

# Hybrid DFT with numeric atomic orbitals

Monday Meeting, 22 April 2024

Florian Knoop ([florian.knoop@liu.se](mailto:florian.knoop@liu.se))



# DFT in a nutshell

Hohenberg-Kohn [1]

Ground state properties of matter determined by density



Walter Kohn  
(1923-2016)

$$\Psi_{\text{el}}(r_1, r_2, \dots, r_{3N}) \iff \rho(x, y, z)$$

- Reduce degrees of freedom by a factor N ( $\sim 10^{23}$ )
- Electron ground state properties: total energy, forces, dipole moments, ...

[1] P. Hohenberg, W. Kohn, *Phys Rev.* **136**, B864 (1964)

[2] W. Kohn, L.J. Sham, *Phys. Rev.* **140**, A1133 (1965)

# DFT in a nutshell

## Kohn-Sham [2]

Map electron density on effective one-particle orbitals



Walter Kohn  
(1923-2016)

$$\rho(r) \implies \sum_i f_i |\psi_i^{\text{KS}}|^2$$

- can be solved **self-consistently** via **Schrödinger-like equation**
- Kohn-Sham orbitals are **not electrons!**

[1] P. Hohenberg, W. Kohn, *Phys Rev.* **136**, B864 (1964)

[2] W. Kohn, L.J. Sham, *Phys. Rev.* **140**, A1133 (1965)

# DFT in a nutshell

## Kohn-Sham [2]

Map electron density on effective one-particle orbitals

$$\rho(r) \Rightarrow \sum_i f_i |\psi_i^{\text{KS}}|^2$$

KS equations:

$$\left( -\frac{1}{2} \nabla^2 + V^{\text{KS}} \left[ \{\psi_i^{\text{KS}}\} \right] \right) \psi_i^{\text{KS}} = \epsilon_i \psi_i^{\text{KS}}$$

- can be solved **self-consistently** via **Schrödinger-like equation**
- Kohn-Sham orbitals are **not electrons!**

[1] P. Hohenberg, W. Kohn, *Phys Rev.* **136**, B864 (1964)

[2] W. Kohn, L.J. Sham, *Phys. Rev.* **140**, A1133 (1965)

# DFT in a nutshell

## Kohn-Sham [2]

Map electron density on effective one-particle orbitals

KS equations:

$$\left( -\frac{1}{2} \nabla^2 + V^{\text{KS}} \left[ \left\{ \psi_i^{\text{KS}} \right\} \right] \right) \psi_i^{\text{KS}} = \epsilon_i \psi_i^{\text{KS}}$$



$$\rho(r) \Rightarrow \sum_i f_i |\psi_i^{\text{KS}}|^2$$

DFT energy

→ DFA(pproximation)

$$E_{\text{total}} = T_0 + \int \rho V_{\text{nuc}} d^3r + \frac{1}{2} \iint \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{r_{12}} d^3\mathbf{r}_1 d^3\mathbf{r}_2 + E_{\text{XC}}, \quad (3)$$

Schrödinger-like equation

[1] P. Hohenberg, W. Kohn, *Phys Rev.* **136**, B864 (1964)

[2] W. Kohn, L.J. Sham, *Phys. Rev.* **140**, A1133 (1965)

# DFT: Basissets used to represent KS states

Basisset: Complete set of functions to represent KS states

$$\psi_i^{\text{KS}}(\mathbf{r}) = \lim_{K \rightarrow \infty} \sum_{k=1}^K c_k \phi_k(\mathbf{r})$$

## Problem

**Complete basis set** = infinite size ( $K = \infty$ )

## Solution

Find *good enough* basis ( $K < \infty$ )

# Popular basissets

## 1) Plane waves (VASP, Quantum Espresso, ...)

basically a Fourier transform, K corresponds to highest energy:  $|\mathbf{k}| \leq K = E_{\text{cut}}/\hbar$

$$\phi_k(\mathbf{r}) \propto \exp(i\mathbf{k} \cdot \mathbf{r})$$

### Advantages

easy integrals, natural choice for periodic systems, single cutoff

### Disadvantages

huge basis needed, cores need extra treatment (pseudopotentials, PAW, ...)

non-periodic systems

# Popular basissets

## 2) Gaussian type orbitals (Gaussian, ORCA, ...)

$$\phi_{k \equiv lmn; \alpha}(\mathbf{r}) \propto x^l y^m z^n \exp(-\alpha r^2)$$

Advantages 

analytic integrals

local → good for molecules

Disadvantages 

more cutoff parameters

# Popular basissets

## 3) Numeric atom-centered orbitals (FHI-aims, DFTB+, ...)

$$\phi_{k \equiv ilm}(\mathbf{r}) \propto \frac{u_i(r)}{r} Y_{lm}(\varphi, \vartheta)$$

### Advantages

**local** → good for molecules, **treat molecules and solids on same footing**

small amount of basis functions needed → **linear scaling**

**all electrons** can be treated with little extra cost (no pseudos or PAW)

### Disadvantages

movable grids → complicated Pulay terms

more difficult to handle analytically than plane waves or Gaussians

# Popular basissets

## 3) Numeric atomic orbitals (NBO, TURBOMOLE, TB+, ...)

Question:

### Advantages

local

small and sparse

all electrons

Answer:  
**Yes.**

### Disadvantages

movable grid

complicated Pulay terms

more difficult to handle analytically than plane waves or Gaussians

# NAOs in FHI-aims

## What should we know?

**light**

Get to know a system, geometry optimization, molecular dynamics, ...

