INVERTING SPARSE OVERLAP MATRICES

Chris-Kriton Skylaris

Acknowledgments

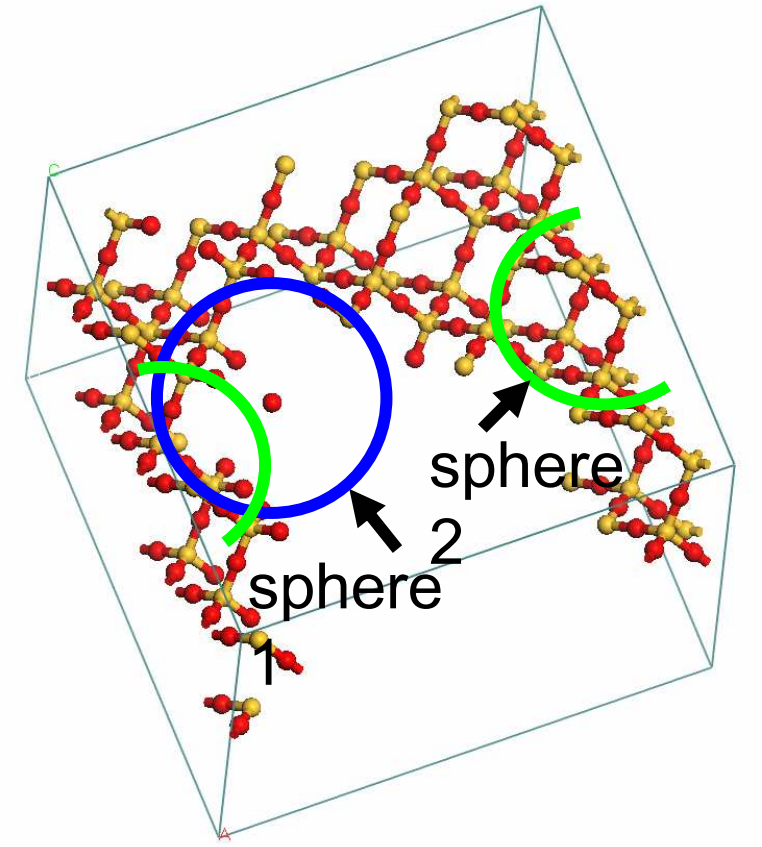
Arash A. Mostofi

Peter D. Haynes

Mike C. Payne

NGWFs are strictly localised in real space by being constrained inside spherical regions of fixed radius.

$$\phi_\alpha(\mathbf{r}) = \sum_n \psi_n(\mathbf{r}) M_{n\alpha}, \; S_{\alpha\beta} = \langle \phi_\alpha | \phi_\beta \rangle \text{ and } \psi_n(\mathbf{r}) = \phi_\alpha(\mathbf{r}) (M^{-1})^\alpha_{\ n}$$



Example:

ϕ_α, ϕ_β, ϕ_γ, ϕ_δ are limited to sphere 1.

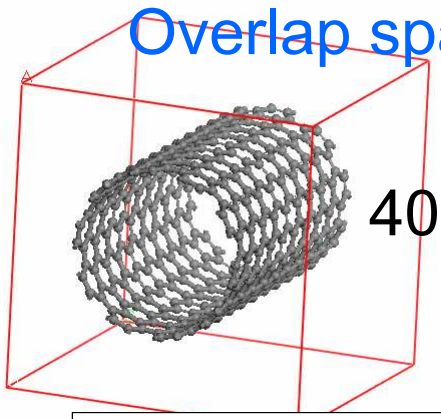
ϕ_ε, ϕ_ζ, ϕ_η are limited to sphere 2. Etc…

