Hartree-Fock exchange and hybrid functionals in ONETEP

James C. Womack

University of Southampton, UK

ONETEP masterclass 2019

University of Warwick
Preamble

Slides will soon be available online at http://www.onetep.org/Main/Workshops

See slides online for bibliography
Hartree-Fock exchange and hybrid functionals

- Hybrid functionals depend on \( \{\psi_i\} \) via \( E_{\text{HFX}} \)
- HFX corrects self-interaction error in Coulomb energy
- The most accurate XC functionals include HFX [1]

\[
E_{\text{HFX}} = - \sum_{i=1,j=1}^{N_{\text{MO}}} z_i z_j (\psi_i \psi_j | \psi_j \psi_i) \\
\int dr dr' \psi_i^*(r) \psi_j^*(r') \frac{1}{|r - r'|} \psi_j(r) \psi_i(r')
\]
Global hybrid functionals

Combine (semi-)local XC and HFX in a constant ratio over all space

e.g. Becke 3-parameter hybrids [4, 5], such as B3LYP [6]

\[ E_{xc}^{B3hyb} = E_{xc}^{LDA} + a_0 \left( E_x^{HFX} - E_x^{LDA} \right) + a_x \Delta E_x^{B88} + a_c \Delta E_c^{GGA} \]

Range-separated hybrids are under development in ONETEP
Linear-scaling Hartree-Fock exchange (LS-HFX)

HFX is challenging for $O(N)$ DFT because ERIs are non-local

- (Semi-)local XC does not require Electron Repulsion Integrals
- BUT ERIs are needed to evaluate HFX for hybrids...

$$(\varphi_\alpha \varphi_\delta | \varphi_\beta \varphi_\gamma)$$

$$\int d\mathbf{r} d\mathbf{r}' \varphi^*_\alpha(\mathbf{r})\varphi_\delta(\mathbf{r}) \frac{1}{|\mathbf{r} - \mathbf{r}'|} \varphi^*_\beta(\mathbf{r}')\varphi_\gamma(\mathbf{r}')$$
LS-HFX with NGWFs

The strictly localized NGWF basis presents unique challenges

\[ E_x^{\text{HFX}} = -K^{\beta\alpha}(\varphi_\alpha \varphi_\delta | \varphi_\beta \varphi_\gamma)K^{\delta\gamma} = -K^{\beta\alpha}X_{\alpha\beta} \]

- NGWFs produce sparse matrices & simplified integrals

\[ \int \left[ \begin{array}{c} \alpha \\ A \\ \bullet \\ \delta \\ D \end{array} \right] \text{dr} \neq 0 \quad \int \left[ \begin{array}{c} \beta \\ B \bullet \quad \gamma \\ C \end{array} \right] \text{dr} = 0 \]

- BUT cannot use conventional ERI evaluation methods
- Straightforward HFX evaluation (via FFTs) is impractical [7, 8]

An alternative scheme is necessary!
LS-HFX with NGWFs

Defining features of the scheme [7]

\[ X_{\alpha\beta} = (\varphi_\alpha \varphi_\delta | f_p V^{pq} (f_q | \varphi_\beta \varphi_\gamma)) K^{\delta\gamma} \]

- Spherical wave (SW) [9]
- Resolution-of-the-identity (RI)
- Distance-based exchange cutoff

Expand NGWF products in SWs

- Strictly localized in a sphere
- Analytic Coulomb potentials
- “2-centre” fitting scheme

\[ f_p(r) = \begin{cases} j_{lp}(q_p r) Z_{lp}m_p (\hat{r}) & r < a \\ 0 & r \geq a \end{cases} \]
Demonstration of $O(N)$ scaling HFX

Polyethylene chains
$H-(\text{CH}_2)_n-H$

Cost in core-hours
$N_{\text{CPU}} \times \text{walltime}$

Data from 2013 [7]
Some improvements since

Linear-scaling is achieved, but with a large prefactor...
A few years later... 

