Linear-scaling, tensorially invariant, self-consistent projector constrained Density Functional Theory in ONETEP

G. Teobaldi, a D. D. O'Regan, b N. D. M. Hine, c A. A. Mostofi d

aStephenson Institute for Renewable Energy, Department of Chemistry, University of Liverpool, UK. Email: g.teobaldi@liv.ac.uk
bTheory and Simulation of Materials, EPFL, Lausanne, Switzerland
cDepartment of Physics, Cavendish Laboratory, Cambridge University, UK
dDepartment of Materials and Physics, Imperial College London, UK
Motivation: need of atomic-level insight in energy materials & interfaces

- Electrochemical **energy storage** (B, SCAP)
- Electrochemical **energy conversion** (FC)
- (Bio-)**fuel production**
- e-(h) migration across **PV interfaces**
- **Catalysis & photo-catalysis**
- (Solar) **light-harvesting**

**Charge (energy) transfer** in extended, heterogeneous, often **buried** interfaces

Need to access to the physico-chemical factors governing the **thermodynamics** and **kinetics** of charge (energy) transfer in such interfaces

**Viable approximations** to the physico-chemical factors governing the **thermodynamics** and **kinetics** of charge (energy) transfer in such interfaces
Electron coupling and e-, h-, energy-transfer (in condensed phases)


Electron Transfer

Donor  Acceptor

Singlet Excitation Energy Transfer

Donor  Acceptor

Triplet Excitation Energy Transfer

Donor  Acceptor

S↔T inter-system crossing


- Selectively localise e/h/e-h pair on ‘a’ and ‘b’

- Optimise (the geometry of the) system on the ‘a’ and ‘b’ diabatic) electronic potential energy surfaces → evaluate $\Delta G^0$, $\lambda$

- Approximate many-body electronic wave-functions $\Psi_a$, $\Psi_b$

- Approximate non-adiabatic coupling $H_{ab} = \langle \Psi_a | H_{el} | \Psi_b \rangle$

($H_{el}$ is not diagonal in the \{a,b\} diabatic representation)
Constrained-DFT [and (approximated) diabatic states: \( \tau_j = \langle \psi_j | \nabla \psi_j \rangle = 0; \quad \nabla \times \tau = \tau \times \tau \)]

\[
\sum \int w_i^\alpha(r) \rho_\alpha(r) \, dr - N_k = 0
\]

\[
W[\rho,\{V_k\}] = E[\rho] + \sum_k V_k \sum_\alpha \int w_i^\alpha(r) \rho_\alpha(r) \, dr - N_k
\]

\[
W[P] = E[P] + W_c(\text{Tr}\{Pw_c\} - N_2)
\]

1. Add constrain on the (e) density [electron/spin-density localisation…]

2. Build a constrained functional \( W \) [via the Lagrange multipliers \( V_k \)]

3. Reformulate \( W \) in terms of the density matrix (\( P \))

4. Maximize \( W \) (concave) WRT \( P \) & \( V_k \)

\[
\frac{\partial^2 W}{\partial V_k \partial V_k} = \sum_{\alpha} \sum_{\lambda} \int \frac{\phi_{\alpha \lambda}(r)}{\phi_{\alpha \lambda}(r')} w_{\alpha \lambda}^\alpha(r') \, dr + cc
\]

\[
= \sum_{\alpha} \sum_{\lambda} \sum_{\lambda'} \phi_{\alpha \lambda}(r) \phi_{\alpha \lambda}(r') \left( \frac{\delta[\phi_{\alpha \lambda}(r')]}{\delta[\phi_{\alpha \lambda}(r)']} \right)
\]

\[
= 2 \sum_{\alpha} \sum_{\lambda} \sum_{\lambda'>\lambda} \langle \phi_{\alpha \lambda} | w_{\lambda'}^\alpha | \phi_{\alpha \lambda'} \rangle \langle \phi_{\alpha \lambda'} | w_{\lambda'}^\alpha | \phi_{\alpha \lambda} \rangle
\]

\[
\nabla A W = \nabla A E + V_c \sum_{\lambda \mu} P_{\lambda \mu} \nabla A w_{c \lambda \mu}
\]

5. atomic forces [various definition of \( w_c \): Löwdin, Becke, Bader, etc population analysis]

• State-of-the art linear-scaling DFT code, capable of treating tens of thousands of atoms.

• Reformulates Kohn-Sham DFT in terms of the single-particle density matrix.

