Electronic energy optimisation in ONETEP

Chris-Kriton Skylaris
c.skylaris@soton.ac.uk

UNIVERSITY OF Southampton
School of Chemistry

ONETEP spring school, 13-16 April 2010
Outline

1. Kohn-Sham DFT calculations
   • Direct energy minimisation versus density mixing
   • Using the density matrix to achieve linear-scaling

2. ONETEP scheme: density kernel and NGWFs
   • Density matrix idempotency and normalisation conditions

3. Linear-scaling functionals for density matrix optimisation
   • Penalty, LNV, PM

4. Optimisation of density kernel
   • Density kernel gradients, tensor properties

5. Optimisation of NGWFs
   • NGWF gradients, preconditioning schemes

6. Overall ONETEP calculation scheme
   • Initialisation of density kernel and NGWFs
   • Flowchart, input keywords
Density Functional Theory (DFT)

• Electronic density

\[ n(r_1) = N \int \cdots \int \Psi(r_1s_1, x_2, \ldots, x_N) \Psi^*(r_1s_1, x_2, \ldots, x_N) ds_1 dx_2 \cdots x_N \]

• Hohenberg-Kohn

\[ E[n] = E_{\text{Kin}}[n] + E_{\text{ext}}[n] + E_{\text{ee}}[n] \]

• Kohn-Sham

\[ E[n] = \sum_i \langle \psi_i | -\frac{1}{2} \nabla^2 | \psi_i \rangle + E_{\text{ext}}[n] + E_{\text{Coul}}[n] + E_{\text{xc}}[n] \]

Non-interacting electrons

\[
\begin{bmatrix}
-\frac{\hbar^2}{2m} \nabla^2 + v_{\text{eff}}(r)
\end{bmatrix}
\psi_i(r) = \varepsilon_i \psi_i(r)
\]

Density of interacting electrons

\[ n(r) = \sum_i |\psi_i(r)|^2 \]
Self-consistency

• Variational principle

\[ \frac{\delta E[\{\psi_i\}]}{\delta \psi_j} = 0 \]

• Constraint

\[ \langle \psi_k | \psi_n \rangle = \delta_{kn} \]

• Eigenvalue equation for molecular orbitals (bands)

\[ \hat{H} \psi_n = \varepsilon_n \psi_n \]

• But the Hamiltonian operator depends on the molecular orbitals

\[ \hat{H}[\{\psi_i^{(p)}\}] \psi_n^{(p+1)} = \varepsilon_n^{(p+1)} \psi_n^{(p+1)} \]
Achieving self-consistency

**Direct energy minimisation**

**Calculate energy**

\[ E^{(p)} = E(\{\psi_i^{(p)}\}) \]

**Vary orbitals**

\[ \{\psi_i^{(p+1)}\} = \{\psi_i^{(p)}\} + \{\delta \psi_i^{(p)}\} \]

So that

\[ E^{(p+1)} < E^{(p)} \]

Are

\[ E^{(p+1)} \approx E^{(p)} \{\psi_i^{(p+1)}\} \approx \{\psi_i^{(p)}\} \]

within tolerance?

- Yes
  - finished

- No
  - Increase \( p \) by 1

**Discontinuous changes**

- Direct search for well-defined minimum

**Diagonalisation and density mixing**

**Solve (diagonalise Hamiltonian)**

\[
\left[ -\frac{1}{2} \nabla^2 + v_{\text{eff}}[n](\mathbf{r}) \right] \psi_i^{(p+1)}(\mathbf{r}) = \varepsilon_i^{(p+1)} \psi_i^{(p+1)}(\mathbf{r})
\]

**Build density**

\[ n^{(p+1)}(\mathbf{r}) = \sum_i f_i |\psi_i^{(p+1)}(\mathbf{r})|^2 \]

Are

\[ n^{(p+1)}(\mathbf{r}) \approx n^{(p)}(\mathbf{r}) + \cdots \]

within tolerance?

