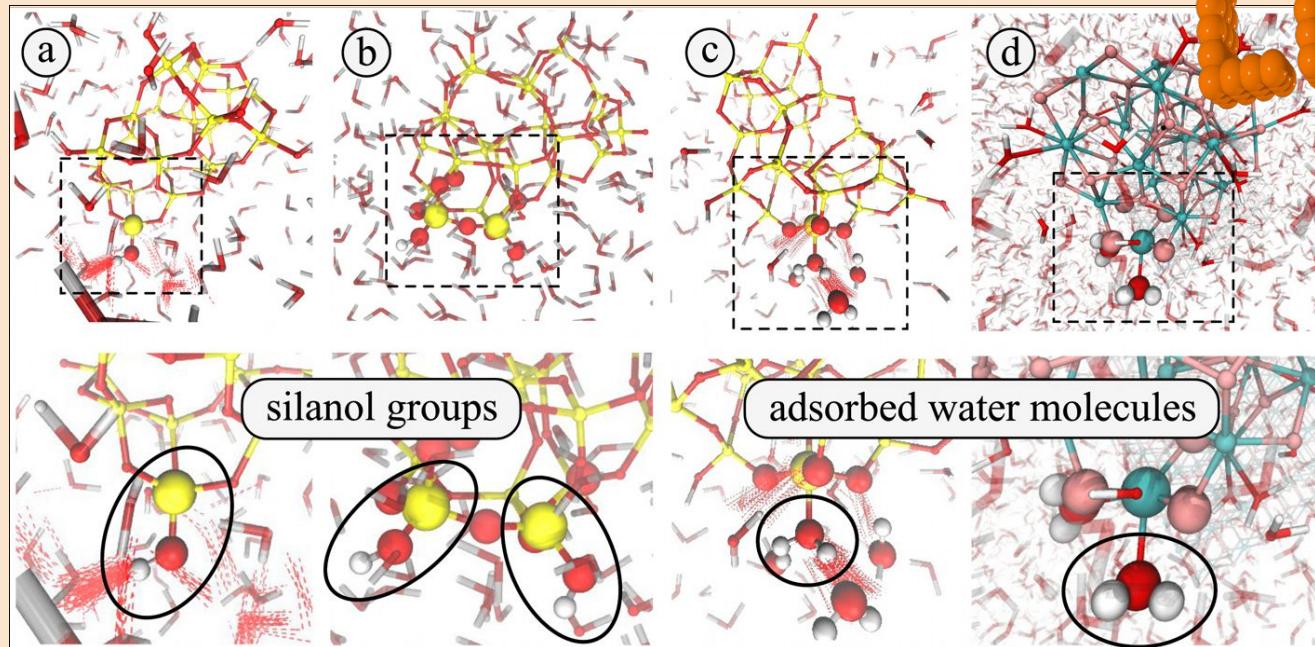


# ARCHER2 training - Introduction to CP2K



***Matt Watkins, University of Lincoln***

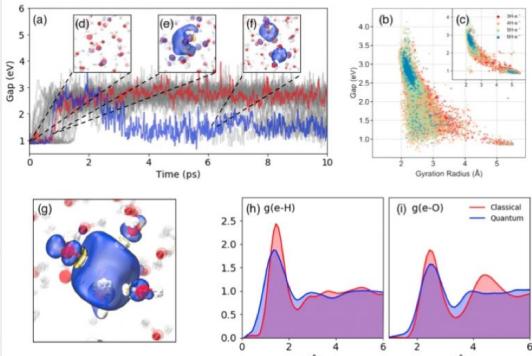
# CP2K - Science

Article | Open Access | Published: 03 February 2021

## Simulating the ghost: quantum dynamics of the solvated electron

Jinggang Lan, Venkat Kapil, Piero Gasparotto, Michele Ceriotti, Marcella Iannuzzi & Vladimir V. Rybkin\*

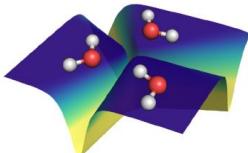
Nature Communications 12, Article number: 766 (2021) | Cite this article



## Acid-base equilibrium

### Microscopic description of acid-base equilibrium

Emanuele Grifoni<sup>a,b</sup>, GiovanniMaria Piccini<sup>b,a</sup>, and Michele Parrinello<sup>a,b,c,1</sup>



Emanuele Grifoni, GiovanniMaria Piccini, & Michele Parrinello; 2019; Microscopic description of acid-base equilibrium. PNAS 116, 4054–4057 (2019)

## Hydrated electrons



### Communications

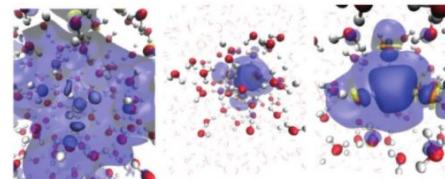


VIP Hydrated Electrons Very Important Paper

International Edition: DOI: 10.1002/anie.201814053  
German Edition: DOI: 10.1002/ange.201814053

### Dynamics of the Bulk Hydrated Electron from Many-Body Wave-Function Theory

Jan Wilhelm, Joost VandeVondele, and Vladimir V. Rybkin\*



Jan Wilhelm, Joost VandeVondele & Vladimir V. Rybkin; 2019; Dynamics of the Bulk Hydrated Electron from Many-Body Wave-Function Theory. Angewandte Chemie International Edition 58, 3890–3893 (2019)

## Micrometre-long covalent organic fibres



Article Published: 27 August 2018

Micrometre-long covalent organic fibres by photoinitiated chain-growth radical polymerization on an alkali-halide surface



Franck Para, Franck Bocquet, Laurent Nony, Christian Loppacher, Michel Féron, Frédéric Chérioux, David Z. Gao, Filippo Federici Canova & Matthew B. Watkins; 2018; Micrometre-long covalent organic fibres by photoinitiated chain-growth radical polymerization on an alkali-halide surface. Nature Chemistry 10, 1112–1117 (2018) Download Citation

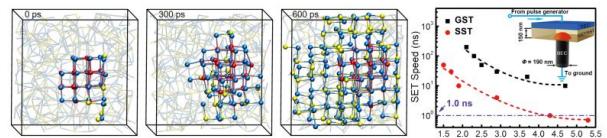
## Science

REPORTS

### PHASE-CHANGE MEMORY

## Reducing the stochasticity of crystal nucleation to enable subnanosecond memory writing

Feng Rao,<sup>1,2\*</sup>† Keyuan Ding,<sup>1,2\*</sup> Yuxing Zhou,<sup>3\*</sup> Yonghui Zheng,<sup>1</sup> Mengjiao Xia,<sup>4</sup> Shilong Lv,<sup>1</sup> Zhitang Song,<sup>1,†</sup> Songlin Feng,<sup>1</sup> Ider Ronneberger,<sup>5</sup> Riccardo Mazzarello,<sup>5</sup> Wei Zhang,<sup>3,†</sup> Evan Ma<sup>3,6</sup>



Cite This! ACS Cent. Sci. 2019, 5, 1663–1675

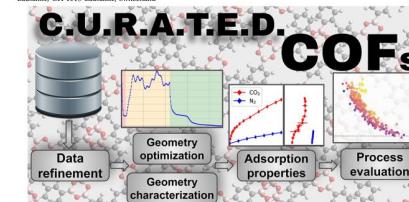
http://pubs.acs.org/journal/acscsn

### Building a Consistent and Reproducible Database for Adsorption Evaluation in Covalent–Organic Frameworks

Danièle Orgari,<sup>1</sup> Aliaksandr V. Yakutovich,<sup>1,2</sup> Leopold Talirz,<sup>1,2</sup> and Berend Smit<sup>1,2,3</sup>

<sup>1</sup>Laboratory of Molecular Simulation (LMSM), Institute des Sciences et Ingénierie Chimiques, École Polytechnique Fédérale de Lausanne (EPFL), Baie de l'Institut 17, Sion, CH-1953 Valais, Switzerland

<sup>2</sup>Theory and Simulation of Materials (THEOS), Faculty des Sciences et Techniques de l'Ingénieur, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland



Danièle Orgari, Aliaksandr V. Yakutovich, Leopold Talirz & Berend Smit; 2019; Building a Consistent and Reproducible Database for Adsorption Evaluation in Covalent–Organic Frameworks. ACS Cent. Sci. 2019, 5, 1663–1675

<https://www.cp2k.org/science>

# CP2K - status paper

paper in JCP special issue on electronic structure methods

AIP The Journal of Chemical Physics

HOME BROWSE INFO FOR AUTHORS COLLECTIONS ACCEPTED MANUSCRIPTS

SIGN UP FOR ALERTS

Home > The Journal of Chemical Physics > Volume 152, Issue 19 > 10.1063/5.0007045

Open • Submitted: 10 March 2020 • Accepted: 22 April 2020 • Published Online: 19 May 2020

CP2K: An electronic structure and molecular dynamics software package - Quickstep: Efficient and accurate electronic structure calculations 

J. Chem. Phys. 152, 194103 (2020); <https://doi.org/10.1063/5.0007045>

 Thomas D. Kühne<sup>1,a)</sup>,  Marcella Iannuzzi<sup>2</sup>, Mauro Del Ben<sup>3</sup>,  Vladimir V. Rybkin<sup>2</sup>, Patrick Seewald<sup>2</sup>, Frederick Stein<sup>2</sup>,  Teodoro Laino<sup>4</sup>,  Rustam Z. Khalilullin<sup>5</sup>, Ole Schütt<sup>6</sup>,  Florian Schiffmann<sup>7</sup>,  Dorothea Golze<sup>8</sup>, Jan Wilhelm<sup>9</sup>, Sergey Chulkov<sup>10</sup>,  Mohammad Hossein Bani-Hashemian<sup>11</sup>, Valéry Weber<sup>4</sup>, Urban Borštník<sup>12</sup>, Mathieu Taillefumier<sup>13</sup>,  Alice Shoshana Jakobovits<sup>13</sup>, Alfio Lazzaro<sup>14</sup>, Hans Pabst<sup>15</sup>, Tiziano Müller<sup>2</sup>,  Robert Schade<sup>16</sup>, Manuel Guidon<sup>2</sup>, Samuel Andermatt<sup>11</sup>, Nico Holmberg<sup>17</sup>,  Gregory K.

more...

<https://www.cp2k.org/science>

# *Ab Initio* DFT: The Basics and The Basis

Kohn-Sham Equation:

$$\mathbf{H}(\rho)\mathbf{C} = \mathbf{C}\Lambda$$

Density:

$$\rho(\mathbf{r}) = \sum_{\alpha} |\psi_{\alpha}(\mathbf{r})|^2$$

$$\psi_{\alpha}(\mathbf{r}) = \sum_i c_{i\alpha} \chi_i(\mathbf{r})$$

Orbital:

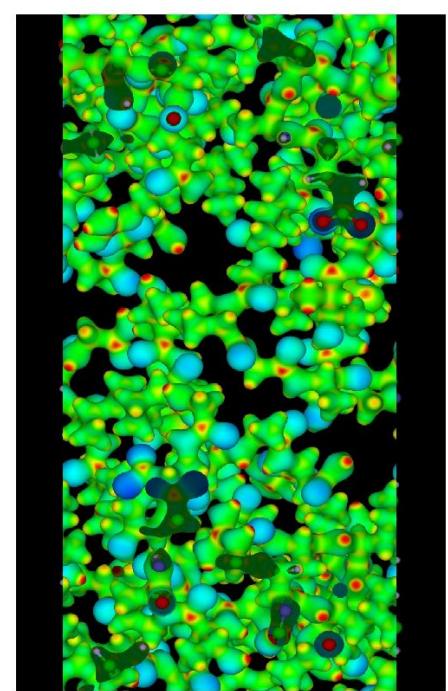
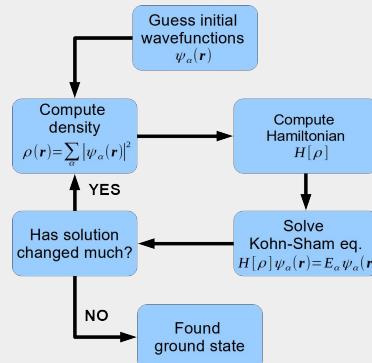
$\{\chi_i\}$  - “The basis set”

Operations:

- From  $\rho$  build  $\mathbf{H}$
- From  $\mathbf{H}$  generate new  $\mathbf{C}$  and hence  $\rho$

Memory Requirements:

- $\mathbf{H}$  - a  $n\_basis \times n\_basis$  matrix
- $\mathbf{C}$  - a  $n\_basis \times n\_electrons$  matrix



# ***Ab Initio - Other Methods***

Higher Accuracy; smaller systems

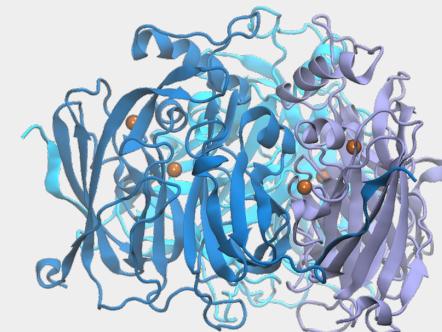
- Double hybrid functionals
- Post DFT - (GW; Green's Function ...)
- Post Hartree-Fock - e.g. MP2; Perturbation Theory

Lower Accuracy; larger systems

- Tight-binding, Semi-empirical

Embedding:

- QM/MM - e.g. Chemshell, CP2K/Gromacs



## Ok we're going to do DFT

### Algorithmic challenges

- construct the Kohn-Sham matrix
  - Hartree potential
  - XC potential
  - HF/exact exchange
  - core/pseudopotential terms
- fast and efficient minimization of the energy functional
- efficient calculation of the density matrix and construction of the Molecular Orbitals (**C**)

Construction of the KS matrix should be (and is in CP2K) linear scaling ( $O(N)$ ) in basis set size. Using dual basis set (GPW) techniques - coming up.

Minimization / diagonalization is harder - linear scaling routines are available but are only more efficient for large systems and require considerable computational resources.

## Gaussian's and Plane Wave method (GPW)

- localized Gaussians, centred at atomic positions
- plane waves

The idea of GPW is to use the plane-waves as an auxiliary basis, primarily to construct the Hartree potential.

This leads to linear scaling KS construction for Gaussian Type Orbitals (GTO)

## GPW ingredients

As the name suggests, this method uses two different types of functions

- Gaussian basis sets (many matrix elements can be done analytically)  
we go a bit further than implied above - to be more accurate, we contract several Gaussians to form approximate atomic orbitals

$$\phi_{\alpha}(\mathbf{r}) = \sum_m d_{m\alpha} g_m(\mathbf{r})$$

where a primitive cartesian Gaussian centred at the origin is given by

$$g_m(\mathbf{r}) = x^{m_x} y^{m_y} z^{m_z} e^{-\alpha_m r^2}$$

and  $m_x + m_y + m_z = l$ , the angular momentum quantum number of the functions.

- Pseudo potentials
- Plane waves auxiliary basis for Coulomb integrals
- Regular grids and FFT for the density
- Sparse matrices (KS and P)
- Efficient screening

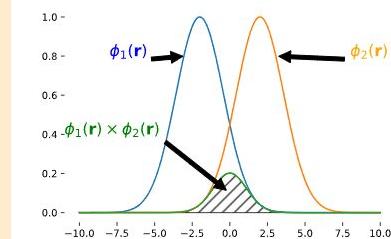
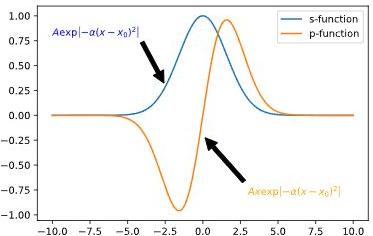
## Gaussian Basis set

- localised atom-centred GTO basis

$$\phi_{\alpha}(\mathbf{r}) = \sum_m d_{m\alpha} g_m(\mathbf{r})$$

- expansion of the density using the density matrix

$$n(\mathbf{r}) = \sum_{\alpha\beta} P_{\alpha\beta} \phi_{\alpha}^*(\mathbf{r}) \phi_{\beta}(\mathbf{r})$$



- operator matrices are sparse. By neglecting overlap elements smaller than some number indirectly specified by `EPS_DEFAULT`, the Gaussian functions are localized with finite support. This means that S, H and P are sparse.

$$S_{\alpha\beta} = \int_r \phi_{\alpha}(\mathbf{r}) \phi_{\beta}(\mathbf{r}) d\mathbf{r}$$

$$H_{\alpha\beta} = \int_r \phi_{\alpha}(\mathbf{r}) V(\mathbf{r}) \phi_{\beta}(\mathbf{r}) d\mathbf{r}$$

# Ab Initio Linear Scaling

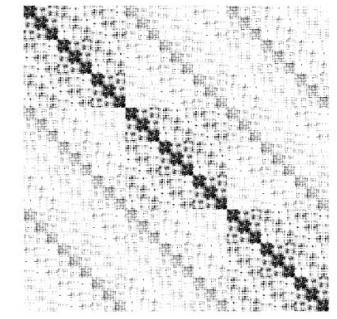
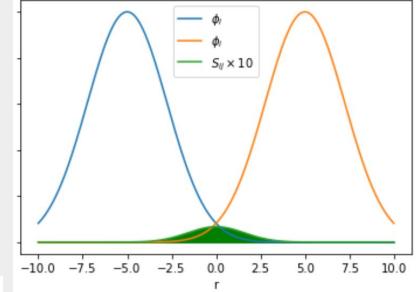
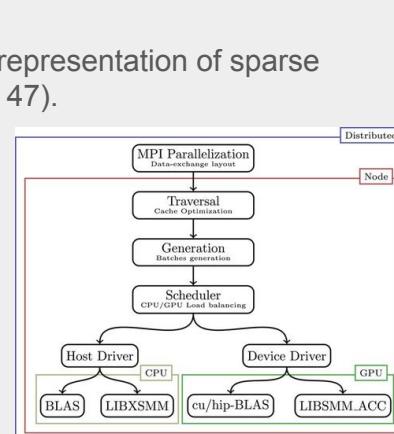
Sparsity is critical:

- Using elements with finite range (Gaussians, numerical, wavelets) - neglect if overlap is smaller than some cutoff
- Overlap **and** Hamiltonian matrices are sparse

$$\phi_\alpha(\mathbf{r}) = \sum_i c_{i\alpha} \chi_i(\mathbf{r}) \quad \chi_i(\mathbf{r}) = N e^{-\alpha r^2} Y_{lm}(\theta, \phi) \quad S_{ij} = \begin{cases} \int d\mathbf{r}^3 \phi_i(\mathbf{r}) \phi_j(\mathbf{r}) \\ \text{or } 0 \text{ if } < \epsilon \end{cases}$$

- e.g, CP2K uses distributed block cyclic sparse row (libDBCSR) representation of sparse matrices (Borštník, Urban, et al. *Parallel Computing* **40**, 5 (2014) 47).
  - Linear scaling matrix - matrix multiplication
  - Complicated to accelerate on GPUs - many block sizes
  - Sparse Linear Algebra
- 
- Direct density matrix minimization
  - Linear scaling Hamiltonian Build - linear scaling Poisson Solver
  - Linear scaling optimization of density matrix - many methods

Example Codes: CP2K, bigDFT, ONETEP



Sparsity pattern of  $\mathbf{H}$  for a 864 H<sub>2</sub>O molecule system

## Basis set libraries

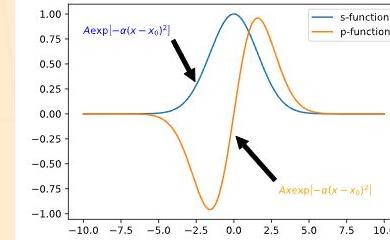
There are two main types of basis sets supplied with CP2K

- GTH\_BASIS\_SETS: atomically optimized sets. These were the first shipped with the code and vary systematically in quality from DZ to QZ for lighter elements. Can be very good for molecular systems, but can be very bad for condensed matter systems.
- BASIS\_MOLOPT: molecularly optimized basis sets. These cover most elements of the periodic table, but only with fairly good quality DZVP-MOLOPT-SR-GTH. Should be a good starting point for most condensed matter calculations.

