Compressed representation of Kohn-Sham orbitals via selected columns of the density matrix

Lin Lin

Department of Mathematics, UC Berkeley;
Computational Research Division, LBNL

Joint work with Anil Damle and Lexing Ying

Fundamental Aspects of DFT,
Oslo, January, 2015

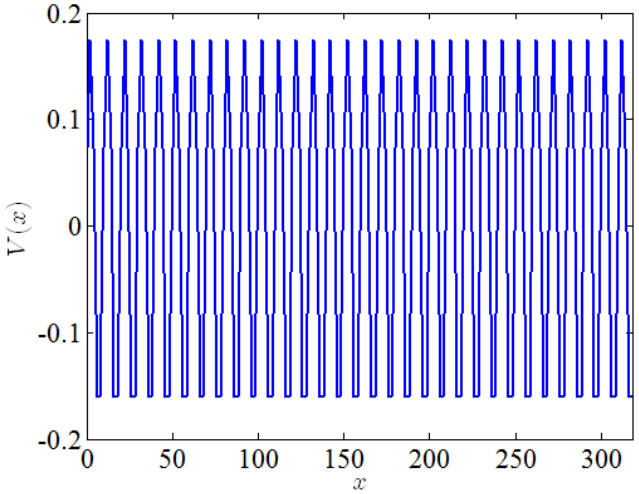
Outline

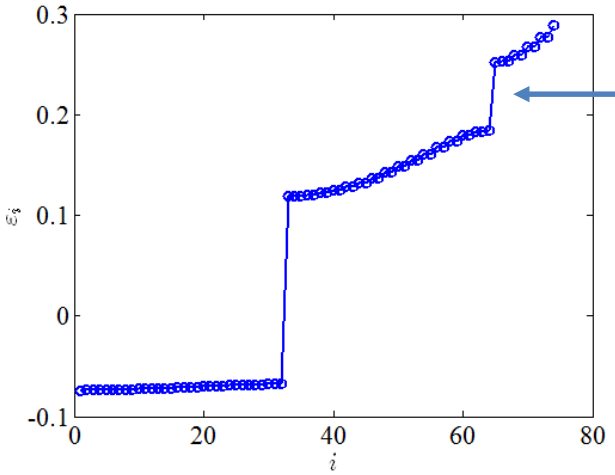
- Introduction
- Previous methods
- Selected columns of density matrix
- Numerical examples
- Conclusion

