Electronic Structure Solvers in SIESTA

Alberto García
ICMAB-CSIC

Diagonalization

$$\psi_i(\boldsymbol{r}) = \sum_\mu \phi_\mu(\boldsymbol{r}) c_{\mu i},$$

$$\sum_{\nu\beta} (H^{\alpha\beta}_{\mu\nu} - E_i S_{\mu\nu}\delta^{\alpha\beta}) c^{\beta}_{\nu i} = 0$$

Generalized eigenvalue problem

$$H^{\alpha\beta}_{\mu\nu} = \langle \phi_\mu | \hat{T} + \hat{V}^{KB} + V^{NA}(\boldsymbol{r}) + \delta V^{H}(\boldsymbol{r}) + V^{\alpha\beta}_{XC}(\boldsymbol{r}) | \phi_\nu \rangle$$

$$S_{\mu\nu} = \langle \phi_\mu | \phi_\nu \rangle$$

$$\rho(\boldsymbol{r}) = \sum_{\mu\nu} \rho_{\mu\nu} \phi^*_\nu(\boldsymbol{r}) \phi_\mu(\boldsymbol{r})$$

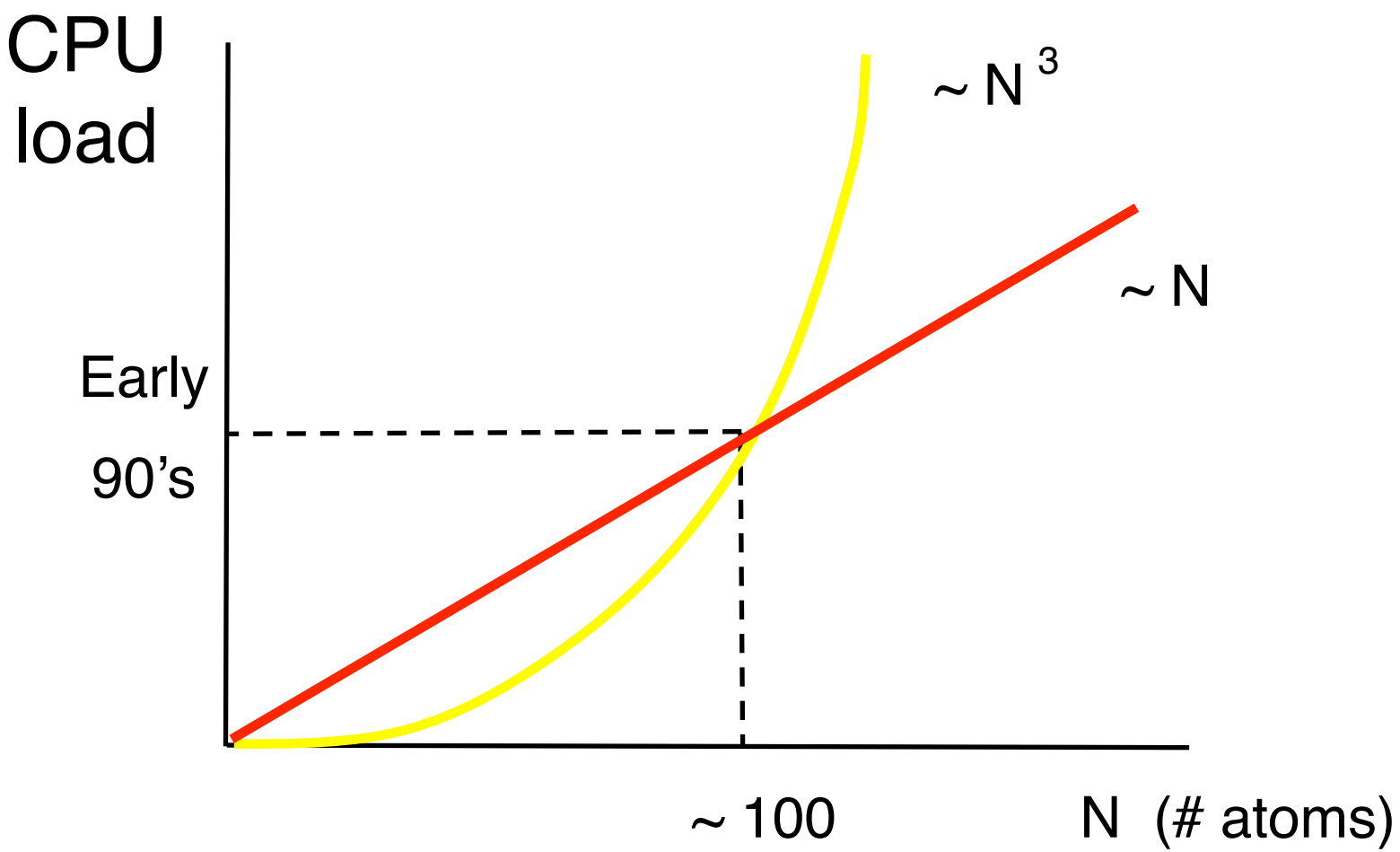
$$E^{BS} = \sum_i n_i \langle \psi_i | \hat{H} | \psi_i \rangle = \sum_{\mu\nu} H_{\mu\nu} \rho_{\nu\mu} = \mathrm{Tr}(H\rho)$$