Benefits from using NGWFs: Sparse matrices

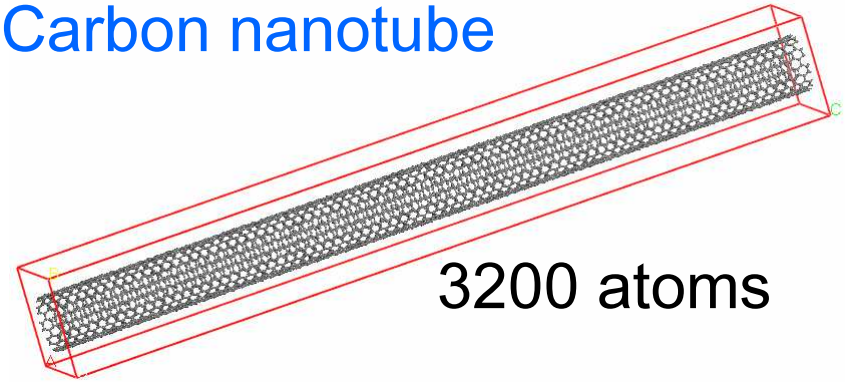
- The elements ($S_{\alpha\beta}$) of the NGWF overlap matrix are nonzero only for pairs of functions whose spheres overlap. Their number scales linearly with system size. The same is true for all operators in the NGWF representation (e.g. the Hamiltonian matrix elements $H_{\alpha\beta}$).

- However, the number of non-zero elements of a sparse matrix must be *less than 1%* for sparse matrix operations to be computationally more efficient than full-matrix operations.

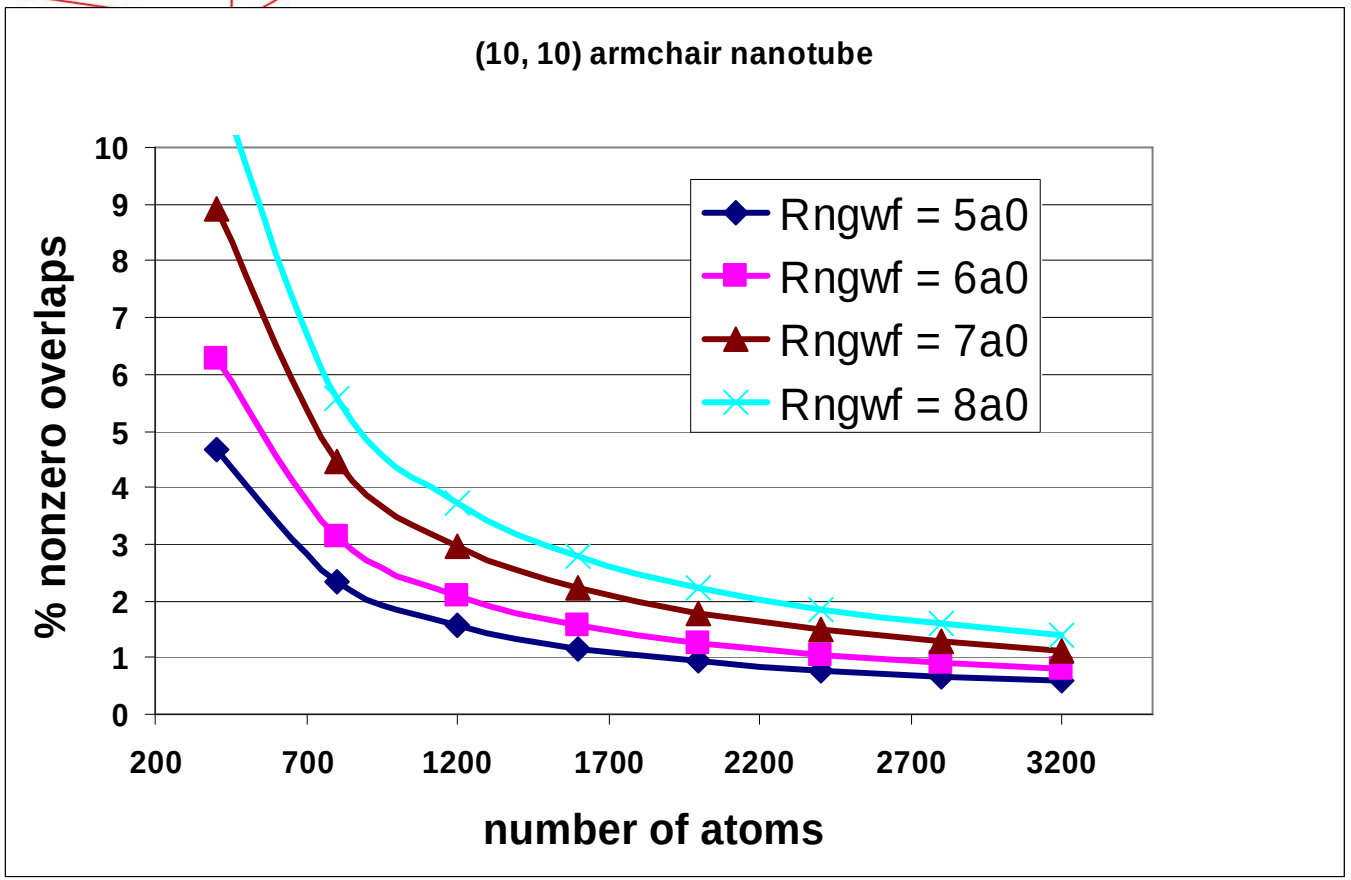
Overlap sparsity example: Carbon nanotube