- Yes
  - finished

- No
  - Mix with previous densities

**Diagonalisation and density mixing**

- Discontinuous changes
- Hamiltonian search (indirect)
Function minimisation

- Iterative procedures
- Need the value of the function and its gradient at each step
- Converge to local minima

- Moves always downhill
- Robust but may need very large number of iterations
- Moves always downhill
- Converges in N steps for N-dimensional quadratic function
- In practice very efficient even for non-quadratic functions
Cubic-scaling computational cost of DFT

Calculation run on 96 cores
Kohn-Sham DFT calculation bottlenecks

Minimise $E$ w.r.t.: \[ \{ \psi_i \} \]

Solve \[ \hat{H} |\psi_n\rangle = \epsilon_n |\psi_n\rangle \]

Build energy and Hamiltonian $H, E[n]$ > $O(N^2)$

Orthogonality $\langle \psi_i | \psi_j \rangle = \delta_{ij}$ $O(N^3)$

Diagonalise Hamiltonian matrix $O(N^3)$

Computational bottlenecks
Density matrix localisation

**Nearsightedness of electronic matter**


In systems with a band gap:

\[ \rho(r, r') \sim e^{-\gamma |r-r'|} \to 0 \quad \text{as} \quad |r - r'| \to \infty \]

Take advantage of this locality condition to make the density matrix contain a **linear** amount of data.

Truncate exponential “tail”, impose:

\[ \rho(r, r') = 0 \quad \text{when} \quad |r - r'| > r_\text{cut} \]
One-particle density matrix in DFT

- **Operator representation**

\[ \hat{\rho} = \sum_n f_n |\psi_n\rangle \langle \psi_n| \]

- **Position representation**

\[ \langle r | \hat{\rho} | r' \rangle = \rho(r, r') = \sum_n f_n \psi_n(r) \psi_n^*(r') \]

- **Conditions**

  - **Idempotency** (from **orbital orthonormality** and **occupancies 1 or 0**)

\[ \rho^2(r, r') = \int \rho(r, r'') \rho(r'', r') dr'' = \sum_n f_n^2 \psi_n(r) \psi_n^*(r') = \sum_n f_n \psi_n(r) \psi_n^*(r') = \rho(r, r') \]

  - **Normalisation** (preserving the **number of electrons**)

\[ \text{tr}[\rho] = \int \rho(r, r) dr = \int \sum_n f_n \psi_n(r) \psi_n^*(r) dr = \sum_n f_n \int |\psi_n(r)|^2 dr = \sum_n f_n = N_e \]
Energy expressions

- With orbitals

\[ E[\{\psi_j\}] = \sum_i \int \psi_i(r) \left( -\frac{1}{2} \nabla^2 \right) \psi_i(r) \, dr + \int V_{\text{ext}}(r)n(r) \, dr + \frac{1}{2} \int \int \frac{n(r)n(r')}{|r - r'|} \, dr \, dr' + E_{\text{xc}}[n] \]

\[ n(r) = \sum_i \left| \psi_i(r) \right|^2 \]

- With density matrix

\[ E[\rho] = \int \left[ -\frac{1}{2} \nabla_r^2 \rho(r', r) \right]_{r' = r} \, dr + \int V_{\text{ext}}(r)n(r) \, dr + \frac{1}{2} \int \int \frac{n(r)n(r')}{|r - r'|} \, dr \, dr' + E_{\text{xc}}[n] \]

\[ \rho(r', r) = \sum_i f_i \psi_i(r') \psi_i^*(r) \quad n(r) = \rho(r, r) \]
Self-consistency

\[ \hat{H}[\{\psi_i^{(p)}\}] |\psi_n^{(p+1)}\rangle = \varepsilon_n^{(p+1)} |\psi_n^{(p+1)}\rangle \] therefore

\[ \hat{H}[\{\psi_i^{(p)}\}] = \sum_n \varepsilon_n^{(p+1)} |\psi_n^{(p+1)}\rangle \langle \psi_n^{(p+1)}| \]

\[ \hat{\rho}[\{\psi_i^{(p)}\}] = \sum_n f_n |\psi_n^{(p)}\rangle \langle \psi_n^{(p)}| \]