$$\rho_{\mu\nu} = \sum_i c_{\mu i} n_i c_{i\nu}$$

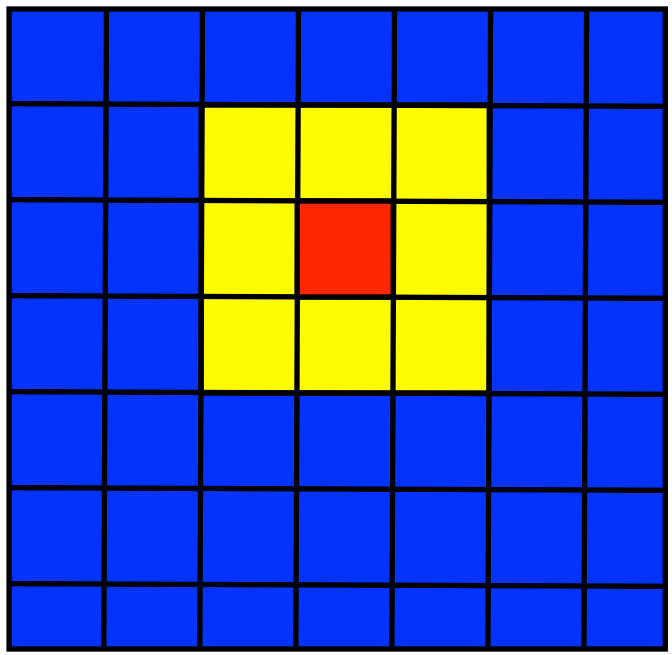
Density matrix

Linear-scaling method

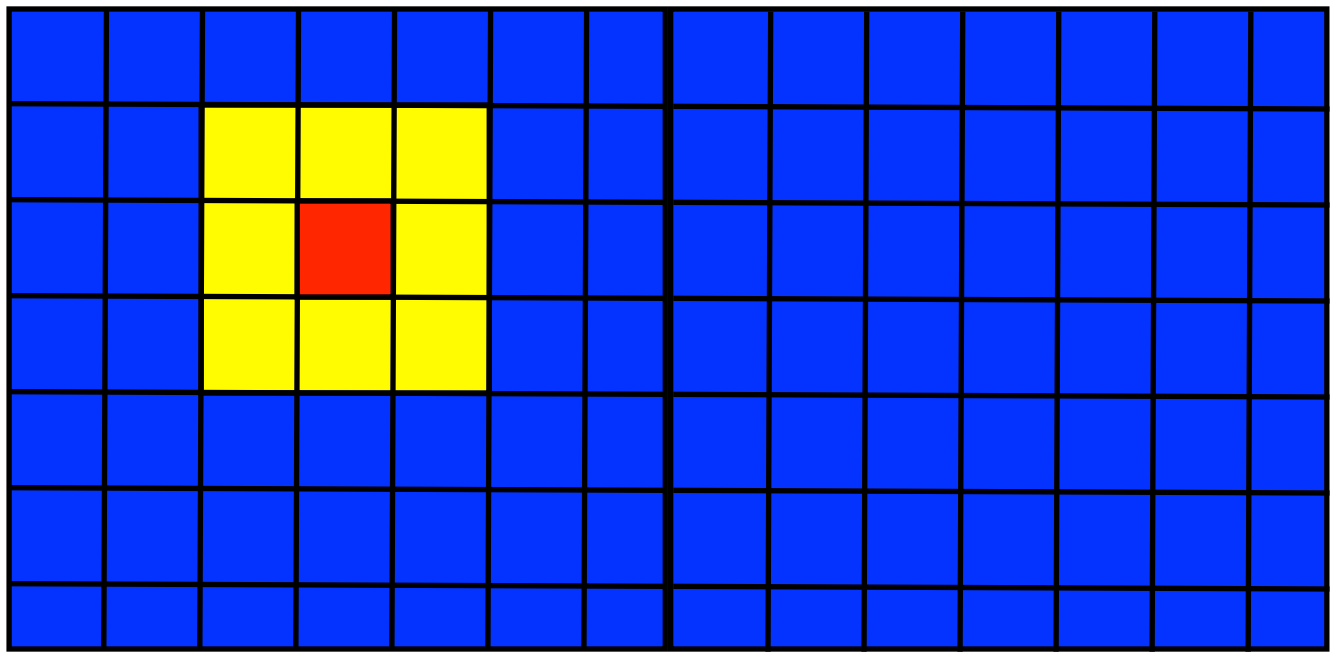
Linear scaling = Order(N)



