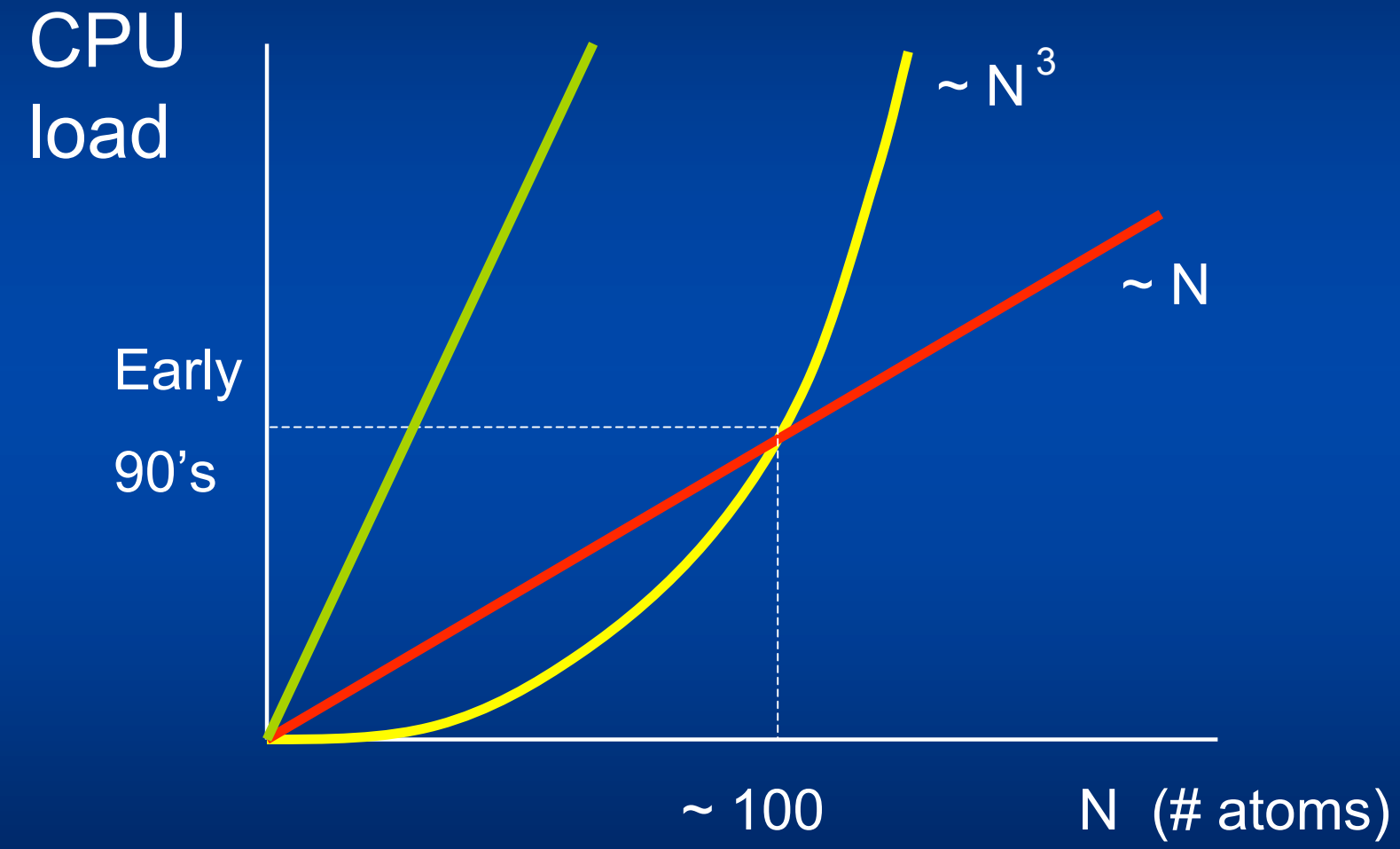
Linear scaling solvers based on Wannier-like functions

P. Ordejón

Institut de Ciència de Materials de Barcelona (CSIC)

Linear scaling = Order(N)



