

# Real-space Machine Learning of Correlation Density Functionals



UNIVERSITÉ DE FRIBOURG  
UNIVERSITÄT FREIBURG

Elias Polak

University of Fribourg  
Vuckovic Group



Workshop on Quantum Theory: Foundations  
and Extensions of Density Functional Theory

Oslo

December 2nd – 6th, 2024

05/12/2024

# Outline

- Introduction
- Second-Order Perturbation Theory based Correlation Energy Density
- Local and Global Energy Training
- Machine Learning the Spin-Component-Scale



# Introduction

## Density Functional Theory (DFT)

- 1. Hohenberg-Kohn Theorem

$$\rho \mapsto v[\rho]$$

- 2. Hohenberg-Kohn Theorem

$$E_0 = \min_{\rho} E_v[\rho]$$

$$\text{s.t. } \rho^{1/2} \in H^1, \rho \geq 0, \|\rho\|_{L^1} = N$$

## Practical considerations of DFT

- Kohn-Sham (KS) DFT:

$$\left( -\frac{1}{2} \nabla^2 + v_s[\rho] \right) \phi_k = \varepsilon_k \phi_k, \rho^{KS}(\mathbf{r}) = \sum_k |\phi_k(\mathbf{r})|^2$$

- Energy Density Functional

$$E_v^{KS}[\rho] = T_s[\rho] + U_H[\rho] + \int v\rho + E_{XC}[\rho]$$

- Density Functional Approximations (DFAs)

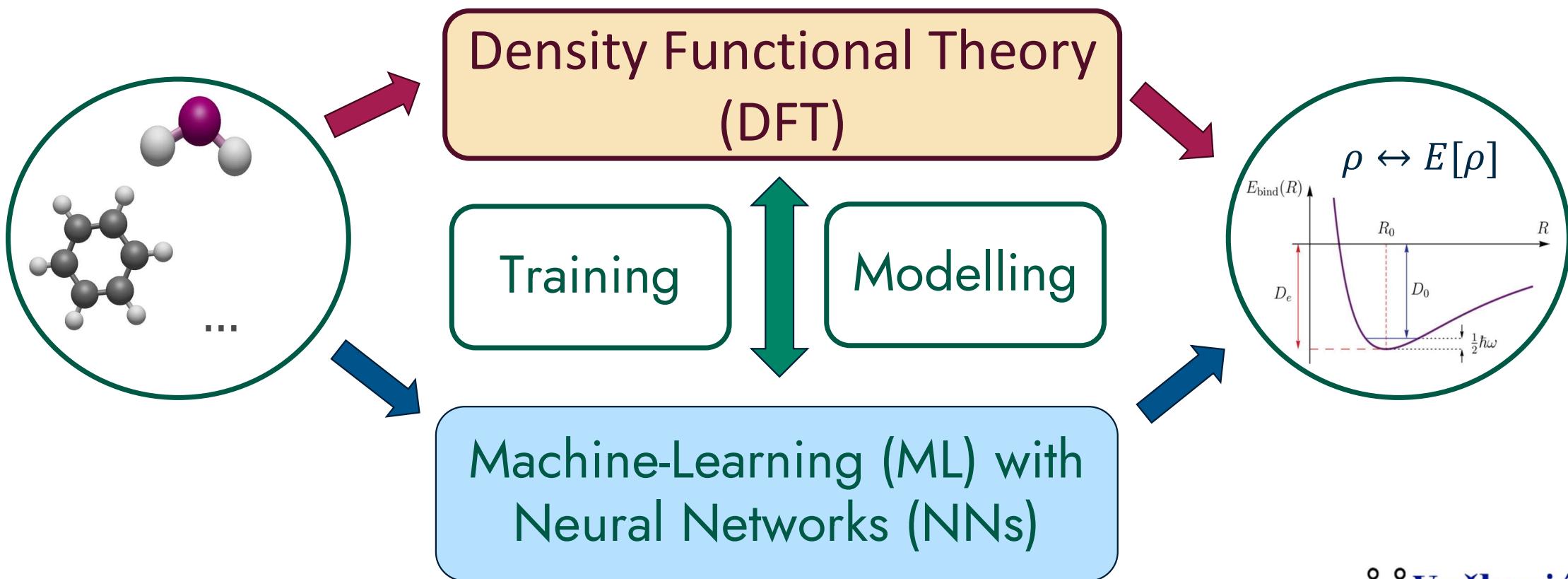
$$E_{XC}[\rho] = E_X[\rho] + \tilde{E}_C[\rho]$$



P. Hohenberg & W. Kohn, *Phys. Rev.* **136** (3B), 1964.

W. Kohn & L. J. Sham, *Phys. Rev.* **140** (4A), 1965.

# Introduction



# Introduction

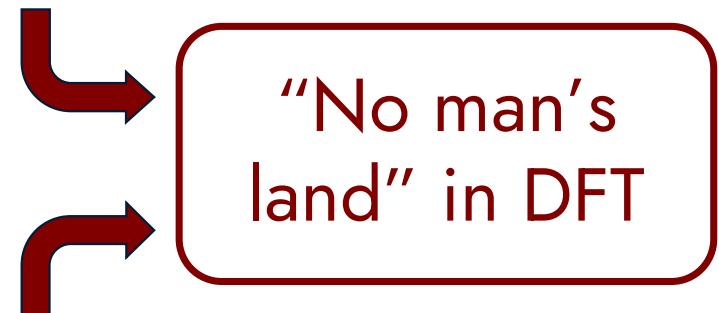
## Training

- DFT reference data:
  - Energies  $E^{DFA}[\rho]$
  - Electronic density  $\rho^{KS}(\mathbf{r})$
  - Forces

## Modelling

- ML models:
  - Energies  $E^{ML}[\rho]$
  - Electronic density  $\rho^{ML}(\mathbf{r})$
  - Potentials  $v^{ML}[\rho](\mathbf{r})$

! Accuracy of DFAs: how reliable is the training on DFT data?



! Transferability to unseen data: does the ML model performance on training set A extrapolate to set B?



# ML DFT

5

- ❖ Transferability key challenge in ML DFT
- ❖ “No man’s land” between old and new school functional approximations
- ❖ Training set affects extrapolation
  
- ❖ Real-space learning of energies
- ❖ Robust training in low-data regime
- ❖ Full discrete space becomes dataset

REVIEW

## The central role of density functional theory in the AI age

Huang et al., *Science* **381**, 170–175 (2023)

QUANTUM CHEMISTRY

Kirkpatrick et al., *Science* **374**, 1385–1389 (2021)

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

DM21

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

Density functional theory describes matter at the quantum level, but all popular approximations suffer from systematic errors that arise from the violation of mathematical properties of the exact functional. We overcame this fundamental limitation by training a neural network on molecular data and on fictitious systems with fractional charge and spin. The resulting functional, DM21 (DeepMind 21), correctly describes typical examples of artificial charge delocalization and strong correlation and performs better than traditional functionals on thorough benchmarks for main-group atoms and molecules. DM21 accurately models complex systems such as hydrogen chains, charged DNA base pairs, and diradical transition states. More crucially for the field, because our methodology relies on data and constraints, which are continually improving, it represents a viable pathway toward the exact universal functional.



Cite this: *Phys. Chem. Chem. Phys.*, 2024, 26, 12289

Received 28th February 2024,  
Accepted 28th March 2024

DOI: 10.1039/d4cp00878b

rsc.li/pccp

## Deep Mind 21 functional does not extrapolate to transition metal chemistry†

Heng Zhao,  Tim Gould  and Stefan Vuckovic  \*<sup>a</sup>

The development of density functional approximations stands at a crossroads: while machine-learned functionals show potential to surpass their human-designed counterparts, their extrapolation to unseen chemistry lags behind. Here we assess how well the recent Deep Mind 21 (DM21) machine-learned functional [Science, 2021, **374**, 1385–1389], trained on main-group chemistry, extrapolates to transition metal chemistry (TMC). We show that DM21 demonstrates comparable or occasionally superior accuracy to B3LYP for TMC, but consistently struggles with achieving self-consistent field convergence for TMC molecules. We also compare main-group and TMC machine-learning DM21 features to shed light on DM21’s challenges in TMC. We finally propose strategies to overcome limitations in the extrapolative capabilities of machine-learned functionals in TMC.

ARTICLE

OPEN



## Completing density functional theory by machine learning hidden messages from molecules

Ryo Nagai<sup>1,2\*</sup>, Ryosuke Akashi<sup>1</sup> and Osamu Sugino<sup>1,2</sup>

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

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



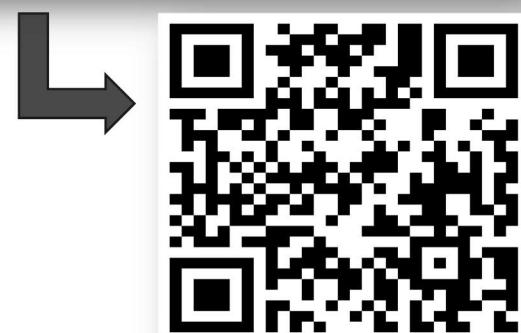
Cite this: DOI: 10.1039/d4sc02358g  
All publication charges for this article have been paid for by the Royal Society of Chemistry

Received 10th April 2024  
Accepted 2nd June 2024  
DOI: 10.1039/d4sc02358g  
rsc.li/chemical-science

## Identifying and embedding transferability in data-driven representations of chemical space†

Tim Gould,  Bun Chan,  Stephen G. Dale  ac and Stefan Vuckovic  \*

Transferability, especially in the context of model generalization, is a paradigm of all scientific disciplines. However, the rapid advancement of machine learned model development threatens this paradigm, as it can be difficult to understand how transferability is embedded (or missed) in complex models developed using large training data sets. Two related open problems are how to identify, without relying on human intuition, what makes training data transferable; and how to embed transferability into training data. To solve both problems for *ab initio* chemical modelling, an indispensable tool in everyday chemistry research, we introduce a transferability assessment tool (TAT) and demonstrate it on a controllable data-driven model for developing density functional approximations (DFAs). We reveal that human intuition in the curation of training data introduces chemical biases that can hamper the transferability of data-driven DFAs. We use our TAT to motivate three transferability principles; one of which introduces the key concept of transferable diversity. Finally, we propose data curation strategies for general-purpose machine learning models in chemistry that identify and embed the transferability principles.



Elias Polak

Real-space Machine Learning of Correlation Density Functionals

05/12/2024

# Energy Density Functionals

$$E^{KS}[\rho] = \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} + T_s[\rho] + U[\rho] + E_X[\rho] + E_C[\rho]$$

Energy functional:  $E[\rho] = \int e(\mathbf{r})\rho(\mathbf{r})d\mathbf{r}$

Energy density:  $e(\mathbf{r}) = e[\rho](\mathbf{r})$

$$\Rightarrow E_C[\rho] = \int e_C(\mathbf{r})\rho(\mathbf{r})d\mathbf{r}$$

 *Correlation Energy Density*

Energy densities are not uniquely defined!



