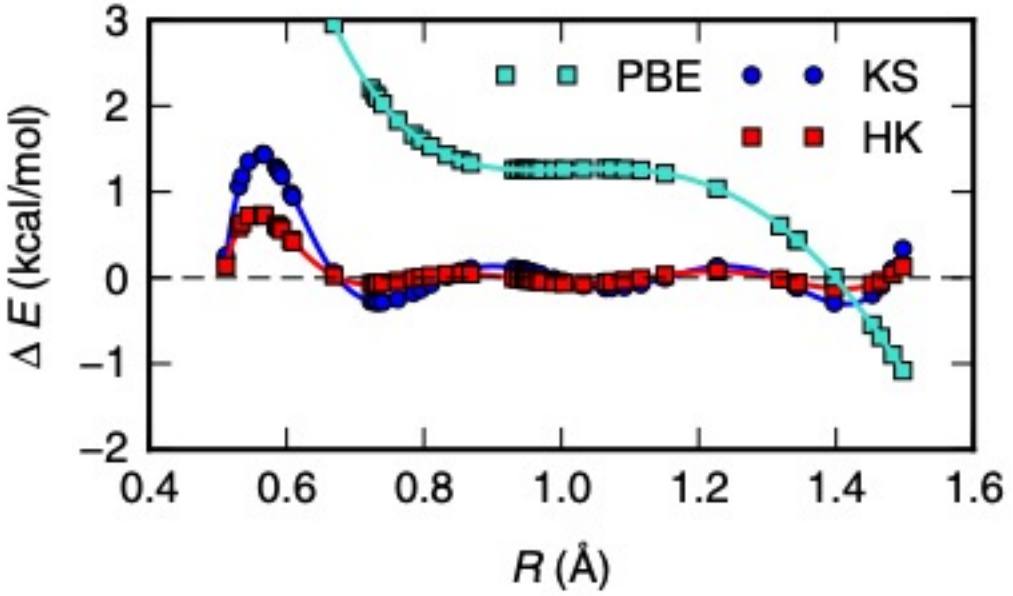
4) Learn energies as a functional of n(x)

$$E^{\mathrm{ML}}[n] = \sum_{i=1}^{M} \alpha_i k(n_i, n),$$

Brockherde, Vogt, Li, Tuckerman, Burke, Müller (**2017**) *Nat. Commun.* 8, 872.

Machine learning for DFT…for H_2

150 training set structures, with 50 test set geometries



ML-KS

$$E^{\mathrm{ML}}[v] = \sum_{i=1}^{M} \alpha_i k(v_i, v)$$

ML-HK

$$E^{\mathrm{ML}}[n] = \sum_{i=1}^{M} \alpha_i k(n_i, n)$$

Lower energy error using densities as an intermediate

M=5 training points

ML-KS error is 1.3 kcal/mol

ML-HK error is **0.7** kcal/mol

Brockherde, Vogt, Li, Tuckerman, Burke, Müller (**2017**) *Nat. Commun.* 8, 872.

Machine learning for DFT: H_2O

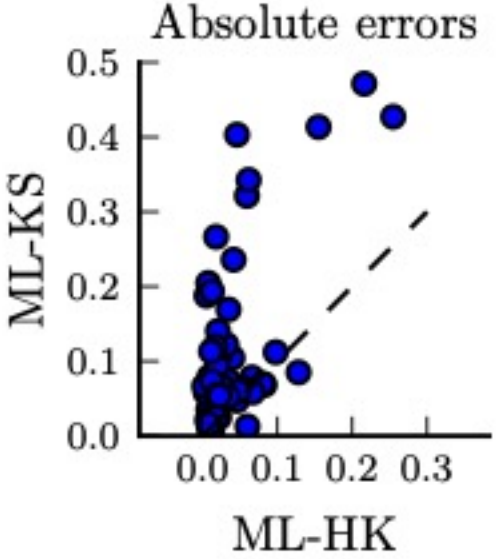
15 of 300 training set structures, with 50 test set geometries