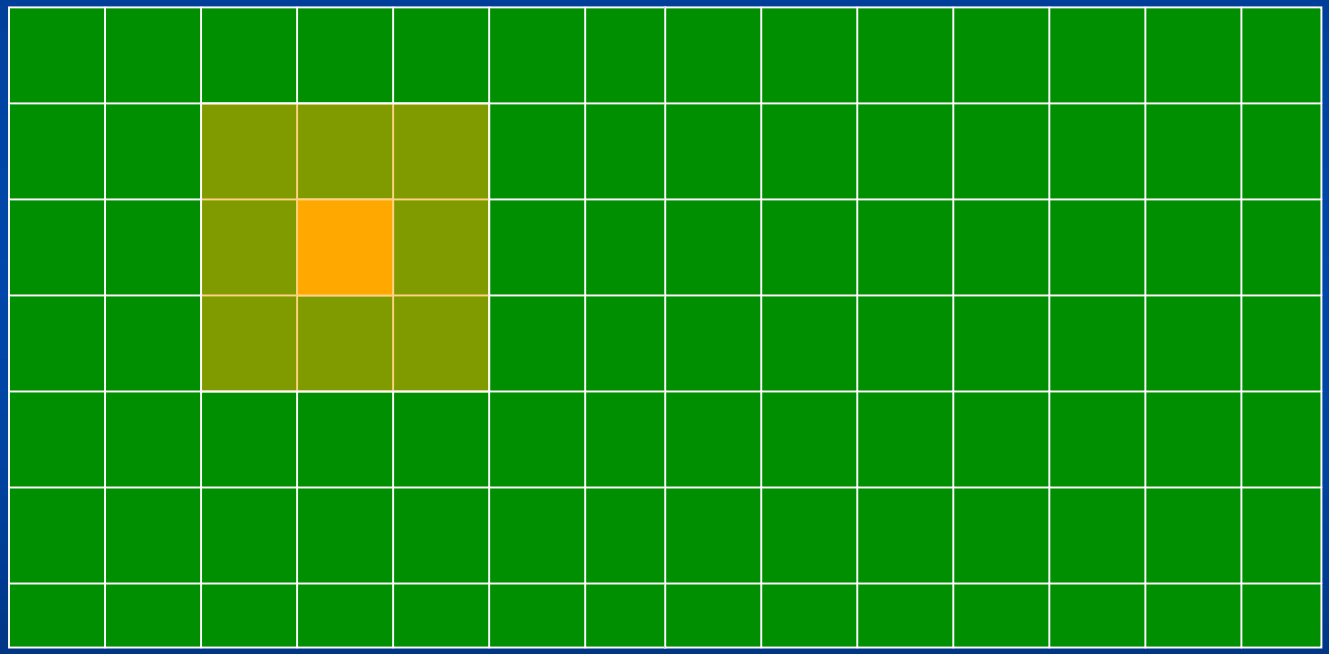
Order-N DFT

1. Find density and hamiltonian (80% of code)
2. Find “eigenvectors” and energy (20% of code)
3. Iterate SCF loop

Steps 1 and 3 spared in tight-binding schemes

Key to O(N): locality

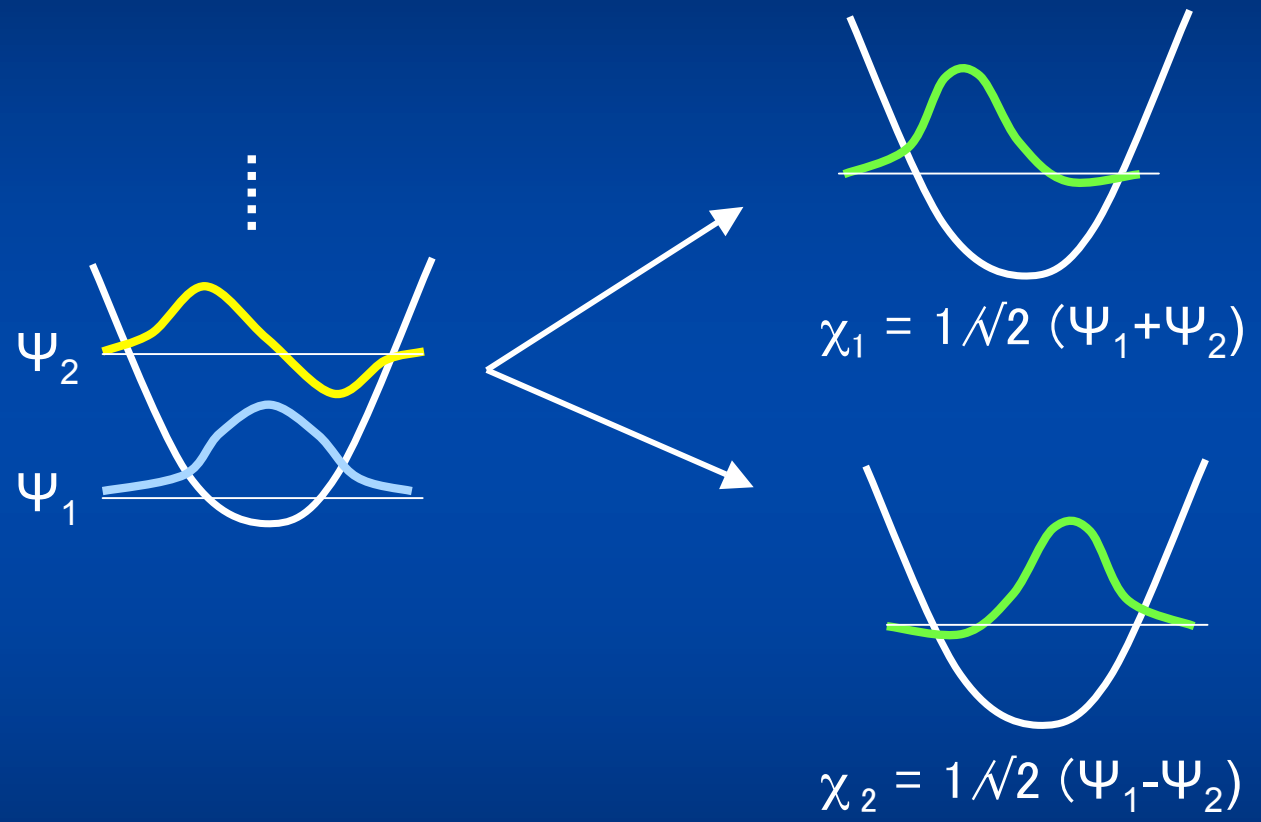
Large system



``Divide and conquer'' *W. Yang, Phys. Rev. Lett. 66, 1438 (1992)*

``Nearsightedness'' *W. Kohn, Phys. Rev. Lett. 76, 3168 (1996)*

Locality of Wave Functions



Wannier functions (crystals)

Localized Molecular Orbitals (molecules)

Locality of Wave Functions

Energy:

$$E = \langle \psi_1 | H | \psi_1 \rangle + \langle \psi_2 | H | \psi_2 \rangle = Tr_{occ}[H]$$

Unitary Transformation: $|\psi_1\rangle, |\psi_2\rangle \rightarrow |\chi_1\rangle, |\chi_2\rangle$

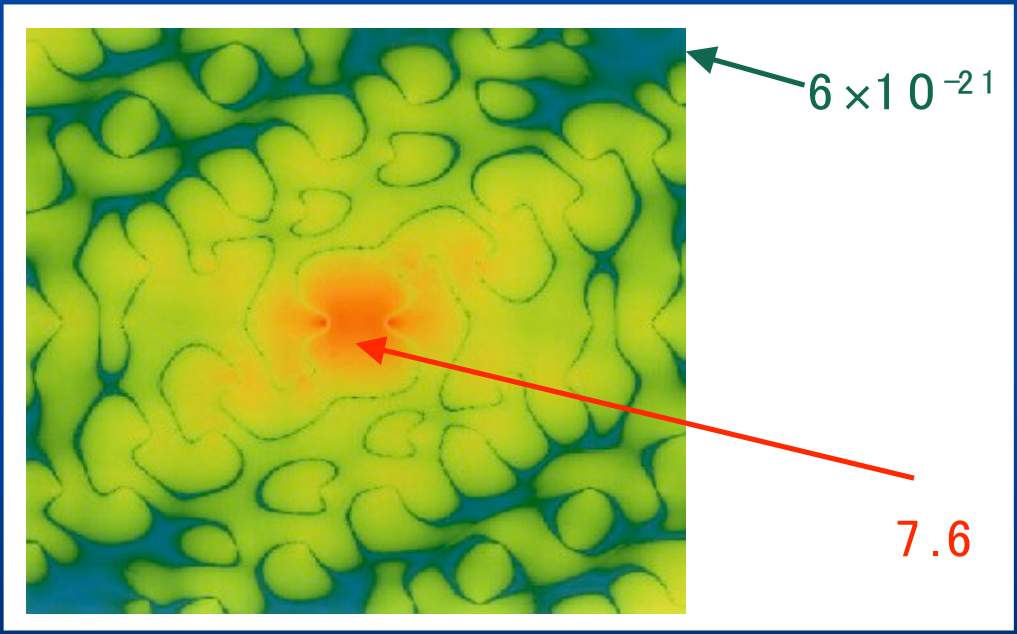
$$E = Tr_{occ}[H] = \langle \chi_1 | H | \chi_1 \rangle + \langle \chi_2 | H | \chi_2 \rangle$$

We do NOT need eigenstates!

We can compute energy with Loc. Wavefuncs.