# A well-defined Correlation Energy Density <sup>7</sup>

Gauge from many-body wave function theory

❑ Electrostatic potential of XC–hole

- Yields exact exchange energy density  $e_X(\mathbf{r})$
- Used as “exact” XC energy density

❑ Second-order perturbation theory framework

- Møller-Plesset (MP2) functional expression
- Local correlation effects from occupied and unoccupied orbitals



$$e_c^{MP2}[\rho](\mathbf{r})$$

F. G. Cruz, K.-C. Lam, and K. Burke, *J. Phys. Chem. A* 1998, **102**, 4911.

K. Burke, F. G. Cruz, and K.-C. Lam, *J. Chem. Phys.* 1998, **109**, 8161.

S. Vuckovic, T.J.P. Irons, A. Savin, A.M. Teale, and P. Gori-Giorgi, *J. Chem. Theory Comput.* 2016, **12**, 2598–2610.

S. Vuckovic, M. Levy, and P. Gori-Giorgi, *J. Chem. Phys.* 2017, **147**, 214107.



# The MP2 Correlation Energy Density

$$e_c^{MP2}(\mathbf{r})$$

$$= -\frac{1}{4\rho(\mathbf{r})} \sum_{ijab} t_{ijab} v_{ijab}(\mathbf{r}),$$

$$\begin{aligned} i, j &= 1, \dots, N \\ a, b &= N, N+1, \dots \end{aligned}$$

$$t_{ijab} = T_{ijab}\delta_{ia}\delta_{jb} - T_{ijba}\delta_{ib}\delta_{ja},$$

$$T_{ijab} = \frac{\langle ij|ab \rangle}{\Delta_{ijab}} \quad \delta_{mn} = \begin{cases} 1, & \sigma_m = \sigma_n, \\ 0, & \sigma_m \neq \sigma_n \end{cases}$$

$$\Delta_{ijab} = \varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j$$

$$v_{ijab}(\mathbf{r}) = V_{ijab}(\mathbf{r})\delta_{ia}\delta_{jb} - V_{ijba}(\mathbf{r})\delta_{ib}\delta_{ja}$$

$$V_{ijab}(\mathbf{r}) = \phi_i(\mathbf{r})\phi_a(\mathbf{r}) \int \frac{\phi_j(\mathbf{r}')\phi_b(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}'$$

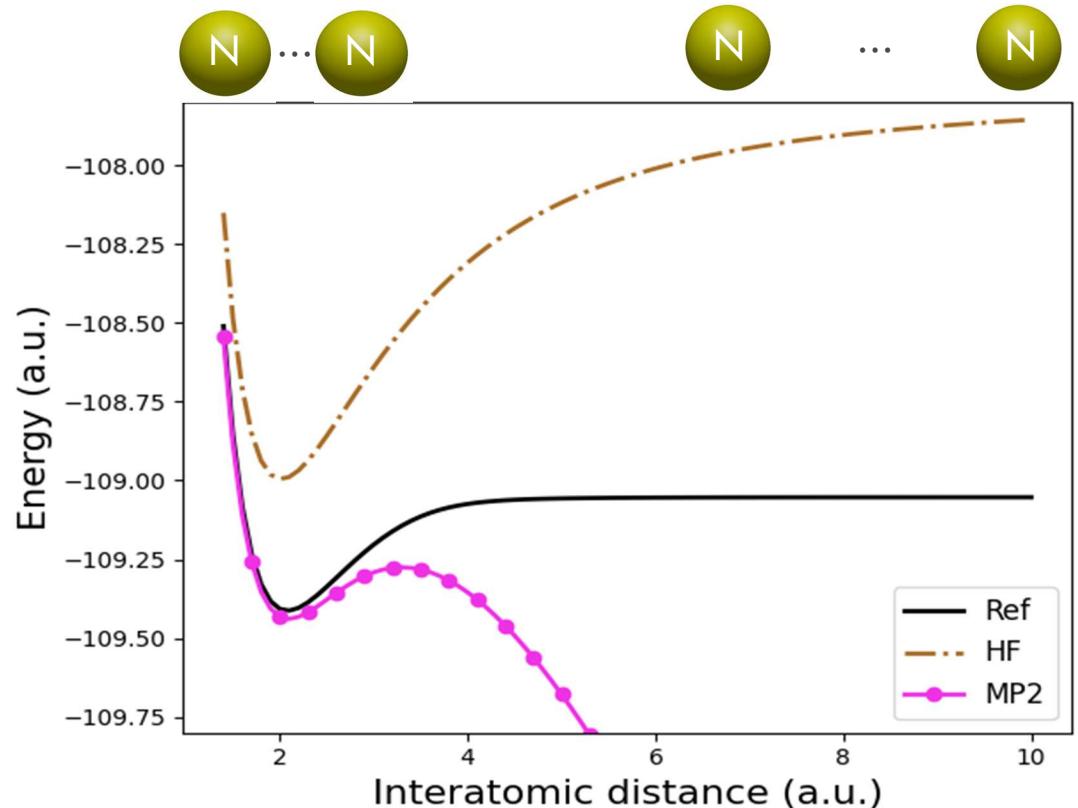
$$\phi_k(\mathbf{r}) = \phi_k[\rho](\mathbf{r})$$



# The MP2 Correlation Energy Density

$$e_c^{MP2}(\mathbf{r})$$

- Evaluation on Hartree-Fock (HF) orbitals
- Singularity when orbital energy gap vanishes
- Wrong electron correlation when breaking bonds



# Regularized MP2 Correlation Energy Density <sup>9</sup>

$$e_c^{\kappa MP2}(\mathbf{r})$$

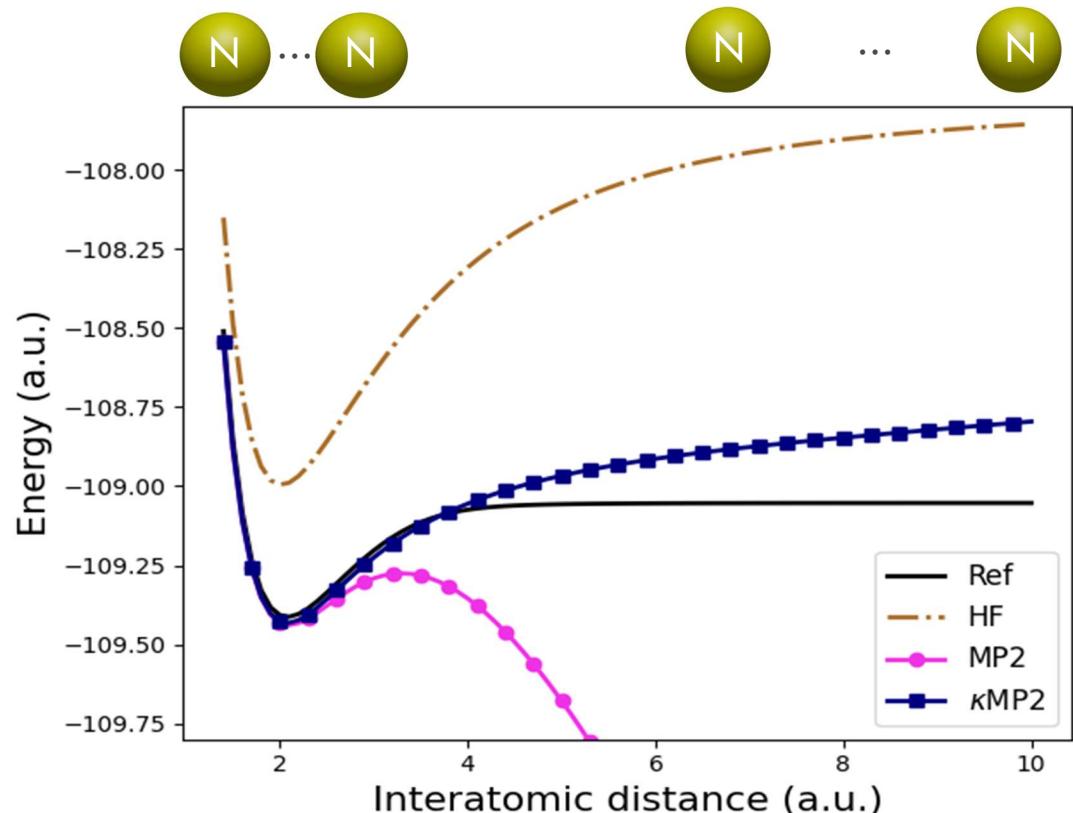
Regularization with  $\kappa > 0$ :

$$T_{ijab}^\kappa = \frac{\langle ij|ab \rangle}{\Delta_{ijab}} \left(1 - e^{-\kappa(\Delta_{ijab})}\right)^2$$

- No singularity when orbital energy gap vanishes
- Possible assessment of crucial correlation effects

J. Lee and M. Head-Gordon, *J. Chem. Theory. Comput.* 2018, **14**, 5203-5219.

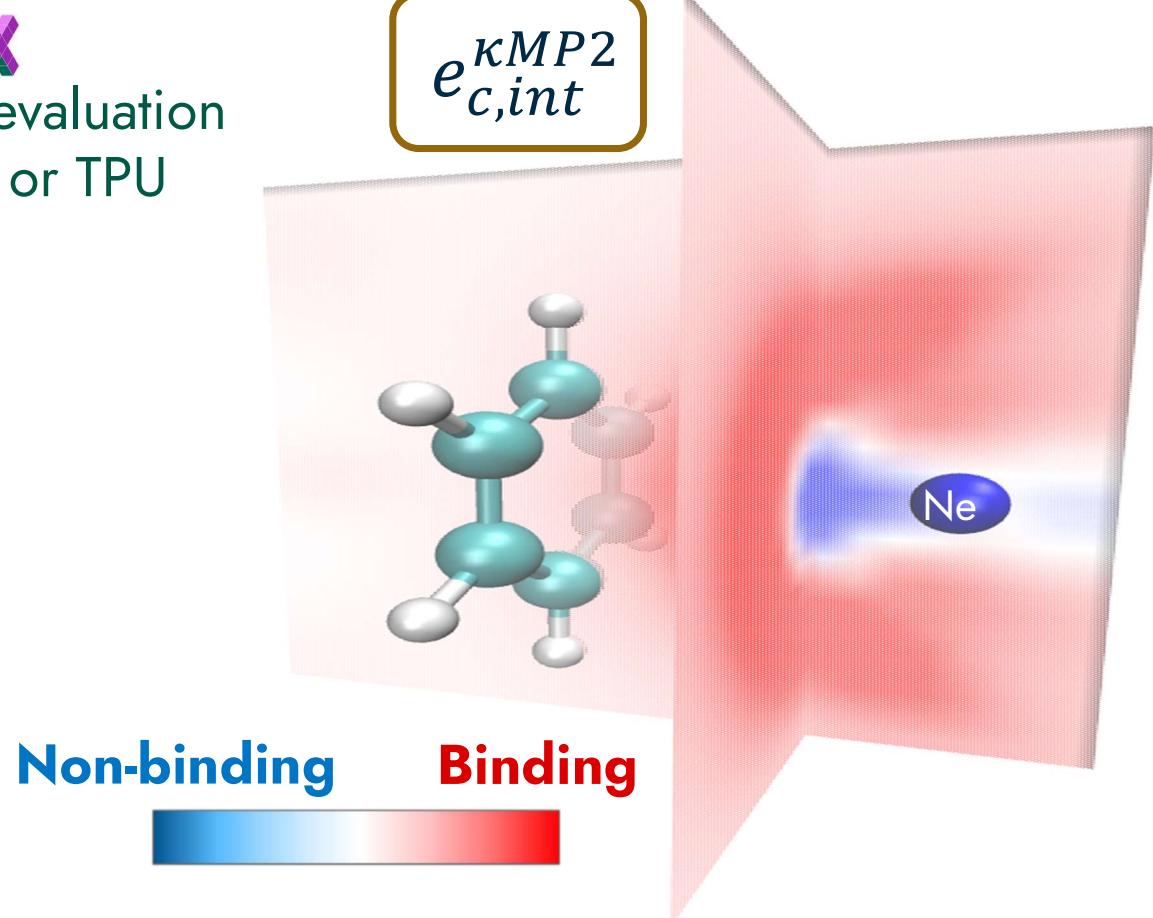
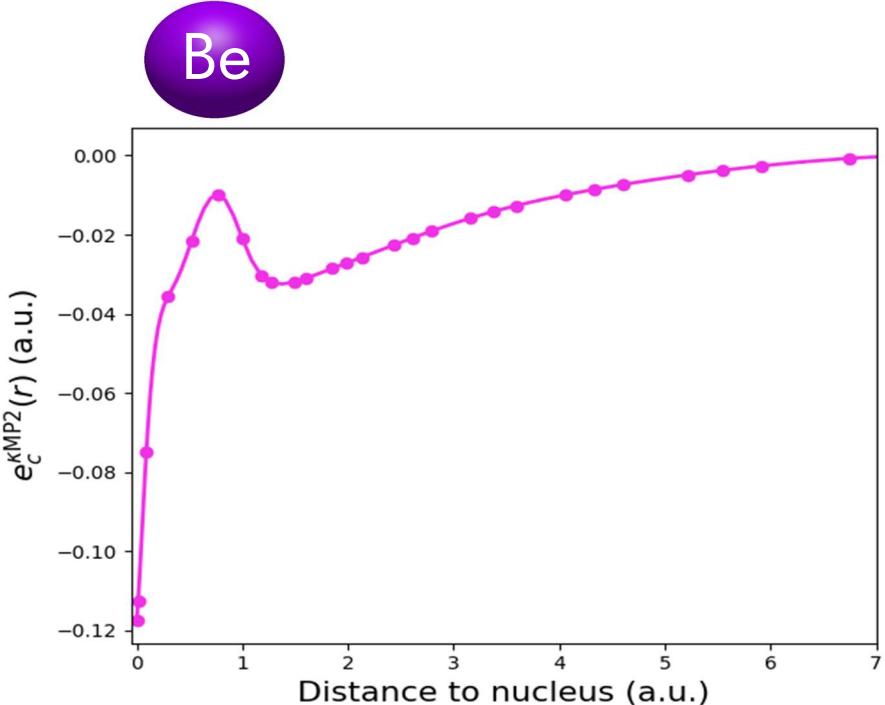
J. Shee et.al., *J. Phys. Chem. Lett.* 2021, **12**(50), 12084–12097.



