Pole Expansion and Selected Inversion Method for Accelerating Electronic Structure Calculation

Mathias Jacquelin, Lawrence Berkeley National Laboratory

PI: Roberto Car, Weinan E (Princeton), Michael Klein, Xifan Wu (Temple), Esmond Ng, Chao Yang, Lin Lin (LBNL)

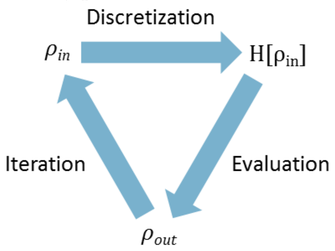
Objective

Develop fast *ab initio* molecular dynamics (AIMD) strategies for large scale disordered systems.

Kohn-Sham density functional theory (KSDFT):

$$\left(-\frac{1}{2}\Delta + V_{ext}\left[\{R_j\}_{j=1}^{M}\right] + \int dx' \frac{\rho(x')}{|x-x'|} + V_{xc}[\rho]\right)\psi_i(x) = \varepsilon_i \psi_i(x)$$

$$\rho(x) = 2\sum_{i=1}^{N/2} |\psi_i(x)|^2, \quad \int dx\, \psi_i^*(x)\psi_j(x) = \delta_{ij}$$



Energy and force

$$E_{tot}\left[\{R_j\}_{j=1}^{M}\right] = \sum_{i=1}^{N/2}\varepsilon_i - \frac{1}{2}\iint dx\, dy\, \frac{\rho(x)\rho(y)}{|x-y|} + E_{xc}[\rho(x)] - \int dx\, V_{xc}[\rho(x)]\rho(x)$$

$$F_i = \frac{-\partial E_{tot}}{\partial R_i}$$

PEXSI: Methodology

PEXSI: Pole EXpansion and Selected Inversion

Objective: Accurate evaluation of the electron density, energy, and atomic force, with at most $O(N^2)$ scaling.

Quasi-1D	$O(N)$
Quasi-2D	$O(N^{1.5})$
3D bulk	$O(N^2)$

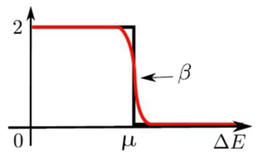
KSDFT in the form of the Fermi operator

$$\rho(x) = 2\sum_{i=1}^{N/2}|\psi_i(x)|^2$$

$$= \begin{pmatrix} \psi_1(x) & \cdots & \psi_{N_z}(x) \end{pmatrix} \begin{pmatrix} \chi(\varepsilon_1-\mu) & & \\ & \ddots & \\ & & \chi(\varepsilon_{N_z}-\mu) \end{pmatrix} \begin{pmatrix} \psi_1(x) \\ \vdots \\ \psi_{N_z}(x) \end{pmatrix}$$

$$= \left\{\chi(H[p]-\mu I)\right\}_{x,x}$$

- μ: Chemical potential such that $\#\{\sigma(H) \le \mu\} = \frac{N}{2}$
- χ: Heaviside function satisfying $\chi(x) = \begin{cases} 2, & x \le 0 \\ 0, & x > 0 \end{cases}$



- Finite temperature, Fermi-Dirac
- Zero temperature, Heaviside

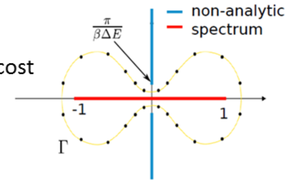
$$\rho = \mathrm{diag}\frac{2}{1+e^{\beta(H[p]-\mu I)}}$$

Pole expansion based on complex analysis

$$\rho \approx \mathrm{diag}\sum_{i=1}^{Q}\frac{\omega_i}{H - z_i I}$$

Optimal expansion cost

$$Q \sim \log \beta\Delta E$$



Selected inversion: compute selected elements of a Green's function without the full inversion

- LDL^T factorization

$$A = \begin{pmatrix} A_{11} & A_{21}^T \\ A_{21} & \hat{A}_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ L_{21} & I \end{pmatrix}\begin{pmatrix} A_{11} & 0 \\ 0 & S_{22} \end{pmatrix}\begin{pmatrix} 1 & L_{21}^T \\ 0 & I \end{pmatrix}$$

$$L_{21} = A_{21}A_{11}^{-1}, \qquad S_{22} = \hat{A}_{22} - A_{21}L_{21}^T$$

- Inversion

$$A^{-1} = \begin{pmatrix} A_{11}^{-1} + L_{21}^T S_{22}^{-1} L_{21} & -L_{21}^T S_{22}^{-1} \\ -S_{22}^{-1}L_{21} & S_{22}^{-1} \end{pmatrix}$$

A^{-1} restricted to the non-zero pattern of L is "self-contained"

