Robust Determination of the Chemical Potential in the Pole Expansion and Selected Inversion Method in PEXSI

Weile Jia, Lin Lin

University of California, Berkeley

Aug 16th 2017

Introduction

- PEXSI method for DFT
- Robust determination of chemical potential
  - Coarse level: inertia counting
  - Fine level: multi-point evaluation of Fermi Operator
- Parallelization
- Numerical results
- Conclusion

Kohn-Sham DFT diagonalization method

Assume $\Phi = [\varphi_1, \cdots, \varphi_N]$ is Kohn-Sham orbitals, H, S are discretized Kohn-Sham Hamiltonian matrix and overlap matrices

$$HC = SC\Lambda \tag{1}$$

$\Lambda = \mathrm{diag}[\lambda_1, \ldots, \lambda_N]$ is a diagonal matrix containing the Kohn-Sham eigenvalues. The single particle density matrix is

$$\Gamma = Cf_\beta(\Lambda - \mu I)C^T. \tag{2}$$

Here $\beta = 1/k_B T$ is the inverse temperature, and

$$f_\beta(x) = \frac{1}{1 + e^{\beta x}} \tag{3}$$

is the Fermi-Dirac distribution. The chemical potential μ chosen to ensure that

$$N_\beta(\mu) = \mathrm{Tr}[f_\beta(\Lambda - \mu I)] = \mathrm{Tr}[S\Gamma] = N_e, \tag{4}$$

where N_e is the number of electrons.

Fermi Operator Expansion(FOE) method

FOE method expands Γ using a rational approximation as

$$\Gamma \approx \sum_{l=1}^{P} \mathfrak{Im} \left(\omega_l^{\rho} (H - (z_l + \mu) S)^{-1} \right). \tag{5}$$

Here $G_l := (H - (z_l + \mu)S)^{-1}$ defines a inverse matrix (or Green's function) corresponding to the complex shift z_l.
Discretized Cauchy contour integral technique can accurately approximate the density matrix with only $\mathcal{O}(\log \beta \Delta E)$ terms. Here $\Delta E := \max_{1 \le i \le N} |\mu - \lambda_i|$ is the spectral radius of the matrix pencil (H, S). The number of poles needed in practice is typically $40 \sim 80$.

KS-DFT : Diagonalization and FOE method

- Dense Hamiltonian matrix H diagonalization to get eigenvalues and eigenvectors
  - Computational complexity $O(N^3)$, expensive for big systems(more than 1000 atoms)
  - Scales up to 5,000 CPU cores for thousands of atoms
- Kohn-Sham DFT with selected inversion:

$$\rho(\Gamma) = \Sigma_{i=1}^{Q} Diag(\frac{w_i}{(H - z_i I)})$$

  - Computational complex 1D system: $O(N_e)$; 2D system: $O(N_e^{1.5})$, 3D system $O(N_e^2)$
  - Compared with linear scaling method, selected inversion method works for both metallic and semi-conductor system and remain accurate at low temperatures.

PEXSI: Pole EXpansion Selected Inversion package

- PEXSI used contour integral poles to expand the Fermi Operator
- Scale up to thousands of CPU cores for single pole
- Needs 40-80 poles in real DFT calculations
- Tow level parallelization: embarrassing parallel over the poles and $Np_{row} \times Np_{column}$ parallel within single pole
- Integrated into BigDFT, CP2K, SIESTA, DGDFT, FHI-aims, QuantumWise ATK
- Typically used for accelerating material simulation with more than $10,000$ atoms
- one challenge is determine the chemical potential μ, so that $|Ne_{computed} - Ne_{exact}|$ within given tolerance.