Locality of Wave Functions

Exponential localization (insulators):

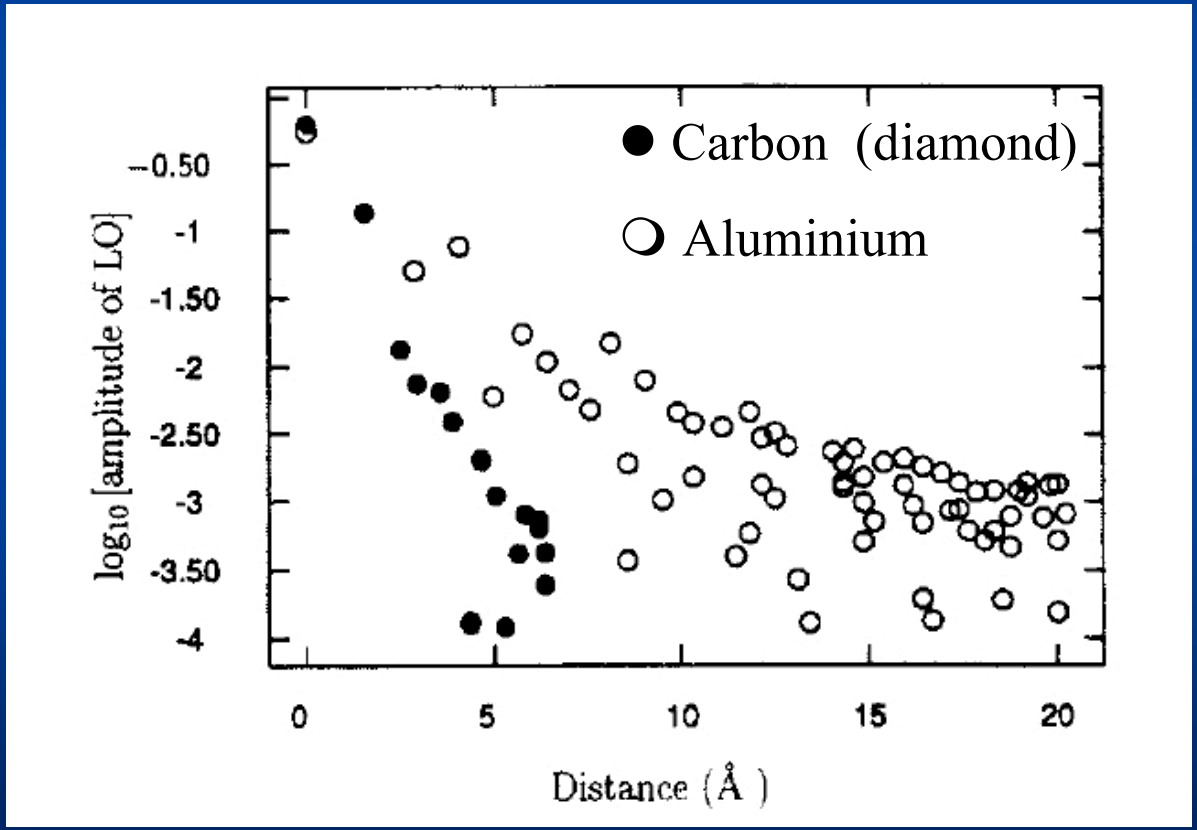


Wannier function in Carbon (diamond)

Drabold et al.

Locality of Wave Functions

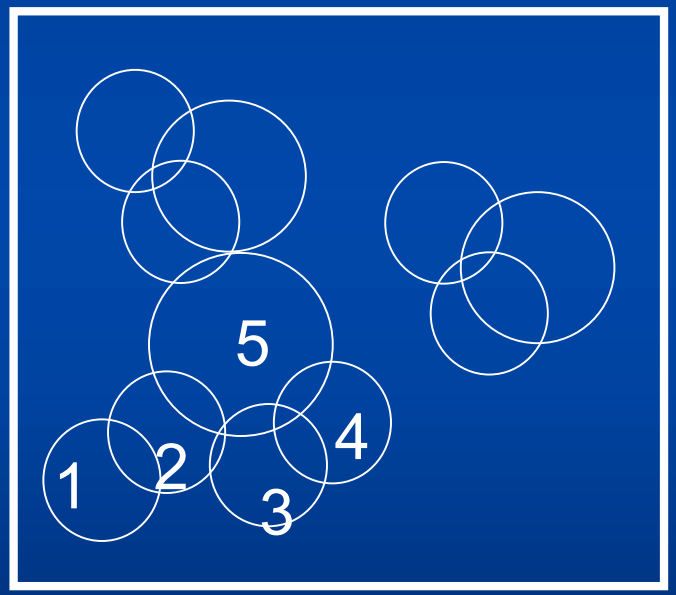
Insulators vs Metals:



Goedecker & Teter, PRB **51**, 9455 (1995)

Linear Scaling

Localization + Truncation



- Sparse Matrices

$$\rho_{ij} = \langle \phi_i | \hat{\rho} | \phi_j \rangle$$

- Truncation errors

$$\rho_{ij} \sim \exp[-\alpha |i-j|]$$

In systems with a gap.
Decay rate α depends on gap E_g

Linear Scaling Approaches

(Localized) object which is computed:

- wave functions 
- density matrix 

Approach to obtain the solution:

- minimization
- projection
- spectral

Reviews on O(N) Methods: Goedecker, RMP ’98

Ordejón, Comp. Mat. Sci.’98

Basis sets for linear-scaling DFT

- *LCAO: - Gaussian based + QC machinery*

  M. Challacombe, G. Scuseria, M. Head-Gordon ...

  - Numerical atomic orbitals (NAO)

  SIESTA

  S. Kenny &. A Horsfield (PLATO)

  OpenMX

- *Hybrid PW – Localized orbitals*

  - Gaussians *J. Hutter, M. Parrinello*

  - "Localized PWs"

