Multiscale Modeling of Atomistic Solids: Linear and Sublinear algorithms

Carlos J. García Cervera
cgarcia@math.ucsb.edu

Mathematics Department
University of California, Santa Barbara

- Joint work with Weinan E (Princeton U.), Jianfeng Lu (Duke U.), Jingrun Chen (UCSB), Weiqing Ren (NUS).
- Supported by NSF CAREER award, NSF DMS 1217315.

Outline

Outline

Outline

Outline

Outline

Kohn-Sham Density Functional Theory

$$E_{KS}[\{\psi_j\}] = 2\sum_{j=1}^{N}(-\frac{1}{2})\int \psi_j\,(\Delta\psi_j)\;dx + \frac{1}{2}\int\int \frac{\rho(x)\rho(y)}{|x-y|}\,dx\,dy$$
$$+ \int V_{ext}(x)\rho(x)\,dx + \int \rho\varepsilon(\rho) + \frac{1}{2}\sum_{i\neq j}\frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|},$$

where

$$\rho(x) = 2\sum_{j=1}^{N}|\psi_j(x)|^2$$

Kohn-Sham DFT:

$$\min_{(\psi_i,\psi_j)=\delta_{ij}} E_{KS}[\{\psi_j\}]$$

Molecular Mechanics Model of a Particle System

At zero temperature, the total energy of a system of N particles is

$$E^{\text{tot}}\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N\} = V(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N) - \sum_{i=1}^{N} f(\boldsymbol{x}_i)\boldsymbol{y}_i,$$

where

$$f(\boldsymbol{x}_i) = \text{external force at } i\text{-th atom},$$

and

$$\boldsymbol{x}_i = \text{position of } i\text{-th atom at undeformed state}$$

$$\boldsymbol{y}_i = \text{position of } i\text{-th atom at deformed state}$$

The atomic configuration is determined by the following minimization problem:

$$\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N\} = \operatorname{argmin} E^{\text{tot}}\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N\},$$

subject to appropriate boundary conditions.

Challenges

- Nonconvexity
- Number of minimizers increases with system size ($\mathcal{O}(\exp(N))$)