Single-point DFT energy with and without HFX

57 atom monomer from organic photovoltaic polymer

BLYP (GGA) 0.3 h
B3LYP (hybrid) 62 h
5 nodes / 120 cores
“Thomas” MMM Hub

Before recent developments to significantly improve performance...
V is central to the SW resolution-of-the-identity

\[ V_{Ap,Bq} = \int dr \, f_p(r_A) g_q(r_B) \quad g_q(r) = \int dr' \, \frac{f_q(r')}{|r - r'|} \]

- Integrand is highly oscillatory
- Closed-form solutions only for same-centre case \((A = B)\)
- Must resort to numerical integration for “off-site” case \(A \neq B\)

Main developer: James Womack
Accelerated SW Coulomb metric matrix evaluation

Old: 3Dc

- Computationally expensive
- Memory: $O(N_{\text{node}}^3)$
- Bottleneck for HFX

New: 2Dn-1Da

- Significantly reduced cost
- Memory: $O(N_{\text{node}}^2)$
- Small fraction of HFX cost
Accelerated SW Coulomb metric matrix evaluation

New scheme also makes much larger systems accessible... 13696 atoms, 1000 cores on Iridis 5: V evaluation took \textbf{30 min}

Myoglobin model
27 atoms

Single inner loop
V evaluation
HFX energy

Iridis 5
1 node (2 \times 20 cores)
10 MPI / 4 OMP
Improved parallelisation

**Old:** poor scaling for > 50 CPUs
- Convoy effects from mixing comms and compute
- Usual atom-to-MPI distribution unsuited for HFX
- OpenMP sections too finely grained / low-level

**New:** scales to 1000s of cores
- “Dry-run” to perform comms in advance
- Atom-pair-to-MPI distribution for HFX
- OpenMP sections moved to higher level
- Aggressive caching in thread-shared memory
- **Cleverness:** intelligently caching most used quantities

Main developer: Jacek Dziedzic
Improved parallelisation

Protein “scoops” on Iridis 5 (April 2019)

Much larger systems are now practically accessible
Setting up a calculation

Key settings when using a hybrid functional

- **Spherical wave resolution-of-the-identity (SWRI)** setup
- **Hartree-Fock exchange (HFX)** setup
- Selecting an **XC functional**

```
%block swri
  for_hfx 3 10 V 12 12 WE2
%endblock swri

%block species_swri-for_hfx
  O
  C
  H
%endblock species_swri-for_hfx

hfx_use_ri : for_hfx
hfx_max_l : 3
hfx_max_q : 10
hfx_cutoff : 20 bohr
hfx_metric : ELECTROSTATIC

cutoff_energy : 800 eV
xc_functional : B3LYP
```
**SWRI setup**

\[ X_{\alpha\beta} = (\varphi_\alpha \varphi_\delta | f_p) V^{pq} (f_q | \varphi_\beta \varphi_\gamma) K^{\delta\gamma} \]

\[ f_p(r) = \begin{cases} j_{l_p}(q pr) Z_{l_p m_p} (\hat{r}) & r < a \\ 0 & r \geq a \end{cases} \]

%block swri
   myname \( l_{\text{max}} \) \( q_{\text{max}} \) metric \( N_i \) \( N_o \) flags
%endblock swri

%block species_swri-mynname
   atom-label-1
   atom-label-2
   ...
%endblock species_swri-mynname

<table>
<thead>
<tr>
<th>Setting</th>
<th>Description</th>
<th>Recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>myname</td>
<td>label for SW basis/metric</td>
<td>short string</td>
</tr>
<tr>
<td>( l_{\text{max}} )</td>
<td>maximum angular momentum in SW basis</td>
<td>( \geq 3 )</td>
</tr>
<tr>
<td>( q_{\text{max}} )</td>
<td>number of spherical Bessels per ( l )</td>
<td>( \geq 10 )</td>
</tr>
<tr>
<td>metric</td>
<td>metric type</td>
<td>( V ) (electrostatic)</td>
</tr>
<tr>
<td>( N_i )</td>
<td>numerical integration intervals</td>
<td>( \geq 12 )</td>
</tr>
<tr>
<td>( N_o )</td>
<td>numerical integration polynomial order</td>
<td>( \geq 12 )</td>
</tr>
<tr>
<td>flags</td>
<td>metric matrix evaluation control</td>
<td>see documentation</td>
</tr>
<tr>
<td>atom-label-X</td>
<td>atoms to include in SWRI</td>
<td>all atom labels</td>
</tr>
</tbody>
</table>
## HFX setup

