TOWARDS DIRECT AND ACCURATE AB INITIO MODELING OF MATERIAL DEFECTS AT LARGE SCALE

Lin Lin

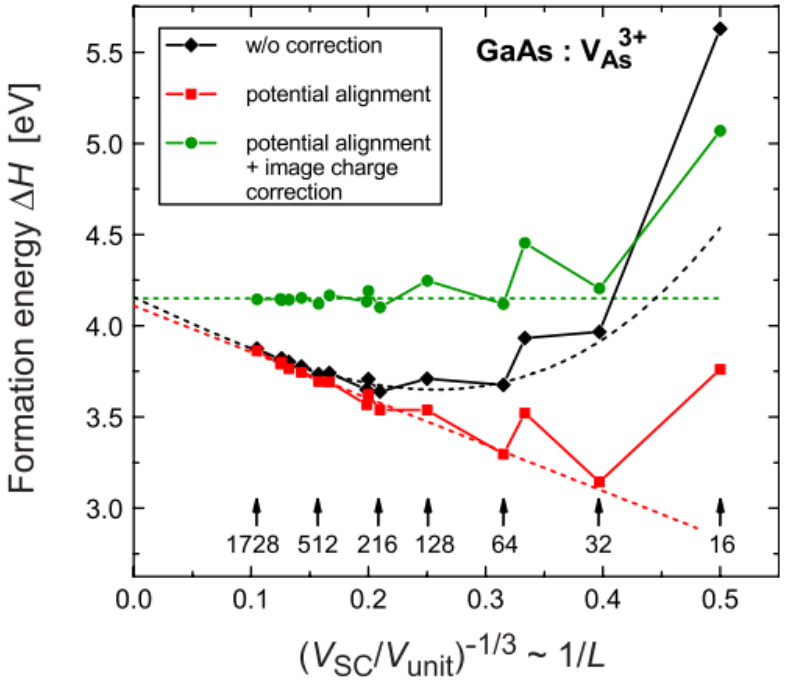
Computational Research Division,
Lawrence Berkeley National Lab

Quantum and Atomistic Modeling of Materials Defects, IPAM, October 2012

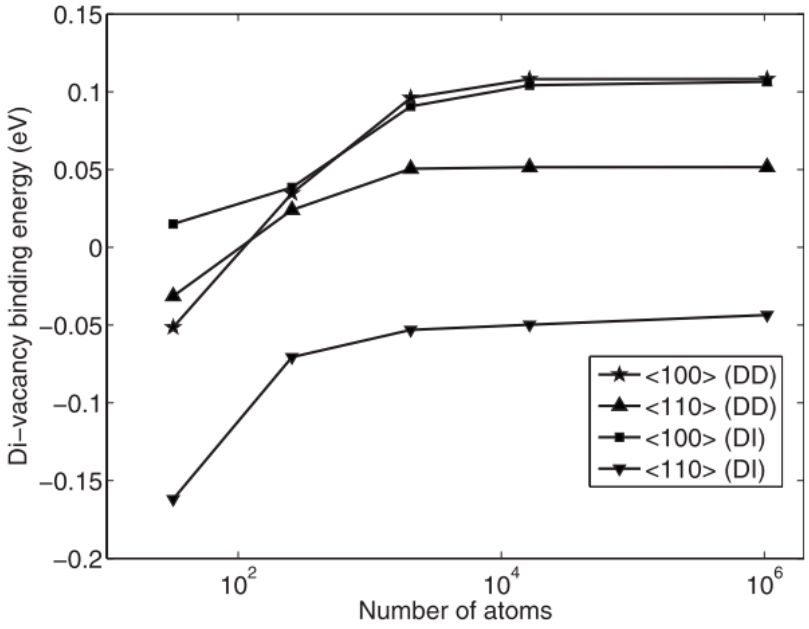
Why large scale systems

- Overcome the finite size effect in quantum mechanical calculation
- Quantum dot and nano system
- Defect formation energy
- Dislocation core
- Solvent, interfaces

Defects



Formation energy of V_{As}^{3+} defect in GaAs
[Lany and Zunger, PRB 2008]



Binding energy divacancies of Al
[Radhakrishnan and Gavini, PRB 2010]

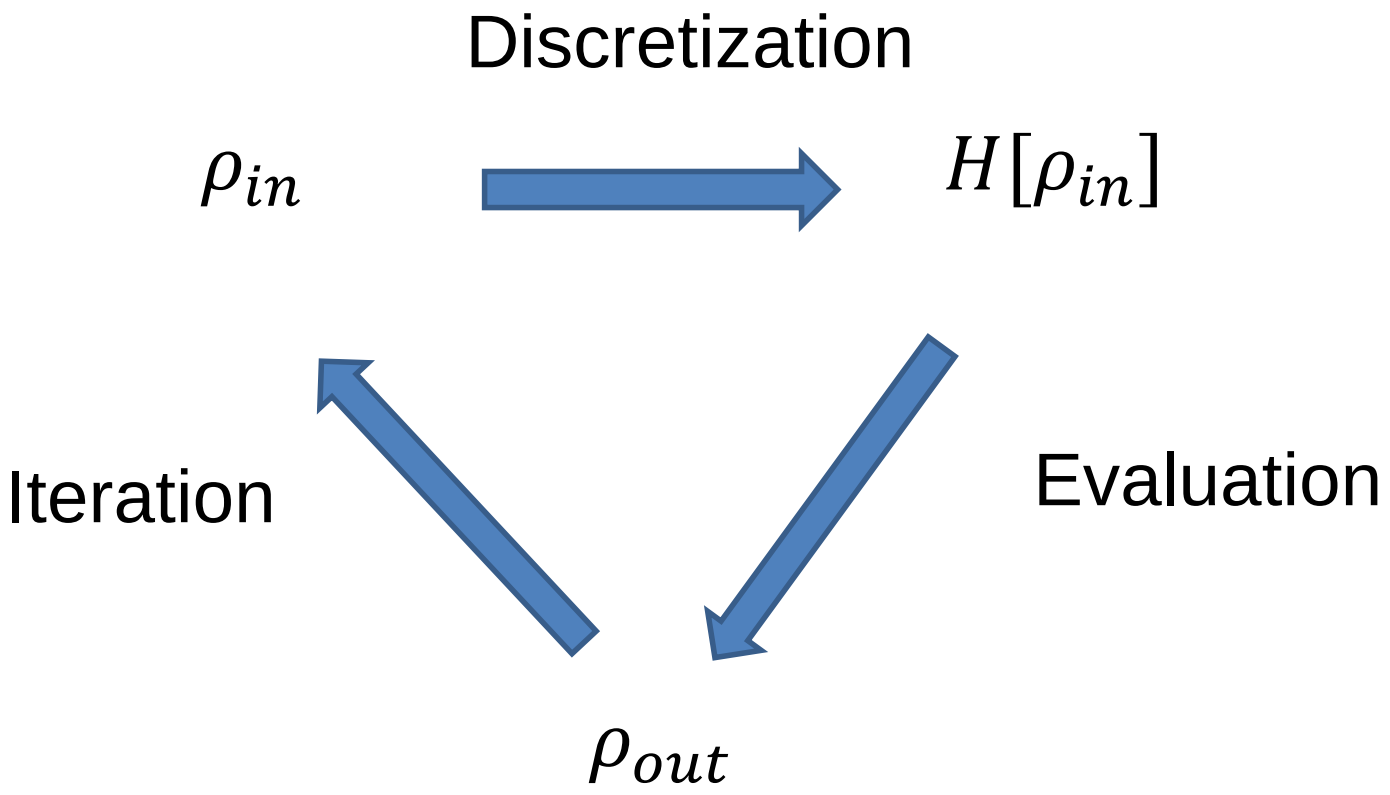
Kohn-Sham density functional theory

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

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

- Efficient: Single particle theory
- Accurate: Exact ground state energy for exact $V_{xc}[\rho]$, [Hohenberg-Kohn,1964], [Kohn-Sham, 1965]

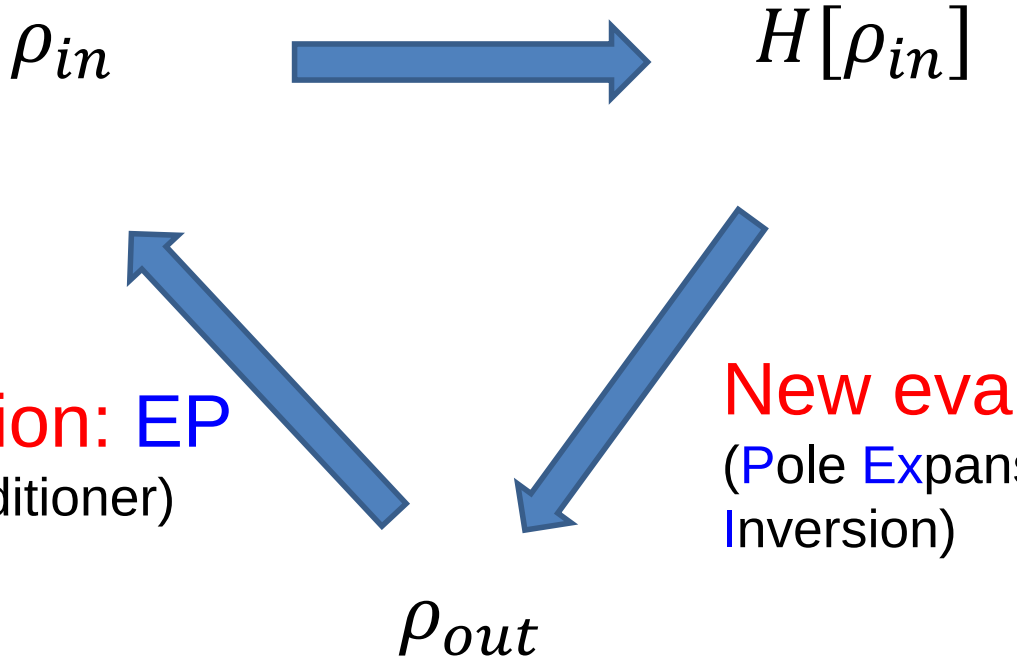
Self Consistent Field Iteration



Self Consistent Field Iteration

New discretization: ALB-DG

(Adaptive Local Basis in a Discontinuous Galerkin framework)