Example of compressed representation

- $H = -\frac{1}{2}\Delta + V(x)$. Periodic potential. Insulating system

$$H\psi_i = \varepsilon_i \psi_i$$

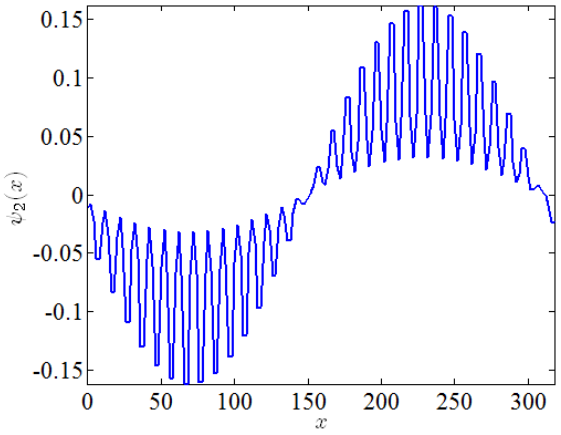


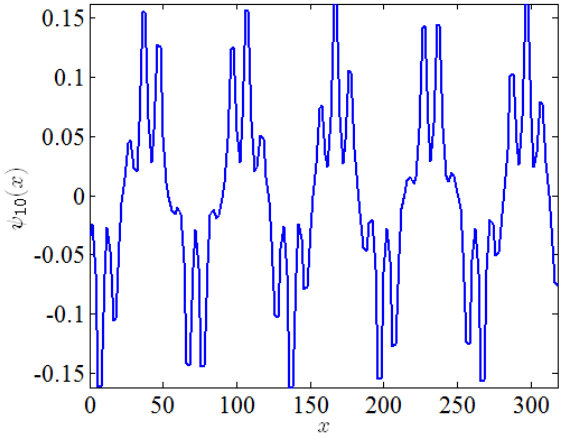


HOMO-LUMO gap

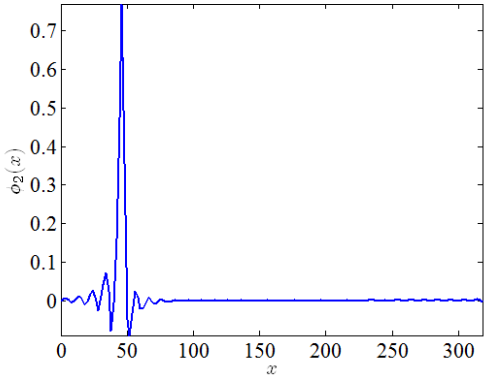
Example of compressed representation

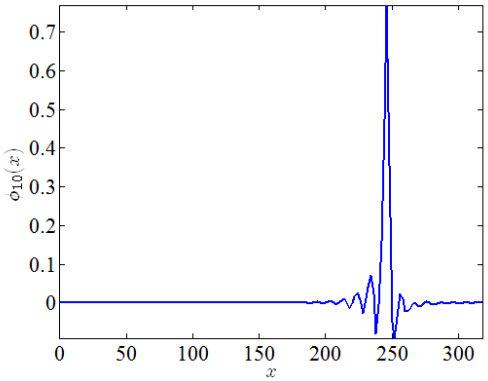
- Delocalized eigenfunction





- Compressed (localized) representation. Unitary transformation of eigenfunctions. Equivalent information.





General concepts of localization

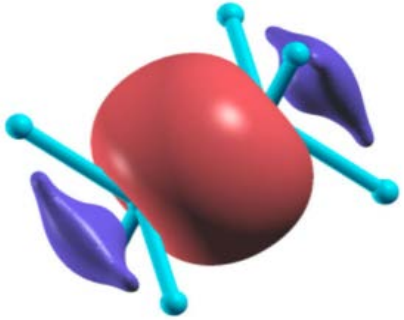
- Eigenfunctions (Molecular orbitals) are in general delocalized across the whole molecule / solids.

- Chemical properties are often localized. Why?

- In quantum physics, a class of sparse representation are given by the Wannier functions [Wannier, 1937], [Marzari et al., Rev. Mod. Phys. 2012] [Kohn, Phys. Rev. Lett. 1996]. In quantum chemistry this is often referred to as the localized molecular orbitals, or Boys orbitals [Boys 1960]

Wannier functions

- Maximally localized Wannier function (MLWF) [Marzari-Vanderbilt, Phys. Rev. B 1997]. Examples below from [Marzari et al. Rev. Mod. Phys. 2012]



Silicon

Graphene

- Reason for the existence of MLWF for insulating systems [Kohn, Phys. Rev. 1959] [Nenciu, Comm. Math. Phys. 1983] Many others

Application of Wannier functions

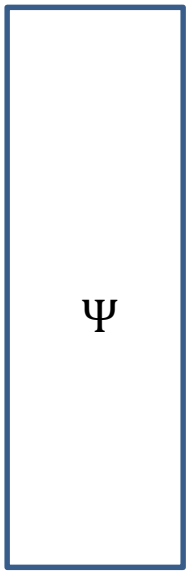
- Analysis of chemical bonding
- Band-structure interpolation
- Basis functions for DFT calculations (representing occupied orbitals ψ_i)
- Basis functions for excited state calculations (representing Hadamard product of orbitals $\psi_i \odot \psi_j$)
- Strongly correlated systems (DFT+U)
- Phonon calculations
- etc

Outline

- Introduction
- Previous methods
- Selected columns of density matrix
- Numerical examples
- Conclusion

(Somewhat non-traditional) setup of the problem

- Ψ: Kohn-Sham orbitals
  dense, unitary matrix of size $m \times n$ $(m \gg n)$



- ε-sparse representation

$$\|\Psi - \Phi C\|_F \leq \varepsilon$$

  - Each column of Φ is sparse.
  - C is an $n \times n$ unitary matrix.

