

Electron Density Prediction

**Bypassing the Kohn-Sham Equations
with Machine Learning**

Niklas Menzel
Coffee Talk, June 18, 2018

Literature Talk



ARTICLE

DOI: [10.1038/s41467-017-00839-3](https://doi.org/10.1038/s41467-017-00839-3)

OPEN

Bypassing the Kohn-Sham equations with machine learning

Felix Brockherde^{1,2}, Leslie Vogt³, Li Li⁴, Mark E. Tuckerman^{3,5,6}, Kieron Burke^{4,7} & Klaus-Robert Müller^{1,8,9}

Last year, at least 30,000 scientific papers used the Kohn-Sham scheme of density functional theory to solve electronic structure problems in a wide variety of scientific fields. Machine learning holds the promise of learning the energy functional via examples, bypassing the need to solve the Kohn-Sham equations. This should yield substantial savings in computer time, allowing larger systems and/or longer time-scales to be tackled, but attempts to machine-learn this functional have been limited by the need to find its derivative. The present work overcomes this difficulty by directly learning the density-potential and energy-density maps for test systems and various molecules. We perform the first molecular dynamics simulation with a machine-learned density functional on malonaldehyde and are able to capture the intramolecular proton transfer process. Learning density models now allows the construction of accurate density functionals for realistic molecular systems.

Introduction and Motivation

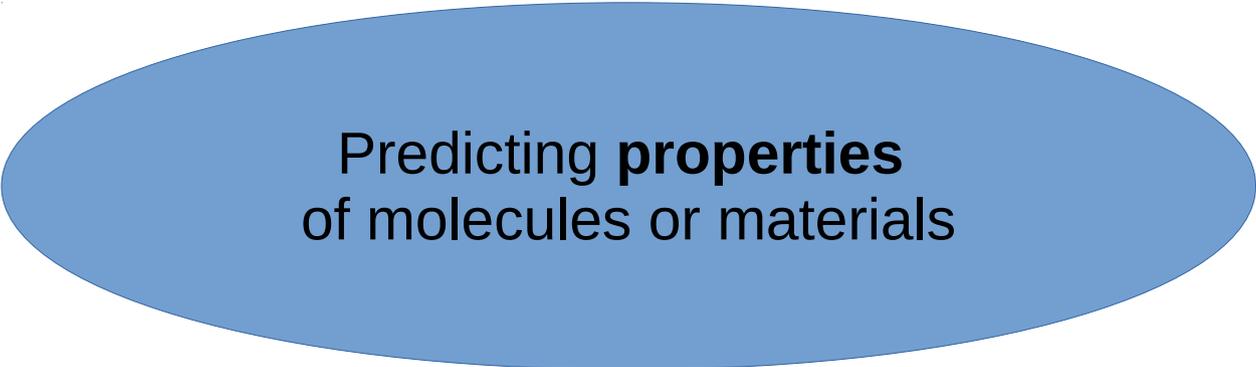
- Kohn-Sham density functional theory (KS-DFT) one of the most popular electronic structure methods: *useful accuracy* at a *moderate cost*

Introduction and Motivation

- Kohn-Sham density functional theory (KS-DFT) one of the most popular electronic structure methods: *useful accuracy* at a *moderate cost*
- Limitations (among others):
System size and simulations involving a large number of KS-DFT calculations, e.g. (ab initio) molecular dynamics (MD) simulations

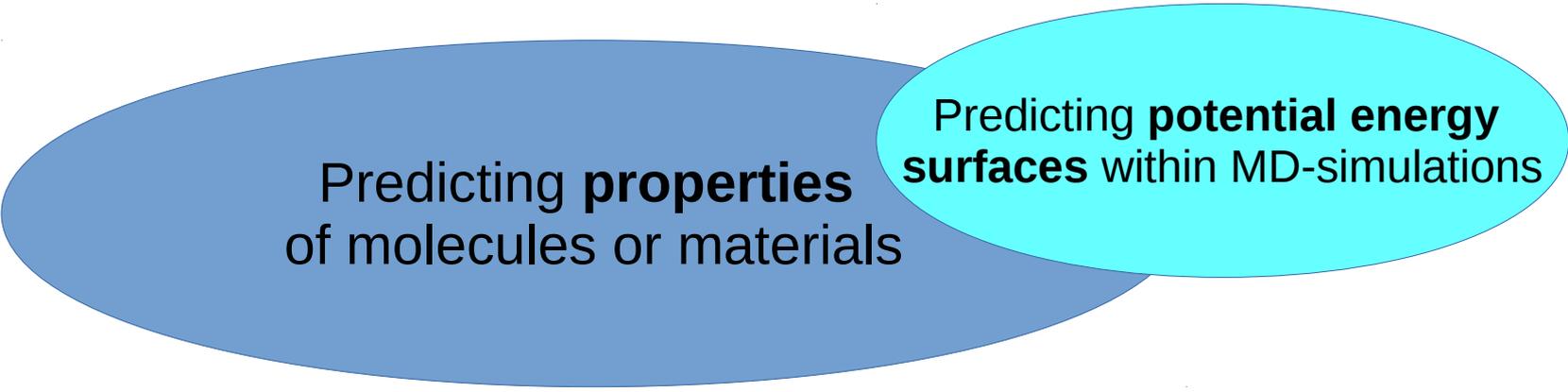
- Using machine learning (ML) in physical sciences holds the promise to **learn mappings** (connections) using data
 - (much) **cheaper** than traditional, model-based methods

- Using machine learning (ML) in physical sciences holds the promise to **learn mappings** (connections) using data
 - (much) **cheaper** than traditional, model-based methods
- Applications:



Predicting **properties**
of molecules or materials

- Using machine learning (ML) in physical sciences holds the promise to **learn mappings** (connections) using data
 - (much) **cheaper** than traditional, model-based methods
- Applications:



Predicting **properties**
of molecules or materials

Predicting **potential energy surfaces** within MD-simulations

- Using machine learning (ML) in physical sciences holds the promise to **learn mappings** (connections) using data
 - (much) **cheaper** than traditional, model-based methods
- Applications:

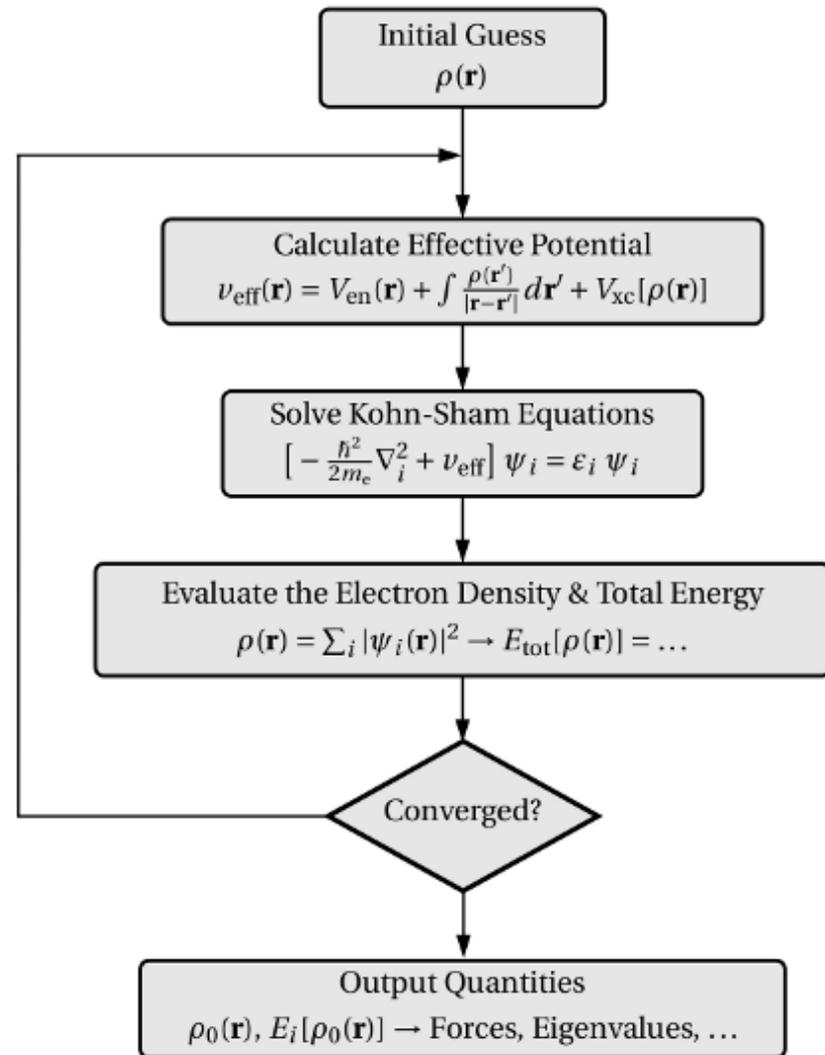
Predicting **properties**
of molecules or materials

