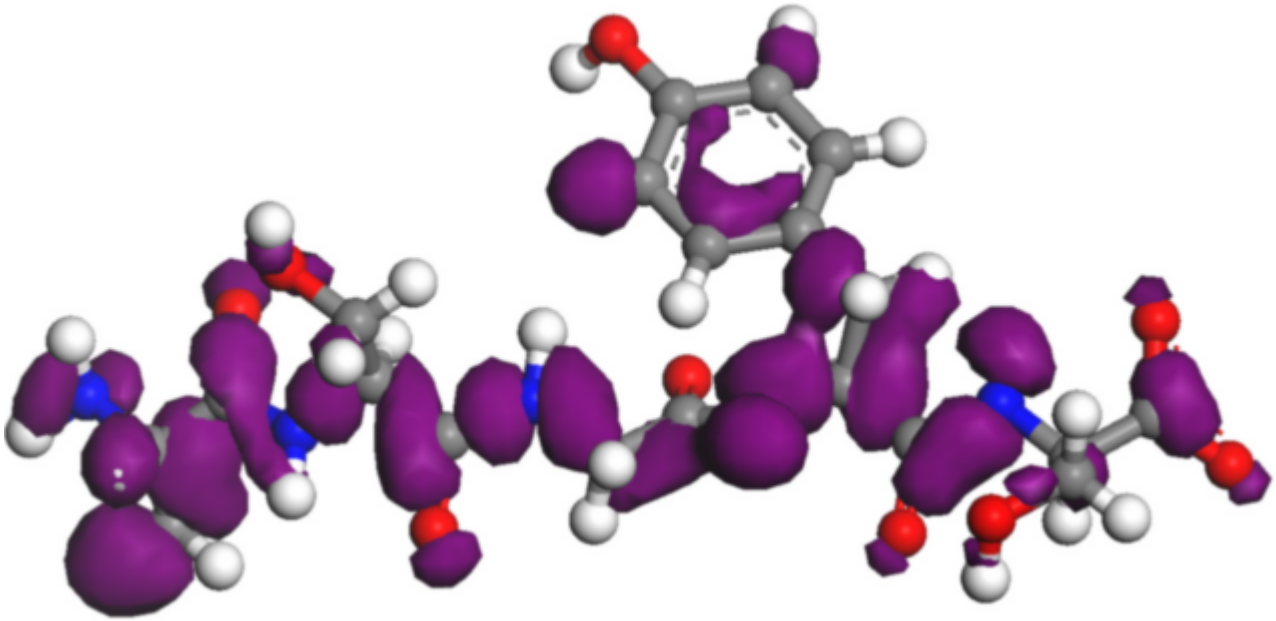
Imperial College London

THE THOMAS YOUNG CENTRE TyC

Introduction to linear-scaling DFT

Peter Haynes

Kohn-Sham equations



$$\left[-\frac{1}{2}\nabla^2 + V_\sigma^{\text{eff}}\right]\psi_{n\mathbf{k}\sigma} = \varepsilon_{n\mathbf{k}\sigma}\psi_{n\mathbf{k}\sigma}$$

The origin of the $O(N^3)$ problem

- Physicists:
  - Typically employ large basis sets of simple functions e.g. plane waves
  - Computational effort dominated by FFTs
  - Asymptotic N^3 scaling from orthogonality constraint

The origin of the $O(N^3)$ problem

- Physicists:
  - Typically employ large basis sets of simple functions e.g. plane waves
  - Computational effort dominated by FFTs
  - Asymptotic N^3 scaling from orthogonality constraint
- Chemists:
  - Typically employ small basis sets of more complicated functions e.g. contracted Gaussians
  - Computational effort dominated by building the Fock matrix

Simplifications

$$\left[-\frac{1}{2}\nabla^2 + V_\sigma^{\text{eff}}\right]\psi_{n\mathbf{k}\sigma} = \varepsilon_{n\mathbf{k}\sigma}\psi_{n\mathbf{k}\sigma}$$

Simplifications

$$\left[-\frac{1}{2}\nabla^2 + V_\sigma\right]\psi_{n\mathbf{k}\sigma} = \varepsilon_{n\mathbf{k}\sigma}\psi_{n\mathbf{k}\sigma}$$

- No self-consistency

Simplifications

$$\left[-\frac{1}{2}\nabla^2 + V\right]\psi_{n\mathbf{k}} = \varepsilon_{n\mathbf{k}}\psi_{n\mathbf{k}}$$

- No self-consistency
- No spin

Simplifications

$$\left[-\tfrac{1}{2}\nabla^2 + V\right]\psi_n = \varepsilon_n \psi_n$$

- No self-consistency
- No spin
- Sample Brillouin zone at Γ only

Simplifications

$$|\psi_n\rangle = \sum_k |\phi_k\rangle c_{kn} \quad \Rightarrow \quad \sum_j H_{ij} c_{jn} = c_{jn} \varepsilon_n$$

- No self-consistency
- No spin
- Sample Brillouin zone at Γ only
- (Localised) orthogonal basis set

Simplifications

$$|\psi_n\rangle = \sum_k |\phi_k\rangle c_{kn} \quad \Rightarrow \quad \sum_j H_{ij} c_{jn} = c_{in} \varepsilon_n$$

- No self-consistency
- No spin
- Sample Brillouin zone at Γ only
- (Localised) orthogonal basis set
  - *M* basis functions → H is *M* x *M* matrix

$$H_{ij} = \langle \phi_i | \left(-\tfrac{1}{2}\nabla^2 + V \right) | \phi_j \rangle = \sum_n c_{in} \varepsilon_n c^*_{jn}$$

→ full diagonalization $O(M^3)$

Simplifications

$$|\psi_n\rangle = \sum_k |\phi_k\rangle c_{kn} \quad \Rightarrow \quad \sum_j H_{ij} c_{jn} = c_{in} \varepsilon_n$$

- No self-consistency
- No spin
- Sample Brillouin zone at Γ only
- (Localised) orthogonal basis set
  - *M* basis functions → H is *M* x *M* matrix
  - *N* lowest states required

  → iterative diagonalization $O(N^2M)$

Total energy methods

- Energy of the Kohn-Sham system:

$$E = \sum_{n}^{\text{occ}} \varepsilon_n$$

Total energy methods

- Energy of the Kohn-Sham system:

$$E = \sum_{n}^{\text{all}} f_n \varepsilon_n$$

- Introduce occupation numbers f_n:
  - 1 for occupied states
  - 0 for unoccupied states

Total energy methods

- Energy of the Kohn-Sham system:

$$E = \sum_{n}^{\text{all}} f_n \varepsilon_n$$

- Introduce occupation numbers f_n:
  - 1 for occupied states
  - 0 for unoccupied states
- Finite temperature:

$$f_n = f(\varepsilon_n) \qquad f(\varepsilon) = \frac{1}{1 + \exp\left(\frac{\varepsilon - \mu}{k_B T}\right)}$$

Total energy methods

- Energy of the Kohn-Sham system:

$$E = \sum_{n}^{\text{all}} f_n \varepsilon_n$$

$$E = \mathrm{tr}\left[\begin{pmatrix} \varepsilon_1 & 0 & 0 & \cdots & 0 \\ 0 & \varepsilon_2 & 0 & \cdots & 0 \\ 0 & 0 & \varepsilon_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \varepsilon_N \end{pmatrix} \begin{pmatrix} f_1 & 0 & 0 & \cdots & 0 \\ 0 & f_2 & 0 & \cdots & 0 \\ 0 & 0 & f_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & f_N \end{pmatrix} \right]$$