When \( |\psi_n^{(p)}\rangle = |\psi_n^{(p+1)}\rangle \) the \( \hat{H} \) and \( \hat{\rho} \) commute

\[ [\hat{H}[\{\psi_i^{(p)}\}], \hat{\rho}[\{\psi_i^{(p)}\}]] = 0 \]

Conditions

- Idempotency

\[ \hat{\rho}^2 = \hat{\rho} \iff f_n = 0 \text{ or } 1 \]

- Normalisation

\[ N_e = 2 \text{ tr}[\rho] = 2 \sum_n f_n \] (spin-unpolarised case)
Calculation bottlenecks with the density matrix

Minimise $E$ w.r.t.:

- Orbitals $\{\psi_i\}$
- Build $E[n] > O(N^2)$
- Orthogonality / solve Hamiltonian $\langle \psi_i | \psi_j \rangle = \delta_{ij}$
  - $O(N^3)$
- Idempotency $\hat{\rho}^2 = \hat{\rho}$
  - $O(N^3)$

Density matrix:

$\rho(r, r') = \sum_i f_i \psi_i(r) \psi_i^*(r')$

Computational bottlenecks

- $n(r) = \sum_i |\psi_i(r)|^2$
- $n(r) = \rho(r, r)$
Linear-scaling DFT using the density matrix

Truncate exponential “tail”, impose:

\[ \rho(\mathbf{r}, \mathbf{r}') = 0 \quad \text{when} \quad |\mathbf{r} - \mathbf{r}'| > r_{\text{cut}} \]

The density matrix

\[ \rho(\mathbf{r}, \mathbf{r}') = \sum_n f_n \psi_n(\mathbf{r}) \psi_n^*(\mathbf{r}') \]

is expanded in non-orthogonal localised functions:

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

- K matrix made sparse by truncation
- S and H in terms of \( \{ \phi_\alpha \} \) also sparse
- Optimise energy with respect to K using algorithms that include only sparse matrix multiplications
- Sparse matrix multiplications can be linear-scaling if the sparsity is suitably exploited
ONETEP aims for linear-scaling with high accuracy

The \( \{ \phi_{\alpha} \} \) are **not fixed**, but are optimised *in situ*, (in addition to \( K \))

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

Molecular orbitals (MOs)

Non-orthogonal Generalised Wannier Functions (NGWFs)
Density matrix in terms of NGWFs

\[ \psi_n(\mathbf{r}) = \phi_\alpha(\mathbf{r}) M_\alpha^n \]

\[ S_{\alpha\beta} = \langle \phi_\alpha | \phi_\beta \rangle \]

\[ \rho(\mathbf{r}, \mathbf{r}') = \sum_n f_n \psi_n(\mathbf{r}) \psi^*_n(\mathbf{r}') = \phi_\alpha(\mathbf{r}) \left( \sum_n M_\alpha^n f_n M_{\beta}^\dagger \phi^*_\beta(\mathbf{r}') \right) = \phi_\alpha(\mathbf{r}) K^{\alpha\beta} \phi^*_\beta(\mathbf{r}') \]

\[ K^{\alpha\beta} = \sum_n M_\alpha^n f_n M_{\beta}^\dagger \]

\[ \rho^2(\mathbf{r}, \mathbf{r}') = \phi_\alpha(\mathbf{r}) K^{\alpha\gamma} S_{\gamma\epsilon} K^{\epsilon\beta} \phi^*_\beta(\mathbf{r}') \]

\[ \text{tr}[\rho] = K^{\alpha\beta} S_{\beta\alpha} = \text{tr}[KS] \]
Density matrix optimisation in ONETEP

\[ E = E[K, \{\phi_\alpha\}] \]

**Inner loop:**
Optimise total (interacting) energy \( E \) w.r.t. \( K \) for fixed \( \{\phi_\alpha\} \) while imposing idempotency and normalisation.

