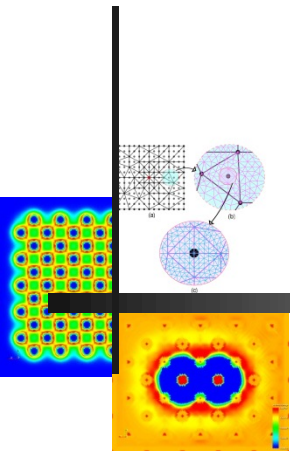


Fast, Accurate and Large-scale Ab-initio Calculations for Materials Simulations



Vikram Gavini

*Department of Mechanical Engineering
Department of Materials Science & Engineering
Program in Applied Physics
University of Michigan, Ann Arbor*

*Collaborators: Sambit Das (U. Mich.), Bikash Kanungo (U. Mich.),
Phani Motamarri (IISc)*

Funding: DoE-BES, TRI, NERSC, OLCF



Quantum Mechanics

- Schrödinger equation - $H\psi = E\psi$

$$H = -\frac{1}{2} \sum_{i=1}^N \nabla_i^2 - \frac{1}{2} \sum_{A=1}^M \frac{1}{M_A} \nabla_A^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|} \\ + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{A=1}^M \sum_{B=1, B>A}^M \frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|}$$

$$\psi = \psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N, \mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_M)$$

- Born-Oppenheimer approximation - Classical treatment of atomic nuclei

$$\psi = \psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$$

- Computational complexity - $\psi \in \mathbf{R}^{3N}$!!
-



Density-functional theory – Kohn-Sham approach

- Ground-state energy is a functional of electron-density !! (Kohn & Sham, 1964-65)

$$\langle \psi | H | \psi \rangle \geq E_0 \quad (\text{Variational statement})$$

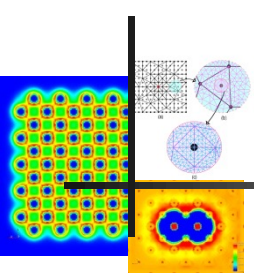
$$\begin{aligned} E_0 &= \min_{\psi} \langle \psi | T + \frac{1}{2} \sum_i' \sum_j' \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_i V_{ext}(\mathbf{r}_i) | \psi \rangle + E_{zz} \\ &= \min_{\psi} \langle \psi | T + \frac{1}{2} \sum_i' \sum_j' \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} | \psi \rangle + \int \rho(\mathbf{r}) V_{ext}(\mathbf{r}) d\mathbf{r} + E_{zz} \\ &= \min_{\rho} \left\{ \underbrace{\left(\min_{\psi \rightarrow \rho} \langle \psi | T + \frac{1}{2} \sum_i' \sum_j' \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} | \psi \rangle \right)}_{F(\rho)} + \int \rho(\mathbf{r}) V_{ext}(\mathbf{r}) d\mathbf{r} \right\} + E_{zz} \end{aligned}$$

$$F(\rho) = T_s(\rho) + E_H(\rho) + E_{xc}(\rho) \longrightarrow \begin{array}{l} \text{Exchange-correlation} \\ \text{functional: Model using} \\ \text{LDA, GGA} \end{array}$$

Kinetic energy of non-interacting electrons:
Computed from wave-functions of the
resulting E-L eqn.



Density Functional Theory



Kohn-Sham eigenvalue problem:

$$\left(-\frac{1}{2}\nabla^2 + V_{\text{eff}}[\rho; \mathbf{R}] \right) \psi_i = \epsilon_i \psi_i$$

↓

Self Consistent Field (SCF) iteration
(Kohn-Sham map)

$$V_{\text{eff}}[\rho; \mathbf{R}] = V_{\text{ext}}(\mathbf{R}) + V_H(\rho) + V_{xc}(\rho) \quad \rho(\mathbf{r}) = 2 \sum_i f_i |\psi_i(\mathbf{r})|^2$$

$$T_s(\Psi) = \frac{1}{2} \sum_i f_i \int |\nabla \psi_i(\mathbf{r})|^2 d\mathbf{r} \quad E_0(\Psi) = T_s(\Psi) + E_{xc}(\rho) + E_H(\rho) + E_{\text{ext}}(\rho) + E_{zz}$$

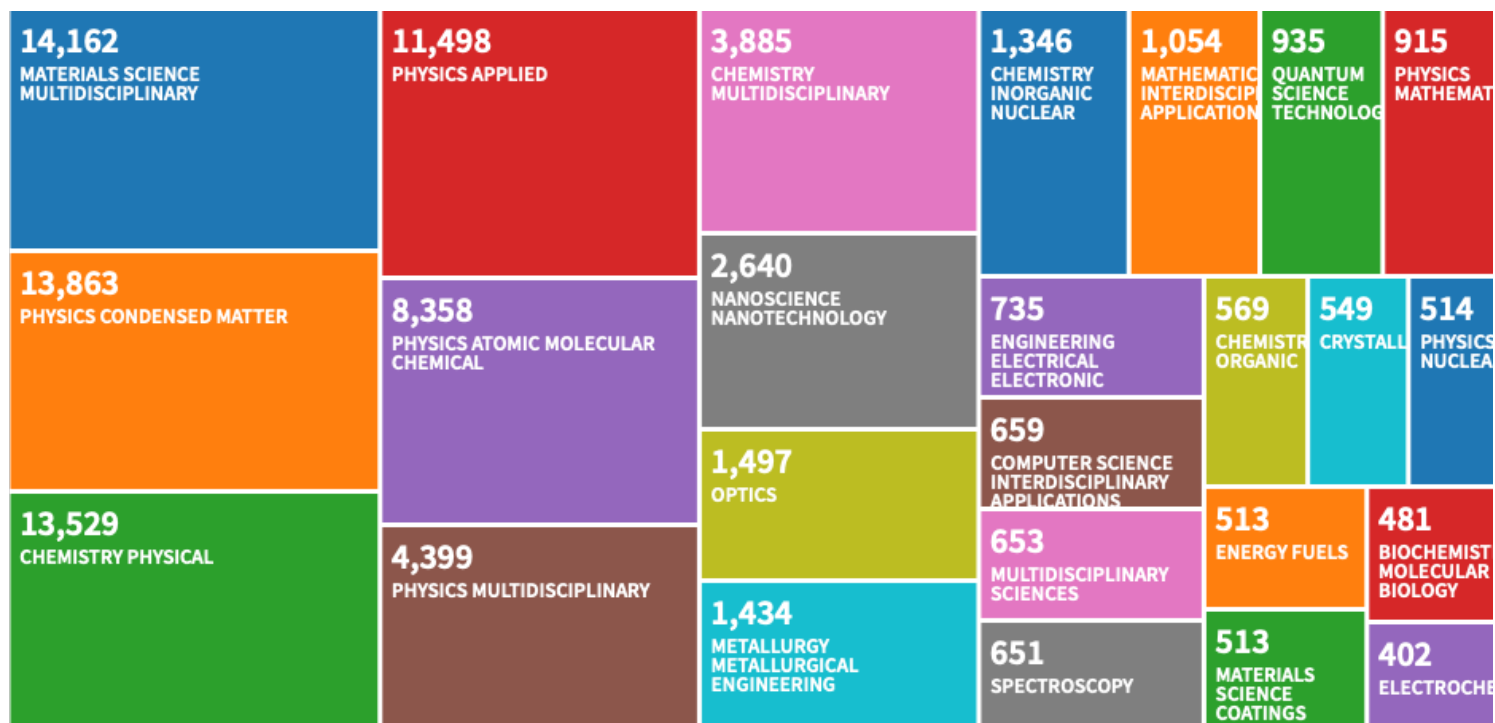
Remarks:

- Ground-state energy; structure -> Range of Material Properties
- The most computational intensive step in each SCF iteration is the solution of the eigenvalue problem
 - ❖ Computational complexity scales as $O(N^3)$



Impact of Density Functional Theory

Citations to seminal work of Walter Kohn (1964,1965)

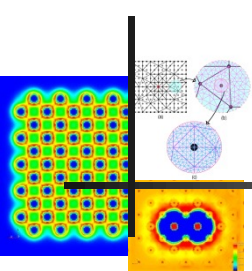


Data compiled from Web of Science

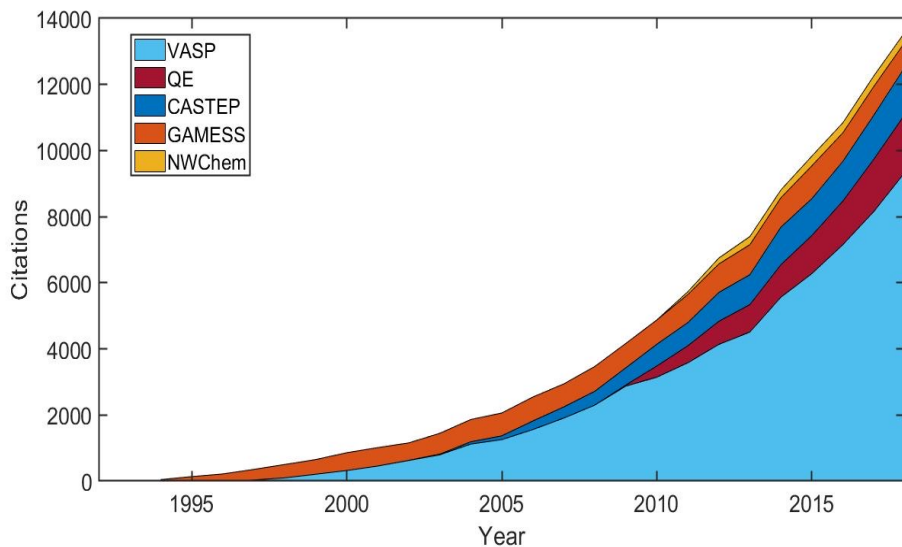
12 of the 100 most-cited papers in scientific literature pertain to DFT!
(Nature **514**, 550 (2014))



DFT codes

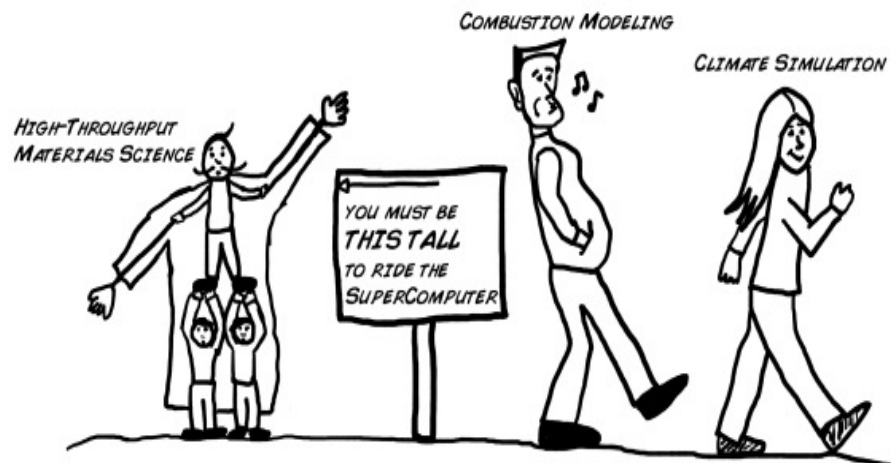


~100 available DFT codes developed since 1980



Data compiled from Web of Science

Relationship to HPC



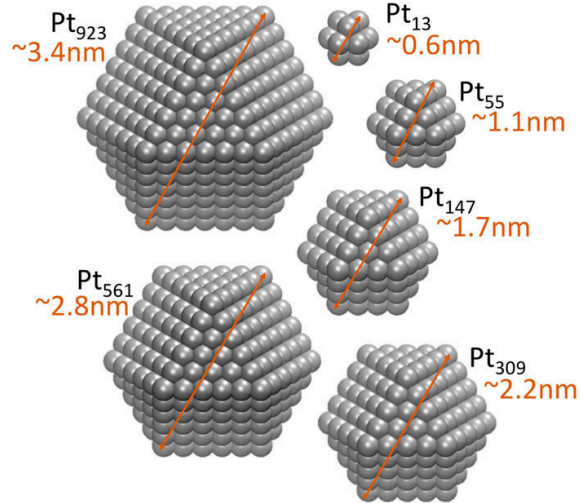
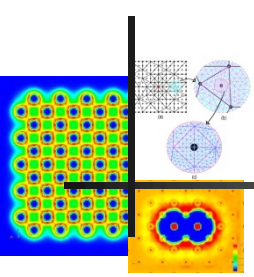
Courtesy: Anubhav Jain

Key Issues

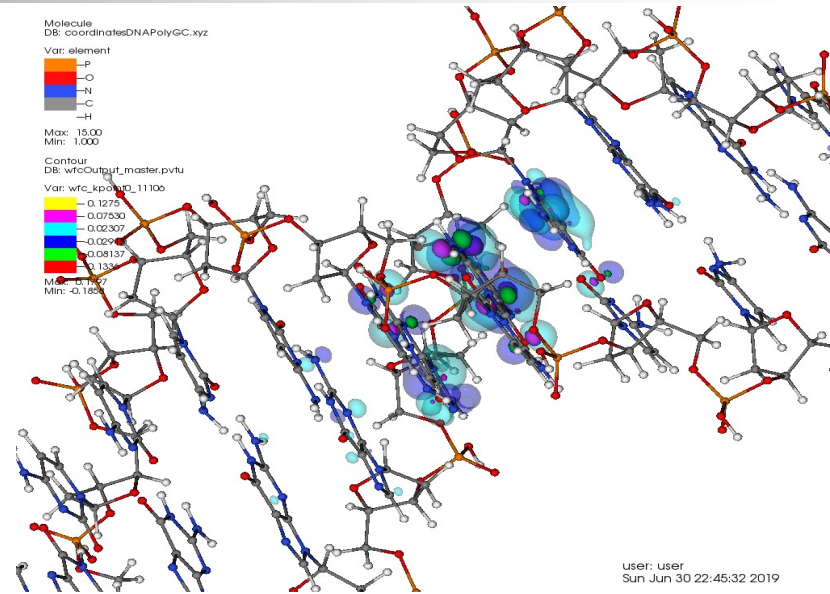
- ❖ Lack of good parallel scalability of existing DFT codes
- ❖ Computational complexity of DFT calculations ($O(N^3)$)



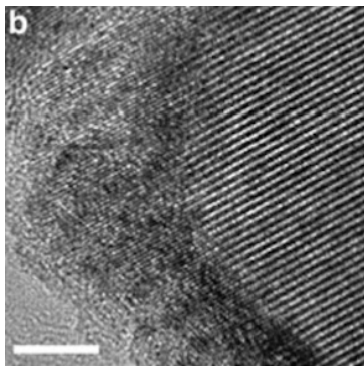
Need for large scale DFT calculations



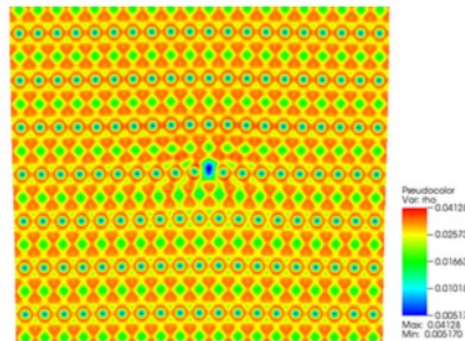
Chemical properties of nanoparticles



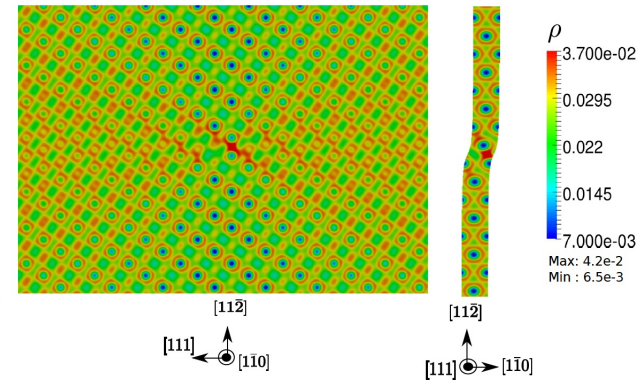
Biological systems



Rocksalt phase formation during Lithiation of Magnetite
He et. al, Nature Comm, 2016



Edge dislocation:
Iyer et al. J. Mech, Phys. Solids (2015)



Screw dislocation:
Das & Gavini J. Mech, Phys. Solids (2017)



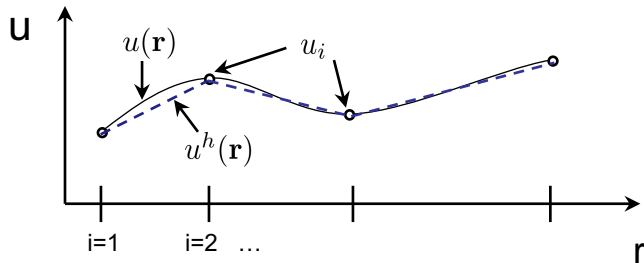
DFT – Finite Element discretization

- Use finite-element basis for computing –

$$\psi_i^h(\mathbf{r}) = \sum_k \psi_{ik} N_k(\mathbf{r}) \quad i = 1, 2, \dots, N$$

ψ_{ik} – Nodal values

$N_k(\mathbf{r})$ – Shape functions



Features of FE basis

- Systematic convergence
 - ❖ Element size
 - ❖ Polynomial order
- Adaptive refinement
- Complex geometries and boundary conditions
- Potential for excellent parallel scalability

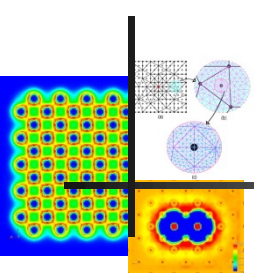
White et al. (1989); Tuschida & Tsukada (1995); Pask et al. (1999); Pask et al. (2001) [and many others]

But ... **huge degree of freedom disadvantage!**

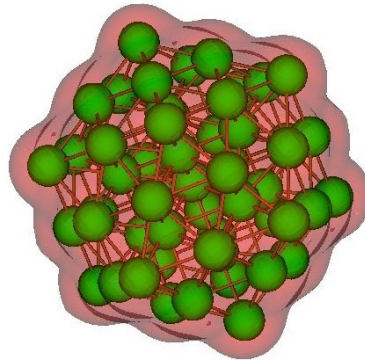
By changing the positioning of the nodes the spatial resolution of basis can be changed/adapted



