

Introduction to density functional theory (DFT)

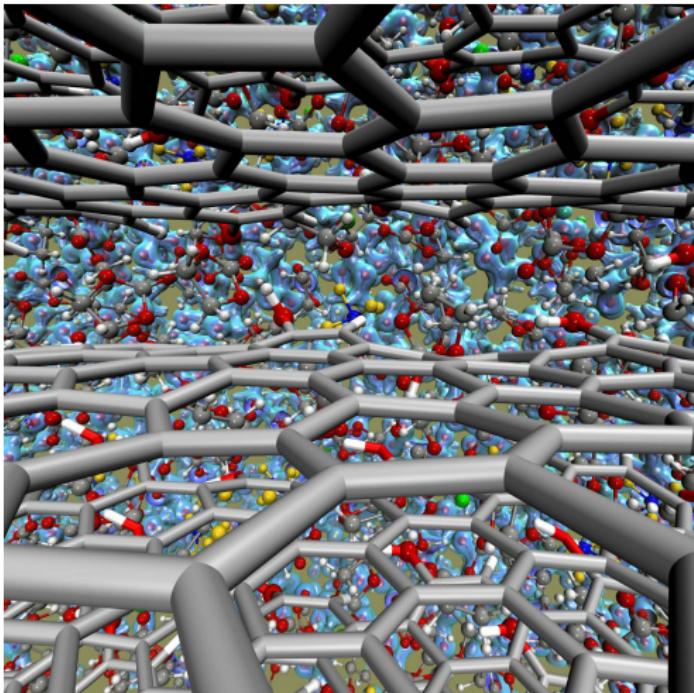
Phanish Suryanarayana



QM/MM and ab initio Molecular Dynamics Summer School
Institute for Computational Molecular Science Education

Ab initio calculations

- Given positions of atoms, calculate the desired properties, e.g., mechanical, electronic, optical, and thermal
- No modeling or empirical input
- Isolated (e.g. molecules) or extended (e.g. crystals, surfaces, and wires) systems
- Typically performed when **electronic structure** (i.e., bonding) plays an important role



Anode electrolyte system (John E Pask)

Schrödinger equation?

- Schrödinger equation (time-independent)

$$H\Psi = E\Psi$$

- H : Hamiltonian, Ψ : Wavefunction, E : Energy
- Linear eigenvalue problem
- Lowest eigenvalue and eigenfunction
- $\Psi(x_1, x_2, \dots, x_{N_e}) \in \mathbb{R}^{3N_e}$, N_e : # electrons
- System with $N_e = 10$ (e.g. 2 Boron atoms)
 - 10 grid points in each direction
 - 10^{30} grid points!
- Computationally intractable



Erwin Schrödinger (Wikipedia)

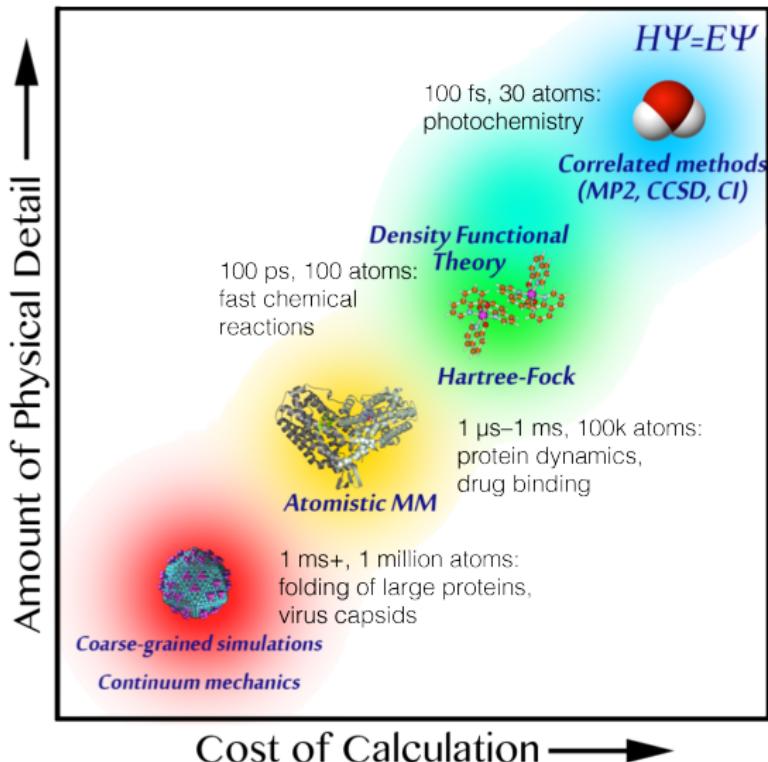
- Density functional theory (DFT)