Key to O(N): locality



Key to O(N): locality



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

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

Generalized DFT functional for O(N) scaling

Electronic states $\{|\psi_i\rangle\}\ (i = 1, \ldots, N)$

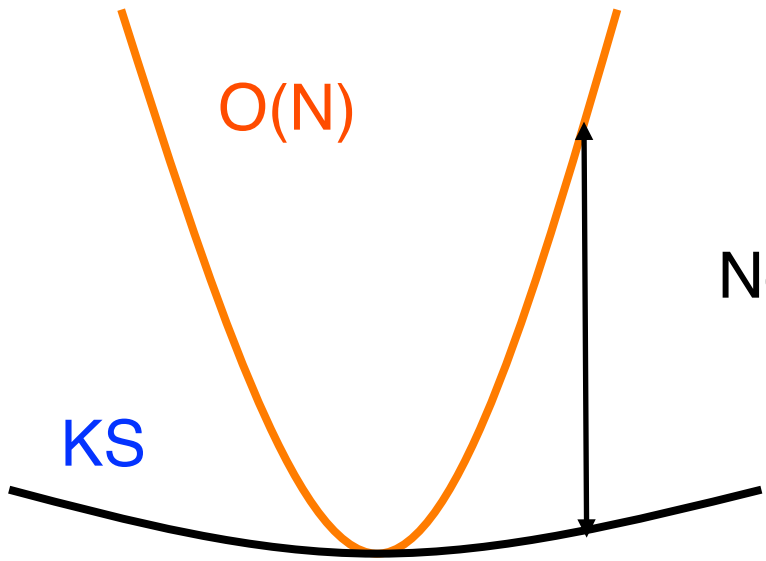
Standard Energy $E = 2\,\mathrm{Tr}(\hat{H}) = 2\sum_{i,j=1}^{N} S_{ij}^{-1} H_{ji},$

$$S_{ij} = \langle\psi_i|\psi_j\rangle,\ H_{ij} = \langle\psi_i|\hat{H}|\psi_j\rangle$$

New energy functional: orthogonality not imposed

$$\tilde{E} = 2\left(\sum_{i=1}^{N} H_{ii} - \sum_{i,j=1}^{N} H_{ji}(S_{ij} - \delta_{ij})\right)$$

O(N) vs. Kohn-Sham functional

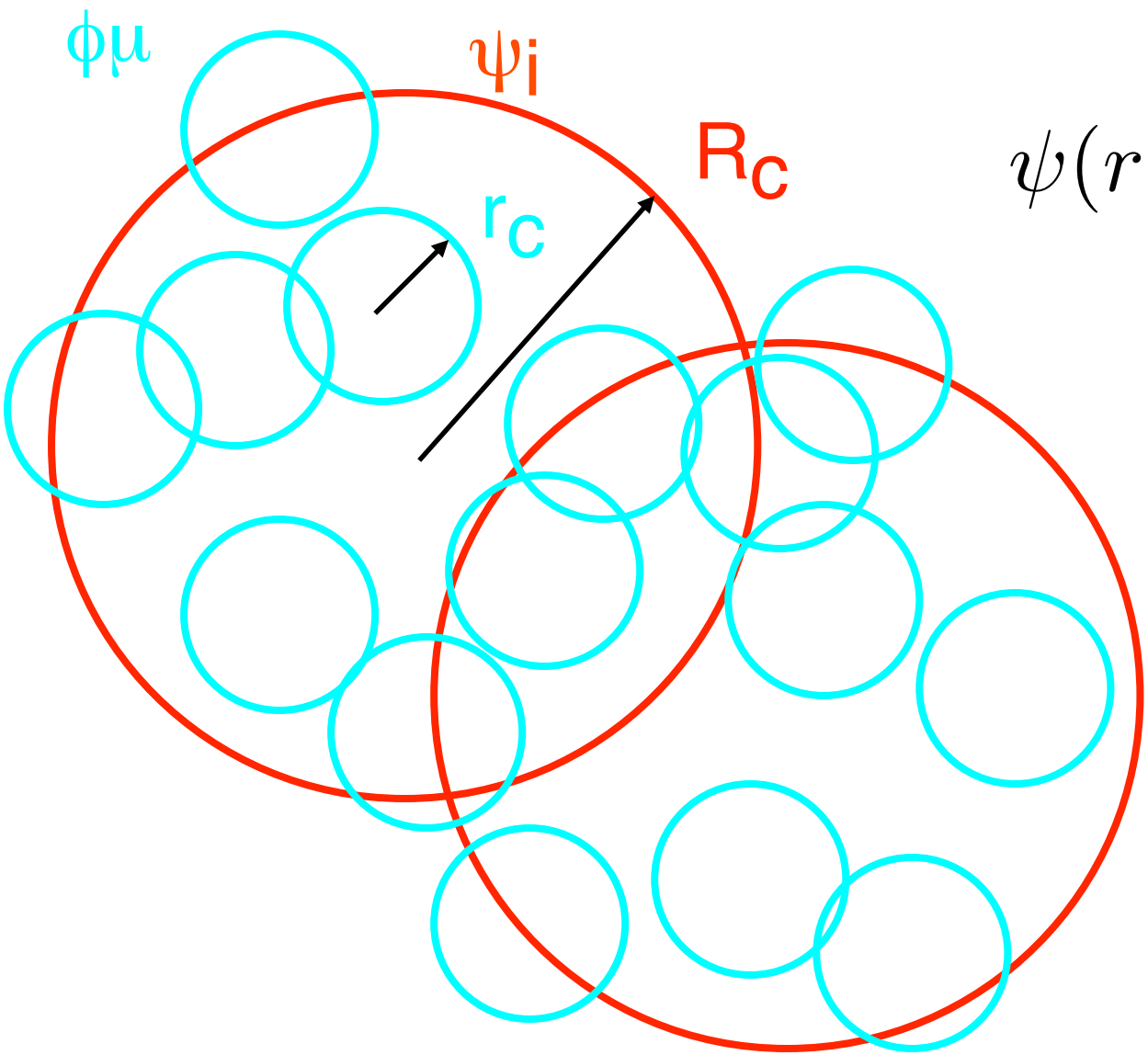


Non-orthogonality penalty

$$S_{ij} = \delta_{ij} \Rightarrow E_{O(N)} = E_{KS}$$

Mauri et al. Ordejon et al (1992)