Predicting **potential energy surfaces** within MD-simulations

Finding **functionals**
of DFT

Recap: KS-DFT

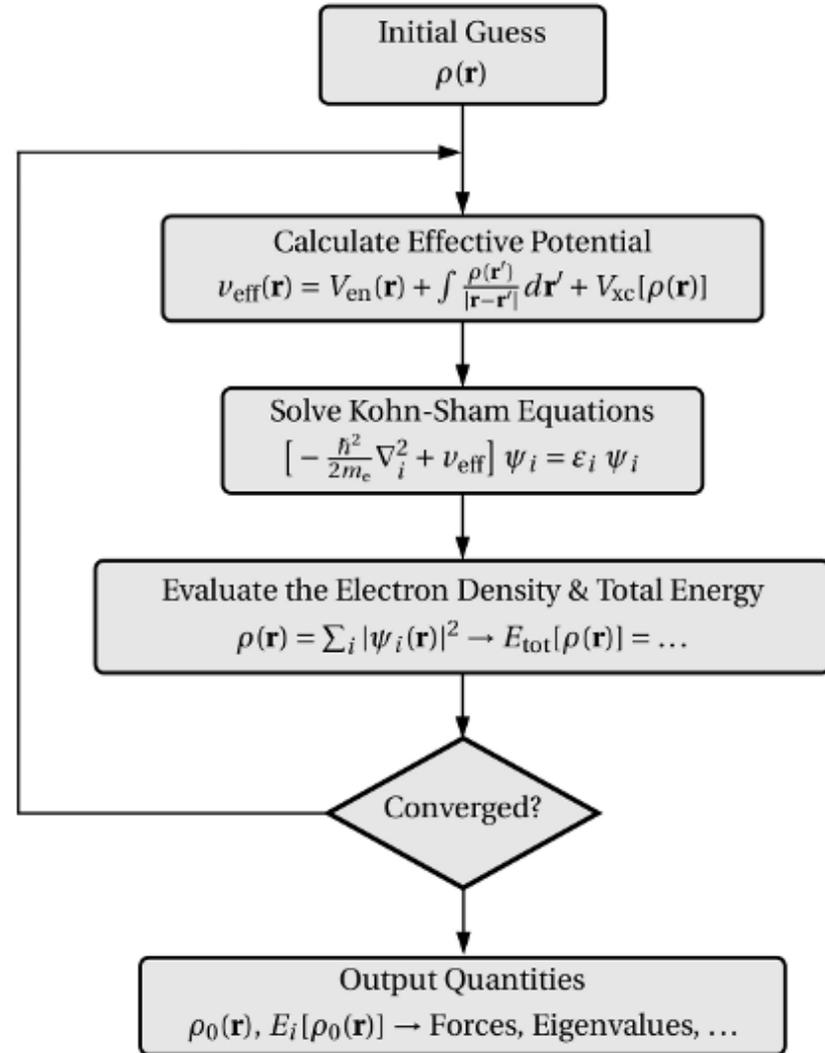
Input: atomic positions / external potential \mathbf{v}_{ext}



Recap: KS-DFT

Input: *atomic positions / external potential* v_{ext}

- ML approaches:*
- $v_{\text{ext}} \rightarrow E_{\text{tot}}(\rho_0)$

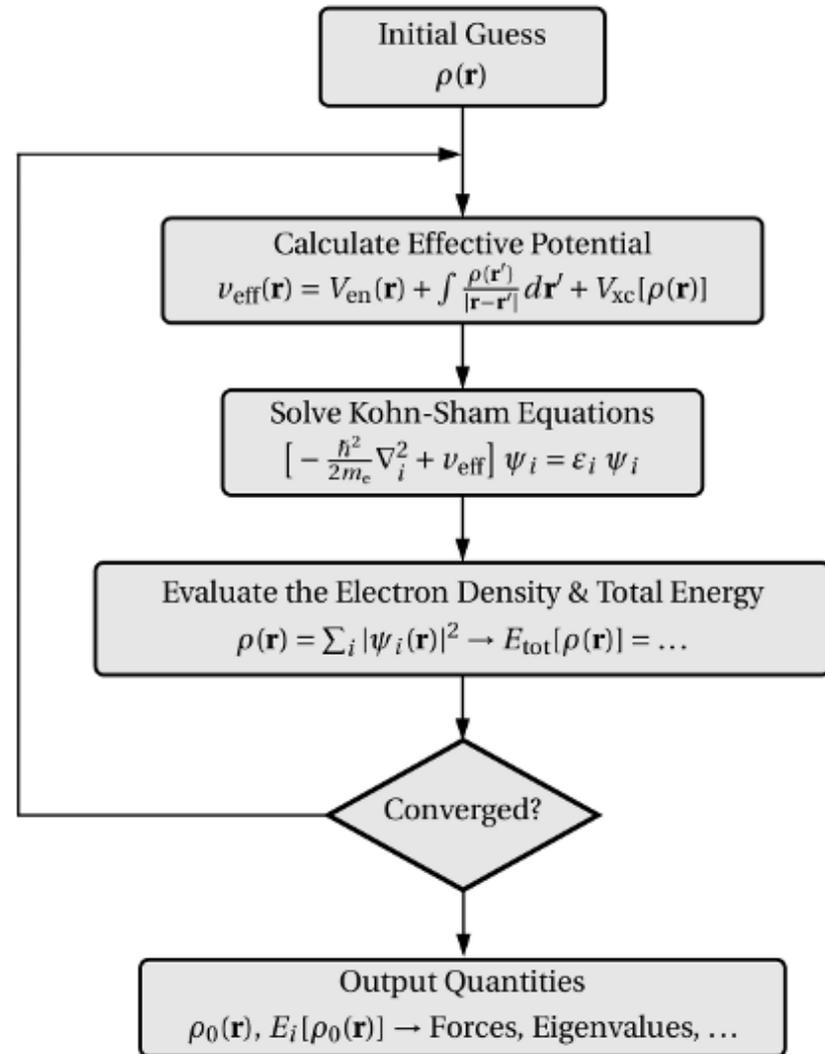


Recap: KS-DFT

Input: *atomic positions / external potential* v_{ext}

ML approaches:

- $v_{\text{ext}} \rightarrow E_{\text{tot}}(\rho_0)$
- $v_{\text{ext}} \rightarrow \rho_0$

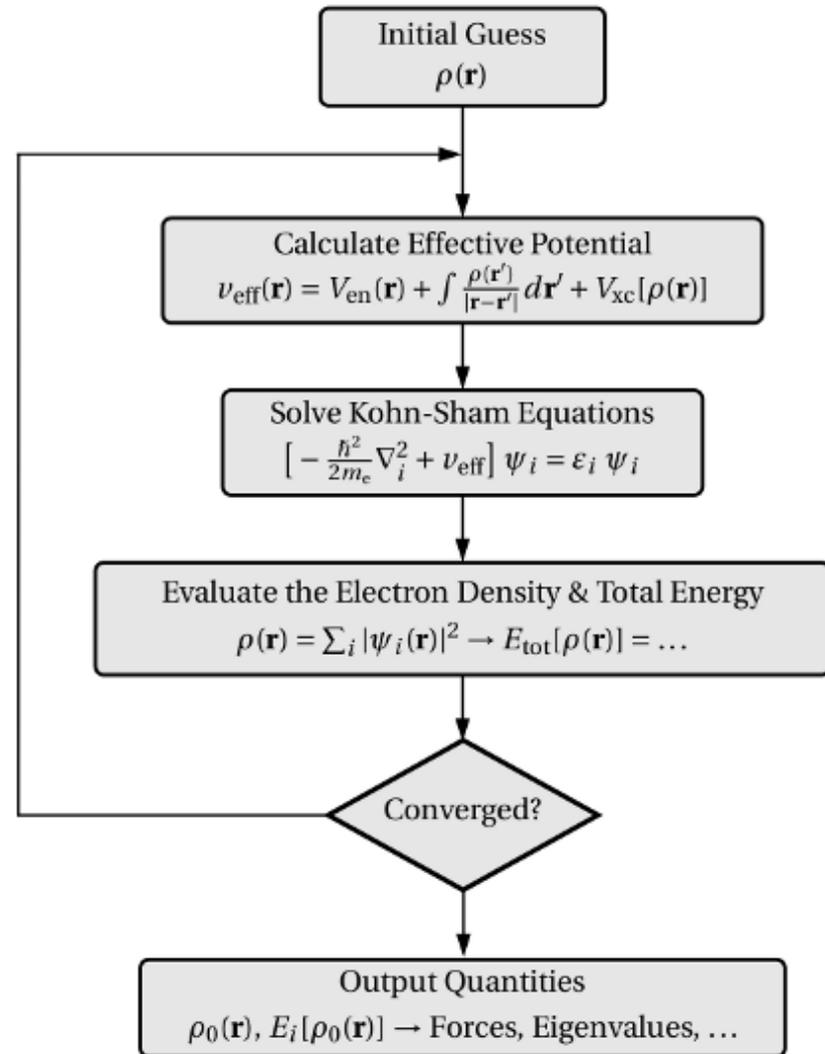


Recap: KS-DFT

Input: *atomic positions / external potential* v_{ext}

ML approaches:

- $v_{\text{ext}} \rightarrow E_{\text{tot}}(\rho_0)$
- $v_{\text{ext}} \rightarrow \rho_0$
- $\rho \rightarrow T_s[\rho]$
- ...