  C. Skylaris, P, Haynes & M. Payne

- *B-splines in 3D grid*

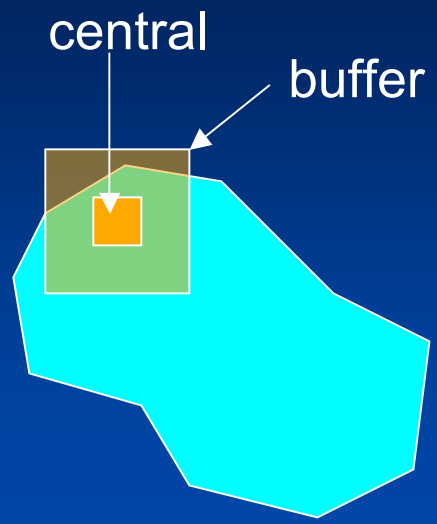
  D. Bowler & M. Gillan

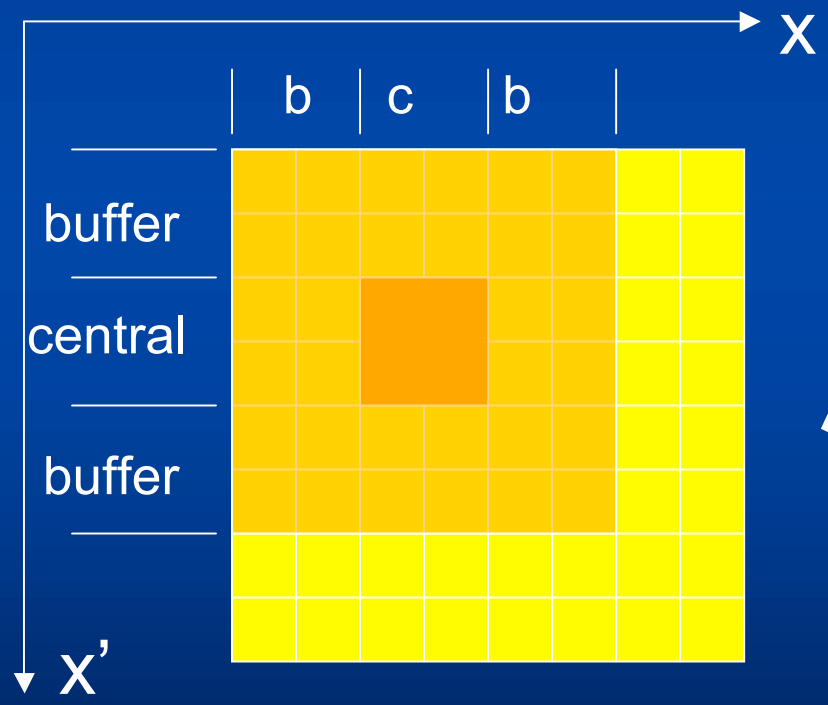
- *Finite-differences (nearly O(N))*

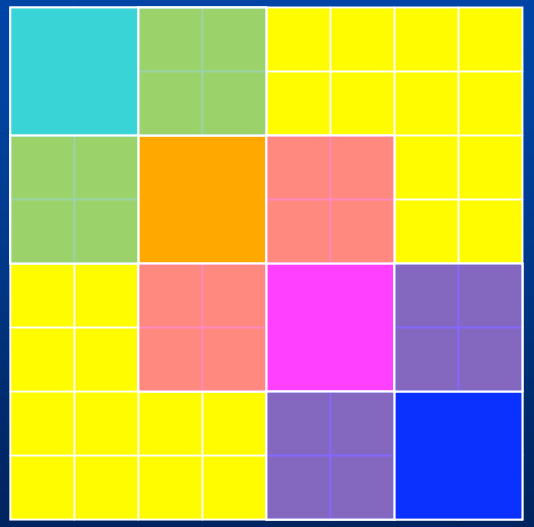
  J. Bernholc

Divide and conquer

Weitao Yang (1992)



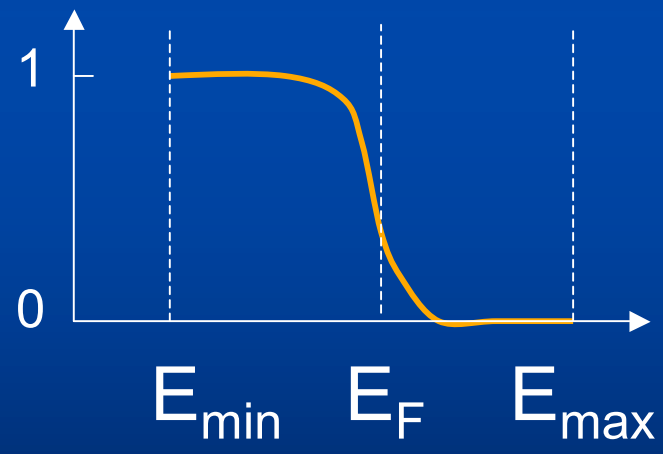




Fermi operator/projector

Goedecker & Colombo (1994)

$$f(E) = 1/(1+e^{E/kT}) \approx \sum_n c_n E^n$$



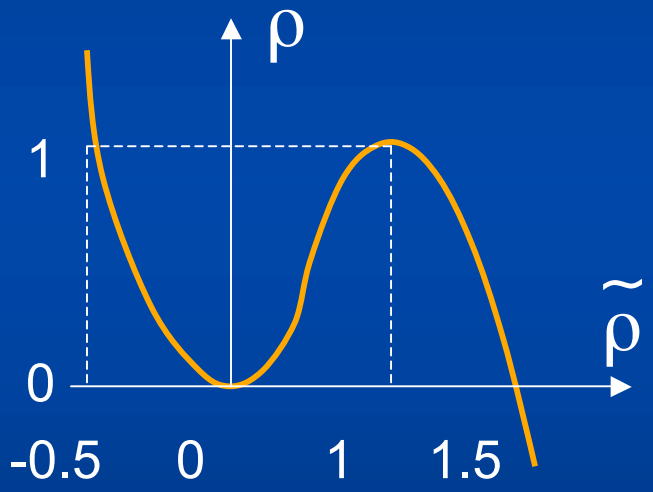
$$\hat{F} \approx \sum c_n \hat{H}^n$$

$$E_{tot} = Tr[\hat{F}\hat{H}]$$

$$N_{tot} = Tr[\hat{F}]$$

Density matrix functional

Li, Nunes & Vanderbilt (1993)



$$\rho_{\mu\nu} = 3\,\tilde{\rho}^{2}_{\mu\nu} - 2\,\tilde{\rho}^{3}_{\mu\nu}$$

$$E_{tot}(\tilde{\rho}_{\mu\nu}) = \Sigma_{\mu\nu}\,\rho_{\mu\nu}H_{\nu\mu} = \min$$

Wannier O(N) functional

- Mauri, Galli & Car, PRB **47**, 9973 (1993)
- Ordejón et al, PRB **48**, 14646 (1993)

$$S_{ij} = \langle \psi_i | \psi_j \rangle \qquad | \psi'_k \rangle = \Sigma_j | \psi_j \rangle S_{jk}^{-1/2}$$

$$E_{KS} = \Sigma_k \langle \psi'_k | \hat{H} | \psi'_k \rangle$$

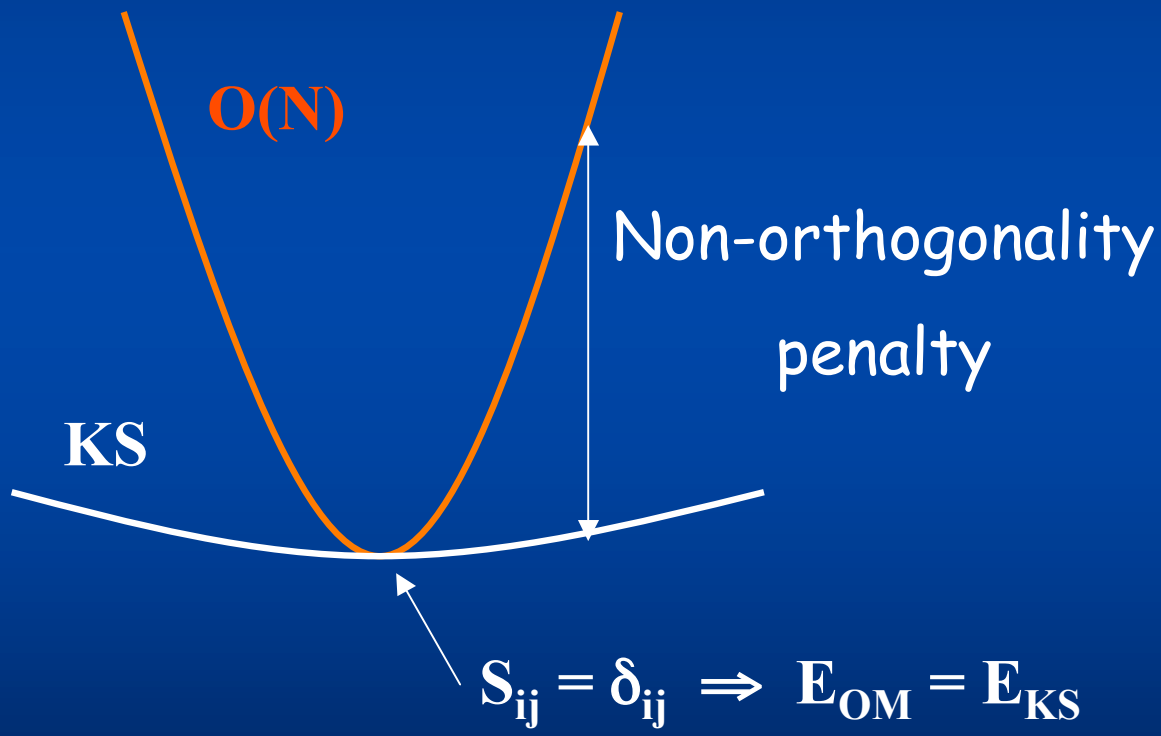
$$= \Sigma_{ijk} S_{ki}^{-1/2} \langle \psi_i | \hat{H} | \psi_j \rangle S_{jk}^{-1/2}$$