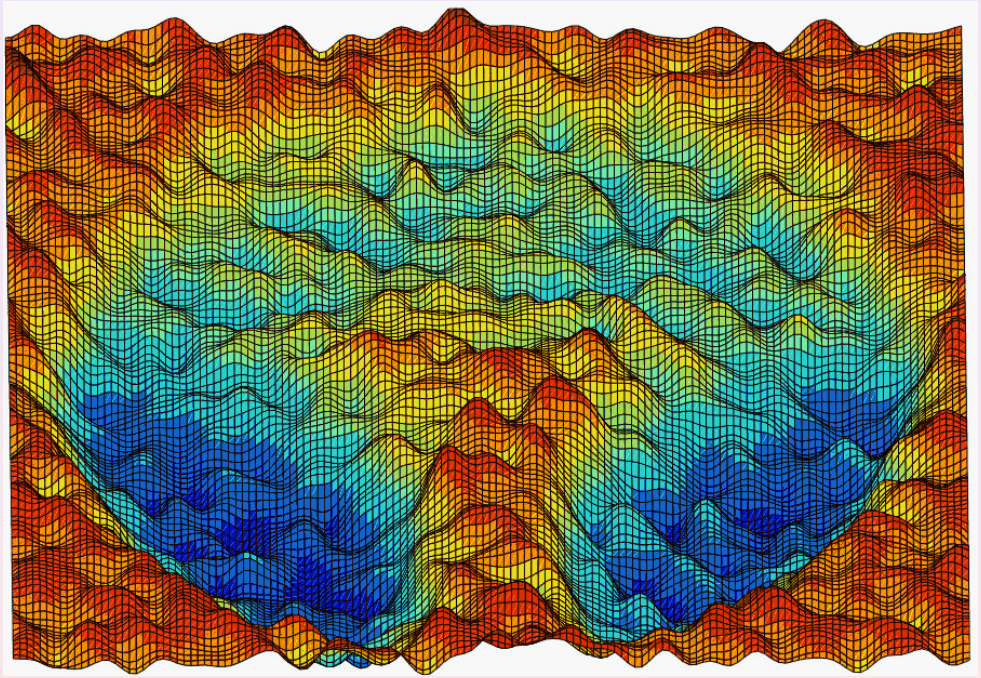
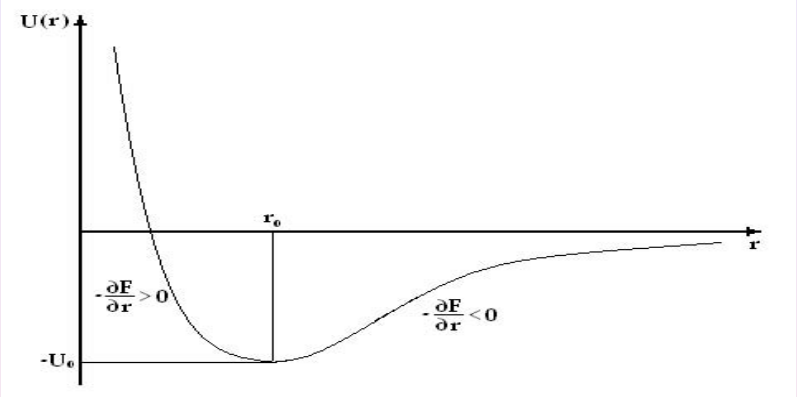
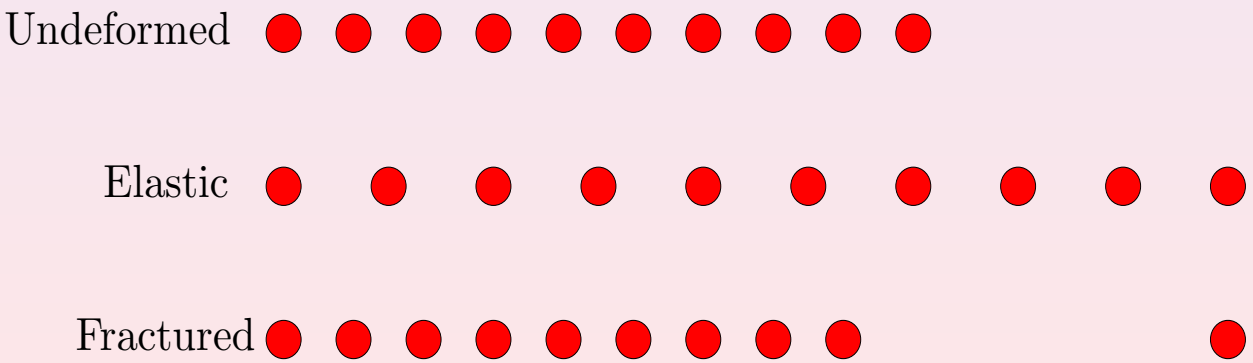


Figure : Local minimizer forest (by courtesy of W. Ren)

Challenges (cont.)

- Physically: Mechanical equilibrium state is a local minimizer and not a global minimizer
- Example: One-dimensional chain in a Lennard-Jones potential:





$$E_{\text{Fractured}} \sim (N-2)U(r_0) + U(D) < E_{\text{Elastic}} \sim (N-1)U(r_0+d)$$

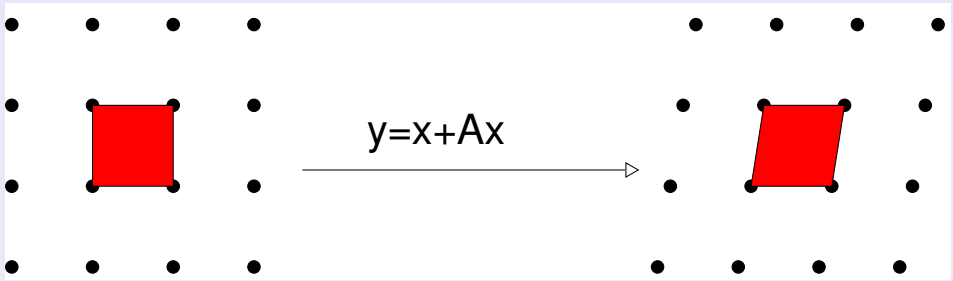
Objectives

Numerically we aim for

- Good convergence properties ? (relevant local minimizer)
- Reliability ? (insensitivity to initial guess and parameters)
- Efficiency ? ($\mathcal{O}(N)$ or $\mathcal{O}(N \log N)$ or sublinear algorithm)