- Goal: find Φ

Maximally Localized Wannier functions

- [Vanderbilt-Marzari, PRB 1997]

- Geometric intuition. Iterative minimization of the sum of second moments of all localized functions.

$$\min_{\substack{\Psi=\Phi C,\\ C^*C=I}} \Omega[\Phi]$$

$$\Omega[\Phi] = \sum_{j=1}^{n} \int |\phi_j(x)|^2 x^2 \, dx - \left(\int \phi_j(x) x dx \right)^2$$

L^1 minimization

$$\min_{\Phi,C} \frac{1}{\mu}\|\Phi\|_1 + \|\Psi - \Phi C\|_F^2$$
$$s.t.\ C^*C = I$$

- L^1 minimization promotes sparsity, and each function can be compactly supported.
- Similar approach taken by [Ozlins-Lai-Caflisch-Osher, PNAS 2013] (see below)

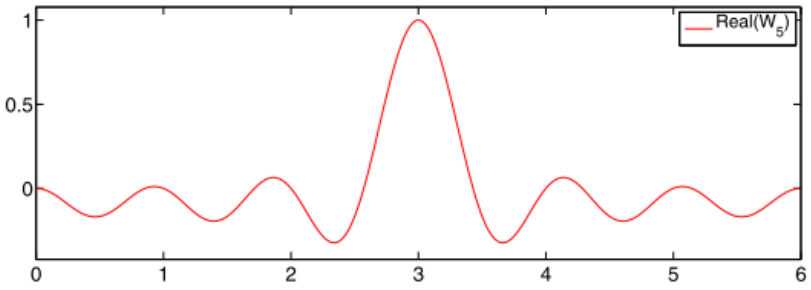


Fig. 1. A quasi-localized Wannier function for 1D Laplace operator.

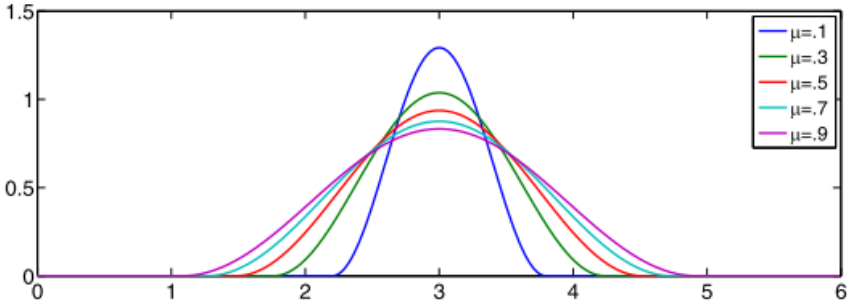


Fig. 2. Theoretical ψ_1 in the 1D free-electron model (Eq. **6**) for different values of μ.

Density matrix

- Ψ is unitary, then

$$P = \Psi\Psi^*$$

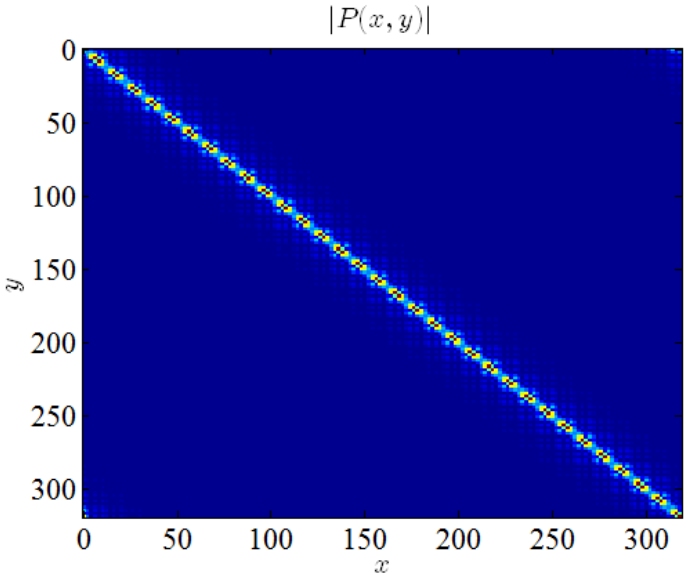
is a projection operator.

$$P = \Psi\Psi^* = \Phi C C^* \Phi^* = \Phi\Phi^*$$

is sparse.

- Decay properties of spectral projector [Benzi-Boito-Razouk, SIAM Rev. 2013]

- Can one construct sparse representation directly from the density matrix?