Orbital localization



$$\psi(r) = \sum_{\mu} c_{i\mu}\phi_{\mu}(r)$$

Non-O(N) orbital minimization method (OMM)

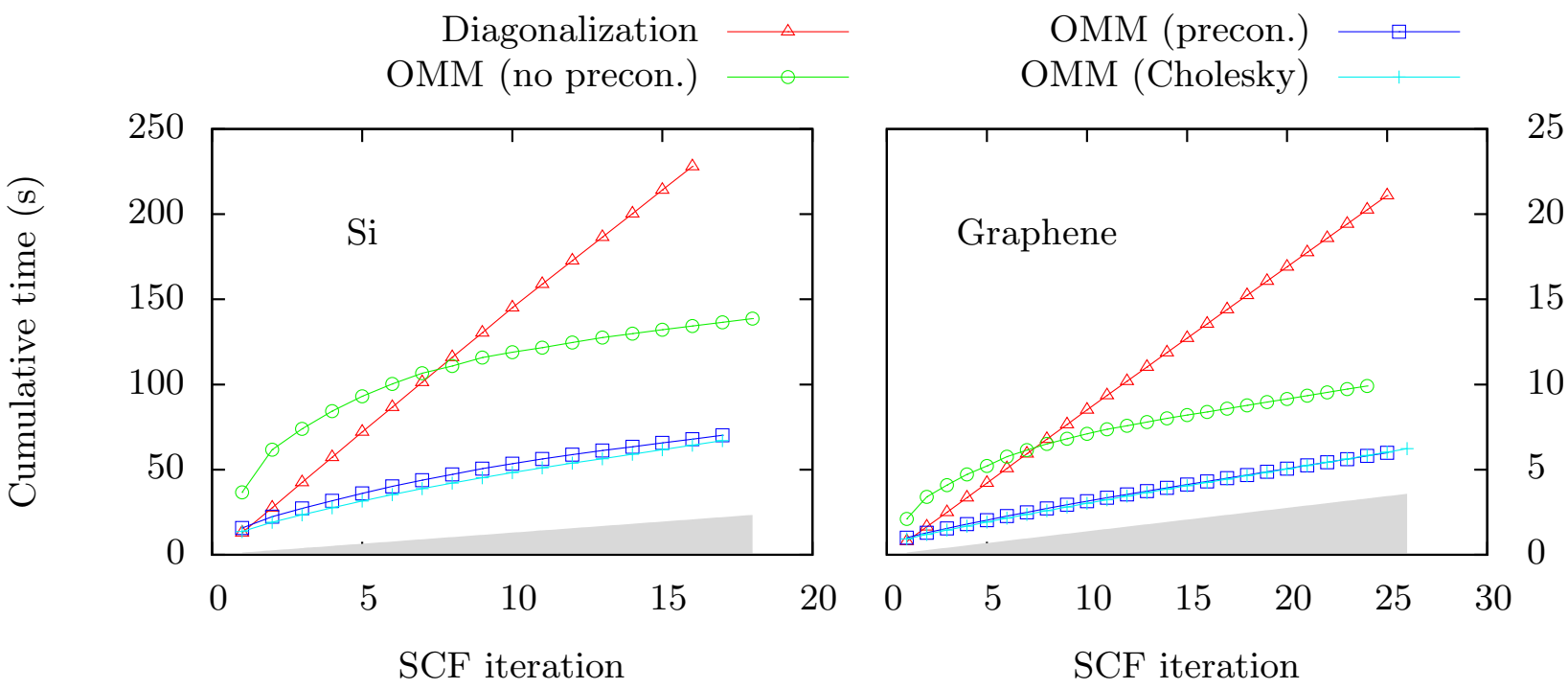
Retains the interesting feature of avoiding orthogonalization

Does not impose the localization condition

It is not O(N), but it is universally applicable

Implemented in SIESTA by Fabiano Corsetti

The values of the solution coefficients in one SCF iteration can be used as starting guesses for the next iteration



The PEXSI solver

Diagonalization

$H|\psi_i >= \epsilon_i S|\psi_i >$ ScaLaPack $O(N^3)$

$$|\psi_i >= \sum_\mu c_i^\mu |\phi_\mu >$$

$$\hat{\rho} = \sum_i f_i |\psi_i >< \psi_i| = \sum_{i\mu\nu} f_i c_i^\mu c_i^\nu |\phi_\mu >< \phi_\nu|$$