**intermediate**

tighter grids, often converged

**tight**

Converged geometries and energies for checking

**really\_tight**

Overcomplete Benchmark basisset for double checking

# NAOs in FHI-aims

## What should we know?

**light**

Get to know a system, geomet

**intermediate**

tighter grids, often converged

**tight**

Converged geometries and ener

**really\_tight**

Overcomplete Benchmark basis



# NAOs in FHI-aims

## Example: Silicon

$$\phi_{k\equivilm}(\mathbf{r}) \propto \frac{u_i(r)}{r} Y_{lm}(\varphi, \vartheta)$$

Si, fcc diamond	light	intermediate	tight	really_tight
$r_{\text{cut}}$ [Å]	5	6	6	6
# Integration points	11,208	35,840	35,840	56,860
Multipole $l_{\max}$	4	6	6	8
Max. # basis functions	36	68	78	78
Error [meV per atom]	28.1	11.6	0.6	0.0
Time per scf step <sup>*</sup> [s]	0.4	3.4	4.7	9.8

# OK fine – why bother?

## Efficient All-electron Hybrid Density Functionals for Atomistic Simulations Beyond 10,000 Atoms

Sebastian Kokott,<sup>1</sup> Florian Merz,<sup>2</sup> Yi Yao,<sup>3</sup> Christian Carbogno,<sup>1</sup> Mariana Rossi,<sup>4</sup> Ville Havu,<sup>5</sup> Markus Rampp,<sup>6</sup> Matthias Scheffler,<sup>1</sup> and Volker Blum<sup>3,7</sup>

<sup>1)</sup> The NOMAD Laboratory at the Fritz Haber Institute of the Max-Planck-Gesellschaft and IRIS Adlershof of the Humboldt-Universität zu Berlin, Germany

<sup>2)</sup> Lenovo HPC Innovation Center, Stuttgart, Germany

<sup>3)</sup> Thomas Lord Department of Mechanical Engineering and Material Science, Duke University, Durham, North Carolina 27708, USA

<sup>4)</sup> MPI for the Structure and Dynamics of Matter, Luruper Chaussee 149, 22761 Hamburg, Germany

<sup>5)</sup> Department of Applied Physics, School of Science, Aalto University, Espoo, Finland

<sup>6)</sup> Max Planck Computing and Data Facility, 85748 Garching, Germany

<sup>7)</sup> Department of Chemistry, Duke University, Durham, North Carolina 27708, USA

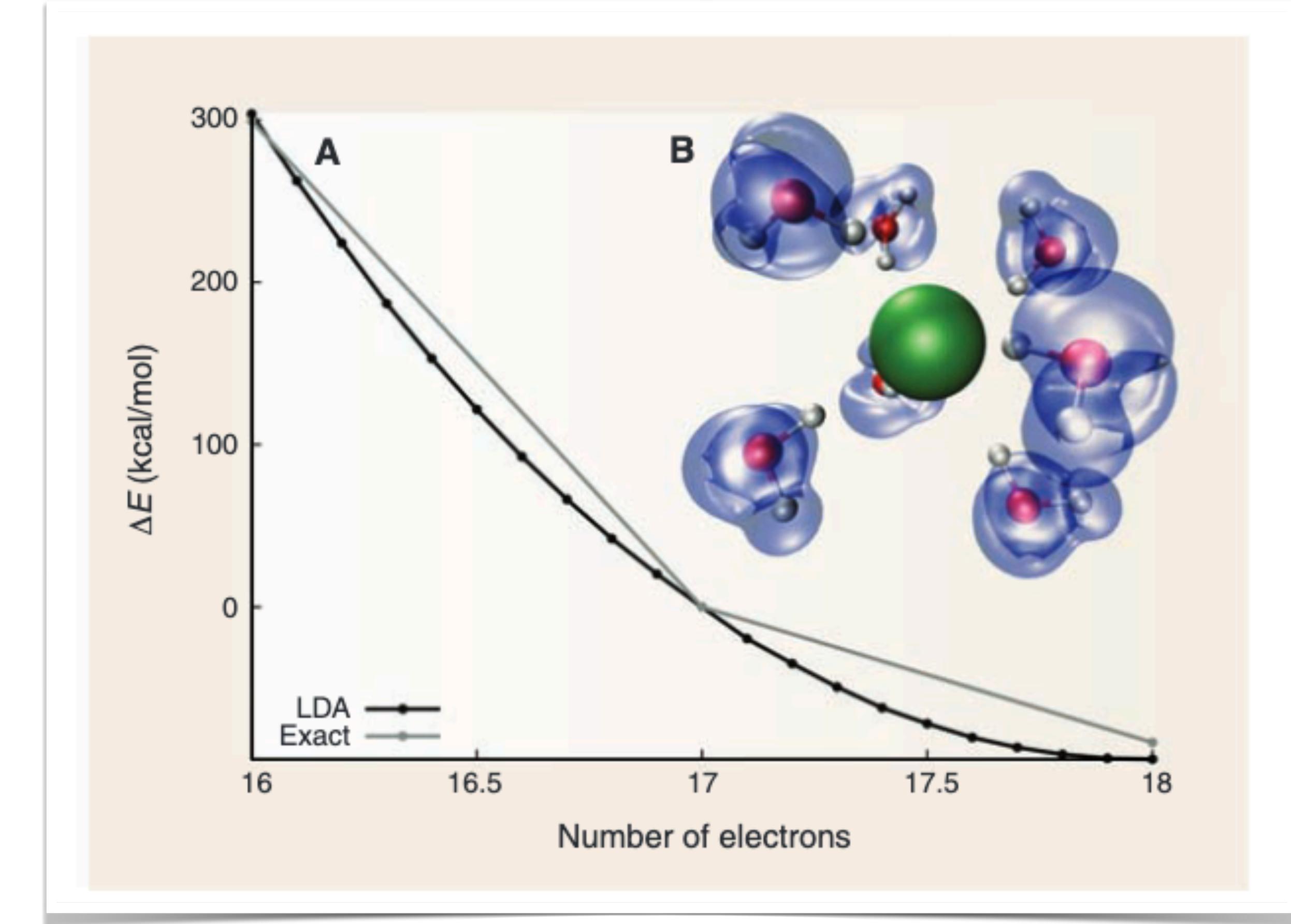
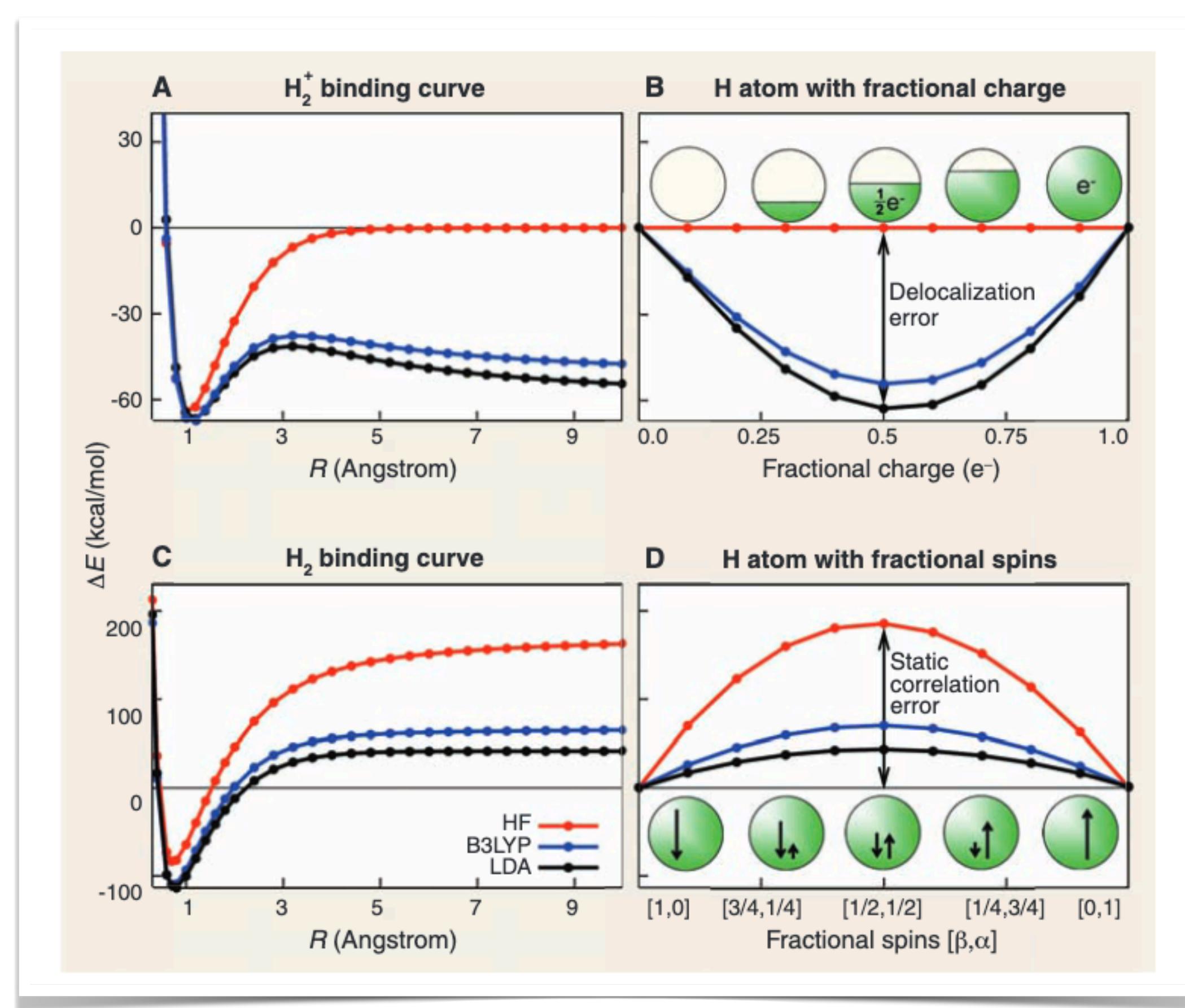
(Dated: 18 March 2024)