The basis set files provide the contraction coefficients ( $d_{m\alpha}$ ) and exponents ( $\alpha_m$ ) of the Gaussian functions.

$$\phi_\alpha(\mathbf{r}) = \sum_m d_{m\alpha} g_m(\mathbf{r})$$

$$g_m(\mathbf{r}) = x^{m_x} y^{m_y} z^{m_z} e^{-\alpha_m r^2}$$



Generally the first column is the exponents ( $\alpha_m$  above) and the later columns give the  $d_{m\alpha}$ , each column being a set  $\alpha$ . Details can be found in the header of the BASIS\_MOLOPT file

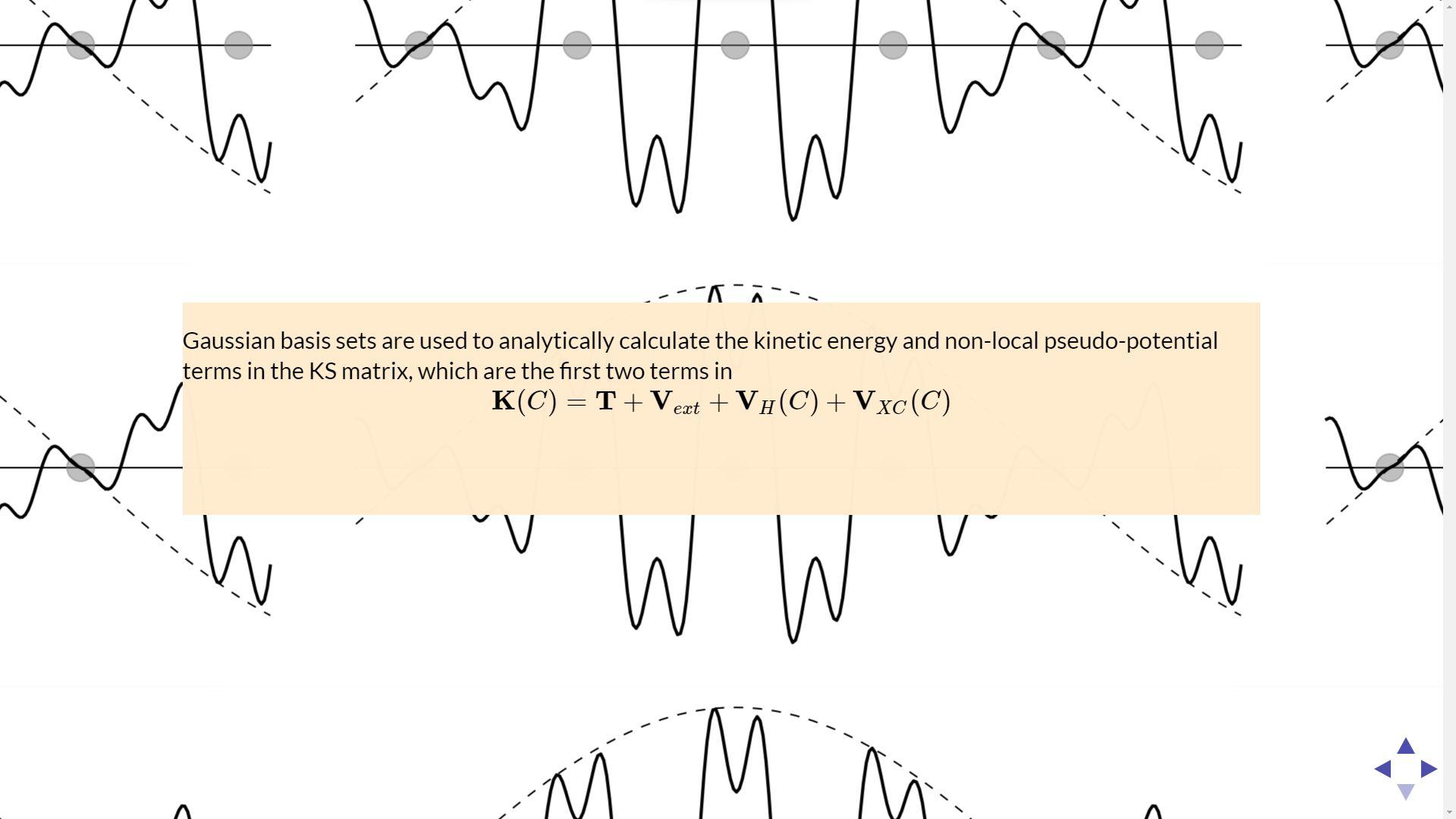
A MOLOPT basis for nitrogen is

```
N DZVP-MOLOPT-SR-GTH DZVP-MOLOPT-SR-GTH-q5
1
2 0 2 5 2 2 1
7.341988051825  0.113789156500  0.077765588400 -0.053744330400 -0.007627243700  0.033688455200
2.542637110957  0.097294516500  0.108655219900 -0.165752516200  0.015163333100  0.109813343200
0.888574967229 -0.445077422600 -0.374125427100 -0.317365165600 -0.129388247500  0.856542971300
0.333802200435 -0.584142233900  0.024021712400 -0.312039675200  0.554905847400  0.509681657500
0.112012109029 -0.139562383500  0.979415132500 -0.117936008100  1.001020469600  0.047030652200
```

Here there are gaussians with five different exponents (in Bohr<sup>-2</sup>). From these 5 sets of functions are built, two *s* functions (2nd and 3rd columns), 2 sets of *p* functions (4th and fifth columns) and one set of *d* functions (last column).

Note that the contraction coefficients are not varied during calculation. For the nitrogen basis above we have  $2 + 2 \times 3 + 1 \times 5 = 13$  variables to optimize for each nitrogen atom in the system.

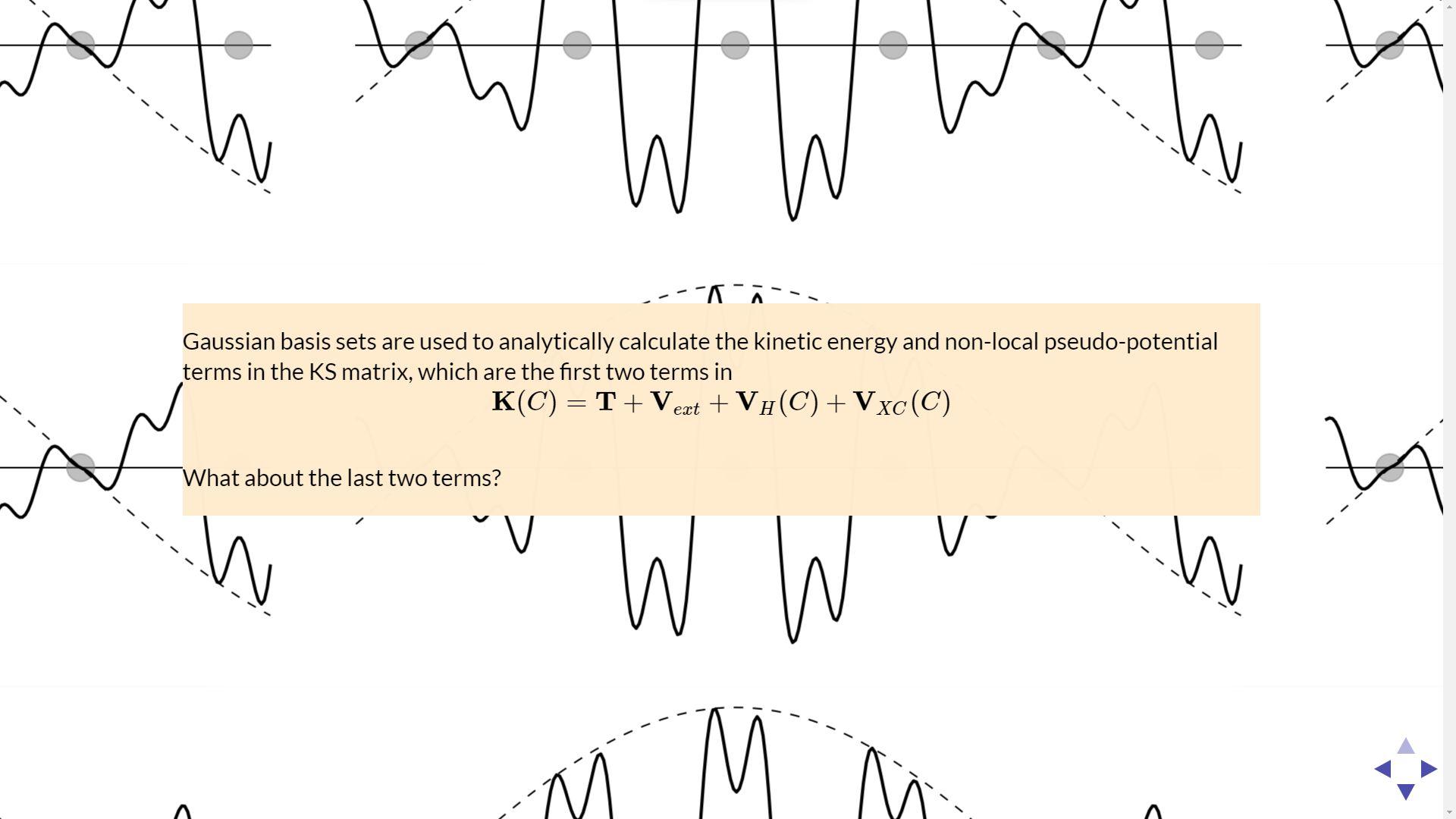
(CP2K tends to use general contractions for efficiency, allowing maximum use of recursion relations to generate matrix elements for high angular momentum functions).



Gaussian basis sets are used to analytically calculate the kinetic energy and non-local pseudo-potential terms in the KS matrix, which are the first two terms in

$$\mathbf{K}(C) = \mathbf{T} + \mathbf{V}_{ext} + \mathbf{V}_H(C) + \mathbf{V}_{XC}(C)$$





Gaussian basis sets are used to analytically calculate the kinetic energy and non-local pseudo-potential terms in the KS matrix, which are the first two terms in

$$\mathbf{K}(C) = \mathbf{T} + \mathbf{V}_{ext} + \mathbf{V}_H(C) + \mathbf{V}_{XC}(C)$$

What about the last two terms?

# *Ab Initio* Plane-Waves

Use a Fourier Expansion to represent the wavefunction

- $\chi_{\mathbf{G}}(\mathbf{r}) = Ne^{i\mathbf{G} \cdot \mathbf{r}}$

Good:

- All points in space equally represented
- Doesn't require an *a priori* "appreciation" of the answer
- Easy to systematically improve - higher frequencies
- Doesn't care if atoms move - basis not tied to atomic positions
- **H** cheap to calculate - use FFT and exploit sparseness in the appropriate space

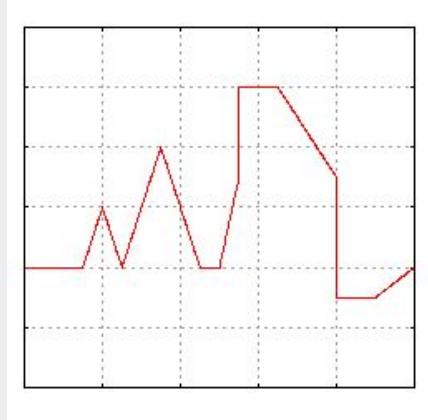
Bad:

- All points in space equally represented - might have to do a lot of work to represent "nothing"
- Do *not* look like the answer so very many are needed - >1,000,000

Algorithms:

- FFT
- Dense linear Algebra - tall, thin matrices

Example 'pure' PW Codes: CASTEP; VASP; ABINIT; Quantum Espresso

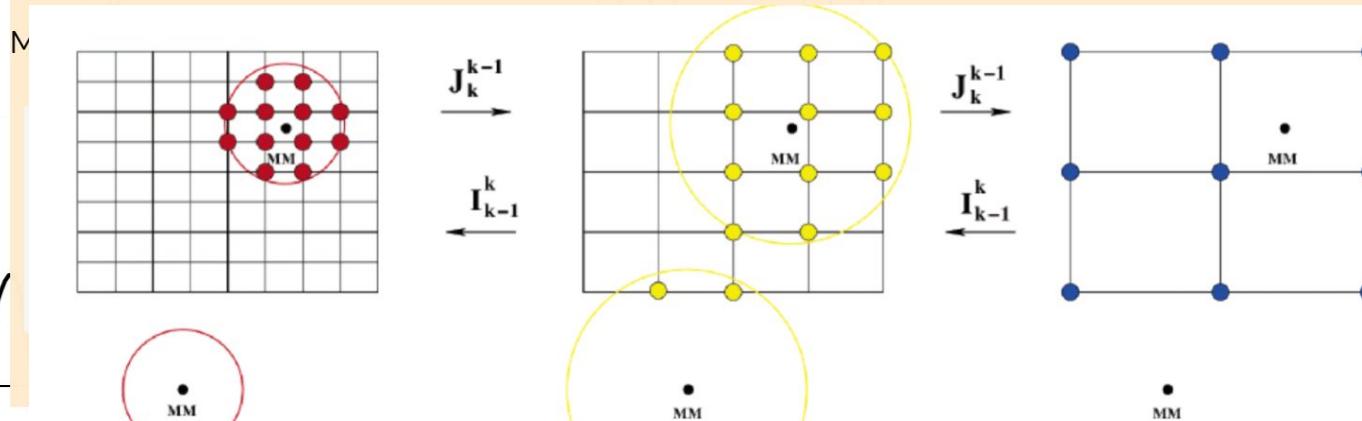


By Scallop7 - Own work; CC BY-SA 4.0;  
<https://commons.wikimedia.org/w/index.php?curid=74597157>

# Multigrids

We collocate the Gaussian representation of the density onto the plane-wave basis set.

When we want to put (collocate) a Gaussian type function onto the realspace grid, we can gain efficiency by using multiple grids with differing cutoff / spacing.



More diffuse Gaussians can be collocated onto coarser grids specified in input as:

```
8MGRID  
CUTOFF 400  
REL_CUTOFF 60  
NGRIDS 5  
&END MGRID
```

## Real space integration

Finite cutoff and simulation box define a realspace grid

- density collocation

$$n(\mathbf{r}) = \sum_{\alpha\beta} P_{\alpha\beta} \phi_{\alpha}^{*}(\mathbf{r}) \phi_{\beta}(\mathbf{r}) \rightarrow \sum_{\alpha\beta} P_{\alpha\beta} \bar{\phi}_{\alpha\beta}(\mathbf{r}) = n(\mathbf{R})$$

where  $n(\mathbf{R})$  is the density at grid points in the cell

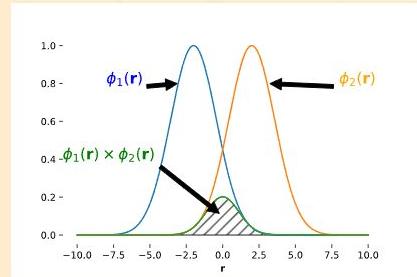
- numerical approximation of the gradient  $n(\mathbf{R}) \rightarrow \nabla n(\mathbf{R})$

- $\epsilon_{xc}$  and derivatives evaluated on the grid

$$v_{XC}[n](\mathbf{r}) \rightarrow V_{XC}(\mathbf{R}) = \frac{\partial \epsilon_{xc}}{\partial n}(\mathbf{R})$$

- real space integration

$$H_{HXC}^{\mu\nu} = \langle \mu | V_{HXC}(\mathbf{r}) | \nu \rangle \rightarrow \sum_{\mathbf{R}} V_{HXC}(\mathbf{R}) \bar{\phi}_{\alpha\beta}(\mathbf{R})$$



## Electrostatics

- long-range term: Non-local Hartree Potential  
Poisson equation solved using the auxiliary plane-wave basis

$$E_H[n_{tot}] = \frac{1}{2} \int_r \int_{r'} \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}$$

where  $n_{tot}$  includes the nuclear charge as well as the electronic.