**Outer loop:**
Optimise total (interacting) energy \( E \) w.r.t. to \( K \) and \( \{\phi_\alpha\} \).

- **Guess** \( K^{\alpha\beta} \) and \( \{\phi_\alpha\} \)
- **Iteratively improve** \( K^{\alpha\beta} \)
- **Iteratively improve** \( \{\phi_\alpha\} \)
- **Converged?**
- **Yes**
  - finished
- **No**
  - **Guess** \( K^{\alpha\beta} \) and \( \{\phi_\alpha\} \)

**Outer loop:**
Optimise total (interacting) energy \( E \) w.r.t. to \( K \) and \( \{\phi_\alpha\} \).

- **Iteratively improve** \( K^{\alpha\beta} \)
- **Converged?**
- **Yes**
  - finished
- **No**
  - **Iteratively improve** \( \{\phi_\alpha\} \)

**Outer loop:**
Optimise total (interacting) energy \( E \) w.r.t. to \( K \) and \( \{\phi_\alpha\} \).

- **Guess** \( K^{\alpha\beta} \) and \( \{\phi_\alpha\} \)
- **Iteratively improve** \( K^{\alpha\beta} \)
- **Converged?**
- **Yes**
  - finished
- **No**
  - **Guess** \( K^{\alpha\beta} \) and \( \{\phi_\alpha\} \)
“Linear-scaling” functionals

• These are not DFT exchange and correlation functionals

• These are functionals of the density matrix which have been designed so that when they are minimised:
  • Produce the total energy
  • Implicitly impose the required conditions (idempotency and normalisation) on the density matrix that minimises them

• The aim of using these functionals is to work with expressions which involve only (sparse) matrix multiplications

• We will examine the functionals implemented in ONETEP
  • Penalty functional
  • Li –Nunes-Vanderbilt (LNV) types of functionals
  • Palser-Manolopoulos approach (not a functional)
Penalty functional

\[ P[\rho] = \text{tr}[(\rho^2 - \rho)^2] = \sum_n (f_n^2 - f_n)^2 \]

- Minimum (P=0) for idempotent density matrices
- Will always converge, regardless how non-idempotent the initial guess
- Constraints to preserve the total number of electrons are needed
Penalty functional derivatives

\[ P[\rho] = \text{tr}[(\rho^2 - \rho)^2] = \sum_n (f_n^2 - f_{n})^2 \]

- Assume a steepest descents “step” to update \( f_k \)

\[ \frac{\partial P}{\partial f_k} = 4f_k^3 - 6f_k^2 + 2f_k \]

\[ f_k^{(p+1)} = f_k^{(p)} + \lambda d_k^{(p)} \quad f_k^{(p+1)} = f_k^{(p)} - \lambda \frac{\partial P}{\partial f_k}^{(p)} \]

- If we apply a quadratic approximation (near the minimum):

\[ f_{\text{new}} = f_{\text{old}} - \frac{1}{2}(4f_{\text{old}}^3 - 6f_{\text{old}}^2 + 2f_{\text{old}}) = 3f_{\text{old}}^2 - 2f_{\text{old}}^3 \]

Purification transformation

Purification transformation

\[ f_{\text{new}} = 3f_{\text{old}}^2 - 2f_{\text{old}}^3 \]

- Improves the idempotency of a nearly idempotent density matrix
- Forces occupancies in interval \([0,1]\) ("weak" idempotency)
- Quadratic convergence. Example:
  - \( f^{(1)} = 1.1 \)
  - \( f^{(2)} = 3 \times 1.1^3 - 2 \times 1.1^2 = 0.968 \)
  - \( f^{(3)} = 3 \times 0.968^3 - 2 \times 0.968^2 = 0.997 \)
  - Etc..
- But, diverges if the initial occupancies are not within certain bounds
Energy with penalty functional


\[ Q[\rho] = E[\rho] + \alpha P[\rho] \]