Off-diagonal representation

$$E = \mathrm{tr}\left[\begin{pmatrix} \varepsilon_1 & 0 & 0 & \cdots & 0 \\ 0 & \varepsilon_2 & 0 & \cdots & 0 \\ 0 & 0 & \varepsilon_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \varepsilon_N \end{pmatrix} \begin{pmatrix} f_1 & 0 & 0 & \cdots & 0 \\ 0 & f_2 & 0 & \cdots & 0 \\ 0 & 0 & f_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & f_N \end{pmatrix} \right]$$

- Trace is invariant under similarity transformation:

$$c_{in} = \langle \phi_i | \psi_n \rangle \quad \Rightarrow \quad F_{ij} = \sum_n c_{in} f_n c^*_{jn}$$

$$E = \mathrm{tr}(FH) = \sum_{ij} F_{ij} H_{ji}$$

Density matrix

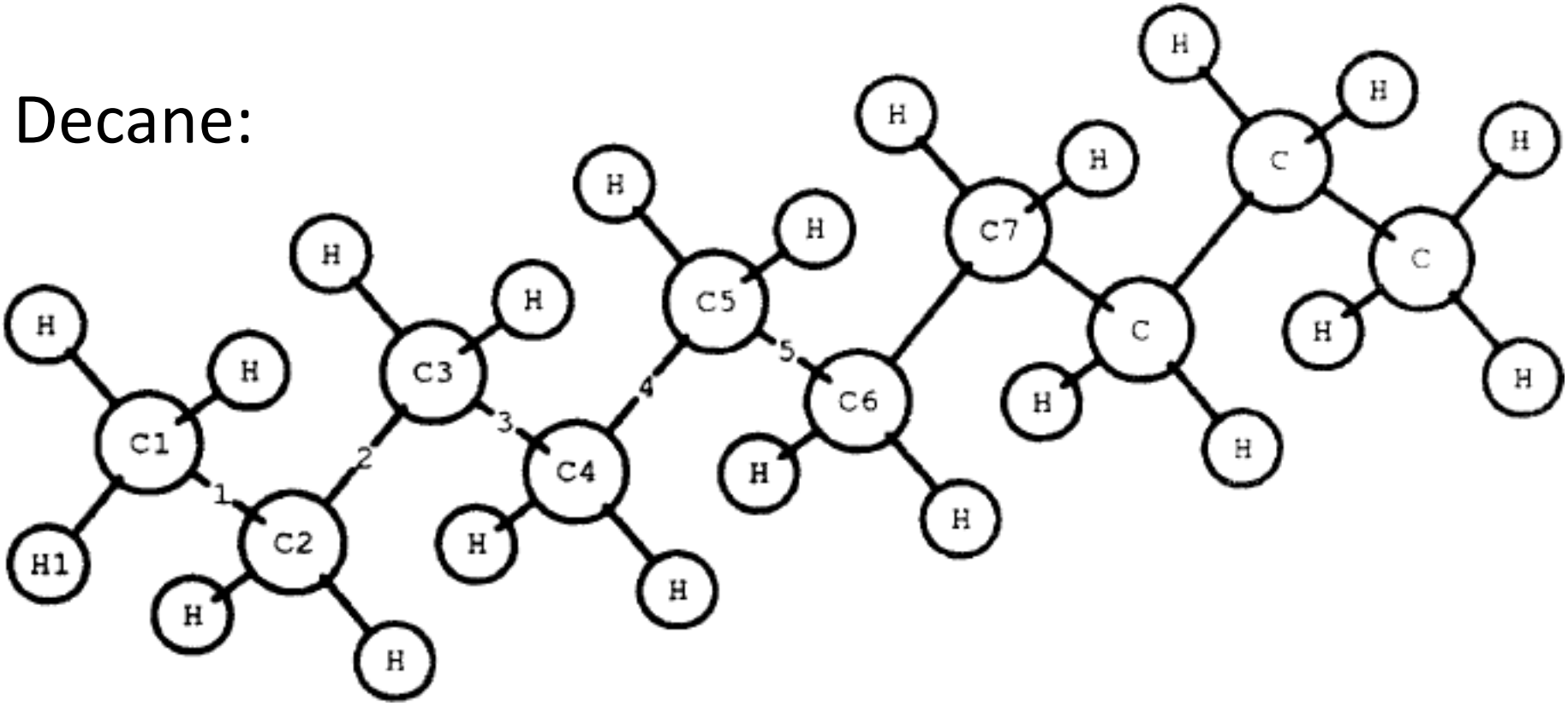
- F is the density matrix
  - F commutes with H (simultaneously diagonalizable)
  - Trace of F is the number of electrons (sum of occupation numbers)
  - At zero temperature F is idempotent: $F^2 = F$

- Solving the Schrödinger equation is equivalent to finding the F that minimizes E subject to the above conditions

Nearsightedness

Hierse & Stechel, *Phys. Rev. B* **50**, 17811 (1994)

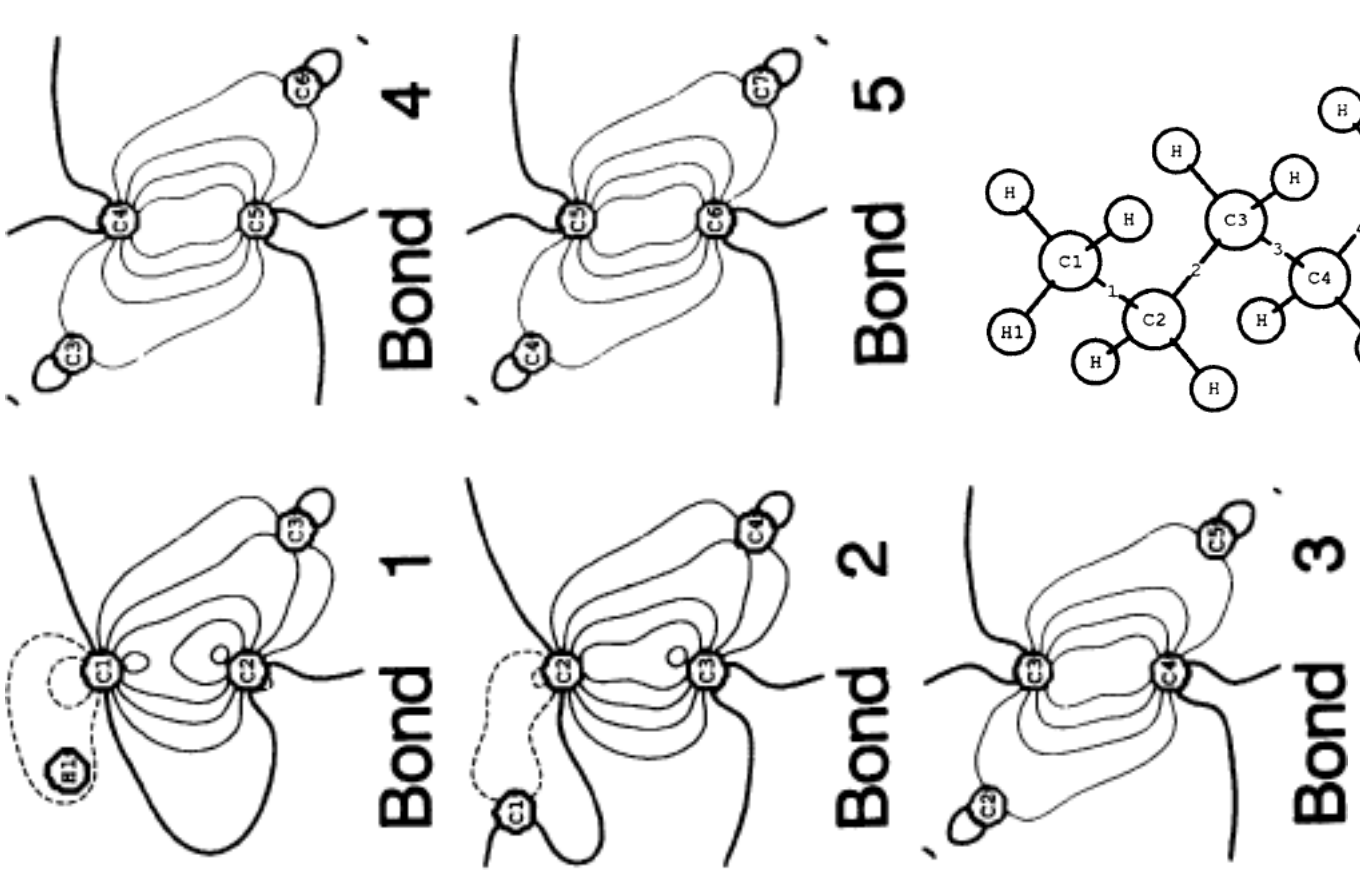
Decane:



DOI:10.1103/PhysRevB.50.17811

Nearsightedness

Hierse & Stechel, *Phys. Rev. B* **50**, 17811 (1994)



DOI:10.1103/PhysRevB.50.17811

Nearsightedness

Hierse & Stechel, *Phys. Rev. B* **50**, 17811 (1994)

TABLE I. Kohn-Sham energy errors for hydrocarbon systems, obtained from self-consistent iteration and/or orbital transfer.

DOI:10.1103/PhysRevB.50.17811

System	ΔE_{KS}/bond	How obtained
C_7H_{16}	0.555 meV	sc iteration
$C_{10}H_{22}$	0.661 meV	sc iteration
$C_{12}H_{26}$	0.703 meV	sc iteration
$C_{12}H_{26}$	0.707 meV	first guess (transfer from $C_{10}H_{22}$)
$C_{12}H_{26}$	0.725 meV	first guess (transfer from C_7H_{16})

Nearsightedness

- Implication for the density matrix:
  - In a local representation it is *sparse*
  - i.e. $F_{ij} \approx 0$ for distant basis functions ϕ_i and ϕ_j

Nearsightedness

- Implication for the density matrix:
  - In a local representation it is *sparse*
  - i.e. $F_{ij} \approx 0$ for distant basis functions ϕ_i and ϕ_j

- In fact the density matrix decays exponentially:
  Brouder *et al.*, *Phys. Rev. Lett.* **98**, 046402 (2007)

Nearsightedness

- Implication for the density matrix:
  - In a local representation it is *sparse*
  - i.e. $F_{ij} \approx 0$ for distant basis functions ϕ_i and ϕ_j

- In fact the density matrix decays exponentially:
  Brouder *et al.*, *Phys. Rev. Lett.* **98**, 046402 (2007)

- Decay rate depends upon
  - Band gap
  - Basis quality

Divide and conquer

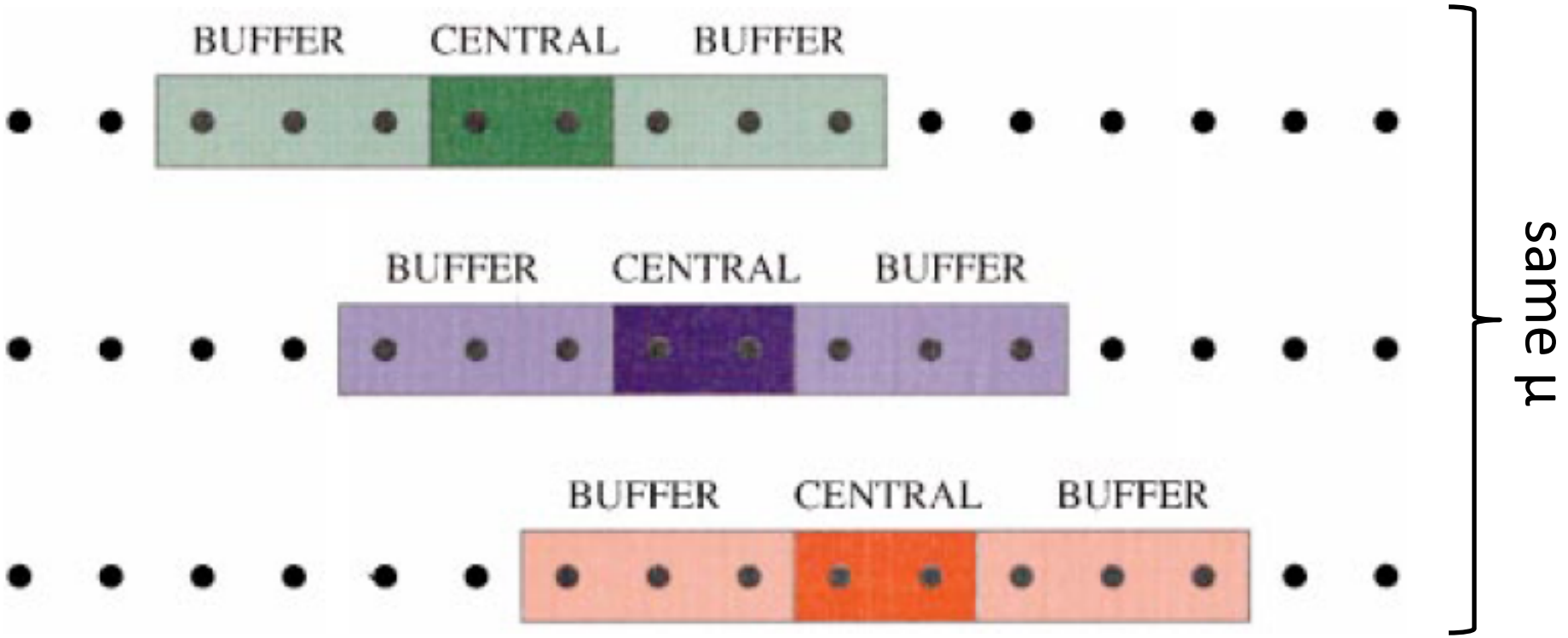
Yang, *Phys. Rev. Lett.* **66**, 1438 (1991)

Yang & Lee, *J. Chem. Phys.* **103**, 5674 (1995)