$\rho_{\mu\nu} = \sum_i f_i c_i^\mu c_i^\nu$ Density matrix

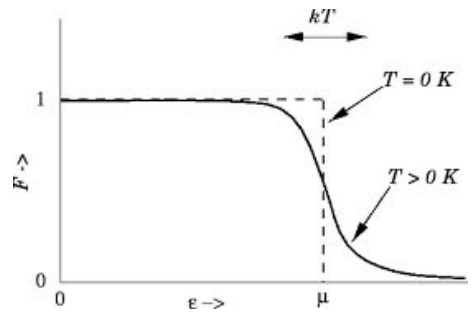
$$N = \mathrm{Tr}[\hat{\rho}S] \qquad E_{BS} = \mathrm{Tr}[\hat{\rho}H]$$

Formal solution to the electronic-structure problem

$$\hat{\rho} = f_\beta(\hat{H} - \mu)$$

$$f_\beta(\epsilon_i - \mu) = \frac{2}{1 + \mathrm{e}^{\beta(\epsilon_i - \mu)}}$$

Fermi-Dirac function



We just need a computationally convenient representation of the Fermi-Dirac function

Polynomial representation of the Fermi-Dirac function: Fermi Operator Expansion (FOE)

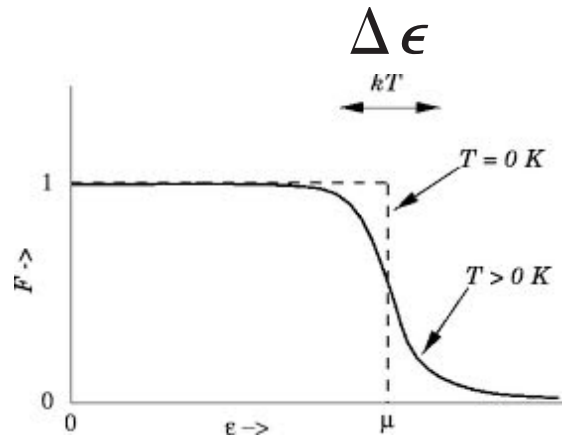
(Chebyshev polynomials used for stability)

$$p(H) = \frac{c_0}{2} I + \sum_{j=1}^{n_{pl}} c_j T_j(H)$$

Calculation of the DM involves only (sparse) matrix multiplications

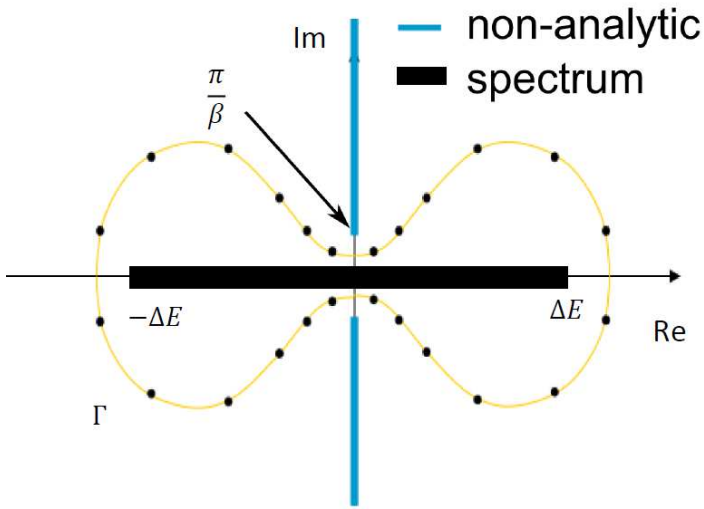
$$n_{pl} \propto \frac{\epsilon_{max} - \epsilon_{min}}{\Delta\epsilon}.$$

Many terms are typically needed



Pole (rational) expansion of the Fermi function

$$f_\beta(\epsilon_i - \mu) = \frac{2}{1 + \mathrm{e}^{\beta(\epsilon_i - \mu)}}$$



$$f_\beta(\epsilon_i - \mu) \approx Im \sum_{l=1}^{P} \frac{\omega_l}{\epsilon_i - (z_l + \mu)}$$

$$\hat{\rho} = \quad Im \left(\sum_{l=1}^{P} \frac{\omega_l}{H - (z_l + \mu)S} \right)$$

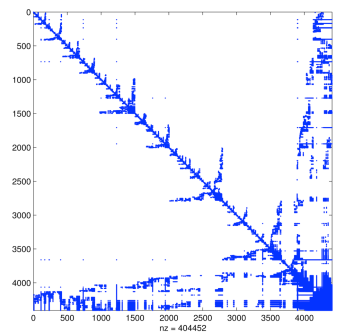
Fewer terms are needed (typically 40 poles are enough)

Pole Expansion plus Selected Inversion

(Lin Lin, Chao Yang, LBNL)

$$\hat{\rho} = Im\left(\sum_{l=1}^{P} \frac{\omega_l}{H-(z_l+\mu)S}\right)$$

One inversion per pole



Only a limited number of elements is needed in the density matrix!!