(The nuclear charge density is (of course) represented as a Gaussian distribution with parameter  $R_I^c$  chosen to cancel a similar term from the local part of the pseudopotential)

- FFT (scaling as  $N\log N$ ) gives

$$\tilde{n}(\mathbf{r}) = \frac{1}{\Omega} \sum_G \tilde{n}(\mathbf{G}) e^{i\mathbf{G} \cdot \mathbf{r}}$$

In the G space representation the Poisson equation is diagonal and the Hartree energy is easily evaluated

$$E_H[n_{tot}] = 2\pi\Omega \sum_G \frac{\tilde{n}(\mathbf{G})^* \tilde{n}(\mathbf{G})}{\mathbf{G}^2}$$

## Multigrids

you can see in the output

### MULTIGRID INFO

count for grid 1:	24832	cutoff [a.u.]	200.00
count for grid 2:	11934	cutoff [a.u.]	66.67
count for grid 3:	6479	cutoff [a.u.]	22.22
count for grid 4:	1976	cutoff [a.u.]	7.41
count for grid 5:	215	cutoff [a.u.]	2.47
total gridlevel count :	45436		

For this system (formaldehyde with an aug-TZV2P-GTH basis) that 45436 density matrix elements ( $\bar{\phi}_{\alpha\beta}$ ) were mapped onto the grids. To be efficient, all grids should be used. To fully converge calculations both "CUTOFF" and "REL\_CUTOFF" need to be increased together otherwise the highest grid may not be used.

## Timings

At the end of the run you'll see timings - these can be very useful for understanding performance. Here we can see the timings for the formaldehyde example. It is in a large box, and the operations on grids dominate the cost on a single processor.

TIMING							
SUBROUTINE	CALLS		ASD		SELF TIME		TOTAL TIME
	MAXIMUM	AVERAGE	MAXIMUM	AVERAGE	MAXIMUM	MAXIMUM	
CP2K		1	1.0	0.013	0.013	24.048	24.048
qs_energies		1	2.0	0.000	0.000	23.417	23.417
scf_env_do_scf		1	3.0	0.000	0.000	22.962	22.962
scf_env_do_scf_inner_loop	14	4.0	0.001	0.001	22.962	22.962	
qs_ks_update_qs_env	14	5.0	0.000	0.000	19.555	19.555	
rebuild_ks_matrix	14	6.0	0.000	0.000	19.554	19.554	
qs_ks_build_kohn_sham_matrix	14	7.0	0.002	0.002	19.554	19.554	
qs_vxc_create	14	8.0	0.000	0.000	8.977	8.977	
xc_vxc_pw_create	14	9.0	0.273	0.273	8.977	8.977	
xc_rho_set_and_dset_create	14	10.0	0.086	0.086	8.147	8.147	
pw_transfer	203	9.2	0.009	0.009	7.539	7.539	
fft_wrap_pw1pw2	203	10.2	0.001	0.001	7.530	7.530	

## Warnings

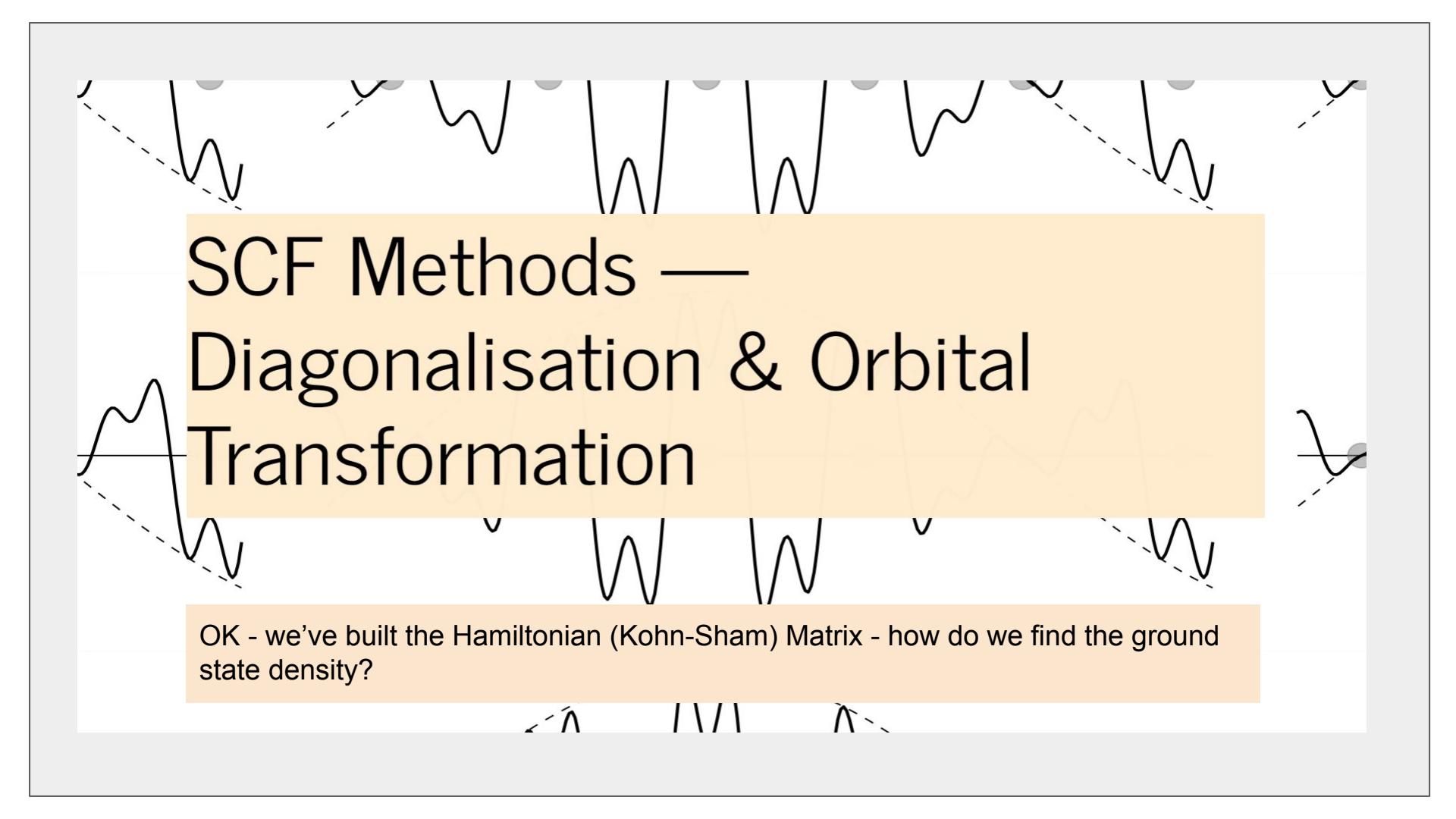
Also check if you get an output like

```
The number of warnings for this run is : 0
```

if there are warnings, you should try and understand why!

## GAPW

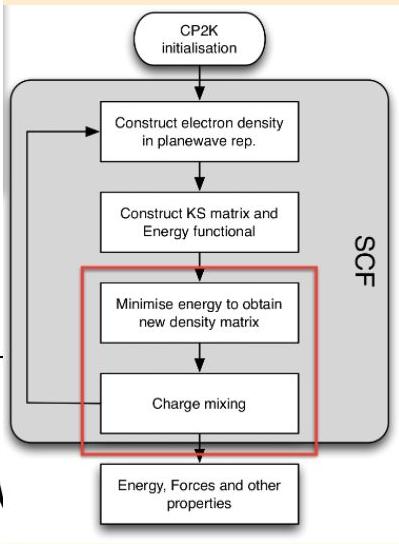
- The Gaussian and Augmented Plane Wave method works similarly to the GPW method. But Gaussians with larger exponents are not mapped onto the grids, instead onto an extra set of atom centered radial grids (there is more to it than this).
- This allows core electrons to be represented and all-electron calculations to be run
- It also allows a smaller plane-wave cutoff to be used.
- Activate GAPW with the METHOD GAPW keywords in the &QS section.
- To run all-electron use POTENTIAL ALL in the &KIND SECTION and an all-electron basis.



# SCF Methods — Diagonalisation & Orbital Transformation

OK - we've built the Hamiltonian (Kohn-Sham) Matrix - how do we find the ground state density?

# Self Consistent Field (SCF)



- Central to the QuickStep (DFT) calculation is the Self-Consistent-Field cycle

$$H[\rho]\phi_n = E_n\phi_n$$
$$\rho(\mathbf{r}) = \sum_n f_n \phi_n^*(\mathbf{r})\phi_n(\mathbf{r})$$

- Key to speed and stability of the calculation:
  - Energy minimisation
  - Charge mixing

## Overview

- Common Methods In CP2K
  - Eigensolvers (diagonlisation)
    - Charge Mixing for Diagonalisation Methods
  - Optimisers
    - Orbital Transformation (OT)
    - Preconditioners
  - Comparison of OT and Diagonalization.

This talk is only an introductory overview.

A much more detailed discussion of the SCF optimisation methods can be found in the [notes from Lianheng Tong](#)

V V V v



## Generalised Eigenvalue Problem

We converted the Kohn-Sham equations into matrix form by introducing basis functions.

$$\mathbf{H}\mathbf{C} = \mathbf{E}\mathbf{S}\mathbf{C}$$

Where

- **H** is the matrix representation of the Kohn-Sham equations
- **C** is the matrix of Molecular Orbital coefficients of the basis functions used
- **S** is the overlap matrix showing how the basis functions overlap (are not orthogonal).
- **E** is the matrix with the eigenenergies of the MOs.

# Diagonalisation

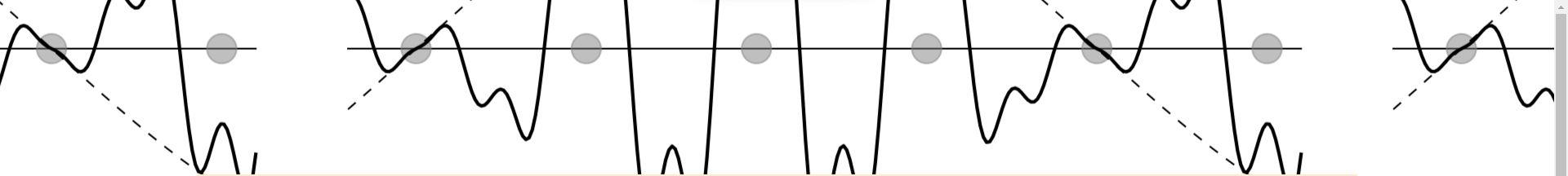
we can diagonalise  $\mathbf{H}'\mathbf{C}' = \mathbf{E}\mathbf{C}'$  to find a new set of MOs given the input Kohn-Sham matrix built from the current density.

Standard methods of diagonalising the matrix can be used - termed 'traditional diagonalisation'.

The new orbitals are used to build

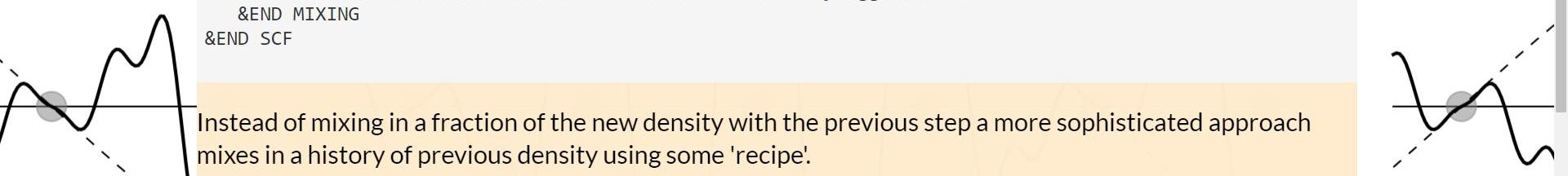
- a new density
- a new Kohn-Sham matrix

Then the process repeats until (if?) it converges - i.e. MOs in are the same as MOs out.

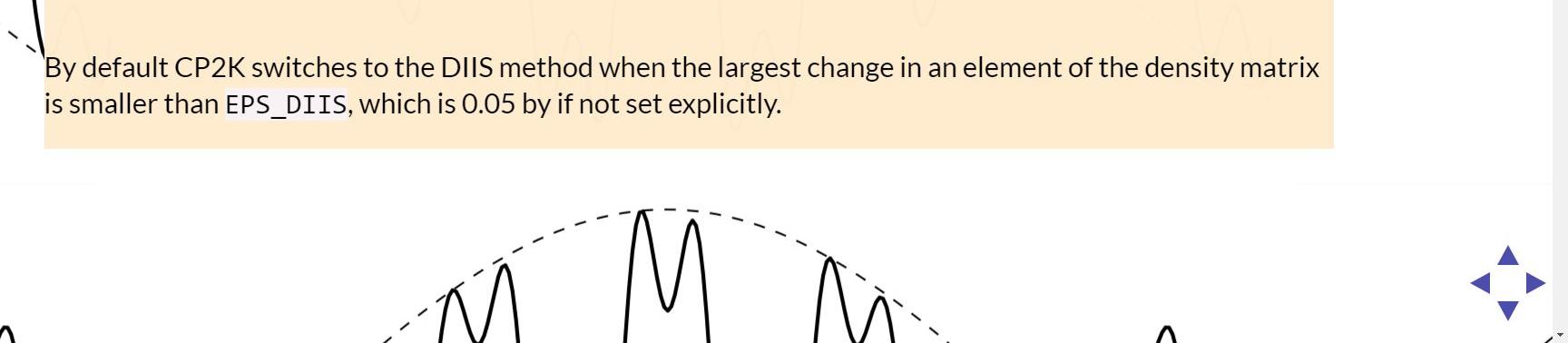


## Input - linear mixing

```
&SCF
  SCF_GUESS ATOMIC
  EPS_SCF 1.0E-06
  MAX_SCF 50
  &MIXING
    ALPHA 0.2 !sensible value, the default 0.4 is very aggressive.
  &END MIXING
&END SCF
```



Instead of mixing in a fraction of the new density with the previous step a more sophisticated approach mixes in a history of previous density using some 'recipe'!



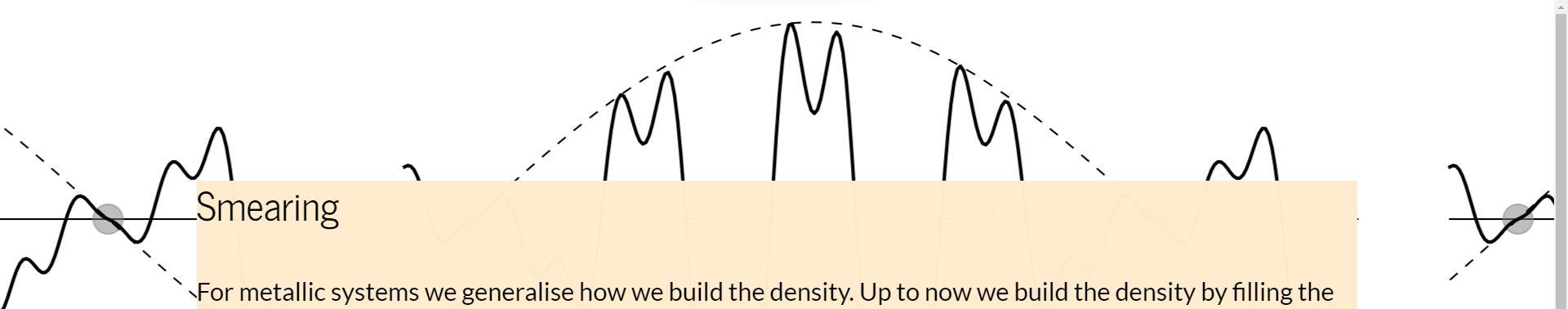
By default CP2K switches to the DIIS method when the largest change in an element of the density matrix is smaller than EPS\_DIIS, which is 0.05 by if not set explicitly.

## Output

should look something like this:

Step	Update method	Time	Convergence	Total energy	Change
1	P_Mix/Diag.	0.50E+00	2.1	0.41056021	-2133.4408435676
2	P_Mix/Diag.	0.50E+00	3.2	0.20432922	-2132.0776002852
3	P_Mix/Diag.	0.50E+00	3.2	0.10741372	-2131.3677551799
4	P_Mix/Diag.	0.50E+00	3.2	0.05420394	-2131.0080867703
5	DIIS/Diag.	0.39E-03	3.2	0.02722180	-2130.8276990683
6	DIIS/Diag.	0.19E-03	3.1	0.00062404	-2130.6473761946
7	DIIS/Diag.	0.84E-04	3.2	0.00050993	-2130.6473778175

note the switch to DIIS when Convergence is < 0.05.