Cauchy-Born rule for a simple lattice[1]

$\mathbf{A} \in \mathbb{R}^{3\times 3}$ $\qquad$ $W(\mathbf{A}) = ?$



- The microscopic state is determined by the macroscopic deformation.
- $W_{\mathrm{CB}}(\mathbf{A})$ is computed by first deforming an infinite crystal uniformly with displacement gradient $\mathbf{A}$, and then setting $W_{\mathrm{CB}}(\mathbf{A})$ to be the energy of the deformed unit cell:

$$W_{\mathrm{CB}}(\mathbf{A}) = \lim_{m\to\infty} \frac{\sum_{y_i,y_j,y_k\in(\mathbf{I}+\mathbf{A})L\cap mD} V(y_i,y_j,y_k)}{|mD|}.$$

Here D is an arbitrary open domain in $\mathbb{R}^d$, L denotes the lattice and $|mD|$ denotes the volume of mD.

[1] M. Born and K. Huang, Oxford University Press, Oxford, 1954

Cauchy-Born rule for a simple lattice (cont.)

- Example: 1d simple lattice with Lennard-Jones potential

$$W_{\mathrm{CB}}(A) = \frac{\zeta^2(6)}{\zeta(12)}\left(|1+A|^{-12} - 2|1+A|^{-6}\right)$$

  $\zeta =$ Riemann-zeta function

- Validity of Cauchy-Born: Under certain lattice stability conditions[1]

$$\|u_{\mathrm{CB}} - \mathbf{y}\| \leq C\epsilon^2,$$

  where $\epsilon =$ lattice constant/diam(Ω).

[1] Theil and Friesecke, *J. Nonl. Sci*, 12, (2002); E, Ming, *Arch. Rat. Mech. Anal.*, 183 (2006); Van Koten and Ortner (2012)

Thermodynamic limit

- More generally: How does one define the energy of a crystal? Thermodynamic limit[1].
- Quantum mechanical description:
  - Does there exist a limit for the energy per unit volume $\frac{1}{N}E_N$ as $N \to \infty$?
  - Does the minimizing electronic density ρ_N approach a limit ρ_∞ in some sense?
  - Does the limit density ρ_∞ have the same periodicity as the assumed periodicity of the nuclei?
- Proved in the context of Thomas-Fermi[2], Thomas-Fermi-von Weiszacker and Hartree-Fock[3].
- Cauchy-Born from Kohn-Sham[4].

---

[1] Blanc, Le Bris, Lions, *Arch. Rat. Mech. Anal.*, 164 (2002)

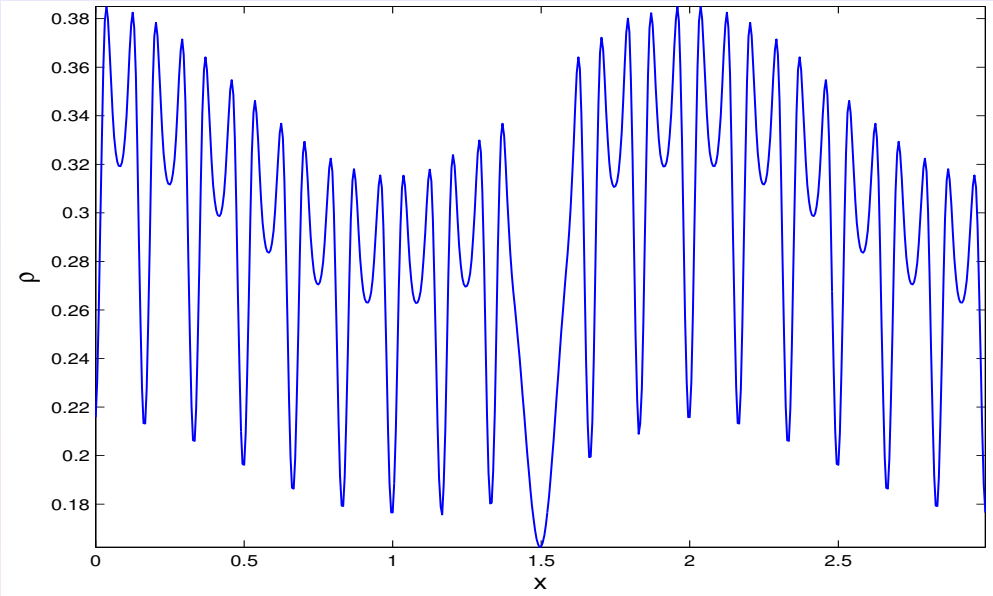
[2] Lieb, Simon, *Adv. Math.*, 23 (1977)