$$= \mathrm{Tr}_{occ}[S^{-1} H] \qquad \textbf{Kohn-Sham}$$

$$E_{OM} = \mathrm{Tr}_{occ}[(2I-S) H] \qquad \textbf{Order-N}$$

$$= \mathrm{Tr}_{occ}[H] + \mathrm{Tr}_{occ}[(I-S) H]$$

Order-N vs KS functionals



Chemical potential

Kim, Mauri & Galli, PRB **52**, 1640 (1995)

$$\rho(r) = 2\Sigma_{ij}\, \psi_i(r)\, (2\delta_{ij}-S_{ij})\, \psi_j(r)$$

$$E_{OM} = Tr_{occ}[\, (2I-S)\, H\,] \qquad \text{\# states = \# electron pairs}$$

$\Rightarrow$ **Local minima**

$$E_{KMG} = Tr_{occ+}[\, (2I-S)\, (H-\eta S)\,] \qquad \text{\# states > \# electron pairs}$$

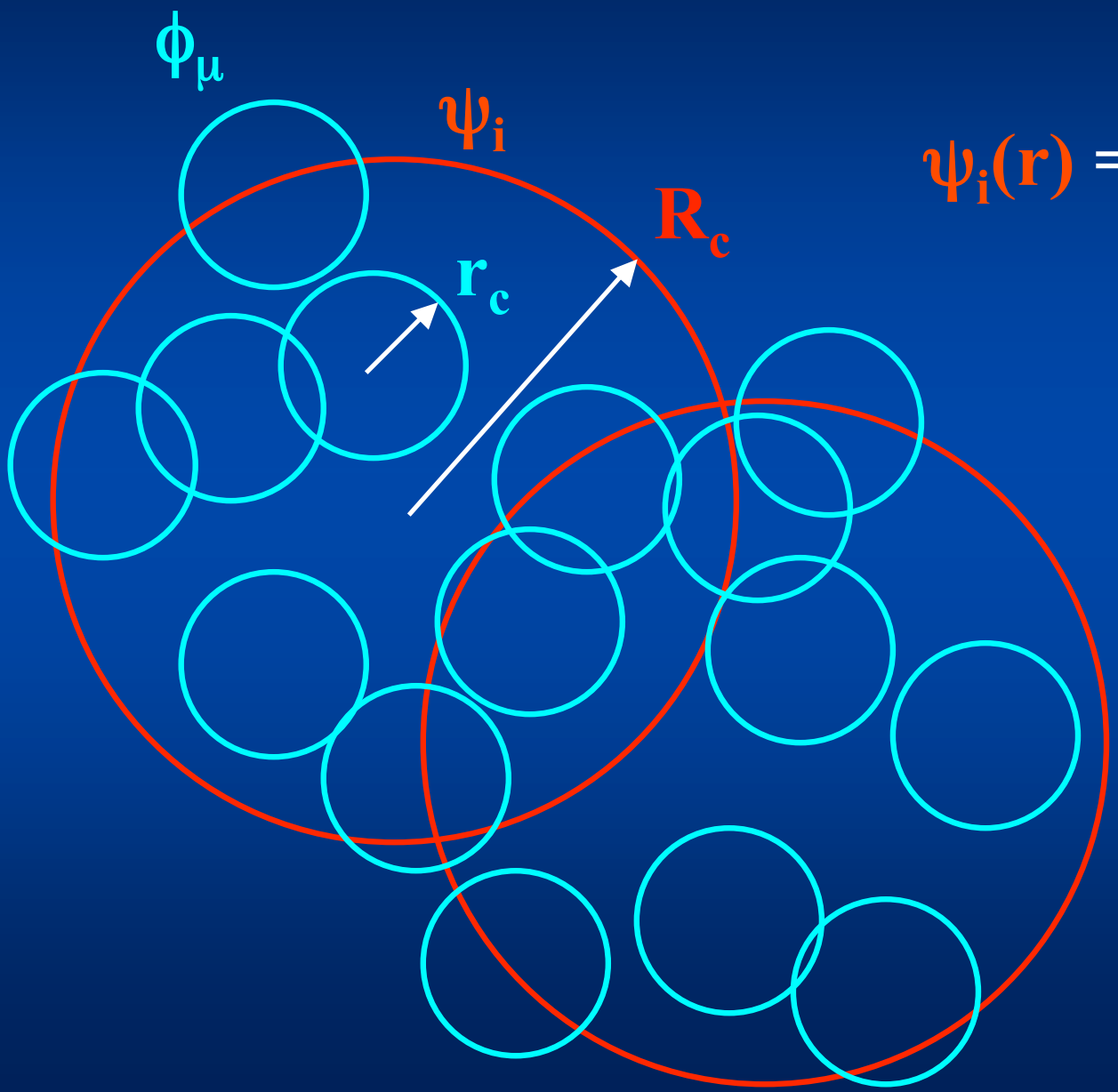
η = **chemical potential (Fermi energy)**

$E_i > \eta \quad \Rightarrow \quad |\psi_i| \approx 0$

$E_i < \eta \quad \Rightarrow \quad |\psi_i| \approx 1$

Difficulties	Solutions
Stability of N(η)	Initial diagonalization / Estimate of η
First minimization of E_{KMG}	Reuse previous solutions