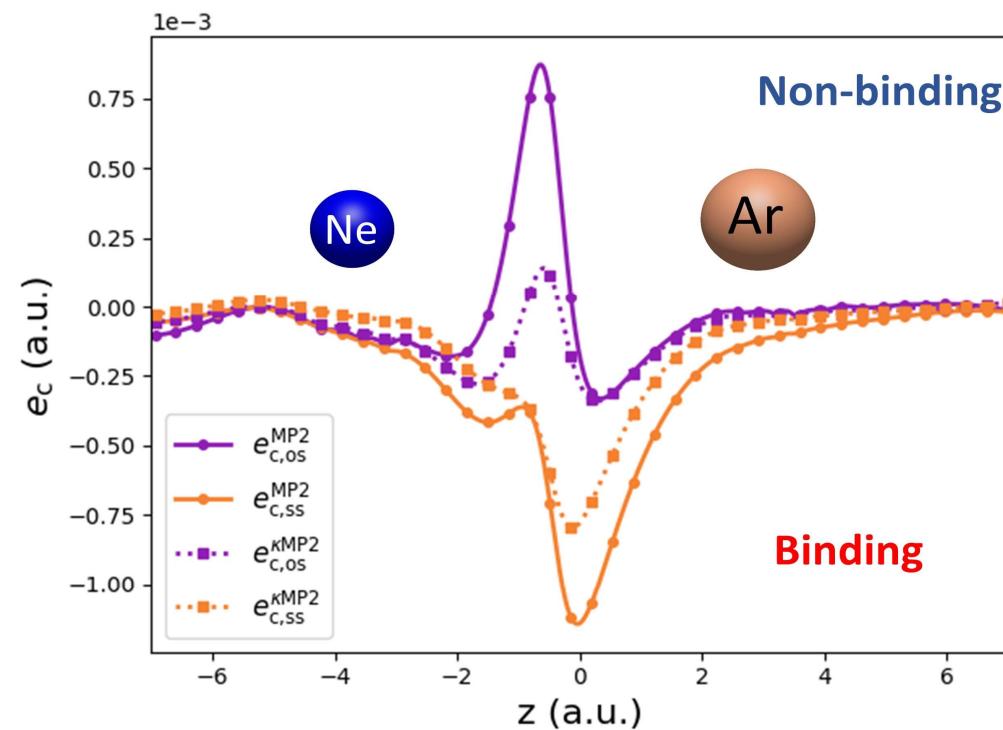
# Regularized MP2 Correlation Energy Density

$$e_c^{\kappa MP2}(\mathbf{r})$$

- ✓ Jax optimized 
- ✓ High performance evaluation
- ✓ Runs on CPU, GPU or TPU

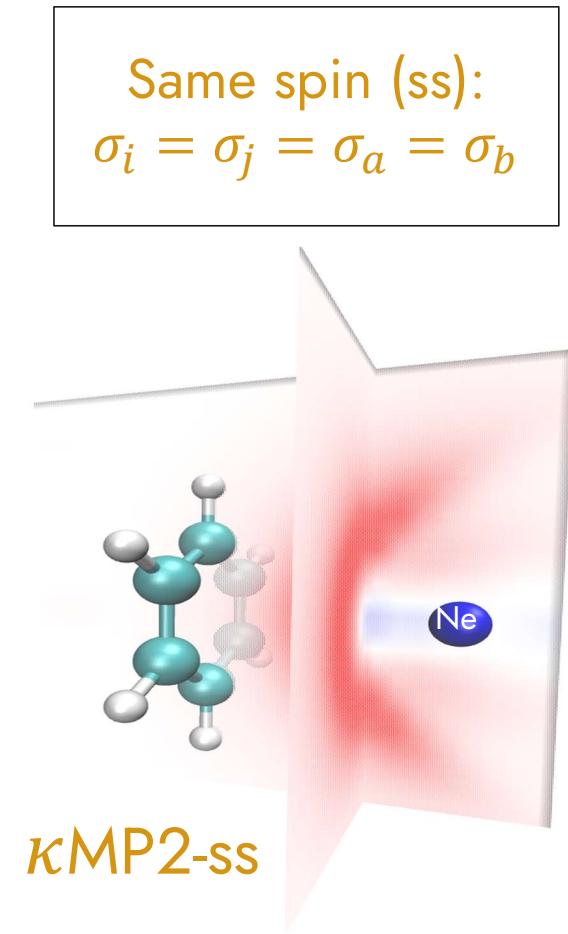
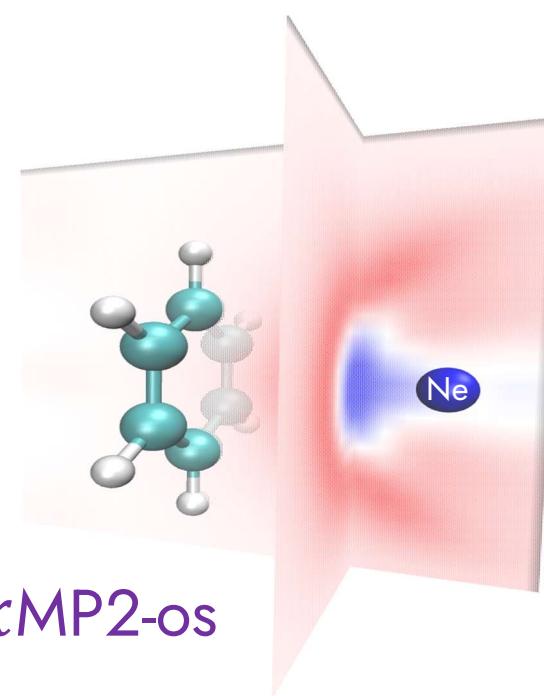


Spin resolution:  $e_c^{MP2} = e_{c,os}^{MP2} + e_{c,ss}^{MP2}$  11

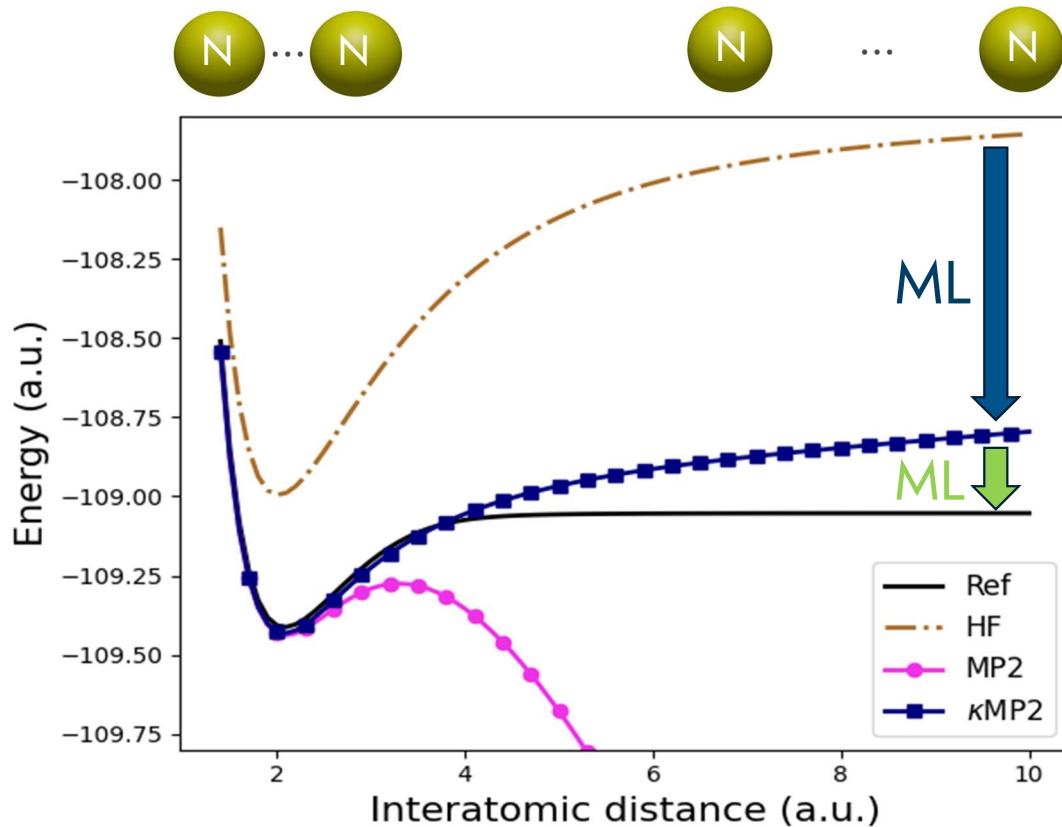


Opposite spin (os):  
 $\sigma_i = \sigma_a \neq \sigma_j = \sigma_b$

Same spin (ss):  
 $\sigma_i = \sigma_j = \sigma_a = \sigma_b$



# ML of Correlation Energies



- Real-space training of energy densities
- $\kappa$ MP2 as proxy reference
- Employing local information from  $\kappa$ MP2
- Learning reference correlation energies



# Local and Global Energy Training

ML of energy functionals:  $E[\rho] = \int e(\mathbf{r})\rho(\mathbf{r})d\mathbf{r}$

Global Energy Loss (GES):

$$\mathcal{L}_{\text{GES}} \sim |E^{\text{ML}}[\rho] - E^{\text{ref}}[\rho]|$$

Local Energy Loss (LES):

$$\mathcal{L}_{\text{LES}} \sim \int |e^{\text{ML}}(\mathbf{r}) - e^{\text{ref}}(\mathbf{r})| \rho(\mathbf{r}) d\mathbf{r}$$