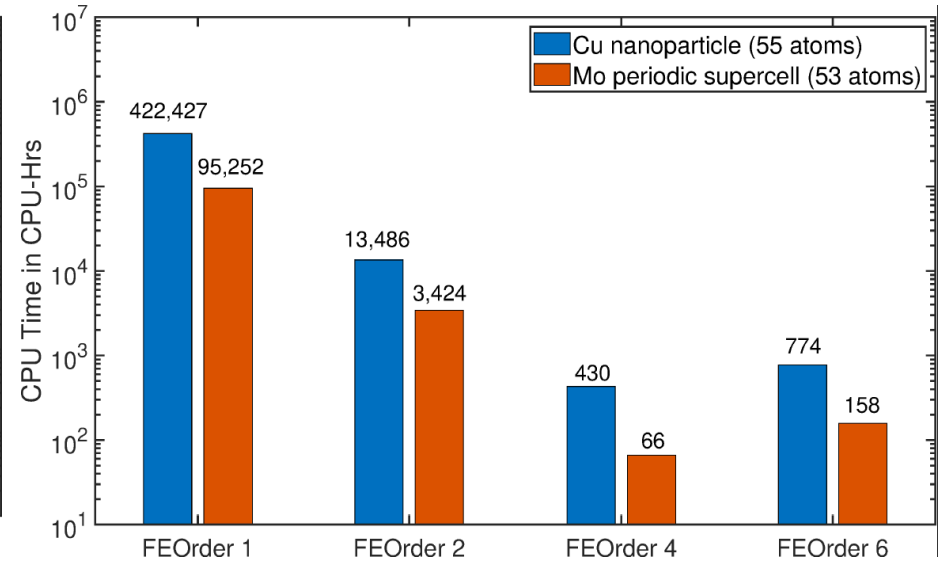
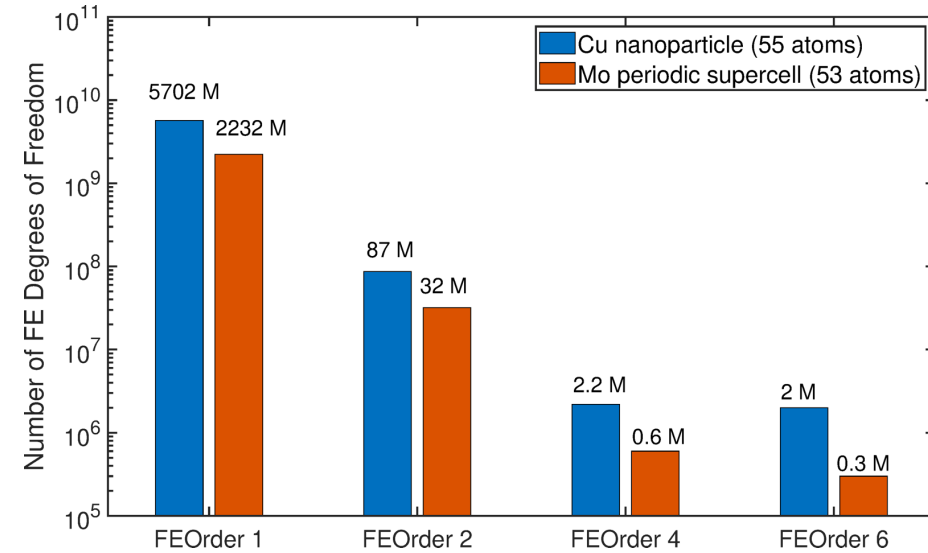
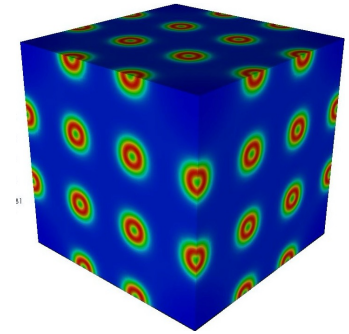
Higher (polynomial) order FE basis



I. Cu nanoparticle
55 atoms



II. Mo periodic
supercell w/ vacancy
53 atoms

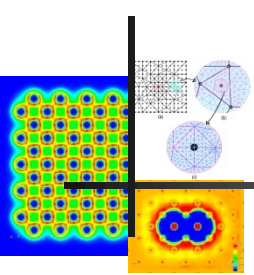


~1000x advantage by using higher-order FE basis !

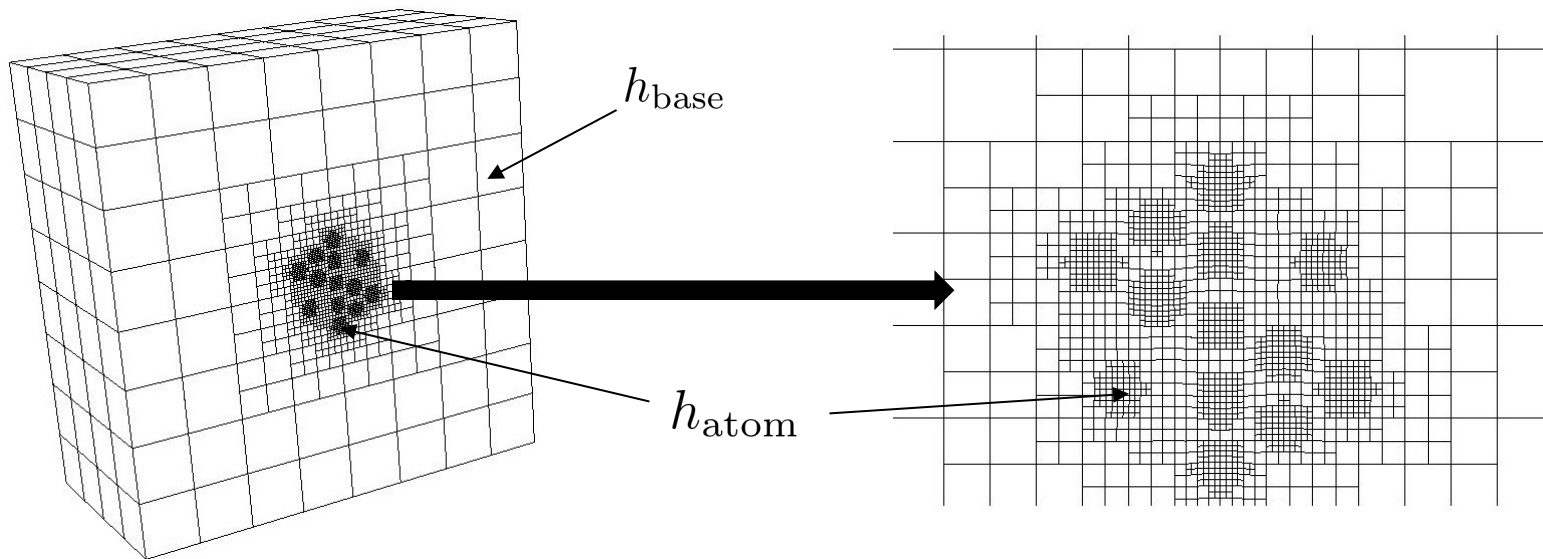


Spatial adaptivity of the FE basis

(Motamarri et al. J Comput Phys. (2013); Motamarri et al. Comput. Phys. Commun. (2020))



- Error Analysis: $|E - E^h| \leq C \left(\sum_i |\bar{\psi}_i - \bar{\psi}_i^h|_{1,\Omega}^2 \right) \leq C \sum_e h_e^{2k} \left[\sum_i |\bar{\psi}_i|_{k+1,\Omega_e}^2 \right]$
- Optimal FE mesh: $\min_h \int_{\Omega} \left\{ h^{2k}(\mathbf{r}) \left[\sum_i |D^{k+1} \bar{\psi}_i(\mathbf{r})|^2 \right] \right\} d\mathbf{r}$ subject to $\int_{\Omega} \frac{d\mathbf{r}}{h^3(\mathbf{r})} = N_E$



System Type pyr II dislocation	DoFs Uniform Mesh	DoFs for Adaptive Mesh
1848 atom Mg	347,206,614	55,112,161
6164 atom Mg	892,047,315	179,034,231



DFT – FE discretization

- Discrete eigenvalue problem:

$$\mathbf{H}\hat{\psi}_k = \varepsilon_k^h \mathbf{M}\hat{\psi}_k$$

$$\mathbf{H}_{ij} = \frac{1}{2} \int_{\Omega} \nabla N_i(\mathbf{r}) \cdot \nabla N_j(\mathbf{r}) d\mathbf{r} + \int_{\Omega} V_{eff}(\mathbf{r}, \mathbf{R}) N_i(\mathbf{r}) N_j(\mathbf{r}) d\mathbf{r}$$

$$\mathbf{M}_{ij} = \int_{\Omega} N_i(\mathbf{r}) N_j(\mathbf{r}) d\mathbf{r}$$

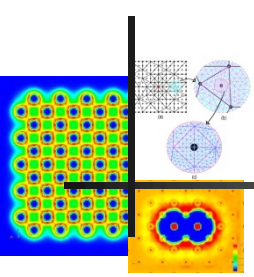
- Transformation to a standard eigenvalue problem:

$$\tilde{\mathbf{H}}\tilde{\psi}_k = \varepsilon_k^h \tilde{\psi}_k \quad \text{where} \quad \tilde{\mathbf{H}} = \mathbf{M}^{-1/2} \mathbf{H} \mathbf{M}^{-1/2} \quad \text{and} \quad \tilde{\psi}_k = \mathbf{M}^{1/2} \hat{\psi}_k$$

- Remark: $\tilde{\mathbf{H}}$ denotes the projection of the Hamiltonian operator into a space spanned by Löwden orthonormalized finite-element basis

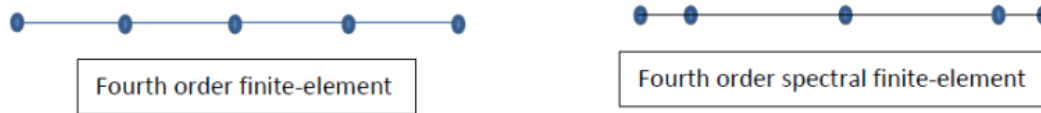


Spectral FE basis and Gauss-Lobatto-Legendre quadrature



➤ Spectral FE basis functions:

- ❖ Constructed from Lagrange polynomials through nodes corresponding to the roots of the derivatives of the Legendre polynomials and boundary nodes (GLL points)



- ❖ Upon using a Gauss-Lobatto-Legendre quadrature rule, the quadrature points coincide with the FE nodes



$$\mathbf{M}_{ij} = \int_{\Omega_e} N_i(\mathbf{r})N_j(\mathbf{r})d\mathbf{r} = C_i\delta_{ij}$$

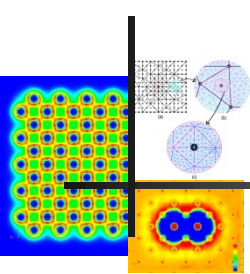
➤ Remarks:

- ❖ Transformation to standard eigenvalue problem is trivial
- ❖ The reduced order quadrature rule is only employed for the computation of the overlap matrix, and the full Gauss quadrature is employed to compute the Hamiltonian matrix.

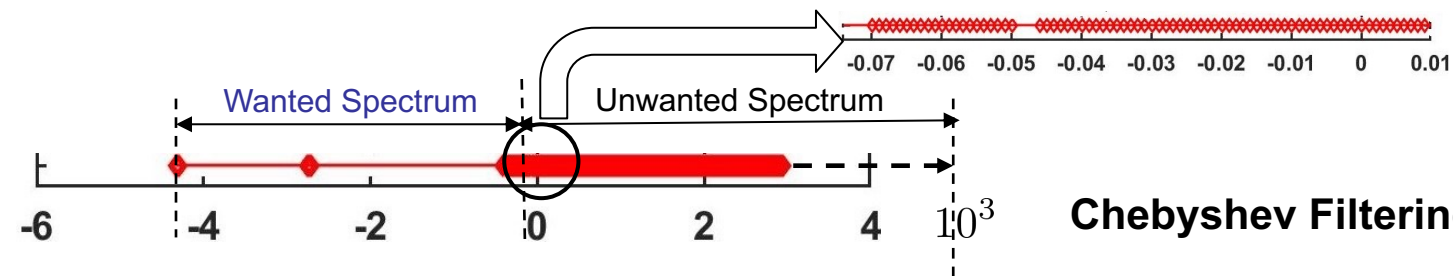


Eigen-space computation: Chebyshev acceleration

(Zhou et al. J. Comput. Phys. 219 (2006); Motamarri et al. J. Comp. Phys. 253, 308-343 (2013))

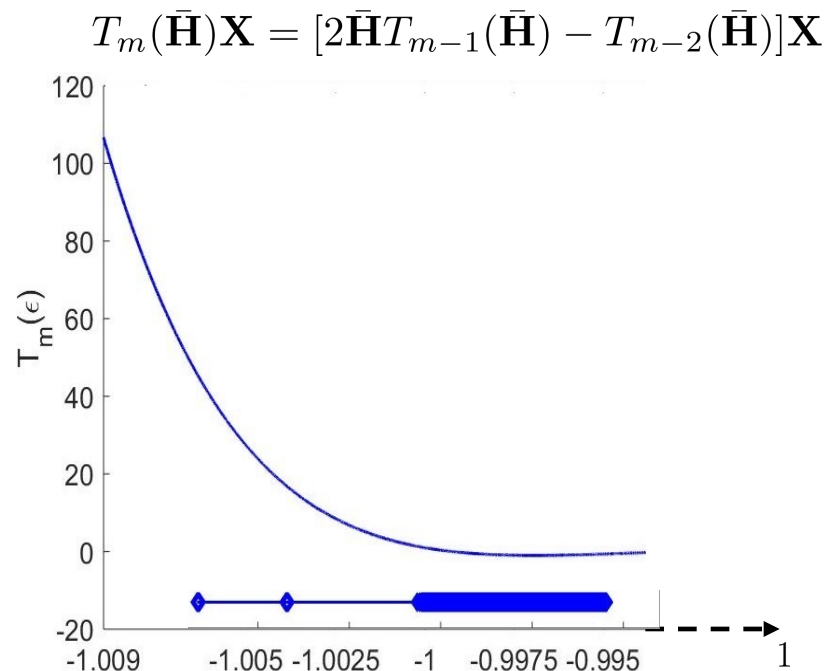
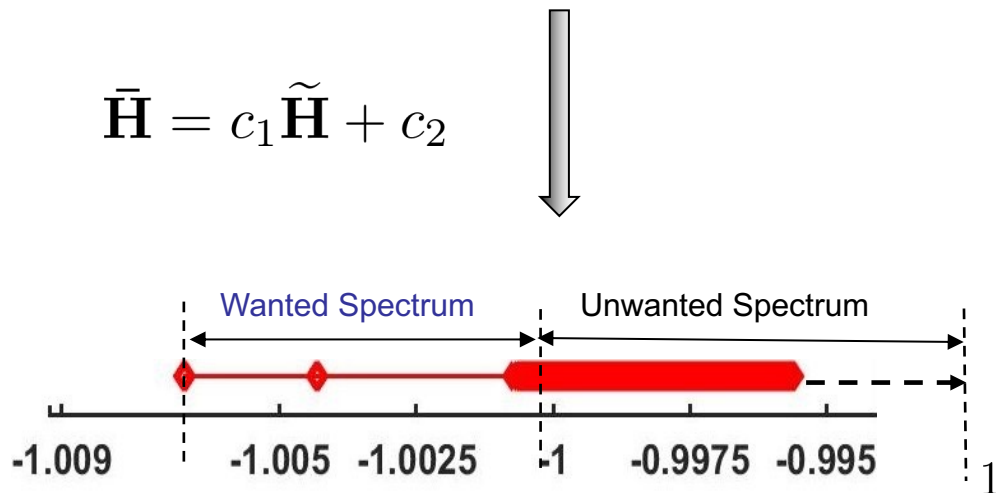


Kohn-Sham eigenvalue problem: $\tilde{\mathbf{H}}\tilde{\psi}_k = \epsilon_k\tilde{\psi}_k$ for $k = 1, 2, \dots, N$ ($N \sim 1.2N_e/2$)



Chebyshev Filtering: $T_m(\bar{\mathbf{H}})\tilde{\Psi} = \tilde{\Psi}_F$

$$\bar{\mathbf{H}} = c_1\tilde{\mathbf{H}} + c_2$$



Numerical algorithm

1. Start with initial guess for electron density $\rho_{in}^h(\mathbf{r}) = \rho_0(\mathbf{r})$ and the initial wavefunctions
2. Compute the discrete Hamiltonian $\tilde{\mathbf{H}}$ using the input electron density ρ_{in}^h
3. **CF:** Chebyshev filtering: $\tilde{\Psi}_F = T_m(\bar{\mathbf{H}})\tilde{\Psi}$
4. **Orthonormalize** CF basis: $\tilde{\Psi}_F \rightarrow \tilde{\Psi}_F^o$
5. **Rayleigh-Ritz procedure:**
 - ❖ Compute projected Hamiltonian: $\hat{\mathbf{H}} = \tilde{\Psi}_F^{o\dagger} \tilde{\mathbf{H}} \tilde{\Psi}_F^o$
 - ❖ Diagonalize $\hat{\mathbf{H}}$: $\hat{\mathbf{H}}\mathbf{Q} = \mathbf{Q}\mathbf{D}$
 - ❖ Subspace rotation: $\tilde{\Psi}^R = \tilde{\Psi}_F^o \mathbf{Q}$
6. Compute electron density $\rho_{out}^h(\mathbf{r}) = 2 \sum_{i=1}^N f(\epsilon_i^h, \mu) |\psi_i^h(\mathbf{r})|^2$
7. If $\|\rho_{out}^h(\mathbf{r}) - \rho_{in}^h(\mathbf{r})\| < tol$, EXIT; else, compute new ρ_{in}^h using a mixing scheme and go to (2).