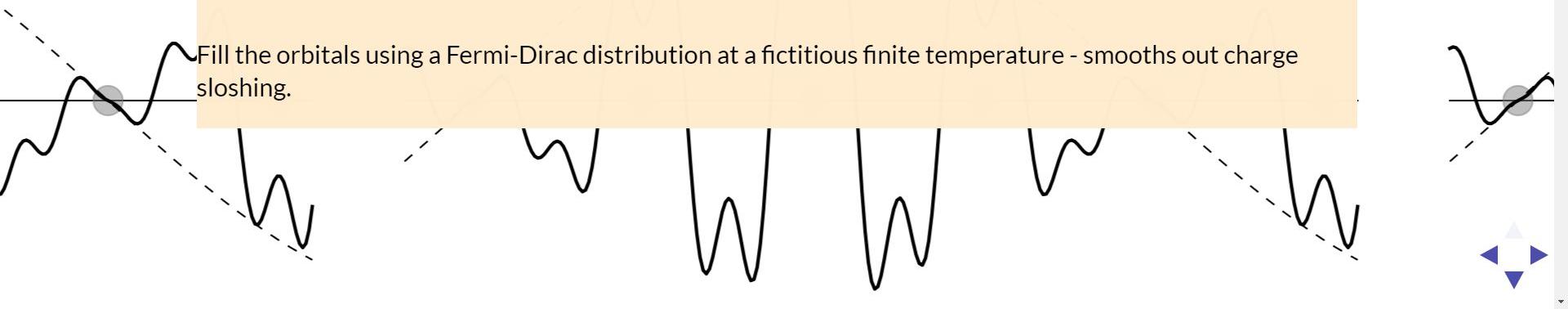
## Smearing

For metallic systems we generalise how we build the density. Up to now we build the density by filling the  $N$  electrons into the  $N$  lowest molecular spin orbitals.

### Problem

If the system is metallic (or has a very small band-gap) this can lead to 'charge sloshing'. The orbitals around the 'fermi energy' can change their ordering, and different ones are occupied from iteration to iteration.

### Solution



Fill the orbitals using a Fermi-Dirac distribution at a fictitious finite temperature - smooths out charge sloshing.

## Smearing input

```
&SCF
  SCF_GUESS ATOMIC
  EPS_SCF 1.0E-6
  MAX_SCF 50
  ADDED_MOS 200
  &SMEAR ON
    METHOD FERMI_DIRAC
    ELECTRONIC_TEMPERATURE [K] 300
  &END SMEAR
  &MIXING
    METHOD BRYODEN_MIXING
    ALPHA 0.2
    NBUFFER 5
  &END MIXING
&END SCF
```

We also use a different mixing scheme, which is probably optimal for metallic systems.

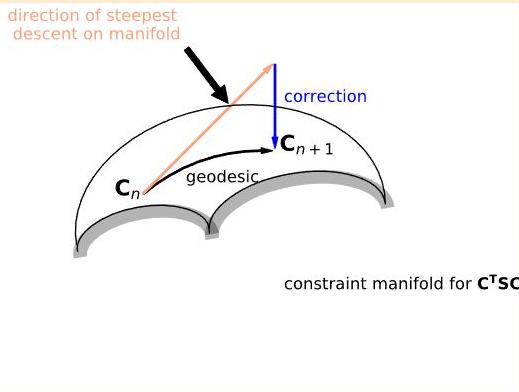
Fermi Temperature is typically between 300 - 3000 K. The larger the value the smoother convergence, but it can affect the physical properties of the system if too large.

## OT

why not just directly minimize the energy functional with respect to the MO coefficients?

- we need our orbitals to remain orthogonal - Pauli principle

So the minimization must be subject to a constraint - on an M dimensional hypersphere!



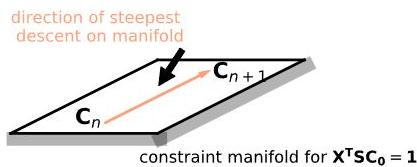
This is built into diagonalisation, as the new vectors are always eigenfunctions of the (current) Kohn-Sham matrix.

change variables in some clever way that builds in the constraint

Work with new variables  $\mathbf{X}$

$$\mathbf{C}(\mathbf{X}) = \mathbf{C}_0 \cos(\mathbf{U}) + \mathbf{X}\mathbf{U}^{-1} \sin(\mathbf{U})$$

$$\mathbf{U} = (\mathbf{X}^T \mathbf{S} \mathbf{X})^{1/2}$$



Can show that this leads to optimization in an  $M-1$  dimensional linear space.

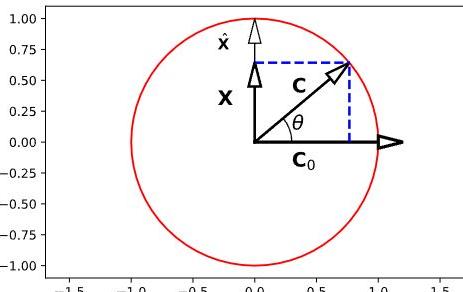
Change variables in some clever way that builds in the constraint.

$$\langle \mathbf{X}, \mathbf{C}_0 \rangle = \mathbf{X}^T \mathbf{S} \mathbf{C}_0 = 0$$

$$\begin{aligned}\mathbf{C} &= \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix} = \mathbf{C}_0 \cos(\theta) \\ &\quad + \hat{\mathbf{X}} \sin(\theta)\end{aligned}$$

$$\begin{aligned}\theta &= \frac{|\mathbf{X}|}{|\mathbf{C}|} = \frac{(\mathbf{X}^T \mathbf{S} \mathbf{X})^{1/2}}{(\mathbf{C}^T \mathbf{S} \mathbf{C})^{1/2}} \\ &= \mathbf{U}\end{aligned}$$

$$\begin{aligned}\mathbf{C}(\mathbf{X}) &= \mathbf{C}_0 \cos(\mathbf{U}) \\ &\quad + \mathbf{X} \mathbf{U}^{-1} \sin(\mathbf{U})\end{aligned}$$



New variables  $\mathbf{X}$  are always orthogonal to the initial vector  $\mathbf{C}_0$  by construction. We can use standard minimization routines and we never have to diagonalise the Kohn-Sham matrix! Far more robust - but requires a band gap - revert to diagonalization and smearing for metallic systems.

## Preconditioning

In minimization problems it is often a good idea change the problem by applying some approximate solution to the problem to make an equivalent set of equations that are

- approximately diagonal
- with diagonal elements of the same size and of order 1.

The OT solver is no exception. There are a variety of preconditioners available, and they can dramatically speed up convergence.

## OT recipe 1 - small-medium systems

```
&SCF
    SCF_GUESS RESTART
    EPS_SCF 1.0E-06
    MAX_SCF 20
    &OT ON
        MINIMIZER DIIS
        PRECONDITIONER FULL_ALL
        ENERGY_GAP 0.001
    &END OT
    &OUTER_SCF
        MAX_SCF 2
    &END OUTER_SCF
&END SCF
```

- This uses the most efficient minimizer, DIIS. Change to CG for difficult systems.
- The most accurate, and expensive to calculate, preconditioner - **FULL\_ALL**.
- The **OUTER\_SCF** restarts the SCF cycle and reapplies a new preconditioner when the original loop finishes.

## OT recipe 2 - pretty large systems

```
&SCF
    SCF_GUESS RESTART
    EPS_SCF 1.0E-06
    MAX_SCF 20
    &OT ON
        MINIMIZER DIIS
        PRECONDITIONER FULL_SINGLE_INVERSE
        ENERGY_GAP 0.1
    &END OT
    &OUTER_SCF
        MAX_SCF 2
    &END OUTER_SCF
&END SCF
```

The FULL\_ALL preconditioner is expensive to apply to large systems (diagonalization of the approximate Hamiltonian is required). The FULL\_SINGLE\_INVERSE is pretty good, and much cheaper for big systems.



## Choice of method

- Use OT if you can, it is faster and converges more reliably
  - disadvantage: requires additional calculations to obtain unoccupied states.
- Use Diagonalisation if you need to work with the occupation numbers of orbitals
  - Use Diagonalisation (with smearing) for metallic systems

The two methods use quite different code paths.

- EPS\_SCF has different meanings for OT (largest derivative of energy wrt MO coefficients) and TD (largest change in the density matrix).

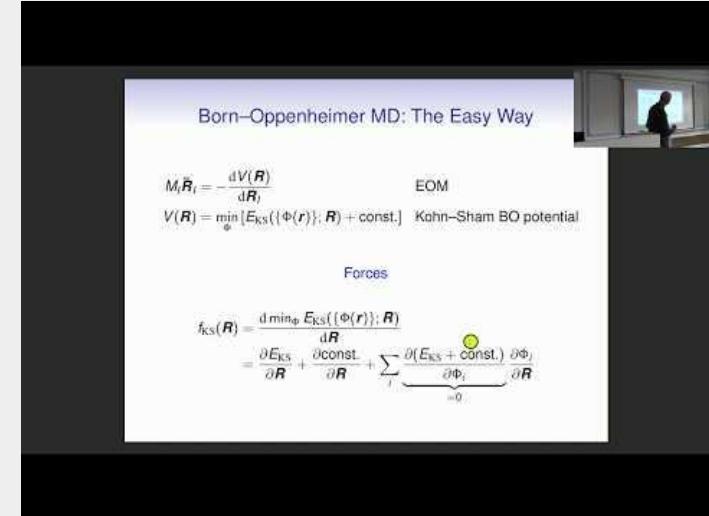
Some options will only work with either OT or TD.

- Most common is MO section in PRINT only works properly with diagonalisation
- MO\_CUBES section in PRINT only works properly with OT.

# Moving atoms

- Geometry optimization (run\_type geo\_opt)
- Cell optimization (run\_type cell\_opt)
- Extensive Molecular Dynamics capability
  - Detailed introduction and discussion of stable Born-Oppenheimer AIMD by Juerg Hutter ->
  - Many ensembles
  - Built in metadynamics
  - Link to PLUMED
- ASE and AIIA integration

All specified in the &Global section of the input

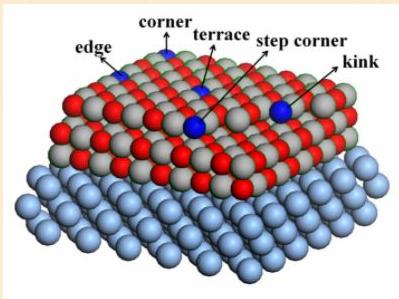


## *Test systems*

- Again files in /work/ta154/shared
- Run.slurm that needs adapting to have correct input names (commented out at the moment)
- H2O-gga.inp a simple water system with GGA functional
- H2O-hyb.inp the same system but running with a hybrid exchange functional (more details about the hybrid settings in a later session)
- Diag\_cu144\_broy.inp which is a metallic copper system.
- I'll run through these examples during the session including a look through the outputs.

## Back to the KS matrix - how to use non-local exchange in CP2K

Hybrid functionals and ADMM



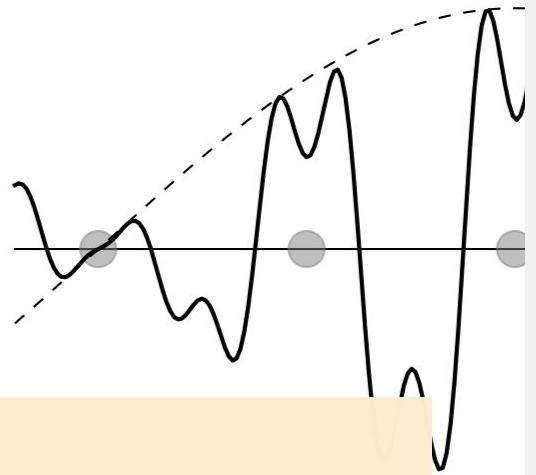
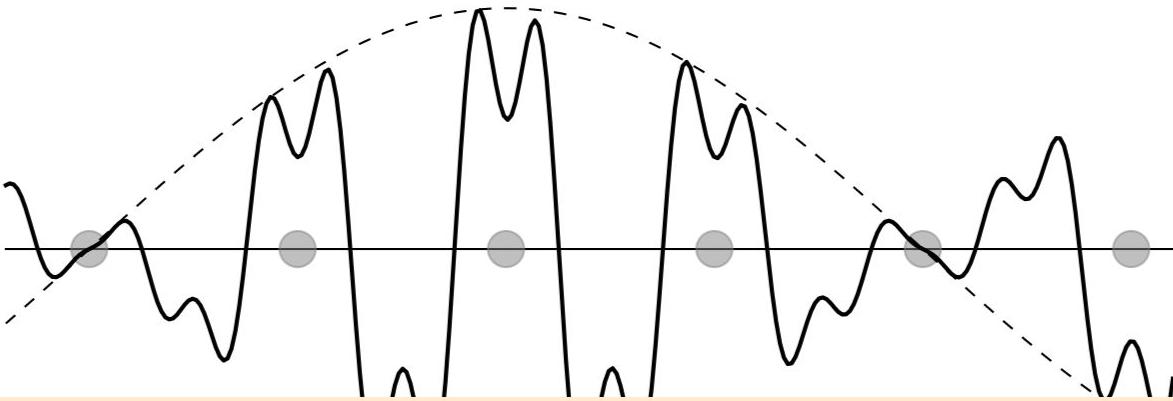
CP2k



UNIVERSITY OF  
LINCOLN

Matt Watkins

School of Mathematics and Physics, University of Lincoln, UK

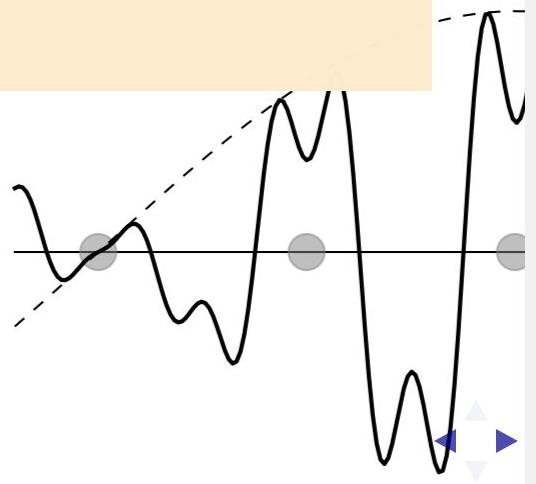
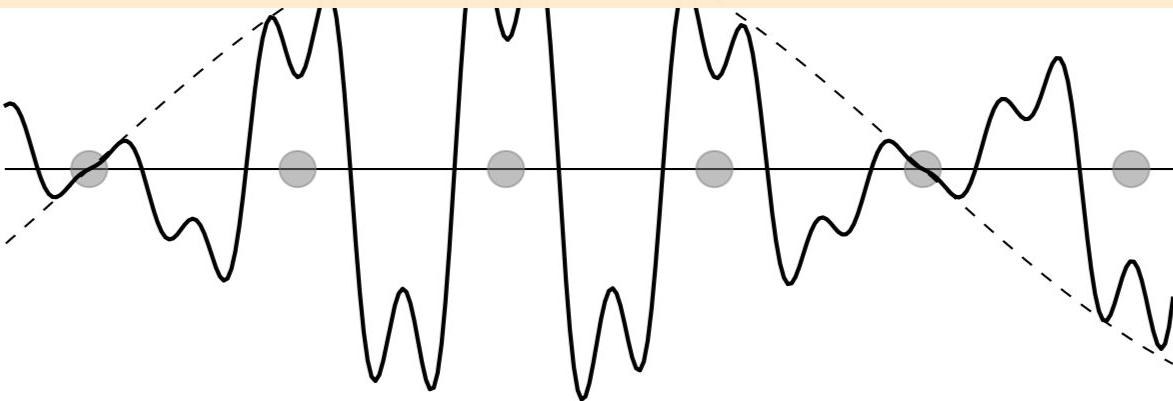


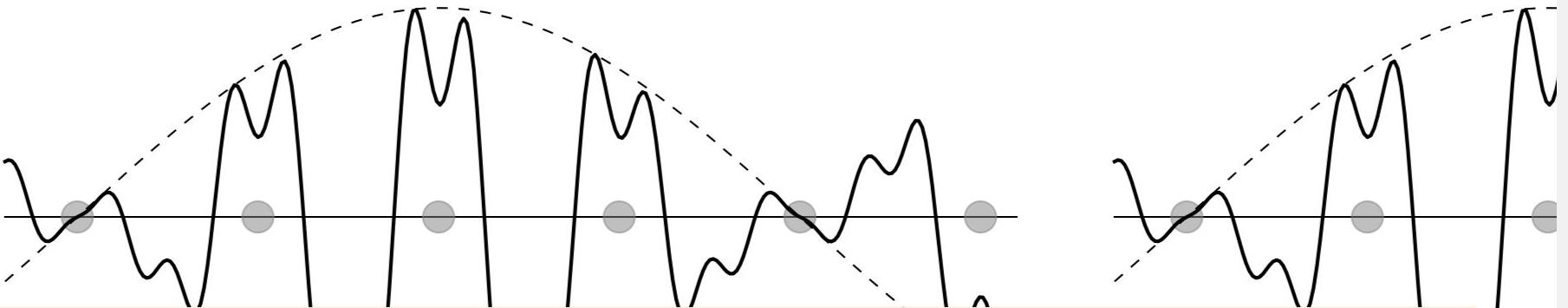
## What are hybrid functionals

Global Hybrid functions

mixing non-local Hartree-Fock exchange with semi-local DFT exchange

$$E_{XC} = \alpha E_X^{HFX}[\psi_i] + (1 - \alpha) E_X^{DFT}[\rho] + E_C^{DFT}[\rho]$$