Hybrid density functional approximations (DFAs) offer compelling accuracy for *ab initio* electronic-structure simulations of molecules, nanosystems, and bulk materials, addressing some deficiencies of computationally cheaper, frequently used semilocal DFAs. However, the computational bottleneck of hybrid DFAs is the evaluation of the non-local exact exchange contribution, which is the limiting factor for the application of the method for large-scale simulations. In this work, we present a drastically optimized resolution-of-identity-based real-space implementation of the exact exchange evaluation for both non-periodic and periodic boundary conditions in the all-electron code FHI-aims, targeting high-performance CPU compute clusters. The introduction of several new refined Message Passing Interface (MPI) parallelization layers and shared memory arrays according to the MPI-3 standard were the key components of the optimization. We demonstrate significant improvements of memory and performance efficiency, scalability, and workload distribution, extending the reach of hybrid DFAs to simulation sizes beyond ten thousand atoms. As a necessary byproduct of this work, other code parts in FHI-aims have been optimized as well, e.g., the computation of the Hartree potential and the evaluation of the force and stress components. We benchmark the performance and scaling of the hybrid DFA based simulations for a broad range of chemical systems, including hybrid organic-inorganic perovskites, organic crystals and ice crystals with up to 30,576 atoms (101,920 electrons described by 244,608 basis functions).

# **Interlude: Hybrid functionals**

# Why hybrid functionals?

(semi-local) DFT can fail dramatically, e.g., **delocalization error [1]**



# Why hybrid functionals?

exchange can be computed exactly → **exact exchange (EXX or HF) [1]**

$$\begin{aligned} E_{\text{total}} = & T_0 + \int \rho V_{\text{nuc}} d^3 \mathbf{r} \\ & + \frac{1}{2} \iint \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{r_{12}} d^3 \mathbf{r}_1 d^3 \mathbf{r}_2 + E_{\text{XC}}, \end{aligned} \quad (3)$$

Naïve:

$$E_{\text{XC}} = E_X + E_C^{\text{LSDA}}, \quad (8)$$

Becke: Use adiabatic connection from noninteracting ( $\lambda=0$ ) to interacting ( $\lambda=1$ )

$$E_{\text{XC}} = \int_0^1 U_{\text{XC}}^\lambda d\lambda, \quad (9)$$

some approximations

$$E_{\text{XC}} \approx \frac{1}{2} E_X + \frac{1}{2} U_{\text{XC}}^{\text{LSDA}}, \quad (12)$$

# How to get a good hybrid?

About 25% exact exchange is good [1]

$$E_{\text{xc}}^{\text{hyb}} = E_{\text{xc}}^{\text{DFA}} + a_0(E_{\text{x}} - E_{\text{x}}^{\text{DFA}}), \quad (5)$$

Perdew, Ernzerhof, Burke (**PBE0**) [1]:

“We will show that

**$a_0 \approx 0.25$  is to be expected for the atomization energies of most molecules,**

but also that

**larger values of  $a_0$  may be more appropriate for the total energies of atoms and molecules,**  
and

**smaller values for atomization energies of molecules with nearly degenerate ground states**  
for the unperturbed ( $\lambda = 0$ ) problem.”