[3] Catto, Le Bris, Lions (2001), Cancès, Lahbabi, Lewin (2012)

[4] Lu, E, *Comm. Math. Sci.*, 5 (2007), García-Cervera, Lu, E, *Comm. Math. Sci.*, 5 (2007), Lu, E, *Memoirs AMS*, (2012)

Cauchy-Born in Quantum Mechanics[1]

Consider the electronic density around a vacancy in an Aluminium crystal:



- The effect of the vacancy is very localized (more on this later).
- Two scales:
  - Fast scale: Interatomic distance (lattice constant).
  - Slow scale: Elastic deformation
- Multiple-scale behavior motivates the use of asymptotic analysis: homogenization.

[1] Lu, E, *Comm. Math. Sci.*, 5 (2007), García-Cervera, Lu, E, *Comm. Math. Sci.*, 5 (2007)

Cauchy-Born (cont.)

We choose units so that the diameter of the domain is $O(1)$, and introduce $\varepsilon = \text{Lattice constant}/diam(\Omega)$. The energy becomes:

$$E[u] = \frac{\varepsilon^2}{2}\int_\Omega |\nabla u|^2 + \varepsilon^2 C_{TF}\int_\Omega u^{10/3} + F_{XC}[\varepsilon^3 u^2] + \frac{\varepsilon}{2}\int_\Omega\int_\Omega \frac{(u^2-m)\cdot(u^2-m)}{|x-y|}.$$

Euler-Lagrange equations:

$$\begin{aligned} -\varepsilon^2\Delta u + \frac{5}{3}\varepsilon^2 u^{7/3} - \phi u + \varepsilon^2 V_{XC}[\varepsilon^3 u] + \lambda u &= 0, \\ -\Delta\phi &= 4\pi\varepsilon(m-u^2), \end{aligned}$$

- λ is a Lagrange multiplier for the normalization constraint.
- ϕ is the Coulomb potential generated by the electrons and the ions.
- Equation set in $\varphi(\Omega)$, with $\varphi : \Omega \to \mathbb{R}^3$ elastic deformation.

Cauchy-Born (cont.)

For smooth elastic deformations:

$$
\begin{aligned}
u &= u(y, \frac{x}{\varepsilon}) = \frac{1}{\varepsilon^{3/2}} u_0\left(y, \frac{x}{\varepsilon}\right) + \frac{1}{\varepsilon^{1/2}} u_1\left(y, \frac{x}{\varepsilon}\right) + \varepsilon^{1/2} u_2\left(y, \frac{x}{\varepsilon}\right) + \cdots, \\
\phi &= \phi\left(y, \frac{x}{\varepsilon}\right) = \phi_0\left(y, \frac{x}{\varepsilon}\right) + \varepsilon\phi_1\left(y, \frac{x}{\varepsilon}\right) + \varepsilon^2\phi_2\left(y, \frac{x}{\varepsilon}\right) + \cdots, \\
\lambda &= \lambda_0 + \varepsilon\lambda_1 + \varepsilon^2\lambda_2 + \cdots
\end{aligned}
$$

- $x = x(y) = \varphi^{-1}(y)$ is the Euler-Lagrange map: It gives the Lagrangian coordinate corresponding to y.
- $u(y, z)$ and $\phi(y, z)$ are functions defined on $\mathbb{R}^3 \times \Gamma$ and are periodic in the second variable on Γ (unit cell).