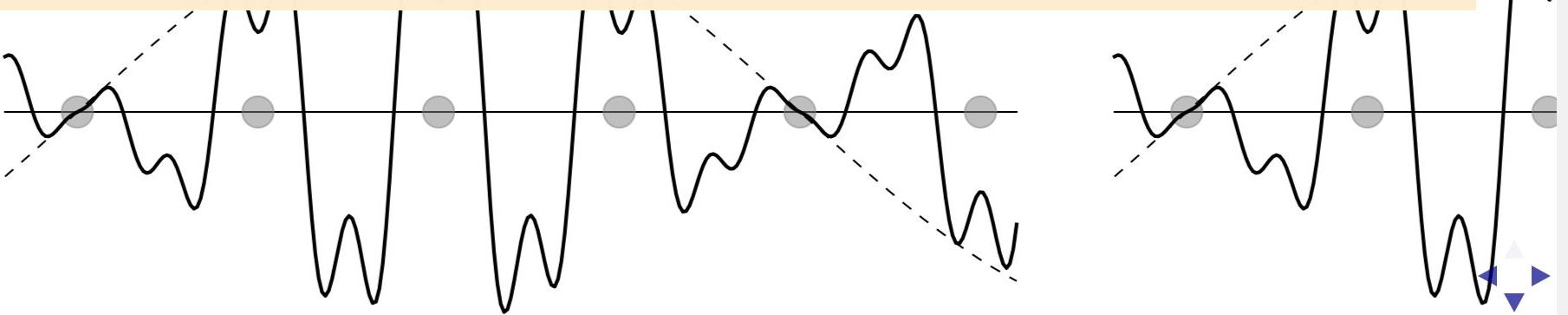




## Why do we need hybrids?

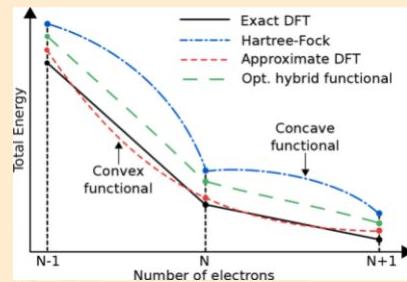
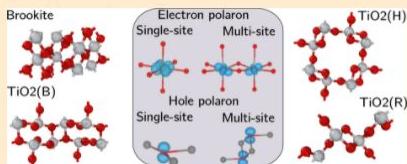
### Practical motivations

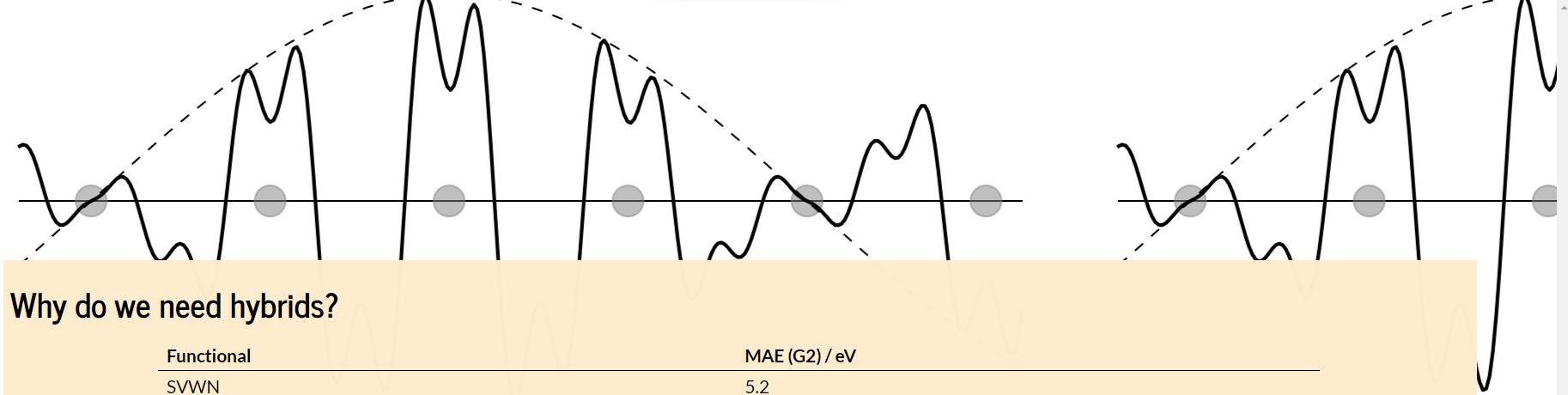
- Improved description of thermochemistry (e.g. atomizations energies, heats of formation)
- Improved description of lattice constants, surface energies, ionization potentials and band gaps
- Qualitatively correct description of localized and correlated states



## Theoretical motivations

- Adiabatic connection
- 1-e energy level discontinuity



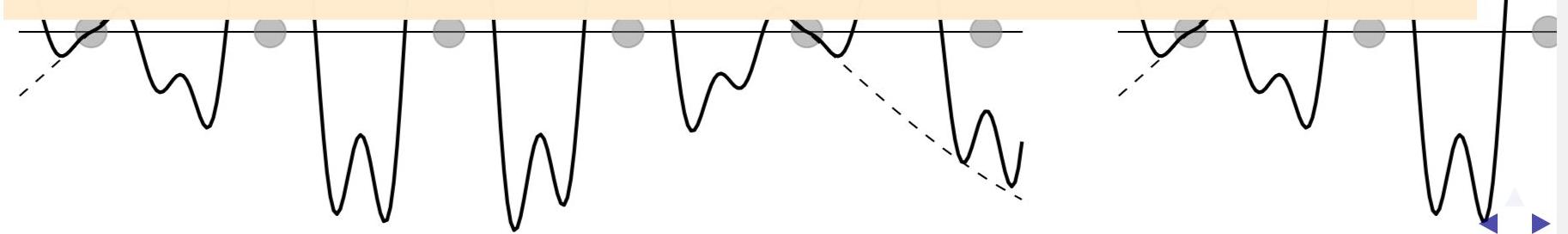


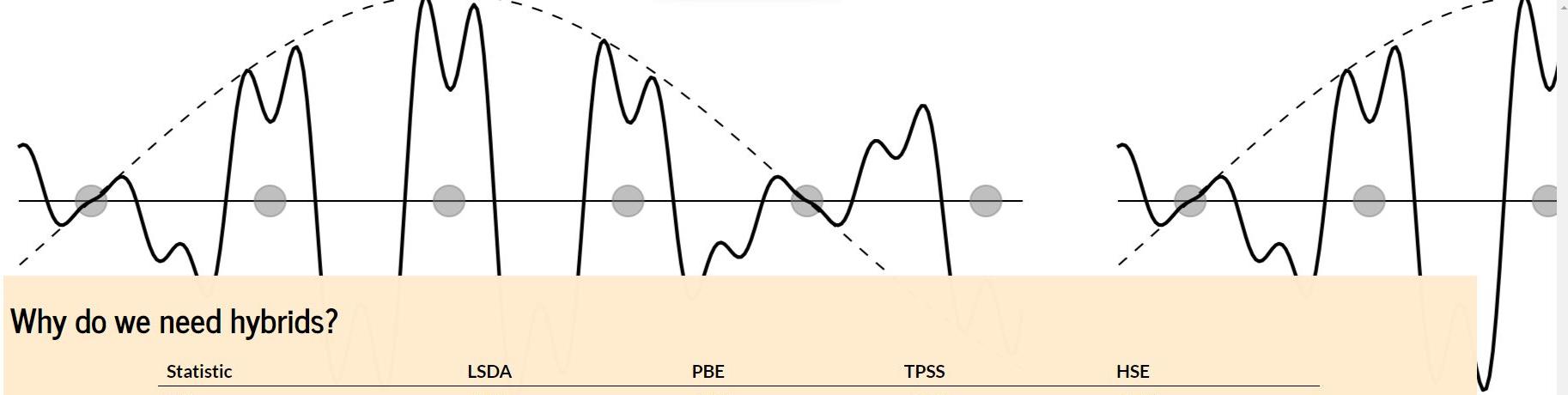
## Why do we need hybrids?

Functional	MAE (G2) / eV
SVWN	5.2
LSD (SVWN5)	3.6
PBE	0.73
BLYP	0.31
B3LYP	0.13
PBEO	0.21

Mean Absolute errors for atomization energies (eV)

Scuseria et al., J. Chem. Phys., 110, 5029 (1999)



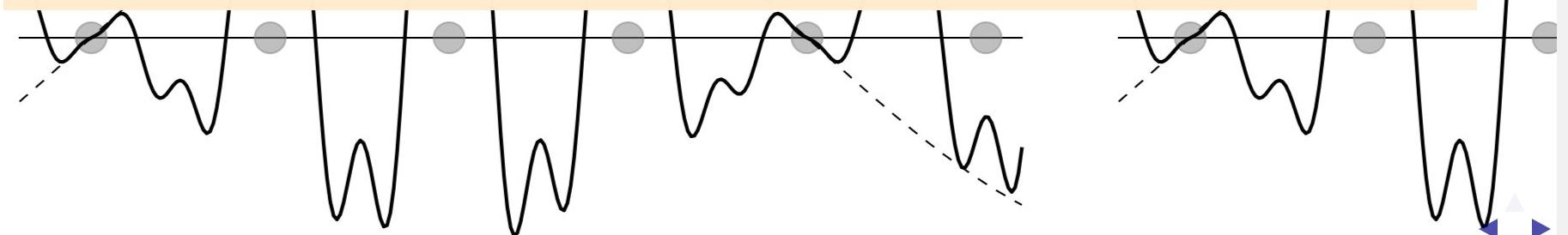


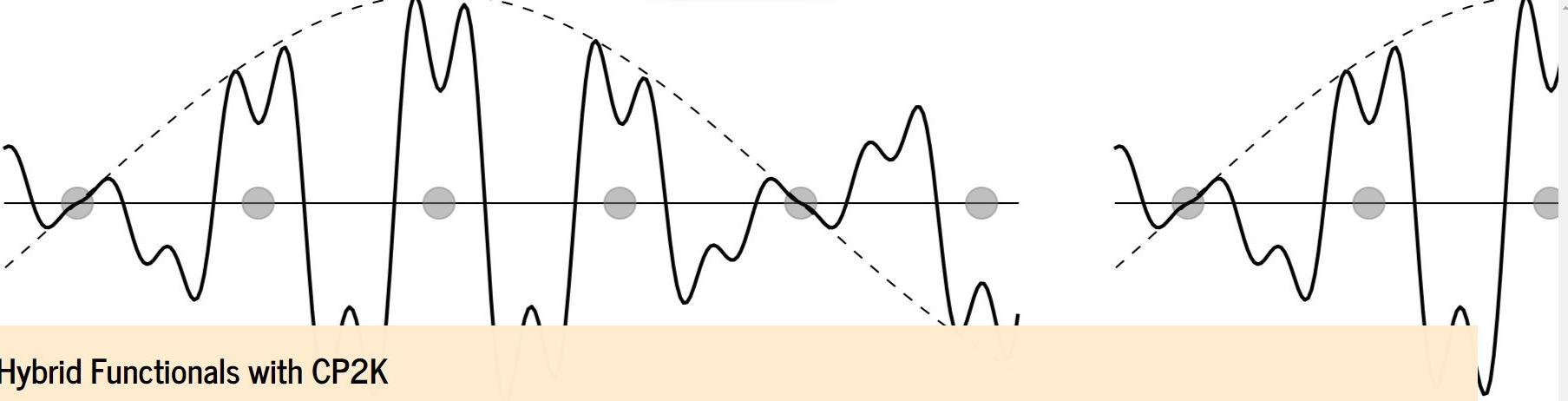
## Why do we need hybrids?

Statistic	LSDA	PBE	TPSS	HSE
ME	-1.14	-1.13	-0.98	-0.17
MAE	1.14	1.13	0.98	0.17
rms	1.24	1.25	1.12	0.34
Max (+)	-	-	-	0.32
Max (-)	-2.30	-2.88	-2.66	-0.72

Mean Absolute errors for Band Gaps (eV)

Scuseria et al., J. Chem. Phys., 123, 174101 (2005)





## Hybrid Functionals with CP2K

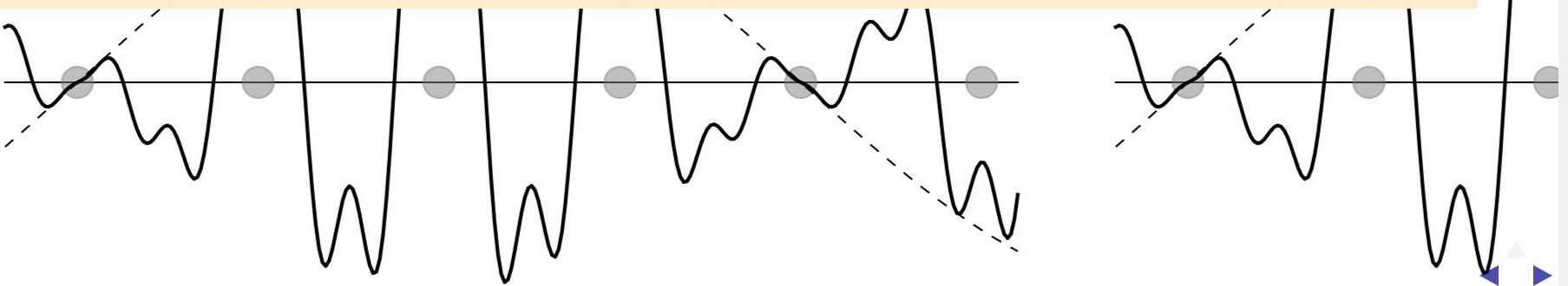
- GGA: total energy functional of the electron density

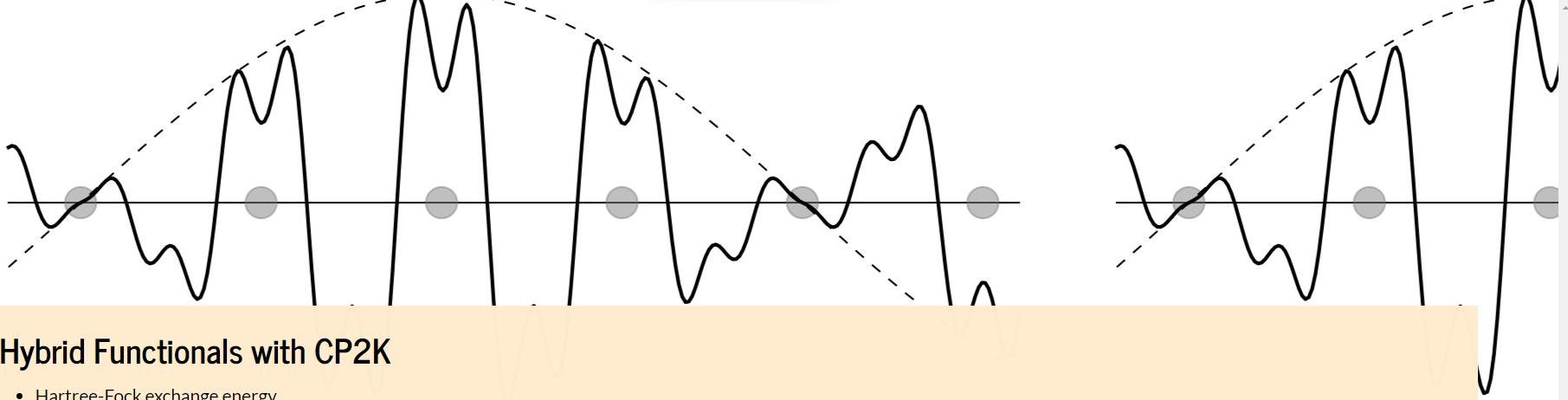
$$E[\rho] = T_S[\rho] + J[\rho] + E_{XC}[\rho] + \int v(\mathbf{r})\rho(\mathbf{r})d\mathbf{r}$$

- hybrid: functional of density and two particle density matrix (orbitals)

$$E_{XC} = \alpha E_X^{HFX}[\psi_i] + (1 - \alpha)E_X^{DFT}[\rho] + E_C^{DFT}[\rho]$$

Guidon, Hutter and VandeVondele, J. Chem. Theory Comput., 6, 2348 (2010)





## Hybrid Functionals with CP2K

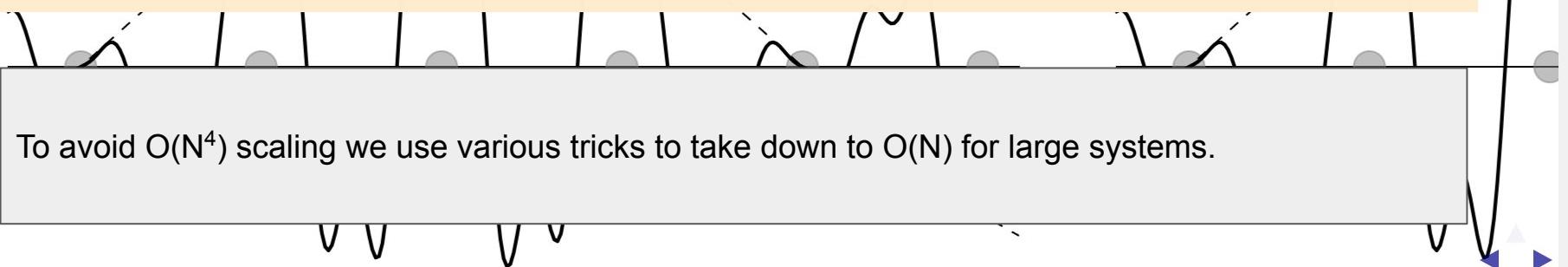
- Hartree-Fock exchange energy

$$E_X^{HFX}[P] = -\frac{1}{2} \sum_{\lambda\sigma\mu\nu} P^{\mu\sigma} P^{\nu\sigma} (\mu\nu|\lambda\sigma)$$