**1 system = 1 energy data point**

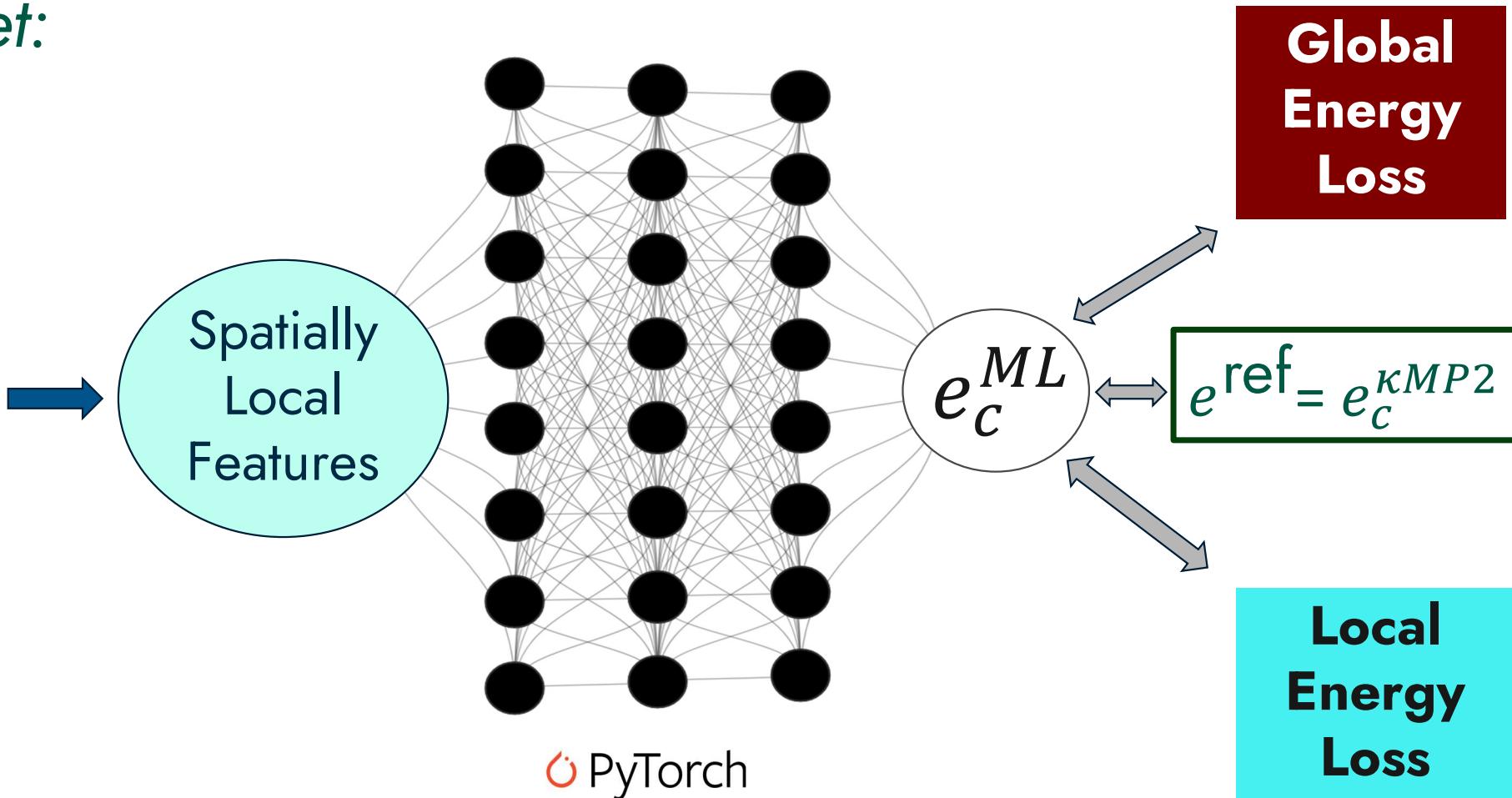
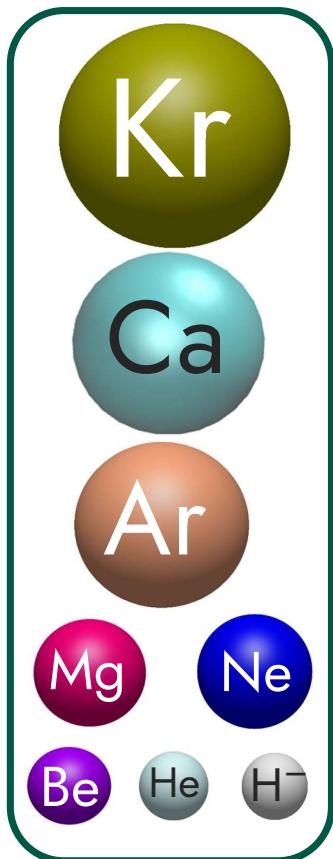
**1 system = 1 energy data set**



# ML of the Correlation Energy Density

14

Training set:



# A scaling invariant ML model

Uniformly scaled density:  $\rho_\gamma(\mathbf{r}) = \gamma^3 \rho(\gamma \mathbf{r})$  for  $\gamma \in \mathbb{R}^+$

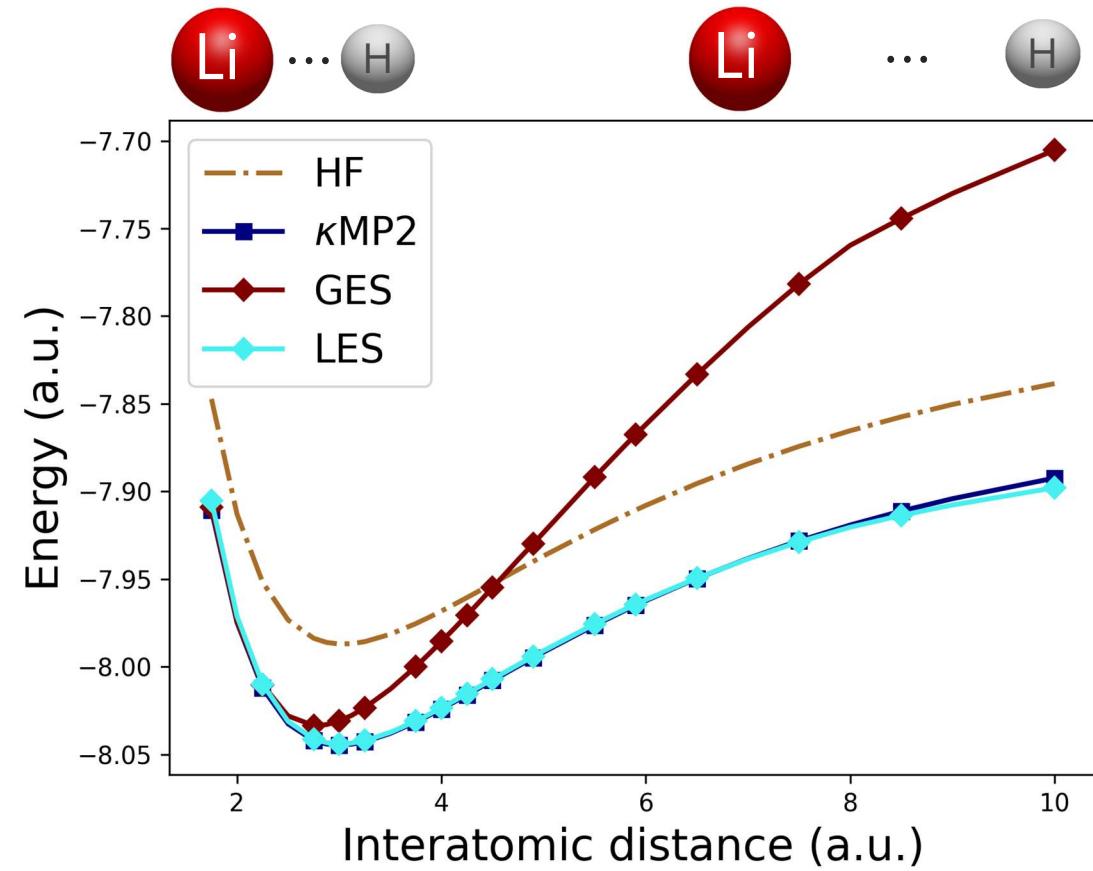
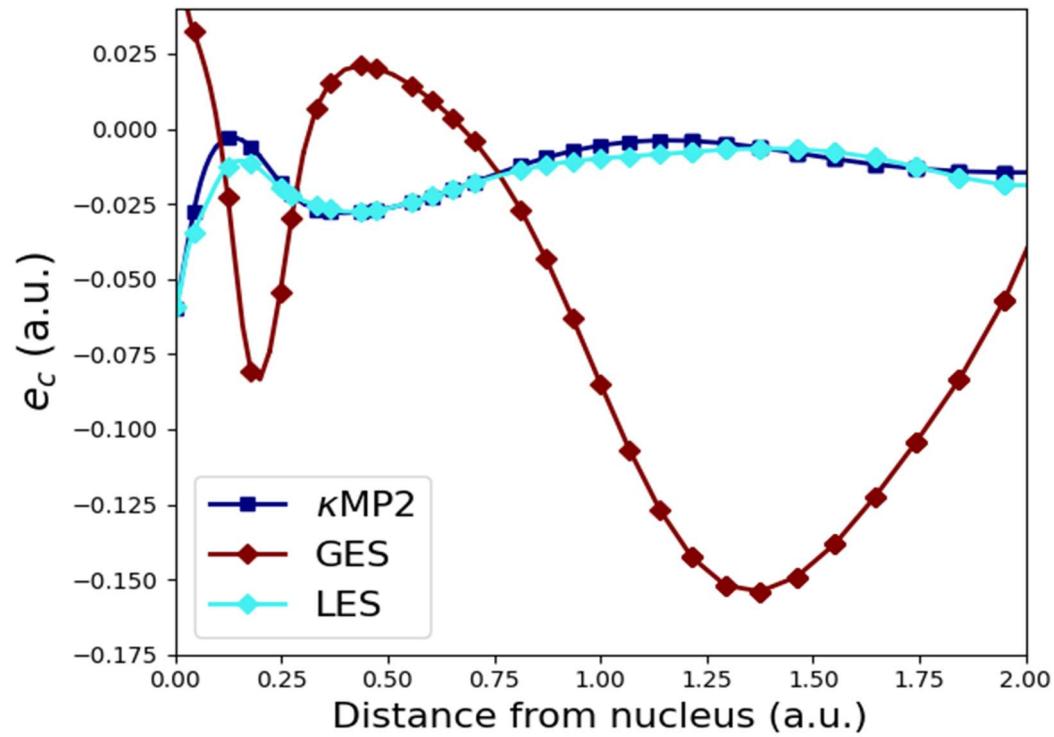
Scaling invariance of MP2:

$$E_C^{MP2}[\rho_\gamma] = E_C^{MP2}[\rho], \quad e_C^{MP2}[\rho_\gamma](\mathbf{r}) = e_C^{MP2}[\rho](\gamma \mathbf{r})$$

- $T_{ijab}[\rho_\gamma] = \gamma^{-1} T_{ijab}[\rho]$  and  $V_{ijab}[\rho_\gamma](\mathbf{r}) = \gamma^4 V_{ijab}[\rho](\gamma \mathbf{r})$
- $e_C^{MP2}[\rho_\gamma](\mathbf{r}) \sim \frac{1}{\rho_\gamma(\mathbf{r})} T_{ijab}[\rho_\gamma] V_{ijab}[\rho_\gamma](\mathbf{r}) = e_C^{MP2}[\rho](\gamma \mathbf{r})$   
 $\Rightarrow e^{ML}(\mathbf{r}) = w^{ML}(\mathbf{r}) e_X(\mathbf{r}) \rho^{-\frac{1}{3}}(\mathbf{r})$