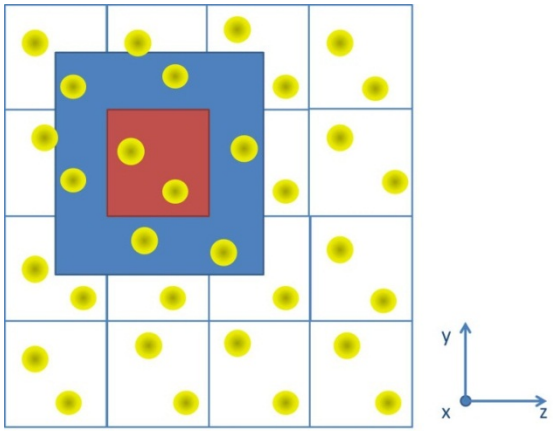
New iteration: EP

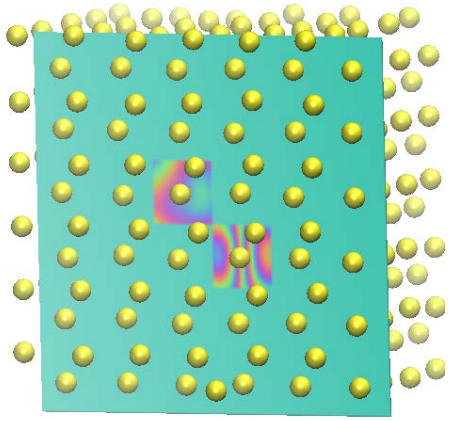
(Elliptic Precoditioner)

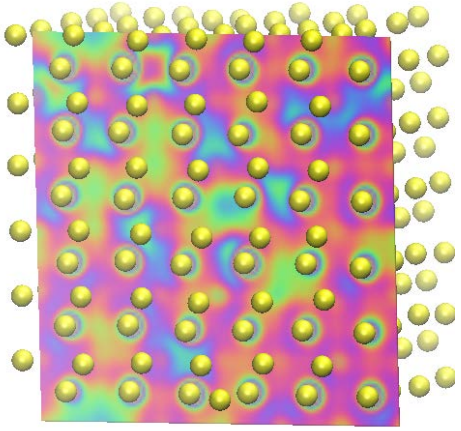
New evaluation: PEXSI

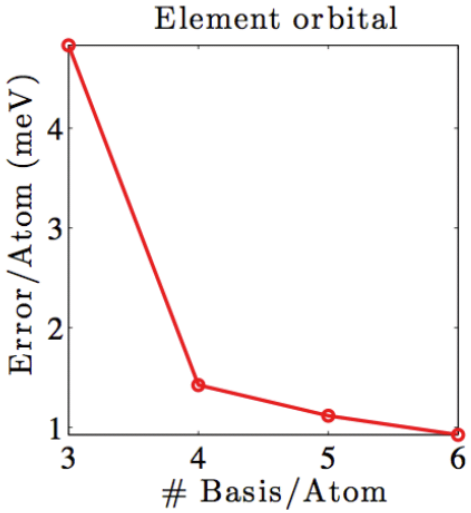
(Pole Expansion and Selected Inversion)

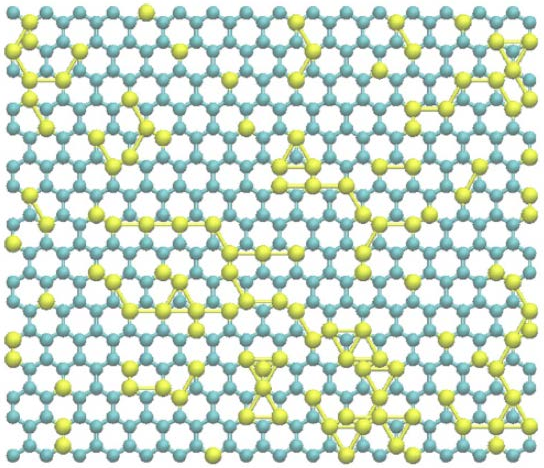
New discretization method





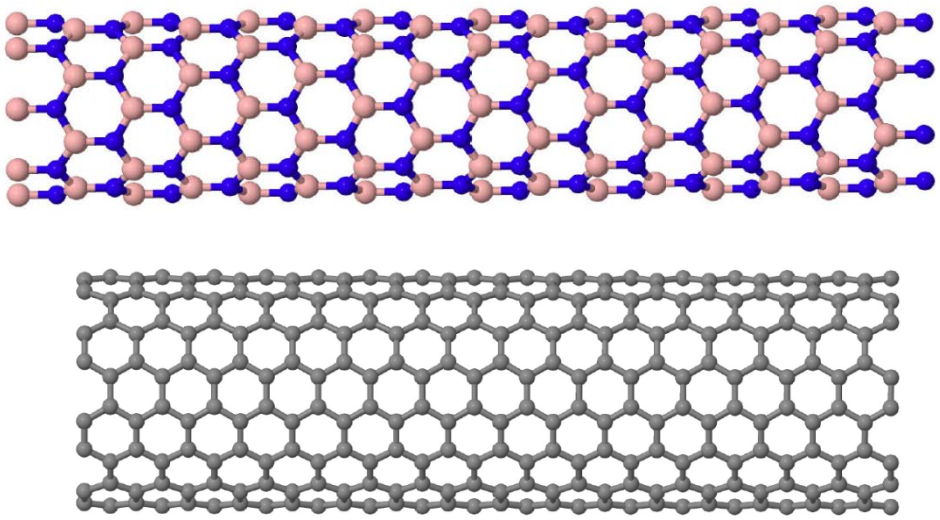


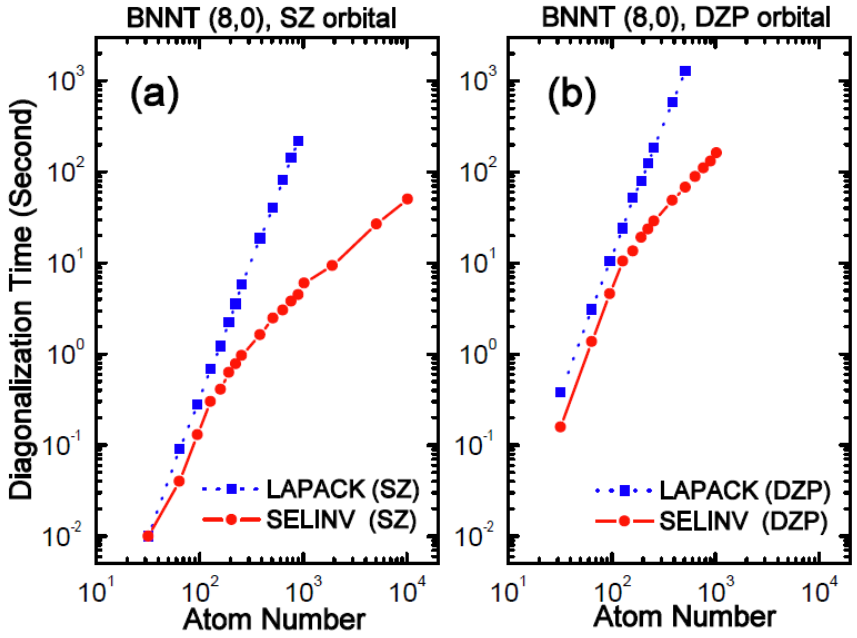




[LL-Lu-Ying-E, JCP 2012 (a)]
[LL-Lu-Ying-E, JCP 2012 (b)]
[LL-Ying, PRB, 2012]

New evaluation method

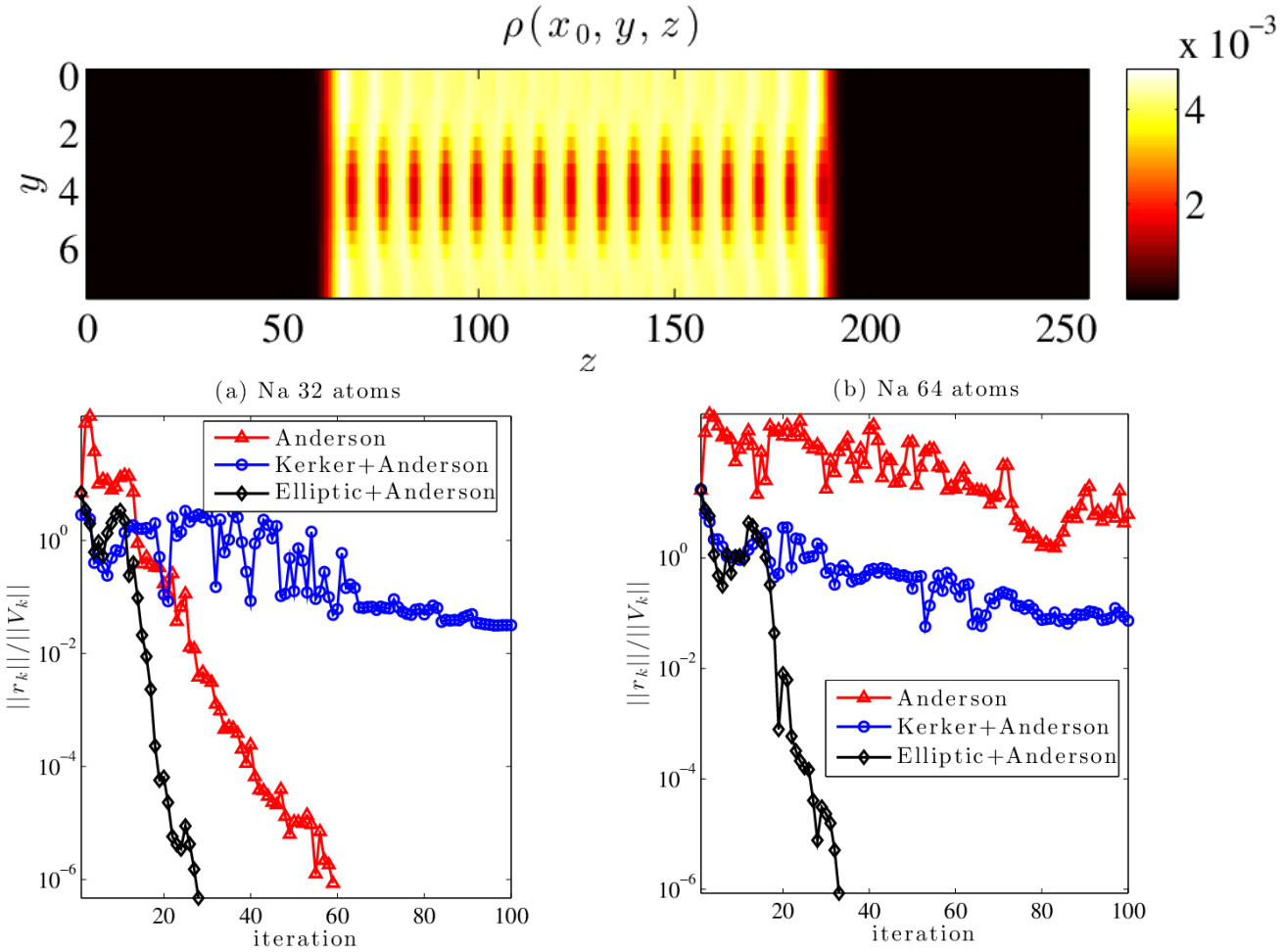




[LL-Lu-Car-E, PRB 2009]
[LL-Lu-Ying-E, CAM 2009]
[LL-Lu-Ying-Car-E, CMS, 2009]