Leading order equations:

$$
\begin{aligned}
-\Delta_2^x u_0 + \frac{5}{3} u_0^{7/3} - \phi_0 u_0 + \lambda_0 u_0 &= 0, \\
-\Delta_2^x \phi_0 &= 4\pi(m_0 - u_0^2).
\end{aligned}
$$

These are the Euler-Lagrange equations for the *periodic* problem, on the deformed cell $\mathbf{A}\Gamma$, with $\mathbf{A} = \nabla\varphi(x)$: Cauchy-Born rule.

Asymptotics in the Kohn-Sham framework

Euler-Lagrange equations:

$\varepsilon =$ Lattice Constant/Diameter of the domain.

$$\begin{aligned} -\frac{\varepsilon^2}{2}\Delta\psi_k + V_{XC}(\varepsilon^3\rho)\psi_k - \phi\psi_k &= \lambda_k\psi_k; \\ -\Delta\phi &= 4\pi\varepsilon(m-\rho), \end{aligned}$$

$$\rho(\times) = 2\sum_{j=1}^{N} |\psi_j(\times)|^2. \tag{1}$$

Asymptotic expansion for $\varepsilon \ll 1$:

$$\begin{aligned} \psi_\alpha\left(y, \frac{x}{\varepsilon}\right) &= \frac{1}{\varepsilon^{3/2}}\psi_{\alpha,0}\left(y, \frac{x}{\varepsilon}\right) + \frac{1}{\varepsilon^{1/2}}\psi_{\alpha,1}\left(y, \frac{x}{\varepsilon}\right) + \varepsilon^{1/2}\psi_{\alpha,2}\left(y, \frac{x}{\varepsilon}\right) + \\ \rho\left(y, \frac{x}{\varepsilon}\right) &= \frac{1}{\varepsilon^3}\rho_0\left(y, \frac{x}{\varepsilon}\right) + \frac{1}{\varepsilon^2}\rho_1\left(y, \frac{x}{\varepsilon}\right) + \frac{1}{\varepsilon^1}\rho_2\left(y, \frac{x}{\varepsilon}\right) + \cdots \\ \phi\left(y, \frac{x}{\varepsilon}\right) &= \phi_0\left(y, \frac{x}{\varepsilon}\right) + \varepsilon\phi_1\left(y, \frac{x}{\varepsilon}\right) + \varepsilon^2\phi_2\left(y, \frac{x}{\varepsilon}\right) + \cdots . \end{aligned}$$

Asymptotics in the Kohn-Sham framework (cont.)

Leading order:

$$\rho_0(y,z) = 2\sum_{\alpha}\sum_{z_j\in L}|\psi_{\alpha,0}(y,z-z_j)|^2.$$

$$\int_{\mathbb{R}^3}\psi^*_{\alpha,0}(y,z-z_i)\psi_{\alpha',0}(y,z-z_j)\,dz = \delta_{\alpha\alpha'}\delta_{ij}/\det(\nabla\varphi(x)).$$

$$-\frac{1}{2}\Delta_2^x\psi_{\alpha,0}(y,z) + V_{XC}(\rho_0)\psi_{\alpha,0}(y,z) - \phi_0(y,z)\psi_{\alpha,0}(y,z) + \sum_{\alpha',z_j\in L}\lambda_{\alpha\alpha',z_j0}\psi_{\alpha',0}(y,z-z_j) = 0;$$

$$-\Delta_2^x\phi_0(y,z) = 4\pi(m_0-\rho_0)(y,z).$$

These are the Euler-Lagrange equations for the *periodic* problem, on the deformed cell: Cauchy-Born rule.

Linear scaling methods for Kohn-Sham DFT

A number of linear scaling methods have appeared in the literature[1]:

1. Orbital Minimization[2].
2. Density Matrix Minimization[3].
3. Fermi Operator Expansion[4].
4. Divide and Conquer[5].

An interesting $O(N^3)$ algorithm:

- **Subspace Iteration**[6].
- New algorithm: Similar to the subspace iteration method of Zhou, Saad, Tiago, and Chelikowsky '06, **but**, we avoid diagonalization and orthogonalization (which is $O(N^3)$)[7].