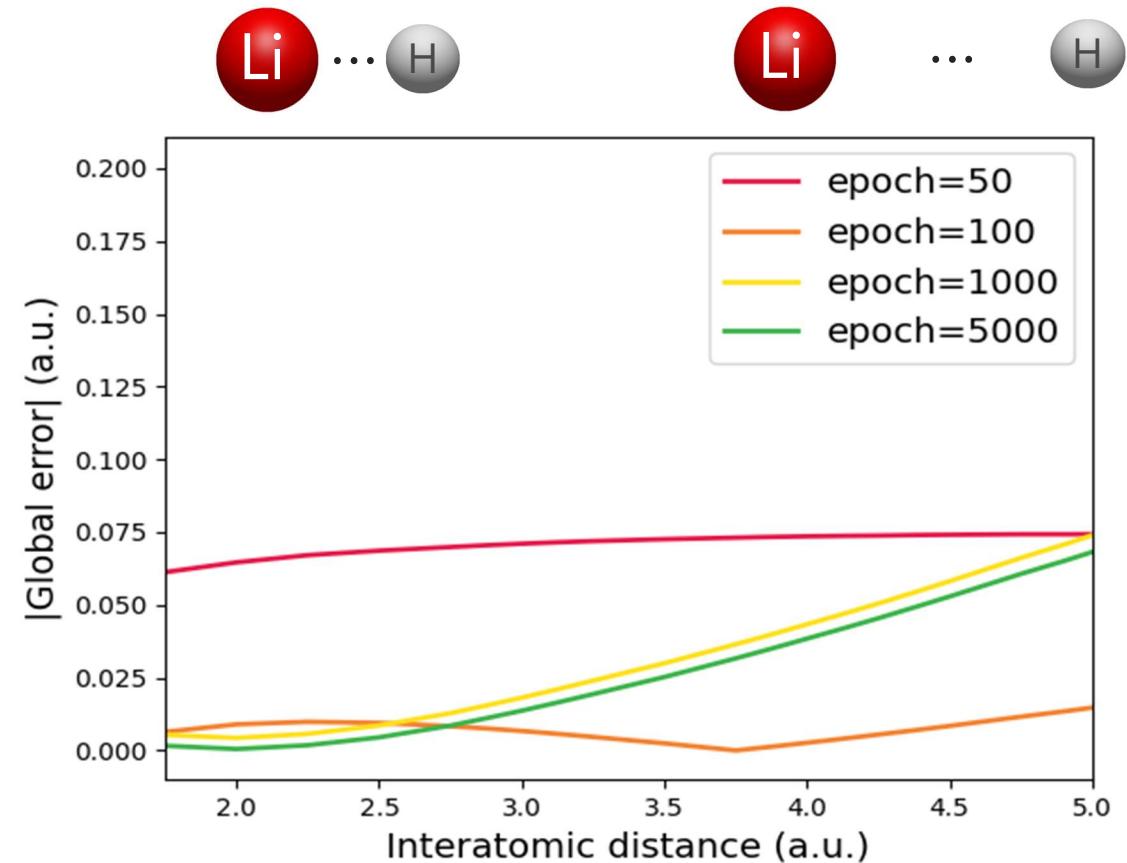
# LES vs. GES



# ML with GES

17

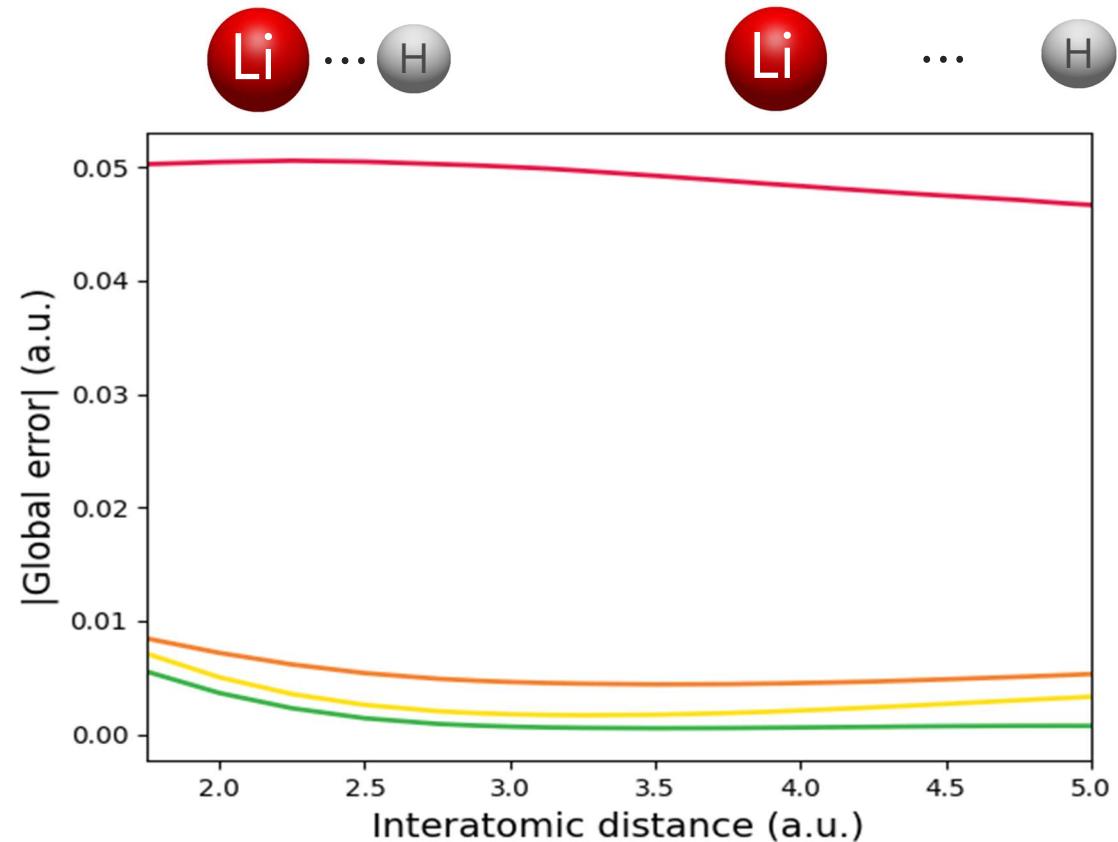
- ❖ Training does not improve ML model prediction
- ❖ Poor transferability even at large amount of epochs



# ML with LES

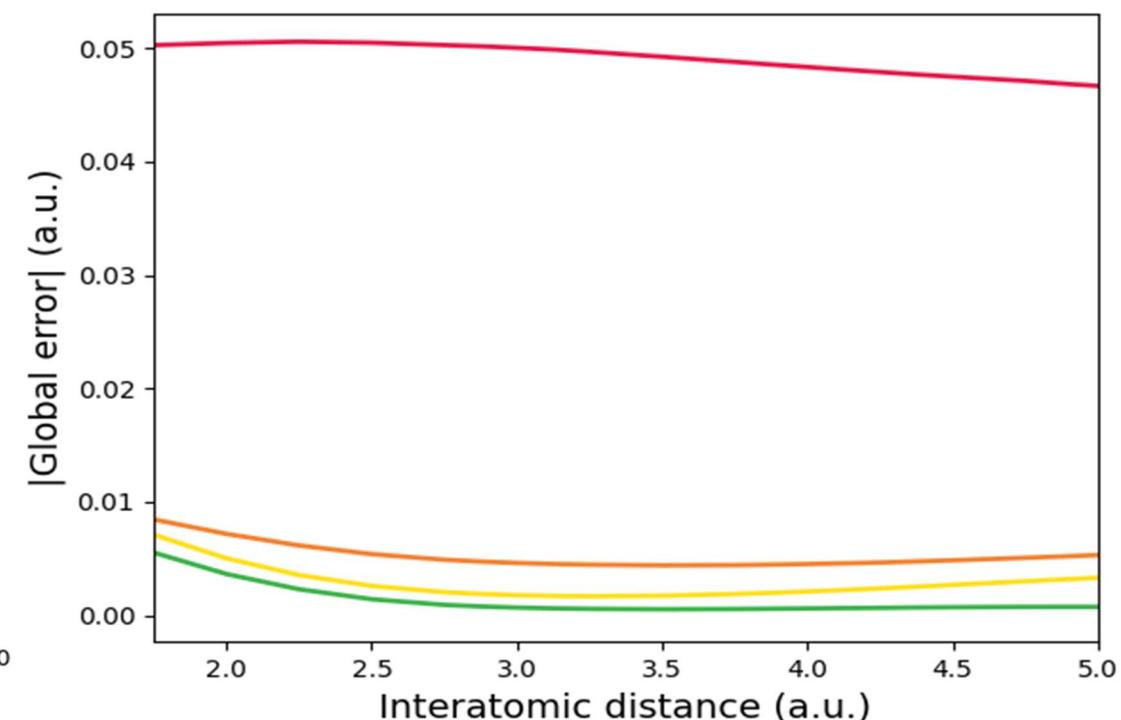
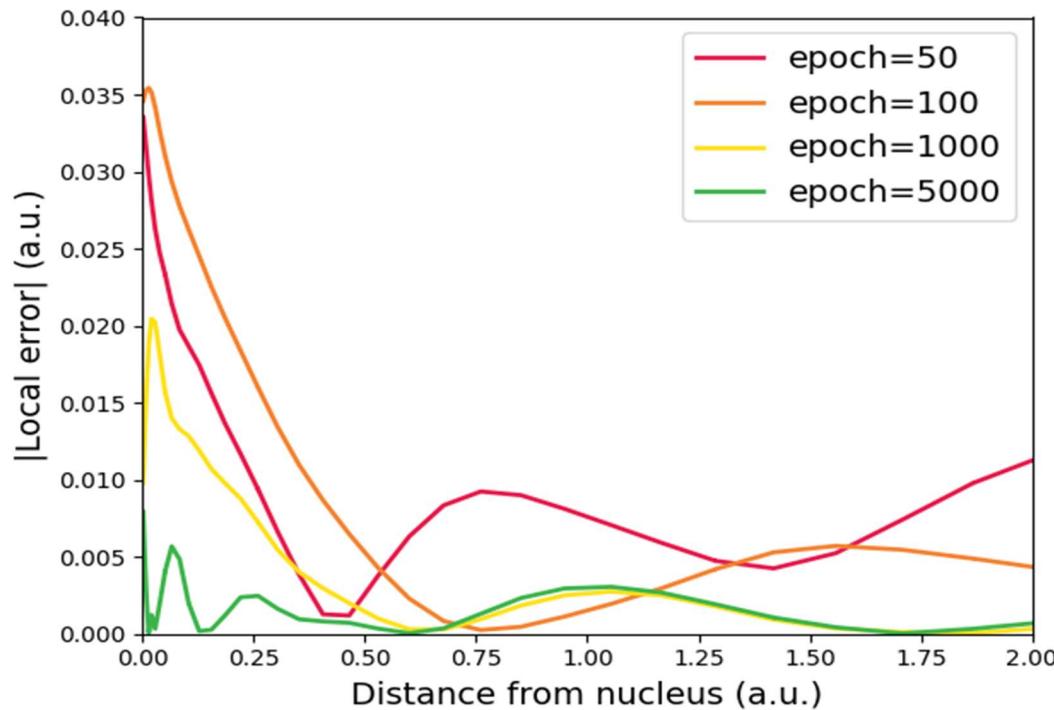
18

- ❖ Training enhances ML based extrapolation
- ❖ Transferability learned epoch by epoch

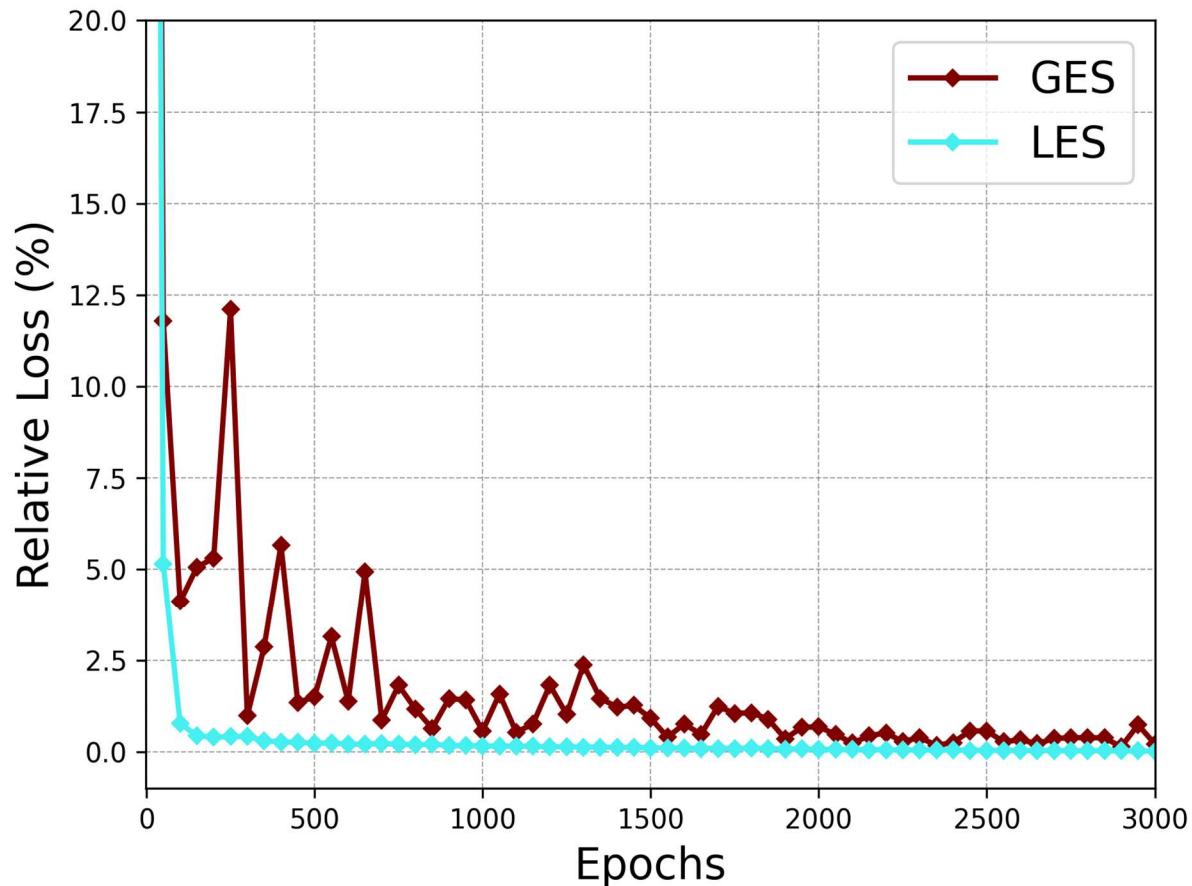


# ML with LES

19



# Learning Convergence



Global Energy Loss (GES):