Force computation: Including both the Hellmann- Feynman force and the Pulay force

$$F_I = -Tr\left[\gamma\frac{\partial H}{\partial R_I}\right] + Tr\left[\gamma^E\frac{\partial S}{\partial R_I}\right]$$

References

L. Lin, M. Chen, C. Yang and L. He, Accelerating atomic orbital-based electronic structure calculation via pole Expansion and selected inversion, J. Phys. Condens. Matter 25, 295501, 2013

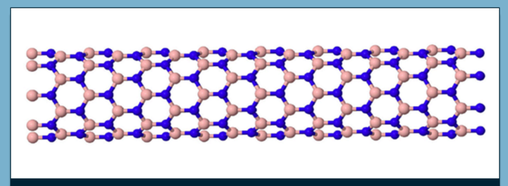
L. Lin, C. Yang, J. Lu, L. Ying and W. E, A Fast Parallel algorithm for selected inversion of structured sparse matrices with application to 2D electronic structure calculations, SIAM J. Sci. Comput. 33, 1329, 2011

L. Lin, C. Yang, J. Meza, J. Lu, L. Ying and W. E, SelInv—An algorithm for selected inversion of a sparse symmetric matrix, ACM Trans. Math. Software 37, 40, 2011

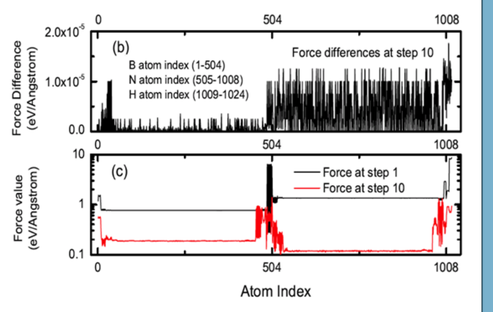
L. Lin, J. Lu, L. Ying and W. E, Pole-based approximation of the Fermi-Dirac function, Chin. Ann. Math. 30B, 729, 2009

L. Lin, J. Lu, L. Ying, R. Car and W. E, Fast algorithm for extracting the diagonal of the inverse matrix with application to the electronic structure analysis of metallic systems, Commun. Math. Sci. 7, 755, 2009

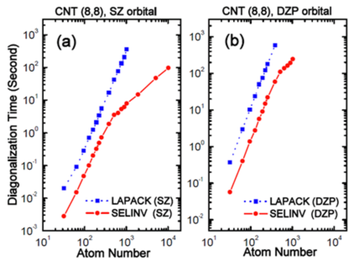
Sequential PEXSI: Results



Boron nitride nanotube (BNNT)



Geometry optimization for BNNT with 1024 atoms (DZP basis, PEXSI solved on a single core)



All on a single core, 80 poles (not parallelized) and 2 iterations for chemical potential.

# Poles	$E_{\mathrm{PEXSI}} - E_{\mathrm{ref}}$ (eV)	MAE Force (eV/Angstrom)
20	5.868351108	0.400431
40	0.007370583	0.001142
60	0.000110382	0.000026
80	0.000000360	0.000002

Efficiency of sequential PEXSI compared to LAPACK

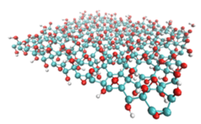
Massively Parallel PEXSI

Distributed memory parallel selected inversion for general matrix (factorization is based on SuperLU_DIST), preliminary version scalable to $64 \sim 256$ procs. More efficient version under progress.

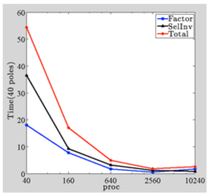
Pole expansion parallelized. With 40 poles used in practice, PEXSI can scale to $256 \times 40 \sim 10,000$ procs.

C++ implementation. Nearly black-box interface, being integrated to SIESTA. CP2K in the future.

Preliminary results

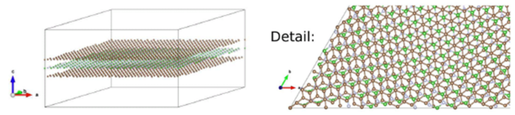


Graphene Oxide with water (320 atoms). Matrix size: 3520. Courtesy of Dr. Limin Liu.



Strong scaling of PPEXSI: Preliminary implementation scales well to 2560 procs even for this small matrix size (with 40 poles)

C-BN-C layered system, weak scaling for more than 10,000 atoms. All examples use $40 \times 256 = 10,240$ procs on hopper.



Number of atoms	Equivalent cells	Matrix dimension	Time per iteration	Scaling	
2532	1 x 1	32916	32	1	
10128	2 x 2	131664	258	8.06	$O(N^{1.2})$ scaling
20256	4 X 2	263328	554	17.3	$O(N)$ scaling

ScaLAPACK performance: 230 sec for 2532 atoms using 768 processors and does not scale beyond.