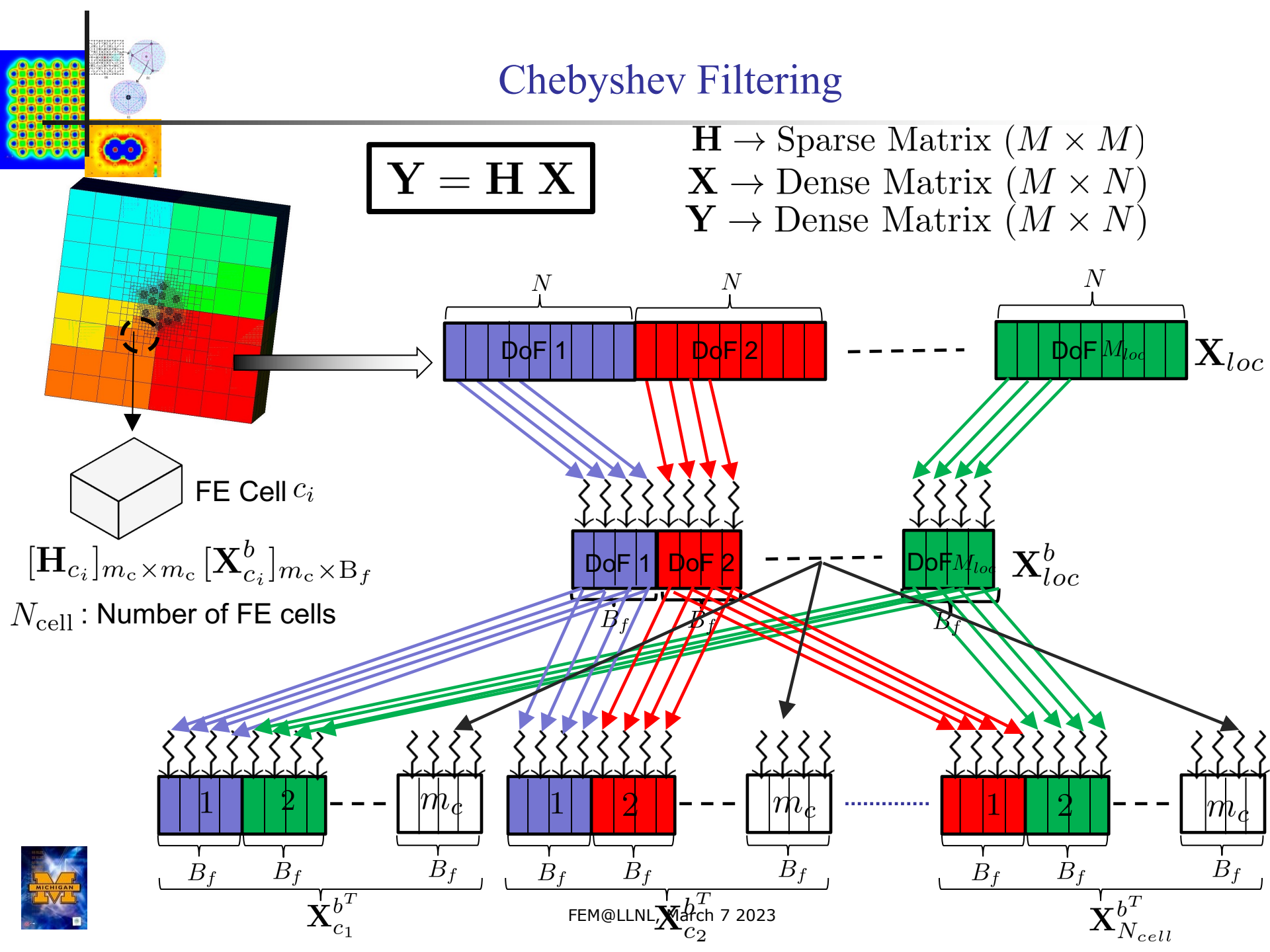
Chebyshev Filtering

$$\mathbf{Y} = \mathbf{H} \mathbf{X}$$

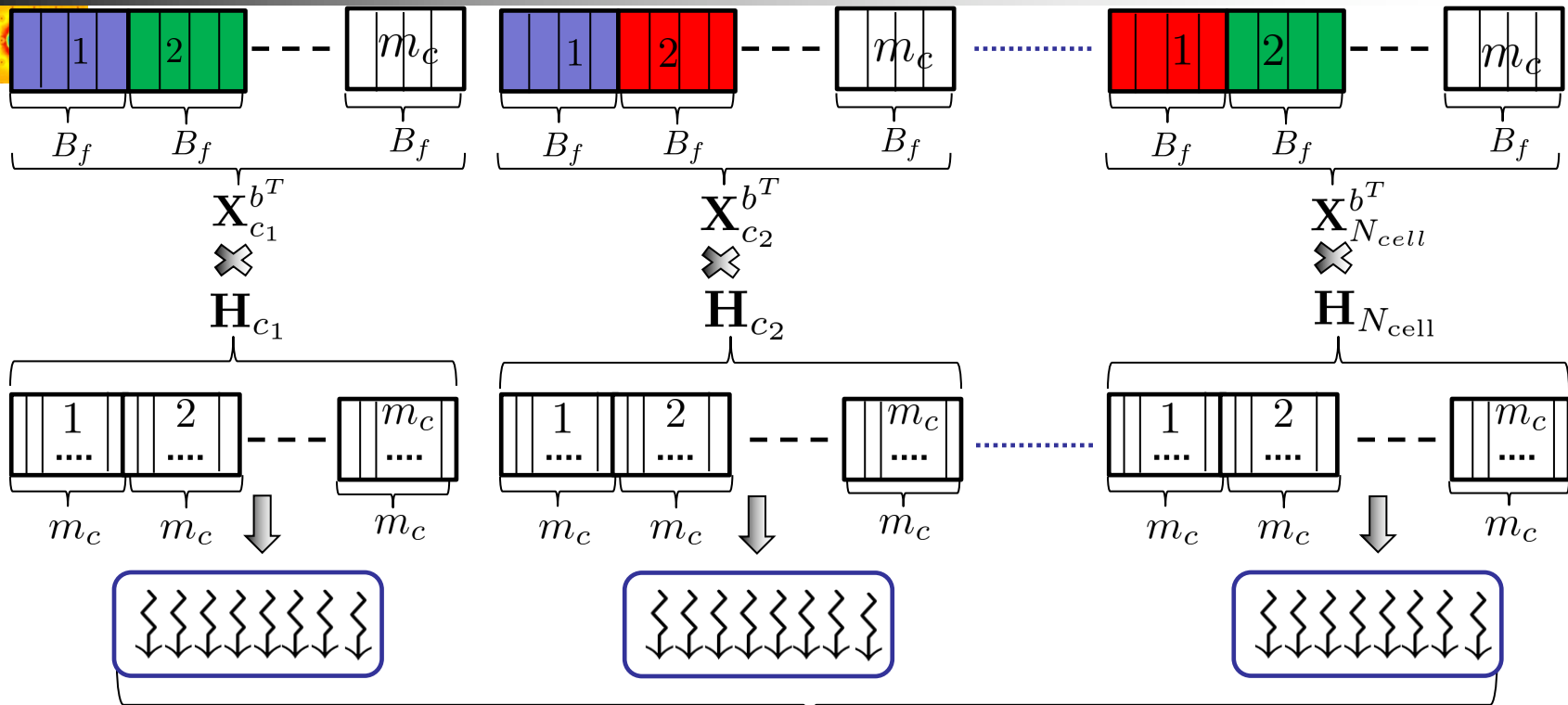
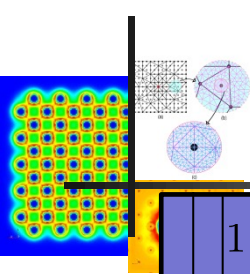
$\mathbf{H} \rightarrow$ Sparse Matrix ($M \times M$)

$\mathbf{X} \rightarrow$ Dense Matrix ($M \times N$)

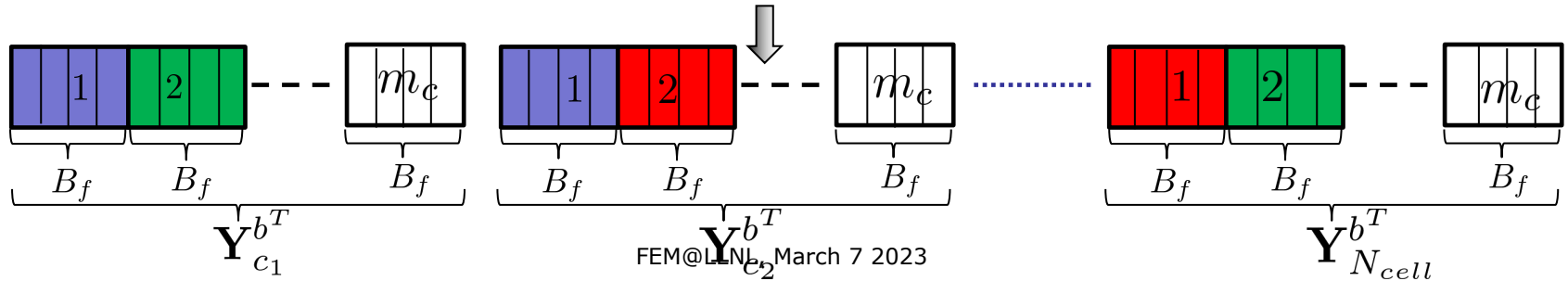
$\mathbf{Y} \rightarrow$ Dense Matrix ($M \times N$)



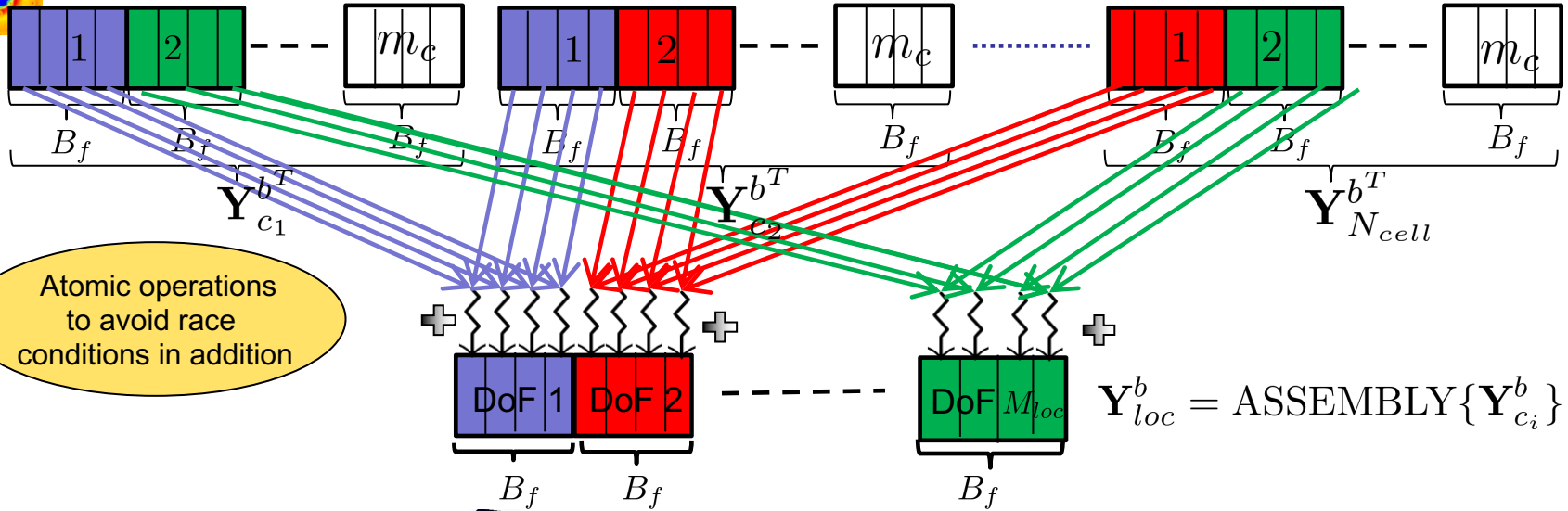
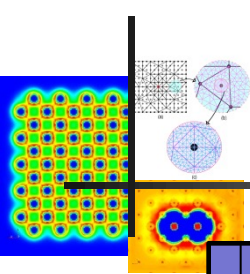
Chebyshev Filtering



Strided Batched xGEMM

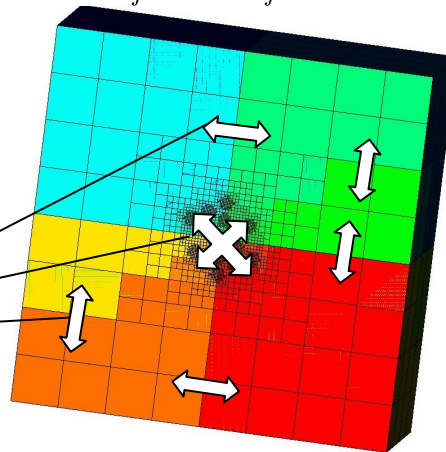


Chebyshev Filtering



$$\mathbf{Y}^b = \text{ASSEMBLY}\{\mathbf{Y}_{loc}^b\}$$

Assembly across processor boundaries: Communication in FP32



Blocked approach allows overlapping compute of one block with communication of another block

$$\mathbf{Y}^b = T_m(\mathbf{H})\mathbf{X}^b = [2\mathbf{H}T_{m-1}(\mathbf{H}) - T_{m-2}(\mathbf{H})]\mathbf{X}^b$$

Repeat for $b = 1 \dots \frac{N}{B_f}$



Performance of Chebyshev filtering (Summit)

Case study: Mg 3x3x3 supercell with a vacancy. (1070 electrons)

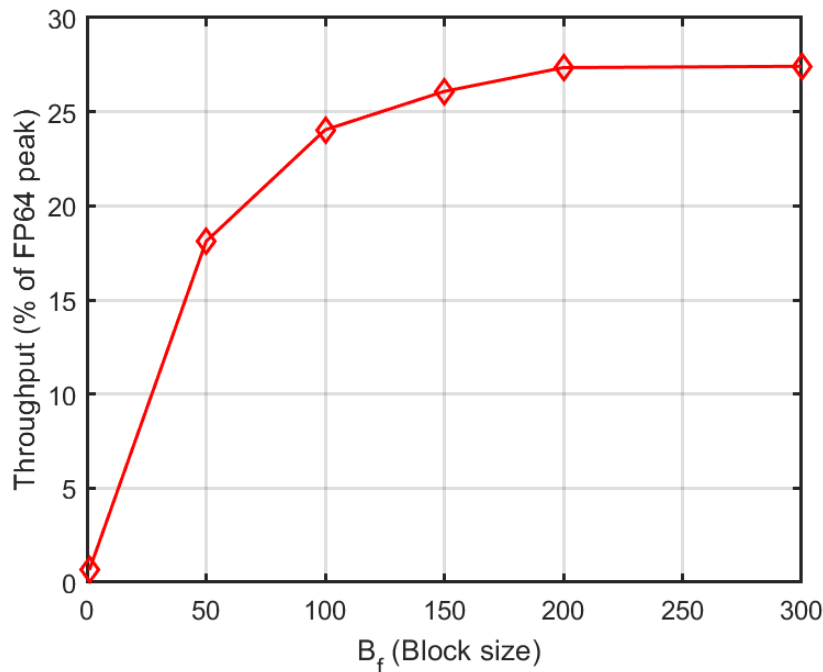


Fig: Chebyshev filtering throughput on 2 Summit nodes using 12 GPUs for various block sizes. FP64 peak of 2 Summit nodes is 87.6 TFLOPS

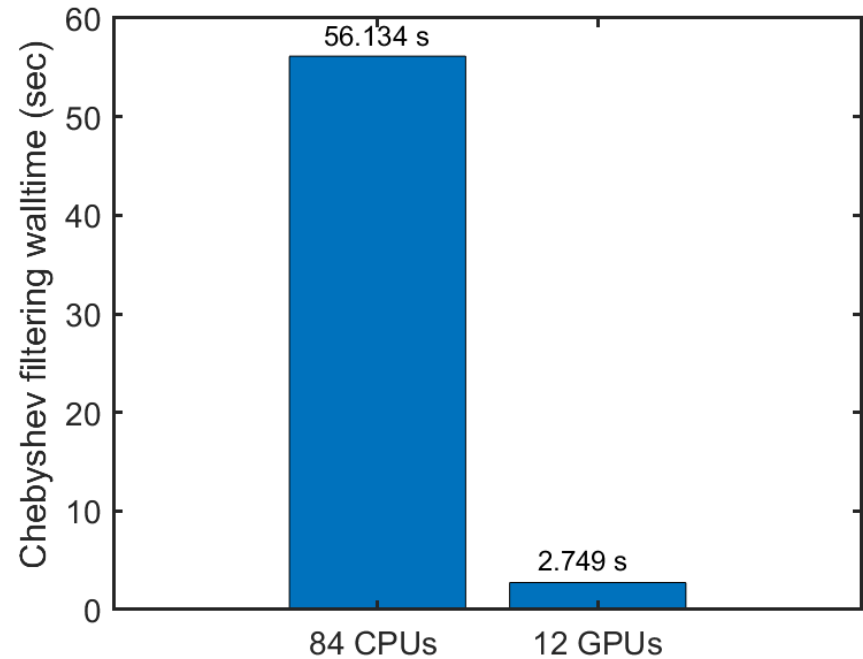


Fig: 20.4x GPU speed up for Chebyshev filtering. CPU run used 2 Summit nodes with 42 MPI tasks per node while GPU run used 2 Summit nodes with 12 GPUs



Orthogonalization: Cholesky Gram-Schmidt

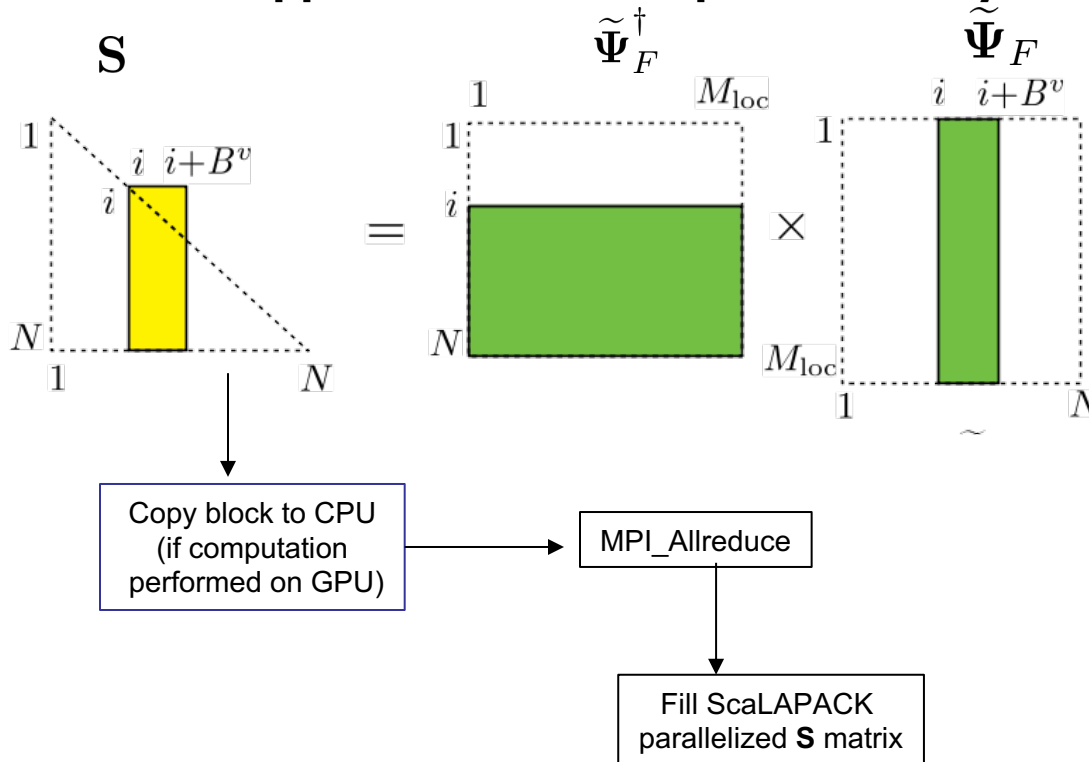
- Cholesky factorization of the overlap matrix: $\mathbf{S} = \tilde{\Psi}_F^\dagger \tilde{\Psi}_F = \mathbf{L}\mathbf{L}^\dagger$. $\mathcal{O}(MN^2)$
- Orthonormal basis construction: $\tilde{\Psi}_F^o = \tilde{\Psi}_F \mathbf{L}^{-1\dagger}$. $\mathcal{O}(MN^2)$

Mixed precision computation for Chol-GS

1. $\mathbf{S} = \text{DP} \{ \mathbf{S}_d \} + \text{SP} \{ \mathbf{S}_{od} \}$
2. $\mathbf{S} = \mathbf{L}\mathbf{L}^\dagger$ in double precision.
3. Orthonormal basis construction:

$$\tilde{\Psi}_F^o = \text{DP} \left\{ \tilde{\Psi}_F \mathbf{L}_d^{-1\dagger} \right\} + \text{SP} \left\{ \tilde{\Psi}_F \mathbf{L}_{od}^{-1\dagger} \right\}$$