• Uses localized functions (NGWFs or Non-orthogonal Generalized Wannier Functions).

• Linear scaling is achieved by exploiting
  – the sparsity of the density kernel, and
  – the exponential decay of the density matrix.

  ... by the introduction of spatial cutoffs for the kernel and the NGWFs.

• Achieves near-complete basis set accuracy.

• Fully parallel, portable.

Single-point energy calculation for an amyloid fibril protein, run on 256 cores of Iridis3 (University of Southampton)

\[ \rho(r, r') = \sum_{\alpha, \beta} \phi_{\alpha}(r) K^{\alpha\beta} \phi_{\beta}^*(r') \]

*Near-complete basis set accuracy*

\[ \Phi_{\alpha} \rightarrow K^{\alpha\beta} \]

\[ \text{LiH} \]

\[ \text{H}_2\text{O} \text{ dimer} \]

---

Projector constrained-DFT in ONETEP

1. For each spin-channel $\sigma$, and cDFT-site $I$, augment the DFT energy by the difference from the targeted electronic population $[(\Delta)N_{(I)(\sigma)}]$ times the constraining potential $(U_{q/s})$

$$E_{cDFT} = E_{DFT} + \sum_{I=1}^{N_{sites}} \sum_{\sigma=1}^{2} U_q^{(I)(\sigma)} (\text{Tr} \left[ n_{(I)(\sigma)} \right] - N_{(I)(\sigma)}) + \sum_{I=1}^{N_{sites}} U_s^{(I)} (\text{Tr} \left[ n_{(I)(\uparrow)} \right] - \text{Tr} \left[ n_{(I)(\downarrow)} \right] - \Delta N_{(I)})$$

where the tensorially invariant $^1$ subspace occupancy of the $I^{th}$ cDFT-site reads:

$$n_{(I)(\sigma)} = n_{(I)(\sigma)m'} = O_{(I)mm''} \langle \varphi_{m''}^{(I)} | \phi_{\alpha} \rangle K^{(\sigma)\alpha\beta} \langle \phi_{\beta} | \varphi_{m'}^{(I)} \rangle$$

with

$$O_{mm'}^{(I)} = \langle \varphi_{m}^{(I)} | \varphi_{m'}^{(I)} \rangle, \quad | \varphi_{m}^{(I)} \rangle = | \varphi_{m'}^{(I)} \rangle O_{m'm}^{(I)}, \quad O_{mm''}^{(I)} O_{m''m'}^{(I)} = \delta_{m'm}$$

2. Augment ($\sigma$-dependent) DFT Hamiltonian with ($\sigma$-dependent) cDFT contributions, calculated as derivative of the cDFT correction with respect to an arbitrary density Kernel:

$$H_{cDFT}^{(\sigma)} = \left[ U_q^{(\sigma)} + (-1)^{1+\sigma} U_s \right] \left[ \partial n_{m}^{(\sigma)} / \partial K^{(\sigma)\alpha\beta} \right] = \left[ U_q^{(\sigma)} + (-1)^{1+\sigma} U_s \right] \left[ V_{\beta m} O_{mm''}^{(I)} W_{m''}\alpha \right]$$

$$V_{\beta m} = \langle \phi_{\beta} | \varphi_{m} \rangle, \quad W_{m''}\alpha = V_{\alpha m''}^\dagger = \langle \varphi_{m''} | \phi_{\alpha} \rangle$$

So, what are we doing in practice?

...(just) adding attractive/repulsive potentials to **constraint** subspace **populations**, population **differences**, and/or magnetic moments (differences)

$$\begin{align*}
E_{cDFT} &= E_{DFT} + \sum_{l=1}^{N_{\text{sites}}} \sum_{\sigma=1}^2 U_q^{(l)}(\sigma) \left( \text{Tr} \left[ n^{(l)(\sigma)} \right] - N_{(l)(\sigma)} \right) + \sum_{l=1}^{N_{\text{sites}}} U_s^{(l)} \left( \text{Tr} \left[ n^{(l)(\uparrow)} \right] - \text{Tr} \left[ n^{(l)(\downarrow)} \right] - \Delta N_{(l)} \right) 
\end{align*}$$
Choosing the projectors for the cDFT-subspaces