Divide and conquer

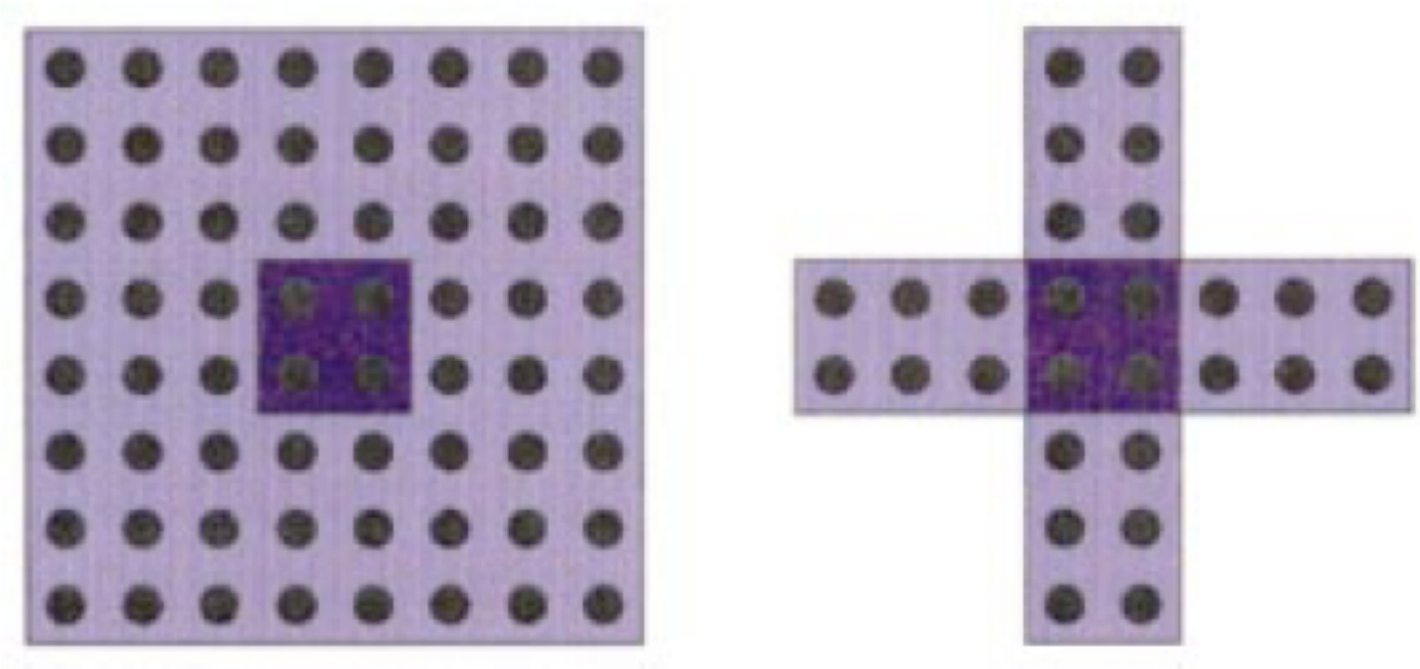
- Consider subvolumes of the whole system
- Calculate contributions to the density (matrix)

DOI:10.1103/RevModPhys.71.1085



Divide and conquer

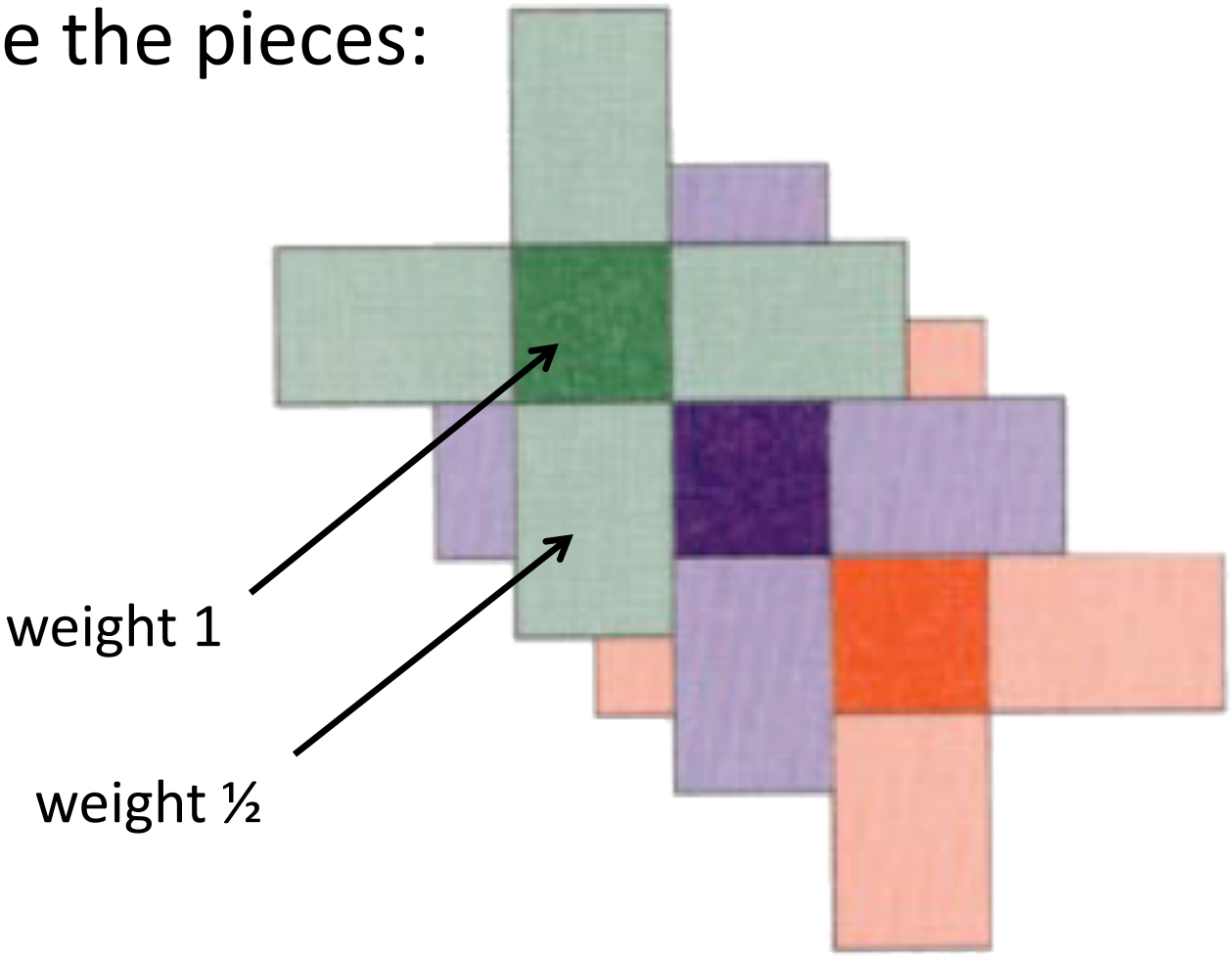
- Trim the corners:



DOI:10.1103/RevModPhys.71.1085

Divide and conquer

- Combine the pieces:



DOI:10.1103/RevModPhys.71.1085

Fermi operator expansion

Goedecker & Colombo, *Phys. Rev. Lett.* **73**, 122 (1994)

Goedecker & Teter, *Phys. Rev. B* **51**, 9455 (1995)

Compatibility

- Need to find F that commutes with H
- Any matrix M always commutes with:
  - The identity I
  - Itself i.e. M
  - Any power of itself e.g. M^2, M^3 etc.