- Ground state depends on the electron density
- $\rho \in \mathbb{R}^3$
- Kohn-Sham formalism: Fictitious non-interacting system
- High accuracy to cost ratio
- Nobel prize in Chemistry (1998): Walter Kohn



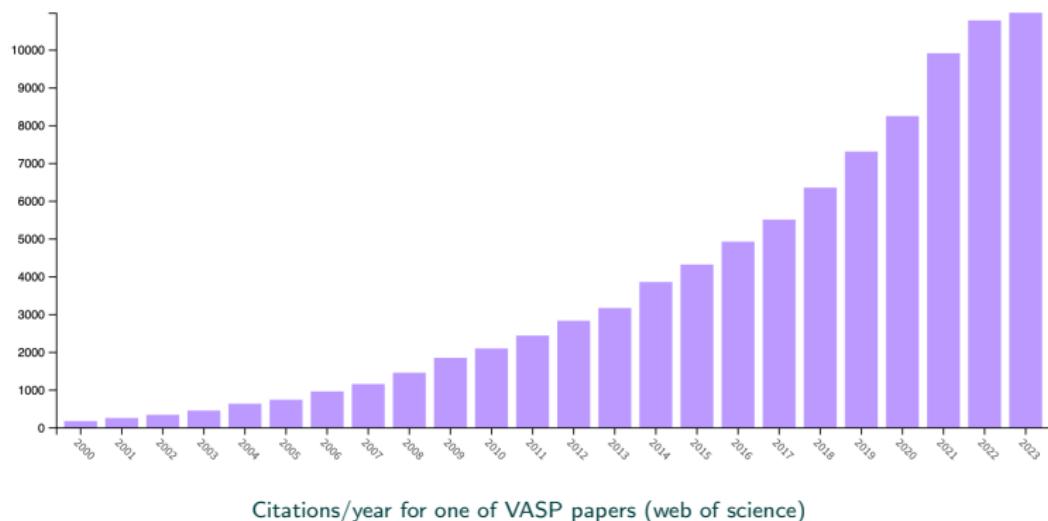
Walter Kohn (Wikipedia)

Cost vs Accuracy



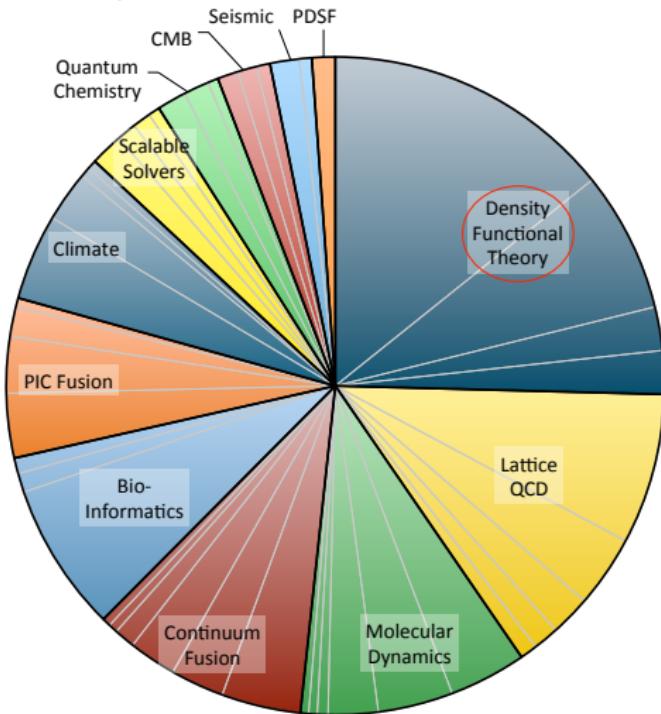
<https://education.molssi.org/mm-tools/01-introduction/index.html>

Popularity of Kohn-Sham DFT



- One of the papers for a plane-wave DFT code VASP
- Other popular DFT codes, e.g., Quantum Espresso, ABINIT, ...