Blocked approach to reduce peak memory



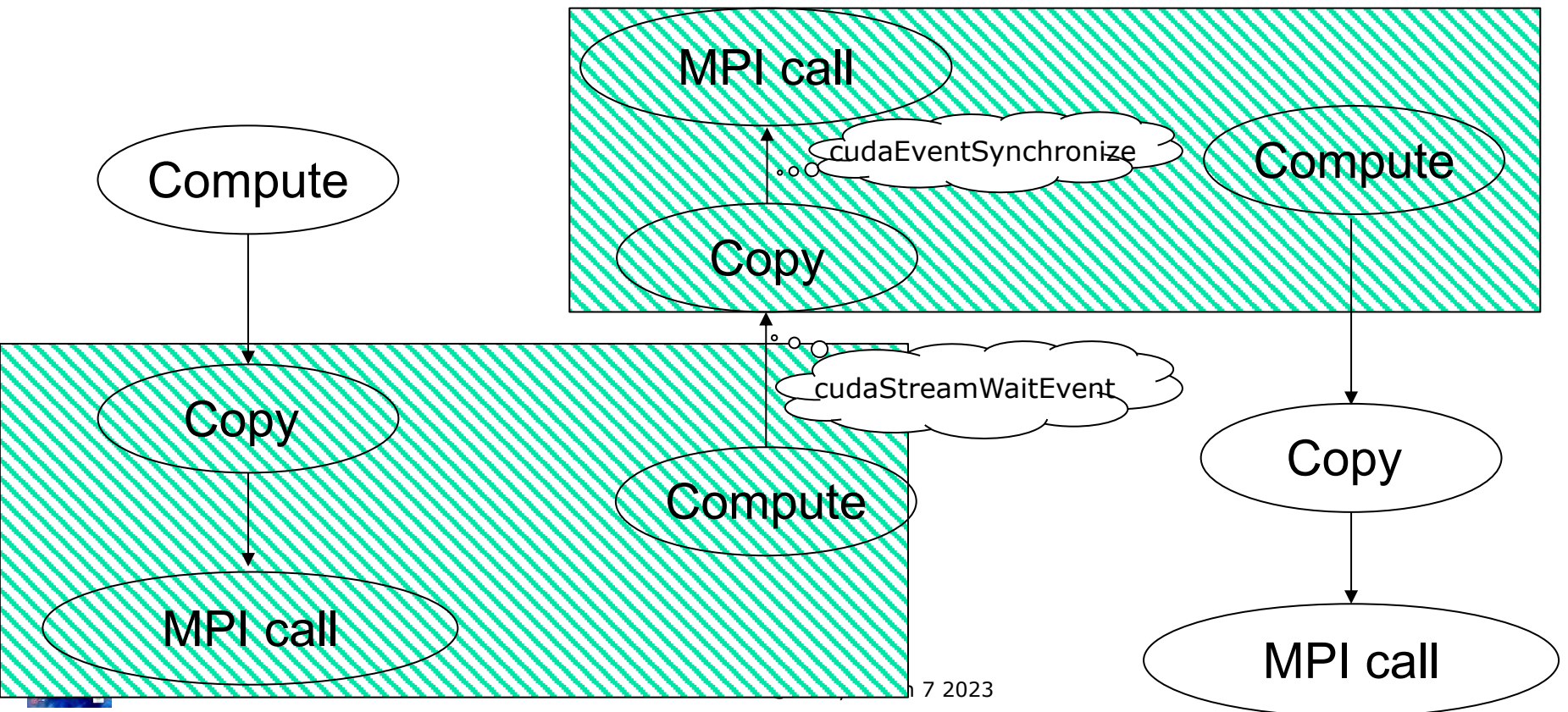
Overlapping compute and data movement on GPUs

Execute copy and MPI calls of current block asynchronously with compute of the successive block.

Block $i-1$

Block i

Block $i+1$

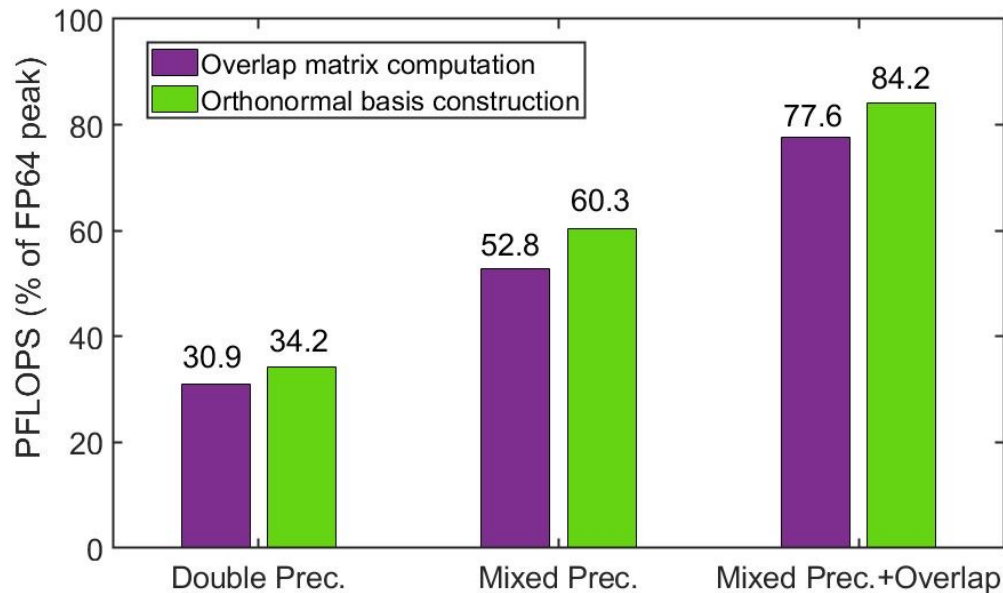


Orthogonalization: Cholesky Gram-Schmidt

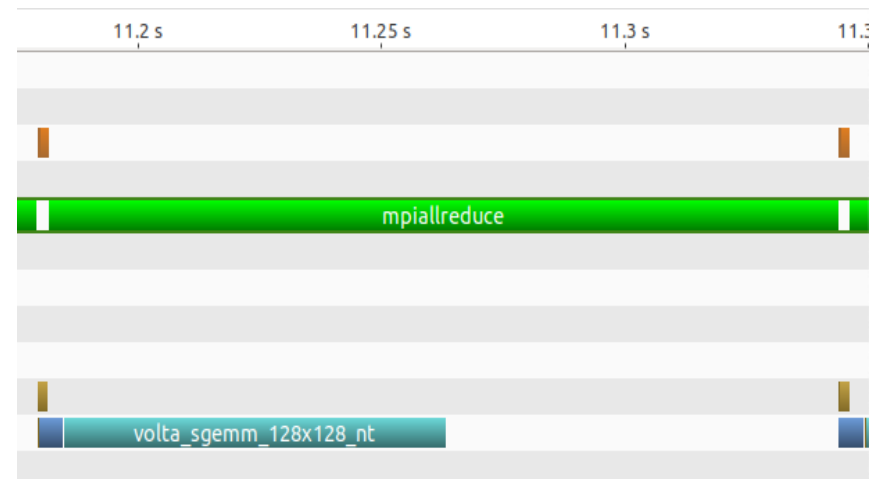
Summit GPU cluster benchmark

Performance improvements due to mixed precision algorithm, and overlapping compute and data movement.

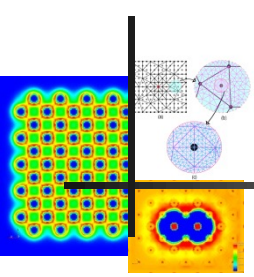
Case study: Mg dislocation system (61,640 electrons) using 1300 Summit nodes



NVVP profile snapshot



Rayleigh-Ritz procedure



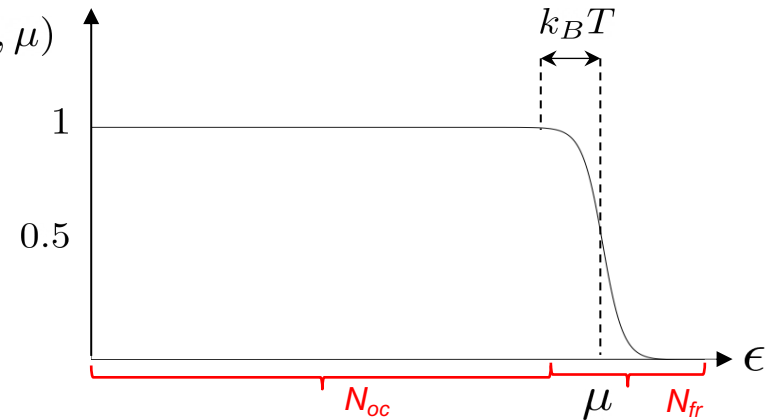
- Compute projected Hamiltonian: $\hat{\mathbf{H}} = \tilde{\Psi}_F^{\circ\dagger} \tilde{\mathbf{H}} \tilde{\Psi}_F^{\circ}$. $\mathcal{O}(MN^2)$
- Diagonalization of $\hat{\mathbf{H}}$: $\hat{\mathbf{H}}\mathbf{Q} = \mathbf{Q}\mathbf{D}$. $\mathcal{O}(N^3)$
- Subspace rotation step: $\tilde{\Psi}^{\mathbf{R}} = \tilde{\Psi}_F^{\circ} \mathbf{Q}$. $\mathcal{O}(MN^2)$

Mixed precision computation for RR

Compute projected Hamiltonian: $f(\epsilon, \mu)$

$$\rho_{\text{out}}^h(\mathbf{x}) = 2 \sum_{i=1}^N f(\epsilon_i^h, \mu) |\psi_i^h(\mathbf{x})|^2$$

$$\tilde{\Psi}_F^{\circ} = \begin{bmatrix} \tilde{\Psi}_{\text{oc}}^{\circ} & \tilde{\Psi}_{\text{fr}}^{\circ} \end{bmatrix}$$



$$\begin{bmatrix} \hat{\mathbf{H}}_{\text{oc-oc}} & \hat{\mathbf{H}}_{\text{oc-fr}} \\ \hat{\mathbf{H}}_{\text{fr-oc}} & \hat{\mathbf{H}}_{\text{fr-fr}} \end{bmatrix} = \begin{bmatrix} \text{SP} \left\{ \tilde{\Psi}_{\text{oc}}^{\circ\dagger} \tilde{\mathbf{H}} \tilde{\Psi}_{\text{oc}}^{\circ} \right\} & \text{SP} \left\{ \tilde{\Psi}_{\text{oc}}^{\circ\dagger} \tilde{\mathbf{H}} \tilde{\Psi}_{\text{fr}}^{\circ} \right\} \\ \text{SP} \left\{ \tilde{\Psi}_{\text{fr}}^{\circ\dagger} \tilde{\mathbf{H}} \tilde{\Psi}_{\text{oc}}^{\circ} \right\} & \text{DP} \left\{ \tilde{\Psi}_{\text{fr}}^{\circ\dagger} \tilde{\mathbf{H}} \tilde{\Psi}_{\text{fr}}^{\circ} \right\} \end{bmatrix}$$

Subspace rotation step: $\tilde{\Psi}^{\mathbf{R}} = \text{DP} \left[\tilde{\Psi}_F^{\circ} \mathbf{Q}_d \right] + \text{SP} \left[\tilde{\Psi}_F^{\circ} \mathbf{Q}_{\text{od}} \right]$

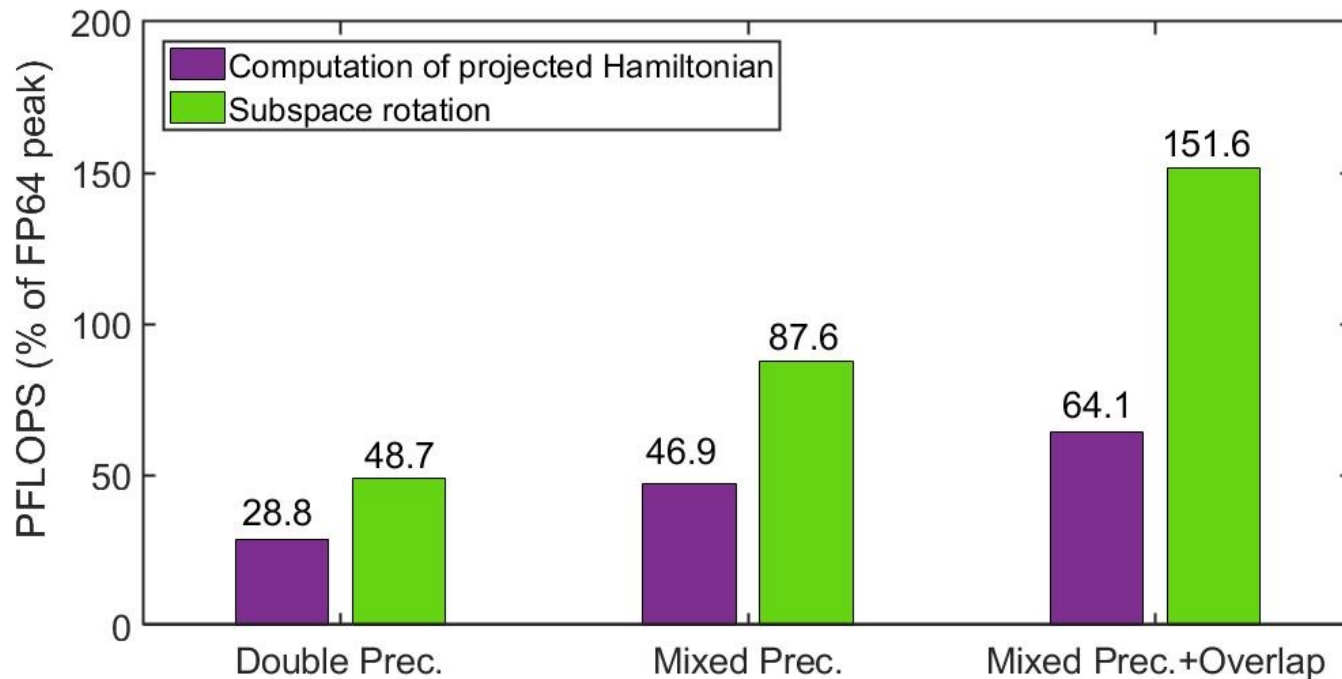


Rayleigh-Ritz procedure

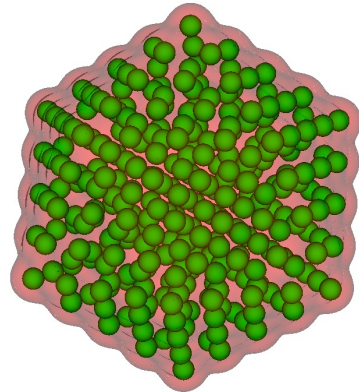
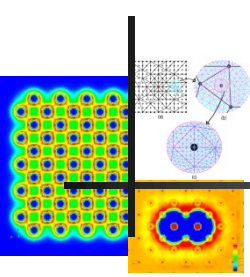
- Blocked approach used for memory optimization
- Compute and communication are overlapped taking advantage of blocked approach.

Summit GPU cluster benchmark

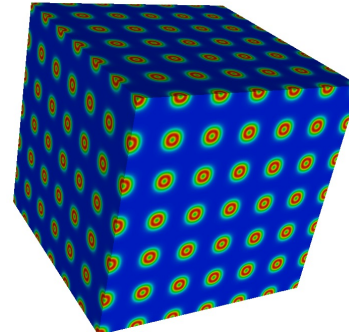
Case study: 61,640 electrons system using 1300 Summit nodes



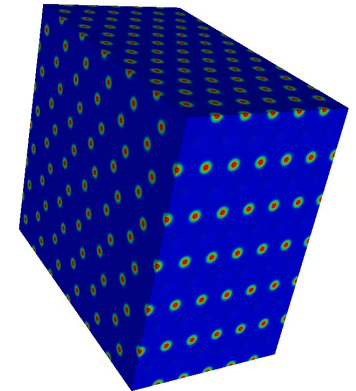
Accuracy and robustness of mixed precision computations



I. Cu nanoparticle
5871 electrons



II. Mo periodic
supercell w/ vacancy
6034 electrons



III. Mg periodic
supercell w/ vacancy
8630 electrons

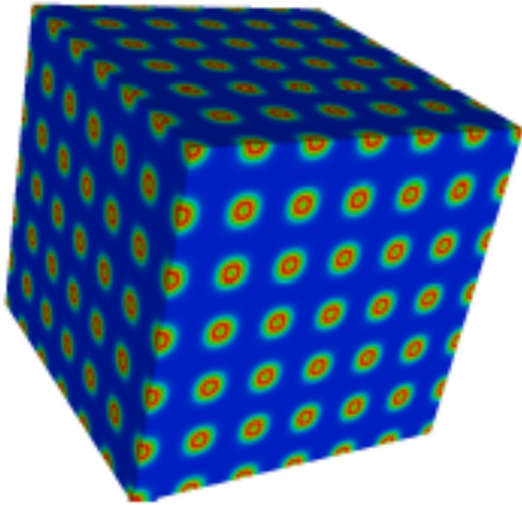
Energy error (Ha/atom)	5×10^{-12}	7×10^{-12}	3×10^{-12}
Max force error (Ha/Bohr)	3×10^{-6}	7×10^{-7}	2×10^{-6}
Total SCFs (DP, MP)	(46,46)	(30,30)	(18,18)



Comparison of DFT-FE & Quantum Espresso

- Vacancy in BCC Mo – periodic calculation ; ONCV pseudopotential
- Accuracy for all calculations $< 0.1 \text{ mHa/atom}$ ($\sim 2 \text{ meV/atom}$)

Compute resources per SCF in Node-Hrs
(NERSC Cori KNL)

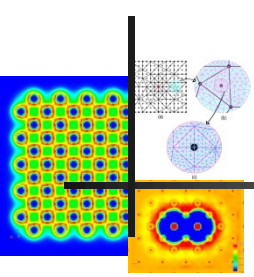


BCCMo6x6x6 monovacancy
(6034 electrons)

System size	Quantum-ESPRESSO (Ecut: 20 Ha)	DFT-FE (h_min: 2.1, p=7)
431 atoms ($N_e = 6034$)	0.56	0.24
1023 atoms ($N_e = 14322$)	22.1	1.4
1999 atoms ($N_e = 27986$)	219.5	7.5



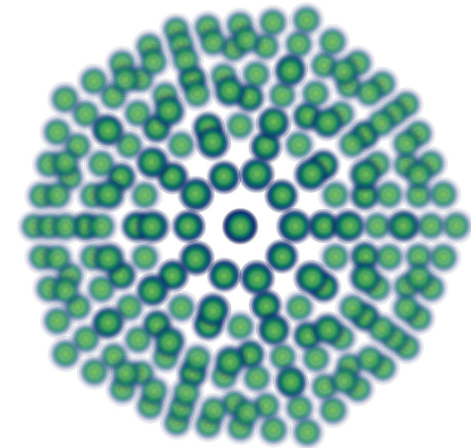
Comparison of DFT-FE & Quantum Espresso



- Cu nanoparticles; ONCV pseudopotential
- Accuracy for all calculations $< 0.1 \text{ mHa/atom}$ ($\sim 2 \text{ meV/atom}$)