# Problems with exact exchange (EXX, HF)

$$E_x = \sum_{\mu\nu,\lambda\sigma} c_{\mu\nu,\lambda\sigma} \langle \mu\nu | \lambda\sigma \rangle$$

$$\langle \mu\nu | \lambda\sigma \rangle = \iint dr_1 dr_2 \psi_\mu(\mathbf{r}_1) \psi_\nu(\mathbf{r}_1) \frac{1}{r_{12}} \psi_\lambda(\mathbf{r}_2) \psi_\sigma(\mathbf{r}_2)$$

two-electron Coulomb integrals

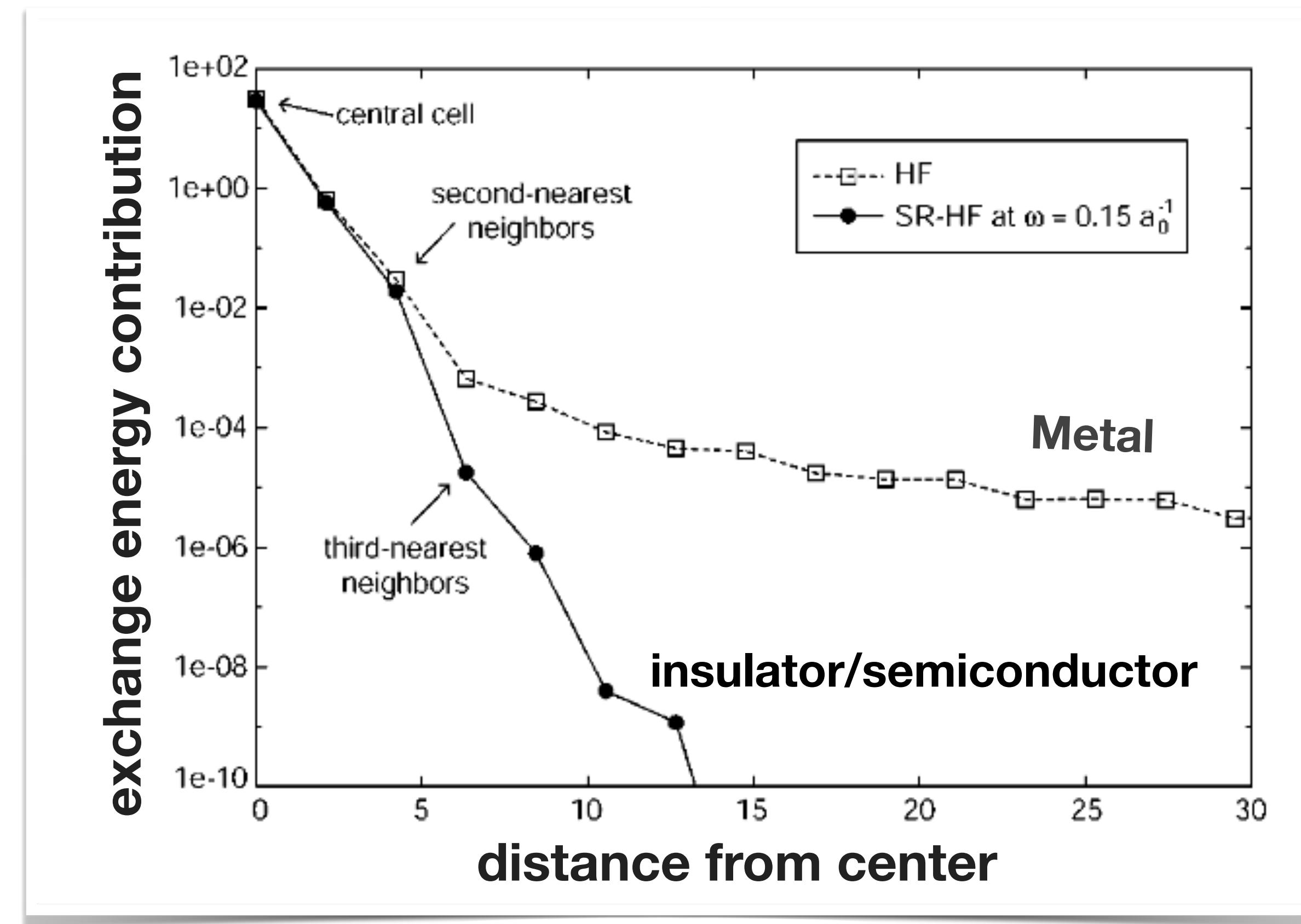
Coulomb interaction

**Problem 1:** Coulomb interaction is long-ranged (= expensive)

**Problem 2:** (unscreened) Coulomb interactions **diverges for metals**

# Solve problem 1: Nearsightedness

Range of exchange decays exponentially for systems with gap [2, 3]



- [1] J. Heyd, G. E. Scuseria, and M. Ernzerhof, J Chem Phys **118**, 8207 (2003).
- [2] E. Prodan and W. Kohn, Proc. Natl. Acad. Sci. **102**, 11635 (2005).
- [3] W. Kohn, Int. J. Quantum Chem. **56**, 229 (1995).