Popularity of Kohn-Sham DFT



NERSC systems usage in 2014 (NERSC workload analysis report)

Energy Functional

$$\mathcal{F}(\Psi, \mathbf{g}, \mathbf{R}) = T_s(\Psi, \mathbf{g}) + E_{xc}(\rho) + K(\Psi, \mathbf{g}, \mathbf{R}) + E_{el}(\rho, \mathbf{R}) - TS(\mathbf{g})$$

where

$$\rho(\mathbf{x}) = 2 \sum_{n=1}^{N_s} g_n \psi_n^2(\mathbf{x}) \quad \text{Electron density}$$

$\Psi = \{\psi_1, \psi_2, \dots, \psi_{N_s}\}$: Orbitals. $\mathbf{g} = \{g_1, g_2, \dots, g_{N_s}\}$: Occupations.

$\mathbf{R} = \{\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N\}$: Atomic positions. T : Electronic temperature.

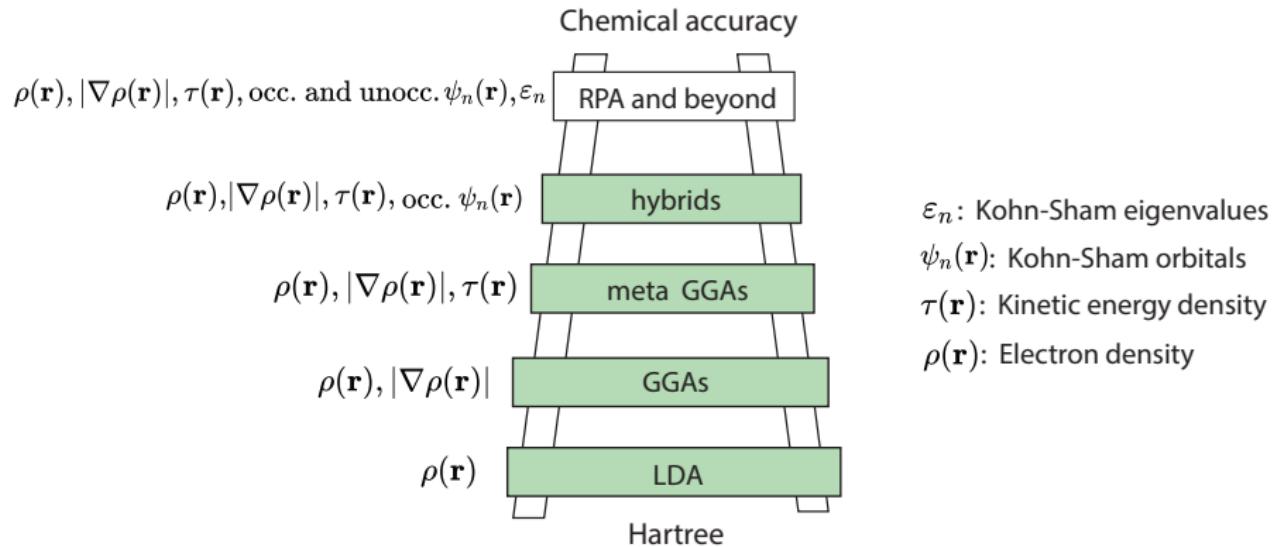
$T_s(\Psi, \mathbf{g})$: Electronic kinetic energy. $E_{xc}(\rho)$: Exchange-correlation energy.

$K(\Psi, \mathbf{g}, \mathbf{R})$: Nonlocal pseudopotential energy. $E_{el}(\rho, \mathbf{R})$: Electrostatic energy.

$S(\mathbf{g})$: Electronic entropy arising from partial orbital occupations.

- Fictitious systems of non-interacting electrons
- Not a functional of the electron density alone
- $E_{xc}(\rho)$ modeled. $E_{xc}(\rho, \nabla\rho, \dots, \Psi)$ models exist.

Exchange-correlation: Jacob's ladder



- Higher rungs → Better accuracy (generally)
- Higher rungs → increased cost.

Exchange-correlation: Alphabet (Peter Elliott)



- Large number of variants
- Machine learned functionals being developed.