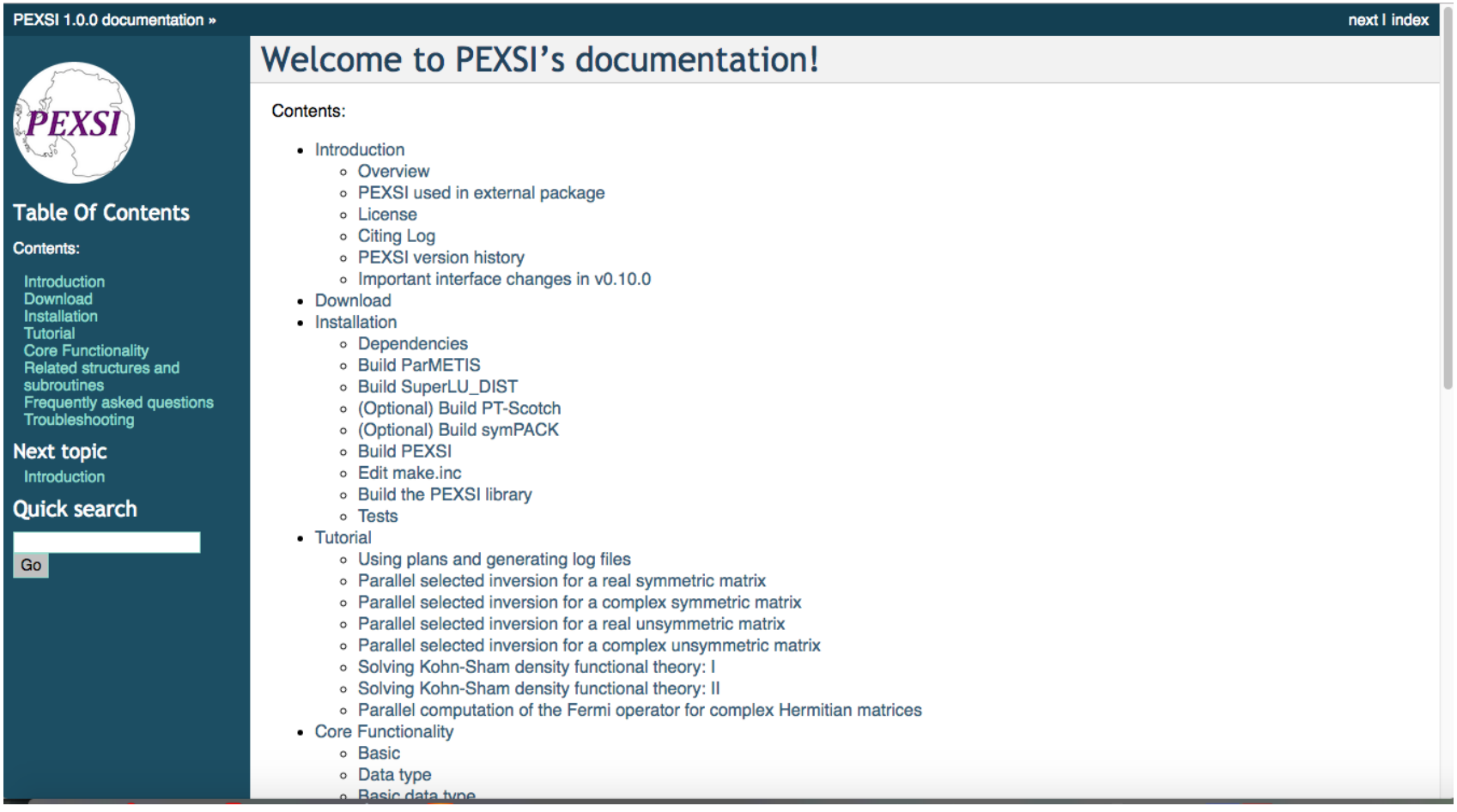
Moussa's optimized pole expansion method

$$\Gamma \approx \sum_{l=1}^{P} \mathfrak{Im}\left(\omega_l^{\rho}(H-(z_l+\mu)S)^{-1}\right). \tag{6}$$

- Pros:
  - modifying the Zolotarev expansion for the sign function through numerical optimization.
  - reduce the number of poles to $10-30$
  - the number of poles only approximately depends on $\beta\Delta E_o$. Here $\Delta E_o := \max_{1\le i\le N}(\mu-\lambda_i)$ is the spectral radius corresponding to the occupied eigenvalues, which can be significantly smaller than the spectral radius of the matrix pencil (H,S).
- Cons:
  - requires extra work for computing energy density matrix
  - can not achieve very high accuracy such as 10^{-9} Hartree/Bohr, which contour integral approach can.

Ref : J. E. Moussa, Minimax rational approximation of the fermi-dirac distribution