Compatibility

- Need to find F that commutes with H
- Any matrix M always commutes with:
  - The identity I
  - Itself i.e. M
  - Any power of itself e.g. M^2, M^3 etc.
- Expand F as a polynomial in H i.e.

$$F \approx c_0 I + c_1 H + c_2 H^2 + \ldots + c_n H^n$$

  - Coefficients are those from a power series expansion of the Fermi-Dirac distribution

Chebyshev polynomials

$$
\begin{aligned}
T_n(x) &= \cos(n \arccos x) \\
T_0(x) &= 1 \\
T_1(x) &= x \\
T_{j+1}(x) &= 2xT_j(x) - T_{j-1}(x)
\end{aligned}
$$

- Defined on [-1,1]
- Bounded between ±1

Chebyshev expansion

- Scale and shift the Hamiltonian so eigenvalues lie on [-1,1]:

$$F \approx \frac{1}{2}c_0 I + \sum_{j=1}^{n} c_j T_j(H)$$

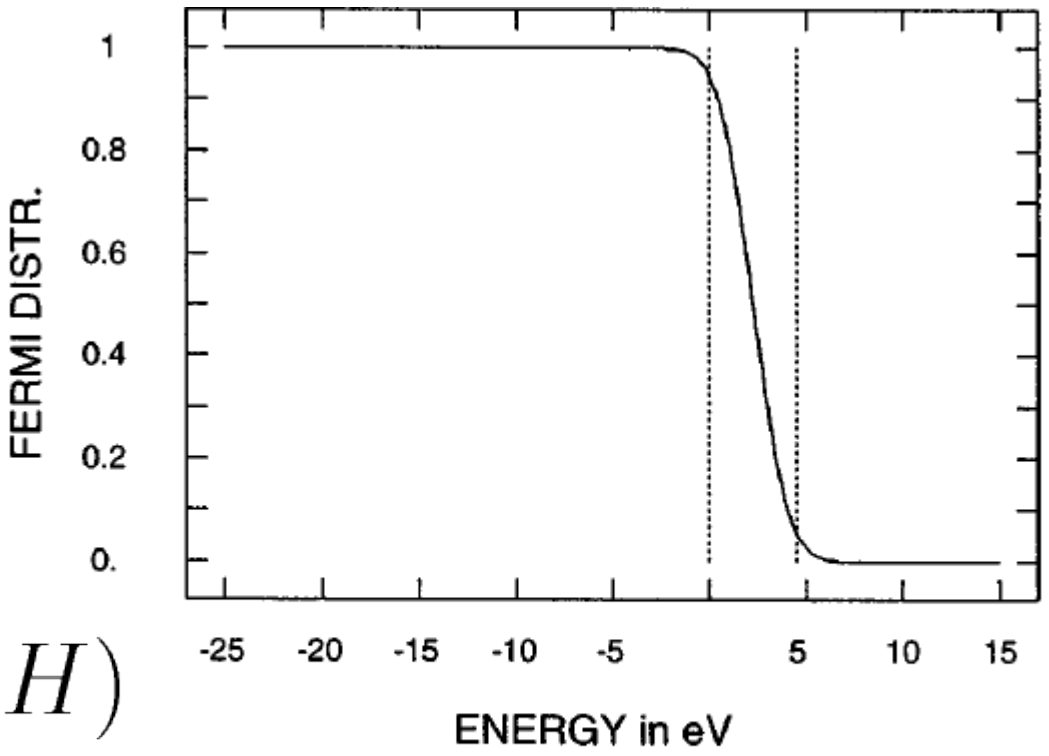


FIG. 7. The Fermi distribution as obtained by a Chebyshev fit of degree 40 in the case of a diamond structure. The band gap is in between the two vertical lines.

DOI:10.1103/RevModPhys.71.1085

Fermi operator expansion

- Region over which expansion changes from 0 to 1 is the energy resolution $\Delta\varepsilon$ (gap)
- Smaller energy resolution requires higher order expansion
- Use finite temperature distribution to avoid Gibbs oscillation
- In practice use error functions instead (decay faster to 0 and 1 away from gap)
- Rational expansion also possible

Density matrix minimization

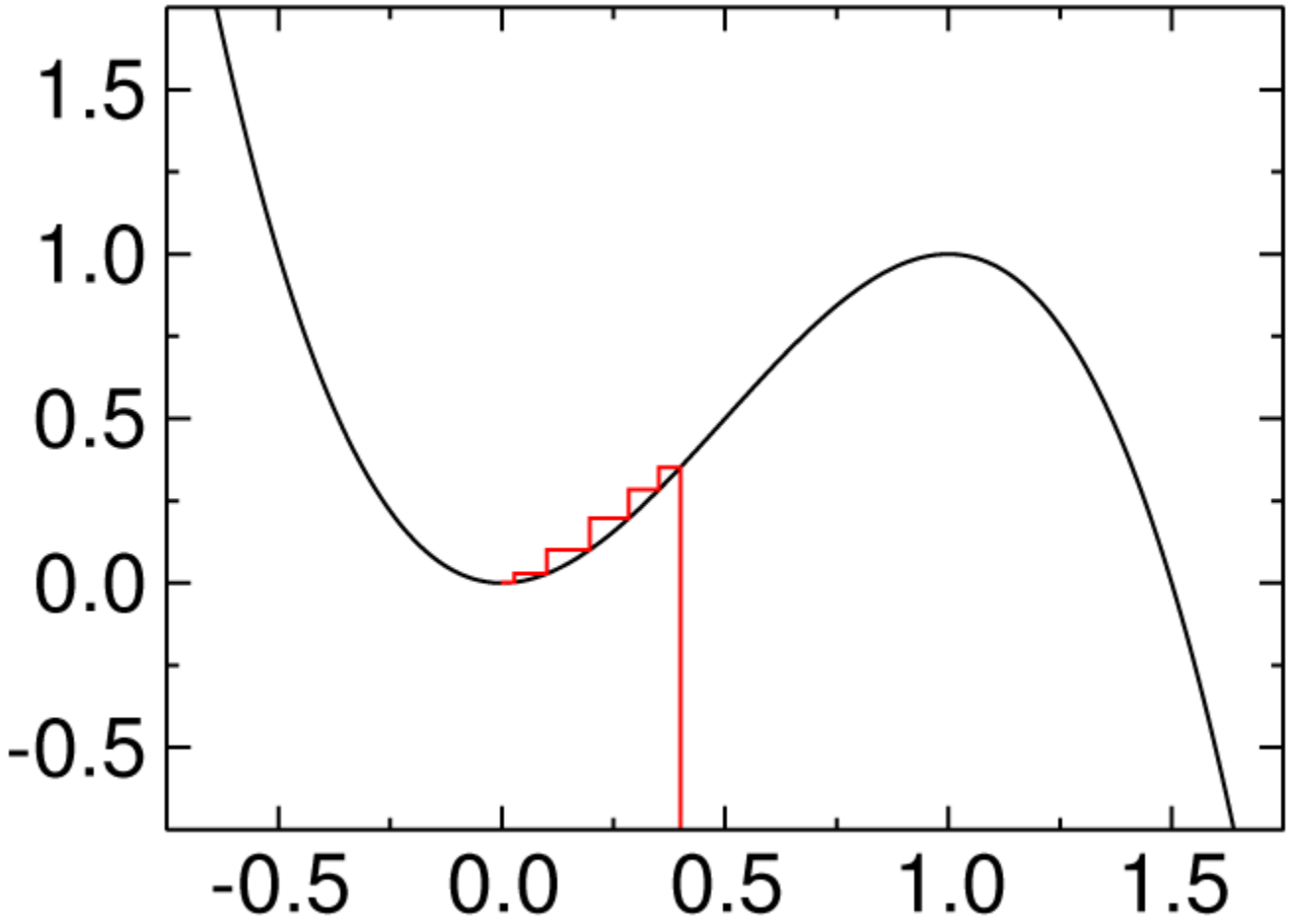
McWeeny, *Rev. Mod. Phys.* **32**, 335 (1960)