$$\mathcal{L}_{\text{GES}} \sim |E^{\text{ML}}[\rho] - E^{\text{ref}}[\rho]|$$

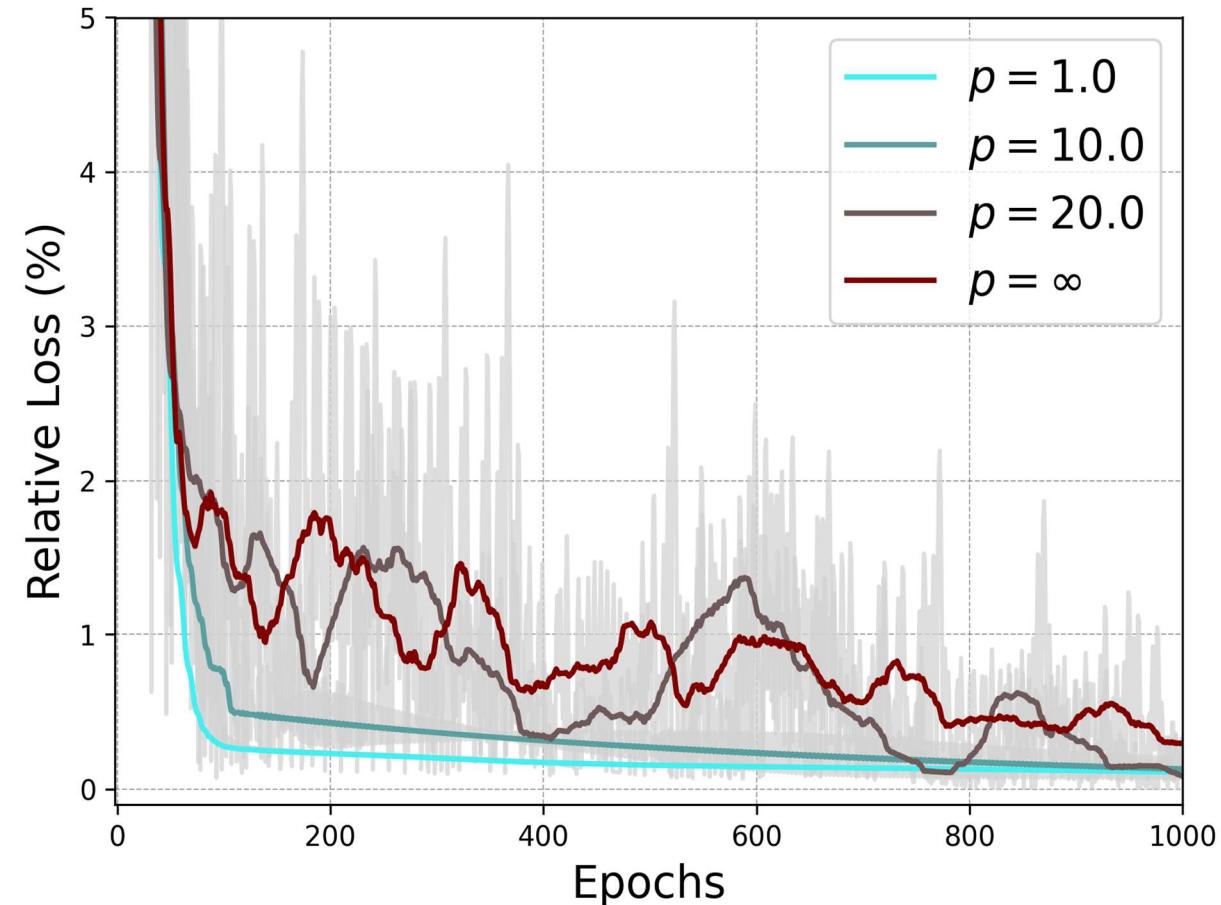
Local Energy Loss (LES):

$$\mathcal{L}_{\text{LES}} \sim \|e^{\text{ML}}(\mathbf{r}) - e^{\text{ref}}(\mathbf{r})\|_{L^p}$$

L →  $\|f\|_{L^p} = \left( \int |f(\mathbf{r})|^p \rho(\mathbf{r}) d\mathbf{r} \right)^{1/p}$



# Learning Convergence



Global Energy Loss (GES):

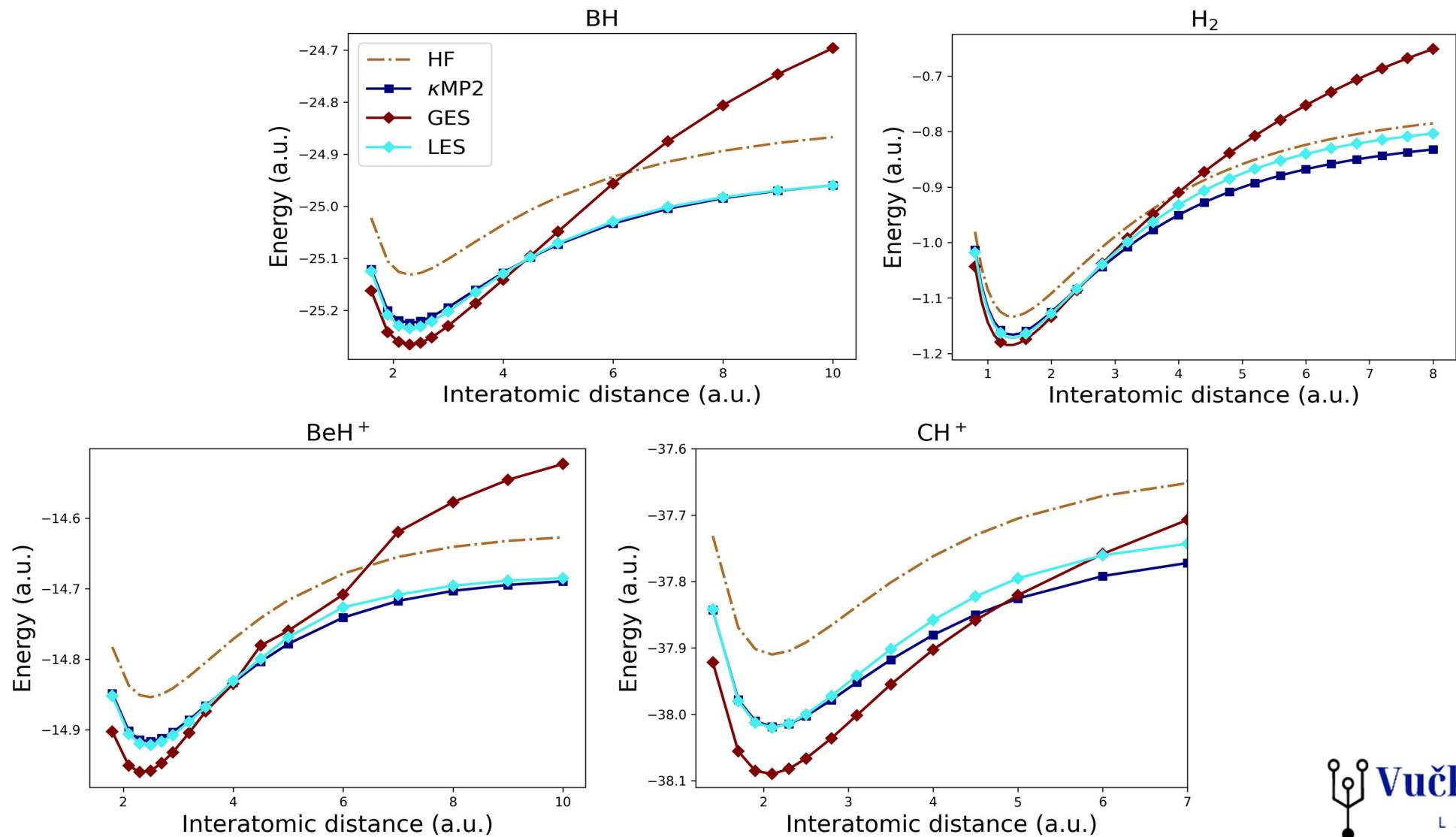
$$\mathcal{L}_{\text{GES}} \sim |E^{\text{ML}}[\rho] - E^{\text{ref}}[\rho]|$$

Local Energy Loss (LES):

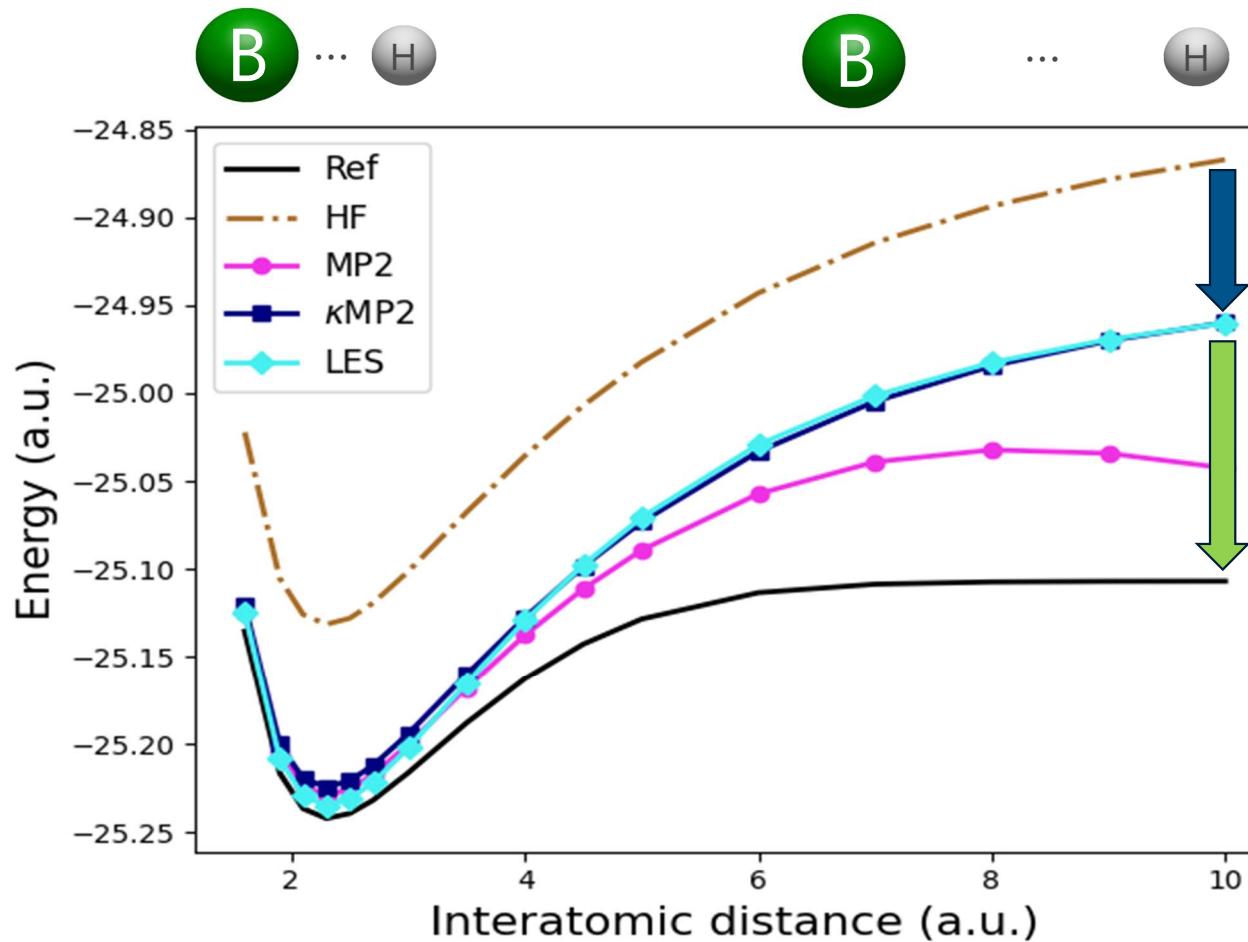
$$\mathcal{L}_{\text{LES}} \sim \|e^{\text{ML}}(\mathbf{r}) - e^{\text{ref}}(\mathbf{r})\|_{L^p}$$