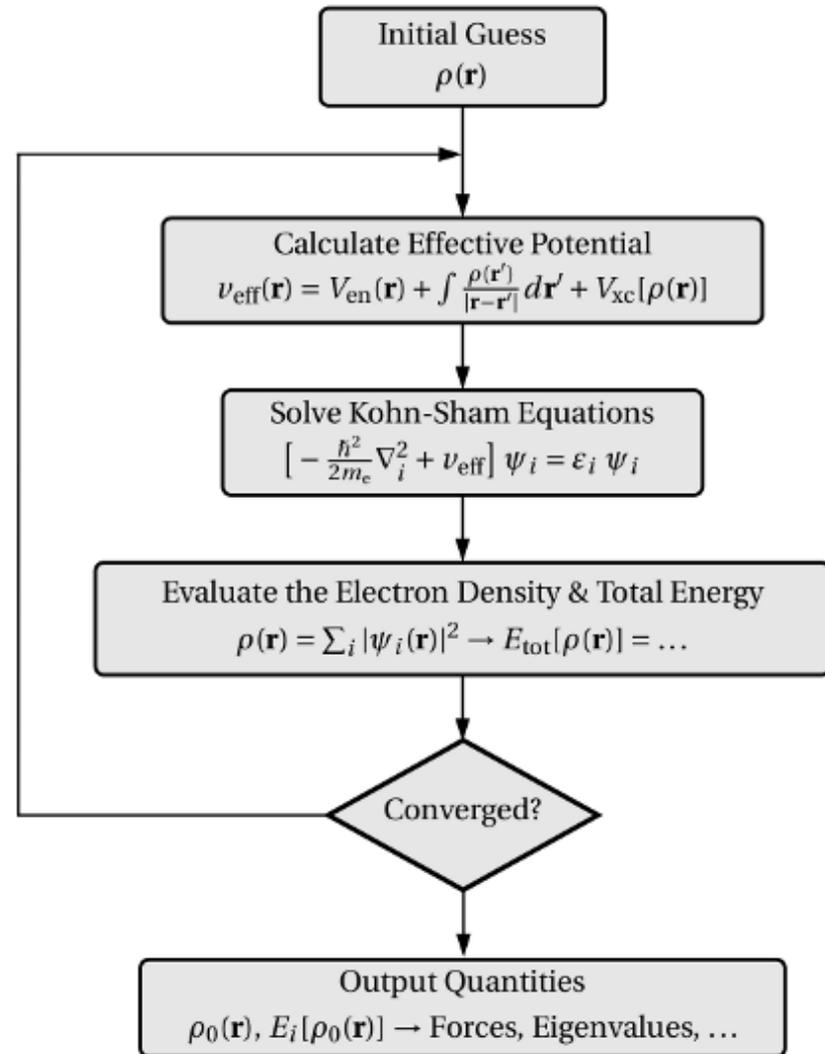


Recap: KS-DFT

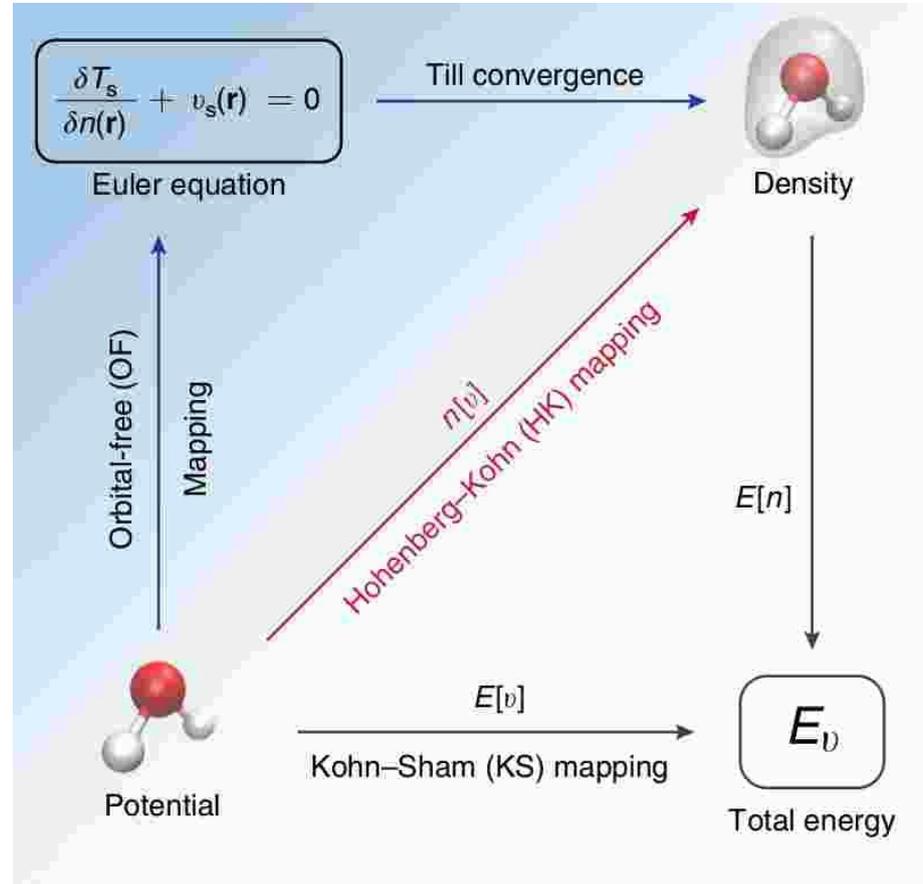
Input: *atomic positions / external potential* \mathbf{v}_{ext}

ML approaches:

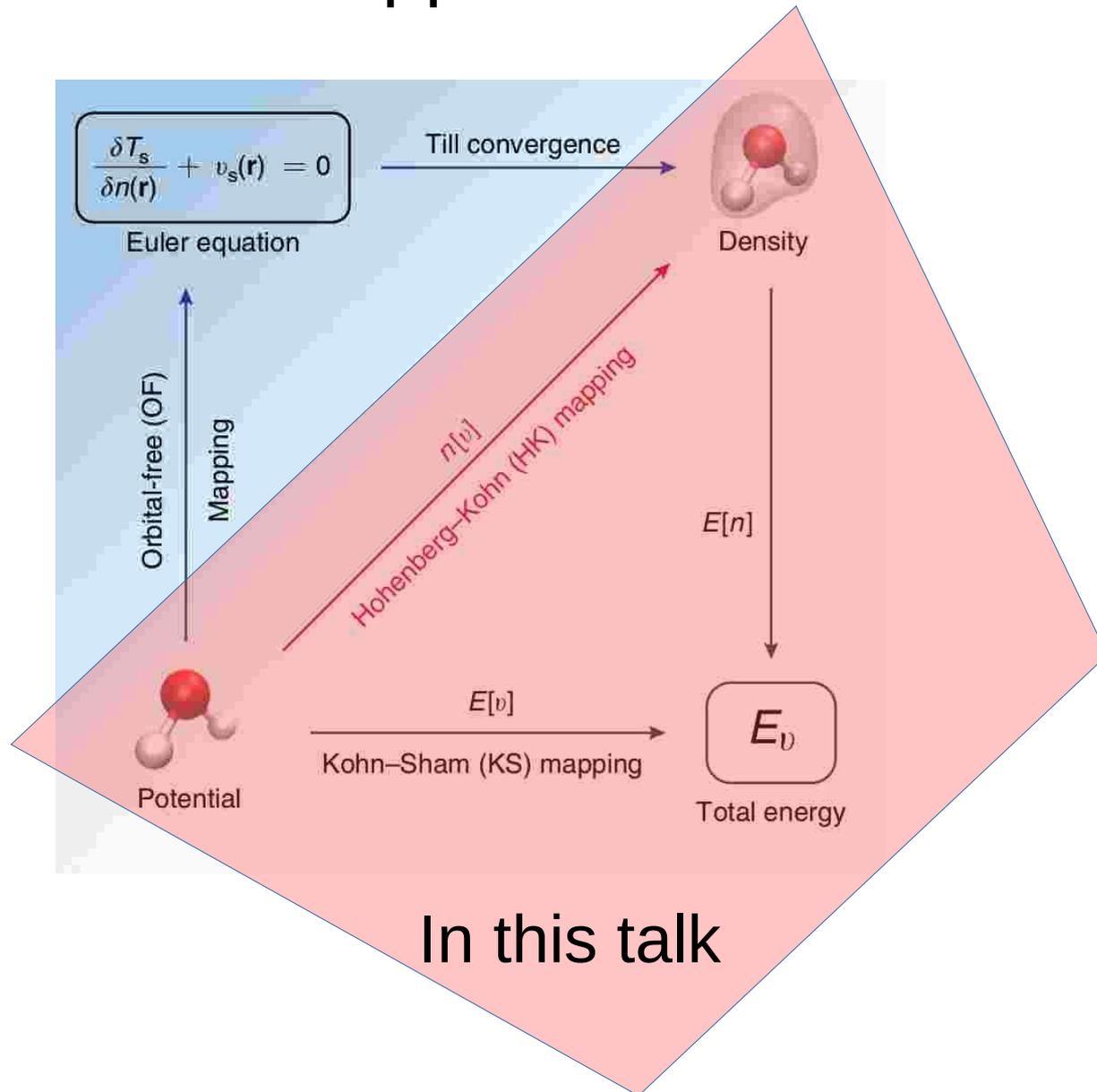
- $\mathbf{v}_{\text{ext}} \rightarrow E_{\text{tot}}(\rho_0)$ ($E^{\text{ML}}[\mathbf{v}]$)
- $\mathbf{v}_{\text{ext}} \rightarrow \rho_0$ ($n^{\text{ML}}[\mathbf{v}]$)
- $\rho \rightarrow T_s[\rho]$
- ...