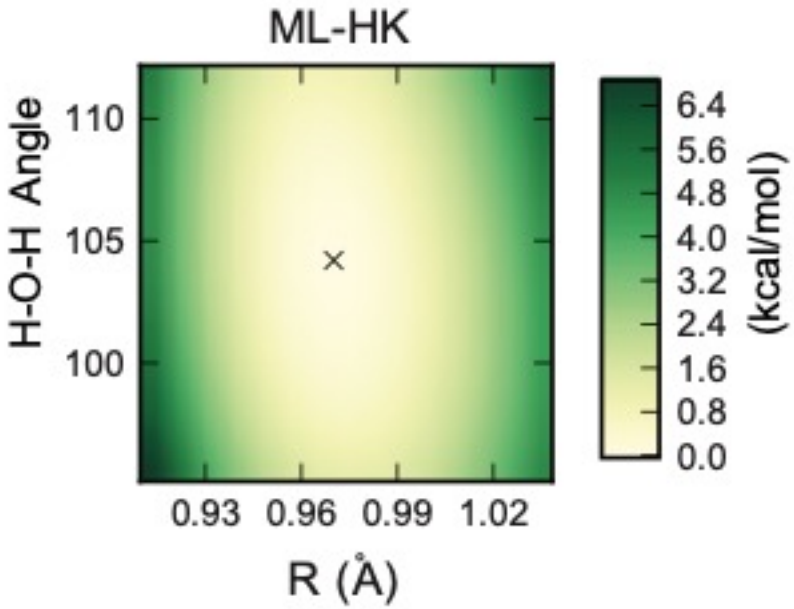
ML-KS

$$E^{\mathrm{ML}}[v] = \sum_{i=1}^{M} \alpha_i k(v_i, v)$$

ML-HK

$$E^{\mathrm{ML}}[n] = \sum_{i=1}^{M} \alpha_i k(n_i, n)$$





For M=15 training points:

ML-KS error is 0.12 kcal/mol (max 0.47 kcal/mol)

ML-HK error is **0.04** kcal/mol (max 0.25 kcal/mol)

The ML-HK energy surface for symmetric conformers around the equilibrium position for PBE (shown by X).

Brockherde, Vogt, Li, Tuckerman, Burke, Müller (**2017**) *Nat. Commun.* 8, 872.

Sampling strategy for training geometries

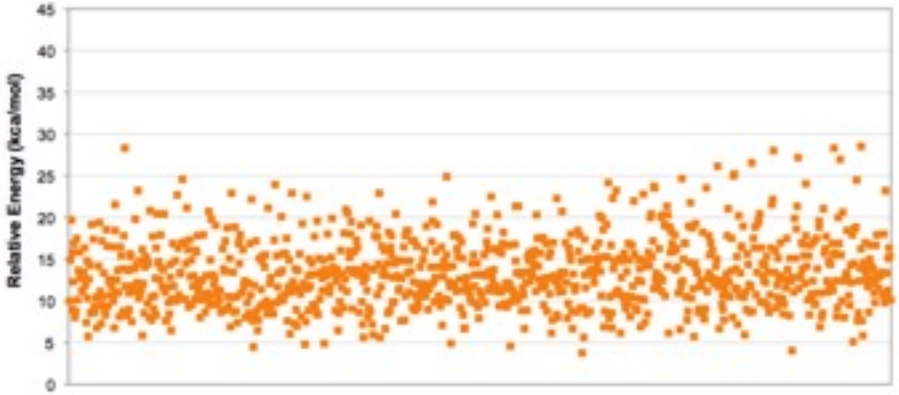
Random Perturbations

Difficult to scale with degrees of freedom

From MD trajectory

Ab initio MD @ 500K

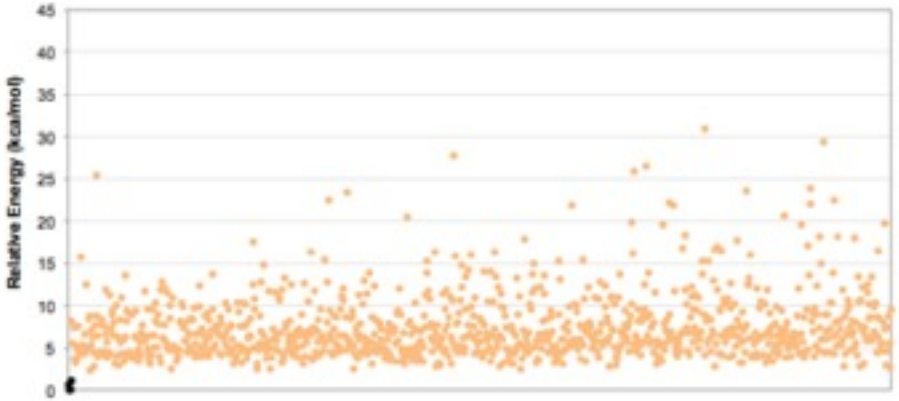
i.e. sGDML [Chmiela *et al.* (2018) Nat. Commun]