Kohn-Sham DFT: Variational Problem

- Kinetic energy of non-interacting electrons

$$T_s(\Psi, \mathbf{g}) = - \sum_{n=1}^{N_s} g_n \int \psi_n(\mathbf{x}) \nabla^2 \psi_n(\mathbf{x}) d\mathbf{x}$$

- Total electrostatic energy $E_{el}(\rho, \mathbf{R})$:

$$\frac{1}{2} \int \int \frac{\rho(\mathbf{x})\rho(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d\mathbf{x} d\mathbf{x}' + \sum_{J=1}^N \int \rho(\mathbf{x}) V_J(\mathbf{x}, \mathbf{R}_J) d\mathbf{x} + \frac{1}{2} \sum_{I=1}^N \sum_{\substack{J=1 \\ J \neq I}}^N \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}$$

Ground-state

$$\mathcal{F}_0 = \min_{\mathbf{R}} \mathcal{F}^*(\mathbf{R})$$

where

$$\mathcal{F}^*(\mathbf{R}) = \min_{\Psi, \mathbf{g}} \mathcal{F}(\Psi, \mathbf{g}, \mathbf{R}) \quad s.t. \quad \int \psi_i(\mathbf{x}) \psi_j(\mathbf{x}) d\mathbf{x} = \delta_{ij}, \quad 2 \sum_{n=1}^{N_s} g_n = N_e$$

Electrostatics: Real-space formulation

- Local reformulation of the electrostatics

$$\begin{aligned} E_{el}(\rho, \mathbf{R}) &= \max_{\phi} \left\{ -\frac{1}{8\pi} \int |\nabla \phi(\mathbf{x})|^2 d\mathbf{x} + \int (\rho(\mathbf{x}) + b(\mathbf{x}, \mathbf{R}))\phi(\mathbf{x}) d\mathbf{x} \right\} \\ &- E_{self}(\mathbf{R}) + E_c(\mathbf{R}) \end{aligned}$$

b : pseudocharge density. $E_{self}(\mathbf{R})$: self energy. $E_c(\mathbf{R})$: correction due to overlapping pseudocharges.

- Cost reduces from $\mathcal{O}(N^2)$ to $\mathcal{O}(N)$
- Compatible with the symmetry of the system
- Simplifies taking of derivatives, e.g., force, stress, phonon, etc

- Suitable space of solutions

$$X = \left\{ \Psi \in (H_0^1(\Omega))^{N_s} \mid \langle \psi_i, \psi_j \rangle_{(L^2(\Omega), L^2(\Omega))} = \delta_{ij}, \quad i, j = 1, 2, \dots, N_s \right\}$$

- Existence of solutions

Theorem

\mathcal{F} possesses a minimizer in X

Electronic ground state and forces

- Governing (Euler-Lagrange) eigenproblem for the electronic ground-state

$$\left(\mathcal{H} \equiv -\frac{1}{2}\nabla^2 + V_{xc} + \phi + V_{nl} \right) \psi_n = \lambda_n \psi_n, \quad n = 1, 2, \dots, N_s,$$

$$g_n = \left(1 + \exp \left(\frac{\lambda_n - \lambda_f}{k_B T} \right) \right)^{-1}, \quad \text{where } \lambda_f \text{ is s.t. } 2 \sum_{n=1}^{N_s} g_n = N_e$$

$$\rho(\mathbf{x}) = 2 \sum_{n=1}^{N_s} g_n \psi_n^2(\mathbf{x}), \quad -\frac{1}{4\pi} \nabla^2 \phi(\mathbf{x}, \mathbf{R}) = \rho(\mathbf{x}) + b(\mathbf{x}, \mathbf{R}),$$