Different ML approaches in DFT



Different ML approaches in DFT



Machine Learning the Hohenberg-Kohn Mapping

Use **basis set representation** of the density and **predict** the basis function **coefficients** $u^{(l)}$ (for each basis function ϕ_l):

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

Orthonormal basis functions

→ coefficients can be predicted **independently**

Here: Discrete Fourier Basis Set (plane waves)

Using kernel ridge regression (**KRR**) the predicted coefficients are:

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

with model parameters

$$\boldsymbol{\beta}^{(l)} = \left(\mathbf{K}_{\sigma^{(l)}} + \lambda^{(l)} \mathbf{I} \right)^{-1} \mathbf{u}^{(l)}, \quad l = 1, \dots, L$$

Gaussian kernel

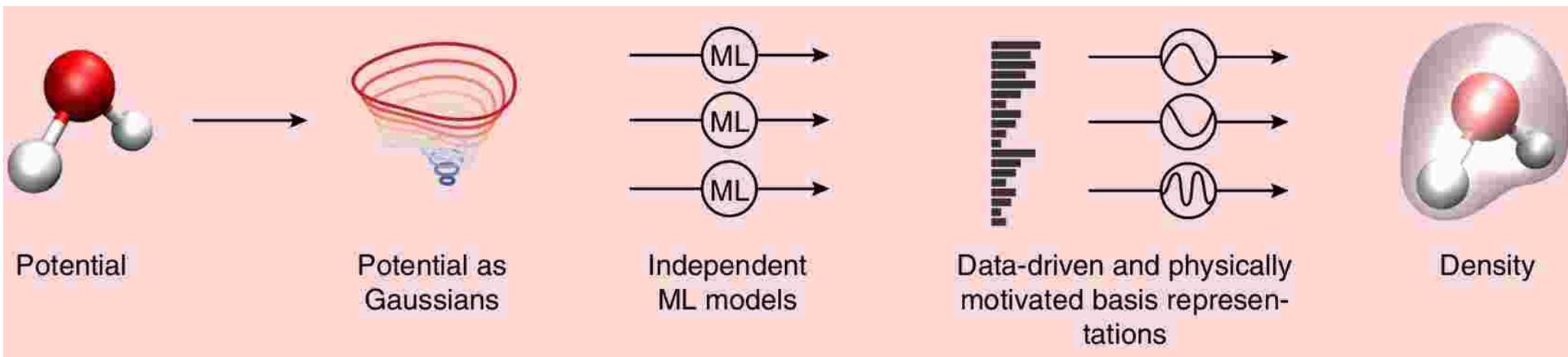
regularization parameter

What representation is used for the external potential?

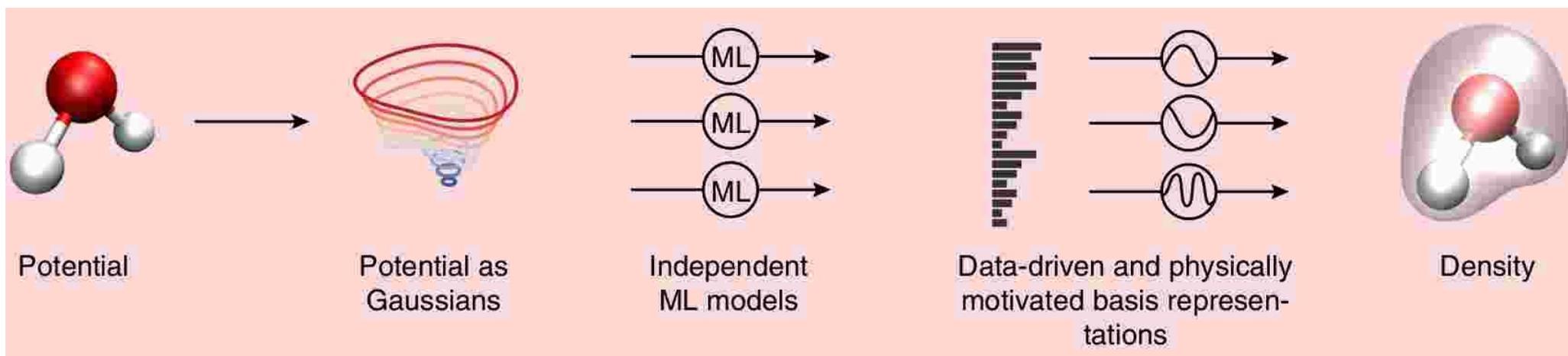
$$v(\mathbf{r}) = \sum_{\alpha=1}^{N^a} Z_{\alpha} \exp\left(\frac{-\|\mathbf{r} - \mathbf{R}_{\alpha}\|^2}{2\gamma^2}\right)$$

Artificial Gaussian potential

Overview of the approach



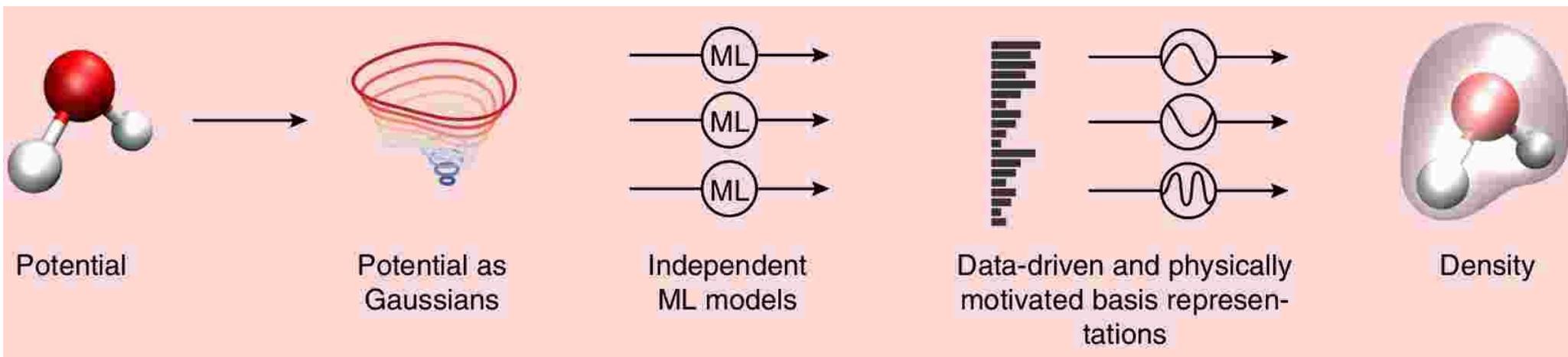
Overview of the approach



Using also the total energies:

Learning the prediction of total energies for given electron density: $n \rightarrow E^{ML}[n]$

Overview of the approach



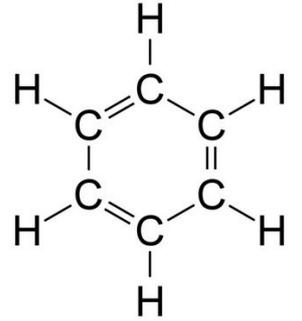
Using also the total energies:

Learning the prediction of total energies for given electron density: $n \rightarrow E^{ML}[n]$

$$\mathbf{v} \rightarrow n^{ML}[\mathbf{v}] \rightarrow E^{ML}[n^{ML}[\mathbf{v}]]$$

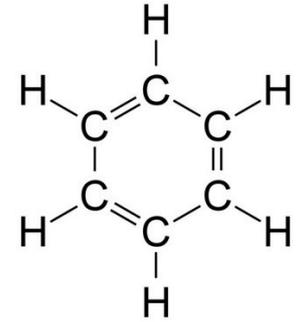
Results

PBE-Density predictions of benzene molecule



Results

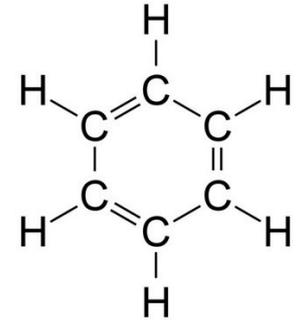
PBE-Density predictions of benzene molecule