Orbital localization

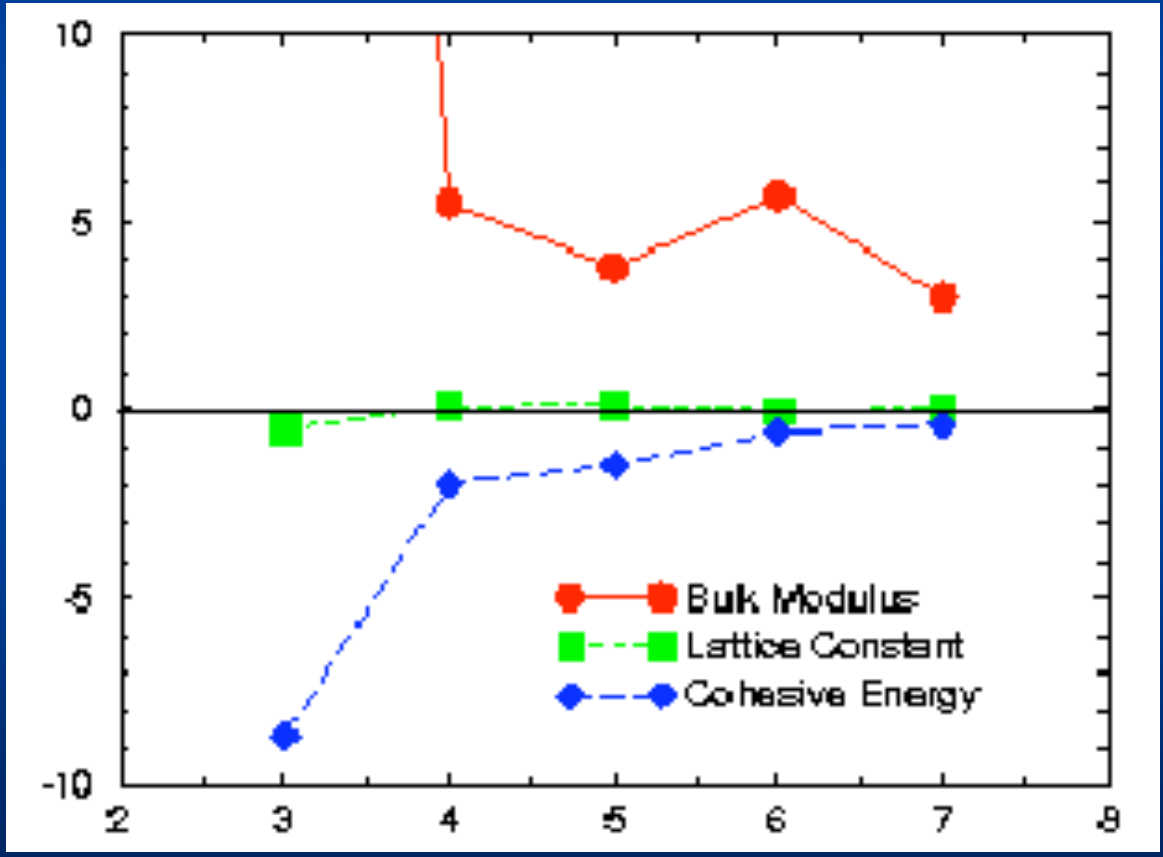


$$\psi_i(\mathbf{r}) = \Sigma_\mu \, c_{i\mu} \, \phi_\mu(\mathbf{r})$$

Convergence with localisation radius

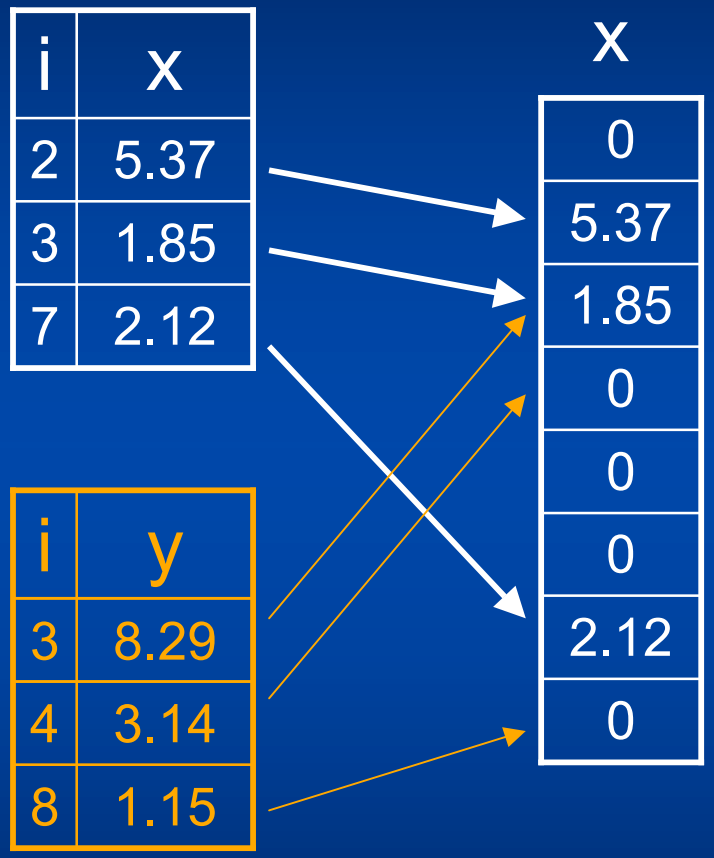
Si supercell, 512 atoms

Relative Error (%)



R_c (Ang)