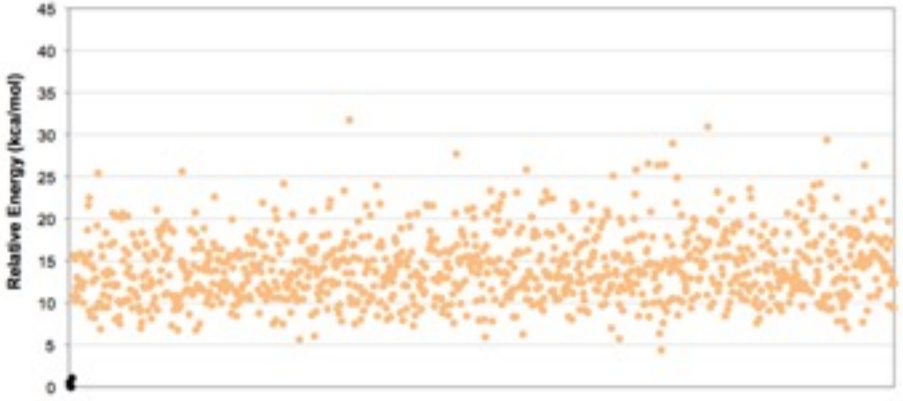
K-means selection to find centers

Classical MD (cheaper!) to sample geometries

K-means centers

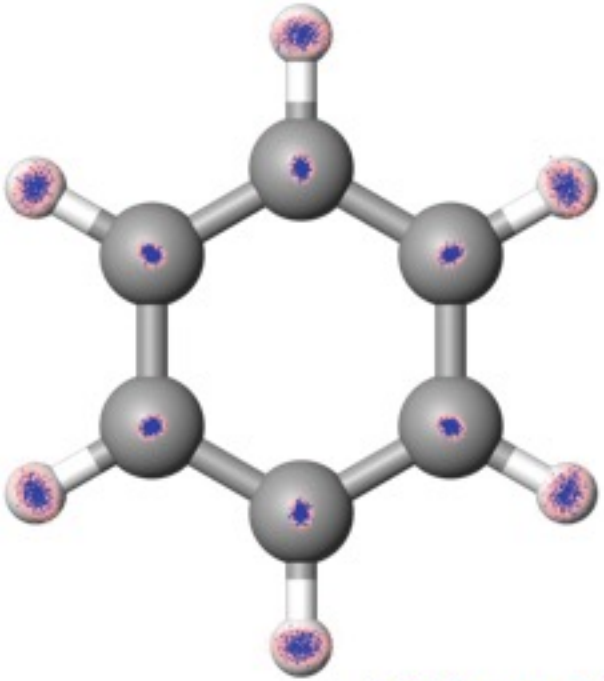


Closest snapshot to K-means centers



PINY_MD: GAFF @ 350 K NVT; CP2K: PBE/DZVP-MOLOPT (900 Ry)

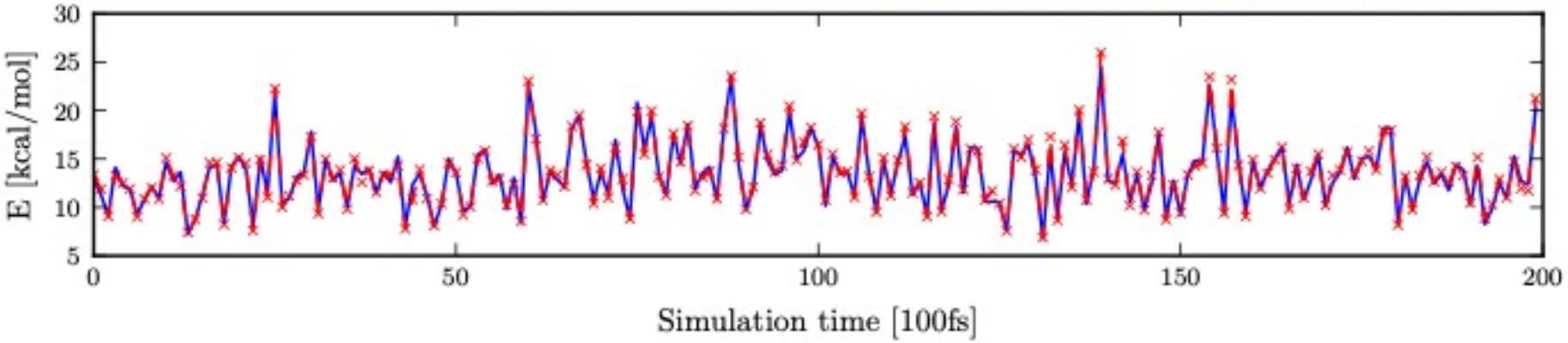
Machine learning for DFT: benzene



Snapshots from classical MD:
Training set (k-means centers @ 350 K)
Test set (random @ 300 K)

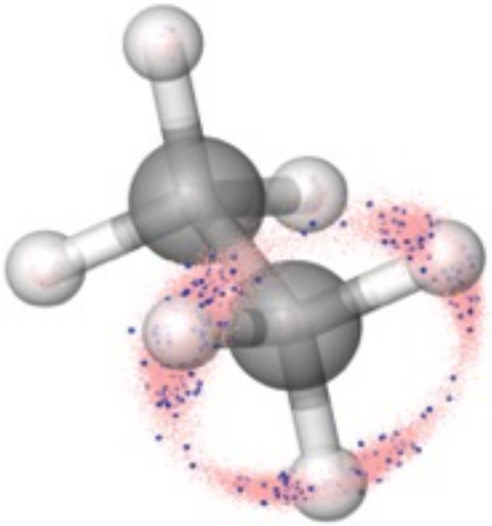
For snapshots from an independent MD trajectory at 300 K, the ML-HK maps have a MAE of only **0.37** kcal/mol on energies that span more than 20 kcal/mol.