400 atoms

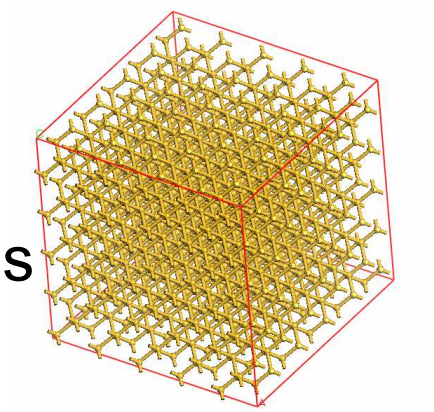


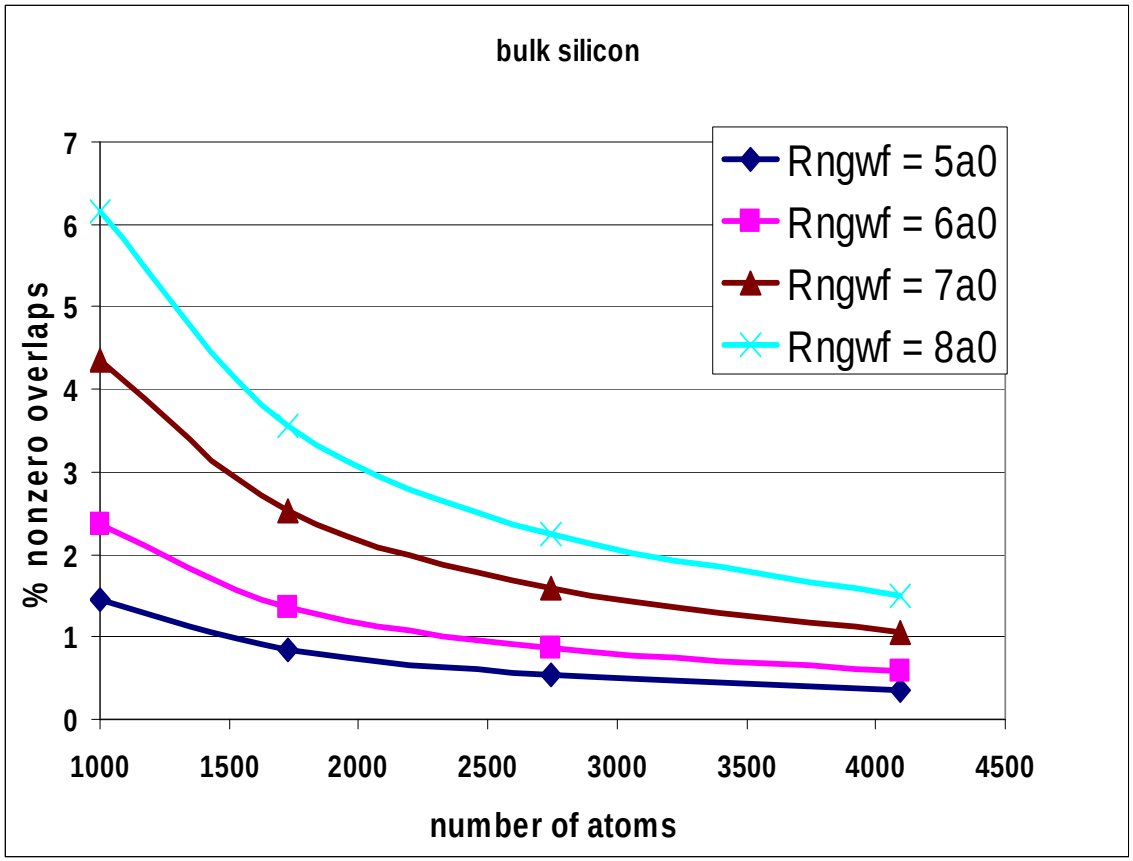
3200 atoms



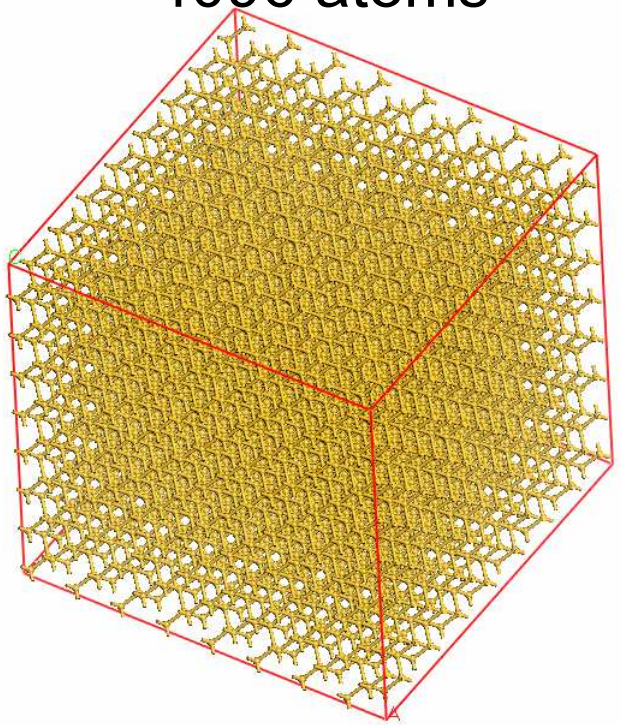
Overlap sparsity example: Bulk silicon

1000 atoms





4096 atoms



Expressing the band energy in terms of NGWFs

$$E[\{\psi_n\}] = 2\sum_n f_n \epsilon_n = 2\sum_n f_n \langle\psi_n|\hat{H}|\psi_n\rangle$$

Band energy becomes

$$E[\{\psi_n\}] = 2\sum_n f_n \langle\psi_n|\hat{H}|\psi_n\rangle = 2\sum_n f_n (M^{-\dagger})_n{}^{\alpha} \langle\phi_\alpha|\hat{H}|\phi_\beta\rangle (M^{-1})^{\beta}{}_n$$

$$= 2H_{\alpha\beta}K^{\beta\alpha} = E[\mathbf{K}, \phi_\alpha] \text{ where } K^{\beta\alpha} = \sum_n (M^{-1})^{\beta}{}_n \, f_n \, (M^{-\dagger})_n{}^{\alpha}$$

$$\rho(\mathbf{r}, \mathbf{r}') = \phi_\alpha(\mathbf{r}) K^{\alpha\beta} \phi^*_\beta(\mathbf{r}') \qquad \text{and} \qquad \lim_{|\mathbf{r}-\mathbf{r}'|\to\infty} \rho(\mathbf{r}, \mathbf{r}') = 0$$

The $\{\phi\}$ are localised by construction. What remains is to make $\mathbf{K}$ sparse by truncating its elements that correspond to spheres separated by more than some cutoff threshold distance.

Optimisation of E w.r.t. K

- Optimise $\mathbf{K}$ using conjugate gradients or steepest descents.
- Constraints on $\mathbf{K}$: Idempotency (i.e. f_n equal to 1 or 0) and constant number of electrons.
- Need search direction

$$-\nabla_{\mathbf{K}} E = -\frac{\partial E}{\partial K^{\beta\alpha}} = 4H_{\alpha\beta} \text{ (a covariant tensor).}$$

- The search direction should be a *contravariant* second order tensor as is $\mathbf{K}$:

$$K_{\beta\alpha}^{(j+1)} = K_{\beta\alpha}^{(j)} - \lambda 4 (S^{-1})^{\beta\gamma} H_{\gamma\delta}^{(j)} (S^{-1})^{\delta\alpha}.$$

E. Artacho, L. M. delBosch, *Phys. Rev. A*, 43 (1991) 5770.

Need to calculate S^{-1} in O(N) operations!