Density matrix could be a more powerful tool

- [Panati, AHP 2007], [Brouder et al, Phys. Rev. Lett. 2007] points out that the exponential decay of Wannier functions cannot hold for Chern insulators (insulators with non-vanishing Chern invariant)

$$\mathbf{C} = \frac{i}{2\pi} \int_{\mathrm{BZ}} d\mathbf{k} \sum_{n}^{\mathrm{occ}} \langle \partial_{\mathbf{k}} u_{n\mathbf{k}} | \times | \partial_{\mathbf{k}} u_{n\mathbf{k}} \rangle$$

- The density matrix may still exhibit exponential decay even for Chern insulators [Thonhauser, Vanderbilt PRB 2006].

Algorithm: Selected columns of the density matrix (SCDM)

Pseudocode (MATLAB. Psi: matrix of size m*n, m>>n)

```
[Q,R,perm] = qr(Psi', 0);
Pc = Psi * Psi(perm(1:n),:)';
```

Rank revealing QR

GEMM

Pc are columns of the density matrix and are localized orbitals!

The following three lines give orthogonal, localized obitals.

```
Pcc = Pc(perm(1:n),:);
L = chol(Pcc, 'lower');
Phi = Pc / L';
```

Very easy to code and to parallelize!

Outline

- Introduction
- Previous methods
- Selected columns of density matrix
- Numerical examples
- Conclusion

Columns of the density matrix

- P is a rank-n matrix of size $m \times m$. $m \gg n$ so the matrix is low-rank.
- In principle, any n linearly independent columns would be sparse and span the same subspace as Ψ.
- Arbitrary choice of columns can be poorly conditioned.

Interpolative Decomposition

- Treat columns of density matrix as an over-complete dictionary for Ψ.

- Find a set of columns $\mathcal{C}$ ($|\mathcal{C}| = n$), permutation Π, transformation matrix T, so that

$$P\Pi = P_{:,\mathcal{C}}[I\ T]$$

  $\|T\|_F$ is small.

- Note that $\mathrm{P}\Pi_{:,1:n}$ corresponds to the columns $P_{:,\mathcal{C}}$. We only need the permutation matrix Π to get the columns $\mathcal{C}$.

Rank-revealing QR

Theorem [Gu-Eisenstat, 1996] Let A be an $m \times n$ matrix, $l = \min(m, n)$,and k is an integer such that $1 \le k \le l$. Then there is a factorization

$$A\Pi = QR$$

where Π is an $n \times n$ permutation matrix, Q is an $m \times l$ matrix with orthonormal columns, and R is an $l \times n$ upper triangular matrix. Furthermore, if

$$Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}, \qquad R = \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix}$$

Q_{11}, R_{11} are of size $k \times k$, then

$$\sigma_k(R_{11}) \ge \sigma_k(A) \frac{1}{\sqrt{1 + k(n-k)}},$$

$$\sigma_1(R_{22}) \le \sigma_{k+1}(A)\sqrt{1 + k(n-k)}$$

$$\left\| R_{11}^{-1} R_{12} \right\|_F \le \sqrt{k(n-k)}$$

Here $\sigma_k(A)$ is the k-th largest singular value of the matrix A.

Rank-revealing QR

- Apply RRQR to the matrix P (size $m \times m$). Since P is exactly of rank n,

$$P\Pi = QR$$

Here Q is a matrix of size $m \times n$, and

$$R = [R_1 \; R_2]$$

is a matrix of size $n \times m$.

$$P\Pi = QR_1[I \;\; R_1^{-1}R_2] \equiv P_{:,\mathcal{C}}[I \; T]$$

$$\|T\|_F = \left\|R_1^{-1}R_2\right\|_F \leq \sqrt{n(m-n)}$$

- Cost: $O(m^2 n)$, which is too costly.

Rank-revealing QR

- Consider

$$\Psi^* \Pi = \tilde{Q} R$$

then

$$P\Pi = \Psi\Psi^*\Pi = \left(\Psi\tilde{Q}\right)\Pi$$

Since

$$\left(\Psi\tilde{Q}\right)^*\left(\Psi\tilde{Q}\right) = \tilde{Q}^*\tilde{Q} = I$$

we find the ID for P via the RRQR decomposition for Ψ^*

- Cost: $O(mn^2)$, affordable since Ψ is a dense matrix to start with.