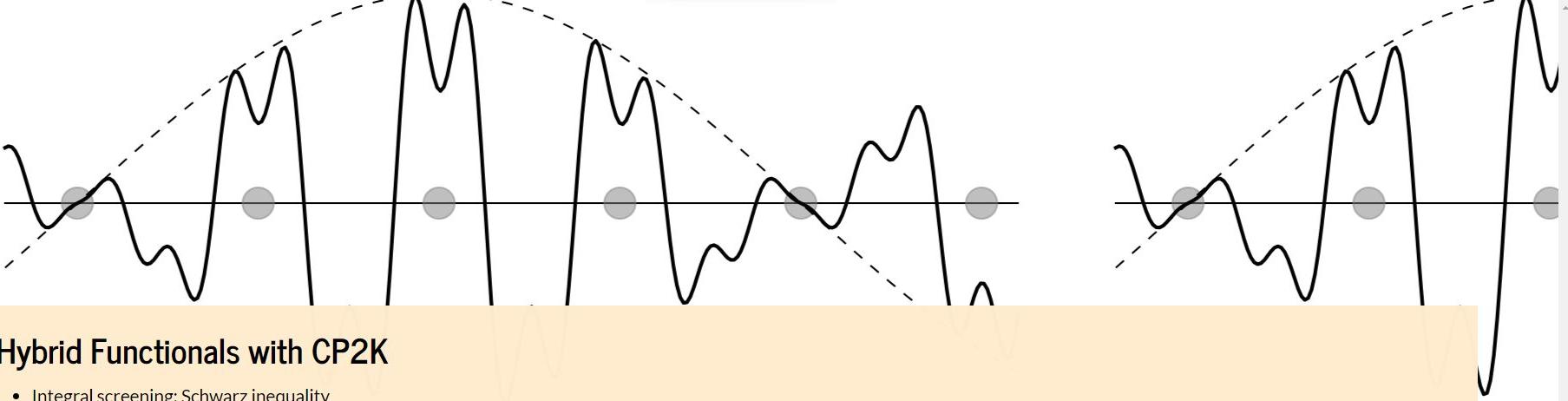
$$P^{\mu\nu} = \sum_i C^{\mu i} C^{\nu i} = CC^T$$

$$(\mu\nu|\lambda\sigma) = \int \int \phi_\mu(\mathbf{r}_1) \phi_\nu(\mathbf{r}_1) g(|\mathbf{r}_1 - \mathbf{r}_2|) \phi_\lambda(\mathbf{r}_2) \phi_\sigma(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2$$

Four-centre two-electron integrals (ERIs) in Mulliken notation: naively  $O(N^4)$  scaling



To avoid  $O(N^4)$  scaling we use various tricks to take down to  $O(N)$  for large systems.



## Hybrid Functionals with CP2K

- Integral screening: Schwarz inequality

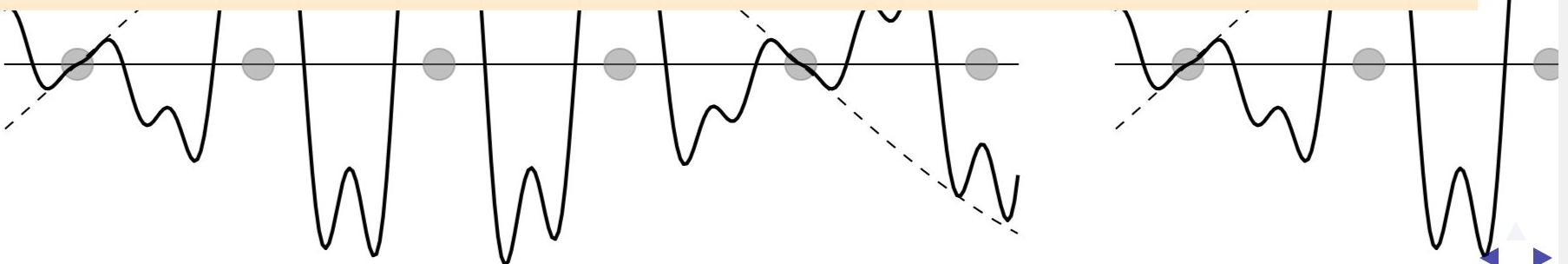
$$|(\mu\nu|\lambda\sigma)| \leq |(\mu\nu|\mu\nu)|^{1/2} |(\lambda\sigma|\lambda\sigma)|^{1/2}$$

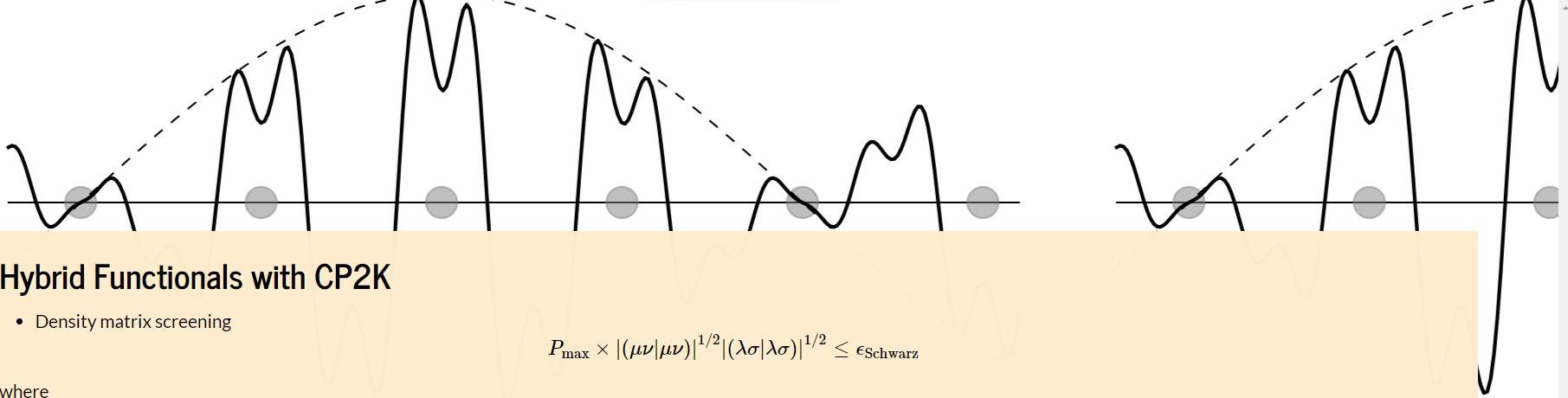
$$\frac{1}{8}O(N^4) \rightarrow O(N^2)$$

Ahlrichs et al., J. Comput. Chem., 10, 104 (1989)

(see [EPS SCHWARZ](#) keyword in [&SCREENING](#) subsection)

Guidon et al., J. Chem. Phys., 128, 214104 (2008)





## Hybrid Functionals with CP2K

- Density matrix screening

$$P_{\max} \times |(\mu\nu|\mu\nu)|^{1/2} |(\lambda\sigma|\lambda\sigma)|^{1/2} \leq \epsilon_{\text{Schwarz}}$$

where

$$P_{\max} = \max\{|P_{\mu\lambda}|, |P_{\mu\sigma}|, |P_{\nu\lambda}|, |P_{\nu\sigma}|\}$$

finally linear scaling

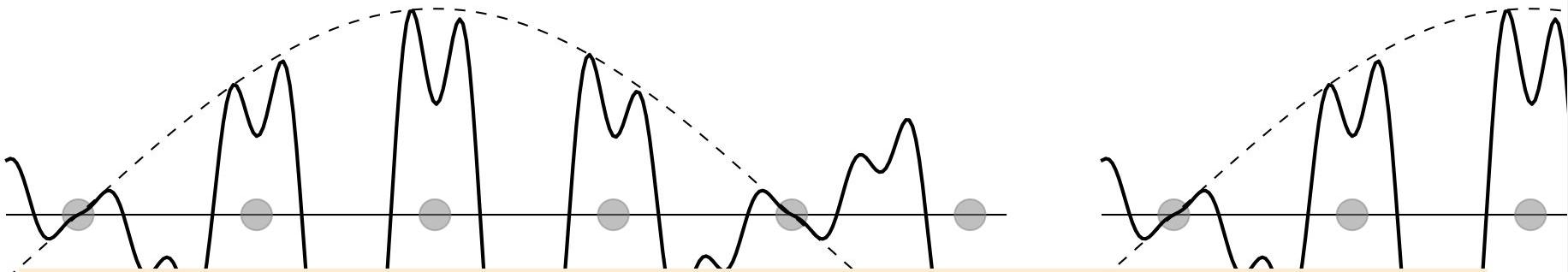
$$O(N^2) \rightarrow O(N)$$

(see [SCREEN ON INITIAL\\_P keyword in &SCREENING subsection](#))

Guidon et al., J. Chem. Phys., 128, 214104 (2008)

$P_{\max}$ : largest density matrix element; usually, the density matrix from the previous SCF iteration is used for the screening in the next iteration, use pre-converged GGA density matrix; very useful for DFT molecular dynamics simulations using hybrid functionals





## Interaction Potential

$$(\mu\nu|\lambda\sigma) = \int \int \phi_\mu(\mathbf{r}_1)\phi_\nu(\mathbf{r}_1)g(|\mathbf{r}_1 - \mathbf{r}_2|)\phi_\lambda(\mathbf{r}_2)\phi_\sigma(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2$$

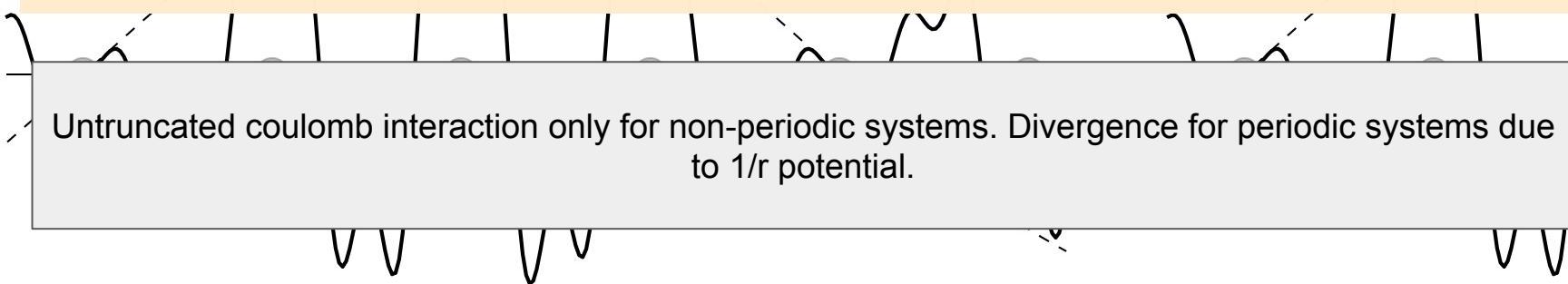
- Standard Coulomb potential gives a global hybrid

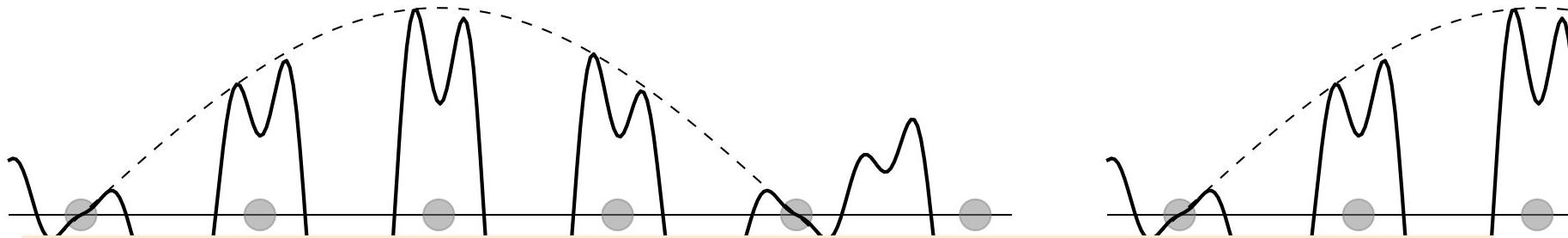
$$g(r) = \frac{1}{r}$$

(see [POTENTIAL TYPE](#) keyword in [&INTERACTION POTENTIAL](#) subsection)

Guidon et al., J. Chem. Phys., 128, 214104 (2008)

For gas-phase, this energy is easy to evaluate; for periodic boundary conditions, it suffers from the integrable singularity, different approaches have been developed to obtain good convergence with respect to the  $k$ -point sum.





## Interaction Potential

$$(\mu\nu|\lambda\sigma) = \int \int \phi_\mu(\mathbf{r}_1)\phi_\nu(\mathbf{r}_1)g(|\mathbf{r}_1 - \mathbf{r}_2|)\phi_\lambda(\mathbf{r}_2)\phi_\sigma(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2$$

- Range-separated Coulomb potential gives a range-separated hybrid

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

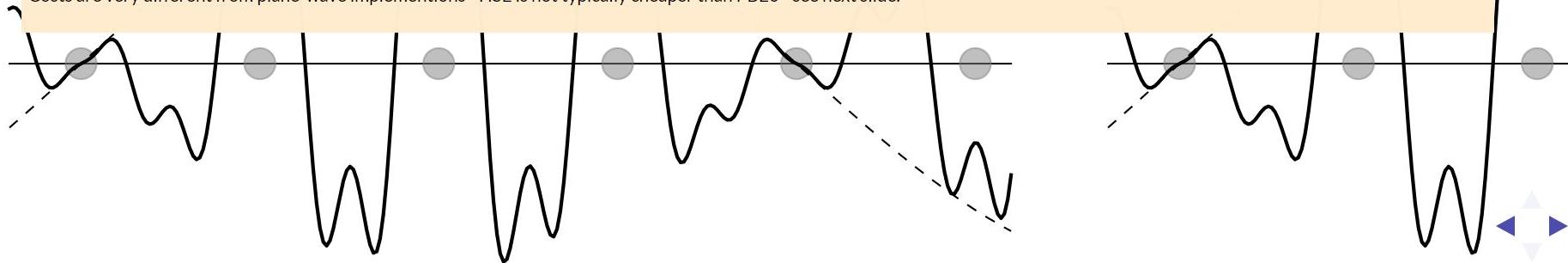
used in HSE06 etc.

(see [POTENTIAL TYPE](#) keyword in [&INTERACTION POTENTIAL](#) subsection)

Guidon et al., J. Chem. Phys., 128, 214104 (2008)

If omega is not large enough very large cells might be needed for HSE calculations at the  $\Gamma$  point.

Costs are very different from plane-wave implementations - HSE is not typically cheaper than PBE0 - see next slide.



## Interaction Potential

$$(\mu\nu|\lambda\sigma) = \int \int \phi_\mu(\mathbf{r}_1)\phi_\nu(\mathbf{r}_1)g(|\mathbf{r}_1 - \mathbf{r}_2|)\phi_\lambda(\mathbf{r}_2)\phi_\sigma(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2$$

- truncated Coulomb potential gives a global hybrid

$$g(r_{12}) = \begin{cases} \frac{1}{r_{12}} & \text{if } r_{12} \leq R_c \\ 0 & \text{otherwise} \end{cases}$$

used in the snappily titles PBE0-TC family of functionals.

$R_C$  is specified by the [CUTOFF RADIUS](#) keyword.

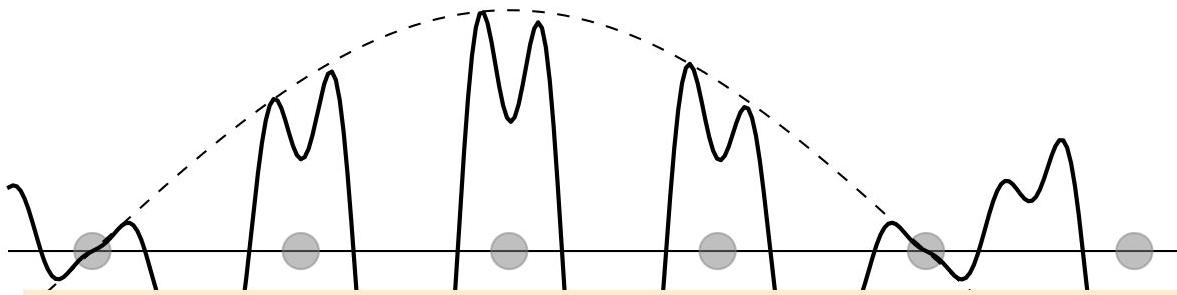
(see [POTENTIAL TYPE](#) keyword in [&INTERACTION POTENTIAL](#) subsection)

Guidon et al., J. Chem. Phys., 128, 214104 (2008)

Truncation is mainly for solid-state environments and avoids numerical problems with  $\Gamma$  point global hybrid calculations.

$R_C$  must be smaller than half the smallest cell dimension.

Requires a `T_C_G_DATA` file - by default in `/cp2k/data` and a modern installation should just find it.