PEXSI Version 1.0



Fermi Dirac distribution of metallic and semi-conductor systems

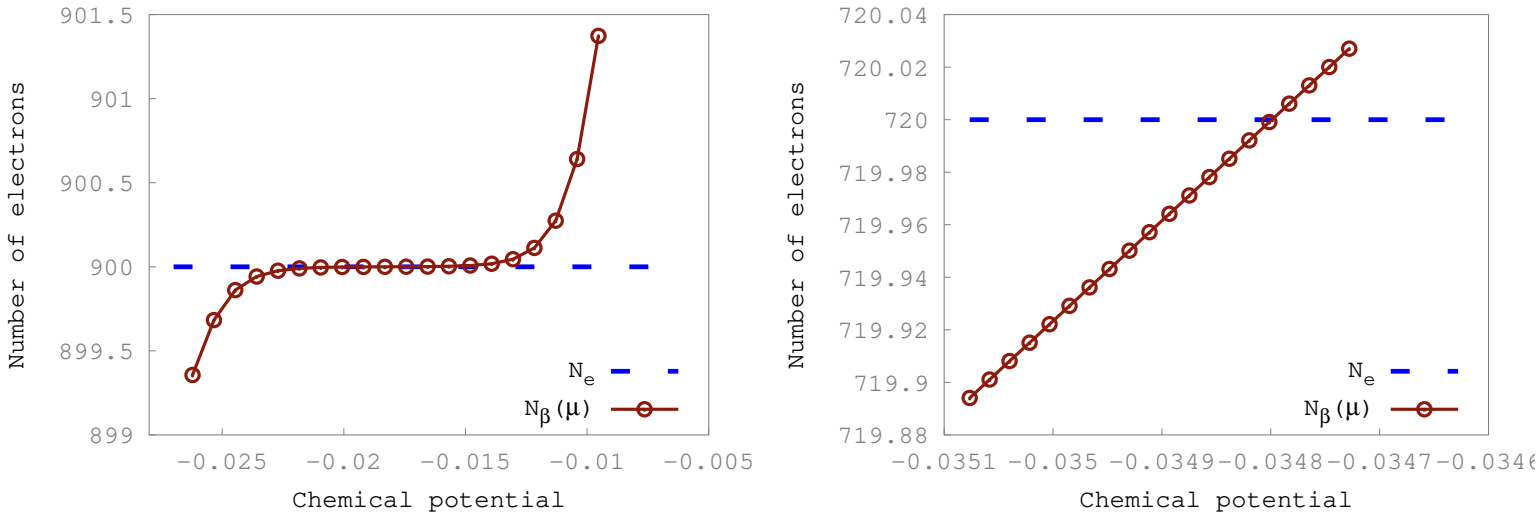


Figure 1: Fermi Dirac distribution for phospherene nanoribbon 180 atoms(left) and graphene 180 atoms (right).

In single SCF step:

- bisection method search for the chemical potential μ.
- use Newton's method to take derivative information into account to accelerate convergence.
- derivative info of $N_\beta(\mu)$ w.r.t μ vanishes when μ is inside a band gap.

Robust chemical potential determination

Instead of calculating a correct μ at each SCF step, we keep a rigorous lower and upper bound of the chemical potential.

- only one PEXSI iteration in each SCF step.
- introduce a point parallelization into the Fermi operator evaluation.
- scaling/interpolate the chemical potential μ, density matrix Γ and the energy density matrix
- Two levels: coarse level - inertia counting; fine level - Fermi operator evaluation.

Coarse level: Inertia counting

$N_\infty(\mu)$, here $\beta = \infty$ means zero temperature limit, can be evaluated by Sylvester's law of inertia. We perform

$$H - \mu S = LDL^T,$$

since the Fermi-Dirac function $f_\beta(x)$ is a non-increasing function and rapidly approaches 1 when $x < 0$ and 0 when $x > 0$, we can select a number τ_β so that we can approximate

$$f_\beta(x) \approx \begin{cases} 1, & x \leq -\tau_\beta, \\ 0, & x \geq \tau_\beta. \end{cases} \tag{7}$$

τ_β is a tunable parameter but is not system dependent. In practice we find that setting $\tau_\beta = 3/\beta = 3k_B T$ is a sufficiently conservative value for the robustness of our method. With this controlled approximation, we have

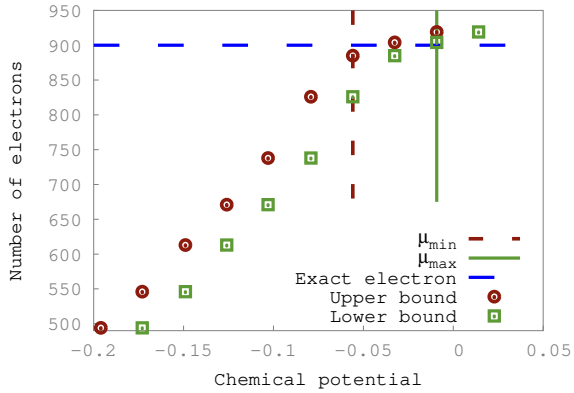
$$N_\beta(\mu - \tau_\beta) \leq N_\infty(\mu) \leq N_\beta(\mu + \tau_\beta). \tag{8}$$

Hence each evaluation of N_∞ provides an upper and a lower bound for N_β at two other energy points

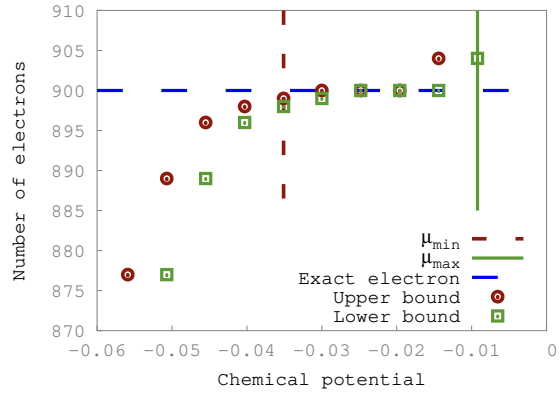
Coarse level: Inertia counting

- PEXSI requires both the sparse factorization and selected inversion, and inertia counting only requires a sparse factorization.
- PEXSI requires evaluations of P Green's functions to obtain one value of $N_\beta(\mu)$, and inertia counting obtains $N_\infty(\mu)$ with one factorization.
- PEXSI requires complex arithmetic, and for real matrices the inertia counting procedure only requires real arithmetic and thus fewer floating point operations. Hence the inertia counting step takes only a fraction of the time by each PEXSI evaluation.
- $N_{point} N_{pole} N_{sparse}$ CPU cores can do $N_{point} N_{pole}$ inertia counting, quickly converge the coarse level (μ_{min}, μ_{max})