# Introduce range separation in HF computation

→ HSE06 [1, 2]

$$\frac{1}{r} = \frac{\text{erfc}(\omega r)}{r} + \frac{\text{erf}(\omega r)}{r}$$

**short-range part**      ↓      **neglect long-range part**

$$\langle \mu \nu | \lambda \sigma \rangle^{\text{SR}} = \iint dr_1 dr_2 \psi_\mu(\mathbf{r}_1) \psi_\nu(\mathbf{r}_1) \frac{\text{erfc}(\omega r_{12})}{r_{12}} \psi_\lambda(\mathbf{r}_2) \psi_\sigma(\mathbf{r}_2)$$
$$\langle \mu \nu | \lambda \sigma \rangle^{\text{LR}} \approx 0$$

$$E_{\text{xc}}^{\text{HSE}} = a E_x^{\text{HF}, \text{SR}}(\omega) + (1 - a) E_x^{\text{PBE}, \text{SR}}(\omega) + E_x^{\text{PBE}, \text{LR}}(\omega) + E_c^{\text{PBE}}$$

- [1] J. Heyd, G. E. Scuseria, and M. Ernzerhof, *J Chem Phys* **118**, 8207 (2003).  
[2] J. Heyd, G. E. Scuseria, and M. Ernzerhof, *J Chem Phys* **124**, 219906 (2006).

# Introduce range separation

→ HSE06 [1]

$$\langle \mu\nu | \lambda\sigma \rangle^{\text{SR}}$$

Bottom line 1:

**Hybrid functionals [1, 2] solve *some* errors  
of (semi-)local DFAs.**

Bottom line 2:

**Range-separated hybrid functionals  
can do the same *but faster*.**

[1] S. Kümmel and L. Kronik, Rev Mod Phys **80**, 3 (2008)

[2] A. M. Teale *et al.*, Phys Chem Chem Phys (2022)

$$\langle \omega \rangle + E_c^{\text{PBE}}$$

- [1] J. Heyd, G. E. Scuseria, and M. Ernzerhof, J Chem Phys **118**, 8207 (2003).  
[2] J. Heyd, G. E. Scuseria, and M. Ernzerhof, J Chem Phys **124**, 219906 (2006).

# Back to FHI-aims

## Efficient All-electron Hybrid Density Functionals for Atomistic Simulations Beyond 10,000 Atoms

Sebastian Kokott,<sup>1</sup> Florian Merz,<sup>2</sup> Yi Yao,<sup>3</sup> Christian Carbogno,<sup>1</sup> Mariana Rossi,<sup>4</sup> Ville Havu,<sup>5</sup> Markus Rampp,<sup>6</sup> Matthias Scheffler,<sup>1</sup> and Volker Blum<sup>3,7</sup>

<sup>1)</sup> The NOMAD Laboratory at the Fritz Haber Institute of the Max-Planck-Gesellschaft and IRIS Adlershof of the Humboldt-Universität zu Berlin, Germany

<sup>2)</sup> Lenovo HPC Innovation Center, Stuttgart, Germany

<sup>3)</sup> Thomas Lord Department of Mechanical Engineering and Material Science, Duke University, Durham, North Carolina 27708, USA

<sup>4)</sup> MPI for the Structure and Dynamics of Matter, Luruper Chaussee 149, 22761 Hamburg, Germany

<sup>5)</sup> Department of Applied Physics, School of Science, Aalto University, Espoo, Finland

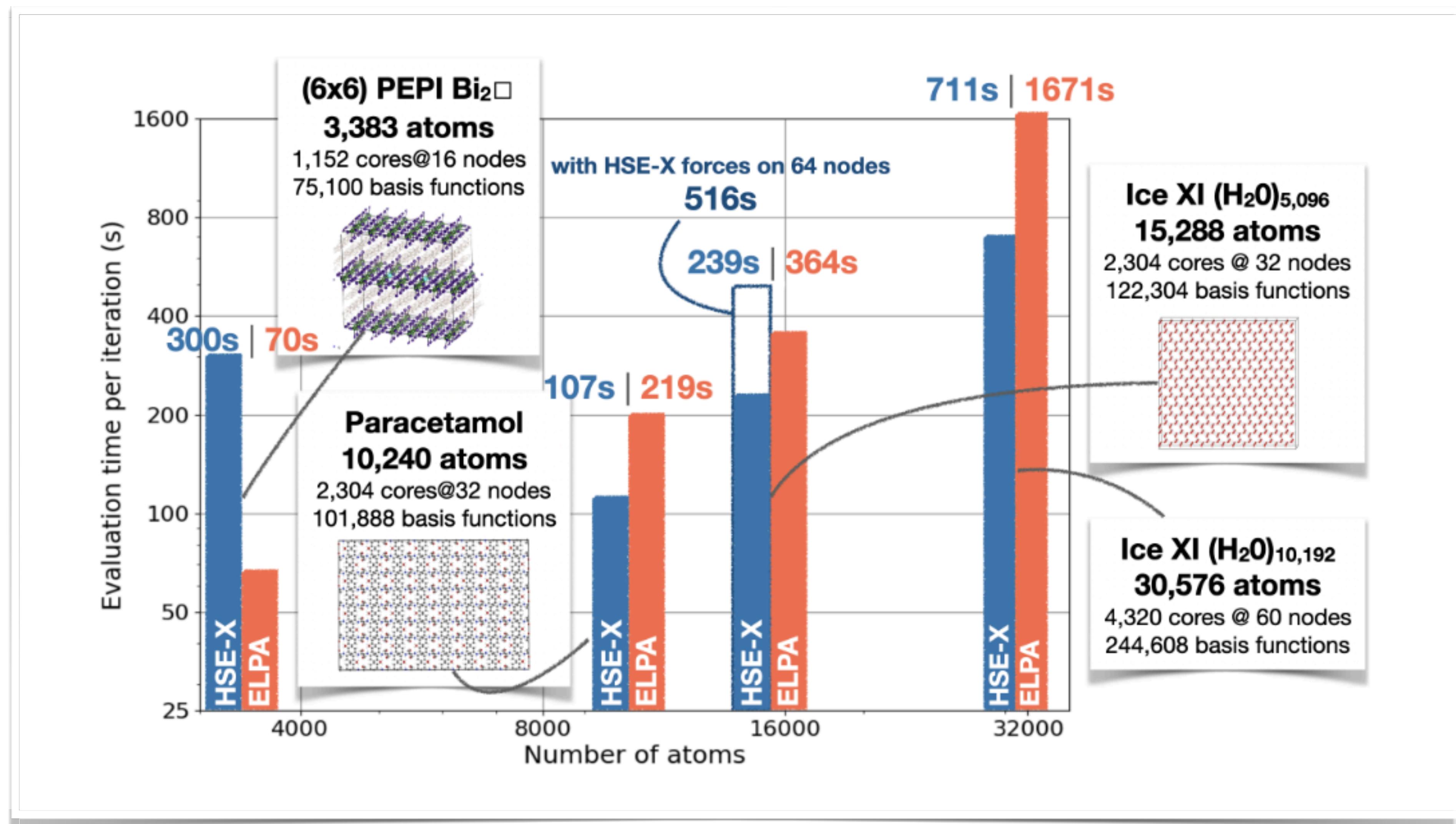
<sup>6)</sup> Max Planck Computing and Data Facility, 85748 Garching, Germany