Compute resources per SCF in Node-Hrs
(NERSC Cori KNL)

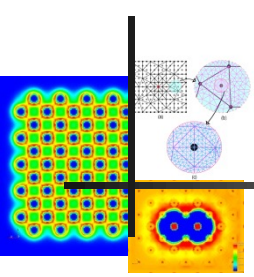
System size	Quantum-ESPRESSO (Ecut: 50 Ha)	DFT-FE (h_min: 0.8; p=6)
147 atoms ($N_e = 2793$)	0.2	0.22
309 atoms ($N_e = 5871$)	5.5	1.3
561 atoms ($N_e = 10569$)	63.4	4.2
923 atoms ($N_e = 17537$)	-	10.9



Cu4shell (5871 electrons)

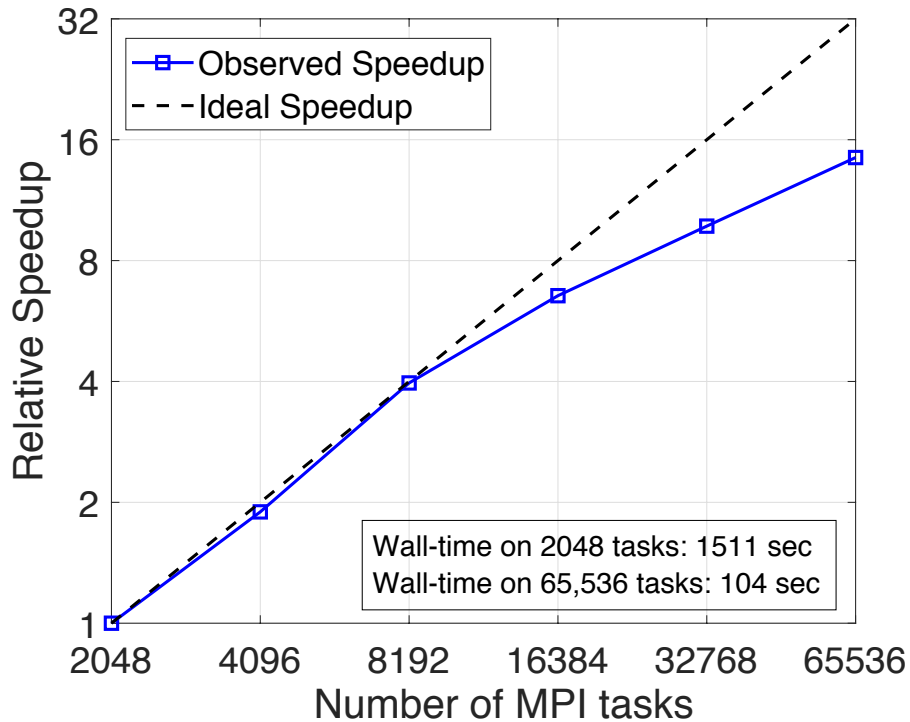


Strong Scaling & Time to solution

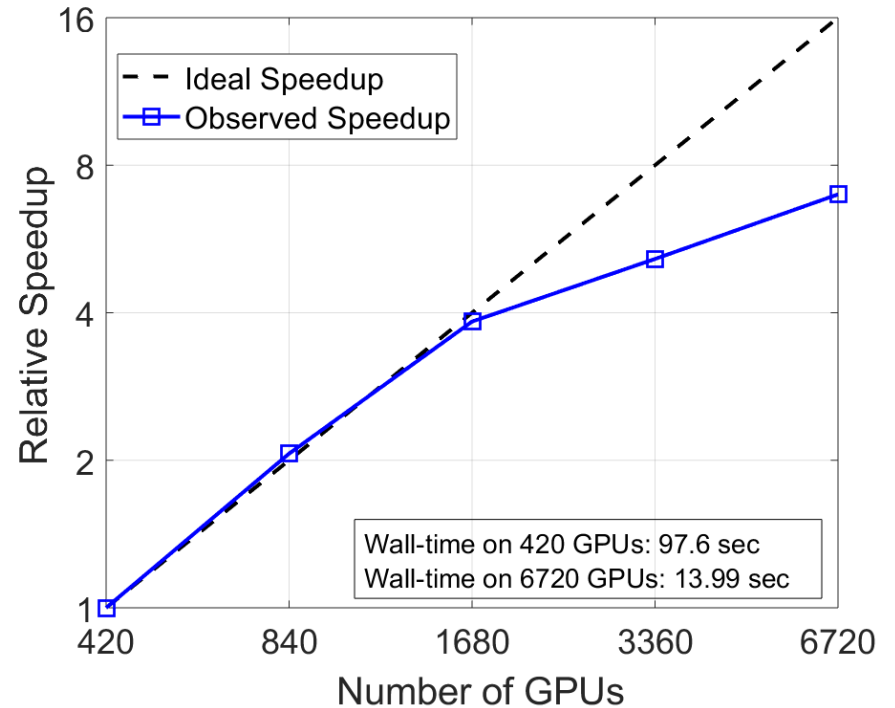


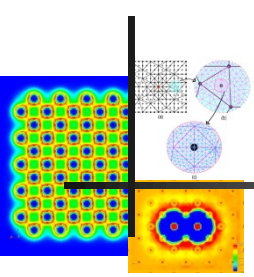
Mg pyr II screw dislocation – 1,848 atoms (18,480 e⁻); 55.11 million FE DoFs

Theta (ALCF)



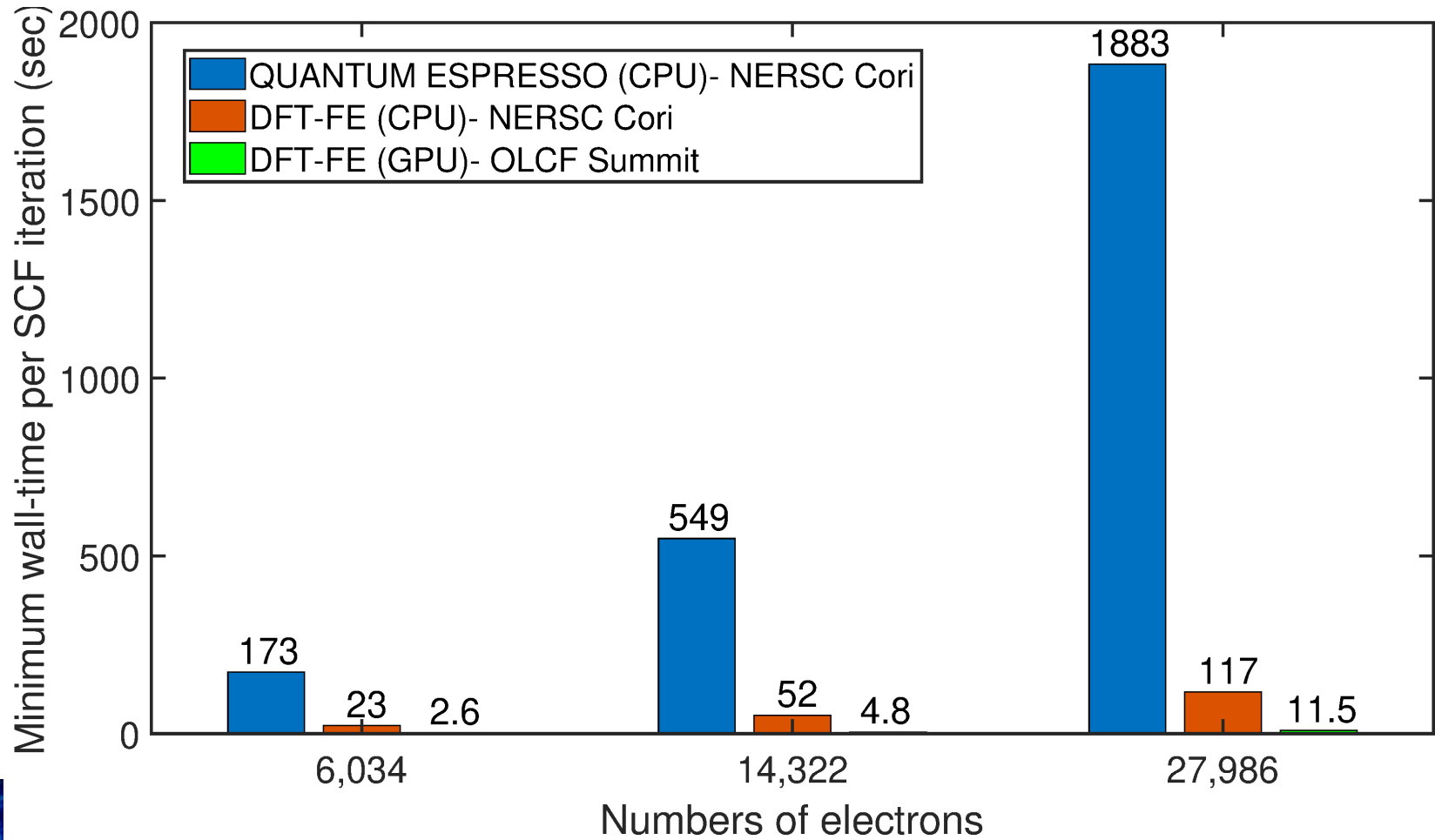
Summit GPUs (OLCF)





Minimum wall-time comparison with plane-wave basis

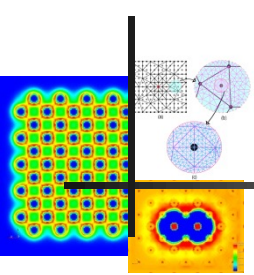
Vacancy in Mo – periodic calculation ; ONCV pseudopotential
Accuracy for all calculations $< 0.1 \text{ mHa/atom}$ ($\sim 2 \text{ meV/atom}$)



DFT-FE – Open source code

(<https://github.com/dftfeDevelopers/dftfe>;

Motamarri et al. Comput. Phys. Commun. (2020); Das et al. Comput. Phys. Commun. (2022))



- Discretization: FE basis; adaptive mesh refinement
- Geometry: Periodic, non-periodic, semi-periodic
- Physics:
 - ❖ Pseudopotential (TM; ONCV) and All-electron calculations (classical FE)
 - ❖ Density based XC functionals
- Calculations:
 - ❖ Ground-state energy
 - ❖ Geometry (ionic and cell) relaxation (Motamarri & Gavini, Phys. Rev. B 97, 165132 (2018))
 - ❖ Ab-initio MD
- Scaling: Tested on Summit, Theta, Stampede, Comet, Cori (up 192,000 cores)
- System sizes: 100,000 electron pseudopotential calculations; 10,000 electron all-electron calculations
- Ported to GPUs:
 - ❖ ~20x speedups (on Summit) in comparison to CPUs on a node-to-node basis
 - ❖ 64PLOPS of sustained performance; ~38% efficiency on Summit
 - ❖ Finalist, 2019 Gordon Bell prize



Application I: Technological challenge of low ductility in Mg

- Magnesium is the lightest structural metal with high strength to weight ratio
 - ❖ 75% lighter than Steel and 30% lighter than Aluminum
- Every 10% reduction in the weight of a vehicle will result in 6-8% increase in fuel efficiency.
 - ❖ Important implications to fuel efficiency and reducing carbon footprint
- Low ductility key issue in the manufacturability of structural components. Main limitation in the adoptability of Mg and Mg alloys in automotive and aerospace sectors. (T.M. Pollock, *Science* **328**, 986-987 (2010))



Courtesy: <https://www.audi-technology-portal.de/en/body>
Current state of art: Hybrid Steel and Aluminum construction

Brittle Mg (pure)

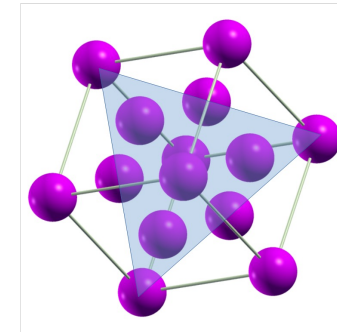
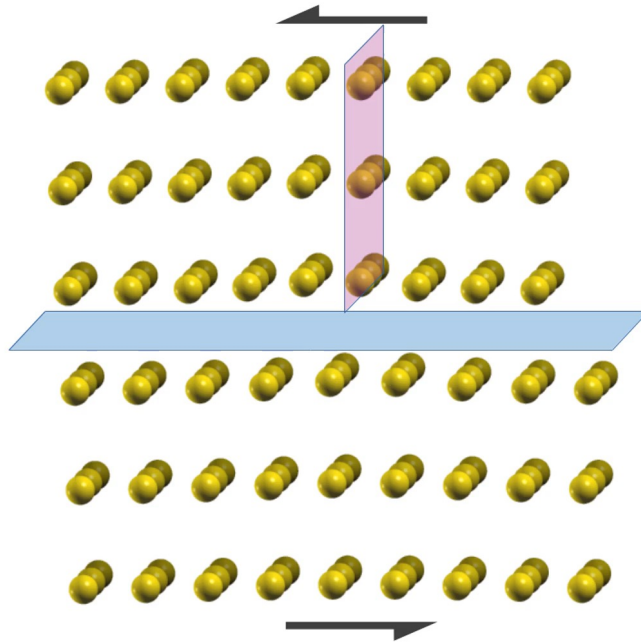
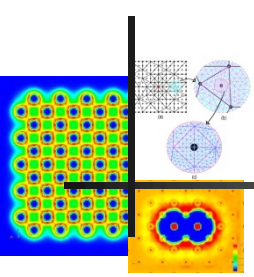


10% CR
CR: cold rolled

S. Sandlöbes et al. *Scientific Reports* 7, 10458 (2017).

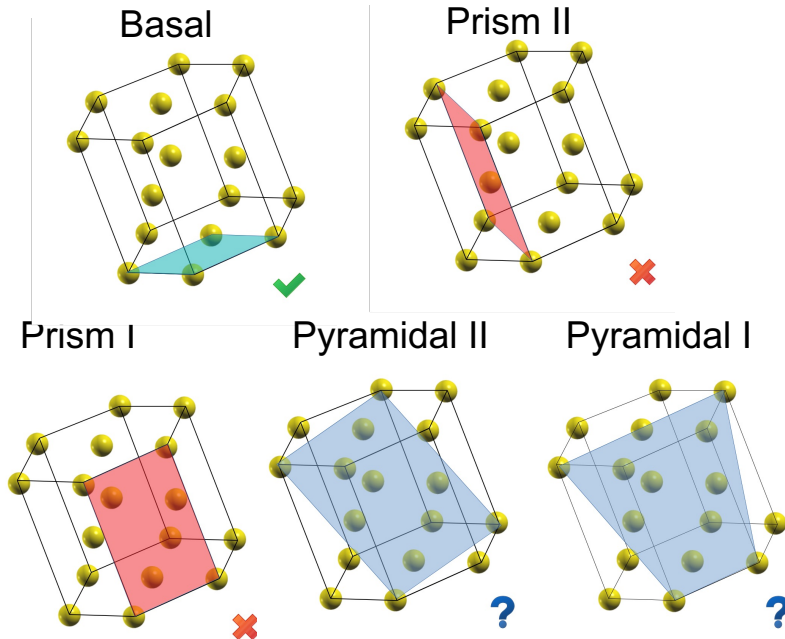


Technological challenge of low ductility in Mg

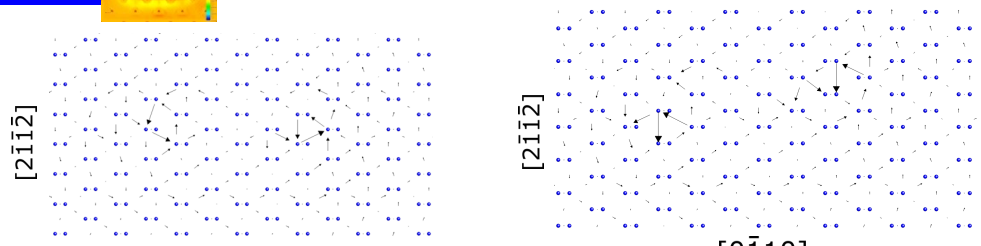
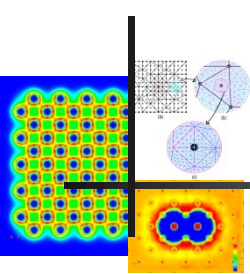


12 slip systems in Face Centered Cubic Crystals → higher ductility

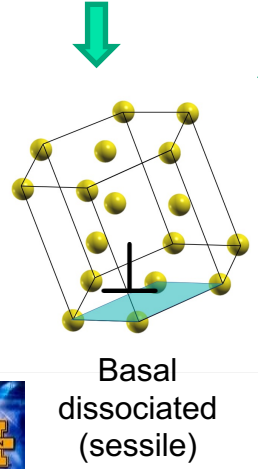
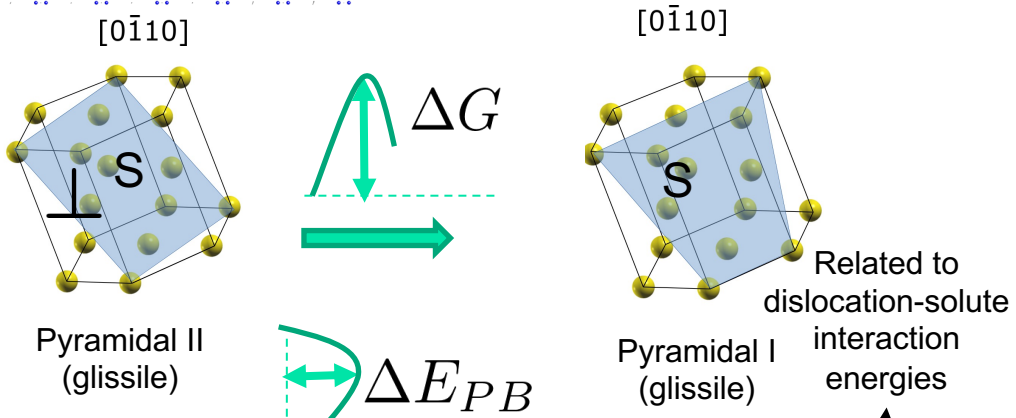
- ❖ Dislocations are energetically more favorable to reside on certain slip systems. (**Energetics**)
- ❖ Dislocation glide occurs after the applied shear stress is greater than the Peierls barrier. (**Activation barrier**)
- ❖ More the number of slip systems where dislocation can glide easily higher is the ductility.