• The cDFT functional is constructed using the occupancy matrix \( n \) of the localised cDFT-projectors \( \varphi_m \)

\[
n^{(I)}(\sigma) = n^{(I)}(\sigma)_{mm'} = O^{(I)}_{mm''} \langle \varphi^{(I)}_{mm''} | \phi_\alpha \rangle K^{(\sigma)\alpha\beta} \langle \phi_\beta | \varphi^{(I)}_{mm'} \rangle
\]

\[
O^{(I)}_{mm'} = \langle \varphi^{(I)}_m | \varphi^{(I)}_{m'} \rangle, \quad | \varphi^{(I)}_m \rangle = | \varphi^{(I)}_{m'} \rangle O^{(I)}_{m'm'}, \quad O^{(I)}_{mm''} O^{(I)}_{m'm''} = \delta_{mm'}
\]

• What to use as cDFT-projectors?
  LCAO basis function, valence pseudo-orbitals, hydrogenic wavefunctions, Maximally Localised Wannier Function (MLWF), in situ optimised DFT-NGWFs, in situ optimised cDFT-NWGFs, self-consistent cDFT-NGWFs?
Choosing between the implemented (8x2x2=32) flavours of cDFT

\[ V \]

\[ \text{SPIN-UP} \]

\[ i^{th} \text{ cDFT-site} \]

\[ \text{SPIN-DOWN} \]

\[ k^{th} \text{ cDFT-site} \]
How to optimise the cDFT-potentials with just one NGWFs-optimisation...

- Kohn-Sham DFT: **minimum** of $E_{\text{DFT}}$ with respect to NGWFs for idempotent $K^{\alpha \beta}$

- cDFT: **maximum** of $E_{\text{cDFT}}$ with respect to constraining potentials ($U_{q/s}$)

Can the problem be solved by alternating NWGFs and $U_{q/s}$ optimisation steps? (having ‘only’ $\partial E_{\text{cDFT}}/\partial \Phi_\alpha$, $\partial E_{\text{cDFT}}/\partial K^{\alpha \beta}$, and $\partial E_{\text{cDFT}}/\partial U_{q/s}$ available)

\[
\frac{\partial E_{\text{cDFT}}}{\partial U^{(i)}_s} = Tr \left[ n^{(t)(i)} \right] - Tr \left[ n^{(i)(i)} \right] - \Delta N^{(i)}
\]

\[
\frac{\partial E_{\text{cDFT}}}{\partial U^{(i)(\sigma)}_q} = Tr \left[ n^{(i)(\sigma)} \right] - N^{(i)(\sigma)}
\]
More on alternate NGWFs-cDFT line search [1/2]

NGWFs-OPT Loop ($E_{\text{DFT+U}_{q/s}}$ minimisation)
- NGWFs_0

Denskern-OPT Loop ($E_{\text{DFT+U}_{q/s}}$ minimisation, fixed $U_{q/s}$)
- NGWFs_1 (fixed $U_{q/s}$)

Denskern-OPT Loop ($E_{\text{DFT+U}_{q/s}}$ minimisation, fixed $U_{q/s}$)

Extrapolate NGWFs_new (NGWFs_0, NGWFs_1)
More on alternate NGWFs-cDFT line search [2/2]

- NGWFs-OPT Loop (\(E_{DFT+Uq/s}\) minimisation)
  - NGWFs_0
    - Uq/s-OPT Loop (\(E_{DFT} + E_c = E_{cDFT}\) MAXIMISATION)
      - Denskern-OPT Loop (\(E_{DFT+Uq/s}\) minimisation, fixed \(Uq/s\))
        - NGWFs_1 (fixed \(Uq/s\), optimised for NGWFs_0)
          - Denskern-OPT Loop (\(E_{DFT+Uq/s}\) minimisation, fixed \(Uq/s\))
  - Extrapolate NGWFs_new (NGWFs_0, NGWFs_1)
In action on $N^{(+)}=N^{(-)}$ (1.1 Å)

**Initial** PAO NGWFs ($\Phi_\alpha$) & cDFT-$\varphi_m$

**Final** cDFT-optimised NGWFs ($\Phi_\alpha$)

**Fast convergence** within the requested $2 \times 10^{-6}$ (NGWFs-RMS) and $10^{-5}$ (cDFT-RMS) thresholds
Do we actually end up with $^*N^+(+) = ^*N^-(−) (1.1 \text{ Å})$? [YES]

\[ \Delta \rho = \rho^\uparrow - \rho^\downarrow \]