Recovering the density matrix

$$P_{:,\mathcal{C}} = \Psi\left(\Psi_{\mathcal{C},:}\right)^*$$

$$P = P_{:,\mathcal{C}} D P_{:,\mathcal{C}}^*$$

Find the matrix D

$$P_{\mathcal{C},\mathcal{C}} D P_{\mathcal{C},\mathcal{C}}^* = P_{\mathcal{C},\mathcal{C}}$$

Here $P_{\mathcal{C},\mathcal{C}}$ is invertible (otherwise we have rank less than n).

Also $P_{\mathcal{C},\mathcal{C}}^* = P_{\mathcal{C},\mathcal{C}}$

$$D = P_{\mathcal{C},\mathcal{C}}^{-1}$$

Recovering the unitary matrix

- $P_{\mathcal{C},\mathcal{C}} = LL^*$ Cholesky factorization
- $\Phi = P_{:,\mathcal{C}}(L^*)^{-1}$

- $P = \Phi\Phi^* = \Psi\Psi^*$, then Φ is unitary (P is a projection operator).

- $\Psi = \Phi C,\ C = \Phi^*\Psi$ is unitary.

Algorithm: Selected columns of the density matrix (SCDM)

Pseudocode (MATLAB. Psi: matrix of size m*n, m>>n)

```
[Q,R,perm] = qr(Psi', 0);
Pc = Psi * Psi(perm(1:n),:)';
```

Rank revealing QR

GEMM

Pc are columns of the density matrix and are localized orbitals!

The following three lines give orthogonal, localized obitals.

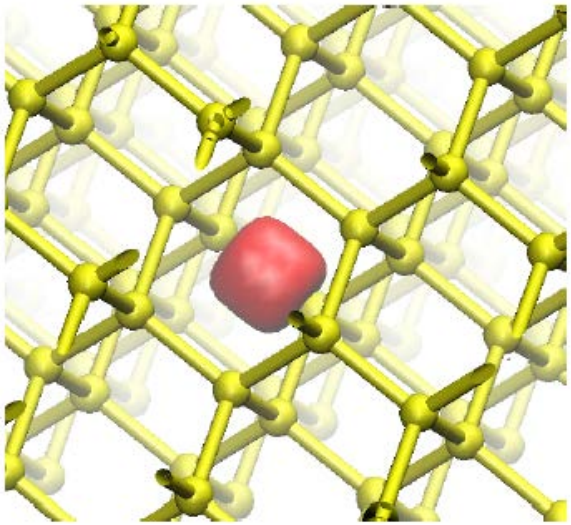
```
Pcc = Pc(perm(1:n),:);
L = chol(Pcc, 'lower');
Phi = Pc / L';
```

Very easy to code and to parallelize!

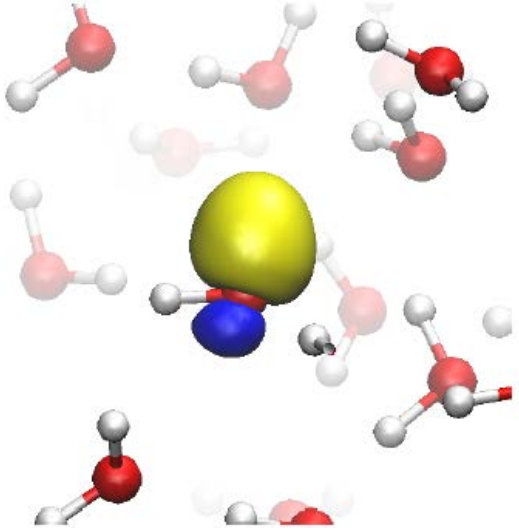
Outline

- Introduction
- Previous methods
- Selected columns of density matrix
- Numerical examples
- Conclusion