Accurate $\langle c+a \rangle$ dislocation core energetics for ductility enhancement in dilute Magnesium alloys



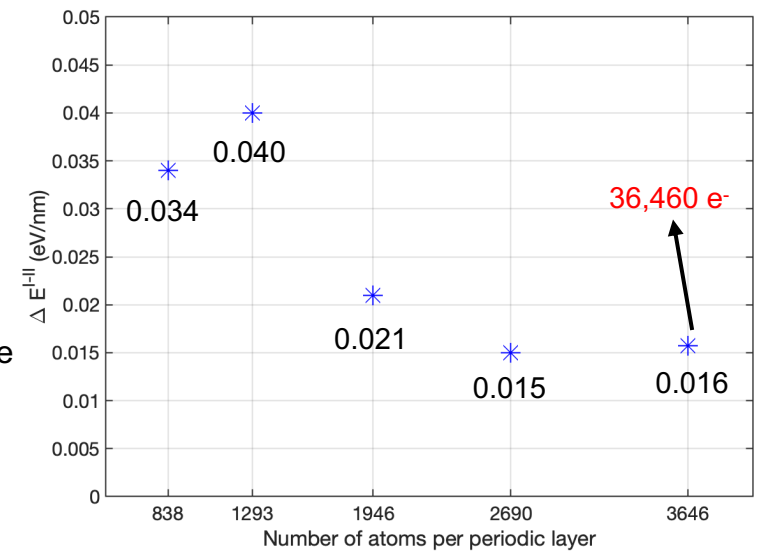
- Chemical accuracy for all calculations ($\langle 0.1\text{mHa/atom}$ energy and $\langle 0.1\text{mHa/Bohr}$ in forces)
- 12 k-points along the dislocation line



$$\Delta G(c) = f_1 \left(\Delta E_{Mg}^{I-II} \right) + f_2(c)$$

The small core energy difference between $\langle c+a \rangle$ screw dislocations on pyramidal I and II planes (ΔE^{I-II}), significantly controls ductility in Mg (*Wu et al., Science., 2018*)

Accurate cell-size converged DFT computation of ΔE^{I-II} has not been possible with DFT codes!

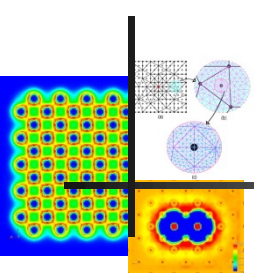


1-2 directions along $[0\bar{1}10]$ – $[2\bar{1}\bar{1}\bar{2}]$

Uniaxial Strains	Slope (eV/nm)
ϵ_{11}	-1.450
ϵ_{22}	-3.507
ϵ_{33}	0.301



Solute grain boundary (GB) segregation energies in Magnesium

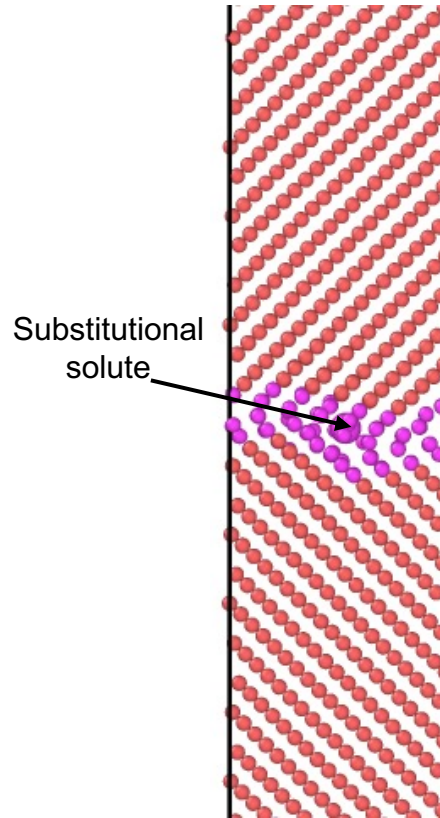


➤ GB solute segregation can enhance non-basal textures in Magnesium through solute drag effects (*Robson 2014, 2016 Met. Mat. Trans. A*)

➤ Understanding equilibrium segregation behavior, which is controlled by ΔE_{seg} , for different GB structures and different solute types with strong chemical interactions

➤ Use DFT computed segregation energies and transport coefficients to predict solute drag on GB

➤ **Challenge: DFT simulations of random GB geometries reaching 4000-5000 atoms with structural relaxation has remained infeasible.**



- Random grain boundary:
 - ❖ Tilt axis: $[1\bar{1}00]$
 - ❖ Tilt angle: 44°
 - ❖ 1932 atoms

E_{GB} Energy with solute atom at a GB site

E_{bulk} Energy with solute atom away from GB

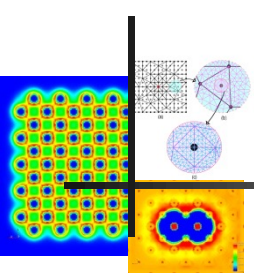
$$\Delta E_{seg} = E_{GB} - E_{bulk}$$

Solute type	ΔE_{seg} (eV)
Al (non-RE)	-0.109
Y (RE)	-0.320

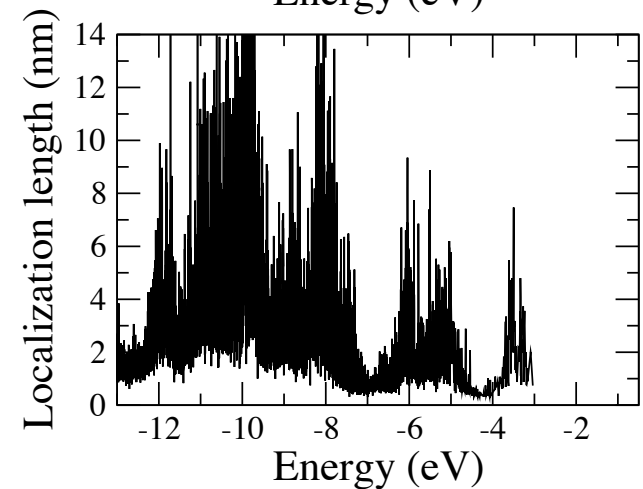
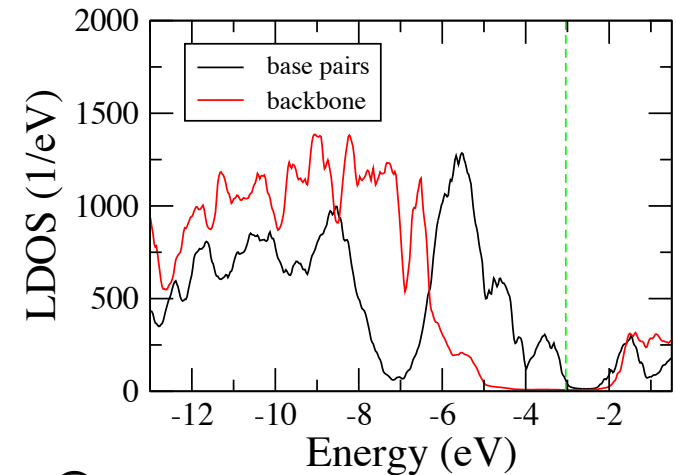
Calculations on larger GB systems with ~4000 atoms ongoing



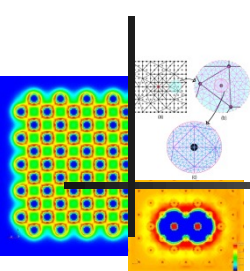
Other Application Studies using DFT-FE



- Understanding electron transport in DNA molecules (*Nature Nanotechnology* 15 836 (2020))
 - ❖ Large-scale simulations involving 100 basis pairs (~6,200 atoms) simulating experiments
 - ❖ Provided new insights into the role of backbone in electron charge transport
- Spin-spin interactions in defects in solids (*npj Computational Materials*, 50 (2021); *Phys. Rev. Mat.* 3 043801 (2019))
 - ❖ Computed spin Hamiltonian parameters that describe electron-electron and electron-nuclear spin interactions
 - ❖ Systematically convergent calculations with all-electron accuracy, possible for the first time
 - ❖ Use mixed pseudopotential and all-electron calculations leveraging the flexibility of the DFT-FE framework

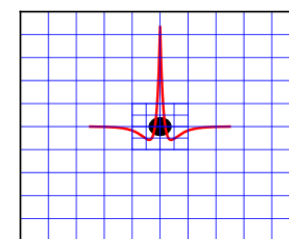
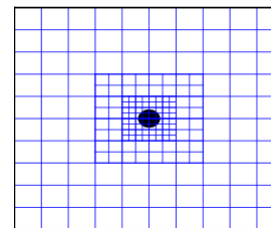
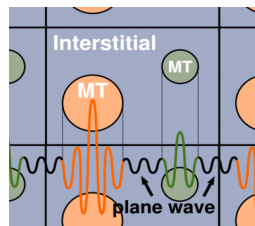
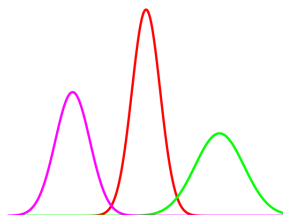


All-electron calculations



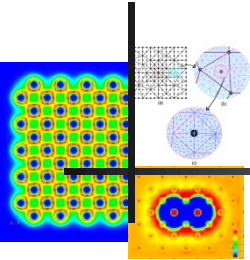
- More stringent basis set requirement to capture the fast oscillations of core electrons
- Need solving for all electrons; though some other approximations to avoid explicit computations can be used.

	Gaussian	APW/LAPW	Classical FE	Enriched FE
Convergence	❌ Incomplete basis, No systematic convergence	✅ Complete basis, Systematic convergence	✅ Complete basis, Systematic convergence	✅ Complete basis, Systematic convergence
Boundary Conditions	❌ using mixed BCs not easy	❌ Only handles periodic BCs	✅ Can handle all BCs	✅ Can handle all BCs
Parallel Scalability	❌ Non-local basis, Poor scalability	❌ Use of FFTs, Poor scalability	✅ Local basis, Good scalability	✅ Local basis, Good scalability
Robustness	❌ Ill-conditioned for large systems.	❌ Sensitive to energy parameters	✅ Well Conditioned	✅ Well Conditioned
Efficiency	✅ 10-100 basis/atom	✅ 100-1000 basis/atom	❌ $10^5 - 10^6$ basis/atom	✅ 10000-50000 basis/atom



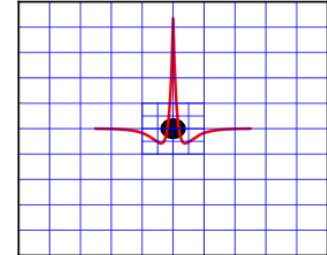
Enriched finite element basis

(Kanungo & Gavini, Phys Rev. B 95 035112 (2017); Rufus, Kanungo & Gavini Phys Rev. B 106 085112 (2021))



- Additional functions appended to the 'Classical' FE basis

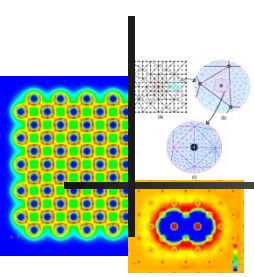
$$\psi^h(\mathbf{x}) = \sum_j N_j^C(\mathbf{x})\psi_j^C + \sum_k N_k^E(\mathbf{x})\psi_j^E$$



(Yamaka & Hodo PRB (2005); Sukumar & Pask IJNME (2009), Extreme Mech. Let. (2017))

- Enriched functions: Radial part computed using 1D radial Kohn-Sham solve, and multiplied by spherical harmonics
 - ❖ Compact support for the enriched functions is obtained by multiplying with a mollifier
- Orthogonalized enrichment: Orthogonalize with respect to the classical FE basis; improves conditioning of the basis
- Integrals computed using an adaptive quadrature (Mousavi et al. (2012))
- Key advantages of enrichment:
 - ❖ Reduced degrees of freedom
 - ❖ Reduced spectral width of the discrete Hamiltonian (especially important for Chebyshev filtering approach for solving the Kohn-Sham problem)

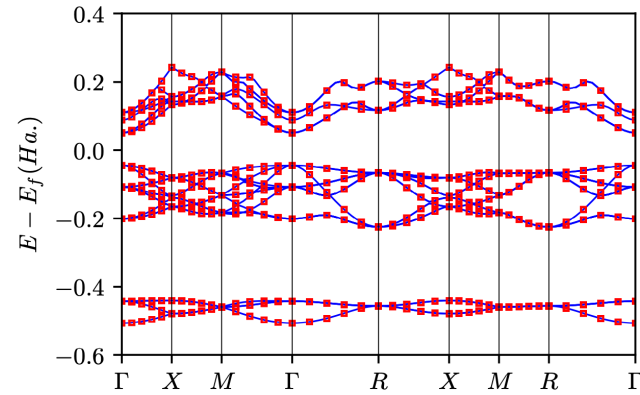
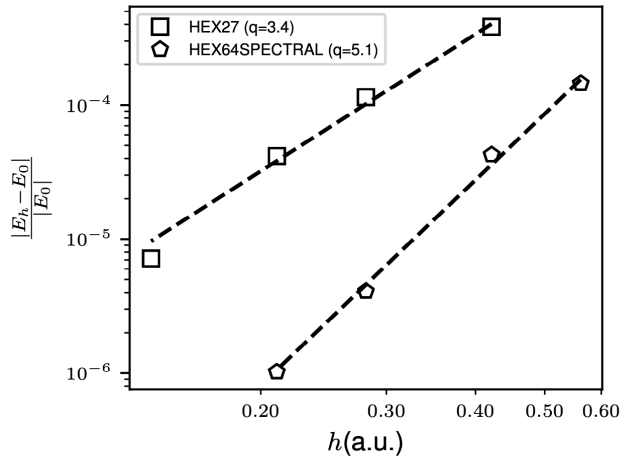




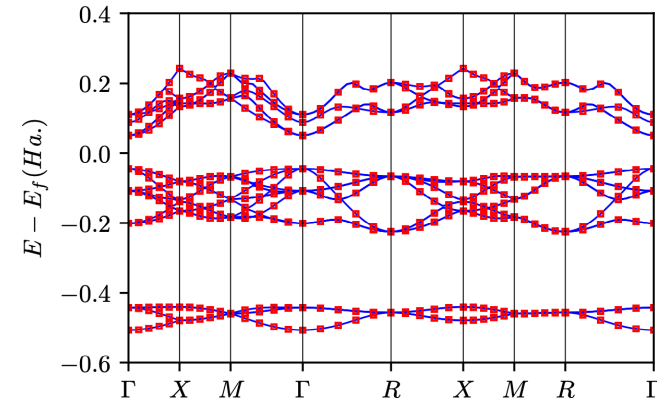
Accuracy of EFE/OEFE basis

(Rufus, Kanungo & Gavini Phys. Rev. B (2021))

Diamond unit cell



Band structure of MgS



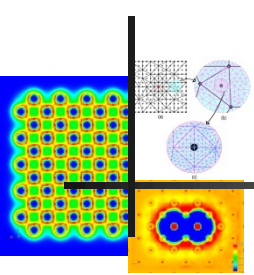
Band structure of Ce

Divacancy in SiC supercells

Supercell	Atoms (Electrons)	OEFE	LAPW+lo	CFE
2 × 2 × 2	62 (620)	-163.1053	-163.1054	-163.1056
3 × 3 × 3	214 (2,140)	-163.1119	-163.1117	-
4 × 4 × 4	510 (5,100)	-163.1133	-	-
5 × 5 × 5	998 (9,980)	-163.1135	-	-



Computational efficiency & Scalability



Divacancy in SiC supercells

Supercell	OEFE		LAPW+lo		CFE	
	C	(c, N)	C	(c, N)	C	(c, N)
$2 \times 2 \times 2$	1.48	(0.08, 12)	1.28	(0.04, 32)	197	(10.9, 18)
$3 \times 3 \times 3$	13.92	(0.76, 14)	45.5	(1.23, 37)	-	-
$4 \times 4 \times 4$	132.6	(6.4, 18)	-	-	-	-
$5 \times 5 \times 5$	1102.5	(45.9, 21)	-	-	-	-

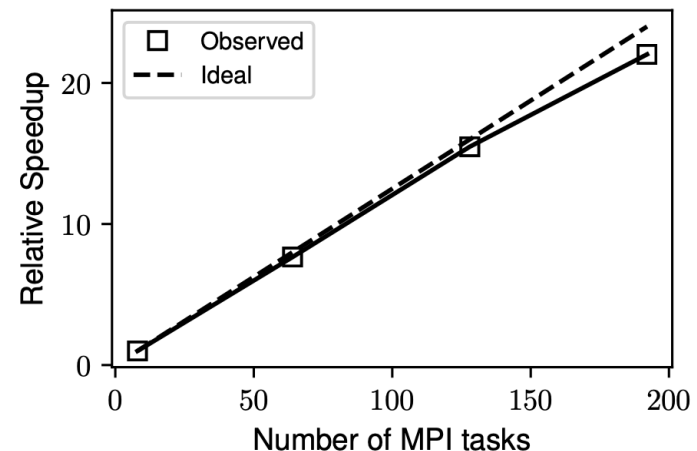
NV Diamond supercells

Supercell	OEFE		LAPW+lo	
	C	(c, N)	C	(c, N)
$2 \times 2 \times 2$	0.19	(0.008, 12)	0.32	(0.02, 16)
$3 \times 3 \times 3$	1.6	(0.071, 16)	15.1	(0.84, 18)
$4 \times 4 \times 4$	16.1	(0.46, 31)	-	-

Cu supercells

Supercell	OEFE		LAPW+lo	
	C	(c, N)	C	(c, N)
$2 \times 2 \times 2$	0.92	(0.033, 24)	0.145	(0.004, 32)
$3 \times 3 \times 3$	20.6	(0.55, 36)	6.46	(0.144, 45)
$4 \times 4 \times 3$	93.12	(2.6, 35)	50.63	(0.92, 55)
$4 \times 4 \times 4$	250.0	(6.0, 41)	-	-

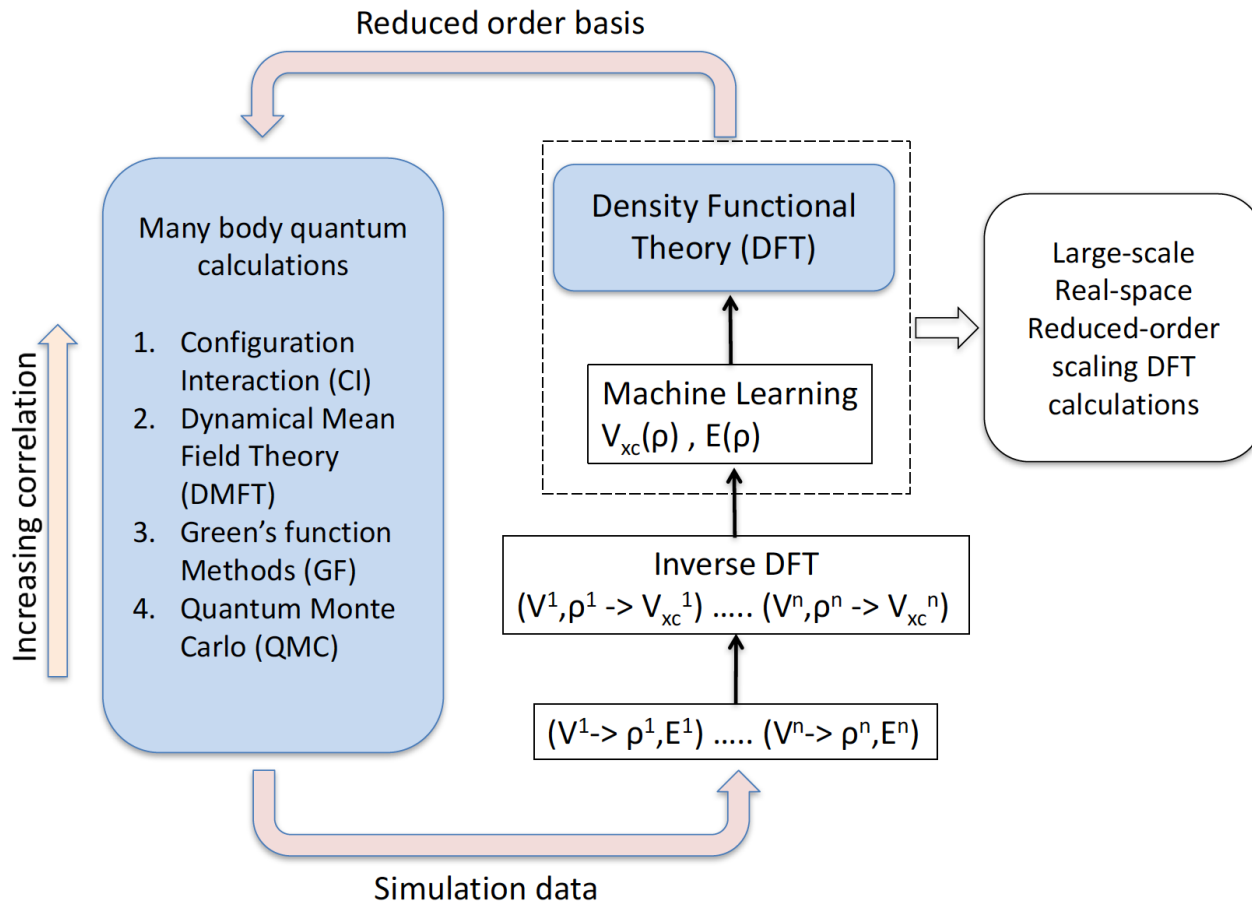
Parallel efficiency Divacancy in 2x2x2 SiC



Large-scale quantum accuracy calculations

- Accuracy of exchange correlation functionals in DFT is not satisfactory for strongly correlated electrons.