- Geometries for training set created via MD simulations using classical force field:

Results

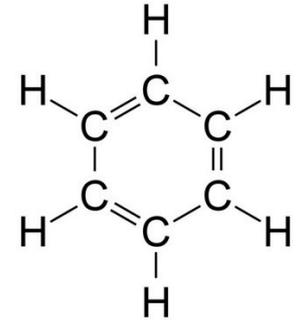
PBE-Density predictions of benzene molecule



- Geometries for training set created via MD simulations using classical force field:
 - Snapshots from isothermal MD simulations at 300 K, 350 K, and 400 K (General Amber Force Field in PINY_MD package) → large set of geometries
 - K-means sampling → 2000 representative points for the grand training set

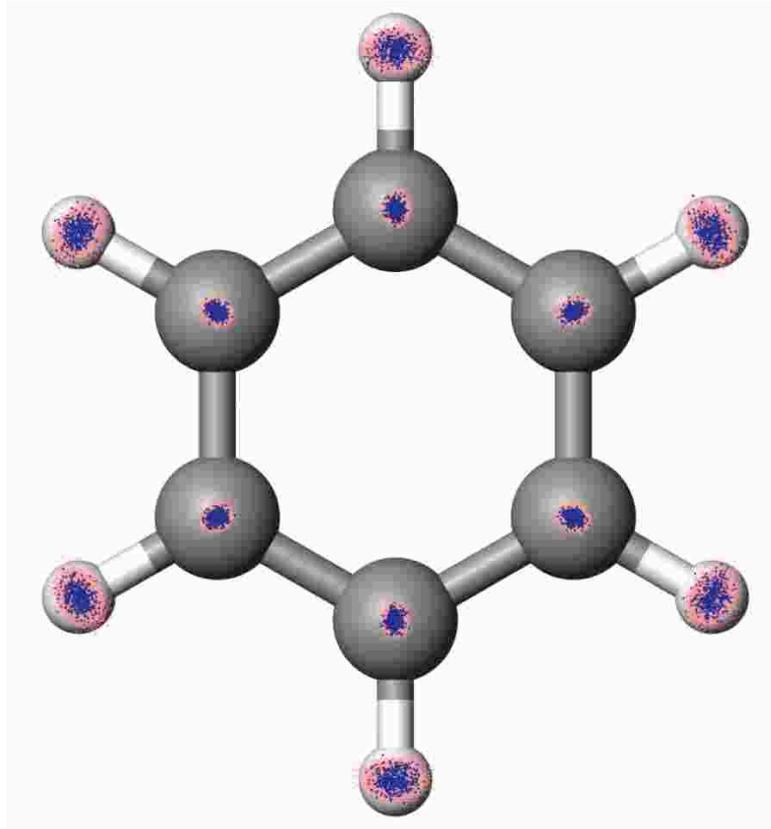
Results

PBE-Density predictions of benzene molecule

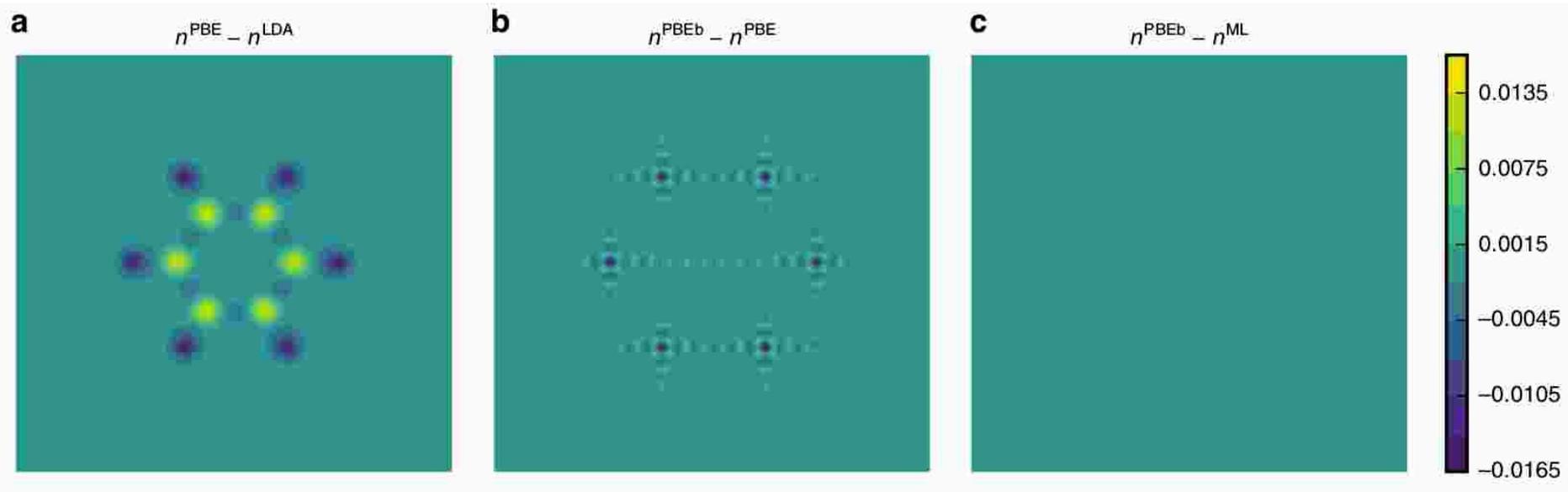


- Geometries for training set created via MD simulations using classical force field:
 - Snapshots from isothermal MD simulations at 300 K, 350 K, and 400 K (General Amber Force Field in PINY_MD package) → large set of geometries
 - K-means sampling → 2000 representative points for the grand training set
- DFT calculations performed for these geometries (Quantum ESPRESSO with PBE-functional)

The extent of benzene conformers:



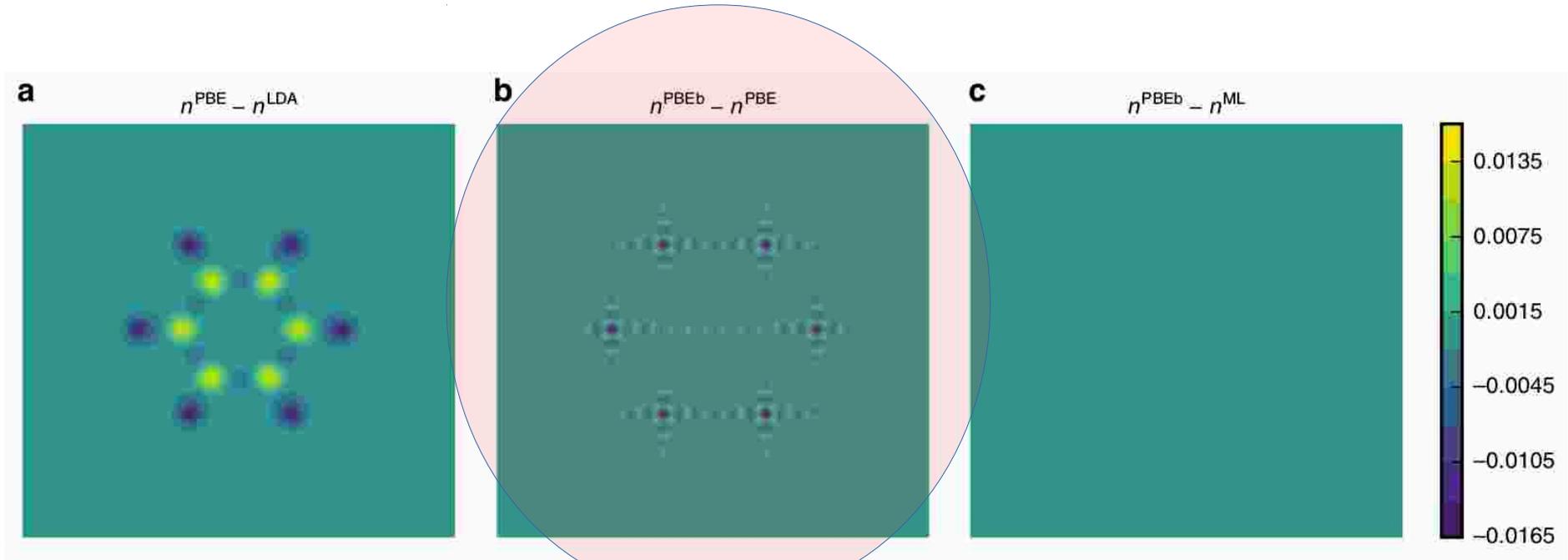
- The conformers generated by MD are displayed in red.
- K-means sampling is used to select 2,000 representative points.
- Test points from an independent trajectory are in blue.