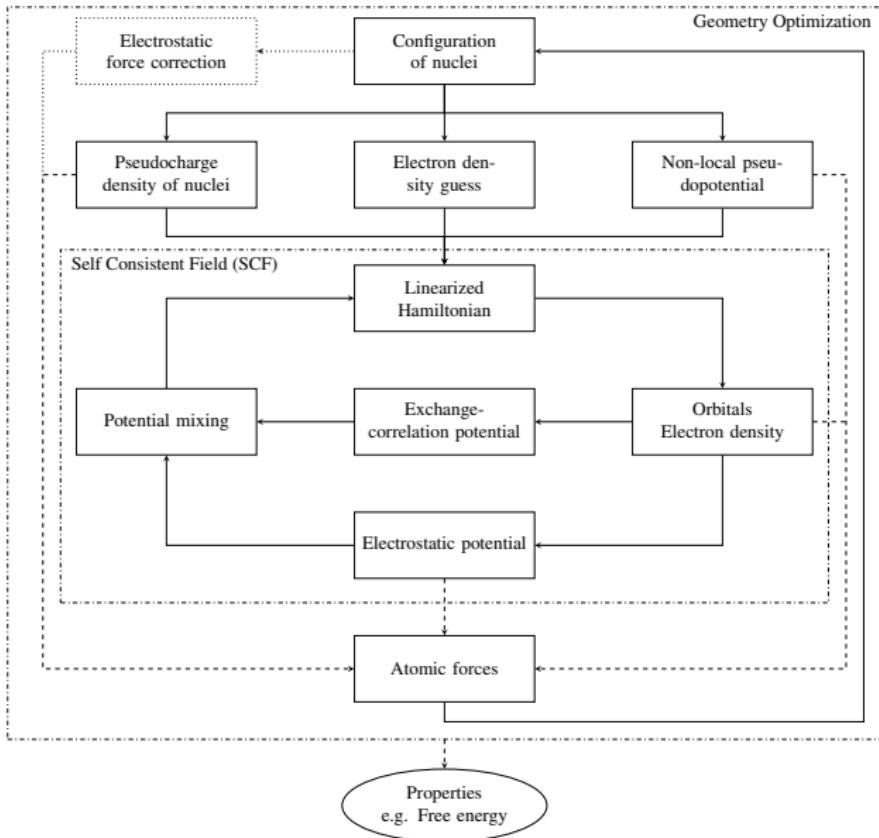
\mathcal{H} : Hamiltonian operator, V_{xc} : Exchange-correlation potential

V_{nl} : Non-local pseudopotential operator, λ_f : Fermi energy.