Can we address this without sacrificing the efficiency of DFT calculations?

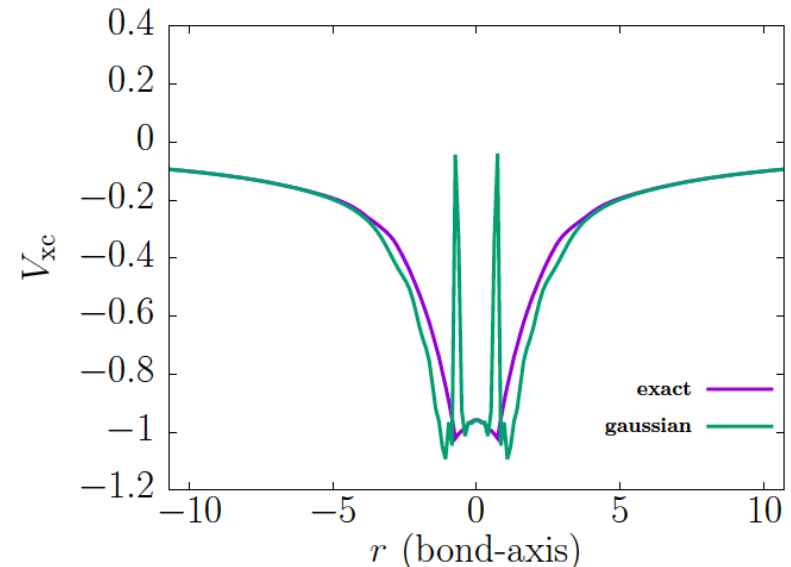
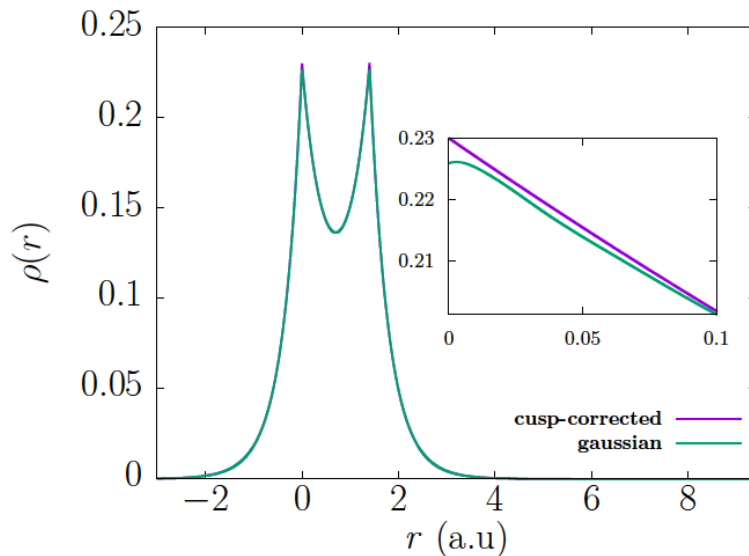


Inverse DFT

Inverse DFT: Determining $V_{xc}(\mathbf{r})$ given $\rho(\mathbf{r})$

Remained an open problem for 25 years – Numerically very challenging:

- Attempts included iterative approaches, constrained optimization approaches
 - ❖ Spurious oscillations; non-unique solutions
 - ❖ Key issues:
 - (i) Many-body QM calculations conducted in incomplete basis (wrong asymptotics)
 - (ii) Inversion in an incomplete basis



Inverse DFT

(Kanungo, Zimmerman & Gavini, Nature Communications **10** 4497 (2019))

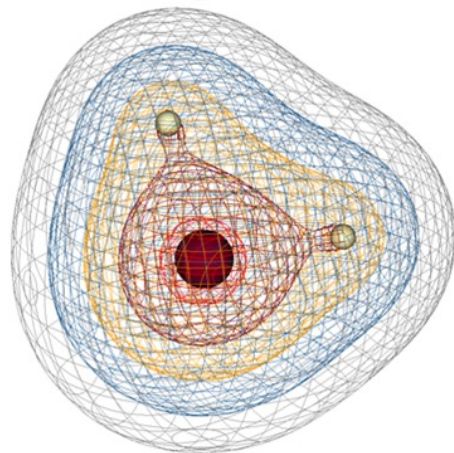
Recent breakthrough in an accurate solution to the inverse DFT problem. Demonstrated on molecular systems that are both weakly and strongly correlated.

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N_e})$$

Many-body wavefunction

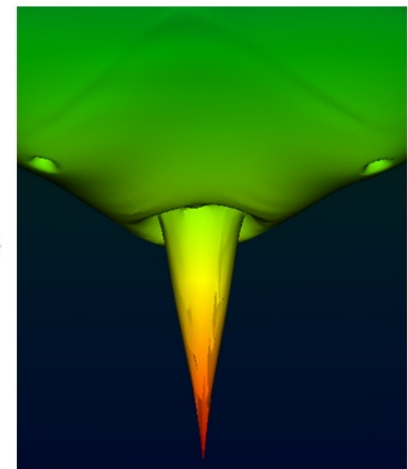
Configuration
Interaction

$$\rho(\mathbf{r}) \xleftrightarrow{v_{xc}[\rho(\mathbf{r})]} v_{xc}(\mathbf{r})$$



**Inverse
DFT**

- ❖ PDE-Constrained Optimization
- ❖ Complete Finite-element basis
- ❖ Cusp Correction
- ❖ Correct Asymptotics



Inverse DFT – Key Ideas

- PDE constrained optimization

$$\arg \min_{v_{xc}(\mathbf{r})} \int w(\mathbf{r}) (\rho_{data}(\mathbf{r}) - \rho(\mathbf{r}))^2 d\mathbf{r}$$

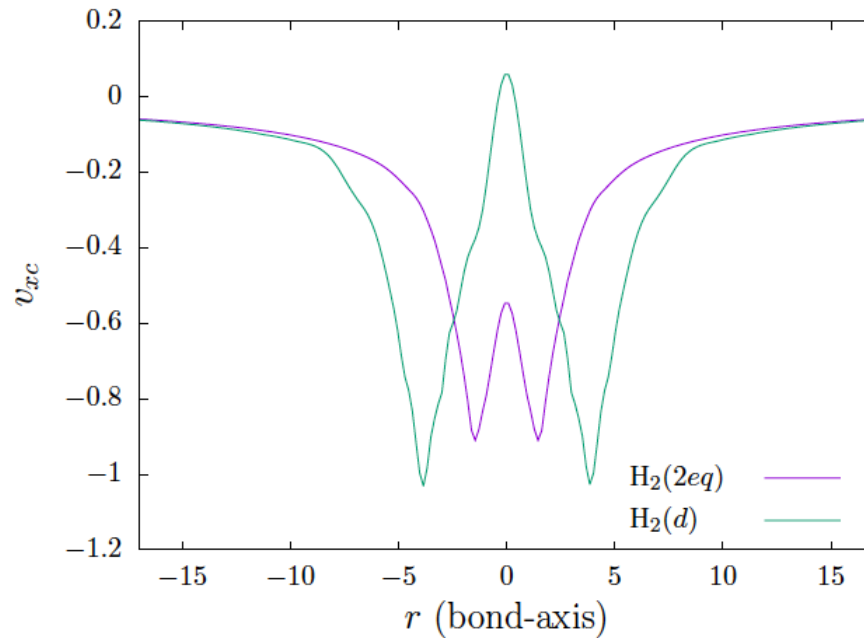
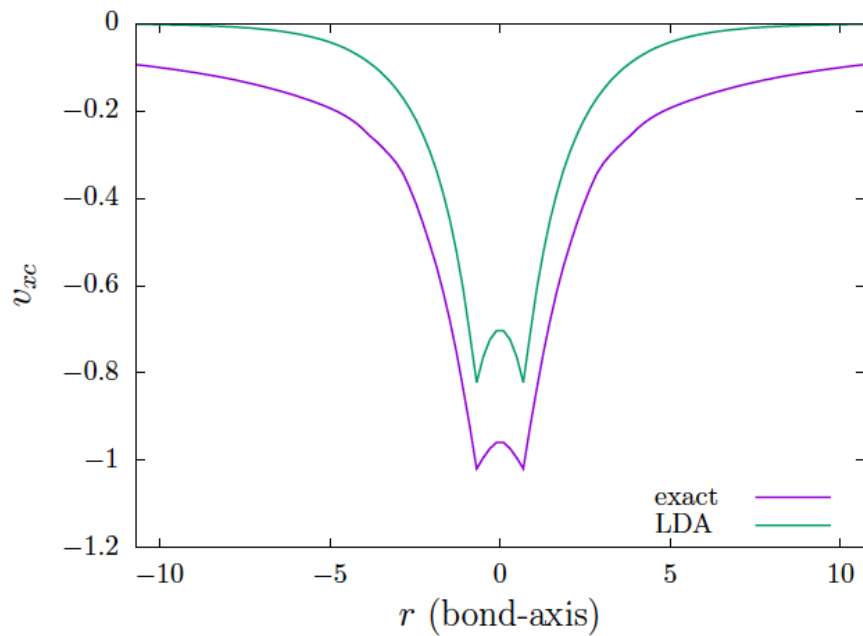
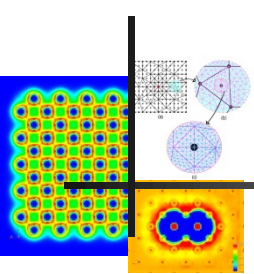
subject to

$$\left(-\frac{1}{2} \nabla^2 + v_{ext}(\mathbf{r}) + v_H(\mathbf{r}) + v_{xc}(\mathbf{r}) \right) \psi_i = \epsilon_i \psi_i, \quad \int |\psi_i(\mathbf{r})|^2 d\mathbf{r} = 1.$$

- Higher-order FE basis for discretization – ensures completeness
- Cusp correction: $\Delta\rho(\mathbf{r}) = \rho_{FE}^{DFT}(\mathbf{r}) - \rho_G^{DFT}(\mathbf{r})$
- Far-field asymptotics: Start with a guess for $V_{xc}(\mathbf{r})$ with correct far-field asymptotics and use homogeneous Dirichlet boundary conditions on the adjoint fields.



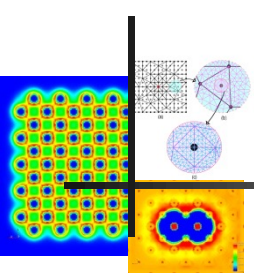
Inverse DFT using ab-initio correlated densities



Good agreement between HOMO eigenvalue and $-I_p$



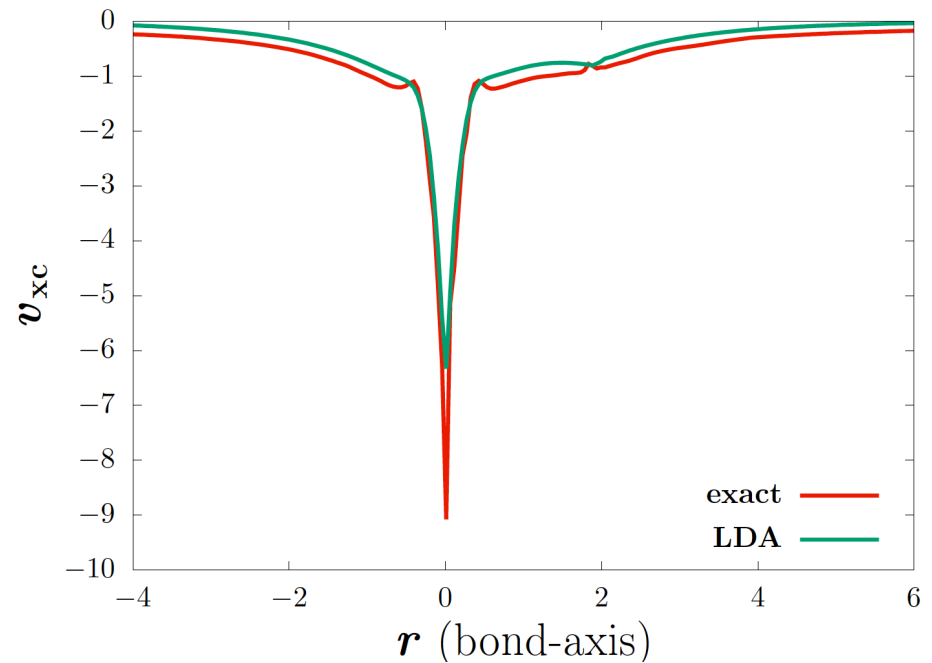
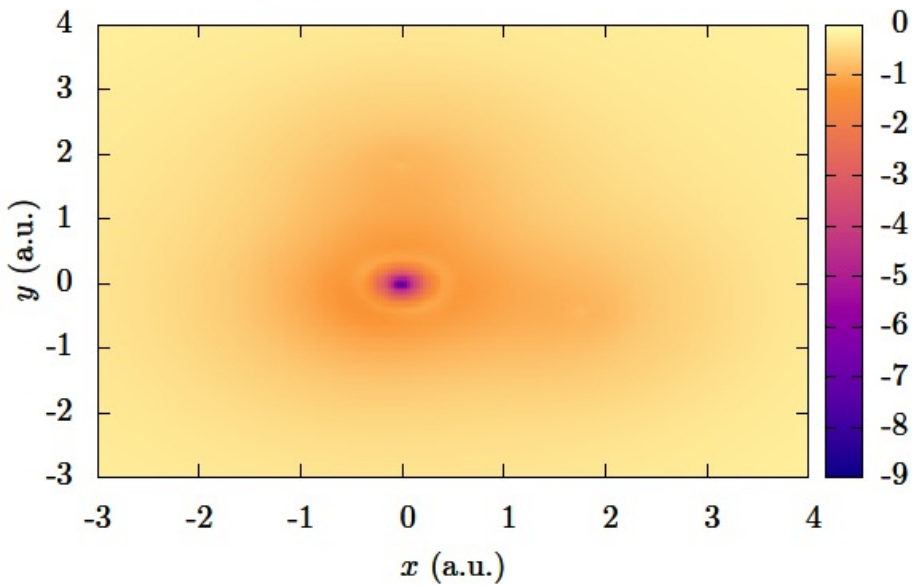
Inverse DFT using ab-initio correlated densities



Materials system H₂O molecule

Exact V_{xc}

(data from full CI calculation)

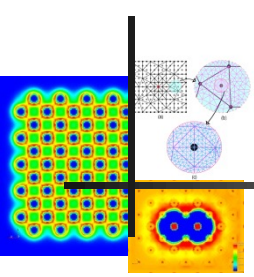


Verification of Koopmans' theorem:

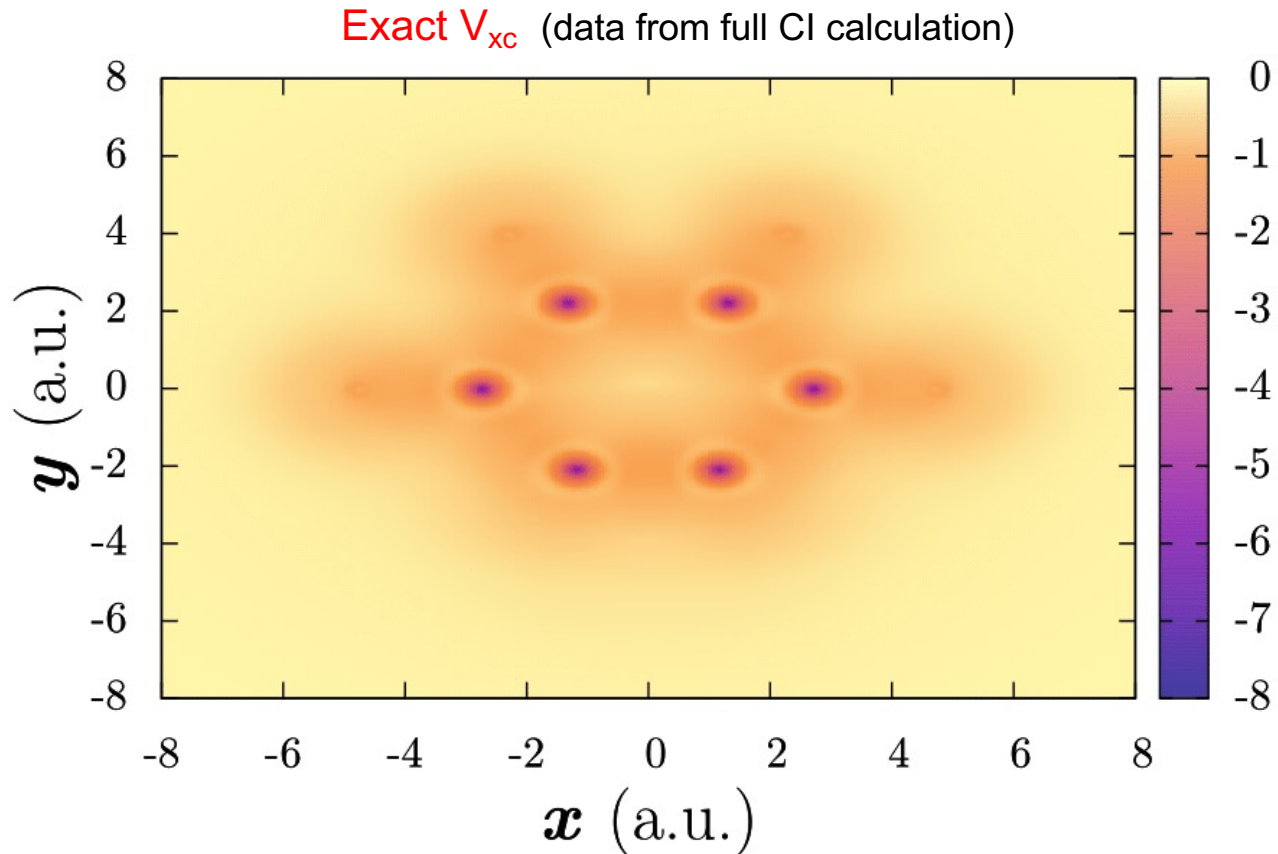
HOMO eigenvalue = -0.452 Ha; $-I_p = -0.454$ Ha