[LL-Yang-Lu-Ying-E, SISC 2011]
[LL-Yang-Meza-Lu-Ying-E, TOMS, 2011]
[LL-Chen-Yang-He, submitted]

New iteration method

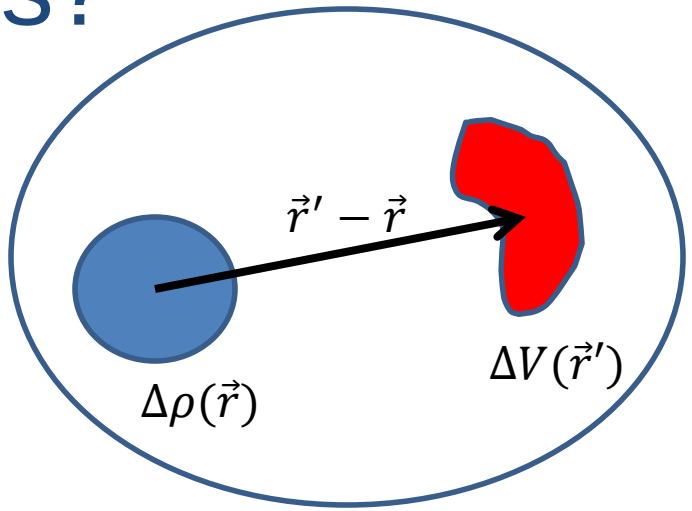


[LL-Yang, arxiv:1206.2225]

Evaluation by diagonalization

- Diagonalization: lowest N/2 eigenvalues and eigenfunctions
- Cubic scaling due to orthogonalization of an $O(N) \times O(N)$ matrix.
- regularly hundreds of atoms, at most ~10,000 atoms

Evaluation: Alternatives?



- Linear scaling algorithms
  - Near-sightedness [Kohn, 1996]
  - Truncation based algorithm:
    low to intermediate accuracy
  - Only applicable to insulators.

[Bowler and Miyazaki, Rep. Prog. Phys 2012]

"…The second challenge is that of metallic systems: there is no clear route to linear-scaling solution for systems with low or zero gaps and extended electronic structure…"

- Difficult task:
  - Accurate and efficient
  - Uniformly applicable to metals as well as insulators.

Outline

PEXSI: Pole EXpansion Selected Inversion

- Pole Expansion
- Selected Inversion
- Quantum chemistry basis set

KSDFT: Matrix point of view

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

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

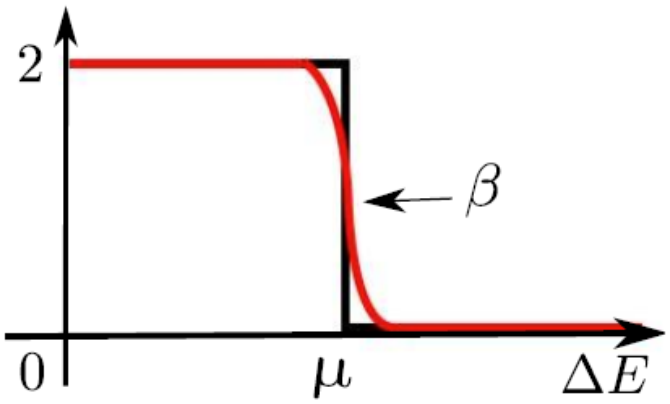
$$= \{\chi(H[\rho] - \mu I)\}_{x,x}$$

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

$$\rho = \operatorname{diag} \chi(H[\rho] - \mu I)$$

Finite temperature: Fermi operator

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



- $\beta = 1/k_B T$: inverse temperature
- μ: Chemical potential

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

Fermi operator expansion

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

$$\approx \mathrm{diag}\left\{\sum_{l=1}^{P} c_l\left(\frac{H[\rho]-\mu I}{\Delta E}\right)^l + \sum_{l=1}^{Q}\frac{\omega_l}{\left(z_l I - \frac{H[\rho]-\mu I}{\Delta E}\right)^{q_l}}\right\}$$

- $\Delta E = \sigma(H-\mu I)$.
- Fermi operator expansion: solving KSDFT without diagonalization

- [Goedecker, 1993], $P \sim O(\beta\Delta E)$
- [Head-Gordon et al, 2004], $P \sim O(\beta\Delta E)$ but with $O(\sqrt{\beta\Delta E})$ operation
- [Ceriotti et al, 2008], $\mathrm{Q} \sim O\left(\sqrt{\beta\Delta E}\right)$; other work

Pole expansion

- [LL, Lu, Ying and E, 2009], $Q \sim O(\log(\beta \Delta E))$

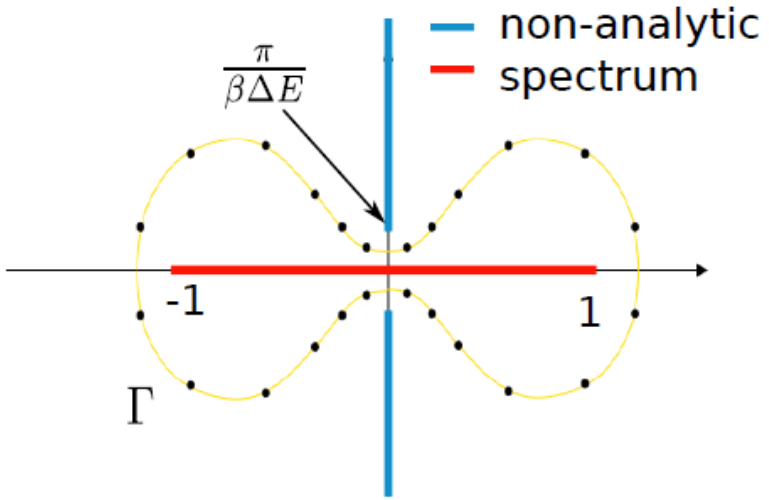
$$\rho \approx \mathrm{diag} \sum_{i=1}^{Q} \frac{\omega_i}{H - z_i I}$$

- $z_i, \omega_i \in \mathbb{C}$ are complex shifts and complex weights

Contour integral technique

Fermi-Dirac

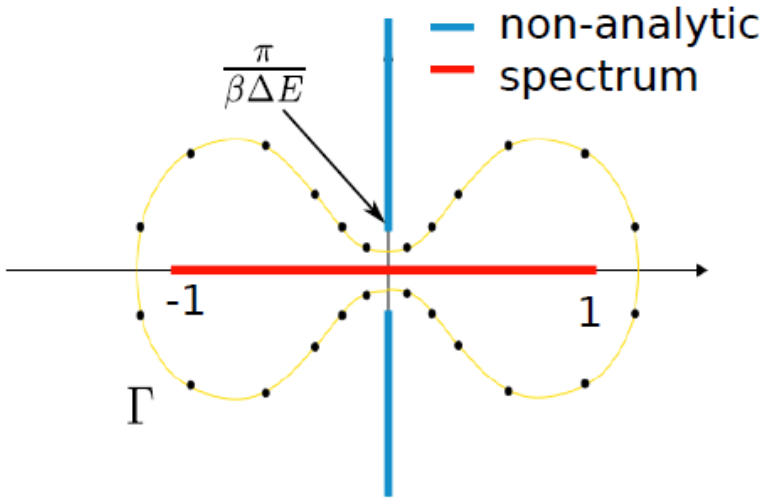
$$\rho(\xi) = \frac{1}{2\pi i}\oint_{\Gamma}\frac{\rho(z)}{z-\xi}dz \approx \frac{1}{2\pi i}\sum_{i=1}^{Q}\frac{\rho(z_i)w_i}{z_i-\xi}$$



Contour integral technique

Fermi-Dirac

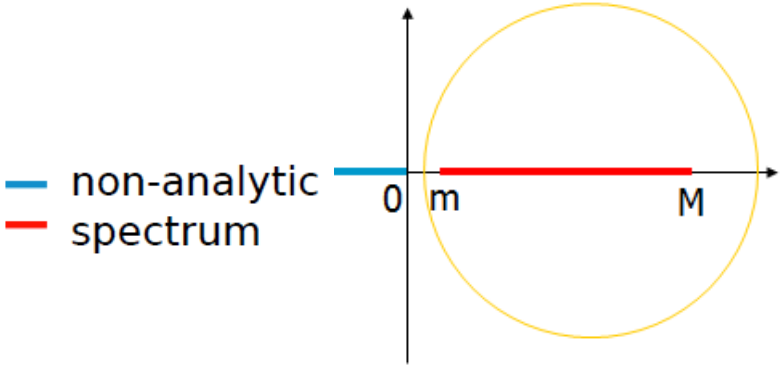
$$\rho(\xi) = \frac{1}{2\pi i}\oint_\Gamma \frac{\rho(\xi)}{z-\xi}dz \approx \frac{1}{2\pi i}\sum_{i=1}^{Q}\frac{\rho(z_i)w_i}{z_i-\xi}$$