$\Rightarrow \|f\|_{L^p} = \left( \int |f(\mathbf{r})|^p \rho(\mathbf{r}) d\mathbf{r} \right)^{1/p}$





# Going beyond $\kappa$ MP2



- ML with LES
- Employing spin-resolution of  $\kappa$ MP2
- Prediction of reference correlation energies



# Spin-Component-Scaled (SCS) $\kappa$ MP2

$$E_c^{\kappa\text{MP2}}[\rho] = C_{os} E_{c,os}^{\kappa\text{MP2}}[\rho] + C_{ss} E_{c,ss}^{\kappa\text{MP2}}[\rho]$$

Real-space analogue:

$$e_c(\mathbf{r}) = c_{os}(\mathbf{r}) e_{c,os}^{\kappa\text{MP2}}(\mathbf{r}) + c_{ss}(\mathbf{r}) e_{c,ss}^{\kappa\text{MP2}}(\mathbf{r})$$

⇒ Real-space, machine-learned and  
regularized extension of SCS  $\kappa$ MP2 (MLS2)

S. Grimme, *J. Chem. Phys.* 2003, **118**, 9095-9102.

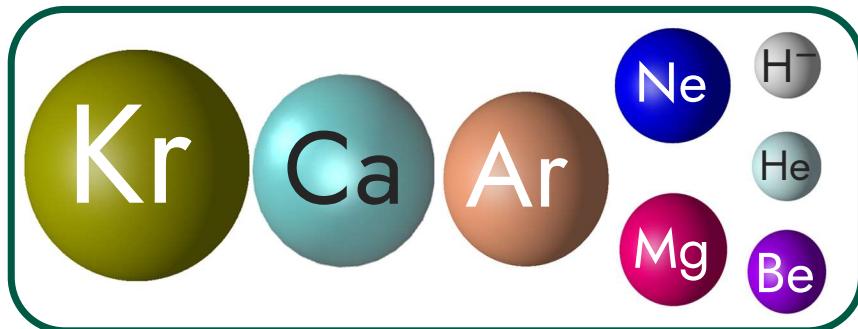
S. Grimme, L. Goerigk and R.F. Fink, *WIREs Comput. Mol. Sci.* 2012, **2**, 886-906.



# ML of real-space SCS $\kappa$ MP2

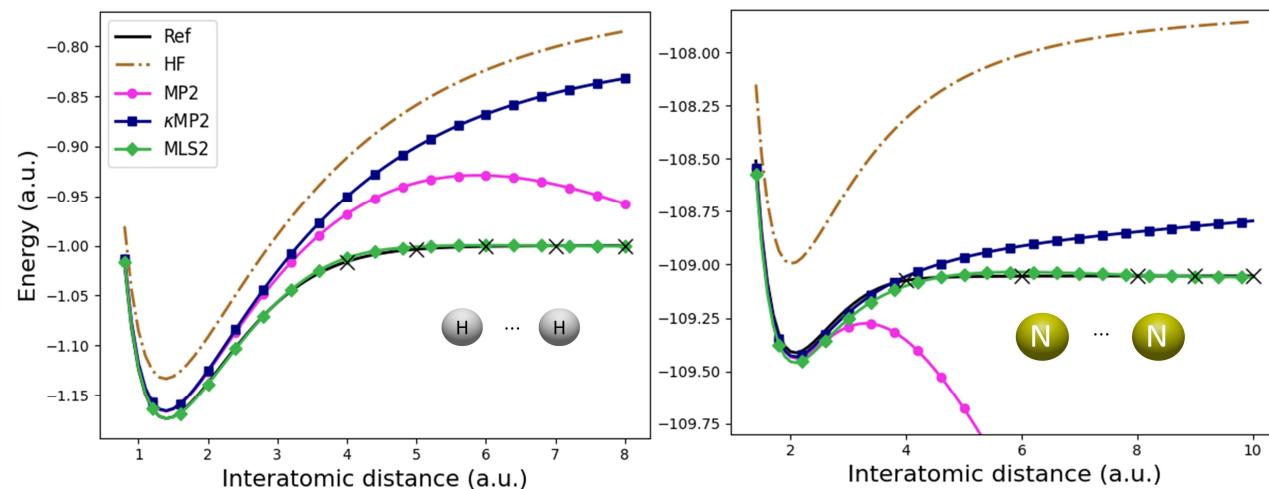
*Training set for MLS2:*

- ❖ 8 atoms/ions



- ❖ 13 additional small complexes (mainly dimers)

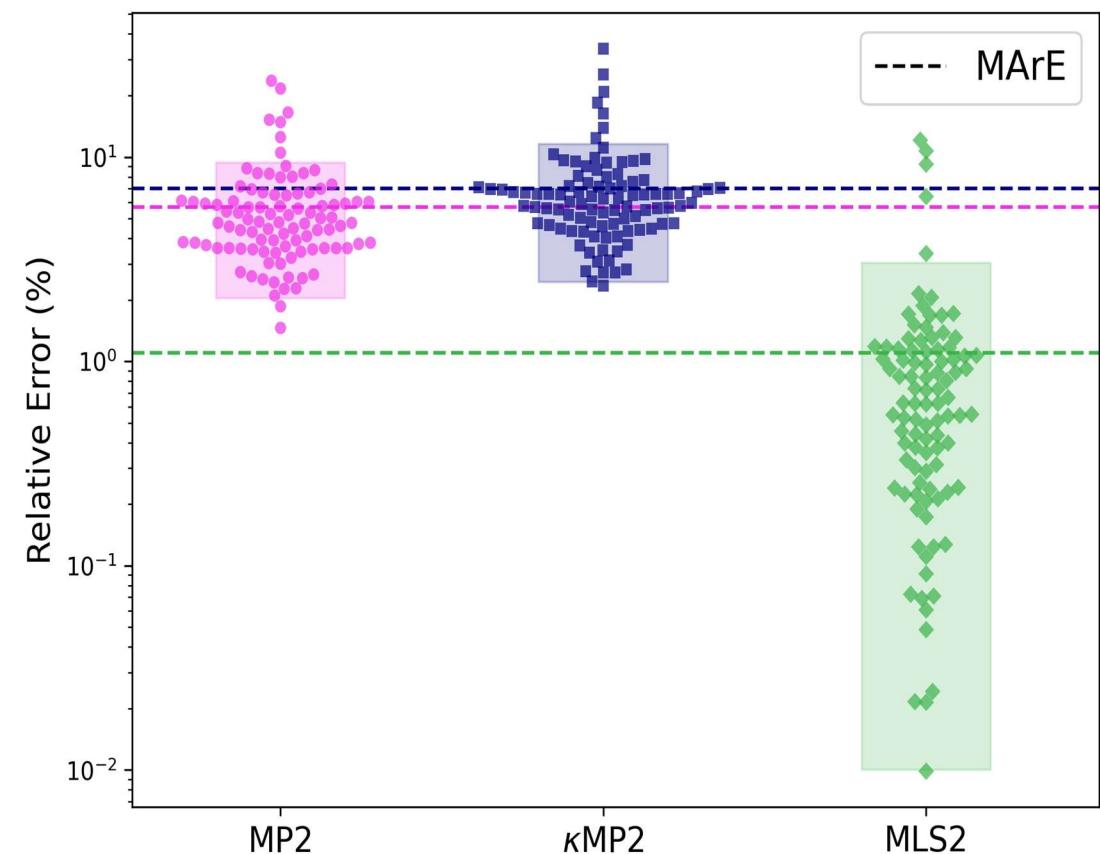
- ❖  $\text{H}_2$ ,  $\text{N}_2$  and  $\text{Li}_2$  at five large interatomic distances



- ❖ Interaction energy data from dispersion-bond dataset RG18

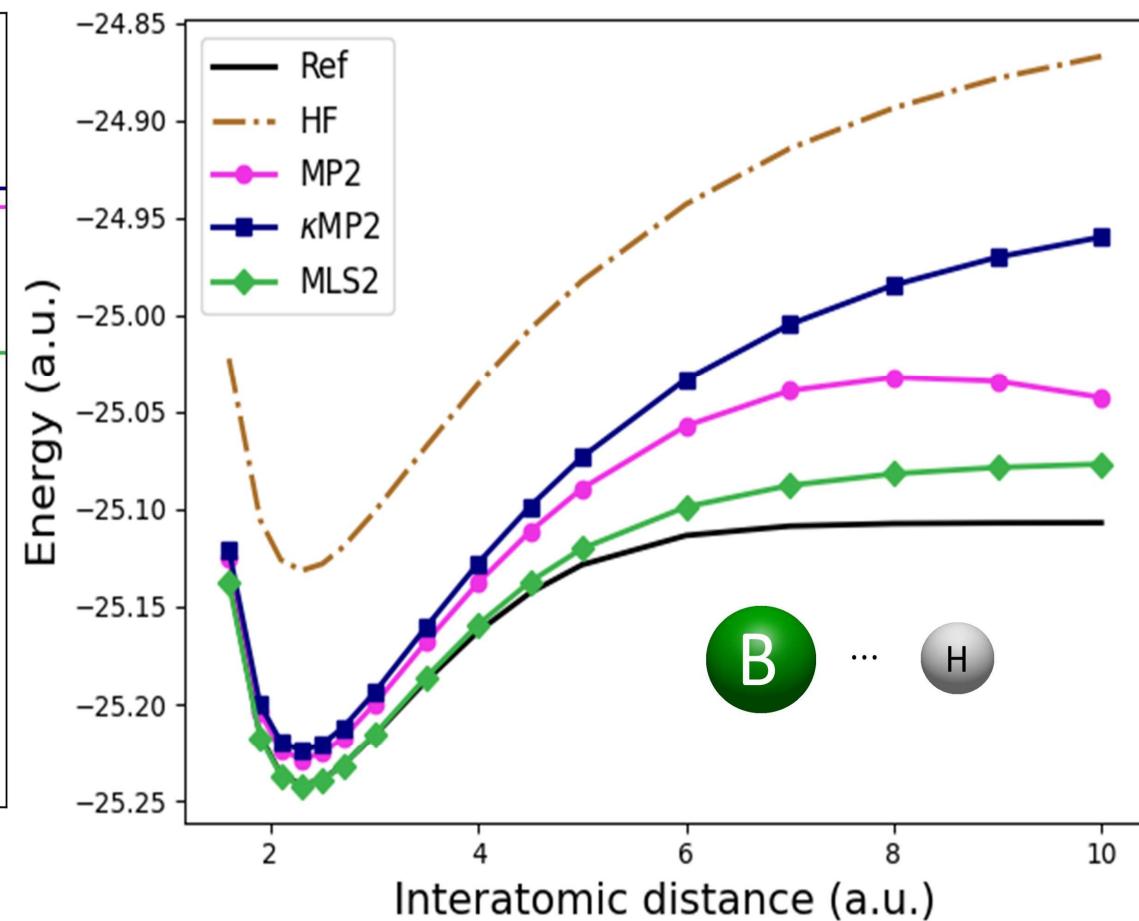


S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.* 2010, **132**, 154104.