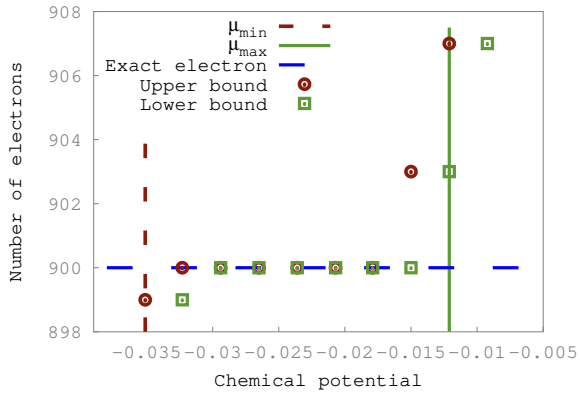
The inertia counting of the PEXSI



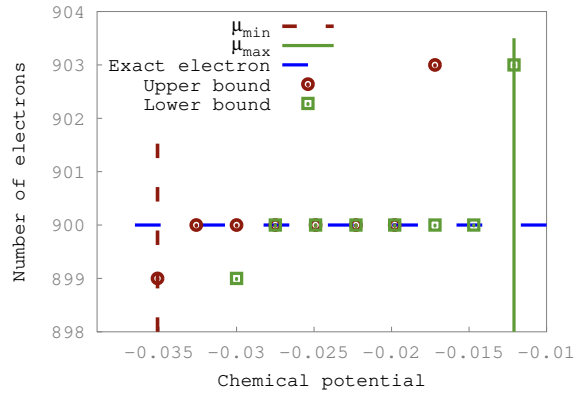
(a) Inertia counting step 1.



(b) Inertia counting step 2.



(c) Inertia counting step 3.



(d) Inertia counting step 4.

Figure 2: Refinement of the bounds for the chemical potential in 4 inertia counting steps for the PNR 180 system. In step 4, the inertia counting procedure stops because the upper and lower bounds can not be further

Fine level: multiple point evaluation of Fermi Operator

In a single SCF step:

- launch N_{point} Fermi operator evaluation. Note that N_{point} are evaluated simultaneously.
- Update the (μ_{min}, μ_{max}) according to the $N_\beta(\mu_g)$, g $= 1...N_{point}$ Hence in each SCF, the (μ_{min}, μ_{max}) will shrink by a factor of $(N_{\text{point}} - 1)^{-1}$

between SCF steps:

- keep a rigorous lower and upper bound of the (μ_{min}, μ_{max}) by updating the search interval with $(\mu_{min} + \Delta V_{min}\ ,\ \mu_{max} + \Delta V_{max})$.

Fine level: interpolate the chemical potential and density matrix

$$\mu_g = \mu_{\text{min}} + \frac{g-1}{N_{\text{point}}-1}(\mu_{\text{max}} - \mu_{\text{min}}), \quad g = 1, \ldots, N_{\text{point}}, \tag{9}$$

Then the chemical potential is found by solving the linear equation

$$N_e = N_\beta(\mu_g^*) + \frac{\mu - \mu_{g^*}}{\mu_{g^*+1} - \mu_{g^*}} \left(N_\beta(\mu_{g^*+1}) - N_\beta(\mu_g^*) \right). \tag{10}$$

Once μ is obtained, the density matrix is linearly mixed similarly as

$$\Gamma = \Gamma(\mu_g^*) + \frac{\mu - \mu_{g^*}}{\mu_{g^*+1} - \mu_{g^*}} \left(\Gamma(\mu_{g^*+1}) - \Gamma(\mu_g^*) \right). \tag{11}$$

Note that the interpolation procedure does not guarantee that μ satisfies the condition (4). However, it will ensure that the search interval is reduced at least by a factor $(N_{\text{point}} - 1)^{-1}$, and hence with k steps of iteration, the convergence rate of the chemical potential is at least $(N_{\text{point}} - 1)^{-k}$.

Fine level: multiple point evaluation of Fermi Operator

$$\mu_g = \mu_{\text{min}} + \frac{g-1}{N_{\text{point}}-1}(\mu_{\text{max}} - \mu_{\text{min}}), \quad g = 1, \ldots, N_{\text{point}}, \quad (12)$$

Then the chemical potential is found by solving the linear equation

$$N_e = N_\beta(\mu_g^*) + \frac{\mu - \mu_{g^*}}{\mu_{g^*+1} - \mu_{g^*}} \left(N_\beta(\mu_{g^*+1}) - N_\beta(\mu_g^*) \right). \quad (13)$$

Once μ is obtained, the density matrix is linearly mixed similarly as

$$\Gamma = \Gamma(\mu_g^*) + \frac{\mu - \mu_{g^*}}{\mu_{g^*+1} - \mu_{g^*}} \left(\Gamma(\mu_{g^*+1}) - \Gamma(\mu_g^*) \right). \quad (14)$$

Note that the interpolation procedure does not guarantee that μ satisfies the condition (4). However, it will ensure that the search interval is reduced at least by a factor $(N_{\text{point}} - 1)^{-1}$, and hence with k steps of iteration, the convergence rate of the chemical potential is at least $(N_{\text{point}} - 1)^{-k}$.