<table>
<thead>
<tr>
<th>Keyword and value</th>
<th>Description</th>
<th>Recommended value</th>
</tr>
</thead>
<tbody>
<tr>
<td>hfx_use_ri myname</td>
<td>label for SW basis/metric to use</td>
<td>short string</td>
</tr>
<tr>
<td>hfx_max_l l_{\text{max}}</td>
<td>max. ang. mom. in SW basis</td>
<td>≥3</td>
</tr>
<tr>
<td>hfx_max_q q_{\text{max}}</td>
<td>number of spherical Bessels per l</td>
<td>≥10</td>
</tr>
<tr>
<td>hfx_cutoff r_{X} [unit]</td>
<td>cutoff for exchange interactions</td>
<td>≥20 bohr</td>
</tr>
<tr>
<td>hfx_metric metricname</td>
<td>metric type (name)</td>
<td>ELECTROSTATIC</td>
</tr>
</tbody>
</table>

\[
E_{X}^{\text{HFX}} = -K^{\beta\alpha}(\varphi_\alpha \varphi_\delta | \varphi_\beta \varphi_\gamma)K^{\delta\gamma} \\
= -K^{\beta\alpha}X_{\alpha\beta}
\]

\[
R_{\alpha\beta} > r_{X} \implies X_{\alpha\beta} = 0
\]
XC functional selection

```
xcc_functional funcname
```

Available hybrid functionals:

```
B1LYP, B1PW91, PBE0, B3LYP, B3PW91, X3LYP
```

For 100% Hartree-Fock exchange, use HF
Understanding the output

A successful calculation

- Similar to local XC calculation with normal output detail
- Local XC ($E_{xc}^{loc}$) and HFX ($E_{x}^{HF}$) energies reported

<table>
<thead>
<tr>
<th>ENERGY COMPONENTS (Eh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kinetic : 27.35982130694516</td>
</tr>
<tr>
<td>Pseudopotential (local) : -126.10605114635337</td>
</tr>
<tr>
<td>Pseudo (non-coul chg cor) : 0.00000000000000</td>
</tr>
<tr>
<td>Pseudopotential (non-local) : 5.13945962391166</td>
</tr>
<tr>
<td>Hartree : 48.59030602370854</td>
</tr>
<tr>
<td>Exchange-correlation : -7.97965813472141</td>
</tr>
<tr>
<td>Hartree-Fock exchange : -1.78484725447100</td>
</tr>
<tr>
<td>Ewald : 14.80743332976961</td>
</tr>
<tr>
<td>Total : -39.97353625121082</td>
</tr>
</tbody>
</table>

Integrated density : 19.99999999999997
Understanding the output

**SWRI details:** metric matrix method, metric type, memory usage, progress evaluating metric matrix

```
SWRI: Initialising module (stage 1)...
SWRI: [hfx] Initialising (stage 1)...
SWRI: [hfx] - Using user-selected metric matrix evaluation scheme (2D numerical-1D analytic).
SWRI: [hfx] - Initialising Bessels (l_max=3)... done.
SWRI: [hfx] - SW batch size (tile width) set to 160.
SWRI: [hfx] - Initialising persistent cache structures... done.
SWRI: [hfx] - Initialising metric matrix... done.
SWRI: [hfx] Done.
SWRI: Stage 1 completed.
Sparse matrix initialisation ... done

SWRI: Initializing module (stage 2)...
SWRI: [hfx] Initialising (stage 2)...
SWRI: [hfx] - NL: Populating neighbour list: SWRI_S_ATOMS... done.
SWRI: [hfx] - Generating and communicating atomblock list.
SWRI: [hfx] 0 atomblocks read from disk.
SWRI: [hfx] 21 atomblocks to process.
SWRI: [hfx] Sorting atomblocks according to proximity to (0.000000, 0.000000, 0.000000).
SWRI: [hfx] - Calculating on-site elements of the metric matrix V... done.
SWRI: [hfx] - Calculating off-site elements of the metric matrix V.
```
Understanding the output