---

[1] Goedecker '99

[2] Mauri, Galli, Car '93, W. Yang '97, C. Yang, J.C. Meza and L.W. Wang '06, Burger and Yang '07, W. Gao and W. E '08

[3] Li, Nunes, Vanderbilt, '93

[4] Goedecker '94; Lin, Lu, Car, E '09

[5] Yang '91, L.W. Wang, Z. Zhao, J. Meza '06, Barrault, Cancès, Hager, Le Bris '07

[6] Zhou, Saad, Tiago, Chelikowsky '06

[7] CJGC, Lu, E, '07, CJGC, Lu, Xuan, E, *Phys. Rev. B*, 2009

Localized Subspace Iteration (LSI)[1]

- Non-orthogonal formulation: emphasis is on the **subspace generated by the wave functions**.
- We want to generate the optimal eigenspace of the self-consistent Hamiltonian: **Filtering out the high end of the spectrum**.
- Localization is key for linear scaling: **We choose a localized representation of the occupied subspace**.
- We use finite differences, and **real-space formulation**.

We **avoid**:

- Diagonalization and orthogonalization.
- Using a basis set.
- Using plane waves.
- Using a *supercell* for non-periodic problems.

[1] CJGC, Lu, E, '07, CJGC, Lu, Xuan, E, *Phys. Rev. B*, 2009

Non-Orthonormal Formulation of Kohn-Sham DFT[1]

Given N linearly independent wave functions, $\{\psi_j\}$, define the overlap matrix:

$$\mathbf{S}_{jk} = \int \psi_j \psi_k.$$

Then,

$$E_{KS}[\{\psi_j\}] = 2 \sum_{j,k} \left(-\frac{1}{2}\right) (\mathbf{S}^{-1})_{jk} \int \psi_j \left(\Delta \psi_k\right) \, dx$$
$$+ \frac{1}{2} \int \int \frac{\rho(x)\rho(y)}{|x-y|} \, dx \, dy + \int V_{ext}(x)\rho(x) \, dx + \int \rho\varepsilon(\rho),$$

where

$$\rho(x) = 2 \sum_{jk} \psi_j(x) (\mathbf{S}^{-1})_{jk} \psi_k(x)$$

[1] Mauri, Galli, Car '93, W. Yang '97

Advantages of the Non-Orthogonal formulation

- $\{\psi_j\}$ not orthogonal.
- Invariant under nonsingular linear transformations: Let $\widetilde{\Psi} = \Psi R$, with $R \in \mathbb{R}^{N\times N}$, invertible. Then,

$$E_{KS}[\widetilde{\Psi}] = E_{KS}[\Psi].$$

- The emphasis is therefore on the subspace spanned by $\{\psi_j\}$.
- Nonorthogonal wave functions have better localization properties.

Filtering Step

Goal: To improve the subspace by removing components in the high end of the spectrum of the Hamiltonian

$$H = -\frac{1}{2}\Delta + V_{eff}[\rho].$$

Power Method

The simplest filter is probably the Power Method (Parlett, '98):

1. Given an initial vector $\mathbf{v}^0$.
2. For $k \geq 0$, define
   2.1
   $$\mathbf{v}^{k+1} = \frac{\mathbf{H}\mathbf{v}^k}{\|\mathbf{H}\mathbf{v}^k\|}, \tag{2}$$
   2.2 $\mu^{k+1} = (\mathbf{v}^{k+1})^T \cdot \mathbf{H}\mathbf{v}^{k+1}$.
3. Repeat until $|\mu^{k+1} - \mu^k| \leq \text{Tolerance}$.

Convergence: If $\mathbf{H}\psi_i = \lambda_i \psi_i$, and $|\lambda_1| \leq |\lambda_2| \leq \cdots \leq |\lambda_N|$, then

$$\frac{1}{\|\mathbf{H}\mathbf{v}\|}\mathbf{H}\mathbf{v} = \psi_N + O\left(\left|\frac{\lambda_{N-1}}{\lambda_N}\right|\right). \tag{3}$$

- Note that when applied to a subspace, the space collapses to a one-dimensional space.

Subspace Iteration

The Subspace Iteration generalizes the Power Method to a subspace (Parlett, '98):

1. Given an initial space V_0 of dimension $M < N$, for each $k \geq 1$:

   1.1 Calculate $W_k = \mathbf{H} V_k$.
   1.2 Orthogonalize the basis (QR decomposition, for example): $W_k = Q_k R_k$.
   1.3 Let $V_k = Q_k$.