Hotteling's recursive improvement of $\mathbf{S^{-1}}$

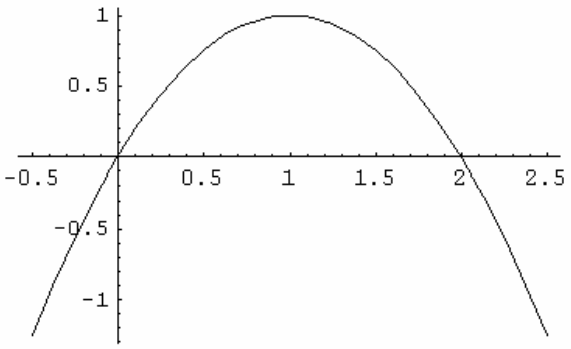
- Definitions

$$\text{residual: } \mathbf{R} = \mathbf{I} - \mathbf{S_0^{-1}S}$$

$$\text{expansion in } \mathbf{R}\text{: } \mathbf{S^{-1}} = (\mathbf{I} - \mathbf{R})^{-1}\mathbf{S_0} = (\mathbf{I} + \mathbf{R} + \mathbf{R^2} + ...)\mathbf{S_0}$$

- Hotelling's recurrence:

$$\mathbf{S^{-1}_{(2n+1)}} = \mathbf{2S_n^{-1} - S_n^{-1}SS_n^{-1}}$$



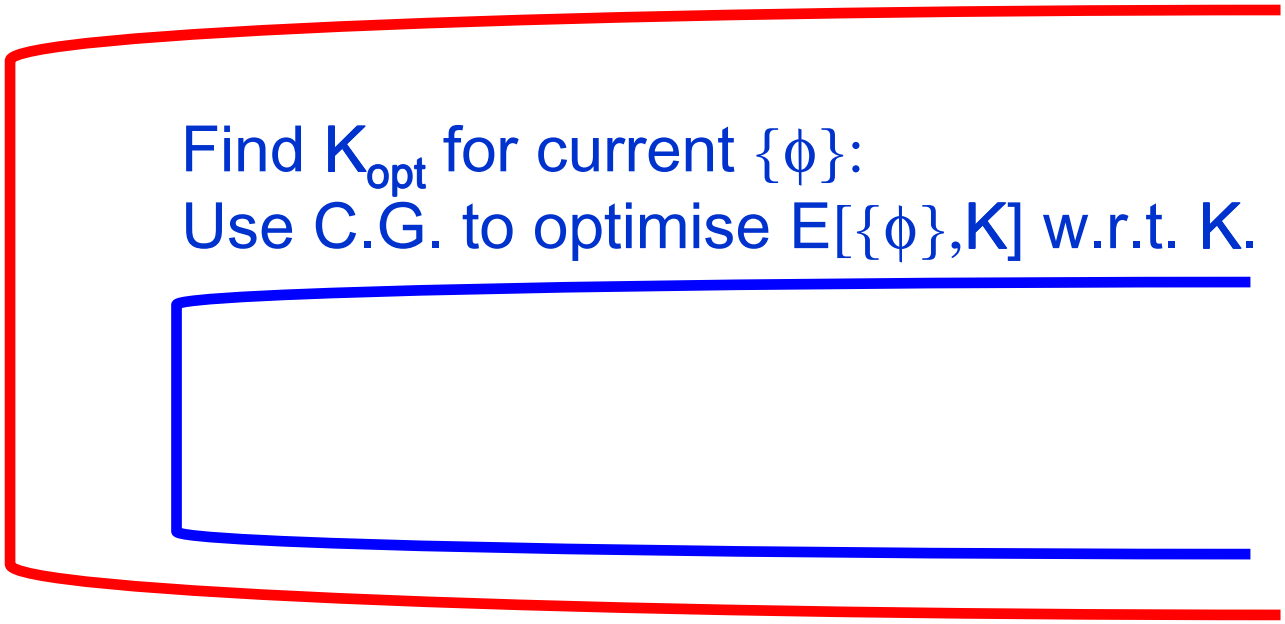
where

$$\mathbf{S_n^{-1}} = (\mathbf{1} + \mathbf{R} + \mathbf{R^2} + ... + \mathbf{R^n})\mathbf{S_0}$$

Involves only matrix multiplications that can be done in O(N) operations when the matrices are sufficiently sparse!