Sparse vectors and matrices



8.29 × 1.85 = 15.34

3.14 × 0 = 0

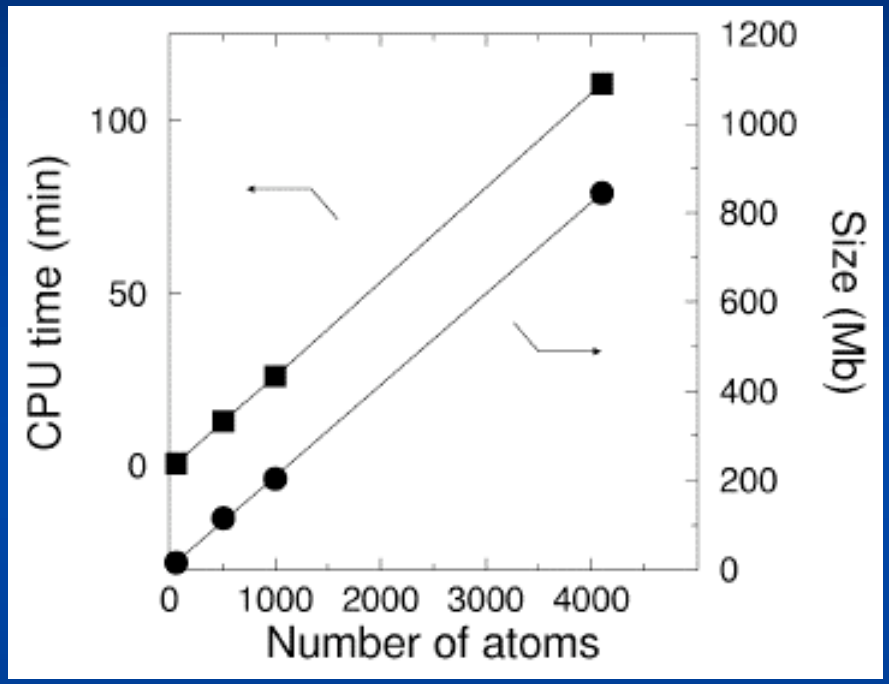
1.15 × 0 = 0

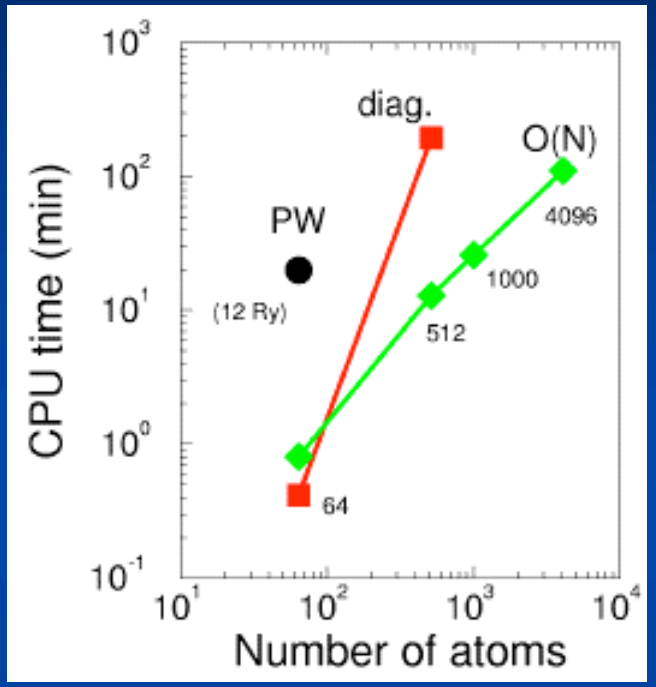
Sum 15.34

Restore to zero $x_i \neq 0$ **only**

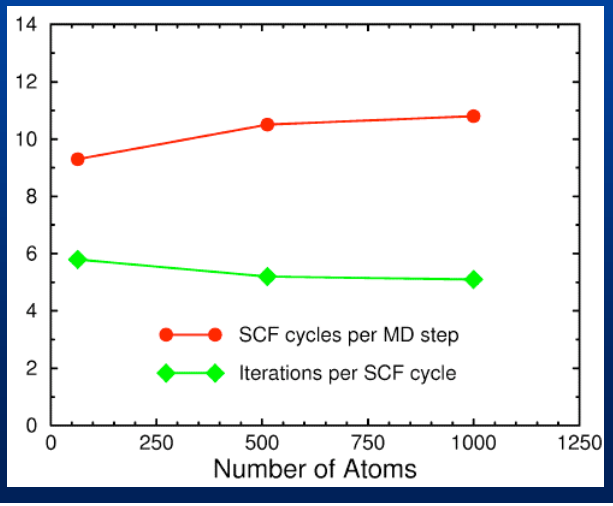
Actual linear scaling

c-Si supercells, single-ζ





Single Pentium III 800 MHz. 1 Gb RAM



132.000 atoms in 64 nodes

Linear scaling solver: practicalities in SIESTA

P. Ordejón

Institut de Ciència de Materials de Barcelona (CSIC)

Order-N in SIESTA (1)

SCF

Calculate Hamiltonian

Minimize E_{KS} with respect to WFs (GC minimization)

Build new charge density from WFs

Energy Functional Minimization

- **Start from initial LWFs** (from scratch or from previous step)

$$\chi_i(r) = \Sigma_\mu \, c_{i\mu} \, \phi_\mu(r)$$

- **Minimize Energy Functional w.r.t. $c_{i\mu}$**

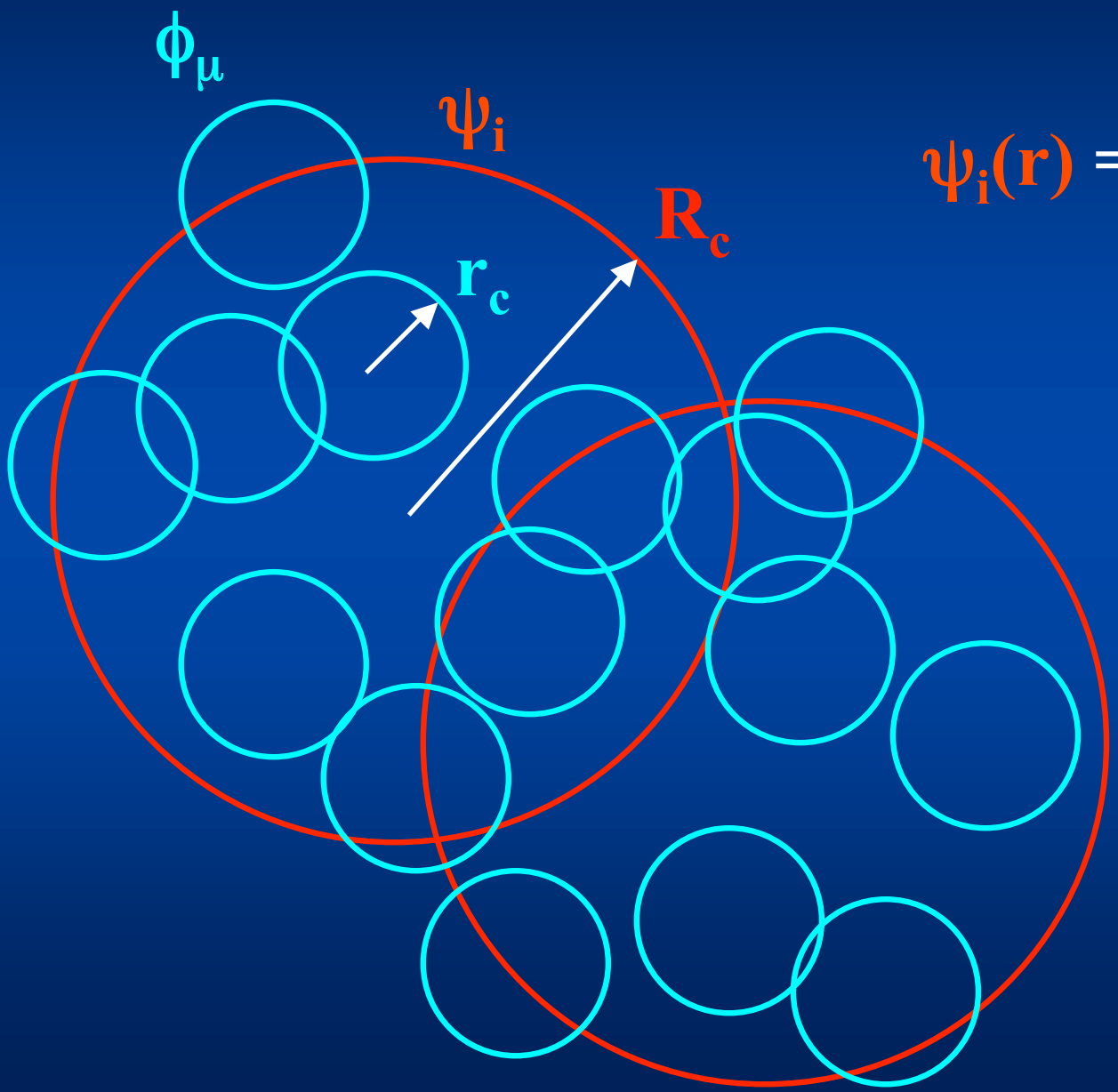
$$E_{OM} = Tr_{occ}[\,(2I-S)\,H\,] \quad \text{or}$$