Li, Nunes & Vanderbilt, *Phys. Rev. B* **47**, 10891 (1993)

Daw, *Phys. Rev. B* **47**, 10895 (1993)

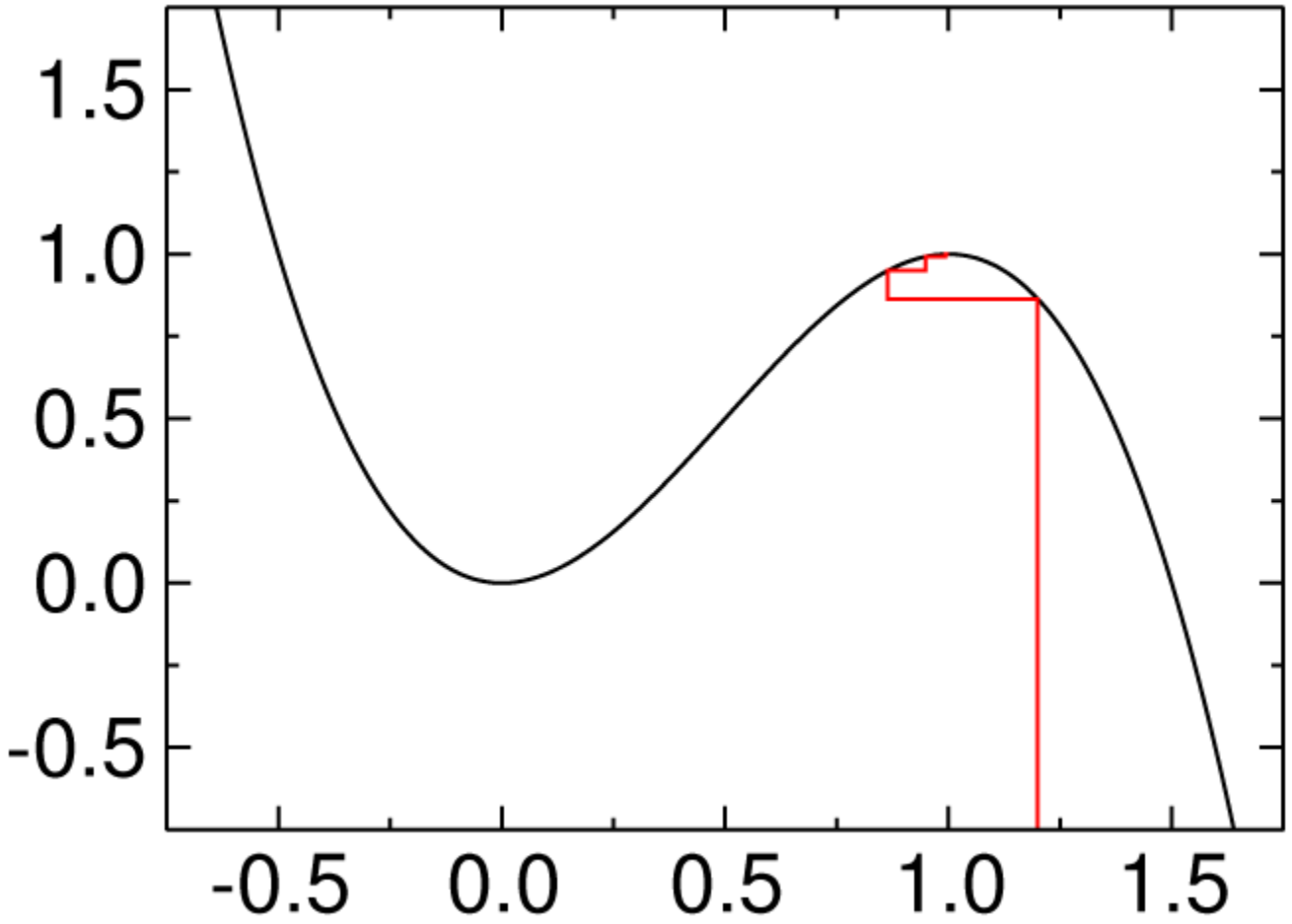
Purifying transformation

$$y = 3x^2 - 2x^3 \quad \Rightarrow \quad x_{k+1} = y(x_k)$$



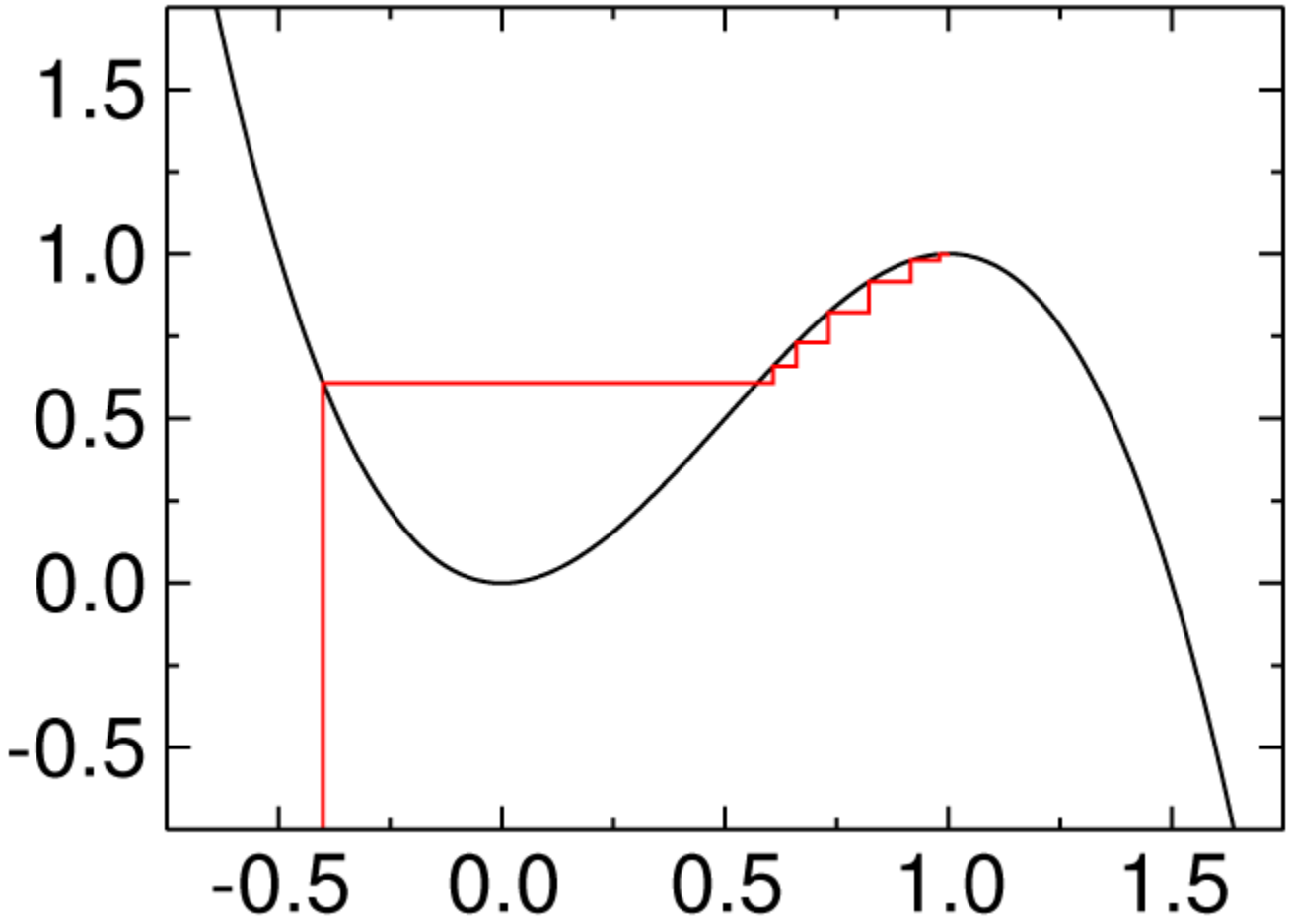
Purifying transformation

$$y = 3x^2 - 2x^3 \quad \Rightarrow \quad x_{k+1} = y(x_k)$$



Purifying transformation

$$y = 3x^2 - 2x^3 \quad \Rightarrow \quad x_{k+1} = y(x_k)$$



Purifying transformation

- Apply it to the density matrix:

$$F_{k+1} = 3F_k^2 - 2F_k^3$$

- Iteration converges to 0 or 1 as long as:

$$f_n \in \left(\frac{1-\sqrt{5}}{2}, \frac{1+\sqrt{5}}{2}\right)$$

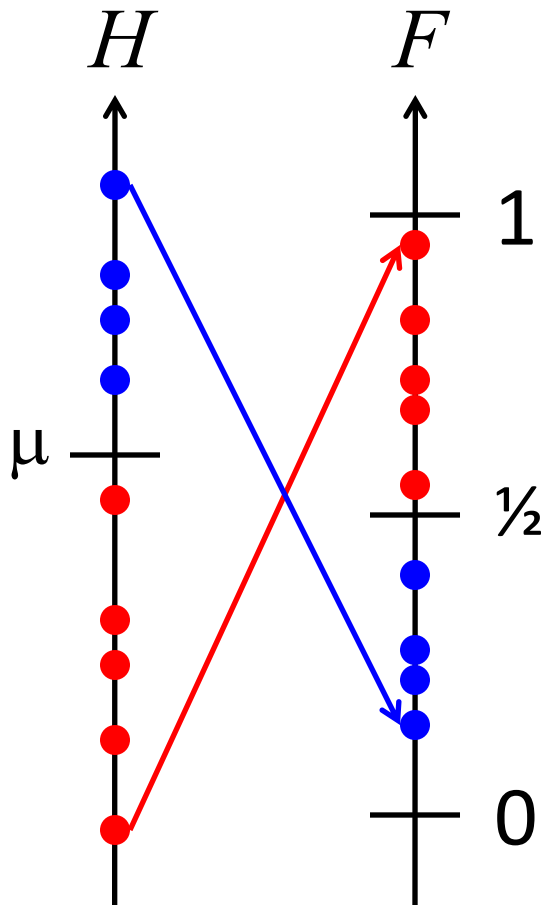
- Converges without “flipping” if:

$$f_n \in \left[\frac{1-\sqrt{3}}{2}, \frac{1+\sqrt{3}}{2}\right]$$

Canonical purification

Palser & Manolopoulos, *Phys. Rev. B* **58**, 12704 (1998)

- Start with Hamiltonian
- Shift, invert and scale so eigenvalues lie in [0,1]
- Apply purification transformation until convergence achieved