Numerical examples



Silicon

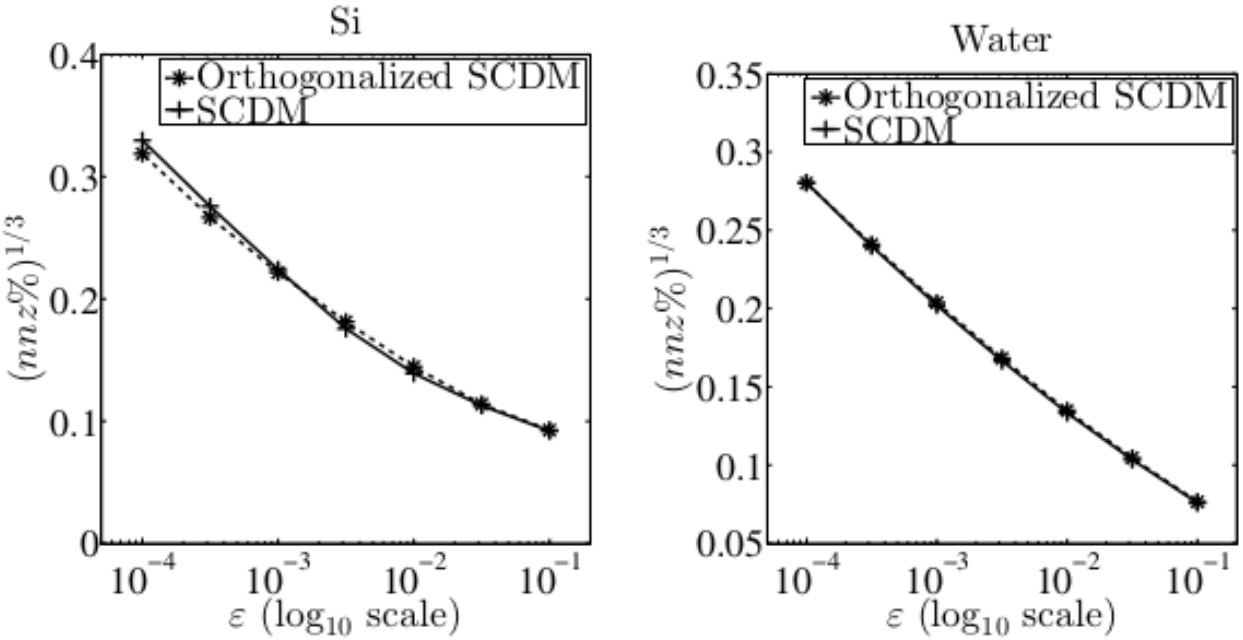


Water

SCDM (orthonormalized) computed using Kohn-Sham density functional theory. Delocalized wavefunctions solved from Quantum ESPRESSO.

$P_{\mathcal{C},\mathcal{C}}$ is well-conditioned. Condition number is 3.18 for Si and 2.83 for Water.