DFT energies for snapshots **ML-HK energies for snapshots**



Brockherde, Vogt, Li, Tuckerman, Burke, Müller (**2017**) *Nat. Commun.* 8, 872.

Machine learning for DFT: ethane

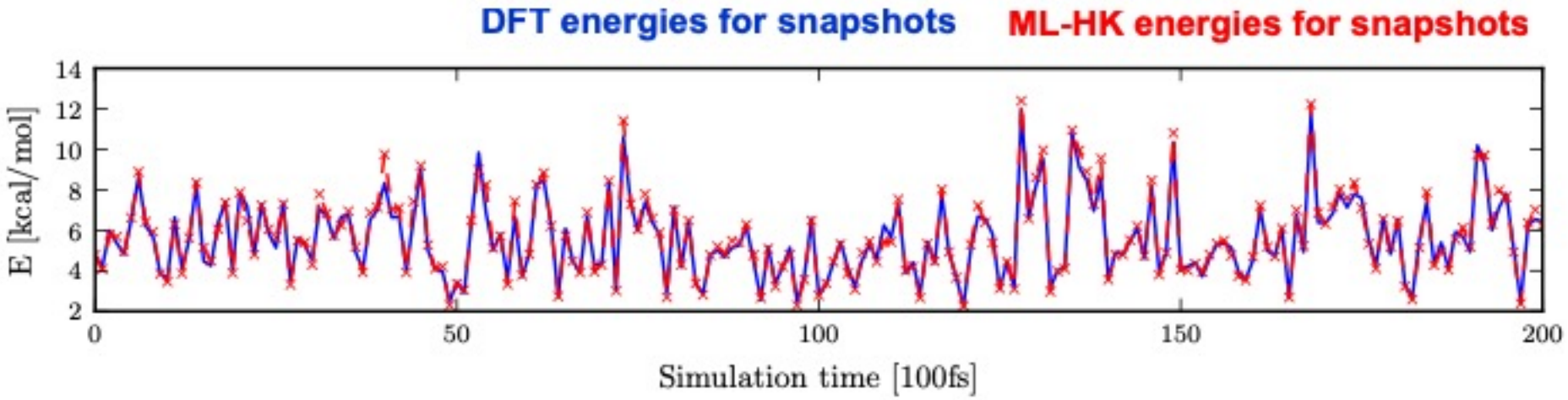


Snapshots from classical MD:

Training set (k-means centers @ 400 K)

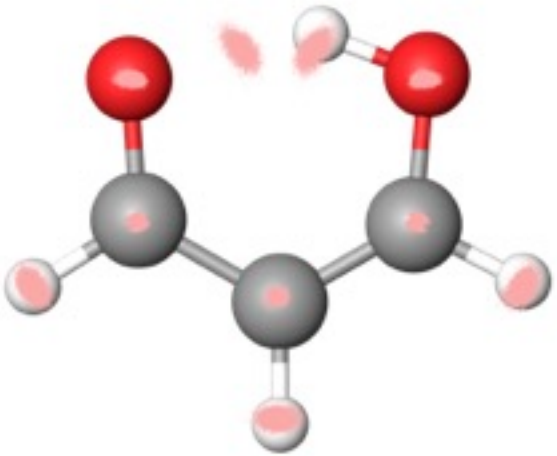
Test set (random @ 300 K)

For snapshots from an independent MD trajectory at 300 K, the ML-HK maps have a MAE of only **0.23** kcal/mol on energies that span more than 10 kcal/mol.



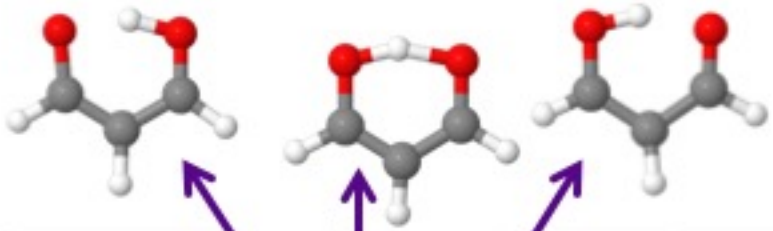
Brockherde, Vogt, Li, Tuckerman, Burke, Müller (2017) *Nat. Commun.* 8, 872.