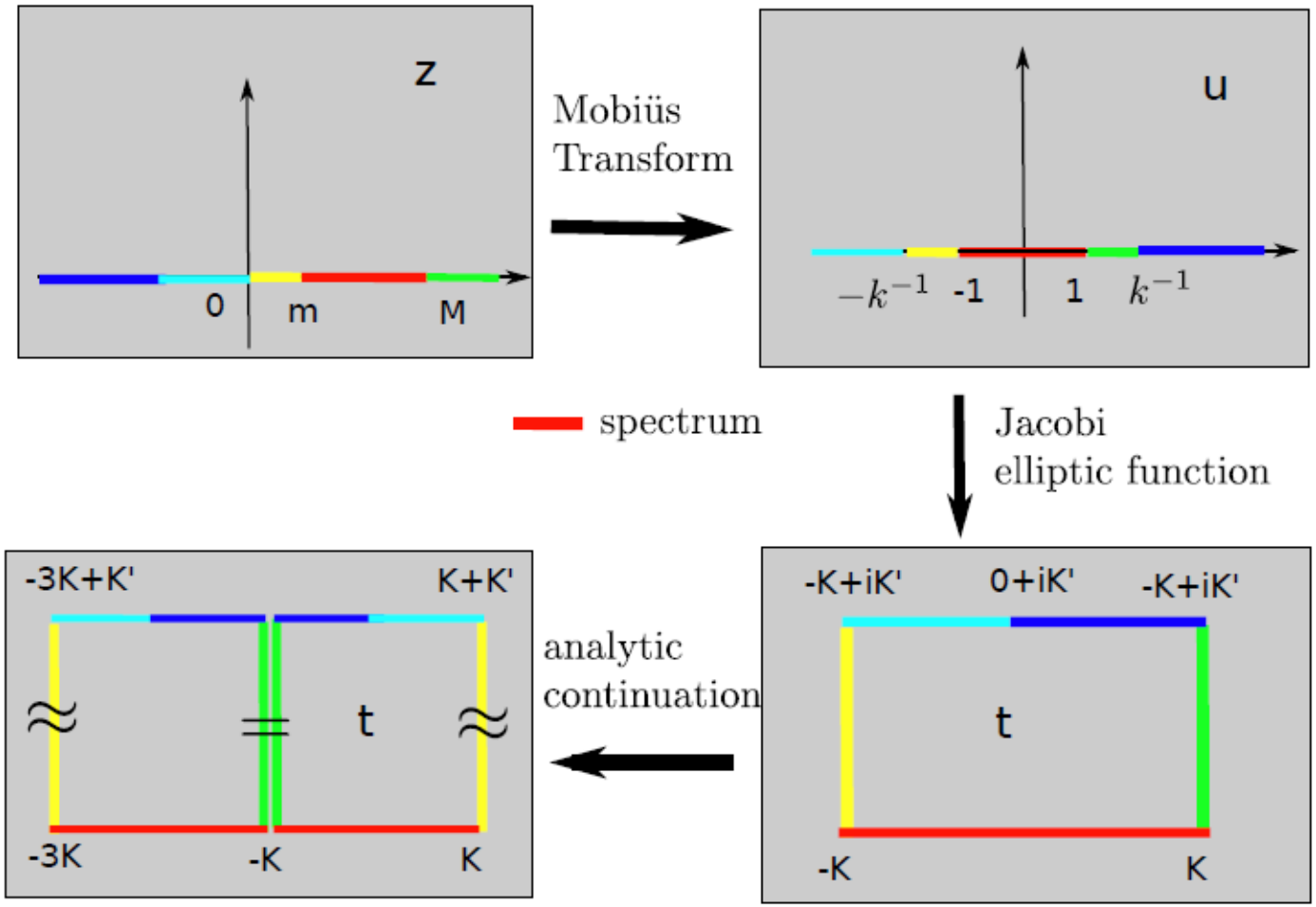
Simpler problem

[Hale, Higham and Trefethen, 2008]

$$\left\|\rho(\xi) - \rho_Q(\xi)\right\| \sim O(e^{-CQ/\log(M/m)})$$

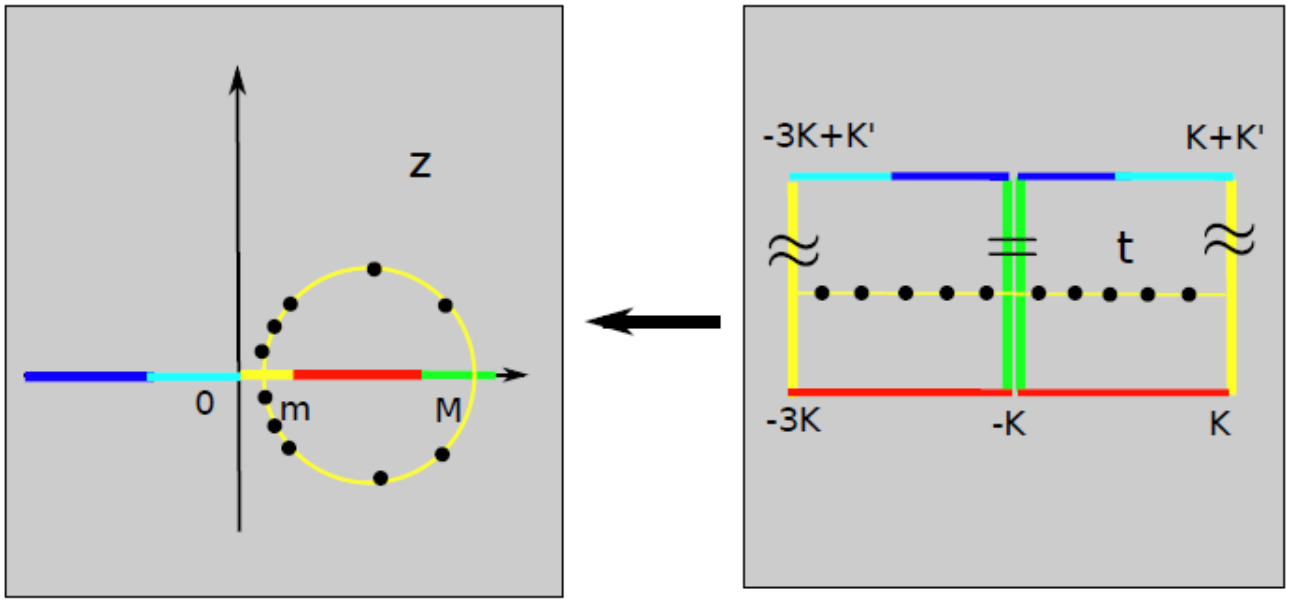


Domain transformation

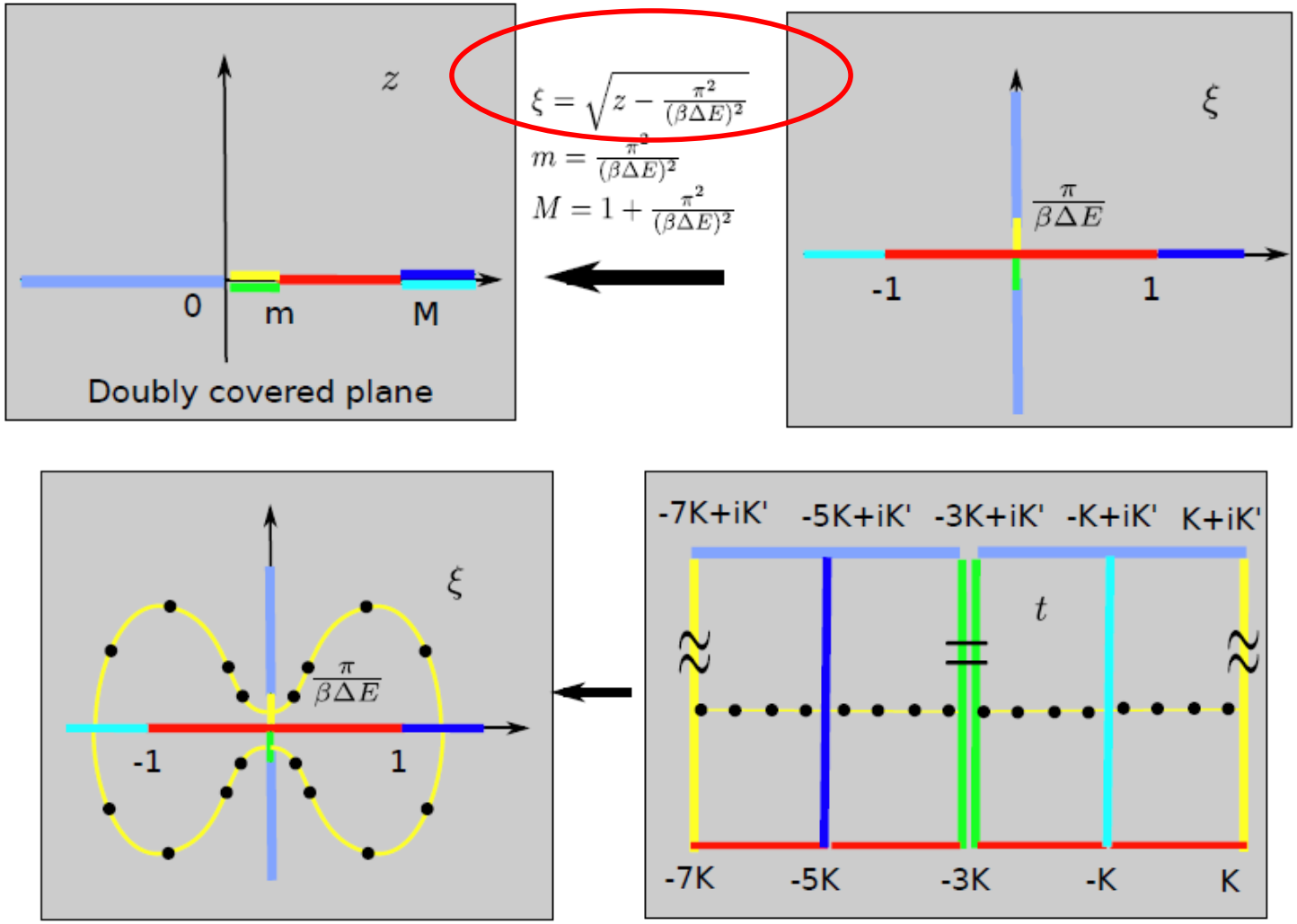


Contour selection

- [Hale, Higham, Trefethen 2008] $\frac{K'}{2K} \sim \frac{1}{\log \frac{M}{m}}$
- Trapezoid rule for periodic function gives geometric convergence