- Balance between minimum energy and minimum penalty
- Near-idempotency, depending on value of $\alpha$
- Correction expression available for the energy
Li-Nunes-Vanderbilt (LNV) functionals

\[ \rho = 3\sigma^2 - 2\sigma^3 \]


\[ \Omega_1[\sigma] = \text{tr}[\rho (H - \mu)] \]


\[ \Omega_2[\sigma] = \text{tr}[\rho H] + \mu' (\text{tr}[\sigma] - N_e) \]

- Energy expressions containing a purification transformation of an “auxiliary” density matrix
- Implicitly enforce idempotency
- Break down if purification transformation breaks
- Minimisation of band structure energy (equivalent to diagonalisation)
- But H also depends on \( \rho \): Self-consistency by density mixing
LNV functionals in ONETEP

\[
L_1[\sigma] = E[\rho] + \mu' (\text{tr}[\sigma] - N_e) = E[3\sigma^2 - 2\sigma^3] + \mu' (\text{tr}[\sigma] - N_e)
\]

\[
L_2[\sigma] = E \left[ \frac{N_e}{\text{tr}[\rho]} \rho \right] = E \left[ \frac{N_e}{\text{tr}[3\sigma^2 - 2\sigma^3]} (3\sigma^2 - 2\sigma^3) \right]
\]

- No density mixing, just **direct minimisation** of interacting energy
- Purification never allowed to break down
  - Occupancy maxima and minima tracked
  - Restored by penalty functional if out of safe range
- Electron number conserved without need to know the chemical potential

Canonical purification

\[ \hat{\rho} = \theta(\mu\hat{I} - \hat{H}) \]

\[ H(r, r') = \sum_i \varepsilon_i \psi_i(r)\psi_i(r') \]

\[ \theta(\mu\hat{I} - \hat{H})(r, r') = \ldots = \sum_i \theta(\mu - \varepsilon_i)\psi_i(r)\psi_i(r') = \sum_i f_i \psi_i(r)\psi_i(r') = \rho(r, r') \]

\[ f_i = \begin{cases} 1, & \text{if } \varepsilon_i < \mu \\ 0, & \text{if } \varepsilon_i > \mu \end{cases} \]

• One can approximate \( \rho \) as a polynomial expansion of \( H \)
• Can do this iteratively with a formula that resembles the purification transformation:


• Always converges
• No need to know \( \mu \)
Density kernel optimisation (inner loop)
Derivatives with respect to the density kernel

Relationship between canonical Kohn-Sham orbitals and NGWFs

\[ \psi_n(\mathbf{r}) = \phi_\alpha(\mathbf{r}) M_n^\alpha \iff \phi_\alpha(\mathbf{r}) = \sum_n \psi_n(\mathbf{r})(M^\dagger)_n^\alpha \]

\[ \langle \psi_n \mid \psi_m \rangle = \delta_{nm} \quad \langle \phi_\alpha \mid \phi_\beta \rangle = S_{\alpha\beta}, \quad (S^{-1})^{\alpha\beta} \]

\[ K^{\alpha\beta} = \sum_n M_n^\alpha f_n(M^\dagger)_n^\beta \]

When varying K:

- Occupancies change
- Kohn-Sham orbitals change
- NGWFs do not change
Example: Gradient of simple LNV functional

$$\Omega[\sigma] = E[3\sigma^2 - 2\sigma^3] = E[3\text{LSL} - 2\text{LSLSL}]$$

Purified density kernel: $$K = 3\text{LSL} - 2\text{LSLSL}$$

Differentiate with respect to the elements of $L$

$$\frac{\partial E}{\partial L^{\alpha\beta}} = 6(\text{SLH} + \text{HLS})_{\beta\alpha} - 4(\text{SLSLH} + \text{SLHLS} + \text{HLSLS})_{\beta\alpha} = G_{\beta\alpha}$$

Tensor correction

- To obtain search directions (in steepest descents or conjugate gradients) a **contravariant** gradient is needed

$$G^{\alpha\beta} = S_{\gamma\epsilon}^{\alpha\gamma} G_{\gamma\epsilon} S_{\epsilon\beta} = 6(\text{LHS}_- + \text{S}_-\text{HL})^{\alpha\beta} - 4(\text{LSLHS}_- + \text{LHL} + \text{S}_-\text{HLSL})^{\alpha\beta}$$