## Interaction Potential

$$(\mu\nu|\lambda\sigma) = \int \int \phi_\mu(\mathbf{r}_1)\phi_\nu(\mathbf{r}_1)g(|\mathbf{r}_1 - \mathbf{r}_2|)\phi_\lambda(\mathbf{r}_2)\phi_\sigma(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2$$

- truncated Coulomb potential gives a global hybrid

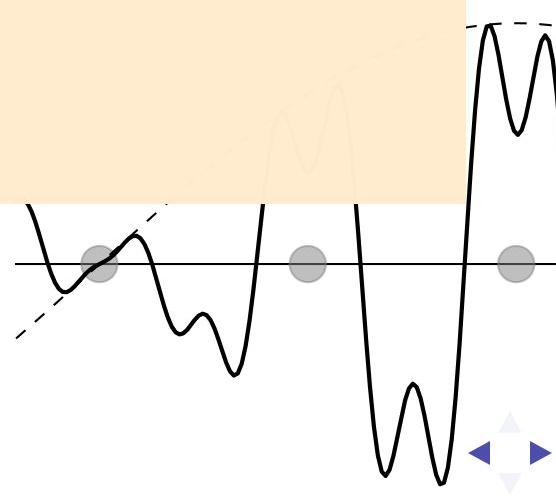
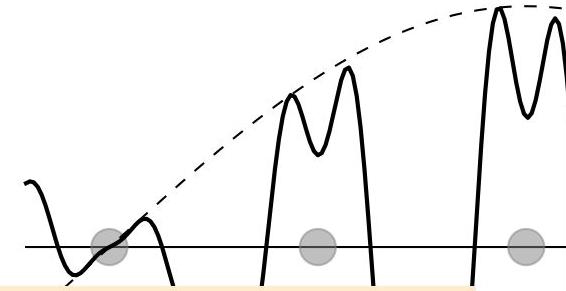
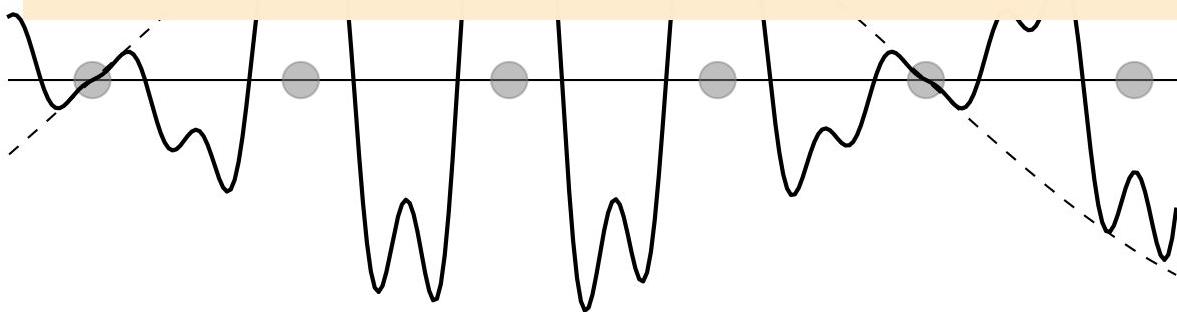
$$g(r_{12}) = \begin{cases} \frac{1}{r_{12}} & \text{if } r_{12} \leq R_c \\ 0 & \text{otherwise} \end{cases}$$

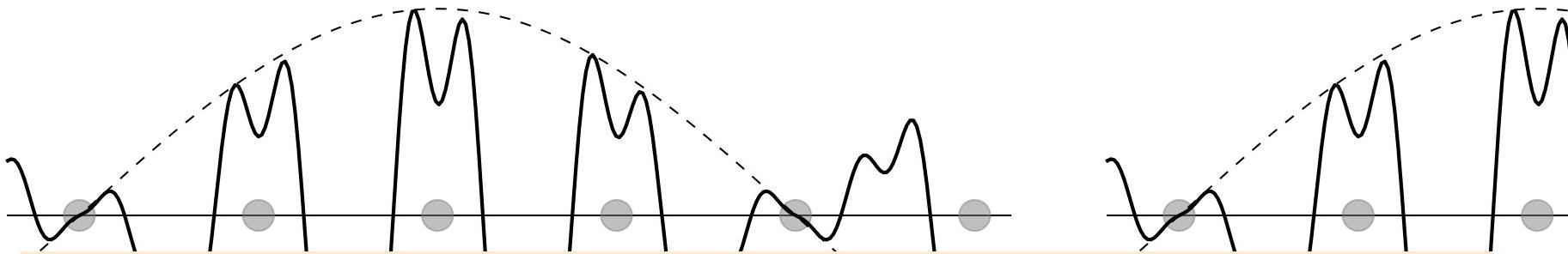
We can add in a long range correction, to replace the exchange energy lost by truncation.

See the [&PBE HOLE T C LR](#) subsection in the [&XC FUNCTIONAL](#) section.

This gives the even more snappily titled PBE0-TC-LRC family of functionals.

Guidon et al., J. Chem. Theory Comput., 5, 3010 (2008)





## Auxiliary Density Matrix Methods (ADMM)

$$E[\rho] = T_S[\rho] + J[\rho] + E_{XC}[\rho, P] + \int v(\mathbf{r})\rho(\mathbf{r})d\mathbf{r}$$

$$E_{XC} = \alpha E_X^{HFX}[P] + (1 - \alpha) E_X^{DFT}[\rho] + E_C^{DFT}[\rho]$$

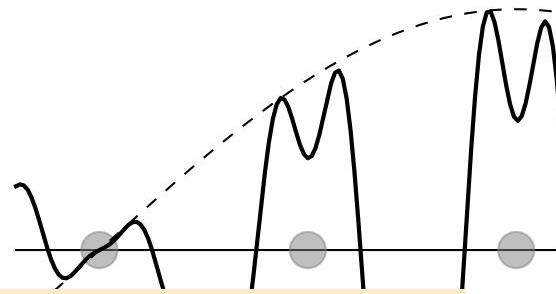
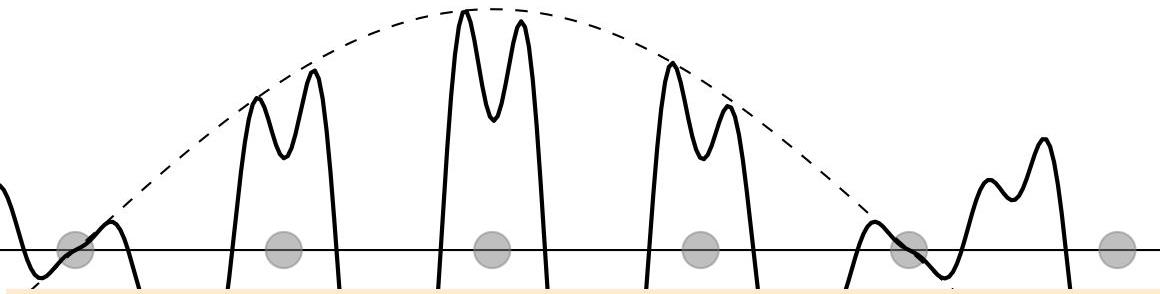
$$E_X^{HFX}[P] = -\frac{1}{2} \sum_{\lambda\sigma\mu\nu} P^{\mu\sigma} P^{\nu\sigma} (\mu\nu|\lambda\sigma)$$

introduce auxiliary density matrix  $\hat{P} \approx P$

$$\begin{aligned} E_X^{HFX}[P] &= E_X^{HFX}[\hat{P}] + E_X^{HFX}[P] - E_X^{HFX}[\hat{P}] \\ &\approx E_X^{HFX}[\hat{P}] + E_X^{DFT}[P] - E_X^{DFT}[\hat{P}] \end{aligned}$$

Guidon, Hutter and VandeVondele, J. Chem. Theory Comput., 6, 2348 (2010)





## ADMM basis sets

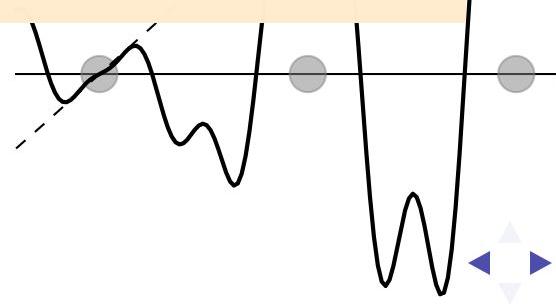
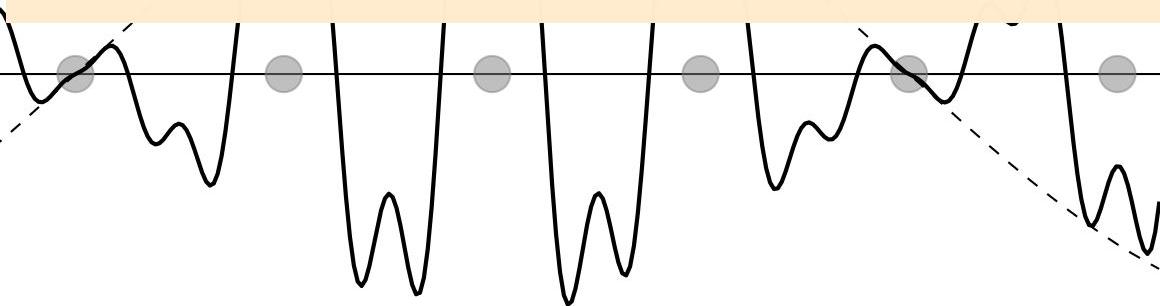
Initially available from H to Cl optimized against atomic calculations

Choice of auxiliary basis set for ADMM

- FIT3: three Gaussian exponents for each valence orbital
- cFIT3: a contraction of FIT3
- pFIT3: FIT3 + polarization functions
- cpFIT3: cFIT3 + polarization functions
- aug-FIT3, aug-cFIT3, aug-pFIT3, aug-cpFIT3: augmented with a "diffuse" function

(see \$CP2K/cp2k/data/BASIS\_ADMM)

Guidon, Hutter and VandeVondele, J. Chem. Theory Comput., 6, 2348 (2010)



## ADMM basis sets

Uncontracted basis sets for transition metals:

- FIT10: 4s + 3p + 3d
- FIT11: 4s + 3p + 3d + 1f
- FIT12: 4s + 3p + 4d + 1f
- FIT13: 4s + 4p + 4d + 1f

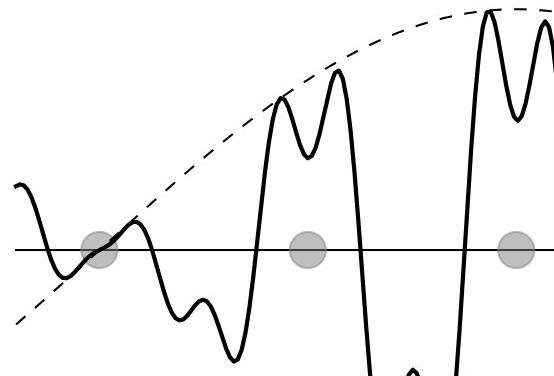
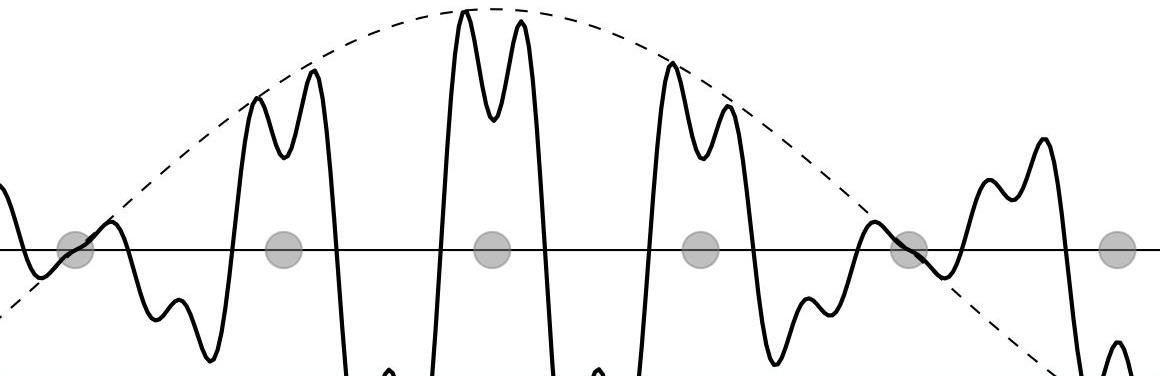
Contracted basis sets (double- $\zeta$  quality)

- cFIT10 / cFIT11 / cFIT12 / cFIT13

Names of the ADMM basis sets for main group elements will differ slightly, and usually the first ADMM basis set will not contain polarisation function.

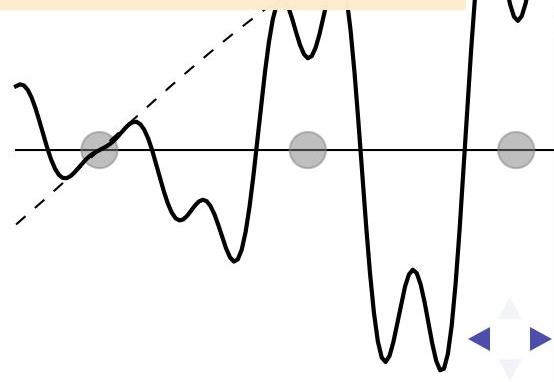
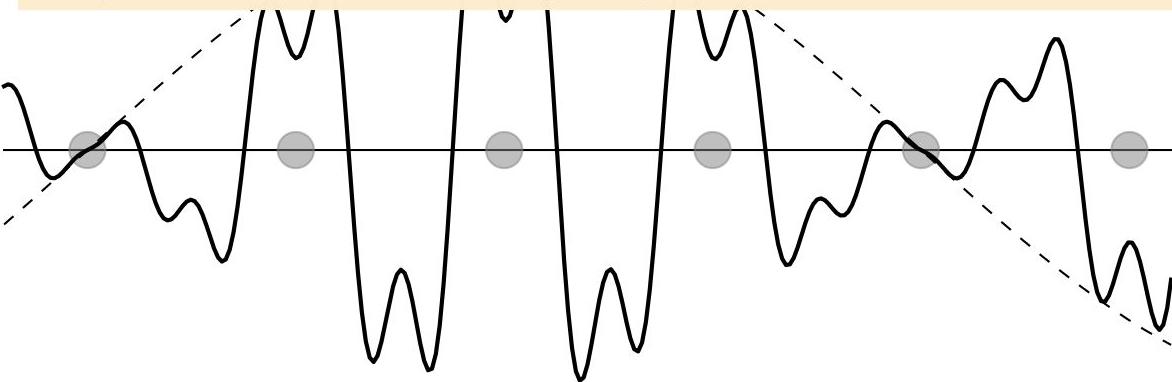
(see \$CP2K/cp2k/data/BASIS\_ADMM\_MOLOPT)

Ling & Slater, unpublished



## General remarks

- Always check the convergence of CUTOFF
- Always check the convergence of properties (e.g. lattice parameters, band gaps) with respect to supercell sizes
- Always start from pre-converged GGA (e.g. PBE) wavefunction and geometry
- Always check the convergence of primary and ADMM basis sets
  - start from a small basis and gradually increase the size
- ADMM has only been implemented for use with GPW
- Only ADMM1 will work with some other functionality (smearing, TDDFPT)



## Example: Diamond band-gap

method	number of integrals	gap [eV]
PBE (PBS)		4.17
PBE (ABS)		4.37
PBEO (PBS)	40,787,850,778,591	6.07
PBEO (ABS)	23,561,509,497	6.25
PBEO ADMM1	24,816,897,009	6.03
PBEO ADMM2	24,795,460,638	6.02

Guidon, Hutter and VandeVondele, J. Chem. Theory Comput., 6, 2348 (2010)



## - Example: Silicon band-gap

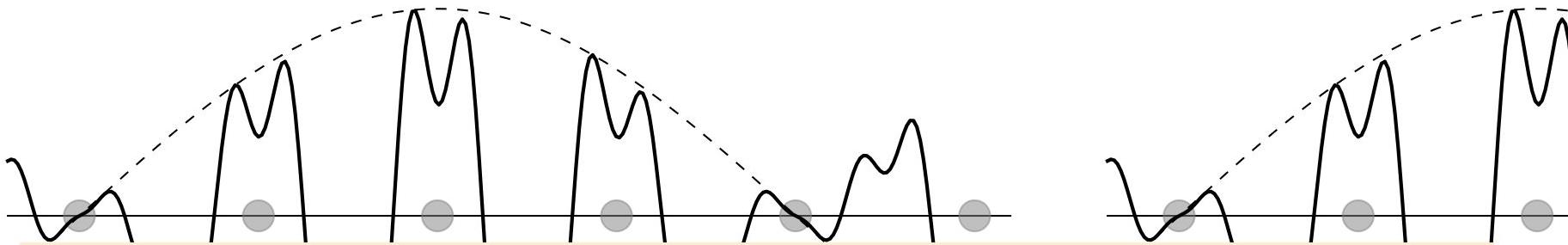
Cutoff Radius [nm]	Band gap [eV]	# Integrals
0.2	1.16 <sup>a</sup>	77799946176 <sup>a</sup>
0.4	1.54 <sup>a</sup>	154325979000 <sup>a</sup>
0.6	1.71 <sup>a</sup>	265868148312 <sup>a</sup>
0.8	1.78 <sup>a</sup>	422457823080 <sup>a</sup>

PBE0-TC-LRC with cFIT3 ADMM basis, 3x3x3 supercell, 216 atoms

ADMM basis	Band gap [eV]	# Integrals
cFIT3	1.16 <sup>a</sup>	422457823080 <sup>a</sup>
FIT3	1.80 <sup>a</sup>	424426850352 <sup>a</sup>
pFIT3	1.98 <sup>a</sup>	1447428361680 <sup>a</sup>
Ref. (VASP)	1.93 <sup>b</sup>	

PBE0-TC-LRC with 8 Å cutoff radius, 3x3x3 supercell, 216 atoms

- <sup>a</sup> Ling & Slater, unpublished; <sup>b</sup> Paier et al., J. Chem. Phys. 124, 154709 (2006)