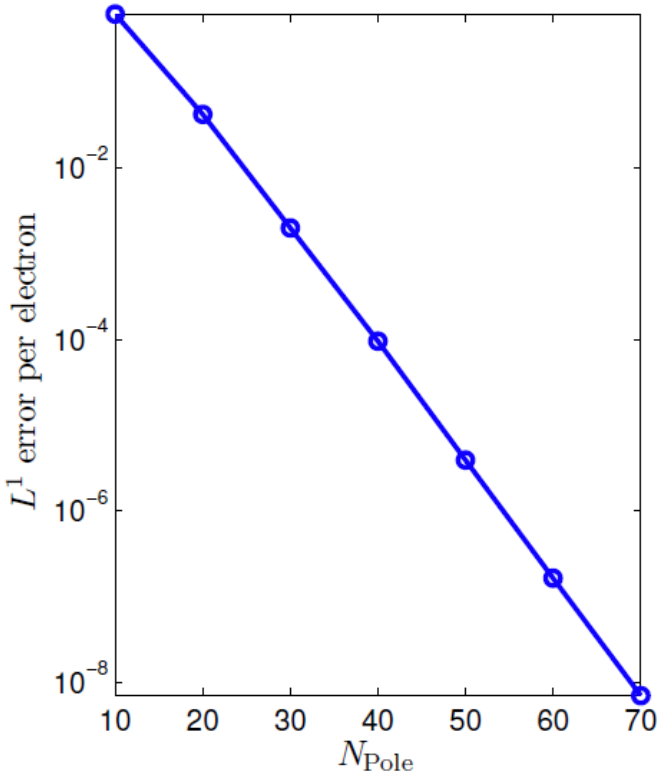


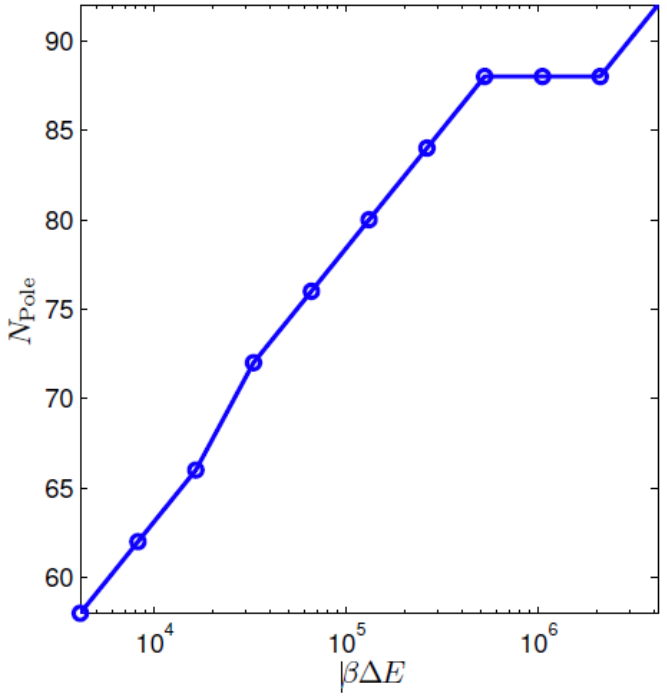
Pole expansion



Numerical result

H: Tight binding model on a 2D grid





Outline

PEXSI: Pole EXpansion Selected Inversion

- Pole Expansion
- Selected Inversion
- Quantum chemistry basis set

Selected inversion

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

- All the diagonal elements of an inverse matrix.
- H is a sparse matrix, but $(H - z_i I)^{-1}$ is a full matrix.
- Naïve approach: $O(N^3)$.
- Need selected inversion.

Selected inversion: basic idea

- 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

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

Observation:

If L_{21} is sparse, $L_{21}^T S_{22}^{-1} L_{21}$ only require rows and columns of S_{22}^{-1} corresponding to the sparsity pattern of L_{21}.

Recursive relation

$$S_{22} = \begin{pmatrix} A_{22} & A_{32}^T \\ A_{32} & \hat{A}_{33} \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 0 \\ L_{21} & I \end{pmatrix} \boxed{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & L_{32} & I \end{pmatrix} \begin{pmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & 0 \\ 0 & 0 & \hat{A}_{33} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & L_{32}^T \\ 0 & 0 & I \end{pmatrix}} \begin{pmatrix} 1 & L_{21}^T \\ 0 & I \end{pmatrix}$$

$$A^{-1} = \begin{pmatrix} A_{11}^{-1} + L_{21}^T S_{22}^{-1} L_{21} & -L_{21}^T S_{22}^{-1} \\ -S_{22}^{-1} L_{21} & \boxed{\begin{matrix} A_{22}^{-1} + L_{32}^T S_{33}^{-1} L_{32} & -L_{32}^T S_{33}^{-1} \\ -S_{33}^{-1} L_{32} & S_{33}^{-1} \end{matrix}} \end{pmatrix}$$

Recursive relation

- $I = \{i | L_{21}(i,1) \neq 0\}, 2 \in I$
- $L_{21}(i,1) \neq 0 \Rightarrow S_{22}(i,j) \neq 0, i,j \in I$

  because $S_{22} = A_{22} - A_{21}L_{21}^T$

  $\Rightarrow L_{32}(i,2) \neq 0, i \in I$

- $A^{-1} = \begin{pmatrix} A_{11}^{-1} + L_{21}^T S_{22}^{-1} L_{21} & -L_{21}^T S_{22}^{-1} \\ -S_{22}^{-1} L_{21} & \begin{matrix} A_{22}^{-1} + L_{32}^T S_{33}^{-1} L_{32} & -L_{32}^T S_{33}^{-1} \\ -S_{33}^{-1} L_{32} & S_{33}^{-1} \end{matrix} \end{pmatrix}$

General result

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

- For experts: calculating A_{ij}^{-1} only requires A_{kl}^{-1} such that k, l are on the critical path of i and j on the elimination tree

[LL, Yang, Meza, Lu, Ying and E, 2011]

Similar work: [Erisman and Tinney, 1975], [Takakashi et al 1973], [Li, Darve et al, 2008], [LL, Lu, Ying, Car and E, 2009]

Complexity

- For Laplacian (or Schrödinger) operator, the cost of evaluating L is $O(N)$ for 1D systems, $O(N^{1.5})$ for 2D systems, and $O(N^2)$ for 3D systems.

- Combined with pole expansion:

  At most $O(N^2)$ scaling for solving Kohn-Sham problem.

Parallel selected inversion

[LL, Yang, Lu, Ying and E, SISC (2011)]

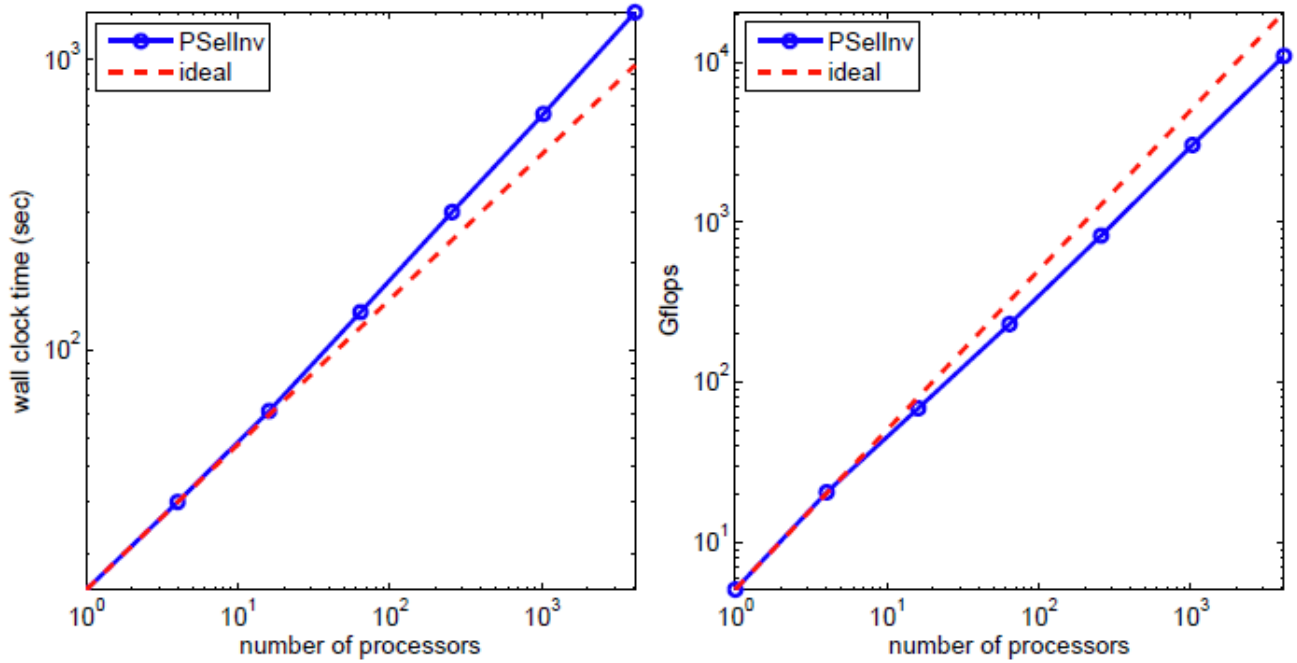
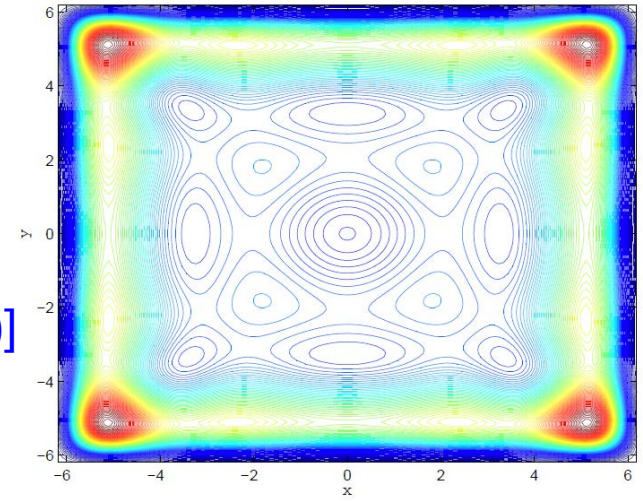


Figure: Log-log plot of total wall clock time and total Gflops with respect to number of processors, compared with ideal scaling. The largest matrix solved has $(4.3 \text{ billion})^2$ degrees of freedom using 4096 processors.

2D Quantum-dot

[LL, Yang, Lu, Ying and E, SISC(2011)]



n_e(#Electrons)	Grid	#proc	Octopus time(s)	New method time(s)
2	31	1	< 0.01	0.01×80
8	63	1	0.03	0.06×80
32	127	1	0.78	0.03×80
128	255	1	26.32	1.72×80
		4	10.79	0.59×80
512	511	1	1091.04	9.76×80
		4	529.30	3.16×80
		16	131.96	1.16×80
2048	1023	1	out of memory	60.08×80
		4	out of memory	19.04×80
		16	7167.98	5.60×80
		64	1819.39	2.84×80

SelInv: Numerical results

SelInv: a selected inversion package for general sparse symmetric matrix written in FORTRAN.
[LL-Yang-Meza-Lu-Ying-E, TOMS, 2011]

Problems from Harwell-Boeing Test Collection and the University of Florida Matrix Collection.

problem	n	selected inversion time	direct inversion time	speedup
bcsstk14	1,806	0.01 sec	0.13 sec	13
bcsstk24	3,562	0.02 sec	0.58 sec	29
bcsstk28	4,410	0.02 sec	0.88 sec	44
bcsstk18	11,948	0.24 sec	5.73 sec	24
bodyy6	0.19,366	09 sec	5.37 sec	60
crystm03	24,696	0.78 sec	26.89 sec	34
wathen120	36,441	0.34 sec	48.34 sec	142
thermal1	82,654	0.44 sec	95.06 sec	216
shipsec1	140,874	17.66 sec	3346 sec	192
pwtk	217,918	14.55 sec	5135 sec	353
parabolic_fem	525,825	20.06 sec	7054 sec	352
tmt_sym	726,713	13.98 sec	$>$ 3 hours	$>$ 772
ecology2	999,999	16.04 sec	$>$ 3 hours	$>$ 673
G3_circuit	1,585,478	218.7 sec	$>$ 3 hours	$>$ 49

Outline

PEXSI: Pole EXpansion Selected Inversion

- Pole Expansion
- Selected Inversion
- Quantum chemistry basis set

Why quantum chemistry basis set?