<sup>7)</sup> Department of Chemistry, Duke University, Durham, North Carolina 27708, USA

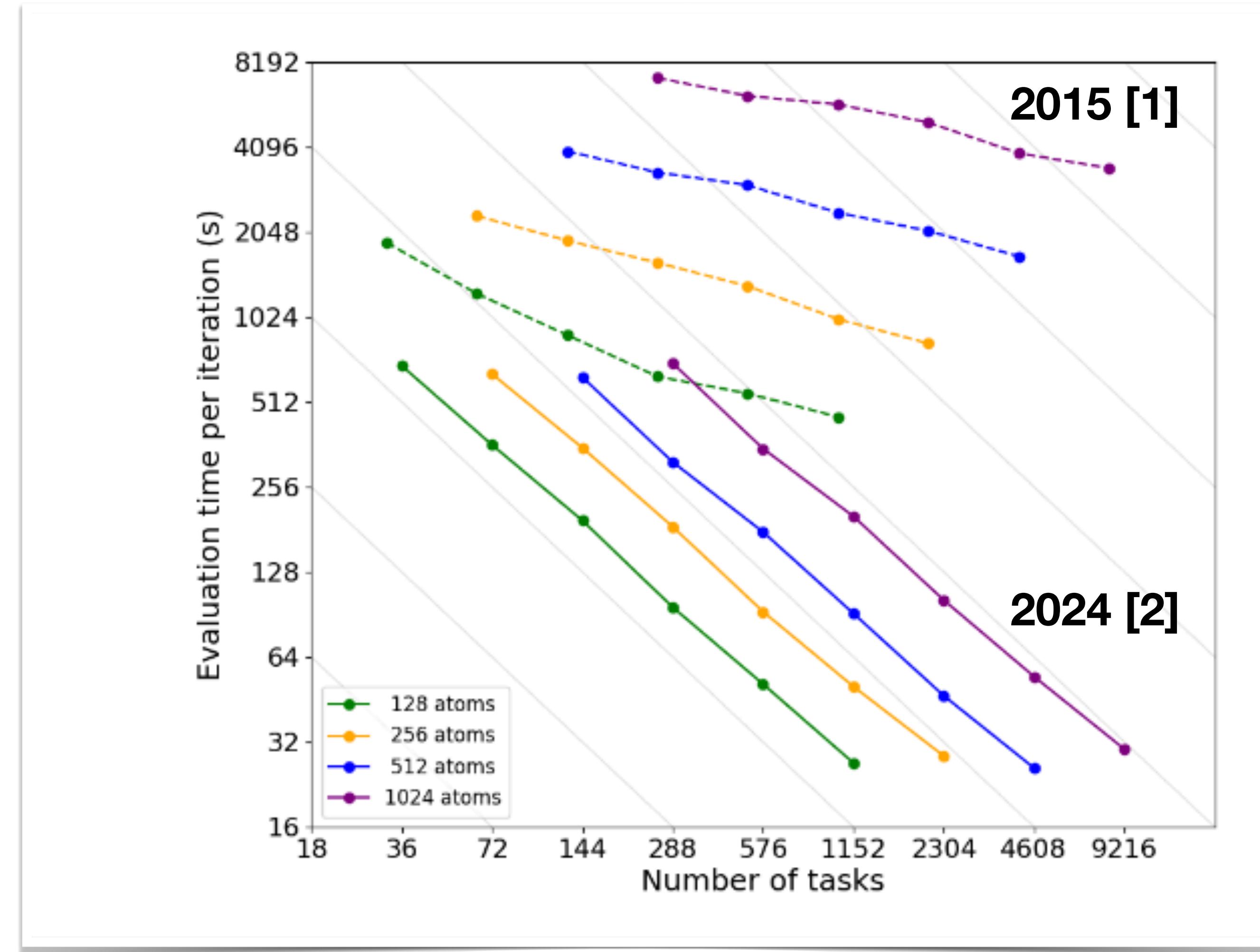
(Dated: 18 March 2024)

Hybrid density functional approximations (DFAs) offer compelling accuracy for *ab initio* electronic-structure simulations of molecules, nanosystems, and bulk materials, addressing some deficiencies of computationally cheaper, frequently used semilocal DFAs. However, the computational bottleneck of hybrid DFAs is the evaluation of the non-local exact exchange contribution, which is the limiting factor for the application of the method for large-scale simulations. In this work, we present a drastically optimized resolution-of-identity-based real-space implementation of the exact exchange evaluation for both non-periodic and periodic boundary conditions in the all-electron code FHI-aims, targeting high-performance CPU compute clusters. The introduction of several new refined Message Passing Interface (MPI) parallelization layers and shared memory arrays according to the MPI-3 standard were the key components of the optimization. We demonstrate significant improvements of memory and performance efficiency, scalability, and workload distribution, extending the reach of hybrid DFAs to simulation sizes beyond ten thousand atoms. As a necessary byproduct of this work, other code parts in FHI-aims have been optimized as well, e.g., the computation of the Hartree potential and the evaluation of the force and stress components. We benchmark the performance and scaling of the hybrid DFA based simulations for a broad range of chemical systems, including hybrid organic-inorganic perovskites, organic crystals and ice crystals with up to 30,576 atoms (101,920 electrons described by 244,608 basis functions).