## Functionals

- PBE0-TC-LRC, Long Range Corrected truncated global hybrid

$$E_{XC}^{PBE0-TC-LRC} = aE_X^{HF,TC}(R_c) + aE_X^{PBE,LRC}(R_c) + (1-a)E_X^{PBE} + E_C^{PBE}$$

Guidon et al., J. Chem. Theory Comput., 5, 3010 (2008) Spencer & Alavi, Phys. Rev. B, 77, 193110 (2008)

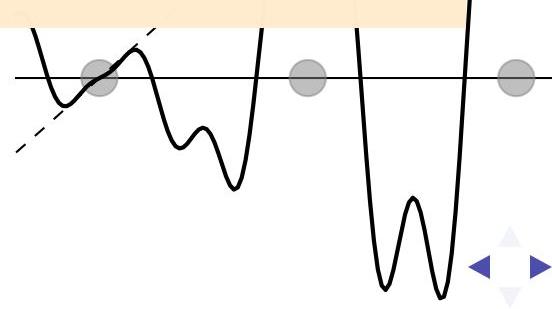
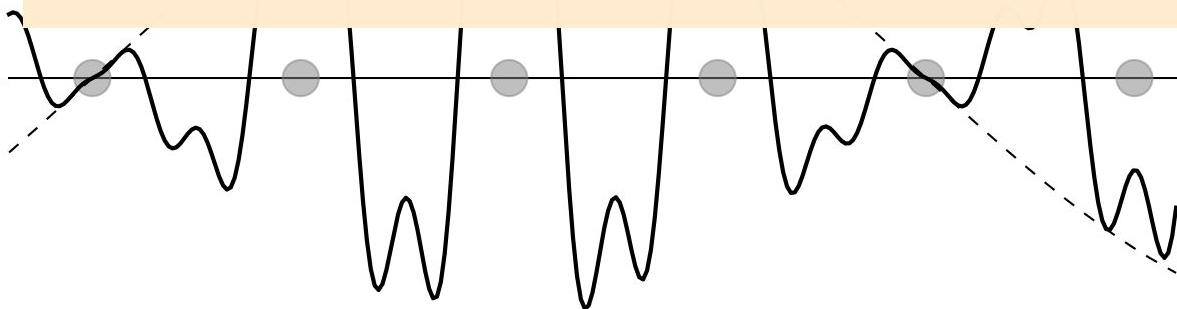
- HSE, Range separated hybrid

$$E_{XC}^{HSE} = aE_X^{HF,SR}(\omega) + (1-a)E_X^{PBE,SR}(\omega) + E_X^{PBE,LR}(\omega) + E_C^{PBE}$$

Scuseria et al., J. Chem. Phys., 125, 224106 (2006)

"Empirical" parameters:  $a$ ,  $R_c$ , and  $\omega$

(note the 0 in PBE0 stands for 0 empirical parameters...)



## - Input structure

Sketch of the key sections for running an ADMM calculation

```
&DFT
  ...
  BASIS_SET_FILE_NAME ./BASIS_MOLOPT
  BASIS_SET_FILE_NAME ./BASIS_ADMM
  WFN_RESTART_FILE_NAME ${project}-RESTART.wfn
  ...
&SCF
  ...
  SCF_GUESS RESTART
  ...
&END SCF
&AUXILIARY_DENSITY_MATRIX_METHOD
  METHOD BASIS_PROJECTION
  ADMM_PURIFICATION_METHOD MO_DIAG
&END AUXILIARY_DENSITY_MATRIX_METHOD
  ...
&XC
  ...
&END XC
```

- Note: The syntax for the AUX basis set changed (after 4.1?) before that it would be AUX\_FIT\_BASIS\_SET \*\*\*

## PBEO vs HSE

- Input structure (see examples in \$CP2K/cp2k/tests/QS/regtest-admm-1/2/3/4)

```
&XC
  &XC_FUNCTIONAL
    &PBE
      SCALE_X 0.75
      SCALE_C 1.0
    &END PBE
    &PBE_HOLE_T_C_LR
      CUTOFF_RADIUS 2.0
      SCALE_X 0.25
    &END PBE_HOLE_T_C_LR
  &END XC_FUNCTIONAL
  &HF
    &SCREENING
      EPS_SCHWARZ 1.0E-6
      SCREEN_ON_INITIAL_P FALSE
    &END SCREENING
    &INTERACTION_POTENTIAL
      POTENTIAL_TYPE TRUNCATED
      CUTOFF_RADIUS 2.0
```

```
&XC
  &XC_FUNCTIONAL
    &PBE
      SCALE_X 0.0
      SCALE_C 1.0
    &END PBE
    &XWPBE
      SCALE_X -0.25
      SCALE_X0 1.0
      OMEGA 0.11
    &END XWPBE
  &END XC_FUNCTIONAL
  &HF
    &SCREENING
      EPS_SCHWARZ 1.0E-6
      SCREEN_ON_INITIAL_P FALSE
    &END SCREENING
    &INTERACTION_POTENTIAL
      POTENTIAL_TYPE SHORTRANGE
```

## Output

Step	Update method	Time	Convergence	Total energy	Change
HFX_MEM_INFO	Est. max. program size before HFX [MiB]:			563	
***	WARNING in hfx_energy_potential.F:600 :: The Kohn Sham matrix is not	***			
***	100% occupied. This may result in incorrect Hartree-Fock results. Try	***			
***	to decrease EPS_PGF_ORB and EPS_FILTER_MATRIX in the QS section. For	***			
***	more information see FAQ: <a href="https://www.cp2k.org/faq:hfx_eps_warning">https://www.cp2k.org/faq:hfx_eps_warning</a>	***			
HFX_MEM_INFO	Number of cart. primitive ERI's calculated:	218851035670			
HFX_MEM_INFO	Number of sph. ERI's calculated:	152193561473			
HFX_MEM_INFO	Number of sph. ERI's stored in-core:	22711518963			
HFX_MEM_INFO	Number of sph. ERI's stored on disk:	0			
HFX_MEM_INFO	Number of sph. ERI's calculated on the fly:	0			
HFX_MEM_INFO	Total memory consumption ERI's RAM [MiB]:	21290			
HFX_MEM_INFO	Whereof max-vals [MiB]:	1516			
HFX_MEM_INFO	Total compression factor ERI's RAM:	8.14			
HFX_MEM_INFO	Total memory consumption ERI's disk [MiB]:	0			
HFX_MEM_INFO	Total compression factor ERI's disk:	0.00			
HFX_MEM_INFO	Size of density/Fock matrix [MiB]:	764			

Number of sph. ERI's calculated on the fly: should ideally be zero.

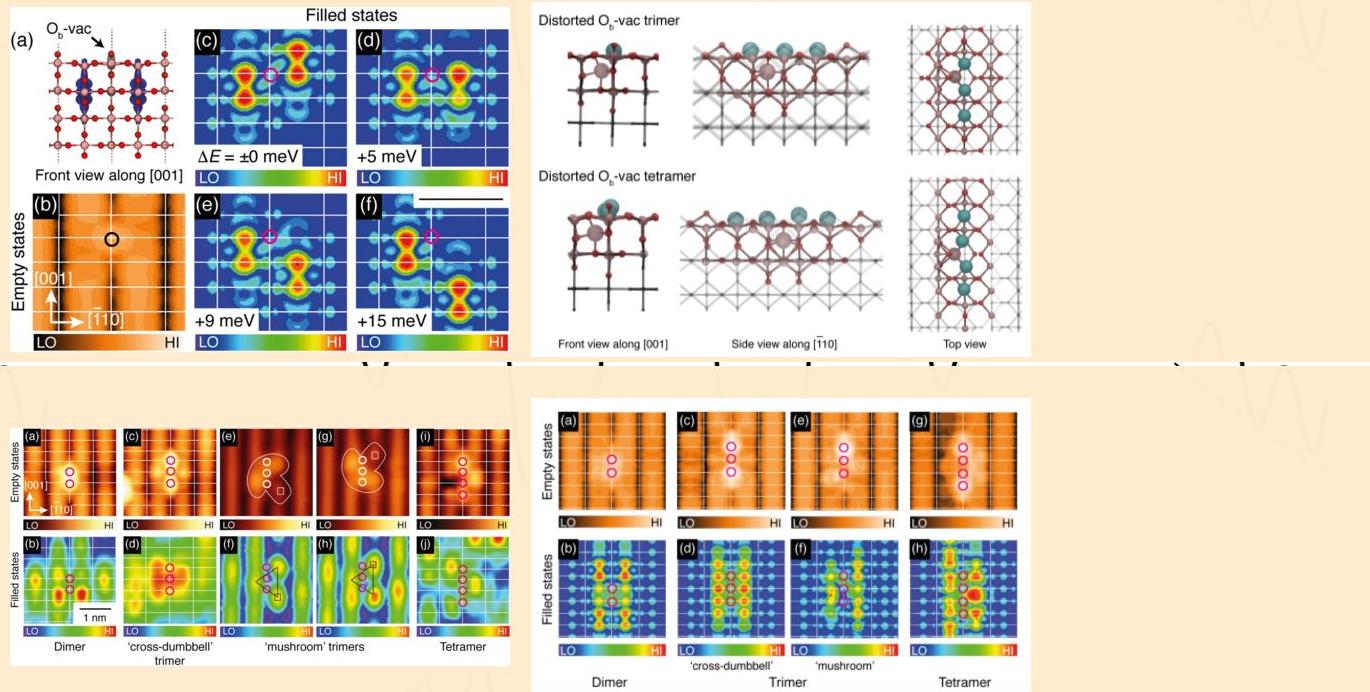
We want to keep ERIs in memory during the SCF loop.

## Extra remarks

- Libint library is used to calculate ERIs and their analytic derivatives, see `$CP2K/cp2k/INSTALL` for more details on installation and linking with CP2K
- Hybrid DFT calculations of large systems are memory demanding, try to increase `MAX_MEMORY` or run the job with more MPI processes. Note `MAX_MEMORY` is the memory per MPI process for ERIs, you must leave space for operating system and rest of the CP2K calculation.
- For extremely large hybrid DFT calculations, try hybrid MPI/OpenMP binary (i.e. `cp2k.psmp`)

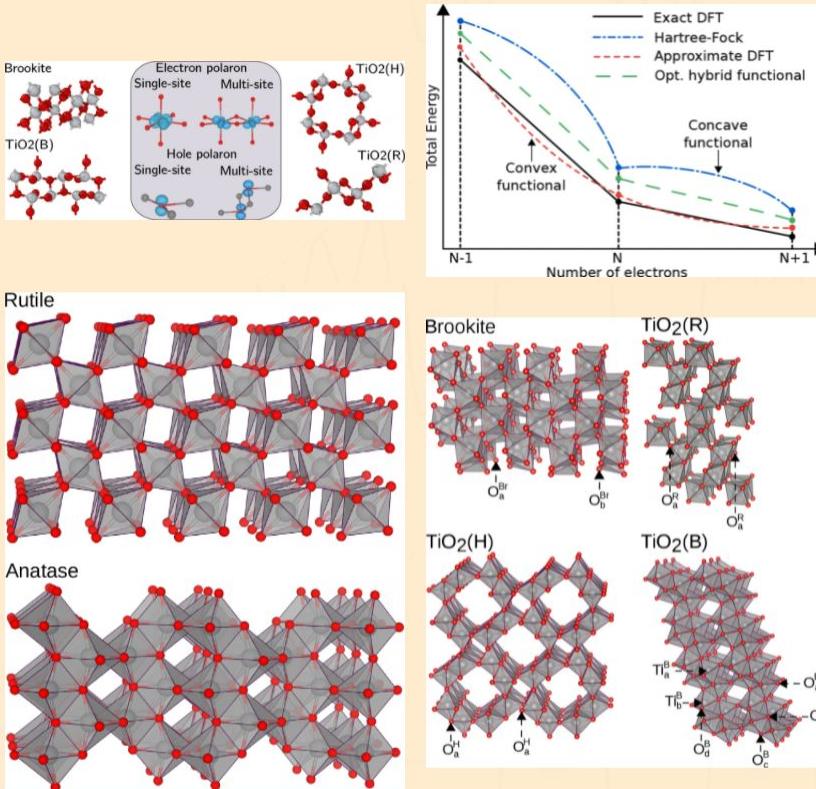
# Examples

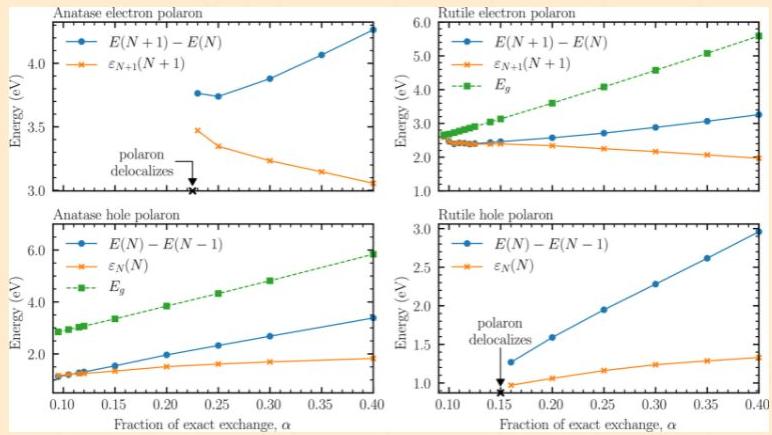
TiO<sub>2</sub> is everyone's favourite material - hybrid calculations with ~1000 atoms and good basis sets using CP2K



Yim et al, Phys. Rev. Lett. 117, 116402 (2016)

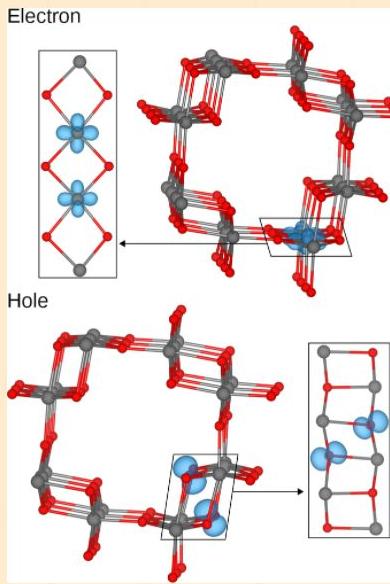
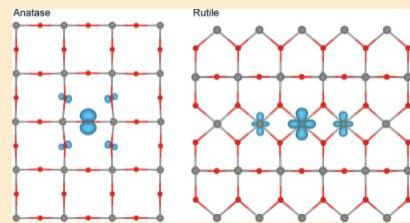
## TiO<sub>2</sub>-Polarons





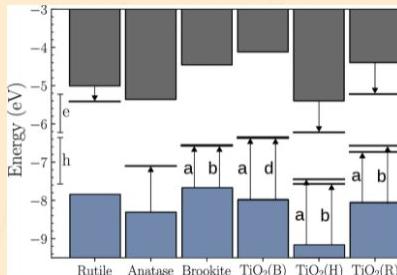
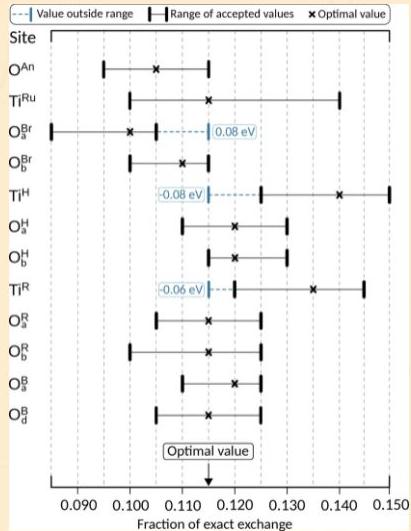
A. R. Elmaslmane et al, J. Chem. Theory Comput., 2018, 14, 3740

## TiO<sub>2</sub>-Polarons



A. R. Elmaslmane et al, J. Chem. Theory Comput., 2018, 14, 3740

## TiO<sub>2</sub>-Polarons



## Up Jacob's Ladder

### Beyond hybrid functionals

Hybrid functions use information from the occupied orbitals - if we use information from the unoccupied orbitals we can construct more advanced functionals - double hybrids, RPA.

Hybrid / ADMM Exercises [here](#)

Note: The syntax for the AUX basis set changed, use 'BASIS\_SET AUX\_FIT cFIT3' instead of 'AUX\_FIT\_BASIS\_SET \*\*\*' in the example.

### Post Hartree Fock wavefunction correlation

CP2K also implements several post-Hartree Fock correlation methods, [MP2](#), [RPA](#), [GW](#)

MP2 exercises [here](#)

GW exercises [here](#)

## Thanks

- Sanliang Ling, University of Nottingham (these slides are basically his reformatted)
- Sanliang's notes also have details of using dispersion corrected functionals [here](#)
- CP2K developers

[link to dispersion corrected functionals notes](#)

# Some Recent developments

## *Updated QMMM methods*

- GEEP method in CP2K - periodic embedding of DFT into electrostatic environment of classical system.
- Recently coupled to Gromacs code as part of [BioExcel](#) project - [ARCHER2 course](#)

## *New fully Linear scaling methods*

- [Massively scaling semi-empirical methods](#)

## *TDDFT and perturbation methods*

- Periodic TDDFT gradients under review
- XAS and [core-level spectroscopy](#)

## *Neural Net potentials*

E.g [Schran et al](#), [Lan et al](#)

# Resources

- Website - [cp2k.org](http://cp2k.org)
- User forum - <https://groups.google.com/g/cp2k>
- Videos - <https://www.youtube.com/c/Cp2kOrgChannel>
- Howtos - <https://www.cp2k.org/howto>