Total energy optimisation strategy: two nested loops.

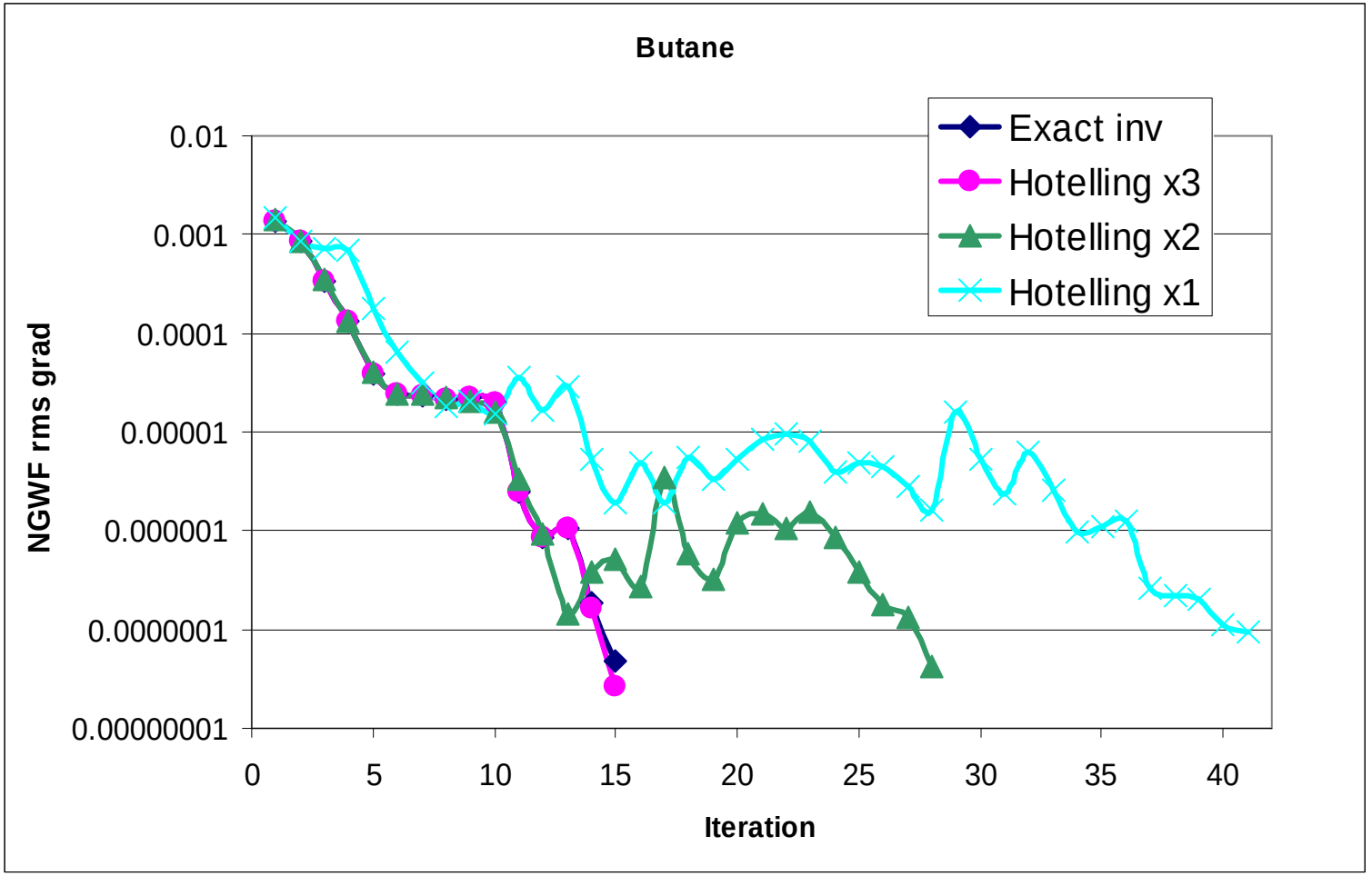
Use C.G. to optimise $E[\{\phi\},K_{opt}]$, w.r.t. $\{\phi\}$ while Hotelling-updating S^{-1} for every change in $\{\phi\}$.