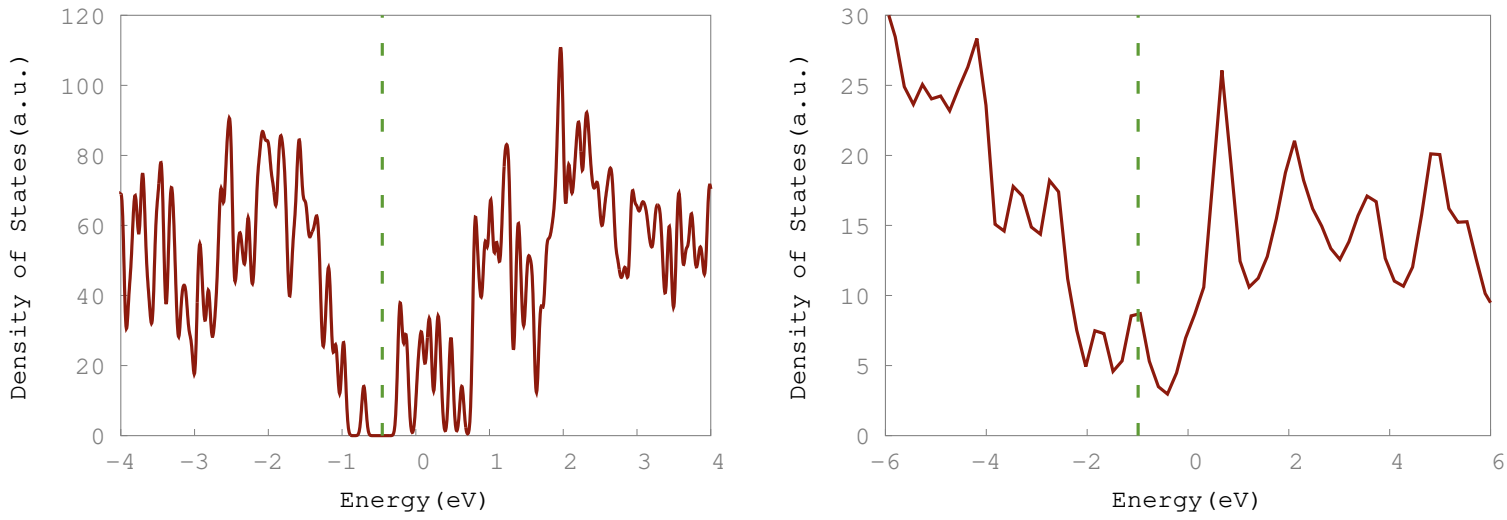
Robust chemical potential search in SCF calculation

Algorithm 1 Robust determination of the chemical potential.

Input:
$\tau^{\mu}_{\text{inertia}}, \tau^{N_e}$, $N_{\text{proc}} = N_{\text{sparse}} N_{\text{point}} N_{\text{pole}}$, $(\mu_{\text{min}}, \mu_{\text{max}})$ as the initial search interval.
Output:
Converged density matrix Γ and chemical potential μ.

while SCF has not converged **do**
  Construct the projected Hamiltonian matrix H and overlap matrix S
  while Inertia counting has not converged **do**
    for $g = 1$ to $N_{pole} N_{\text{point}}$ **do**
      Evaluate $N_{\infty}(\mu_g)$ for the processor group associated with μ_g
    Construct the upper and lower bounds for $N_{\beta}(\mu_g)$ for each point μ_g.
    Update the search interval $(\mu_{\text{min}}, \mu_{\text{max}})$.
  for $g = 1$ to N_{point} **do**
    for $l = 1$ to N_{pole} **do**
      Evaluate $(H - (z_l + \mu)S)^{-1}$ for the processor group associated with l
    Evaluate μ and Γ for processor group associated with μ_g if needed.
  Perform linear interpolation for μ and Γ
  Evaluate the density and the new potential.
  Compute the difference of the potential ΔV.
  Update the search interval $(\mu_{\text{min}}, \mu_{\text{max}}) \leftarrow (\mu_{\text{min}} + \Delta V_{\text{min}}, \mu_{\text{max}} + \Delta V_{\text{max}})$.