NGWF optimisation (outer loop)

- Guess $K^{\alpha\beta}$ and $\{\phi_\alpha\}$
- Iteratively improve $K^{\alpha\beta}$
- Converged?
  - Yes
    - Iteratively improve $\{\phi_\alpha\}$
    - Converged?
      - Yes: finished
      - No: Iteratively improve $K^{\alpha\beta}$
  - No: Converged?
Derivatives with respect to NGWFs

\[ \phi_\alpha(r) = \sum_m D_m(r) C_{m\alpha} \]

Gradient of LNV functional with respect to psinc expansion coefficients of NGWFs

\[ g_m^\alpha = \frac{\partial L}{\partial C^*_m} = 2w \left[ (\hat{H}\phi_\beta)(r) K^{\beta\alpha} + \phi_\beta(r) Q^{\beta\alpha} \right]_{r=r_m} \]

Tensor-corrected covariant gradient

\[ g_{m\alpha} = g_m^\beta S_{\beta\alpha} = 2w \left[ (\hat{H}\phi_\beta)(r) K^{\beta\gamma} S_{\gamma\alpha} + \phi_\beta(r) Q^{\beta\gamma} S_{\gamma\alpha} \right]_{r=r_m} \]

Linear-scaling calculation of gradient with:
- FFT box technique
- Sparse Hamiltonian and overlap matrices
- Sparse density kernel

Maintaining localisation: Advantages of orthogonal basis set

Localisation by truncation of basis set expansion

$$\phi(r) = \sum_i D_i(r)c^i$$

What happens with a non-orthogonal basis:

$$\frac{\partial E}{\partial c^*i} = g_i$$

$$S_{ij} = \langle D_i(r)|D_j(r)\rangle$$

$$c^i_{\text{new}} = c^i_{\text{old}} - \lambda \sum_j (S^{-1})^{ij} g_j$$

Iterative update contains gradients from outside of localisation region.
Occupancy preconditioning

- Derivative of $E$ w.r.t. orbitals
  \[
  \frac{\delta E}{\delta \psi_n^*(\mathbf{r})} = 2f_n[\hat{H}\psi_n(\mathbf{r}) - \epsilon_n\psi_n(\mathbf{r})]
  \]

- Derivative of $E$ w.r.t. NGWFs
  \[
  \frac{\delta E}{\delta \phi_{\alpha}^*(\mathbf{r})} = \int \frac{\delta E}{\delta \phi_{\alpha}(\mathbf{r}')}(\frac{\delta \phi_{\alpha}(\mathbf{r}')}{\delta \psi_n(\mathbf{r})})^* \, d\mathbf{r}' = \frac{\delta E}{\delta \phi_{\alpha}(\mathbf{r})}M_{\alpha n}
  \]
  \[
  \frac{\delta E}{\delta \phi_{\alpha}^*(\mathbf{r})} = \frac{\delta E}{\delta \psi_n^*(\mathbf{r})}(M^\dagger)_{n\beta}(S^{-1})^{\beta\alpha}
  \]

Relation between Kohn-Sham orbitals and NGWFs

\[
\phi_{\alpha}(\mathbf{r}) = \sum_n \psi_n(\mathbf{r})(M^\dagger)_n\alpha
\]

\[
\psi_n(\mathbf{r}) = \phi_{\alpha}(\mathbf{r})M_{\alpha n}^\alpha \quad \langle \psi_n|\psi_m \rangle = \delta_{nm}
\]

\[
\langle \phi_{\alpha}|\phi_{\beta} \rangle = S_{\alpha\beta} , \quad (S^{-1})^{\alpha\beta}
\]

\[
K^{\alpha\beta} = \sum_n M_{\alpha n}f_n(M^\dagger)_n^{\beta}
\]