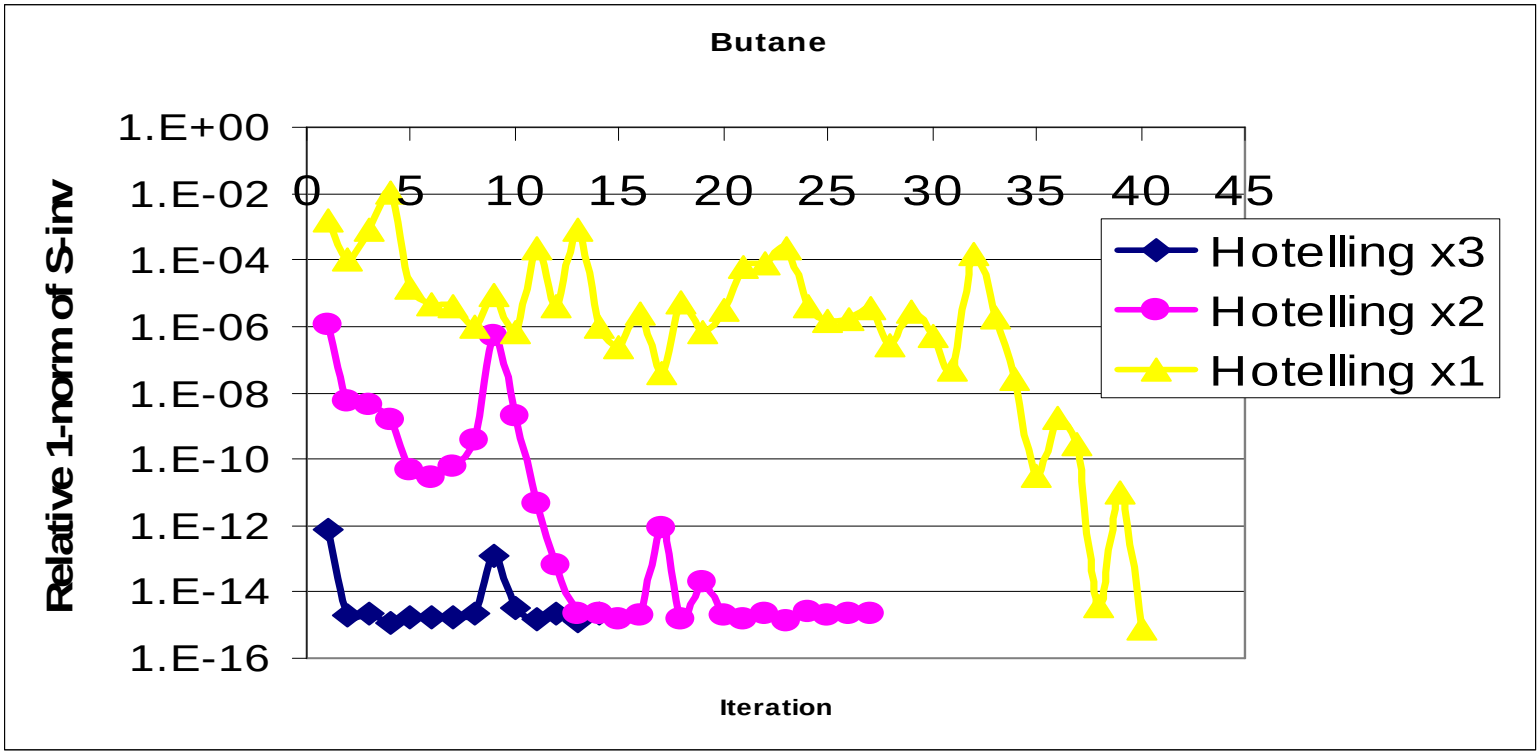


Impose on S^{-1} same sparsity pattern as K.

Example: Calculations on butane C_4H_{10}

Convergence of NGWFs with CG iterations





Approximation	Total Energy (Eh)
Exact S^{-1}	-28.674482206
Hotelling x3	-28.674482199
Hotelling x2	-28.674478544
Hotelling x1	-28.674485824