Li-Nunes-Vanderbilt

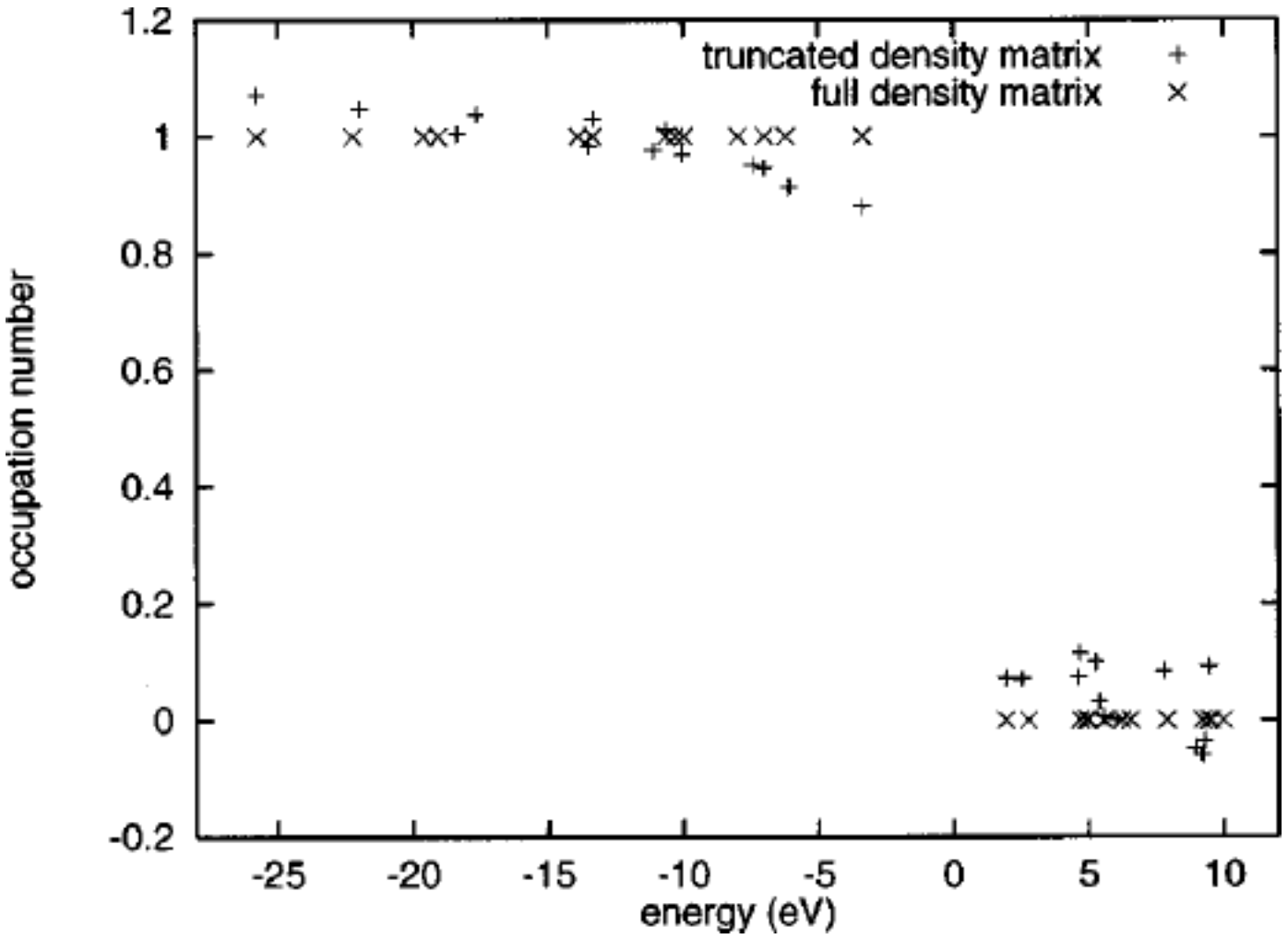
- Define a purified density matrix P

$$P = 3F^2 - 2F^3$$

- Minimize $E = \mathrm{tr}(PH)$ with respect to F

- Truncate F to obtain linear scaling

Li-Nunes-Vanderbilt



Goedecker, *Rev. Mod. Phys.* **71**, 1085 (1999)

DOI:10.1103/RevModPhys.71.1085

Orbital minimization

Mauri *et al.*, *Phys. Rev. B* **47**, 9973 (1993)

Ordejón *et al.*, *Phys. Rev. B* **48**, 14646 (1993)

Mauri & Galli, *Phys. Rev. B* **50**, 4316 (1994)

Ordejón *et al.*, *Phys. Rev. B* **51**, 1456 (1995)

Kim *et al.*, *Phys. Rev. B* **52**, 1640 (1995)

Orbital minimization

- Works with Wannier functions rather than density matrix
- Imposes the orthogonality constraint by expanding the inverse overlap matrix about the identity:

$$\begin{aligned} S^{-1} &= [I-(I-S)]^{-1} \\ &= I+(I-S)+(I-S)^2+\ldots \\ &\approx 2I-S \end{aligned}$$

Orbital minimization

- Leads to a generalized functional:

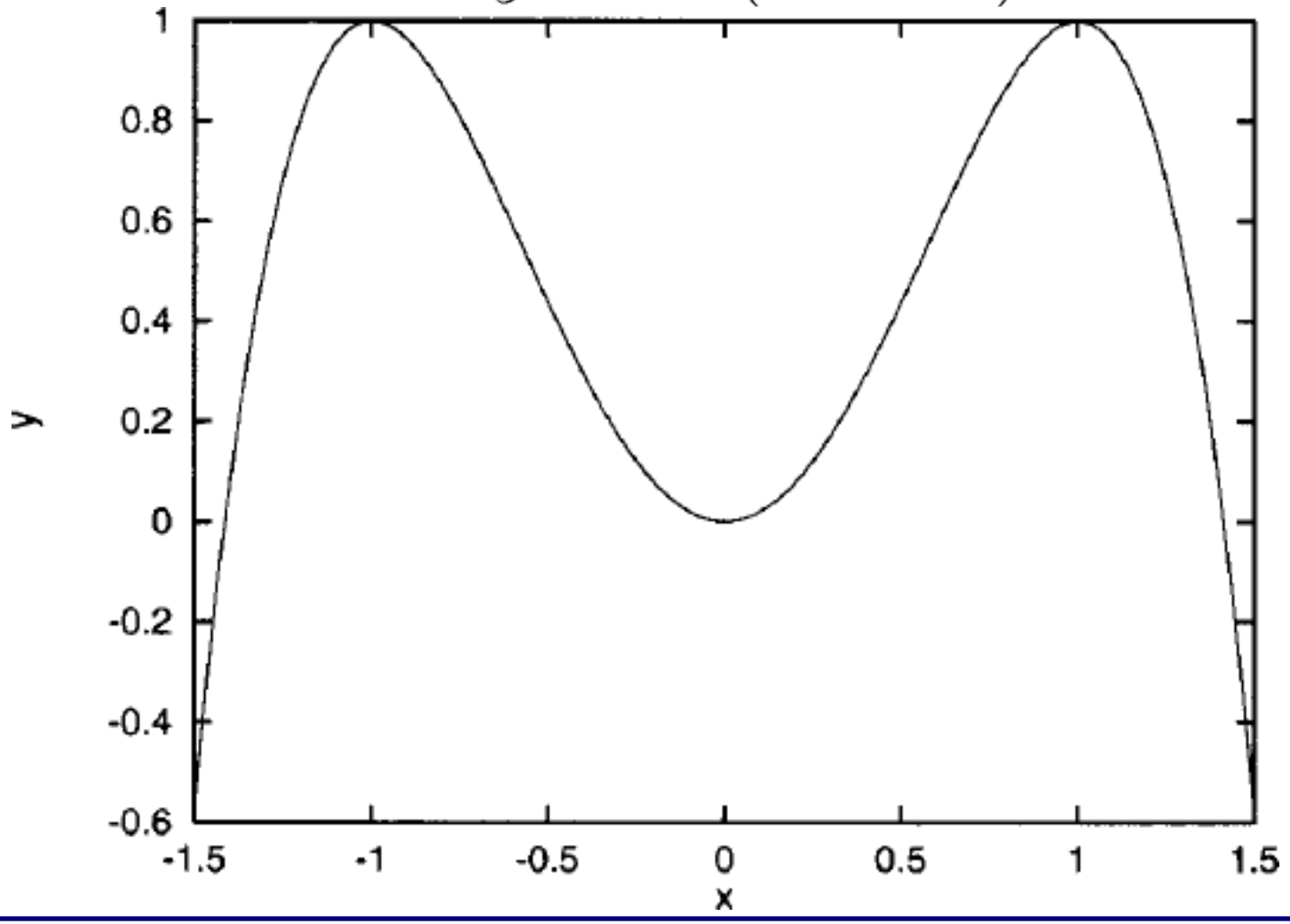
$$\Omega = 2\sum_{n}\sum_{ij} c_{in}^{*} H'_{ij} c_{jn} - \sum_{nm}\sum_{ij} c_{in}^{*} H'_{ij} c_{jm} \sum_{k} c_{km}^{*} c_{kn}$$

  – where $H' = H - \mu I$

- Quartic in the coefficients c
- Solve for localized orbitals to obtain linear scaling

Orbital minimization

$$y = x^2(2 - x^2)$$



Goedecker, *Rev. Mod. Phys.* **71**, 1085 (1999)
DOI:10.1103/RevModPhys.71.1085

Orbital minimization

- With localization constraints:
  - Large number of iterations required
  - Atom-centred Wannier functions can break symmetry
  - Local minimum so runaway solutions possible
  - Problems conserving electron number

Yang, *Phys. Rev. B* **56**, 9294 (1997)

More information

Stefan Goedecker
"Linear scaling electronic structure methods"
Rev. Mod. Phys. **71**, 1085 (1999)

David Bowler and Tsuyoshi Miyazaki
"*O*(*N*) methods in electronic structure calculations"
Rep. Prog. Phys. **75**, 036503 (2012)