Remove ill-conditioning due to $f_n \sim 0$

\[
\frac{\delta E}{\delta \phi_{\alpha}(\mathbf{r})} \quad f_n \text{ in } [0,1]
\]

\[
2[\hat{H}\phi_{\beta}(\mathbf{r})K^{\beta\alpha} - \phi_{\beta}(\mathbf{r})K^{\beta\gamma}H_{\gamma\eta}(S^{-1})^{\eta\alpha}]
\]

make all $f_n = 1$ \quad occ\_mix = 1.0

\[
2[\hat{H}\phi_{\beta}(\mathbf{r})(S^{-1})^{\beta\alpha} - \phi_{\beta}(\mathbf{r})(S^{-1})^{\beta\gamma}H_{\gamma\eta}(S^{-1})^{\eta\alpha}]
\]

Covariant \quad Use for search directions

ONETEP spring school, 13-16 April 2010
Kinetic energy preconditioning

Length scale ill-conditioning

- Convergence rate falls with increasing ratio of max/min eigenvalues of the Hamiltonian
- High energy eigenstates are dominated by kinetic energy
- Need to reduce the contribution of kinetic energy in the NGWF gradient

- Pre-multiply NGWF gradient in reciprocal space with a function that behaves as the inverse of the kinetic energy at high wave vectors and approaches 1 at low wavevectors:

\[ \tilde{g}_{k\alpha} \rightarrow \tilde{f}_k \tilde{g}_{k\alpha} \]

Example:

\[ \tilde{f}_k = \frac{k_0^2}{k_0^2 + k^2} \]

Fast convergence
True linear-scaling

H-bond (7 atoms)

Crystalline silicon (1000 atoms)

ZSM5 zeolite (576 atoms)

Protein (988 atoms)

Calculation of $S^{-1}$

The Inverse of

$$S_{\alpha\beta} = \langle \phi_{\alpha} | \phi_{\beta} \rangle$$

Is a **contravariant** tensor, like the density kernel

Can be iteratively generated with the Hotelling formula

$$S_{\text{new}}^{-1} = 2S_{\text{old}}^{-1} - S_{\text{old}}^{-1} S S_{\text{old}}^{-1}$$

- Similar to purification transformation
- Converges rapidly
- Can generate $S^{-1}$ with linear-scaling cost from the beginning
- Can be used to update $S^{-1}$ when NGWFs are updated with minimal computational effort
- Takes advantage of kernel truncation
Overall ONETEP calculation scheme
Initialisation of density kernel and NGWFs

1. **Guess** \{\phi\} as Gaussians (\texttt{AUTO}) or local PAOs (\texttt{.fbl} file input)
2. Guess charge density as sum of atomic densities
3. Build Hamiltonian in \{\phi\} representation
4. Generate K by canonical purification (\texttt{maxit_palser_mano})
5. Improve K by minimisation of energy with penalty functional (\texttt{maxit_pen, pen_param})
6. Proceed to K and \{\phi\} optimisation
NGWF and density kernel optimisation

Start with initialised $K$ and $\{\phi\}$

Do LNV optimisation step for $K$
- $\text{lrv\_threshold\_orig}$,
- $\text{minit\_lrv}$,
- $\text{maxit\_lrv}$

Converged?

Yes

Do optimisation step for $\{\phi\}$
- $\text{ngwf\_threshold\_orig}$,
- $\text{maxit\_ngwf\_cg}$,
- $\text{k\_zero}$

Converged?

No

Yes

Properties, geometry optimisation, etc.
Key points

• Direct energy minimisation, no density mixing
• All quantities (Hamiltonian, energy, gradients) built with linear-scaling cost
  - NGWF localisation in real space, FFT-box technique in reciprocal space
  - Sparse Hamiltonian and overlap matrixes
  - Sparse density kernel ($\text{kernel\_cutoff}$)
• Linear-scaling iterative algorithms using sparse matrix algebra
• Matrix products preserve sparsity patterns, e.g. $\textbf{KSK}$ is less sparse than $\textbf{K}$
• Tensorially correct gradients
• Preconditioned NGWF optimisation