**SWRI details:** metric matrix method, metric type, **memory usage**, progress evaluating metric

<table>
<thead>
<tr>
<th>Spherical wave res. of identity Chebyshev engine</th>
<th>Estimated memory requirement per MPI rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>workspace atomblock                             : 200.00 kB</td>
<td></td>
</tr>
<tr>
<td>vtile                                            : 200.00 kB</td>
<td></td>
</tr>
<tr>
<td>recv tile                                       : 200.00 kB</td>
<td></td>
</tr>
<tr>
<td>workers_done                                    : 24.00 B</td>
<td></td>
</tr>
<tr>
<td>atomblocklist                                   : 420.00 B</td>
<td></td>
</tr>
<tr>
<td>tilelist (root only)                            : 252.00 B</td>
<td></td>
</tr>
<tr>
<td>tile hash table (upper bound)                   : 1.17 MB</td>
<td></td>
</tr>
<tr>
<td>nodes_template                                  : 162.00 kB</td>
<td></td>
</tr>
<tr>
<td>swop_values                                     : 648.00 kB</td>
<td></td>
</tr>
<tr>
<td>swy_values                                      : 648.00 kB</td>
<td></td>
</tr>
<tr>
<td>swhome_all_coeffs                               : 25.31 MB</td>
<td></td>
</tr>
<tr>
<td>swpot_batch_coeffs                              : 25.31 MB</td>
<td></td>
</tr>
<tr>
<td>all_atomblocks (indices)                        : 252.00 B</td>
<td></td>
</tr>
<tr>
<td>my_atomblocks (indices)                         : 252.00 B</td>
<td></td>
</tr>
<tr>
<td>num_q_per_l                                     : 48.00 B</td>
<td></td>
</tr>
<tr>
<td>atomblock rotation matrix                       : 200.00 kB</td>
<td></td>
</tr>
</tbody>
</table>

| Estimated peak total per MPI rank : 54.00 MB |

---

**SWRI:** [hfx] 21 tiles to process (21 atomblocks x 1 batches)
SWRI: [hfx] Tiles completed: 0, left: 15, in progress: 6.
SWRI: [hfx] Tiles completed: 18, left: 0, in progress: 3.
SWRI: [hfx] - Filling the remainder by symmetry.
SWRI: [hfx] - Writing for_hfx.vmatrix to file "h_bond_B3LYP_2Dn-1Da_modified_for_hfx.vmatrix"...done
SWRI: Done.
SWX: Initialising module... done.
Tips for maximising performance

HFX is . . .

▶ Much more expensive than local XC
▶ Very memory-hungry (RAM)

Improve performance by . . .

▶ Always using the new metric matrix (2Dn-1Da) scheme (default)
▶ Choosing a sensible exchange cutoff $r_X$
▶ Increasing the $\frac{N_{\text{OMP}}}{N_{\text{MPI}}}$ ratio (more RAM per MPI process)
▶ Increasing per process cache limits to fully utilise node memory

```plaintext
  cache_limit_for_swops     cache_limit_for_expansions
  cache_limit_for_ps        cache_limit_for_ngwfs
  cache_limit_for_dknblks   cache_limit_for_coeffs
```

These apply to version of HFX in v5.2 and earlier
Further information

For practical usage, see online documentation

“Spherical-wave resolution of identity (SWRI), Distributed Multipole Analysis (DMA), and Hartree-Fock exchange (HFx)”

For theory and technical details, see Ref. 7


Publications on recent developments coming soon. . .
## Acknowledgements

### Funding and compute resources

- EPSRC-funded CCP9 flagship ([EP/P02209X/1](https://www.epsrc.ac.uk/))
- UK MMM Hub “Thomas” ([EP/P020194/1](https://www.epsrc.ac.uk/))
- IRIDIS HPC Facility (Southampton)

### CCP9 flagship collaborators

- Nick Hine (Warwick, PI)
- Peter Haynes (Imperial, Co-I)
- Arash Mostofi (Imperial, Co-I)
- Chris-Kriton Skylaris (Southampton, Co-I)
- Jacek Dziedzic (Southampton/Gdańsk, RCo-I)
- Jolyon Aarons (Warwick, PDRA)
- Joseph Prentice (Imperial, PDRA)