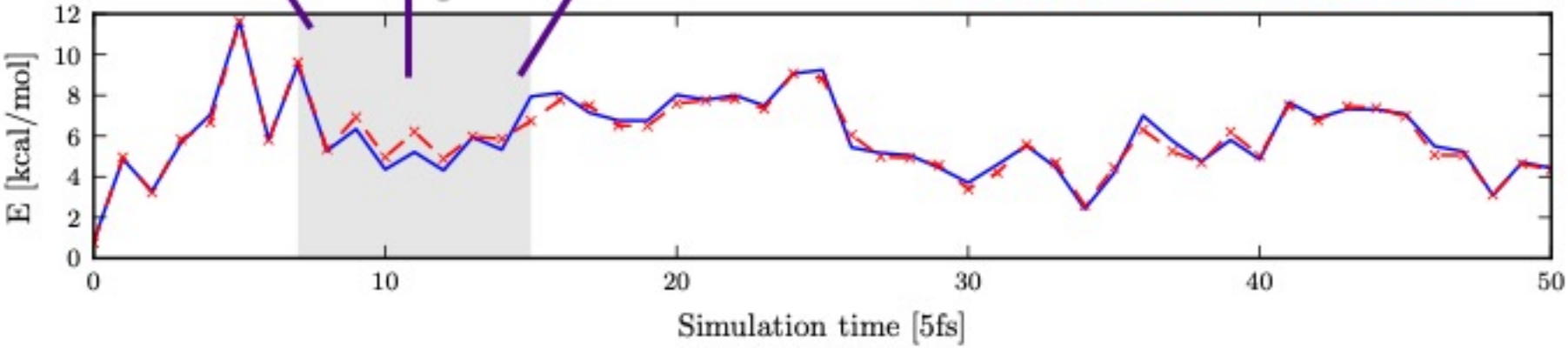
Machine learning for DFT: malonaldehyde



The ML-HK model can predict energies during an intramolecular proton transfer in the enol form of malonaldehyde, after training on classical trajectories with fixed tautomers!

Energy errors for snapshots from an *ab inito* trajectory are **0.27** kcal/mol.