2. Repeat until convergence.

- The orthogonalization step is necessary in order to ensure the linear independence of the vectors in the new space.
- If $\mathbf{H}\psi_i = \lambda_i \psi_i$, and $|\lambda_1| \leq |\lambda_2| \leq \cdots \leq |\lambda_N|$, then the subspace iteration converges with rate of convergence

$$\tau = \frac{\lambda_M}{\lambda_{M+1}} < 1. \tag{4}$$

Polynomial Filtering

- Filtering improves the rate of convergence of the subspace iteration.
- If the polynomial P splits the spectrum of $\mathbf{H}$, in the sense that

$$P(\lambda_i) \leq P(\lambda_M), \quad i = 1, \ldots, M, \tag{5}$$
$$P(\lambda_j) \geq P(\lambda_{M+1}), \quad j = M+2, \ldots, M, \tag{6}$$

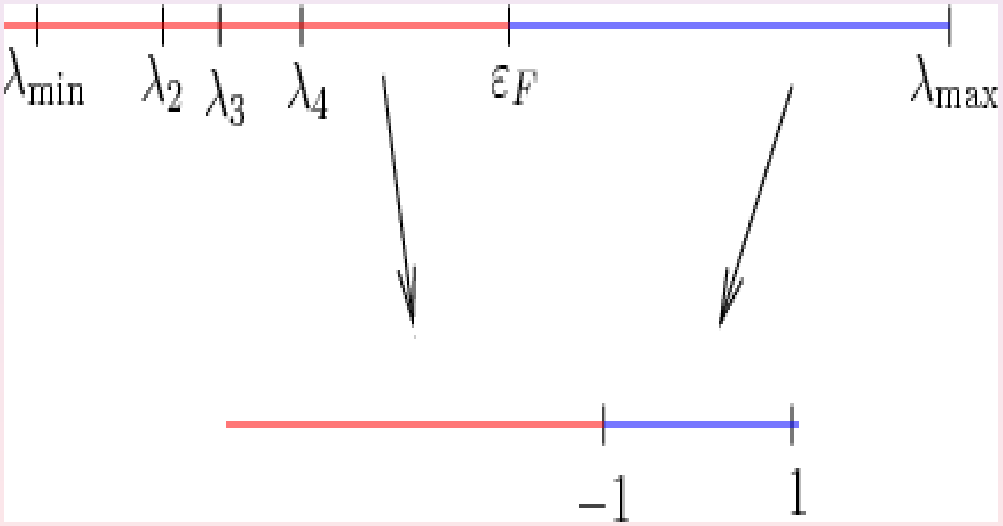
  the rate of convergence of the polynomial filtered subspace iteration is

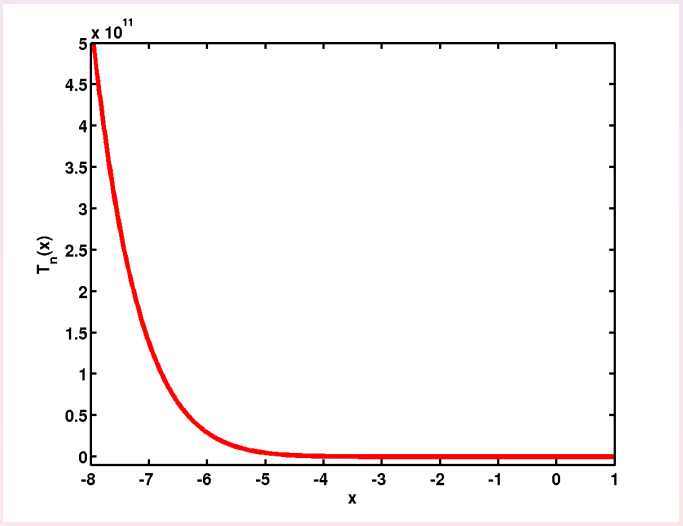
$$\kappa = \left| \frac{P(\lambda_M)}{P(\lambda_{M+1})} \right|. \tag{7}$$

- No diagonalization necessary.