Valence density
difference between
PBE and LDA functionals
(at PBE optimized geometry)

Fourier basis
representation
error

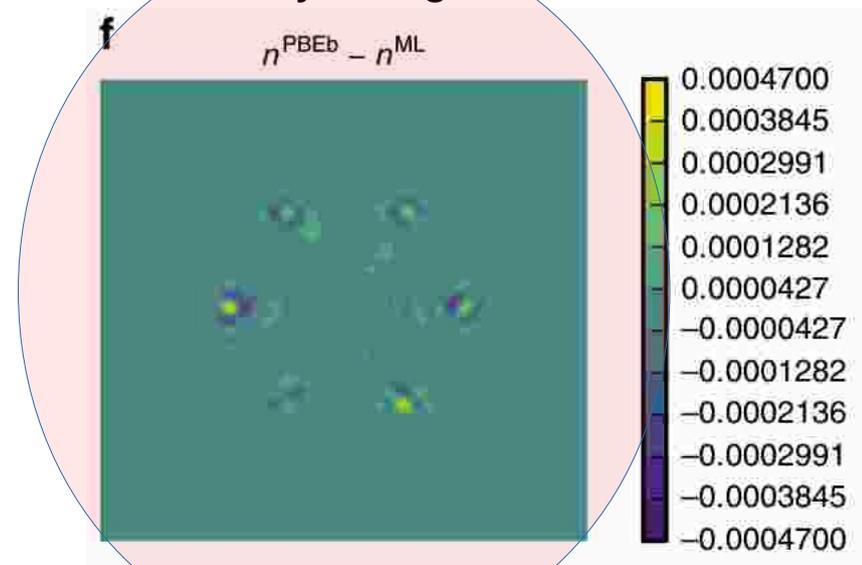
Machine learning
density fitting error



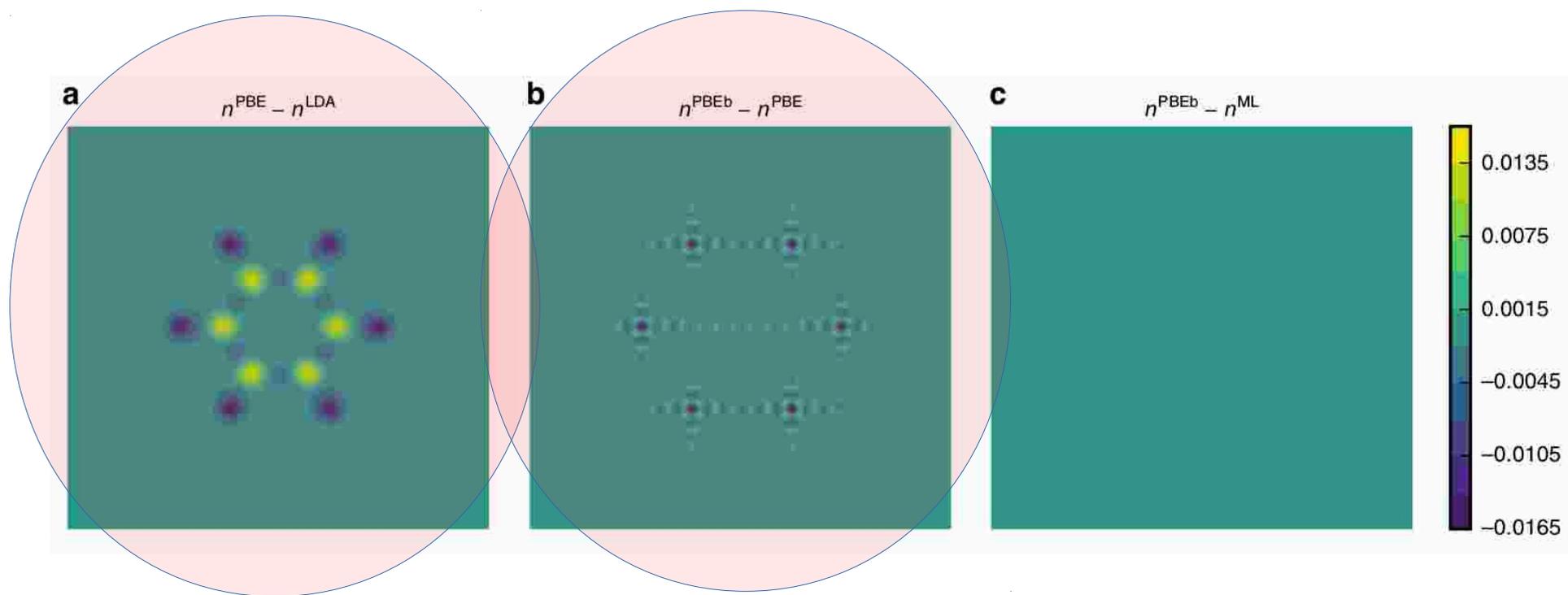
Valence density difference between PBE and LDA functionals (at PBE optimized geometry)

Fourier basis representation error

Machine learning density fitting error



- Errors introduced by the **Fourier basis representation** **two orders of magnitude larger** than those introduced by the **ML HK-map**

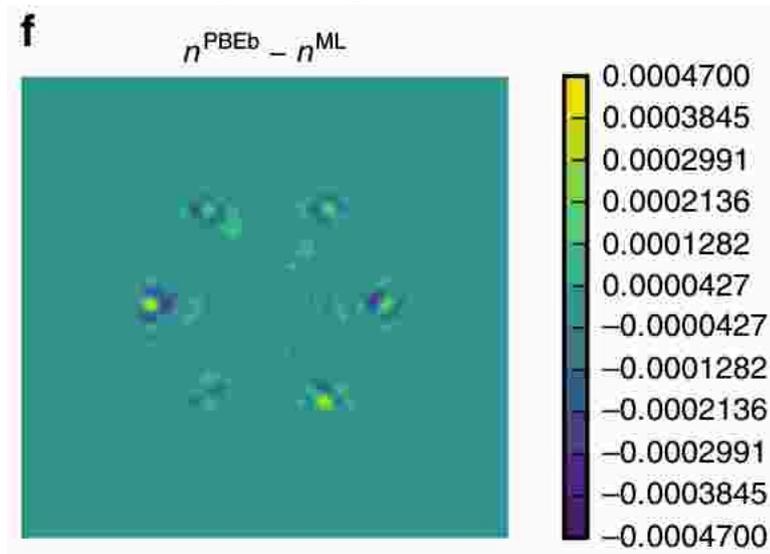


Valence density difference between PBE and LDA functionals (at PBE optimized geometry)

Fourier basis representation error

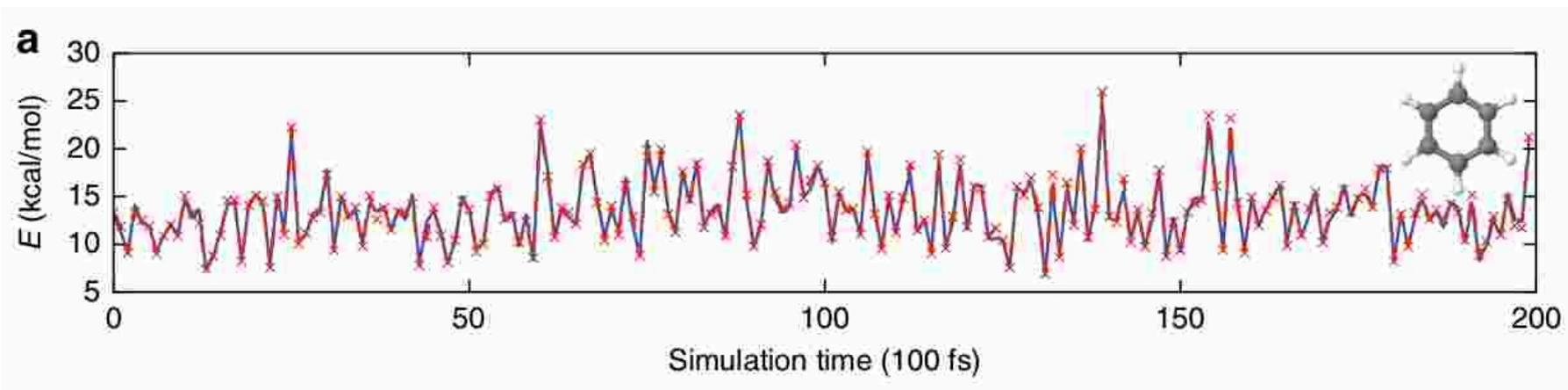
Machine learning density fitting error

- ML-HK errors considerably smaller than difference in PBE and LDA densities
 - ML-HK map specific to densities used to train in the model
 - should be able to **differentiate between densities** generated with **other approaches**