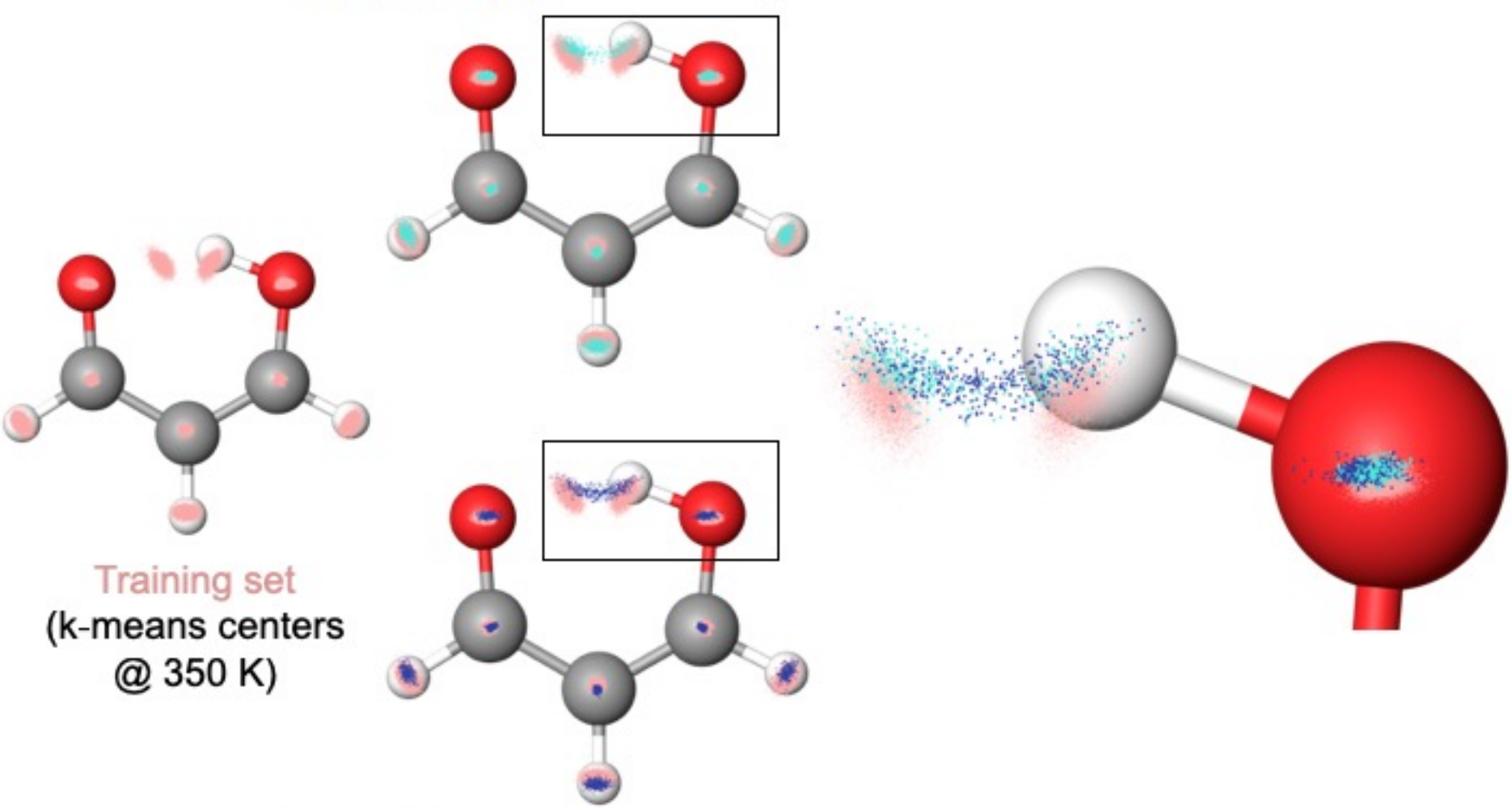




Brockherde, Vogt, Li, Tuckerman, Burke, Muller (**2017**) *Nat. Commun.* 8, 872.

Overlap of test and training data

External Check : DFT MD @ 300 K

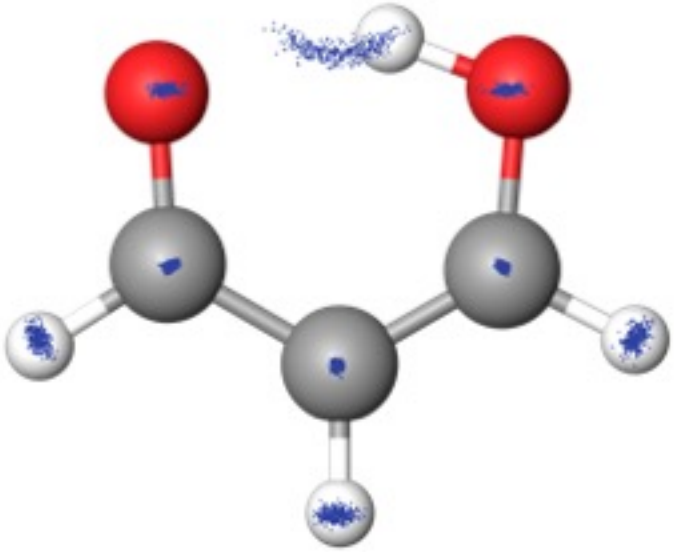


Training set
(k-means centers
@ 350 K)

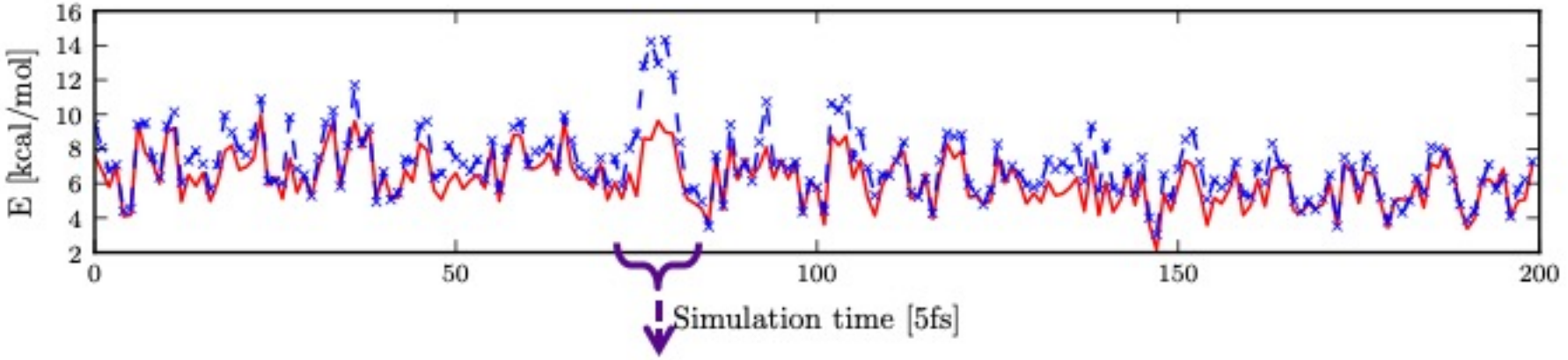
Internal Check : ML-HK MD @ 300 K

Brockherde, Vogt, Li, Tuckerman, Burke, Muller (**2017**) *Nat. Commun.* 8, 872.