- Noticeably sharp separation of $\rho^\uparrow$ and $\rho^\downarrow$ in spite of 5Å-radius NGWFs and PAO-$\varphi_m$
- How is it possible? Because $K^{\alpha\beta}$, the NGWFs $\{\phi_\alpha\}$, and $U_q$ have been optimised for the given cDFT-problem.
(Fixed-\(\varphi_m\)) cDFT atomic-forces

Instantaneous cDFT-site population

\[ E_{cDFT} = E_{DFT} + U_q(Tr[n] - N) \]

\[ F = -\frac{\partial E_{cDFT}}{\partial R} = -\nabla_R E_{cDFT} = -\nabla_R E_{DFT} - \nabla_R U_q(Tr[n] - N) = \]

\[ = -\nabla_R E_{DFT} - (\nabla_R U_q)(Tr[n] - N) - U_q \nabla_R Tr[n] + U_q Tr[n] \nabla_R N = \]

\[ \to 0 \text{ after cDFT-solution} \]

\[ = 0 \text{ (N is constant)} \]

(Fixed-$\varphi_m$) cDFT atomic-forces and geometry optimisation for $N^{(+)}=N^{(-)}$

- Fast-convergence of BFGS optimisation algorithm: $F_{\text{max}} < 0.01 \text{ eV/Å}$ in 4 BFGS iterations
- Spin-decoupling of two $\pi$-electrons in $^\uparrow N^{(+)}=^\uparrow N^{(-)} \rightarrow$ elongation of optimum bond-distance
- 4 PAO-$\varphi_m$: 1.10 Å $\rightarrow$ 1.35 Å
- 9 PAO-$\varphi_m$: 1.10 Å $\rightarrow$ 1.33 Å
- Different set of cDFT-projectors ($\varphi_m$) $\rightarrow$ different constraints $\rightarrow$ different cDFT-solutions
- How to reduce cDFT-ambiguities?

$\rho^+ - \rho^+ = +1.5 \times 10^{-3} \text{ e Å}^{-3}$
$\rho^+ - \rho^+ = -1.5 \times 10^{-3} \text{ e Å}^{-3}$
Optimised cDFT-NGWFs ($\Phi_\alpha$) as new cDFT-projectors ($\varphi_m$)

- Use cDFT-optimised NGWFs ($\Phi_\alpha$) at iteration (i) as new cDFT-projectors ($\varphi_m$) for iteration (i+1), and re-solve the cDFT problem by optimising the cDFT-NGWFs ($\Phi_\alpha$) for iteration (i+1).
- Iterate until $E_{\text{cDFT}}$ (i.e. $\Phi_\alpha$, $K^{\alpha\beta}$, and $\varphi_m$) do not change within given threshold.
Optimised cDFT-NGWFs ($\Phi_\alpha$) as new cDFT-projectors ($\varphi_m$)

Initial PAO NGWFs ($\Phi_\alpha$) & cDFT-$\varphi_m$

Final cDFT-optimised NGWFs ($\Phi_\alpha$)

$\varphi_m^{(i+1)} = \phi_\alpha^{(i)}$

Initial cDFT-optimised NGWFs ($\Phi_\alpha$) & cDFT-$\varphi_m$

Self-consistent cDFT-optimised NGWFs ($\Phi_\alpha$) & cDFT-$\varphi_m$
How fast/tight is the convergence of the cDFT-projectors ($\varphi_m$)?

- **Fast convergence** ($\leq 7$ iterations), even for a tightly-constrained problem [N-N: 1.35 Å (1.33 Å)]
- **Tight convergence** ($\leq 10^{-12}$ eV), even for reasonable kinetic-energy cutoffs (1000 eV)
- As the cDFT-projectors ($\varphi_m$) reach self-consistency, energy of cDFT-solution is lowered
- **Best cDFT-projectors** as those yielding the lowest-energy cDFT solution [for the same targeted constraint]
How large is the computational overhead?

- **4 PAO-$$\phi_m$$**: 1st cDFT-solution: **10** NWGFS-opt iterations; sc-$$\phi_m$$ procedure: **12**
- **9 PAO-$$\phi_m$$**: 1st cDFT-solution: **8** NGWFs-opt iterations; sc-$$\phi_m$$ procedure: **13**

- ~$$x2$$ increase of computation cost, for a convergence tighter than $$10^{-12}$$ eV
- ~$$x0.5$$ increase of computation cost, for a convergence tighter than $$10^{-4}$$ eV
Self-consistency of system-geometry ($R$) and cDFT-projectors ($\varphi_m$)

- sc-cDFT-projectors ($sc-\varphi_m$) depend on system-geometry $R$

- $\frac{\partial \varphi_m^{SC}}{\partial R}$?!? no idea(!)