$$E_{KMG} = Tr_{occ+}[\,(2I-S)\,(H-\eta S)\,]$$

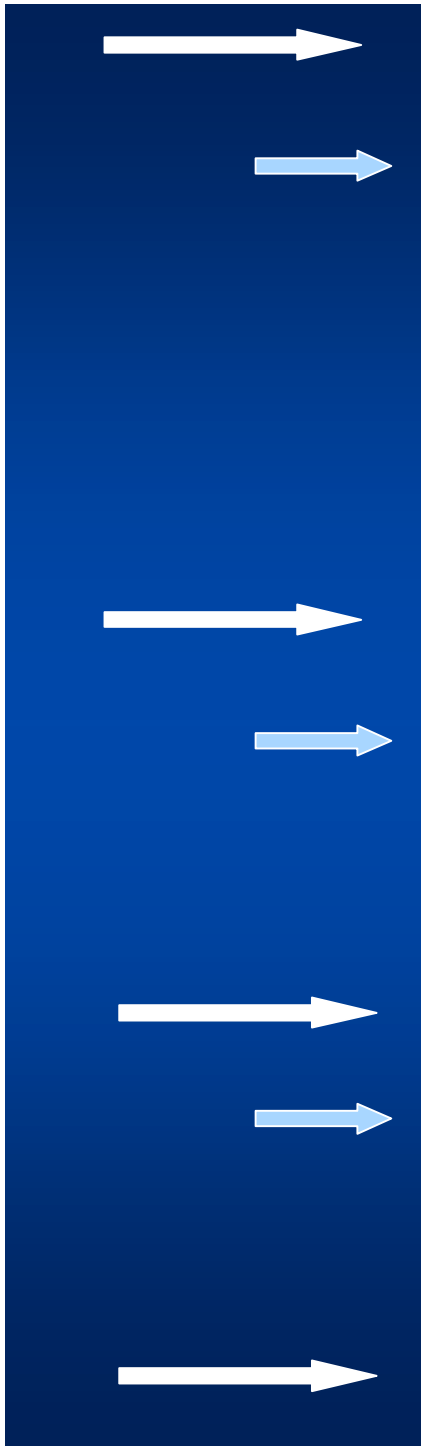
- **Obtain new density**

$$\rho(r) = 2\Sigma_{ij} \, \psi_i(r) \, (2\delta_{ij}-S_{ij}) \, \psi_j(r)$$

Orbital localization



$$\psi_i(\mathbf{r}) = \Sigma_\mu \, c_{i\mu} \, \phi_\mu(\mathbf{r})$$



```
siesta: iscf =    7
Eharris(eV) =      -1868.7292   E_KS(eV) =      -1868.9207   dDmax =  0.0072


ordern: enum =     32.0000
cgwf: iter =    1      grad =          -0.000017      Eb(Ry) =     -16.103689
cgwf: iter =    2      grad =          -0.000041      Eb(Ry) =     -16.103690
cgwf: iter =    3      grad =          -0.000005      Eb(Ry) =     -16.103690
cgwf: iter =    4      grad =          -0.000008      Eb(Ry) =     -16.103691
cgwf: iter =    5      grad =          -0.000006      Eb(Ry) =     -16.103691
cgwf: iter =    6      grad =          -0.000001      Eb(Ry) =     -16.103691
cgwf: iter =    7      grad =          -0.000001      Eb(Ry) =     -16.103691
cgwf: iter =    8      grad =          -0.000001      Eb(Ry) =     -16.103691
cgwf: iter =    9      grad =          -0.000001      Eb(Ry) =     -16.103691
cgwf: iter =   10      grad =           0.000000      Eb(Ry) =     -16.103691

cgwf:  CG tolerance reached

denmat: qtot (before DM normalization) =      32.0000
ordern: qtot (after  DM normalization) =      32.0000

siesta: iscf =    8
Eharris(eV) =      -1868.7291   E_KS(eV) =      -1868.8286   dDmax =  0.0046


ordern: enum =     32.0000
cgwf: iter =    1      grad =          -0.000001      Eb(Ry) =     -16.100366
cgwf: iter =    2      grad =          -0.000001      Eb(Ry) =     -16.100366
cgwf: iter =    3      grad =           0.000000      Eb(Ry) =     -16.100366

cgwf:  CG tolerance reached

denmat: qtot (before DM normalization) =      32.0000
ordern: qtot (after  DM normalization) =      32.0000

siesta: iscf =    9
Eharris(eV) =      -1868.7290   E_KS(eV) =      -1868.6741   dDmax =  0.0030


ordern: enum =     32.0000
cgwf: iter =    1      grad =          -0.000001      Eb(Ry) =     -16.096840
cgwf: iter =    2      grad =          -0.000002      Eb(Ry) =     -16.096840
cgwf: iter =    3      grad =           0.000000      Eb(Ry) =     -16.096840

cgwf:  CG tolerance reached

denmat: qtot (before DM normalization) =      32.0000
ordern: qtot (after  DM normalization) =      32.0000

siesta: iscf =   10
Eharris(eV) =      -1868.7290   E_KS(eV) =      -1868.7273   dDmax =  0.0003
```

Order-N in SIESTA (2)

- Practical problems:
  - Minimization of E versus WFs:
    - First minimization is hard!!! (~1000 CG iterations)
    - Next minimizations are much faster (next SCF and MD steps)
    - ALWAYS save SystemName.LWF and SystemName.DM files!!!!
  - The Chemical Potential (in Kim's functional):
    - Data on input (ON.Eta). Problem: can change during SCF and dynamics.
    - Possibility to estimate the chemical potential in O(N) operations
    - If chosen ON.Eta is inside a band (conduction or valence), the minimization often becomes unstable and diverges
    - Solution I: use chemical potential estimated on the run
    - Solution II: do a previous diagonalization

Example of instability related to a wrong chemical potential

```
cgwf: iter =    1      grad =           -1.328841      Eb(Ry) =      -2.582441
cgwf: iter =    2      grad =           -3.867501      Eb(Ry) =      -2.710455
cgwf: iter =    3      grad = -3309930340.938110      Eb(Ry) =   -6532.475547
cgwf: iter =    4      grad = ******************      Eb(Ry) = **************
cgwf: iter =    5      grad = -NaN                    Eb(Ry) = **************
```

Order-N in SIESTA (3)

- SolutionMethod OrderN
- ON.Functional
  Ordejon-Mauri or Kim (def)
- ON.MaxNumIter
  Max. iterations in CG minim. (WFs)
- ON.Etol
  Tolerance in the energy minimization
  $2(E_n-E_{n-1})/(E_n+E_{n-1}) < \text{ON.Etol}$
- ON.RcLWF
  Localisation radius of WFs

Order-N in SIESTA (4)

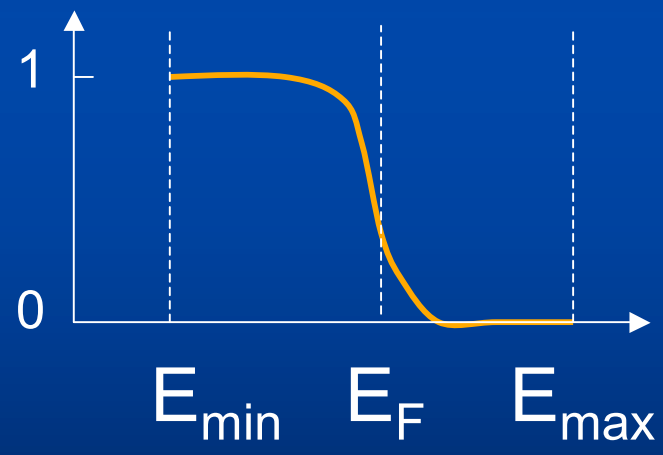
- ON.Eta (energy units)
  Chemical Potential (Kim)
  Shift of Hamiltonian (Ordejon-Mauri)

- ON.ChemicalPotential
- ON.ChemicalPotentialUse
- ON.ChemicalPotentialRc
- ON.ChemicalPotentialTemperature
- ON.ChemicalPotentialOrder

Fermi operator/projector

Goedecker & Colombo (1994)

$$f(E) = 1/(1+e^{E/kT}) \approx \sum_n c_n E^n$$



$$\hat{F} \approx \sum c_n \hat{H}^n$$

$$E_{tot} = Tr[\hat{F}\hat{H}]$$

$$N_{tot} = Tr[\hat{F}]$$