Use of a very efficient Selected Inversion algorithm

$$A = \begin{pmatrix} a & b^T \\ b & \hat{A} \end{pmatrix}$$

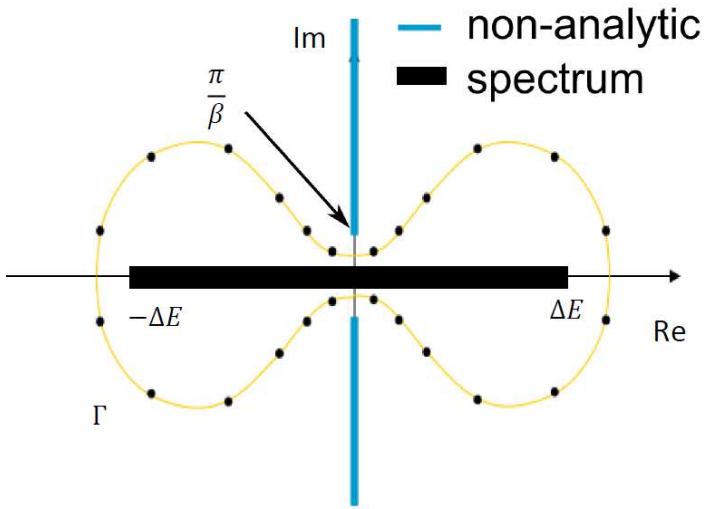
$$(A =)LDL^T = \begin{pmatrix} 1 & \\ l & \hat{L} \end{pmatrix} \begin{pmatrix} \alpha & \\ & \hat{A}-bb^T/\alpha \end{pmatrix} \begin{pmatrix} 1 & l^T \\ & \hat{L} \end{pmatrix}$$

There are no DM locality approximations, as in O(N) methods

For sufficiently big problems

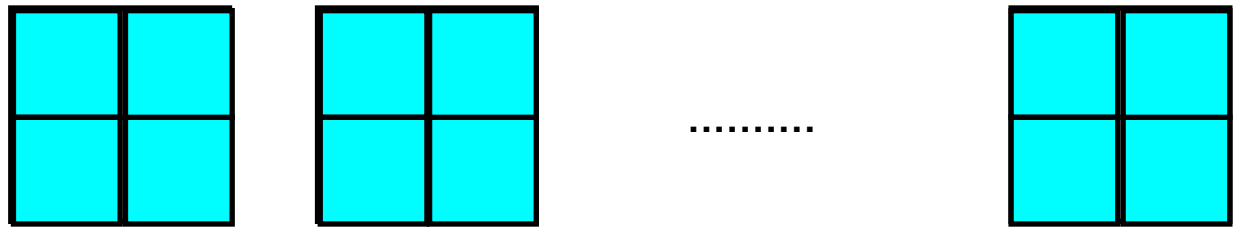
$$\begin{aligned} \text{(quasi-)1D:} &\quad \mathcal{O}(N) \\ \text{(quasi-)2D:} &\quad \mathcal{O}(N^{3/2}) \\ \text{3D:} &\quad \mathcal{O}(N^2) \end{aligned}$$

(Due to sparsity of the target density matrix)



The method is applicable to metals using low effective temperatures

Trivially parallel over poles, with perfect load balancing



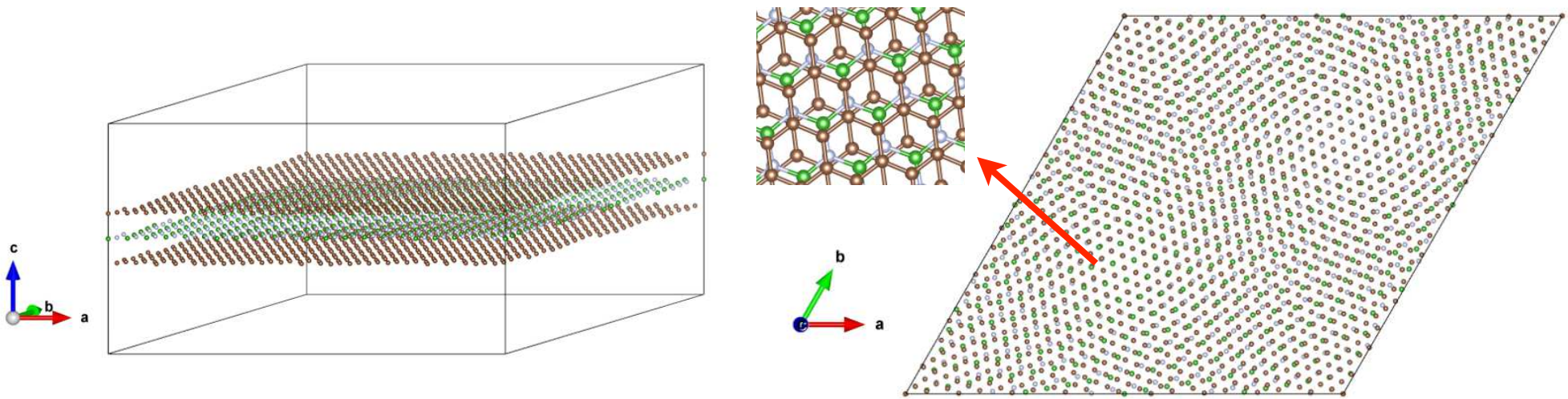
4 processors per pole x 40 poles :: 160 processors

The number of processors per pole is a parameter of the calculation.
Lower limit: as needed to fit the problem in memory.