DOS of the PNR and GRN 180 atoms system

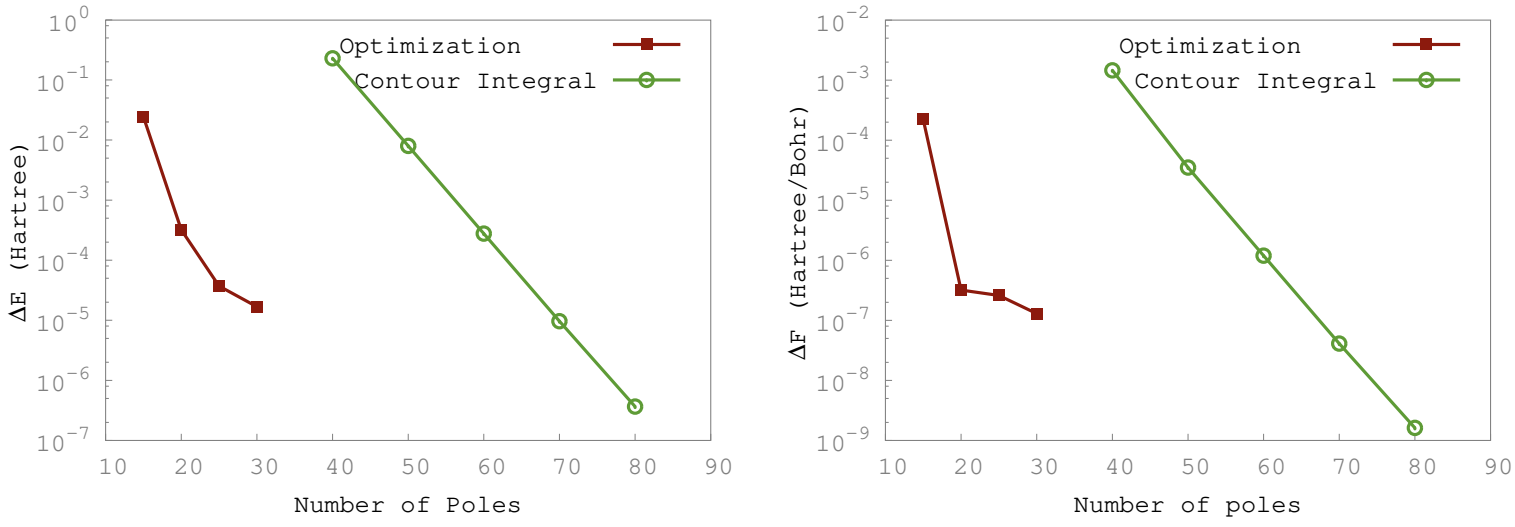


(a) PNR 180 (b) GRN 180

Figure 3: The total densities of states(DOS) of PNR and GRN 180 atoms, respectively. The fermi levels are marked by the green dash line.

The tests are performed on Edison, NERSC, with our DGDFT package.

Pole Expansion Accuracy: Moussa's optimized and Contour integral pole expansion

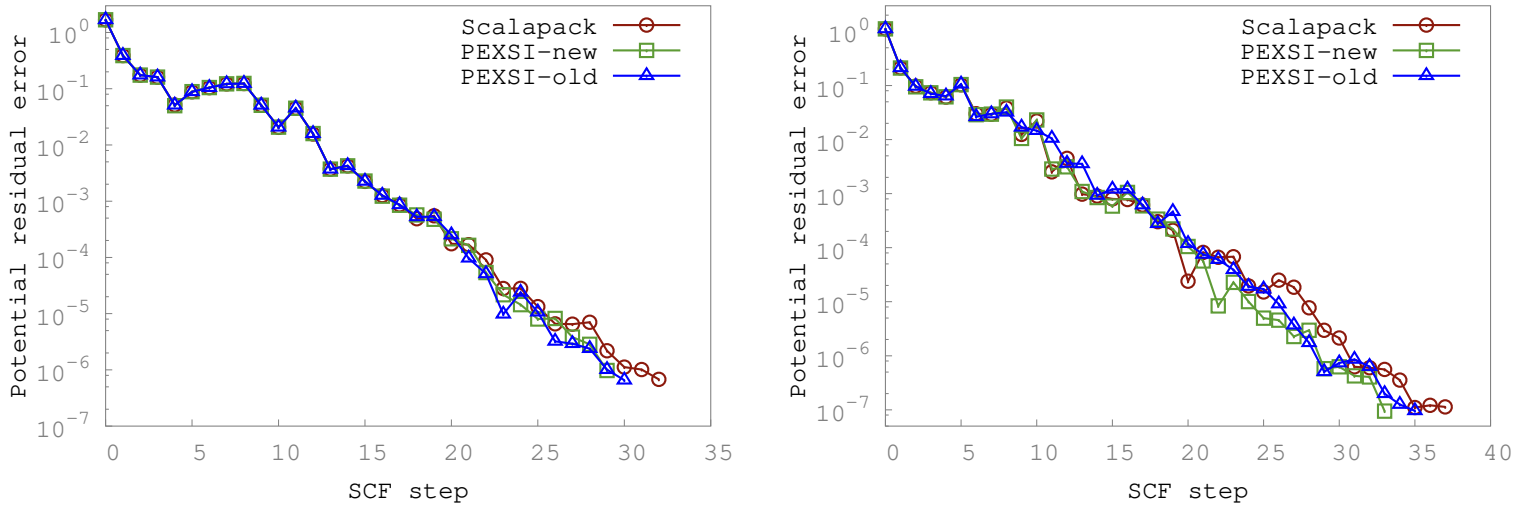


(a) Energy

(b) Force

Figure 4: Error of the energy and force of the PEXSI method with respect to the number of poles for the graphene 180 system.