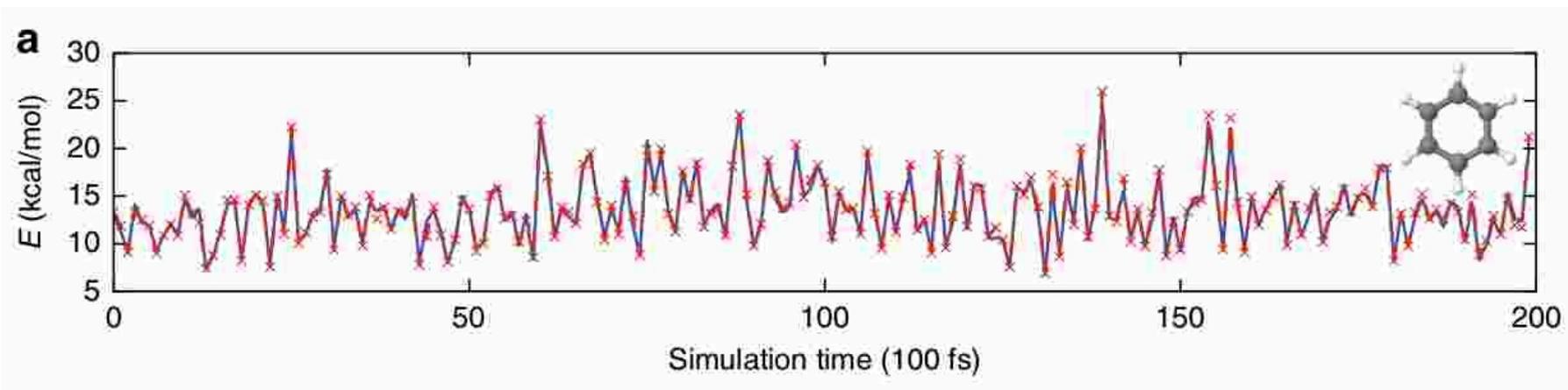
- Errors introduced by the **Fourier basis representation two orders of magnitude larger** than those introduced by the **ML HK-map**

Final ML model tested on 200 conformational snapshots of independent MD trajectory (300K)



PBE-values in blue, ML-HK values in red

Final ML model tested on 200 conformational snapshots of independent MD trajectory (300K)



PBE-values in blue, ML-HK values in red

Mean absolute error (**MAE**)
and max absolute error (**Max**)

Table 3 Energy and density-driven errors of the ML-HK approach on the MD data sets

Molecule	Training trajectories	ΔE		ΔE_D^{ML}	
		MAE	Max	MAE	Max
Benzene	300 K	0.42	1.7	0.32	1.5
	300 + 350 K	0.37	1.8	0.28	1.5
	300 + 400 K	0.47	2.3	0.30	1.8
Ethane	300 K	0.20	1.5	0.17	1.3
	300 + 350 K	0.23	1.4	0.19	1.1
	300 + 400 K	0.14	1.7	0.098	0.62
Malonaldehyde	300 + 350 K	0.27	1.2	0.21	0.74

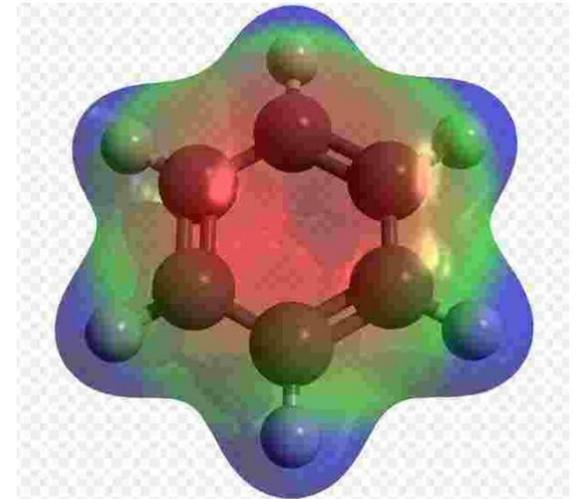
Errors are given in kcal/mol for different training trajectory combinations

Discussion

- Representation of external potential (expressed as sum of Gaussians) not rotational invariant
 - more difficult to construct a predictive model
 - try other representations? (e.g. SOAP, MBTR, ...)
- Given enough data, is it possible to learn the ‘entire’ HK-map / DFT-functional?
- Training on much more accurate quantum chemical densities and energies → construction of (nearly) exact density functionals (significantly reducing the computational cost)
- Prediction of density matrix (i.e. basis set coefficients for each molecular orbital) can be used for training hybrid functionals

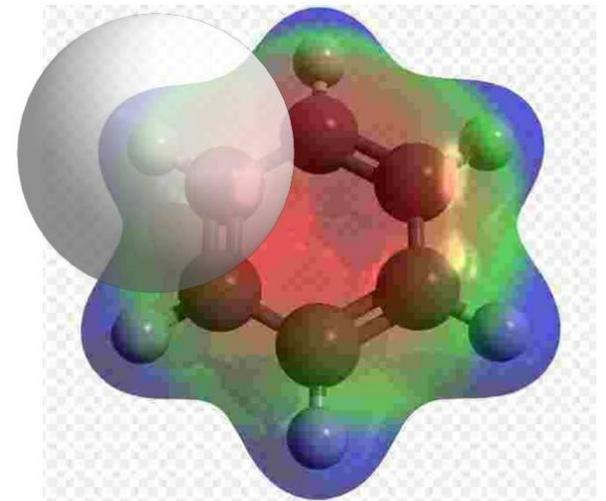
Discussion

- Instead of predicting directly the entire electron density:



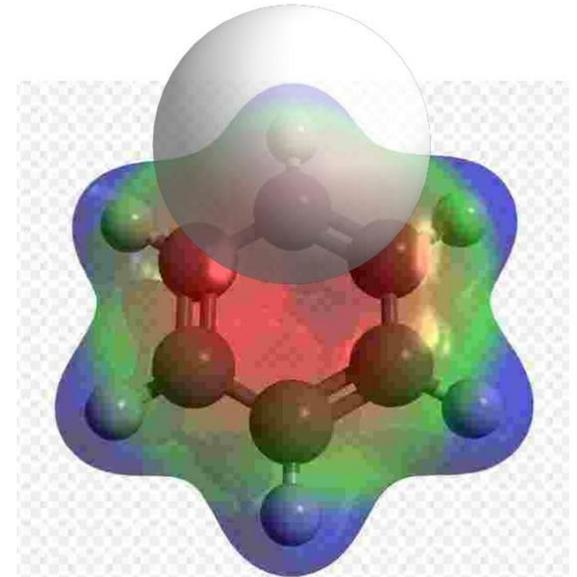
Discussion

- Instead of predicting directly the entire electron density:
 - Predict local electron densities centered at each atom



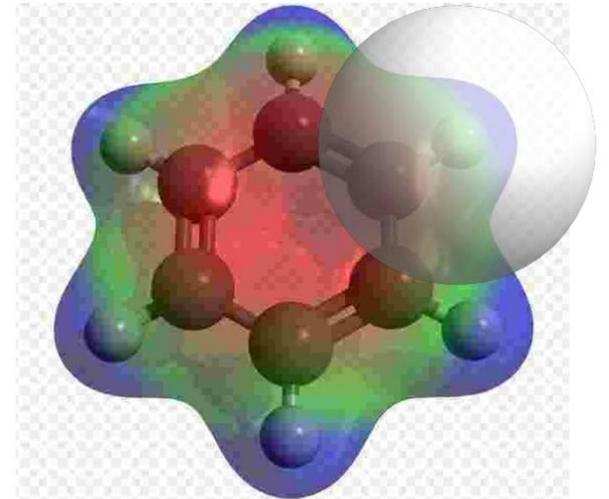
Discussion

- Instead of predicting directly the entire electron density:
 - Predict local electron densities centered at each atom



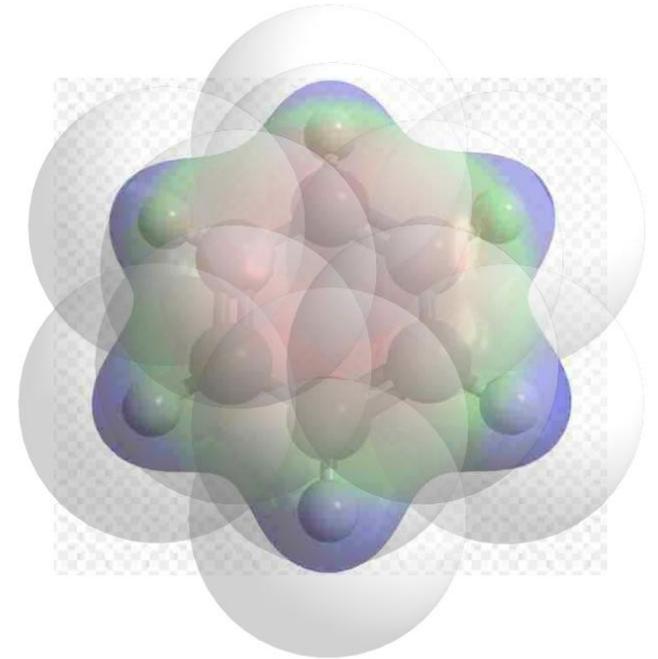
Discussion

- Instead of predicting directly the entire electron density:
 - Predict local electron densities centered at each atom