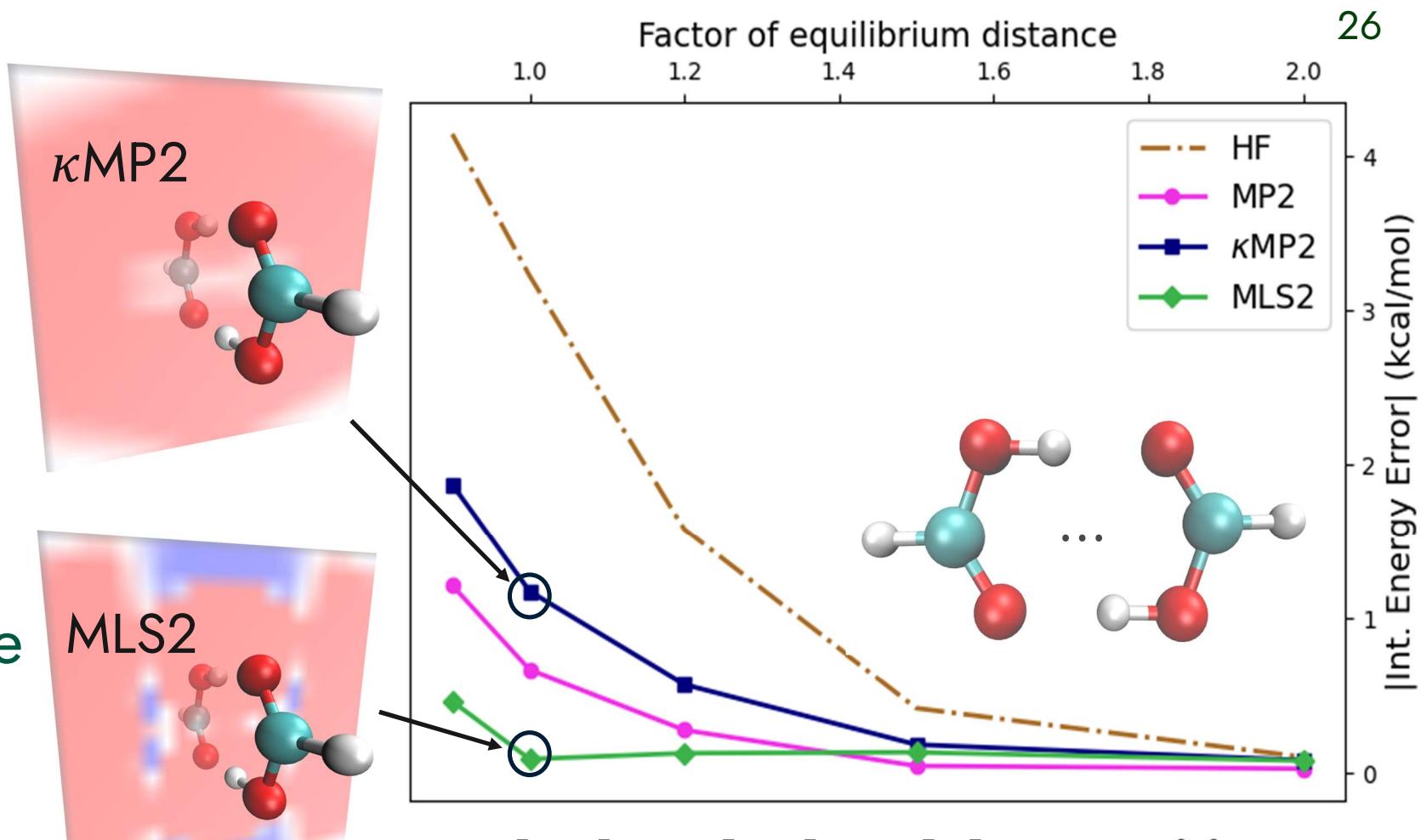


Test on 96 correlation energies  
from the W4-11 database

A. Karton, S. Daon and J.M. Martin, *Chem. Phys. Lett.* 2011, **510**, 165.



Extrapolation to hydrogen-bonded systems  
 ➤ Formic acid dimer example



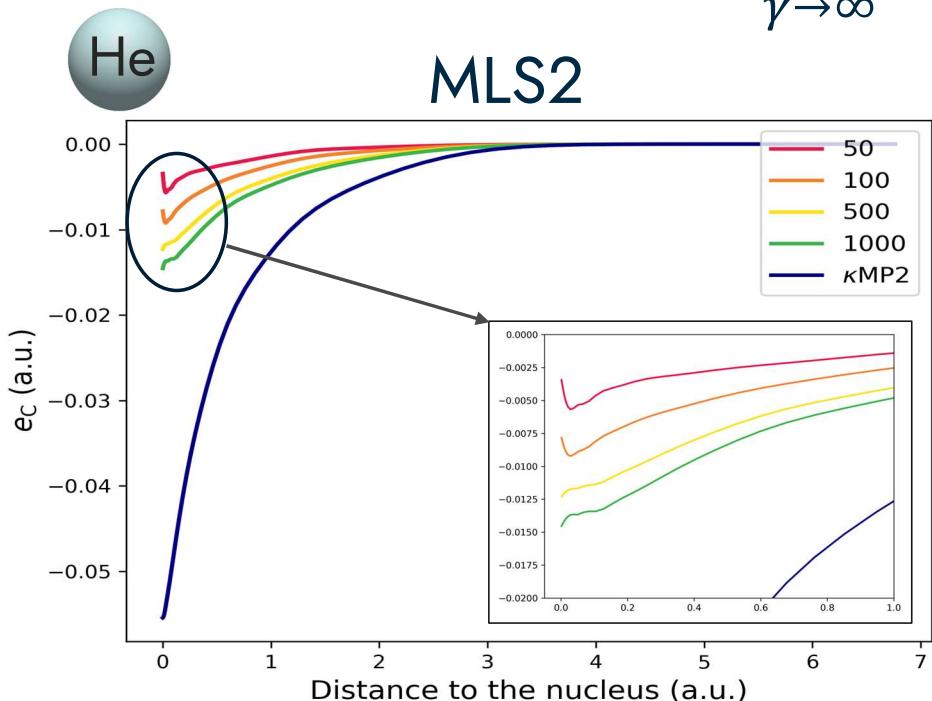
$$E_{int}[AB] = E[AB] - E[A] - E[B]$$

Vučković  
L A B

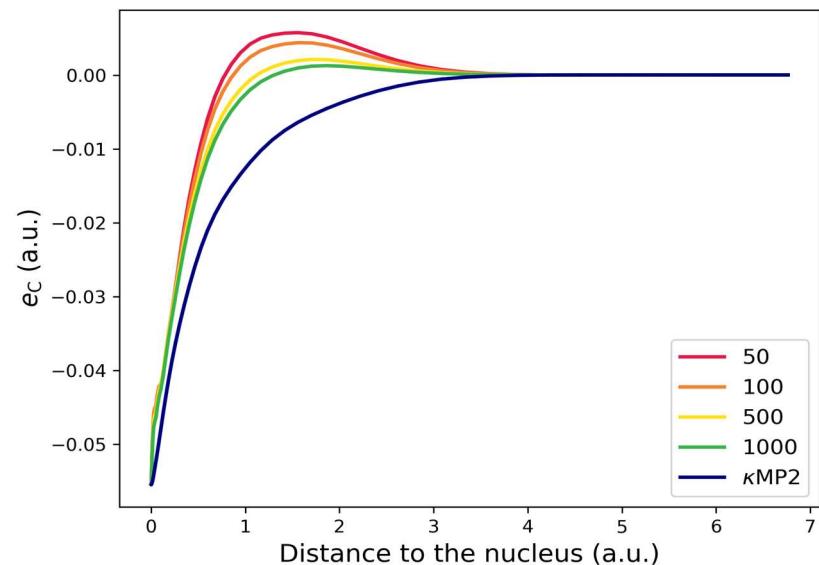
# Constrained MLS2

High-density limit with Görling–Levy (GL) perturbation theory:

$$\lim_{\gamma \rightarrow \infty} E_C[\rho_\gamma] = E_C^{GL2}[\rho]$$



MLS2 with constraint on  $c_{OS}(\mathbf{r})$  and  $c_{SS}(\mathbf{r})$ :  
 $c_{OS}(\mathbf{r}), c_{SS}(\mathbf{r}) \rightarrow 1$  for  $\gamma \rightarrow \infty$ .



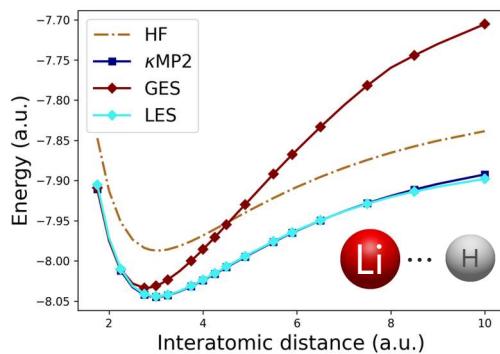
A. Görling & M. Levy, *Phys. Rev. B* 1993, **47**(20), 13105–13113.

A. Görling & M. Levy, *Phys. Rev. A* 1994, **50**(1), 196–204.

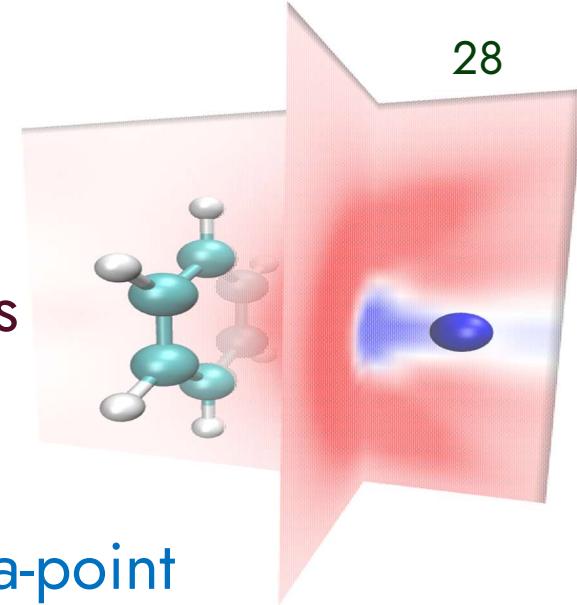
**Vučković**  
L A B

# Conclusions

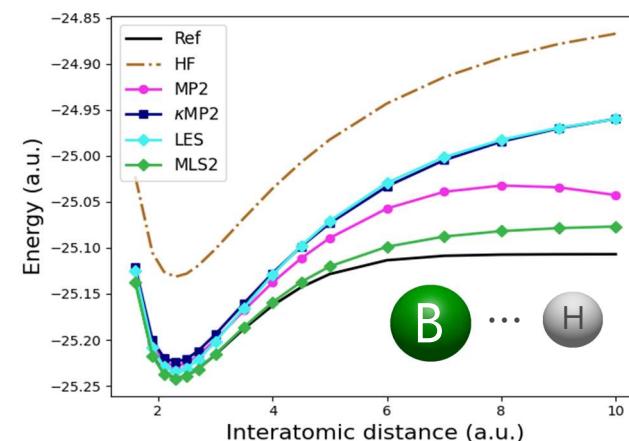
- $\kappa$ MP2 correlation energy density generator
  - ✓ Sound candidate for real-space learning of energies

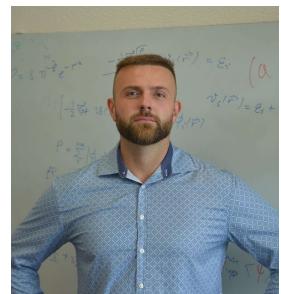


- LES based ML enables transferability
  - ✓ Every point in space becomes a data-point
  - ✓ Enhanced learning convergence



- ML the real-space extension of SCS regularized MP2
  - ✓ Bridging the gap to reference data
  - ✓ Opening up new ML DFAs avenues

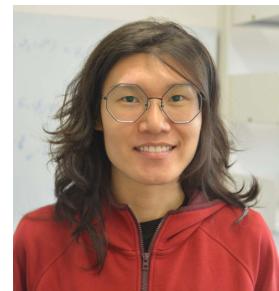




Prof. Stefan  
Vuckovic



Dr. Kimberly J.  
Daas (UC Irvine)

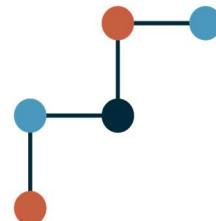


Heng Zhao

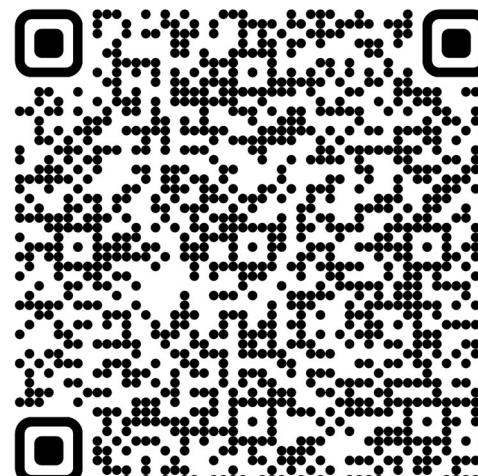


Etienne Palos  
(UC San Diego)

# Acknowledgments



**Swiss National  
Science Foundation**



ChemRxiv



[Elias.Polak@unifr.ch](mailto:Elias.Polak@unifr.ch)



UNIVERSITÉ DE FRIBOURG  
UNIVERSITÄT FREIBURG

**Thank  
you**