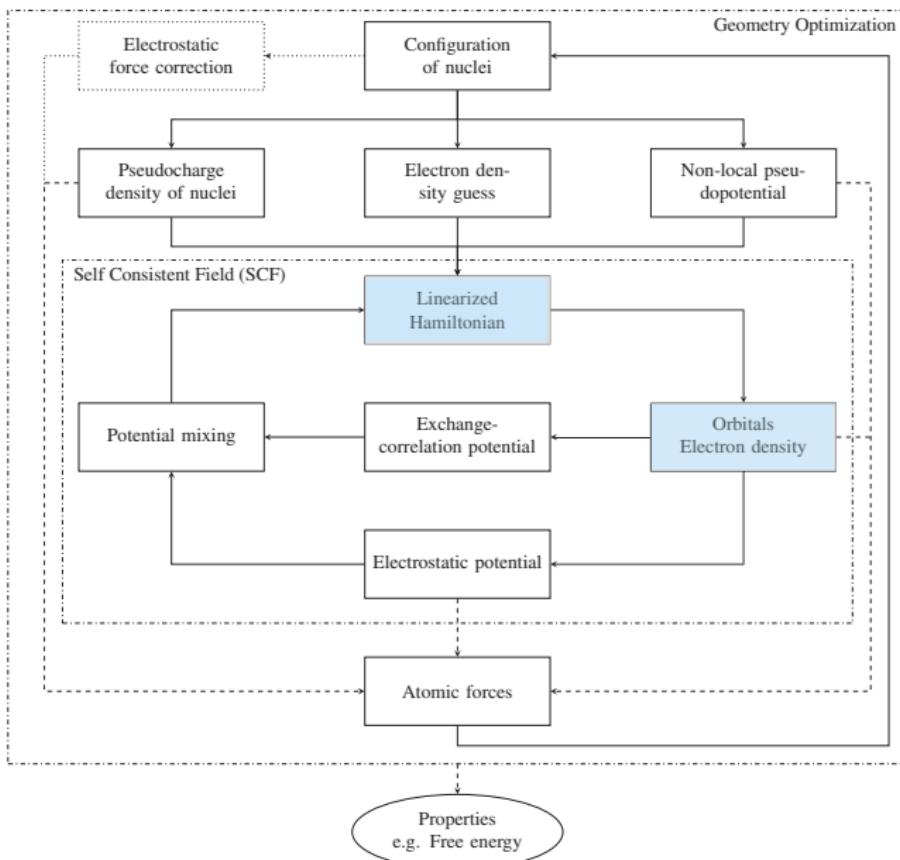
- Atomic forces

$$\begin{aligned} \mathbf{f}_J &= -\frac{\partial \mathcal{F}^*(\mathbf{R})}{\partial \mathbf{R}_J} = \int \nabla b_J(\mathbf{x}, \mathbf{R}_J) \phi(\mathbf{x}, \mathbf{R}) d\mathbf{x} + \mathbf{f}_{J,c}(\mathbf{R}) \\ &\quad - 4 \sum_{n=1}^{N_s} g_n \sum_{lm} \gamma_{Jl} \left(\int \psi_n(\mathbf{x}) \chi_{Jlm}(\mathbf{x}, \mathbf{R}_J) d\mathbf{x} \right) \left(\int \nabla \psi_n(\mathbf{x}) \chi_{Jlm}(\mathbf{x}, \mathbf{R}_J) d\mathbf{x} \right). \end{aligned}$$

Outline of DFT calculation



Electron density calculation



Eigensolver

- Electron density calculated using partial diagonalization, e.g., CheFSI (Zhou et. al., 2006)

- Chebyshev polynomial filtering ($\mathcal{O}(N_s N_d)$)

$$\Psi_f = p_m(\mathbf{H})\Psi$$

- Solve projected (onto the filtered basis Ψ_f) eigenvalue problem ($\mathcal{O}(N_s^2 N_d) + \mathcal{O}(N_s^3)$):

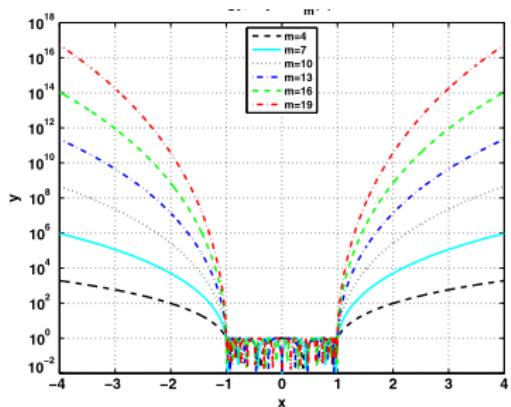
$$\mathbf{H}_s \mathbf{y}_n = \lambda_n \mathbf{M}_s \mathbf{y}_n, \quad n = 1, 2, \dots, N_s$$

- Subspace rotation ($\mathcal{O}(N_s^2 N_d)$)

$$\Psi = \Psi_f \mathbf{Y}$$

where the columns of the matrix $\mathbf{Y} \in \mathbb{R}^{N_s \times N_s}$ contain the eigenvectors \mathbf{y}_n .

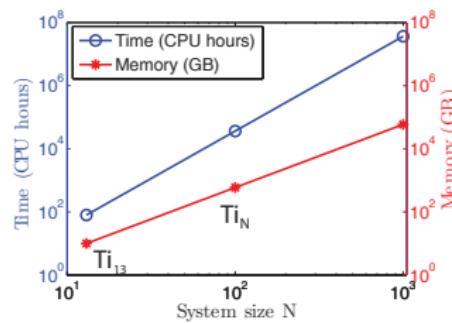
- Degree of Chebyshev polynomial ~ 20
- Partial diagonalization
- Can utilize the orbitals from the previous SCF iteration as guess.
- Minimizes the number of global communications.
- Scales as $\mathcal{O}(N^3)$. However, extremely small prefactor.