Main test systems

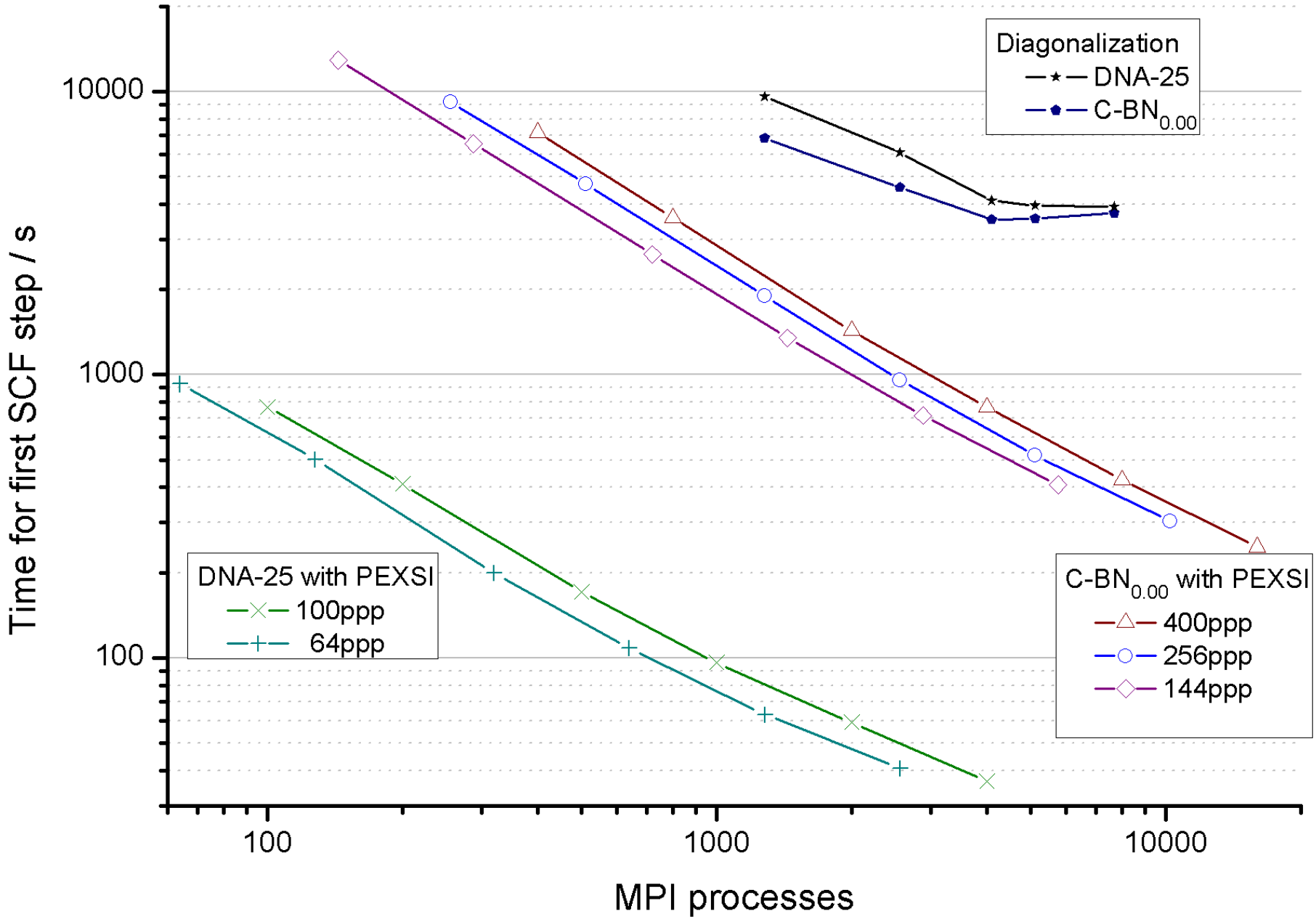
A repeated DNA strand (up to 18000 atoms) (1D)