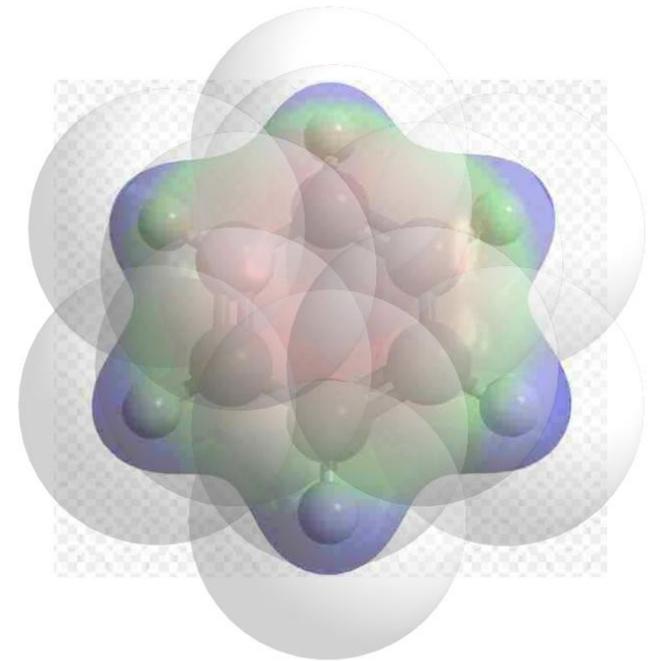
Discussion

- Instead of predicting directly the entire electron density:
 - Predict local electron densities centered at each atom
 - Finally: combining all local densities



Discussion

- Instead of predicting directly the entire electron density:
 - Predict local electron densities centered at each atom
 - Finally: combining all local densities

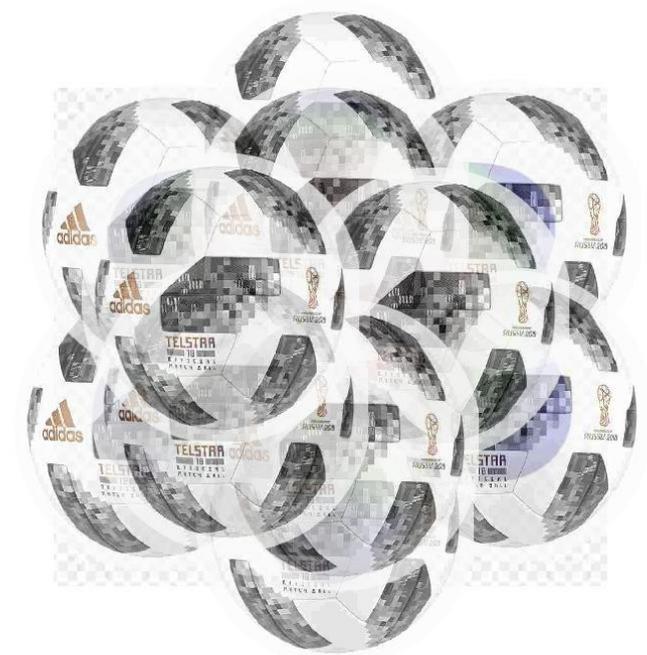


NAO
representation
→ Predicting
NAO basis
functions

Discussion

- Instead of predicting directly the entire electron density:
 - Predict local electron densities centered at each atom
 - Finally: combining all local densities

**Thank you for your
attention**



**NAO
representation**
→ Predicting
NAO basis
functions

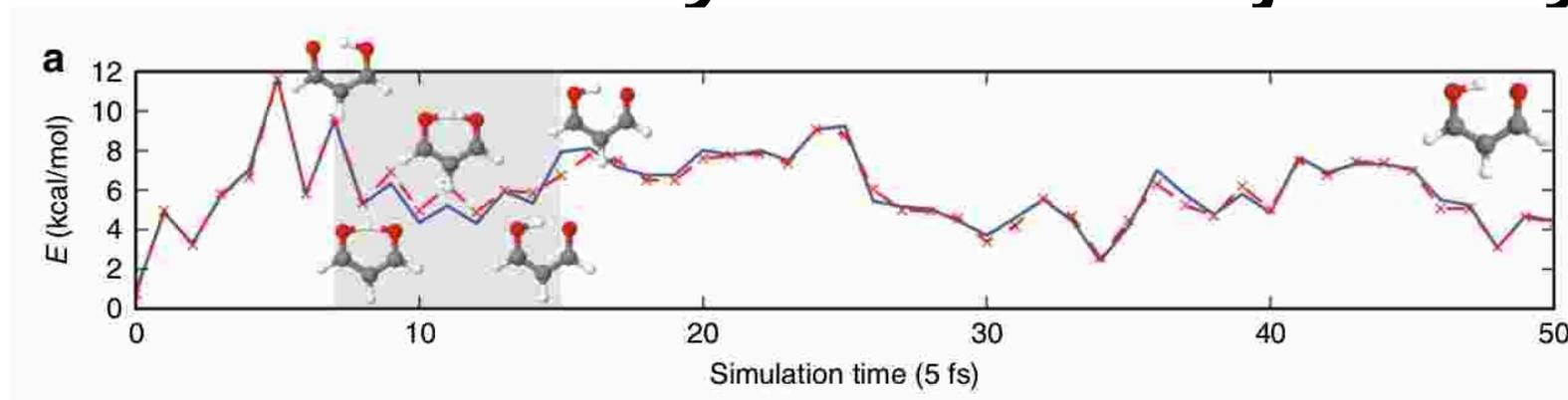
Comparison ML-HK and ML-KS

Table 2 Prediction errors on H₂ and H₂O

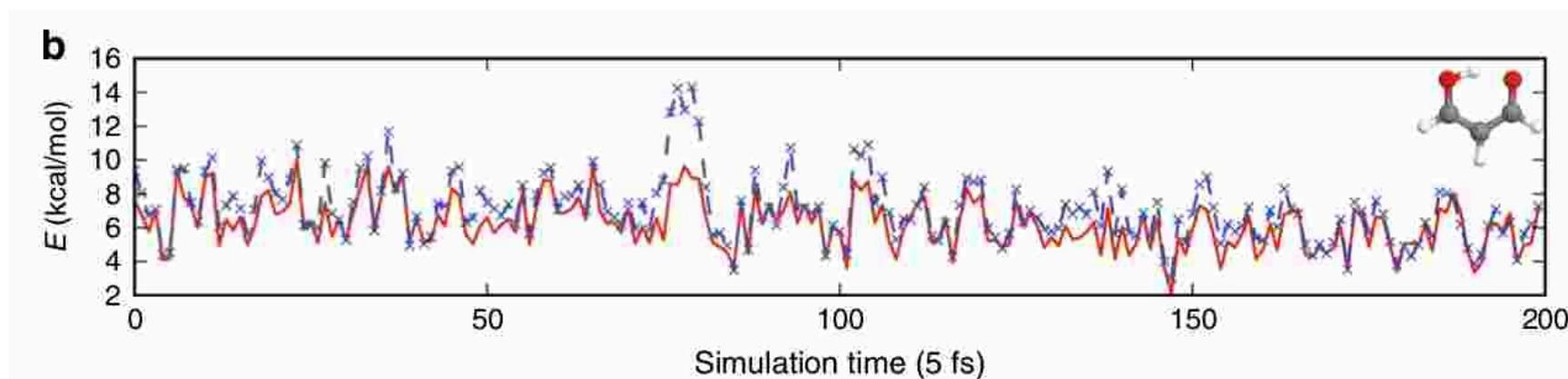
Molecule	<i>M</i>	ML-KS				ML-HK					
		ΔE		ΔR_o	$\Delta\theta_o$	ΔE		ΔE_D^{ML}		ΔR_o	$\Delta\theta_o$
		MAE	Max			MAE	Max	MAE	Max		
H ₂	5	1.3	4.3	2.2	—	0.70	2.9	0.18	0.54	1.1	—
	7	0.37	1.4	0.23	—	0.17	0.73	0.054	0.16	0.19	—
	10	0.080	0.41	0.23	—	0.019	0.11	0.017	0.086	0.073	—
H ₂ O	5	1.4	5.0	2.1	2.2	1.1	4.9	0.056	0.17	2.3	3.8
	10	0.27	0.93	0.63	1.9	0.12	0.39	0.099	0.59	0.12	0.38
	15	0.12	0.47	0.19	0.41	0.043	0.25	0.029	0.14	0.064	0.23
	20	0.015	0.064	0.043	0.16	0.0091	0.060	0.011	0.058	0.024	0.066

Errors are shown for increasing numbers of training points *M* for the ML-KS and ML-HK approaches. In addition, the estimated density-driven contribution to the error for the ML-HK approach (Eq. (9)) is given. Energies are given in kcal/mol, bond-lengths in pm, and angles in degrees

Malonaldehyde MD trajectory



Energy errors of ML-HK along a 0.25 ps ab initio MD trajectory of malonaldehyde. PBE values in blue, ML-HK values in red.



Energy errors of ML-HK along a 1 ps MD trajectory of malonaldehyde generated by the ML-HK model. ML-HK values in red, PBE values of trajectory snapshots in blue.

Total Energy Errors from ML-HK generated trajectory snapshots