Challenges in Kohn-Sham DFT

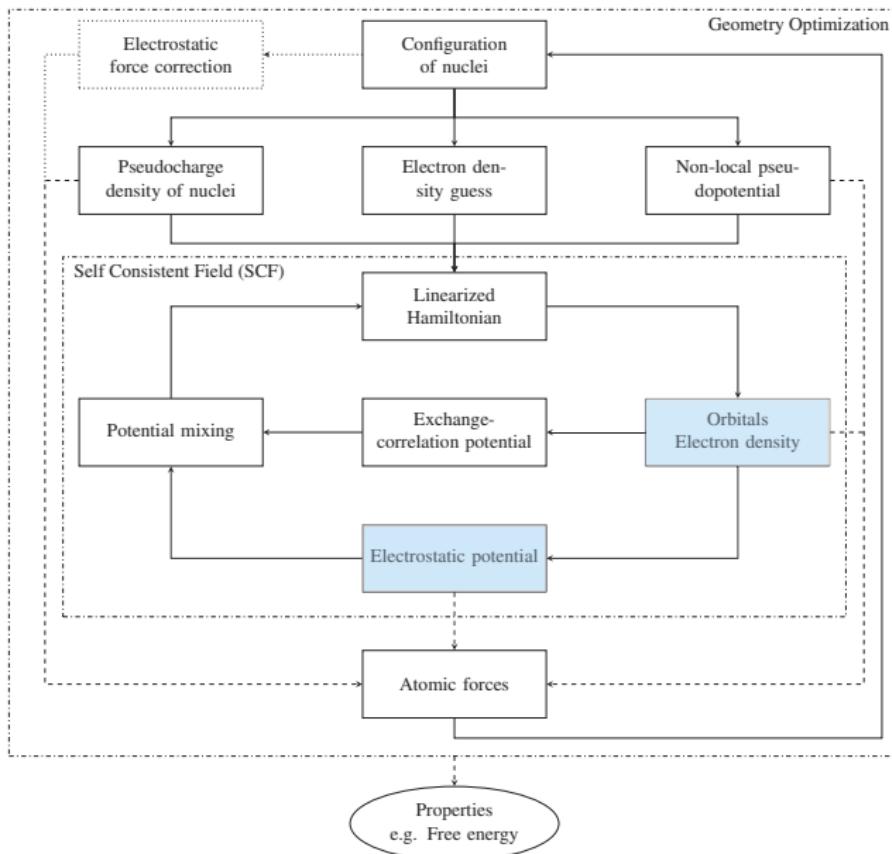
The DFT problem is challenging!

- Nonlinear problem
- Accuracy
 - ~ 0.00001 relative accuracy.
 - Material properties: Derivatives of energy
 - Pseudopotential
 - Exchange-correlation, ...
- Large number of degrees of freedom
 - $O(N^2)$ storage
 - $\mathcal{O}(400 - 30,000)$ basis functions/atom
 - Basis functions per atom: ~ 1000
 - Storage for 10,000 atoms: 1500 GB
- Orthonormal constraint on the orbitals
 - $\mathcal{O}(N^3)$ computational cost
 - 100 atom calculation: 10 min (fast!)
 - 10,000 atom calculation: $\sim 150,000$ hours



Extrapolated computational time and memory for Titanium nanostructures taken by ABINIT planewave code. The numbers for Ti_{13} are used as reference. Ti_{1000} would require 4000 CPU years and 60 TB memory!

Electrostatic potential calculation



Poisson problem: linear solver

- Need to solve the Poisson problem

$$-\frac{1}{4\pi} \nabla^2 \phi(\mathbf{x}, \mathbf{R}) = \rho(\mathbf{x}) + b(\mathbf{x}, \mathbf{R})$$