quasi-2D C-BN structures, up to 13000 atoms

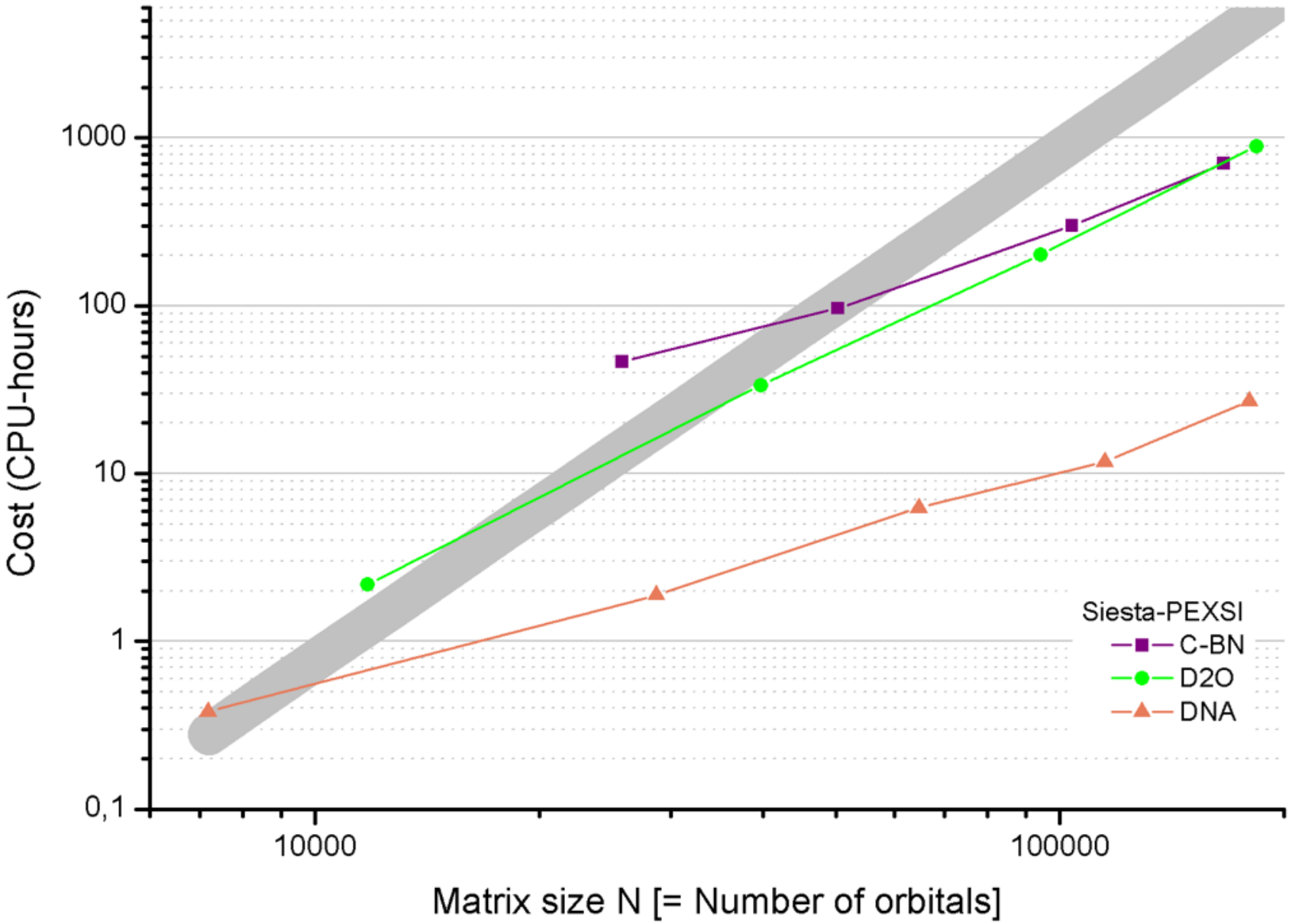


A box of water molecules (3D)

Strong scaling



Weak scaling



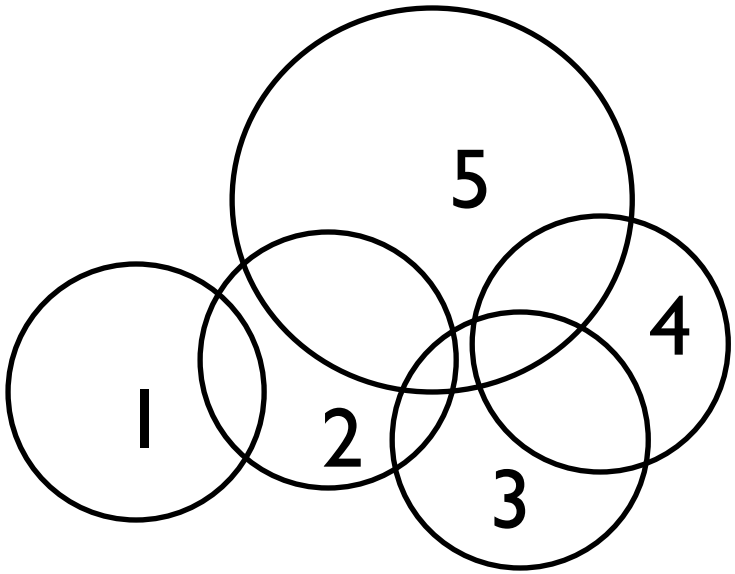
The PEXSI solver will be included soon in the Siesta distribution

SIESTA-PEXSI: Massively parallel method for efficient and accurate ab initio materials simulation without matrix diagonalization

Lin Lin, Alberto García, Georg Huhs, Chao Yang

arXiv:1405.0194 [physics.comp-ph] (Accepted in J. Phys. Cond. Matter)

Sparsity



1 with 1 and 2

2 with 1,2,3, and 5

3 with 2,3,4, and 5

4 with 3,4 and 5

5 with 2,3,4, and 5

$S_{\mu\nu}$ and $H_{\mu\nu}$ are sparse

$\rho\mu\nu$ is not strictly sparse

but only a sparse subset is needed

Non-overlap interactions