- Much smaller degrees of freedom (DOF) per atom

Basis	Example	DOF / atom
Uniform basis	Planewave Finite difference Finite element	500~10000 or more
Quantum chemistry basis	Gaussian orbitals atomic orbitals	4~100

- Difficulty
  - Nonorthogonal basis
  - Not mentioned yet: Energy, Free energy and Force
- [LL-Chen-Yang-He, submitted]

Setup

- Basis $\Phi = [\varphi_1(x) \cdots \varphi_N(x)]$
- Kohn Sham orbitals $\psi_i(x) = \sum_{j=1}^{N} \varphi_j(x) c_{ji}$ or $\Psi = \Phi C$
- Projection matrix $\mathrm{H}_{ij} = \int dx\, dx'\, \varphi_i(x) \hat{H}(x, x') \varphi_j(x')$
- Overlap matrix $S_{ij} = \int dx\, \varphi_i(x) \varphi_j(x)$

Kohn-Sham problem (Ξ is a diagonal matrix)

$$HC = SC\Xi$$

Electron density in the real space

$$\rho(x) = \Psi(x) f(\Xi - \mu) \Psi^{\mathrm{T}}(x) = \Phi(x) C f(\Xi - \mu) C^T \Phi^T(x)$$

Pole expansion for a diagonal matrix

$$f(\Xi - \mu) \approx \sum_{i=1}^{Q} \frac{\omega_i}{\Xi - z_i I}$$

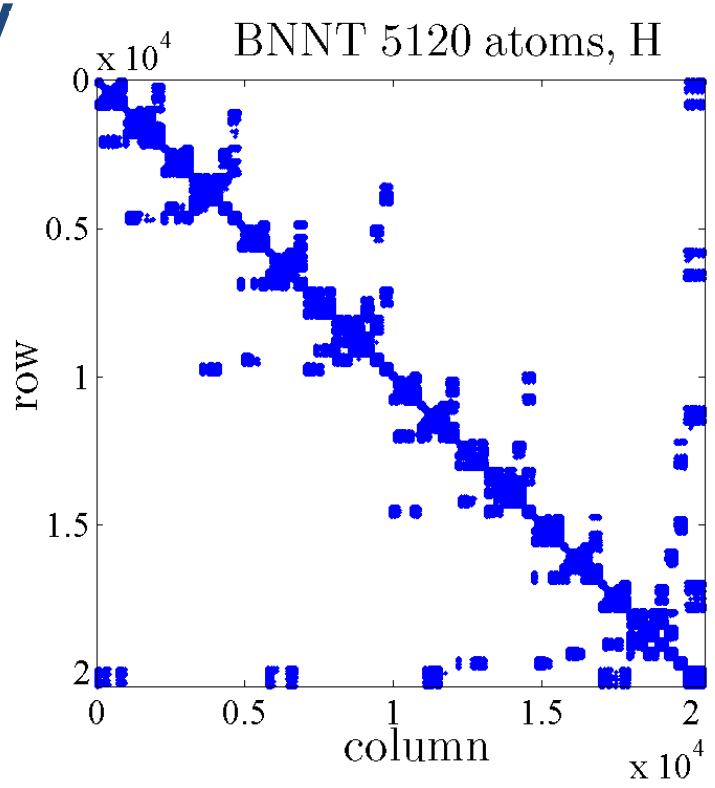
$$\rho(x) = \Phi(x) \sum_{i=1}^{Q} \frac{\omega_i}{\mathrm{C}^{-T} \Xi \mathrm{C}^{-1} - z_i C^{-T} C^{-1}} \Phi^{\mathrm{T}}(x)$$

$$= \Phi(x) \sum_{i=1}^{Q} \frac{\omega_i}{H - z_i S} \Phi^{\mathrm{T}}(x)$$

$$= \sum_{i,j} \varphi_i(x) \gamma_{ij} \varphi_j(x)$$

Sparsity is the key

- H and S are sparse



- $\varphi_i(x)\varphi_j(x) \neq 0 \Rightarrow S_{ij} \neq 0 \Rightarrow H_{ij} \neq 0 \Rightarrow \gamma_{ij} \neq 0$
- Selected elements: $\{(i,j)|H_{ij} \neq 0\}$
- Selected inversion for $\frac{\omega_i}{H - z_i S}$

Total energy

$$E_{tot} = Tr[\gamma H] - \frac{1}{2}\iint dxdy \frac{\rho(x)\rho(y)}{|x-y|} + E_{xc}[\rho(x)] - \int dx\, V_{xc}[\rho](x)\rho(x)$$

$$Tr[\gamma H] = \sum_{ij} \gamma_{ij} H_{ji}$$

Selected inversion

Helmholtz free energy

- Finite temperature effect [Mermin, 1965, Alavi et al, 1994]
- Energy and Entropy

$$F_{tot} = Tr[\gamma^{\mathcal{F}} S] + \mu N_e - \frac{1}{2}\iint dxdy \frac{\rho(x)\rho(y)}{|x-y|} + E_{xc}[\rho(x)] - \int dx\, V_{xc}[\rho](x)\rho(x)$$

Free energy density matrix

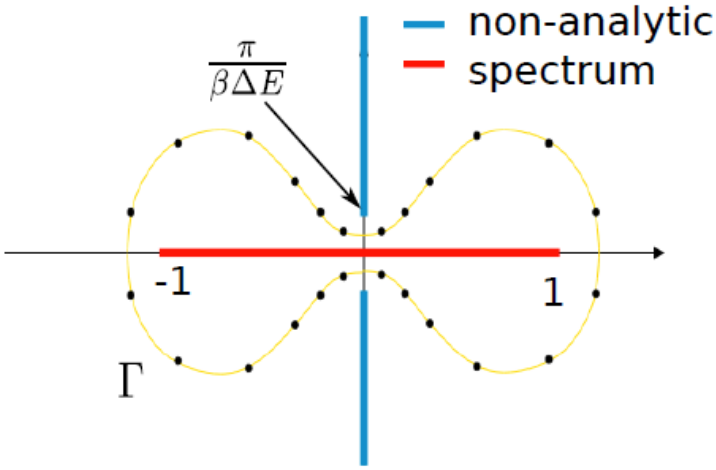
$$\gamma^{\mathcal{F}} = C f^{\mathcal{F}}(\Xi - \mu) C^T$$

$$f^{\mathcal{F}}(x-\mu) = -2\beta^{-1}\log(1 + e^{\beta(\mu - x)})$$

Helmholtz free energy

- $f^{\mathcal{F}}(x-\mu) = -2\beta^{-1}\log(1+e^{\beta(\mu-x)})$
- Same analytic structure as $f(x-\mu)$
- $\gamma^{\mathcal{F}} \approx \sum_{i=1}^{Q} \frac{{\omega_i}^{\mathcal{F}}}{H - z_i S}$

- Pole expansion with the same shift but different weight
- The same selected elements of $(H - z_i S)^{-1}$



Force

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

Hellmann-Feynman force

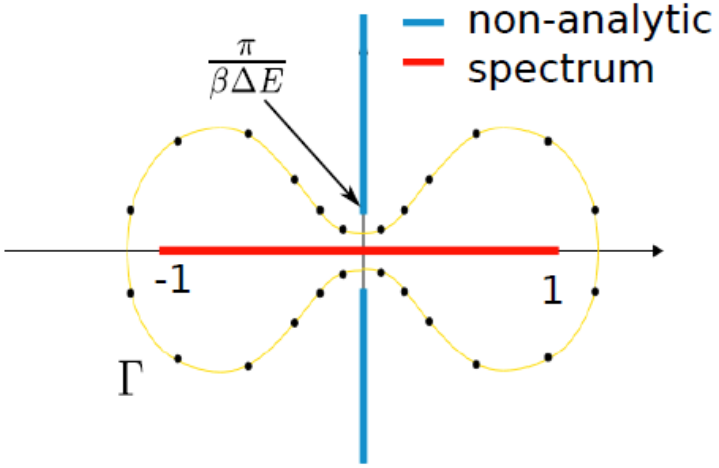
Pulay force

- Energy density matrix

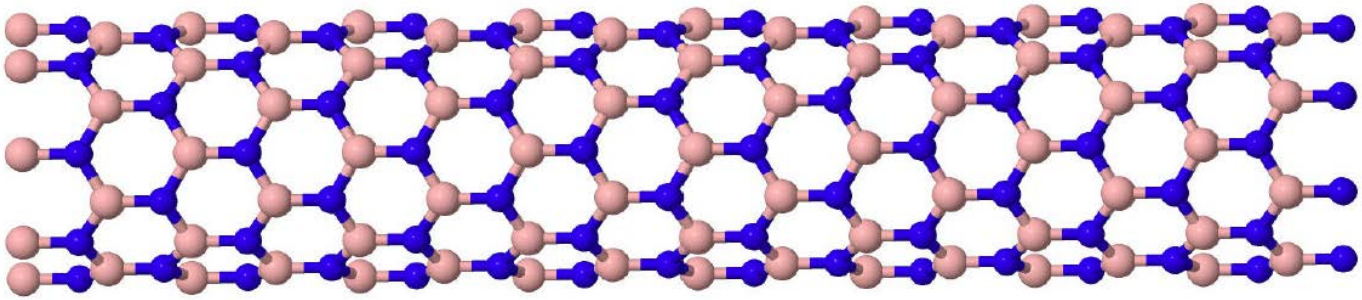
$$\gamma^E = C f^E(\Xi - \mu) C^T$$

$$f^E(x-\mu) = x f(x-\mu)$$

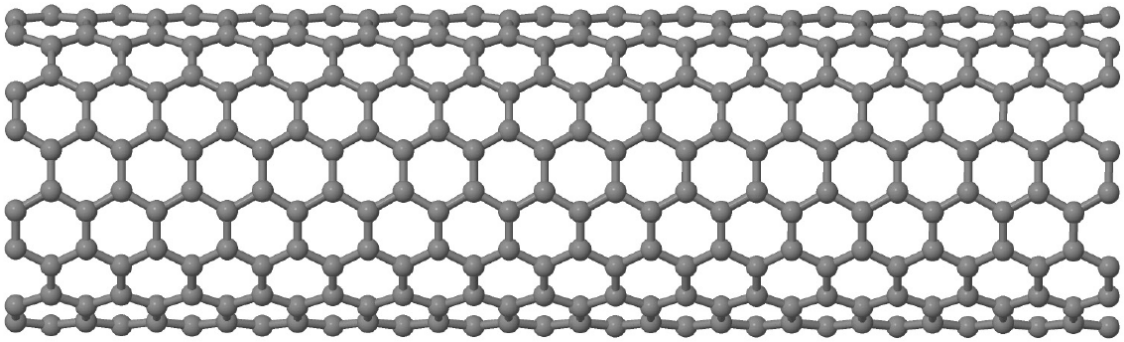
- Pole expansion with the same shift but different weight
- The same selected elements of $(H - z_i S)^{-1}$