- ‘Solution’: perform optimisation of both $\varphi_m$ and $R$, alternating self-consistency cDFT-projector ($\varphi_m$) optimisation (at fixed-$R$) and geometry ($R$) optimisation (with fixed-$sc-\varphi_m$) loops.

$R_0, \varphi_m^0$ $\downarrow \varphi_m$ sc-loop $\uparrow R$-opt loop $\rightarrow R_{opt}$ $sc-\varphi_m(R_{opt})$

$R_0, \varphi_m^{0,SC}(R_0)$ $\rightarrow R_{opt}[\varphi_m^{0,SC}(R_0), \varphi_m^{0,SC}(R_0)]$ $\downarrow \varphi_m$ sc-loop $\uparrow R$-opt loop $\rightarrow R_{opt}[\varphi_m^{0,SC}(R_0), \varphi_m^{0,SC}(R_{opt})]$ $\rightarrow \ldots$ and so on...
How large is the computational overhead?

- **4 \( \varphi_m \):** 1\textsuperscript{st} sc-\( \varphi_m \) loop: 22 NGWFs optimisation iterations
  2\textsuperscript{nd} sc-\( \varphi_m \) loop: 7 NGWFs optimisation iterations

- **9 \( \varphi_m \):** 1\textsuperscript{st} sc-\( \varphi_m \) loop: 21 NGWFs optimisation iterations
  2\textsuperscript{nd} sc-\( \varphi_m \) loop: 6 NGWFs optimisation iterations

- **Largest computational overhead from extra R\(_{opt}\) loop** (4 BFGS-iterations)
  [if interested in improving convergence of \( E_{cDFT} \) from \( 10^{-3} \) to \( 10^{-5} \) eV]
Intra-molecular e-transfer in Q-TTF-Q\(^{(-)}\)

HOMO\(_a\) \hspace{1cm} \lambda = \lambda_i + \lambda_s \hspace{1cm} \text{HOMO}\(_b\)

- 0.03 eV Å\(^{-1}\) geometry optimisation threshold (Pulay corrected\(^1\))
- \(\lambda_i\)\(_{\text{ONETEP}}\) (BLYP, \(\varphi_m\): DFT-\(\varphi_a\)) = 16.5 kcal/mol
- \(\langle\lambda_i\rangle\) (BLYP, PW, Hirshfeld)\(^{2a}\) = 5.1 kcal/mol [averaged over MD trajectory in \(H_2O\)]
- \(\lambda_i\) (B3LYP/6-31+G(d), Löwdin population)\(^{2b}\) = 13.1 kcal/mol
- \(\lambda_i\) (B3LYP/6-31+G(d), Becke population)\(^{2c}\) = 16.9 kcal/mol
- \(\lambda_{\exp}(\text{ethyl acetate:t-butyl alcohol 10:1}) = \lambda_i + \lambda_s \sim 22\) kcal/mol\(^{2b}\)

e-transfer in metal-organic systems: Ferrocenium-Ferrocene complexes

\[
\begin{align*}
\Delta \rho &= \rho^\uparrow - \rho^\downarrow = 10^{-2} \text{ e Å}^{-3} \\
\text{• 0.05 eV Å}^{-1} \text{ geometry optimisation threshold (Pulay corrected)} \\
\text{• } \Delta G^\ddagger_{\text{ONETEP}}(\text{PBE}, \varphi_m; \text{DFT-}\phi_a) &= 5.1 \text{ kcal/mol} \\
\text{• } \Delta G^\ddagger(\text{B3LYP/6-31G**/LANL2DZ, Becke population}) &= 7.5 \text{ kcal/mol} \\
\text{• } \Delta G^\ddagger_{\text{ONETEP}}(\text{PBE}, \varphi_m; \text{DFT-}\phi_a) &= 1.7 \text{ kcal/mol} \\
\text{• } \Delta G^\ddagger(\text{B3LYP/6-31G**/LANL2DZ, Becke population}) &= 1.4 \text{ kcal/mol}
\end{align*}
\]

To summarise

\[ \Phi_\alpha \quad K^{\alpha\beta} \]

\[ \varphi_m \quad U_q \quad K^{\alpha\beta} \]

\[ \rho^+ - \rho^+ > 0 \]
\[ \rho^+ - \rho^+ < 0 \]