Machine learning for DFT: malonaldehyde



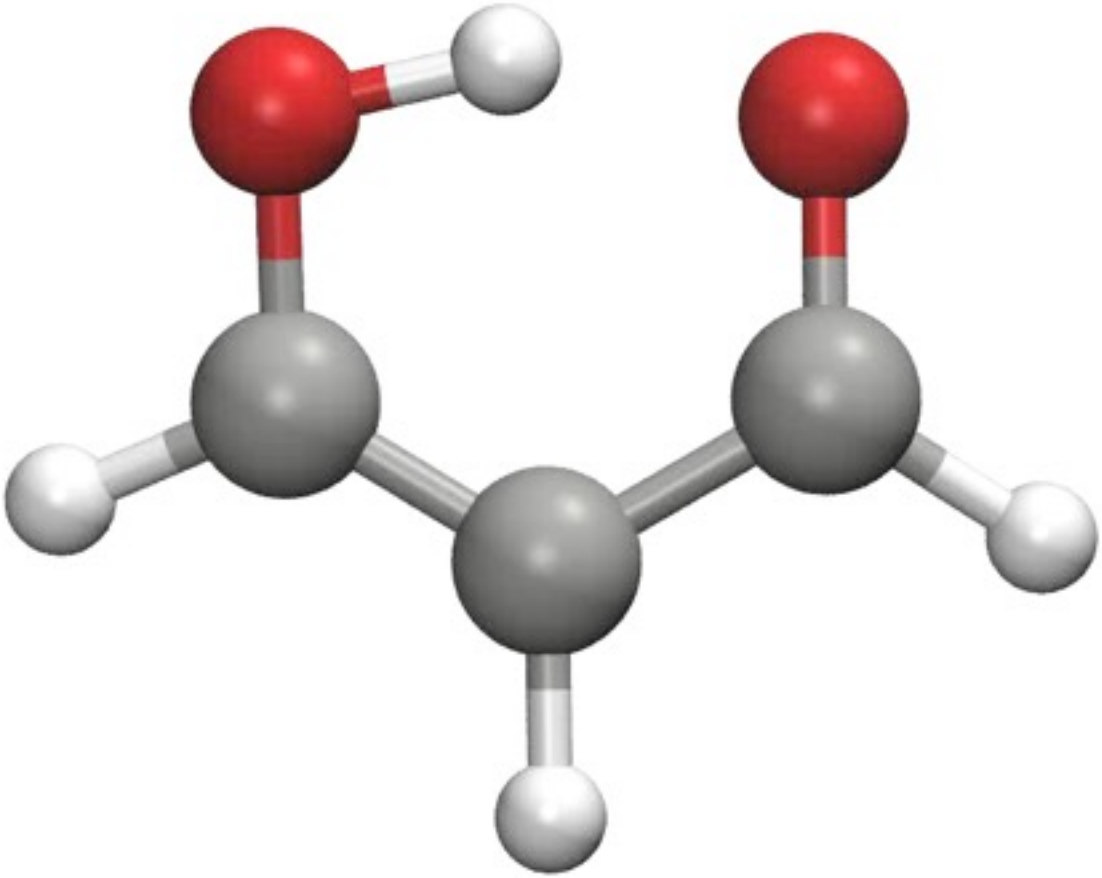
For a 300 K MD trajectory ***generated using the ML-HK model***, the MAE relative to PBE is **0.77** kcal/mol.



(Some configurations need to be added to the training set!)