Numerical examples with atomic orbitals



Boron Nitride Nanotube



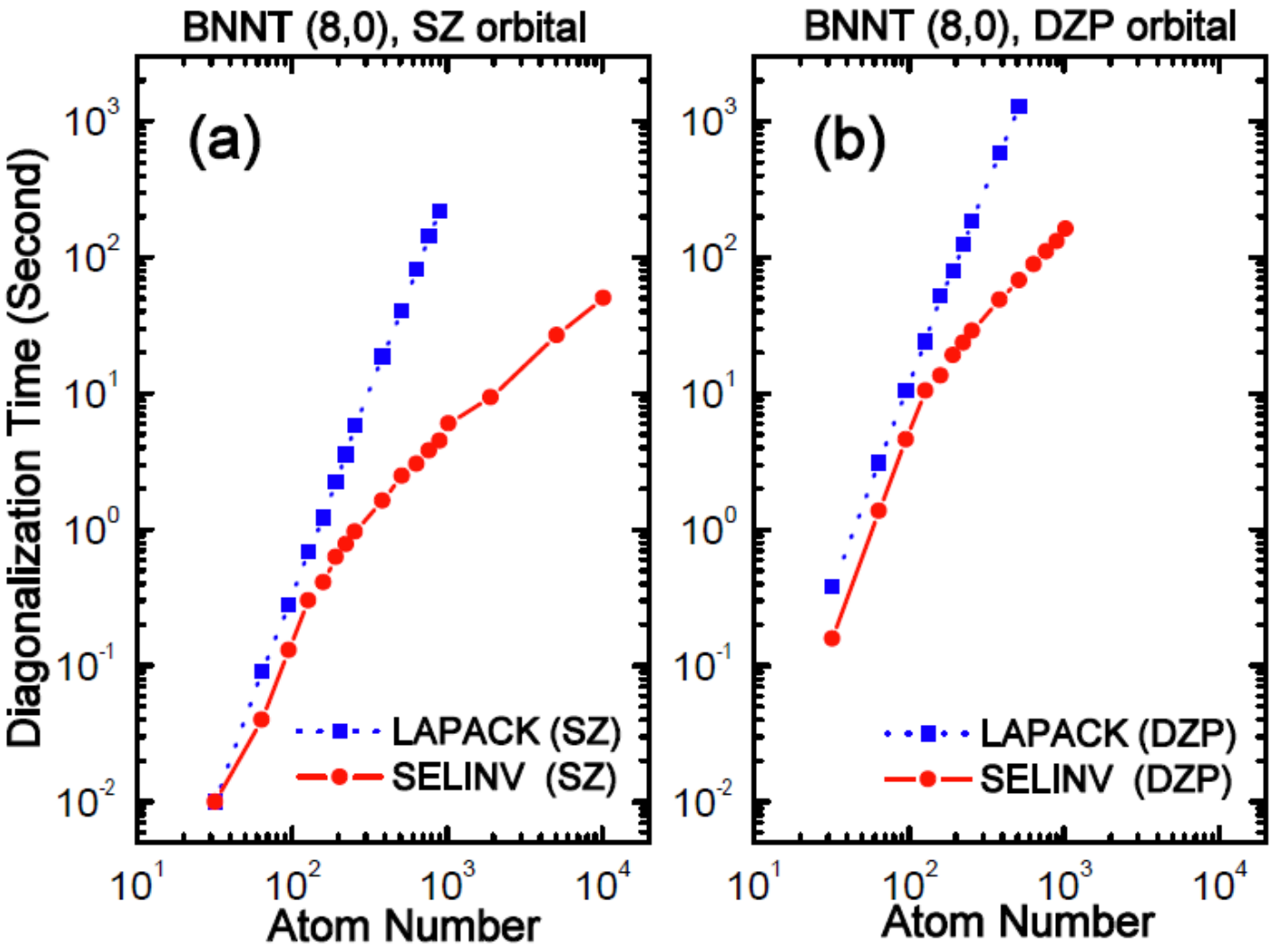
Carbon Nanotube

Atomic orbitals by [Chen et al, 2010, 2011]

Accuracy of the pole expansion

# Poles	$E_{\text{PEXSI}} - E_{\text{ref}}$ (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 the selected inversion

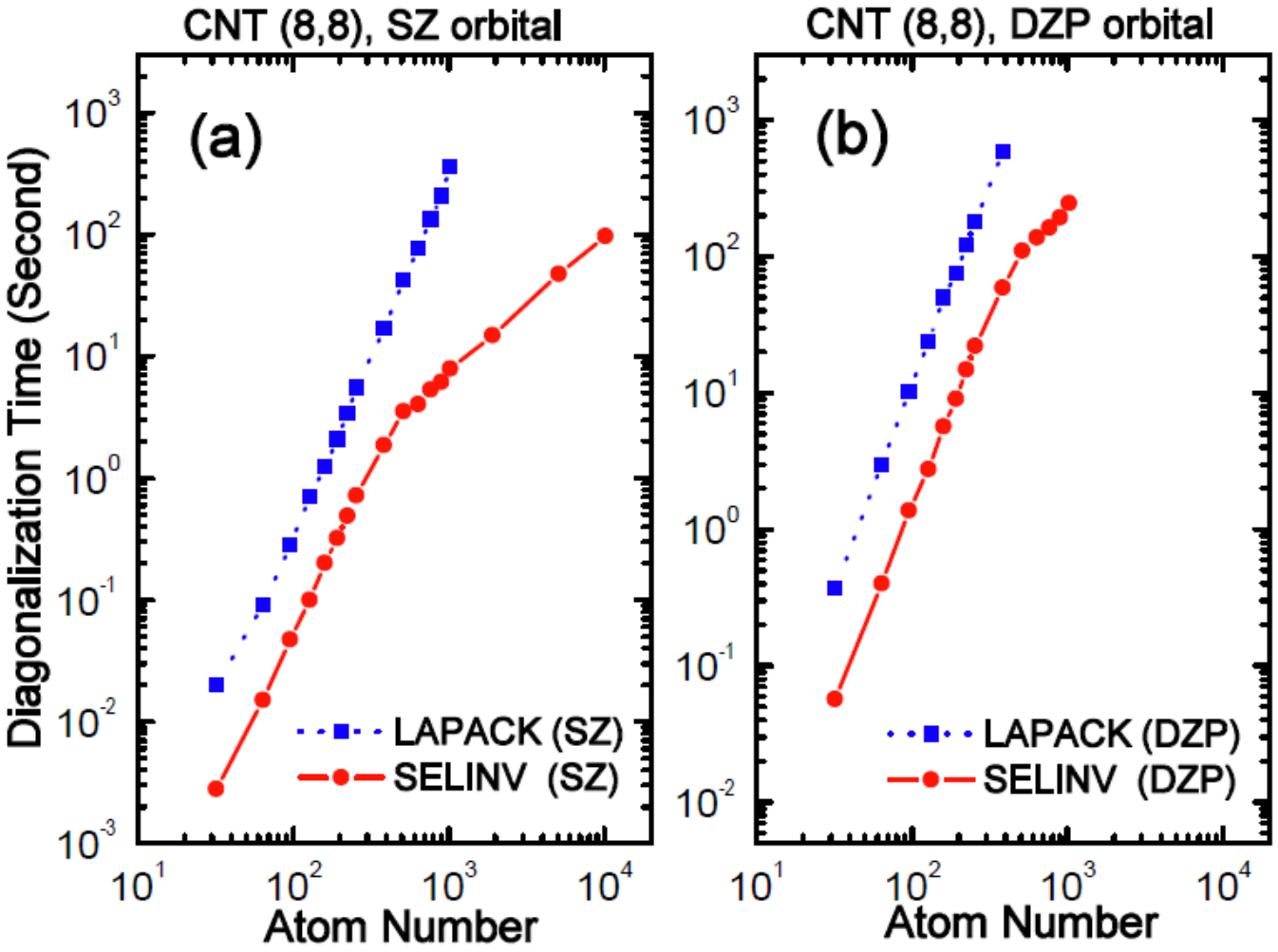


Boron Nitride

SZ: single-zeta (4 basis per atom)

DZP: Double-zeta with polarization (13 basis per atom)