Pole Expansion SCF Accuracy: PNR and GRN 180 system

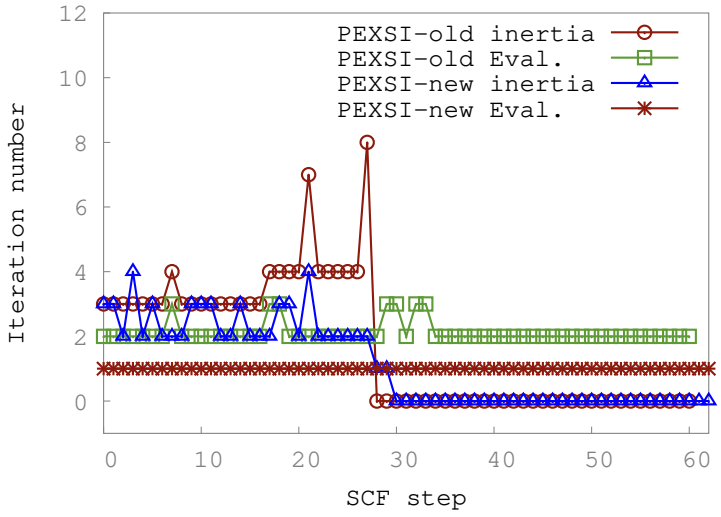


(a) PNR 180.

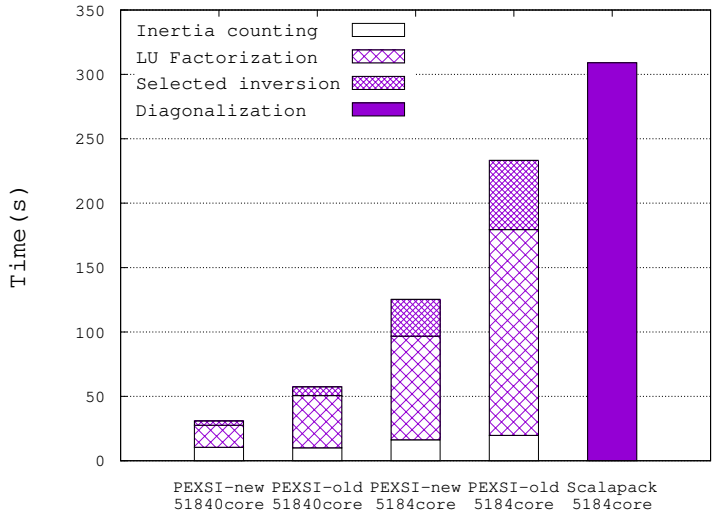
(b) GRN 180.

Figure 5: SCF convergence with ScaLAPACK and PEXSI-old and PEXSI-new along the SCF steps for PNR 180 and GRN 180 systems.

PEXSI Timing



(a) Number of inertia counting step and evaluation of the Fermi operator steps.



(b) The timing for PEXSI-new, PEXSI-old and diagonalization per SCF step.

Figure 6: Comparison of the performance of PEXSI-new, PEXSI-old and diagonalization for the GRN 6480 system.