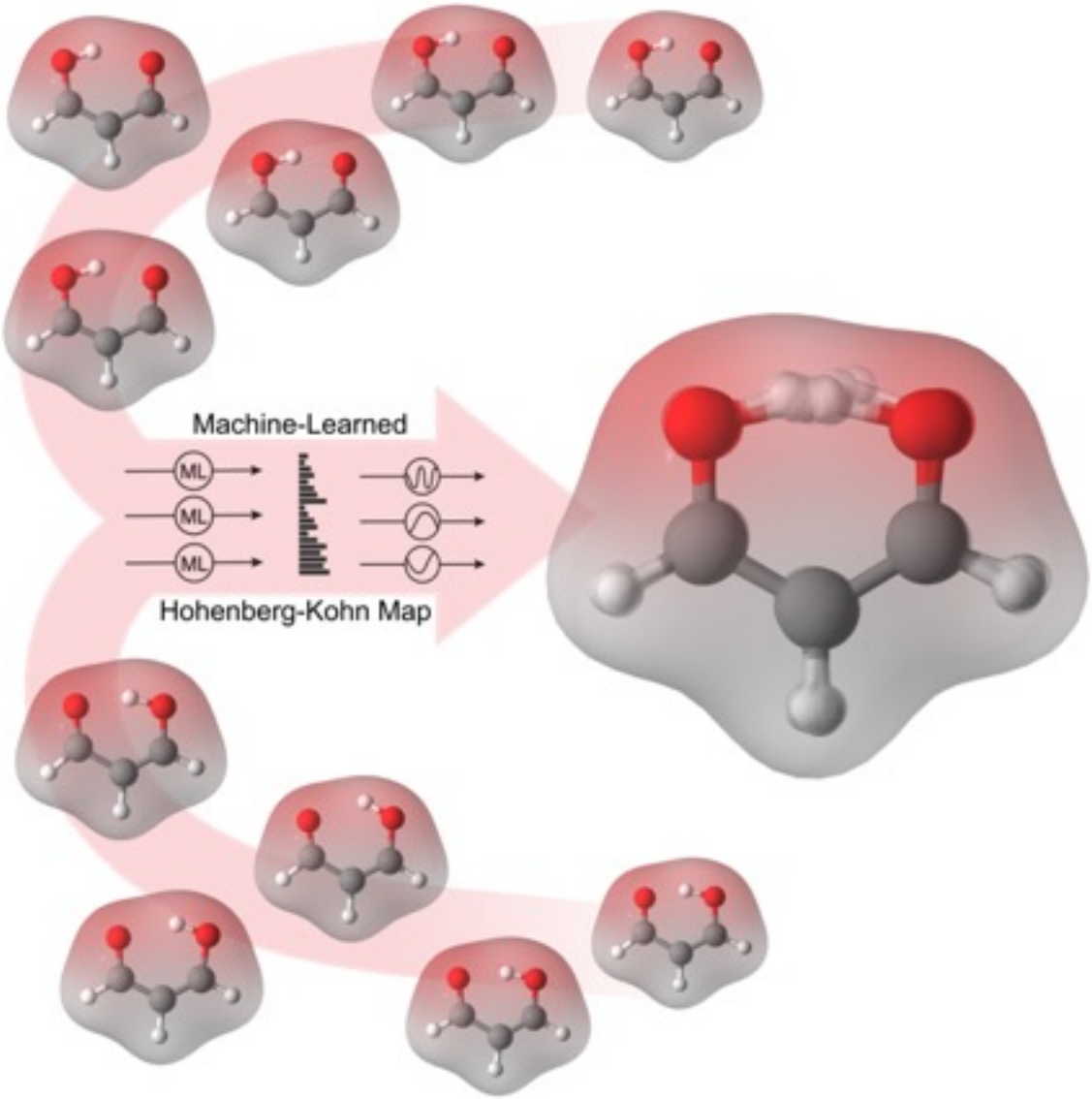
Brockherde, Vogt, Li, Tuckerman, Burke, Muller (**2017**) *Nat. Commun.* 8, 872.

Machine learning for DFT: malonaldehyde



Brockherde, Vogt, Li, Tuckerman, Burke, Müller (**2017**) *Nat. Commun.* 8, 872.

Machine learning electron densities



Brockherde, Vogt, Li, Tuckerman, Burke, Muller (**2017**) *Nat. Commun.* 8, 872.

Future directions (happening now!)

Active learning

Adding new conformers based on ML model sampling

Transfer learning

Related families of molecules, energy schemes, bonding environments

Periodic solids

Atomic and molecular systems

Improved energy descriptions

Learning alternate DFT and CCSD(T) energies

Sampling FES for observable properties

Molecular dynamics including path integrals

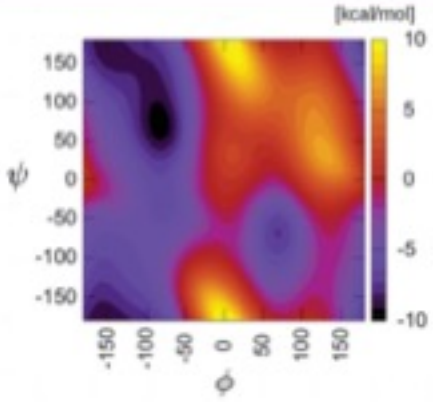
Exciting new applications of ML to dynamic and reactive molecular systems!

Acknowledgements

New York University

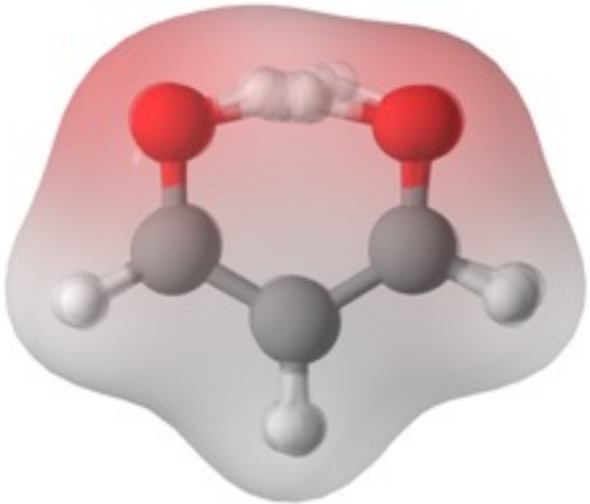
Mark Tuckerman

- Elia Schneider
- Joey Cendagorta
- High School Students (LD, NP)



The Cooper Union

Robert Topper



University of California, Irvine

Kieron Burke

- Li Li

Technical University Berlin

Klaus-Robert Müller

- Felix Brockherde
- Mihail Bogojeski

MRSEC

NYU High Performance Computing