Efficiency of the selected inversion



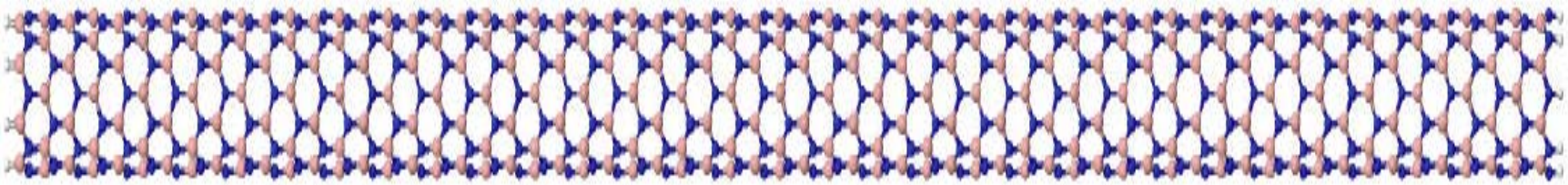
Carbon nanotube

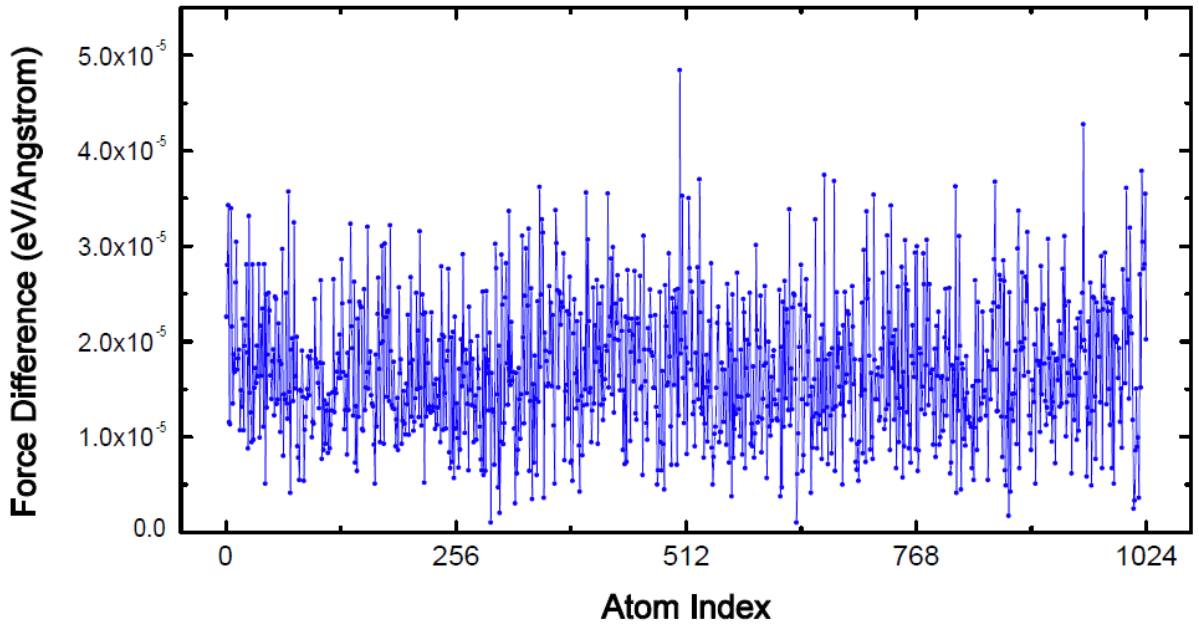
SZ: single-zeta (4 basis per atom)

DZP: Double-zeta with polarization (13 basis per atom)

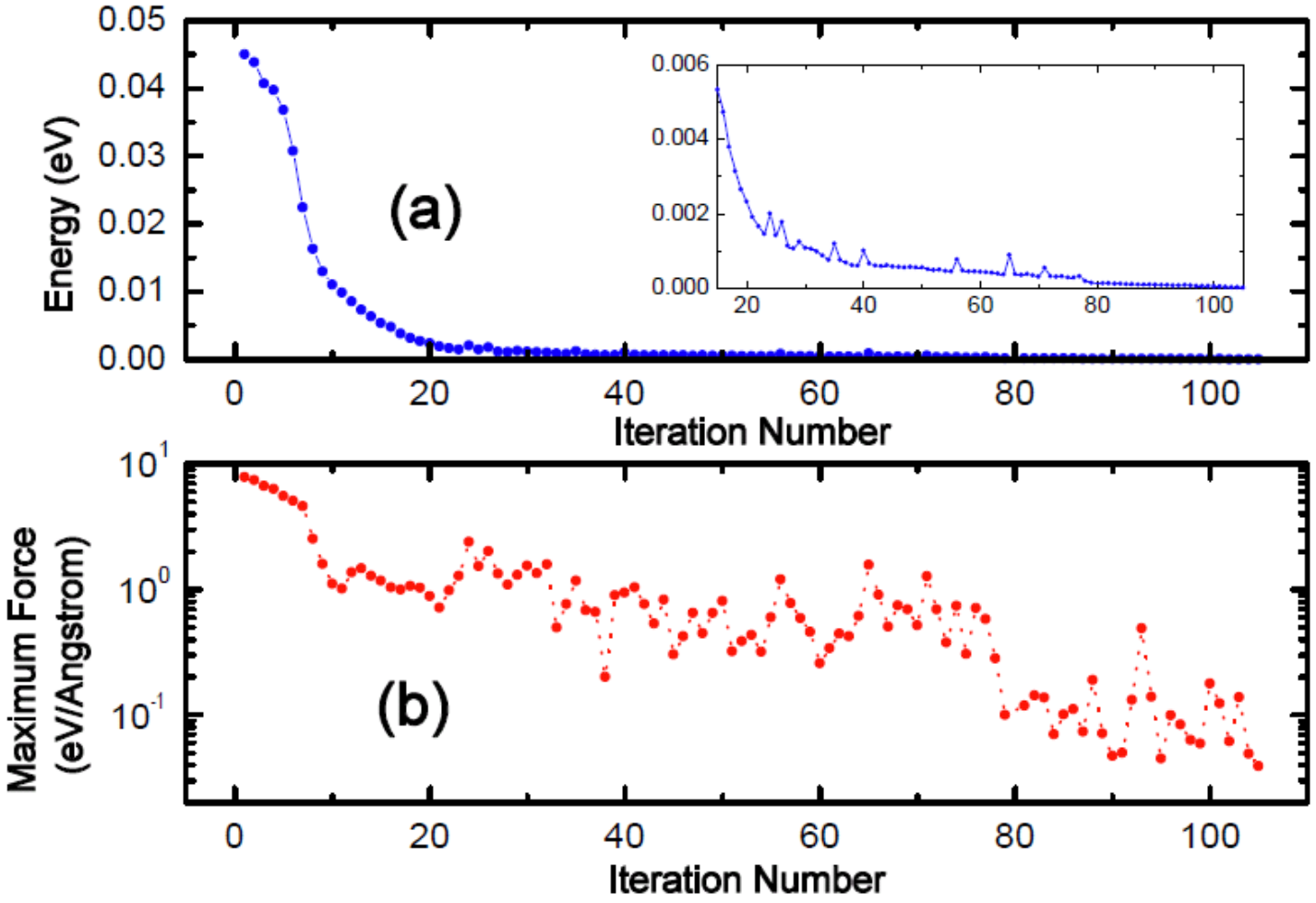
Geometry optimization: BNNT

Truncated BNNT. 504 B atoms, 504 N atoms, 16 H atoms

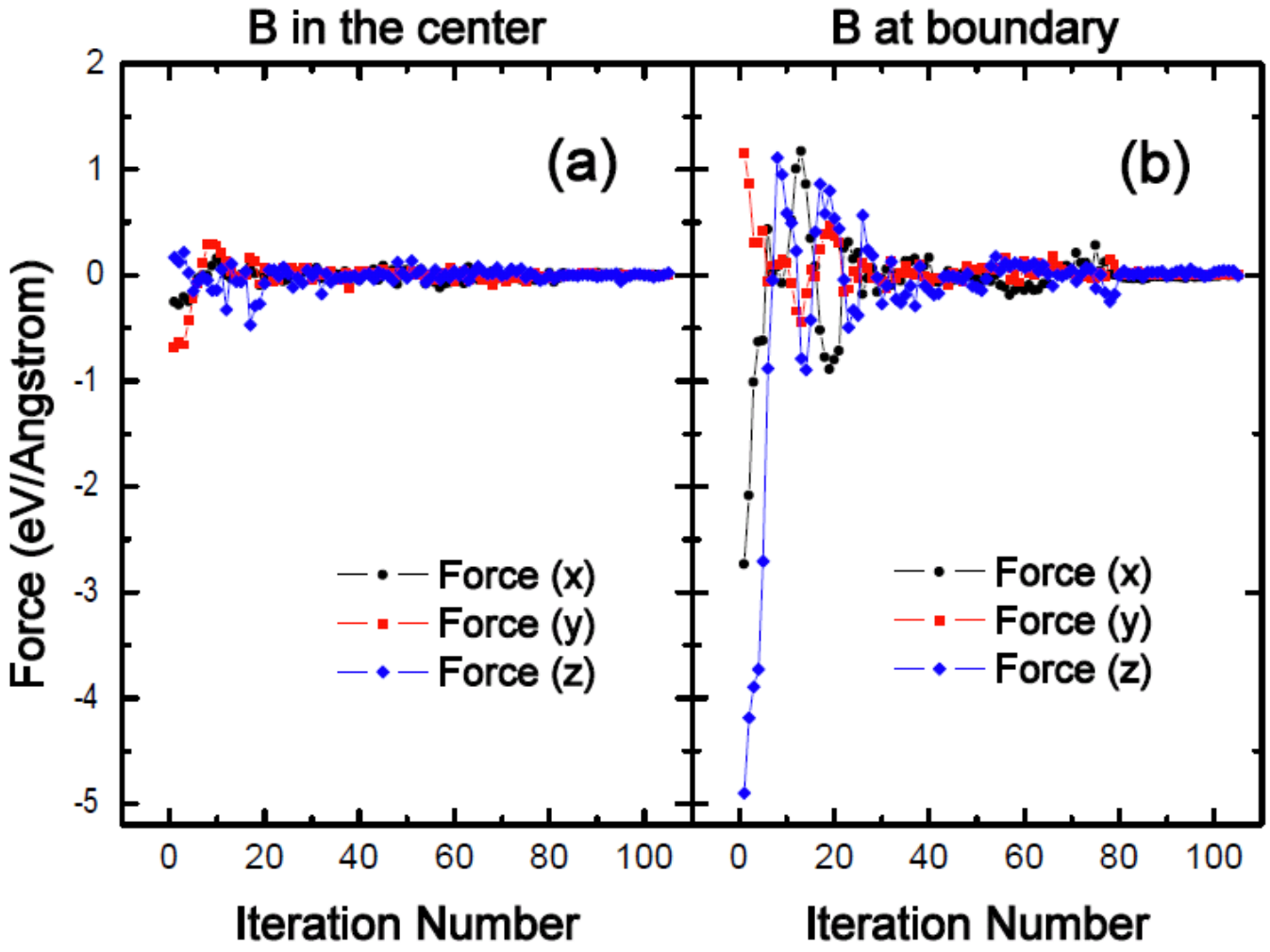




Geometry optimization: BNNT



Geometry optimization: BNNT



Conclusion

- Pole Expansion and Selected Inversion (PEXSI) method for Kohn-Sham density functional theory
- Accurate calculation of density, total energy, free energy and force.
- Selected inversion: $O(N)$ for quasi-1D system, $O(N^{1.5})$ for quasi-2D system, and $O(N^2)$ for 3D bulk systems.
- Quantum chemistry basis set (Gaussian orbitals, atomic orbitals etc)
- Black-box: suitable for all codes with quantum chemistry basis set

Acknowledgment

Luis Alvarez fellowship supported by LBNL and DOE. SciDAC.

Collaborators:

- Prof. Roberto Car, Princeton University
- Dr. Mohan Chen, University of Science and Technology in China
- Prof. Weinan E, Princeton University and Peking University
- Prof. Lixin He, University of Science and Technology in China
- Prof. Juan Meza, UC Merced
- Prof. Jianfeng Lu, Duke University
- Dr. Chao Yang, Lawrence Berkeley National Lab
- Prof. Lexing Ying, University of Texas at Austin

Thank you for your attention!