# HSE06 for many atoms



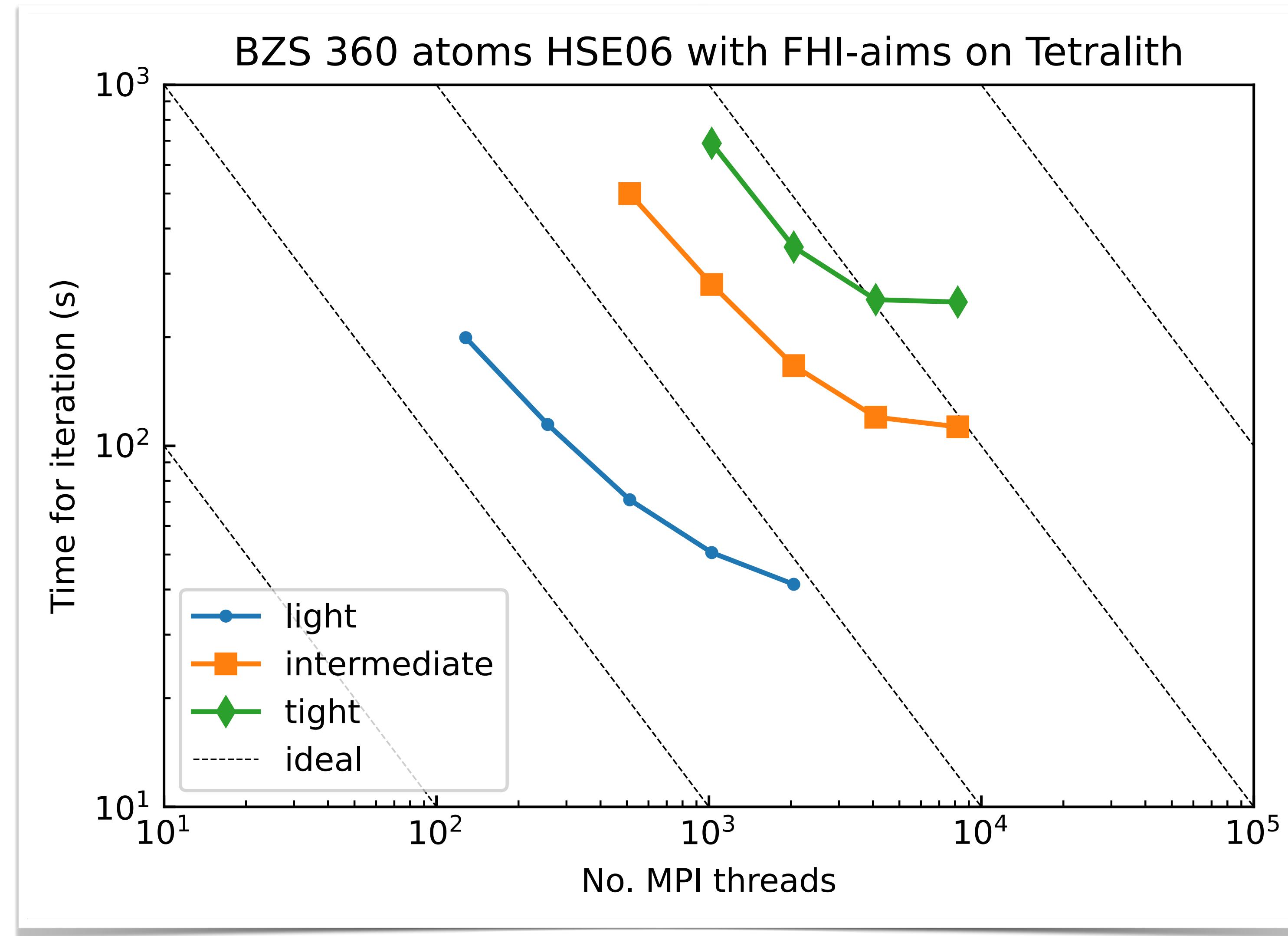
# Strong scaling



[1] S. V. Levchenko *et al.*, Comput Phys Commun **192**, 60 (2015)

[2] S. Kokott *et al.*, arXiv 2403.10343 (2024)

# Strong scaling on NSC's tetralith



# Strong scaling on NSC's tetralith

Dardel?

**Unfortunately no. Cray MPI libraries are too buggy :().**

PDC is aware.