Inverse DFT using ab-initio correlated densities



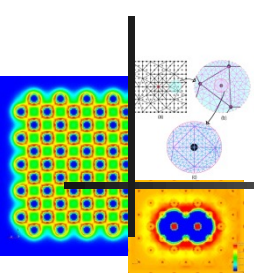
Materials system C_6H_4 (*ortho*-Benzyne) – strongly correlated system



Verification of Koopmans' theorem:

HOMO eigenvalue = -0.354 Ha; $-I_p = -0.355$ Ha

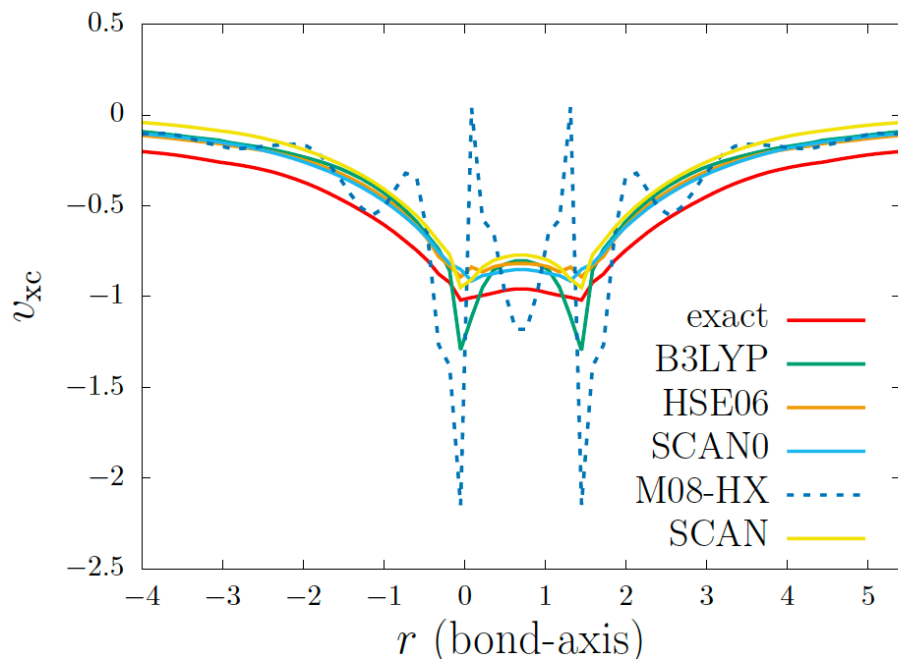




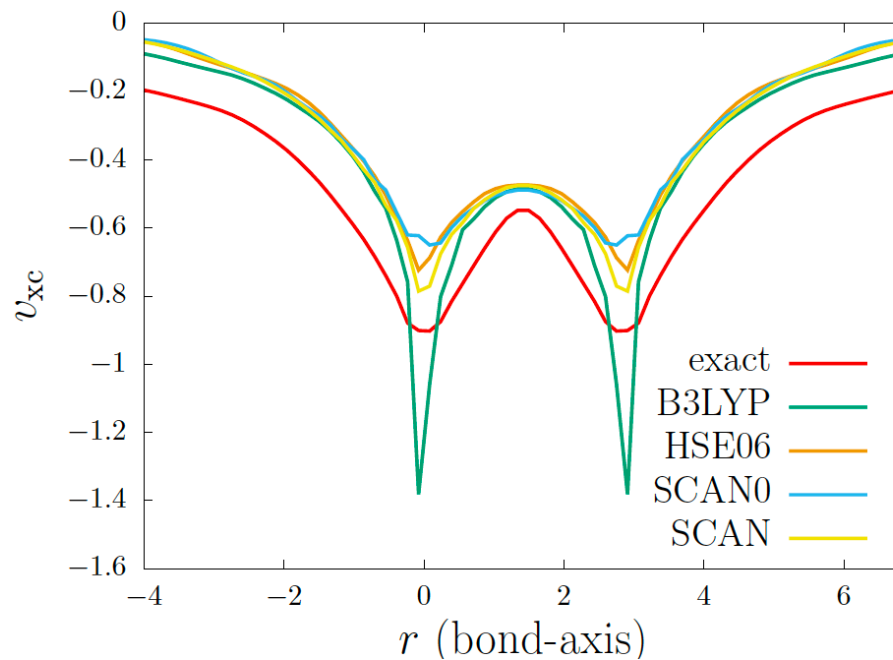
Comparison of exact XC with model XC potentials

(Kanungo, Zimmerman & Gavini, J. Phys. Chem. Lett. **12** 12012 (2021))

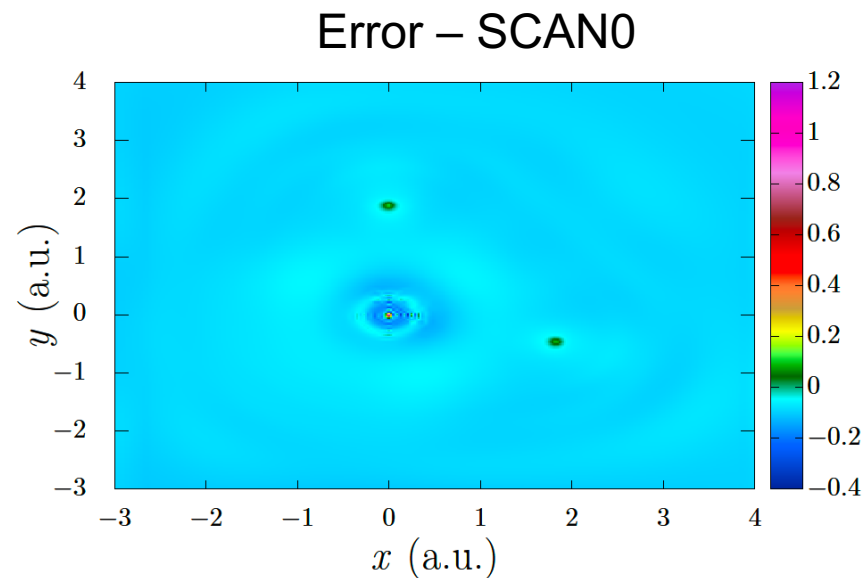
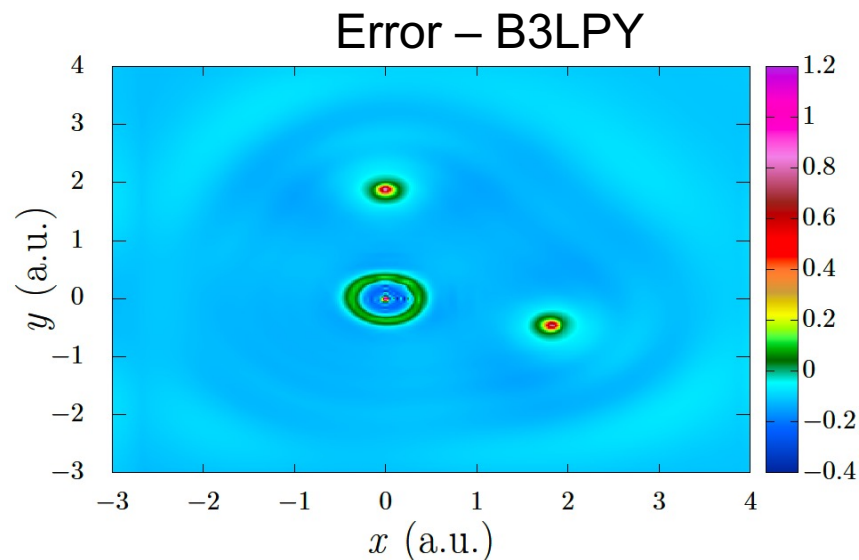
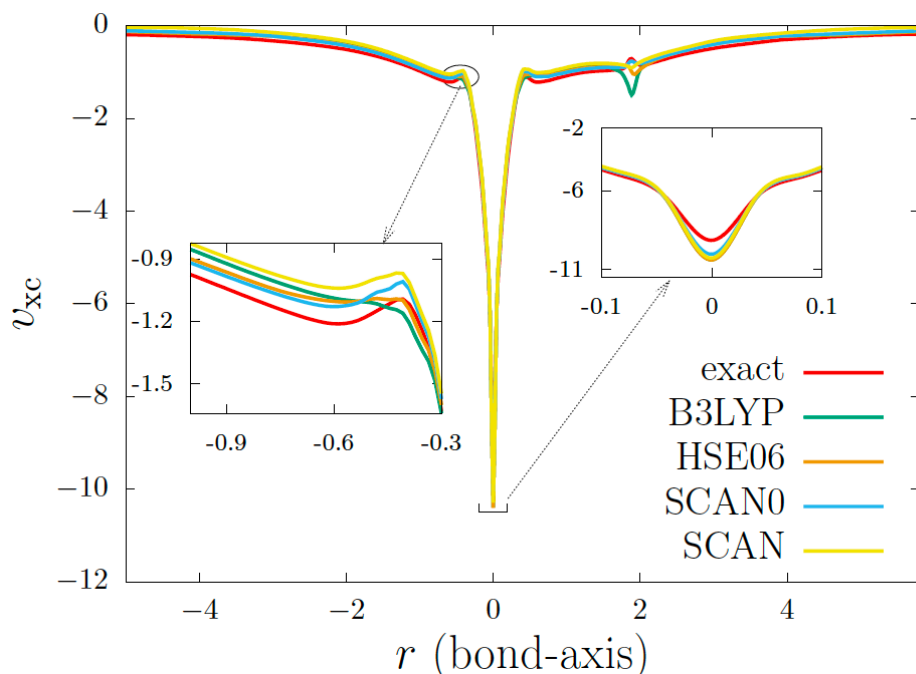
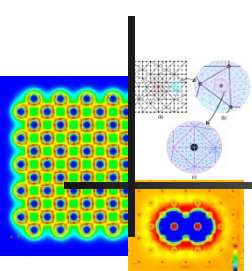
H₂ – equilibrium bond length



H₂ – twice equilibrium bond length



Comparison of exact XC with model XC potentials (H₂O)



Comparison of exact XC with model XC potentials (H₂O)

Relative errors in XC potentials – $O(10^{-1} - 10^{-0})$

Model	H ₂ (eq)		H ₂ (0.8eq)		H ₂ (2eq)		LiH		H ₂ O		C ₆ H ₄	
	e ₁	e ₂	e ₁	e ₂	e ₁	e ₂	e ₁	e ₂	e ₁	e ₂	e ₁	e ₂
B3LYP	0.192	1.657	0.216	1.125	0.272	1.377	0.107	1.862	0.041	0.309	0.049	0.071
HSE06	0.162	0.285	0.202	0.309	0.277	0.448	0.094	1.398	0.042	0.311	–	–
SCAN0	0.145	0.240	0.181	0.167	0.273	0.302	0.087	1.010	0.030	0.227	0.045	0.098
SCAN	0.233	0.427	0.207	0.429	0.254	0.517	0.092	1.391	0.044	0.298	–	–
M08-HX	0.248	4.784	0.249	5.873	0.291	5.109	0.094	1.800	0.065	0.545	–	–
PBE	0.261	1.100	0.241	0.857	0.262	1.251	0.119	2.013	0.058	0.427	0.059	0.105
PW92	0.288	0.267	0.267	0.278	0.297	0.420	0.152	0.264	0.133	0.354	0.145	0.407

$$e_1 = \frac{\|\rho_{\text{data}} \delta v_{\text{xc}}\|_{L_2}}{\|\rho_{\text{data}} v_{\text{xc}}^{\text{exact}}\|_{L_2}}$$

$$e_2 = \frac{\|\rho_{\text{data}} |\nabla \delta v_{\text{xc}}|\|_{L_2}}{\|\rho_{\text{data}} |\nabla v_{\text{xc}}^{\text{exact}}|\|_{L_2}}$$

Relative errors in density – $O(10^{-3} - 10^{-2})$

Model	H ₂ (eq)		H ₂ (0.8eq)		H ₂ (2eq)		LiH		H ₂ O		C ₆ H ₄	
	f ₁	f ₂	f ₁	f ₂	f ₁	f ₂	f ₁	f ₂	f ₁	f ₂	f ₁	f ₂
B3LYP	0.011	0.025	0.010	0.022	0.039	0.045	0.006	0.009	0.003	0.002	0.004	0.004
HSE06	0.004	0.006	0.004	0.006	0.054	0.069	0.005	0.006	0.002	0.001	–	–
SCAN0	0.004	0.006	0.003	0.006	0.061	0.082	0.003	0.006	0.001	0.001	0.002	0.001
SCAN	0.006	0.012	0.005	0.011	0.043	0.054	0.004	0.008	0.002	0.001	–	–
M08-HX	0.018	0.047	0.019	0.049	0.062	0.080	0.006	0.011	0.002	0.003	–	–
PBE	0.010	0.020	0.010	0.017	0.031	0.034	0.007	0.010	0.003	0.002	0.004	0.004
PW92	0.023	0.025	0.025	0.027	0.057	0.077	0.022	0.024	0.010	0.014	0.014	0.018

$$f_1 = \frac{\|\rho_{\text{data}}^{\text{exact}} - \rho_{\text{data}}^{\text{model}}\|_{L_2}}{\|\rho_{\text{data}}^{\text{exact}}\|_{L_2}}$$

$$f_2 = \frac{\|\nabla(\rho_{\text{data}}^{\text{exact}} - \rho_{\text{data}}^{\text{model}})\|_{L_2}}{\|\nabla \rho_{\text{data}}^{\text{exact}}\|_{L_2}}$$



Learning E_{xc}

➤ Express $E_{xc}[\rho] = \int e_{xc}[\rho](\mathbf{r})d\mathbf{r}$ where $e_{xc}[\rho](\mathbf{r})d\mathbf{r}$: energy density

➤ Local/semi-local models

➤ $e_{xc}^{\text{LDA}}(\mathbf{r}) = e_{xc}^{\text{ML}}[\rho](\mathbf{r})$

➤ $e_{xc}^{\text{GGA}}(\mathbf{r}) = e_{xc}^{\text{ML}}[\rho, \nabla\rho](\mathbf{r})$

➤ $e_{xc}^{\text{near-sighted}}(\mathbf{r}) = e_{xc}^{\text{ML}}[\rho, \nabla\rho, \nabla^2\rho, g](\mathbf{r})$ $g(\mathbf{r}) = \int \rho(\mathbf{r}')k(\mathbf{r}, \mathbf{r}')d\mathbf{r}'$

➤ Since we do not have exact $e_{xc}[\rho](\mathbf{r})$

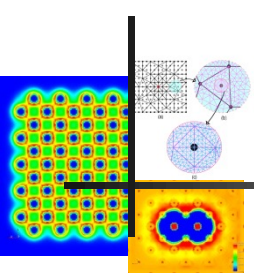
➤ $\mathcal{L} : \sum_I \int (v_{xc}^{\text{exact}}[\rho_I](\mathbf{r}) - v_{xc}^{\text{ML}}[\rho_I](\mathbf{r}))^2 d\mathbf{r} + c \sum_I (E_{xc}^{\text{exact},I} - E_{xc}^{\text{ML},I})^2$

➤ $v_{xc}^{\text{ML}}(\mathbf{r}) = \frac{\delta E_{xc}^{\text{ML}}[\rho]}{\delta\rho(\mathbf{r})} = \int \frac{\delta e_{xc}^{\text{ML}}(\mathbf{r}')}{\delta\rho(\mathbf{r})} d\mathbf{r}'$

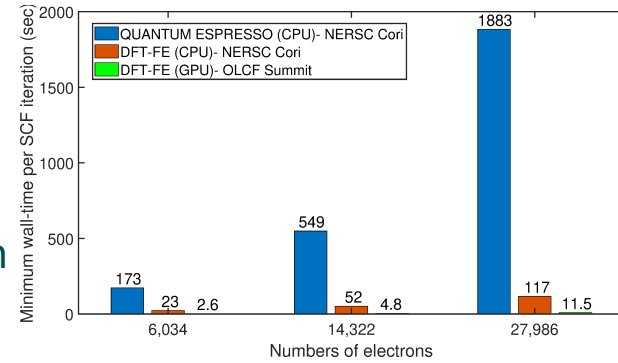
➤ Learning is on $e_{xc}[\rho](\mathbf{r})$, but optimization is on $v_{xc}(\mathbf{r})$



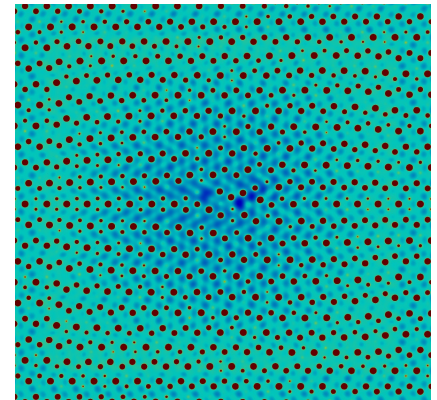
Concluding remarks



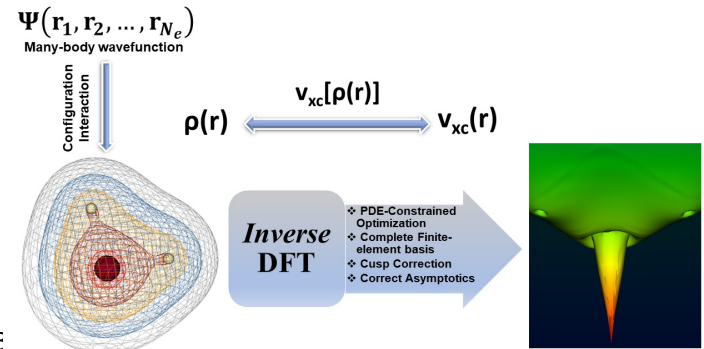
- Large-scale Kohn-Sham DFT calculations
 - ❖ Higher-order FE discretization
 - ❖ Algorithmic and HPC aspects of KS eigenvalue problem
 - ❖ Fast and accurate large-scale calculations possible
 - ❖ Extensions to All-electron calculations

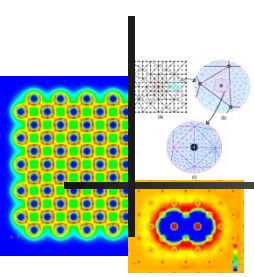


- Some applications
 - ❖ Energetics of pyramidal I & II dislocations in Mg
 - ❖ Electronic structure of DNA molecules
 - ❖ Spin Hamiltonian parameters



- Tackling XC approximations in DFT
 - ❖ Inverse DFT to compute exact XC potentials
 - ❖ System identification / M-L





THANK YOU!