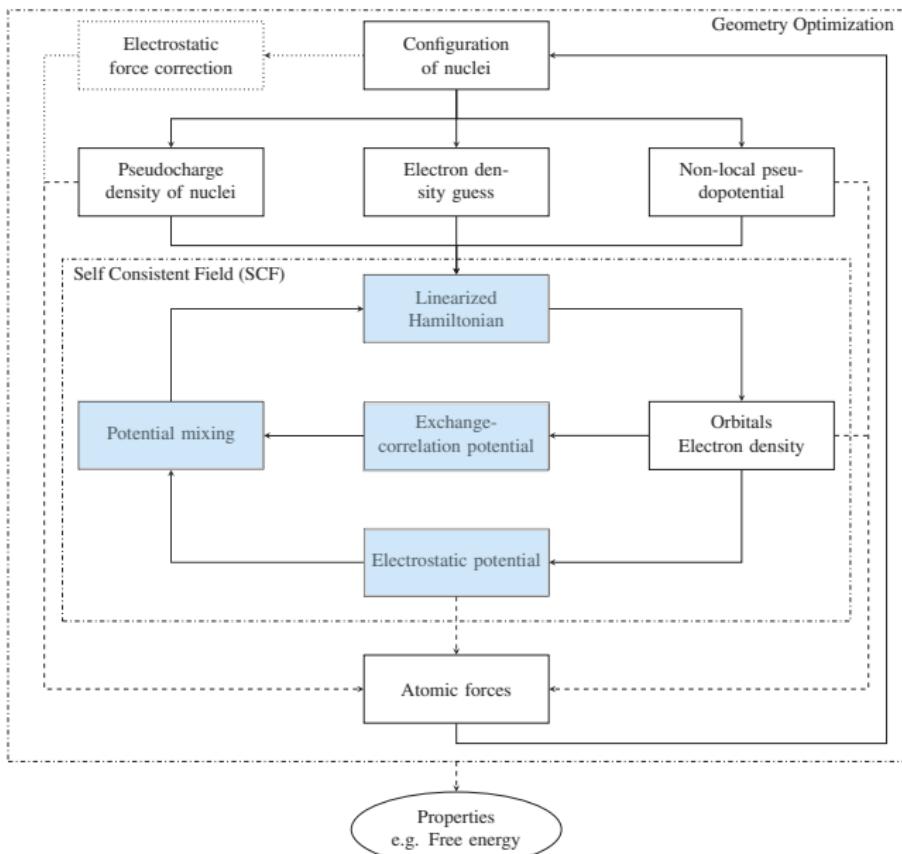
- Linear systems of the form

$$\mathbf{A}\mathbf{x} = \mathbf{b},$$

$$\mathbf{A} \in \mathbb{R}^{N \times N}, \quad \mathbf{x} \in \mathbb{R}^{N \times 1} \text{ and } \mathbf{b} \in \mathbb{R}^{N \times 1},$$

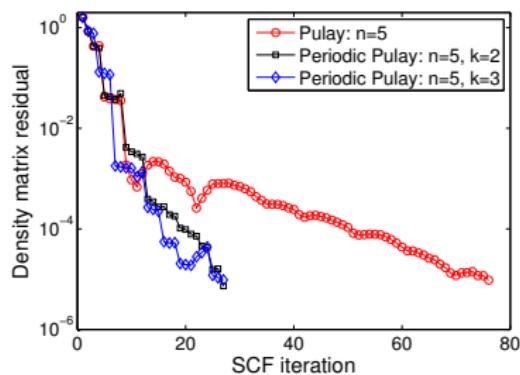
- Large number of linear solves for other variants of DFT
- Scalability of the linear solver?
- Multigrid scales as $\mathcal{O}(N)$, but challenging to parallelize to extremely large scales
- Basic fixed-point techniques? e.g. Richardson/Jacobi method.
 $(\mathbf{x}_{k+1} = \mathbf{x}_k + \omega(\mathbf{b} - \mathbf{A}\mathbf{x}_k))$
 - Large prefactors and poor scaling with system size
 - Unattractive compared to Krylov subspace approaches
 - Potential for massive parallelization!
- Idea: Significantly accelerate the convergence of the Richardson/Jacobi method, while maintaining its underlying computational locality and simplicity to the maximum extent possible.

Accelerating SCF fixed-point iteration: Mixing

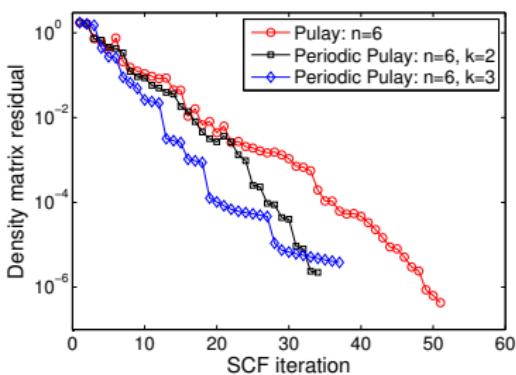


Mixing schemes

- SCF iteration: $\rho = f(g(\rho))$
- Cost of DFT simulations \propto to number of SCF fixed-point iterations
- Number of iterations increases with system size
- Pulay mixing (Anderson extrapolation) most commonly used approach for accelerating convergence



(a) Pd-bulk



(b) SiC-slab

Comparison of SCF convergence for the Pulay and Periodic Pulay mixing schemes.

**Thank You!
Questions?**