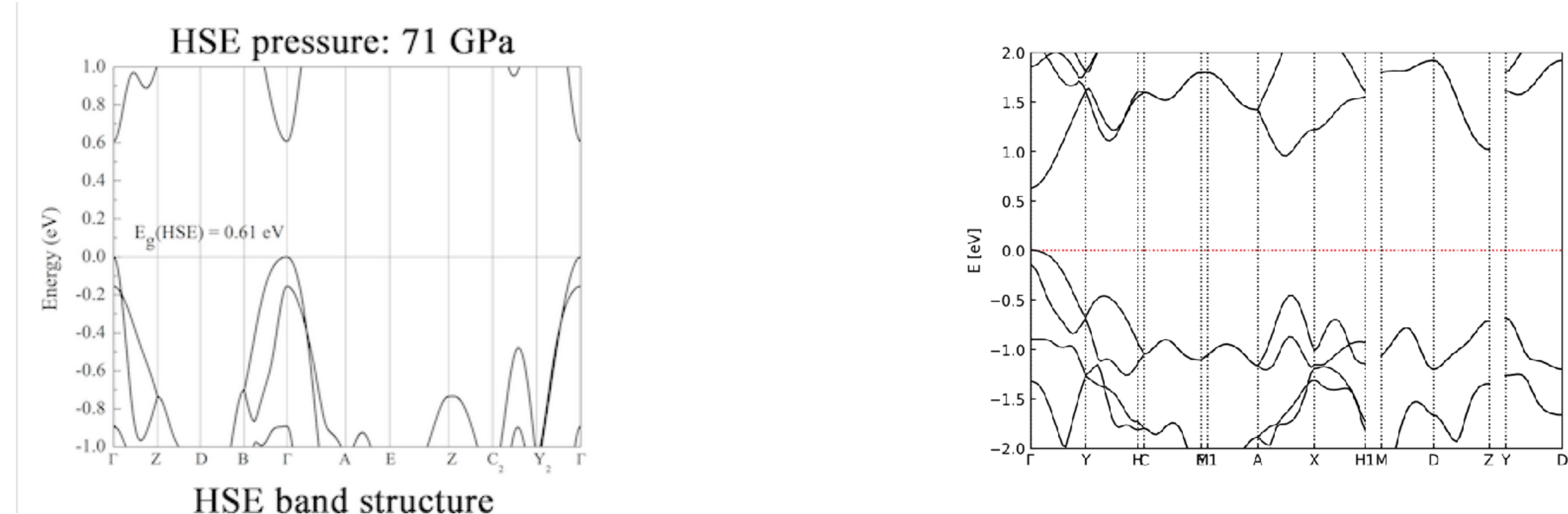
10<sup>3</sup>

10<sup>4</sup>

10<sup>5</sup>

No. MPI threads

# Bandstructures in reasonable time



**VASP:**  
4 nodes, 16 cores per node, 72h walltime

**FHI-aims:**  
2 nodes, 32 cores per node, 1.5h walltime

# How does it work?

**Structures: geometry.in**

```
>>> ase convert POSCAR geometry.in
```

# How does it work?

## CLI: CLIMS uses ASE

```
>>> pip install clims  
  
>>> clims-configure --species-path \  
~/local/FHIaims/species_defaults/defaults_2020/  
  
>>> clims-prepare-run --species light --relax
```

# How does it work?

Calculation parameters: control.in

```
xc                                pbe
relativistic                      atomic_zora scalar
k_grid                            8 8 8
relax_geometry                     trm 5e-3
relax_unit_cell                   full

# plus more input for the basissets
...
```

# How does it work?

## Calculation parameters

xc  
relativistic  
k\_grid  
relax\_geometry  
relax\_unit\_cell  
  
# plus more input  
...

```
#-----  
# FHI-aims file: control.in  
# Created using the Atomic Simulation Environment (ASE)  
# Fri Apr 19 16:55:22 2024  
#-----  
xc pbe  
relativistic atomic_zora scalar  
k_grid 8 8 8  
relax_geometry trm 5e-3  
relax_unit_cell full  
#-----  
#####  
#  
# FHI-aims code project  
# VB, Fritz-Haber Institut, 2009  
#  
# Suggested "light" defaults for Pt atom (to be pasted into control.in file)  
# Be sure to double-check any results obtained with these settings for post-processing,  
# e.g., with the "tight" defaults and larger basis sets.  
#  
#####  
species Pt  
# global species definitions  
nucleus 78  
mass 195.084  
#  
l_hartree 4  
#  
cut_pot 3.5 1.5 1.0  
basis_dep_cutoff 1e-4  
#  
radial_base 72 5.0  
radial_multiplier 1  
angular_grids specified  
division 0.4222 50  
division 0.9557 110  
division 1.2477 194  
division 1.5393 302  
# division 1.9382 434  
# division 2.0887 590  
# division 2.1534 770  
# division 2.2208 974  
# division 2.6985 1202  
# outer_grid 974  
outer_grid 302  
#####  
#  
# Definition of "minimal" basis  
#  
#####
```

ca scalar

# How does it work?

## CLI: CLIMS

```
>>> clims-prepare-run --species light --bands --dos -hse06
```

# How does it work?

Calculation parameters: `control.in`

```
xc                                hse06 0.11
hse_unit                           bohr
hybrid_xc_coeff                   0.25
relativistic                      atomic_zora scalar
k_grid                             9 9 9
exx_band_structure_version        1
output                            band  0.00000  0.00000  0.00000  0.00000  0.00000  0.50000  18 G  Y
output                            band  0.00000  0.00000  0.50000 -0.46222  0.00000  0.54058  18 Y  H
output                            band -0.46222  0.00000  0.54058 -0.50000  0.00000  0.50000  2 H  C
output                            band -0.50000  0.00000  0.50000 -0.50000  0.50000  0.50000  19 C  E
output                            band -0.50000  0.50000  0.50000 -0.53778  0.50000  0.45942  2 E  M1
output                            band -0.53778  0.50000  0.45942 -0.50000  0.50000  0.00000  17 M1 A
output                            band -0.50000  0.50000  0.00000 -0.50000  0.00000  0.00000  19 A  X
output                            band -0.50000  0.00000  0.00000 -0.53778  0.00000  0.45942  17 X  H1
output                            band -0.46222  0.50000  0.54058  0.00000  0.50000  0.50000  18 M  D
output                            band  0.00000  0.50000  0.50000  0.00000  0.50000  0.00000  18 D  Z
output                            band  0.00000  0.00000  0.50000  0.00000  0.50000  0.50000  19 Y  D

# plus more input for the basissets
```

# How does it work?

## Run the job

```
#!/bin/bash -l

ulimit -s unlimited
export OMP_NUM_THREADS=1

srun /path/to/aims/build/aims.231208.scalapack.mpi.x
```

# Final words



- ✓ **FHI-aims is easy to install** on tetralith/dardel → contact me
- ✓ **we have an FHI-aims license** → Florian Trybel
- ⚙️ **we could do a small code tutorial?** → contact me

# References



- [1] V. Blum *et al.*, Comput Phys Commun **180**, 2175 (2009).
- [2] S. V. Levchenko *et al.*, Comput Phys Commun **192**, 60 (2015).
- [3] A. C. Ihrig *et al.*, New J Phys **17**, 093020 (2015).
- [4] S. Kokott *et al.*, arXiv 2403.10343 (2024)

Thanks!

More info, code tutorials, etc:

[fhi-aims.org](https://fhi-aims.org)



Sebastian Kokott  
FHI / MSIP e.V.



Volker Blum  
Duke University