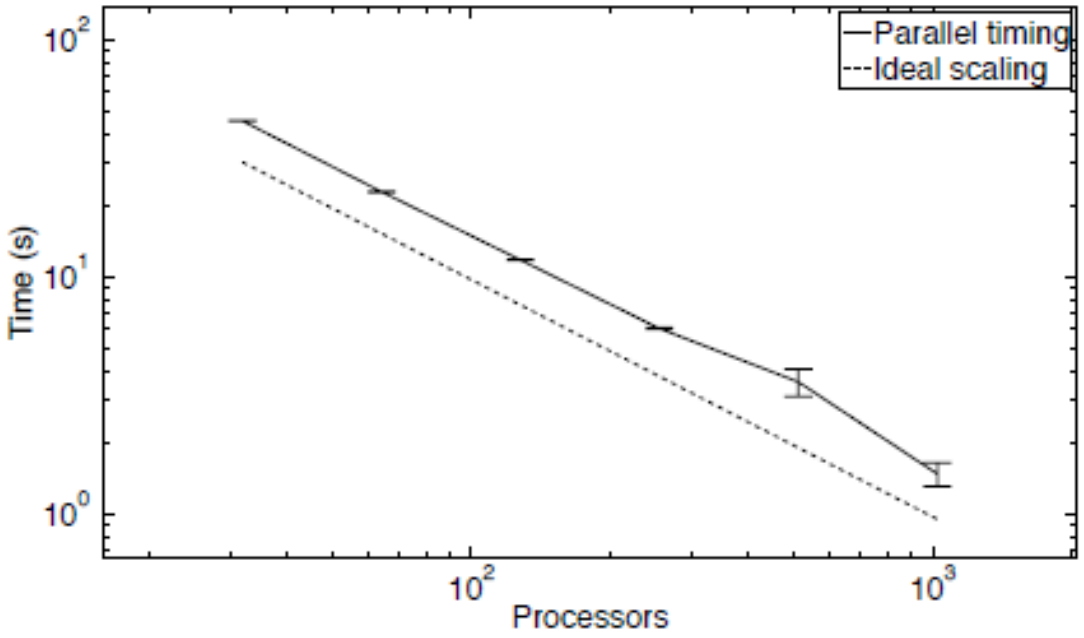
Exponential decay



$nnz\%$: percentage of non-zero entries
ε: relative truncation error

Exponential decay: $\exp(-Cd) \sim \varepsilon,\ d \sim (nnz\%)^{1/3}$

Parallelization

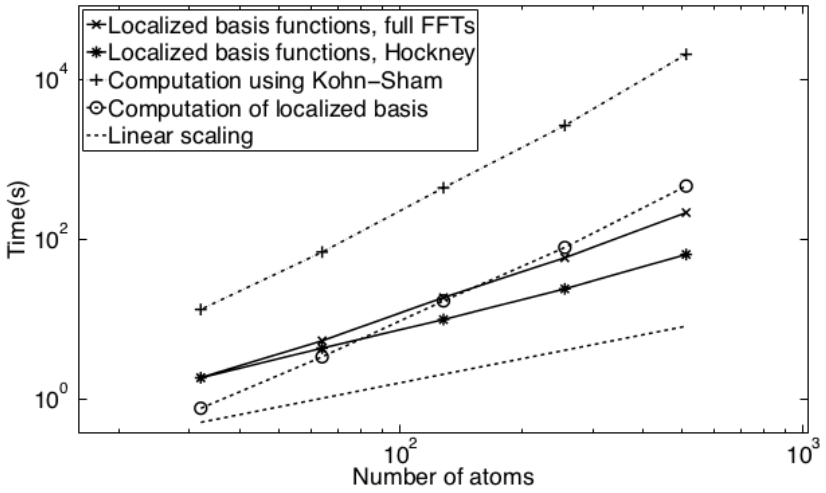


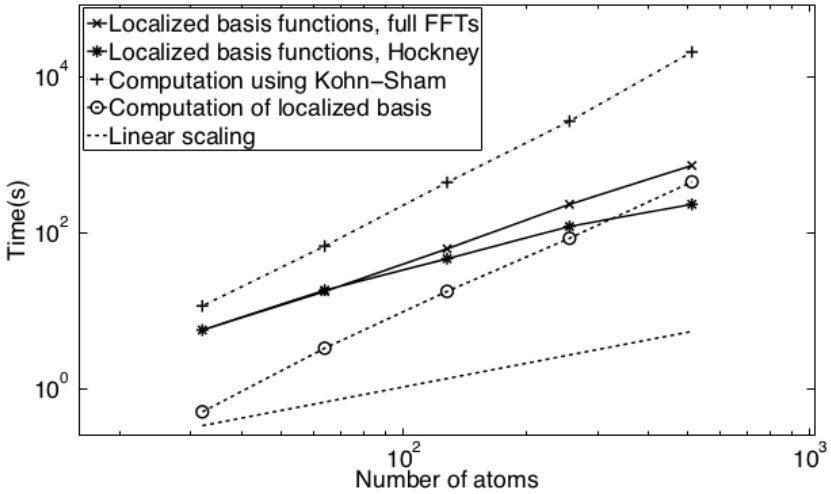
Time for computing the orthogonalized SCDM for a Si system (dimension: 777600*1024)

Other QR strategies available [Demmel-Grigori-Gu-Xiang, 2013]

Linear scaling Hartree-Fock calculation

$$E_x = -\frac{1}{2}\sum_{i,j=1}^{N_e}\iint \frac{\phi_i(x)\phi_j(x)\phi_j(y)\phi_i(y)}{|x-y|}\,\mathrm{d}x\,\mathrm{d}y.$$





5% error in energy calculation
300 times speed-up

0.4% error in energy calculation
90 times speed-up

Conclusion

- Selected columns of the density matrix (SCDM) is a simple, robust and deterministic method to find localized orbitals. Alternative method to the Maximally Localized Wannier functions.

- Straightforward implementation when Kohn-Sham orbitals are represented on real space grid. It can be easily used in any electronic structure code

- Very easy to parallelize and highly parallelizable. Can accelerate Hartree-Fock and hybrid functional calculations.

Conclusion

[Damle-LL-Ying, Compressed representation of Kohn-Sham orbitals via selected columns of the density matrix, J. Chem. Theory Comput. 11, 1463, 2015]

Thank you for your attention!

Support from DOE Alvarez fellowship, SciDAC, CAMERA funding