Convergence of the PEXSI interpolation

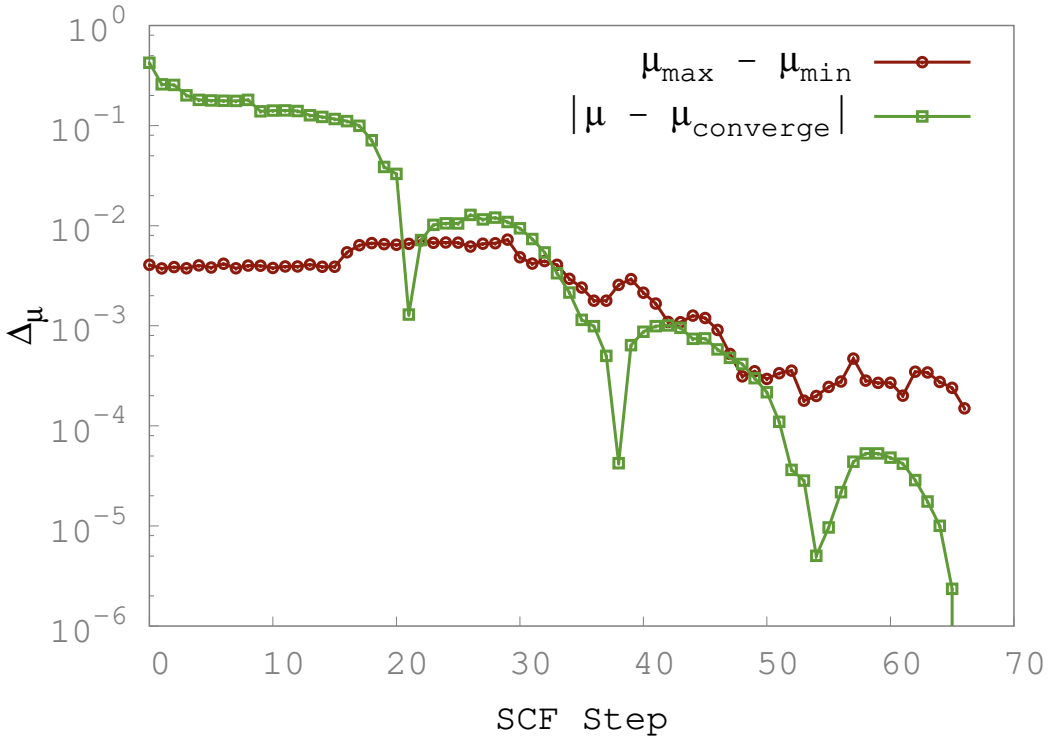
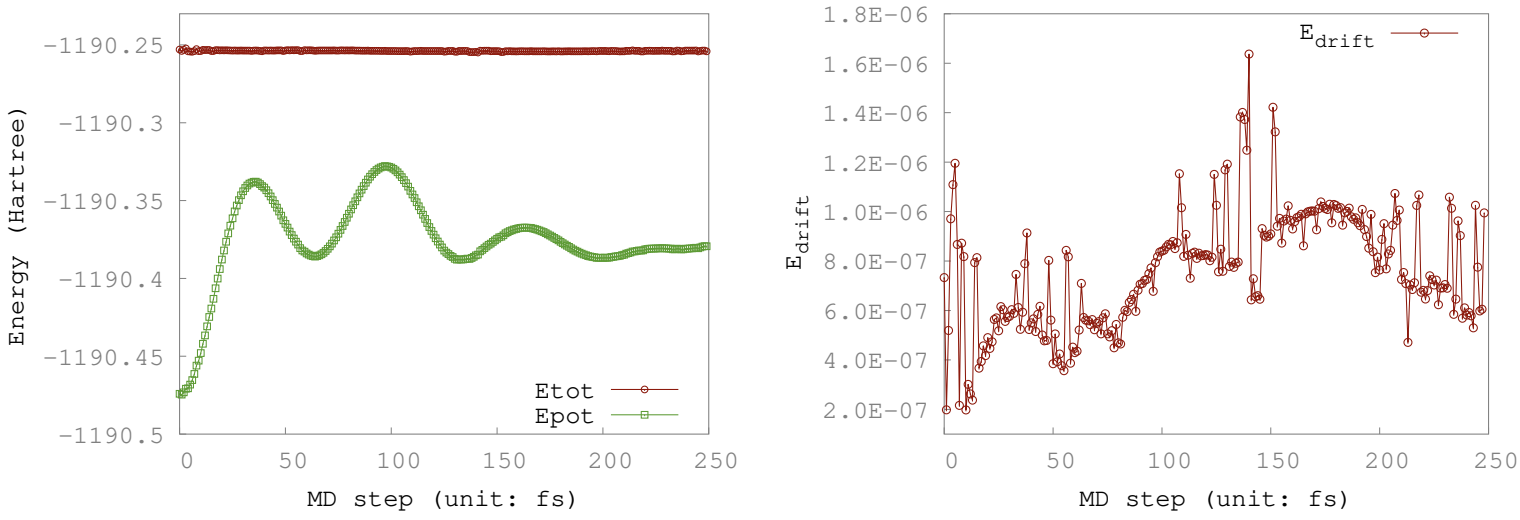


Figure 7: $\mu_{max} - \mu_{min}$ and $|\mu - \mu_{convergence}|$ with respect to SCF iteration

MD simulation



(a) Total and potential energy

(b) Energy drift

Figure 8: The potential energy, total energy and the drift of the total energy using the PEXSI-new method.

Conclusion

- We present a robust and efficient algorithm in determine the chemical potential in the PEXSI for solving KSDFT.
- Dynamically update the lower and upper bounds of the chemical potential at each SCF.
- Linear interpolation of the chemical potential μ and charge density matrix Γ
- Only one Fermi operator evaluation is performed in a single SCF step. The chemical potential converges with the SCF iterations.
- Moussa's optimized pole expansion is set as default in PEXSI 1.0
- Two points parallelization is recommended for PEXSI 1.0
- Tests show that we have a speedup of 2x for Graphene 6480 atoms.

Thank you for your attention

Acknowledgement