Chebyshev Filter

- Optimal choice: Chebyshev polynomial. $T_n(H)$.
  - Recursive: $T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$.
  - $$T_n(x) = \begin{cases} \cos(n \cos^{-1} x) & if\, |x| \leq 1, \\ (-1)^n \cosh(n \cosh^{-1} |x|) & if\, |x| \geq 1, \end{cases}$$





Chebyshev Filter

- In the context of electronic structure analysis, subspace iteration has been used by Zhou, Saad, Tiago, and Chelikowsky, '06.
- The orthogonalization step leads to an $O(N^3)$ method.
- We replace the orthogonalization step with a localization step, achieving $O(N)$.
- The Fermi energy must be estimated (no diagonalization is used).

Estimation of the Fermi energy

- Given the wave functions Ψ, we know that $\Phi = \Psi \mathbf{S}^{-1/2}$ are orthogonal (*Löwdin transformation*).
- The Ritz matrix is

$$\mathbf{R} = \Phi^T \mathbf{H} \Phi = \mathbf{S}^{-1/2} \Psi^T \mathbf{H} \Psi \mathbf{S}^{-1/2}$$

- We estimate the Fermi energy by the maximum eigenvalue of the Ritz matrix.
- The eigenvalues of $\mathbf{R}$ are the same as the eigenvalues of $\mathbf{S}^{1/2} \mathbf{R} \mathbf{S}^{-1/2} = \Psi^T \mathbf{H} \Psi \mathbf{S}^{-1}$.
- We can use the Power method.
- Note that we do not need $\mathbf{S}^{-1}$, only $\mathbf{w} = \mathbf{S}^{-1} \mathbf{v}$, which can be obtained by solving

$$\mathbf{S} \mathbf{w} = \mathbf{v}.$$

- $\mathbf{S}$ is sparse and localized, and $\Psi^T \mathbf{H} \Psi$ is sparse: The Fermi energy can be estimated in $O(N)$.

Optimally localized wave functions[2]

Given $\{\psi_j\}$, define

$$V = span\{\psi_j\}.$$

The optimally localized wave function, or generalized non-orthogonal Wannier function, ϕ^*, is the minimizer of

$$\inf_{\phi \in V, \|\phi\|_2 = 1} \int w(x)|\phi(x)|^2 \, dx.$$

- Generalizes the Maximally Localized Wannier Functions of Marzari and Vanderbilt (1997).
- An alternative procedure is the *Frobenius Localization* (Weiguo Gao and Weinan E '08).
- One can show that the non-orthogonal wave functions have exponential decay[1].
- The *best* weight function: $w(x) = |x - c|^{2p}$.

[1] Jianfeng Lu '08

[2] Weinan E, Tiejun Li, and Jianfeng Lu, *PNAS* 107 (2010)

Algorithm for Localization[1]

1. Given a set of wave functions, $\{\psi_j\}_{j=1}^N$, centered at the locations $\{b_j\}_{j=1}^N$, respectively.
2. We obtain an optimally localized basis by minimizing

$$F[\phi] = \frac{\int_{\mathbb{R}^3} |x - b_j|^{2p} |\phi(x)|^2 \, dx}{\int_{\mathbb{R}^3} |\phi(x)|^2 \, dx},$$

   among functions ϕ of the form

$$\phi(y) = \sum_{k=1}^{r} \alpha_k \psi_k(x).$$

3. Minimization leads to

$$\mathbf{W}a = \lambda \mathbf{S}a.$$

- **Only a fixed number r of functions involved: $O(N)$**
- The localized functions span the same space.

[1] CJGC, Jianfeng Lu, Weinan E, '07; CJGC, Jianfeng Lu, Yulin Xuan, Weinan E, '09

Computation of the Electronic Density

$$\rho(x) = 2 \sum_{jk} \psi_j(x) (\mathbf{S}^{-1})_{jk} \psi_k(x)$$

- Computing $\mathbf{S}^{-1}$ directly is $O(N^3)$.
- Instead, we use the Newton-Schultz iteration to solve

$$\mathbf{DSD} - \mathbf{D} = 0.$$

- $\mathbf{S}$ and $\mathbf{S}^{-1}$ are localized near the diagonal.
- Exploiting sparsity, computation is $O(N)$ [1].
- Alternatively, a pseudoinverse can be used[2].

[1] Jansik, Host, Jorgensen, Olsen, and Helgaker '07, Rubensson and Salek '05

[2] W. Yang '97

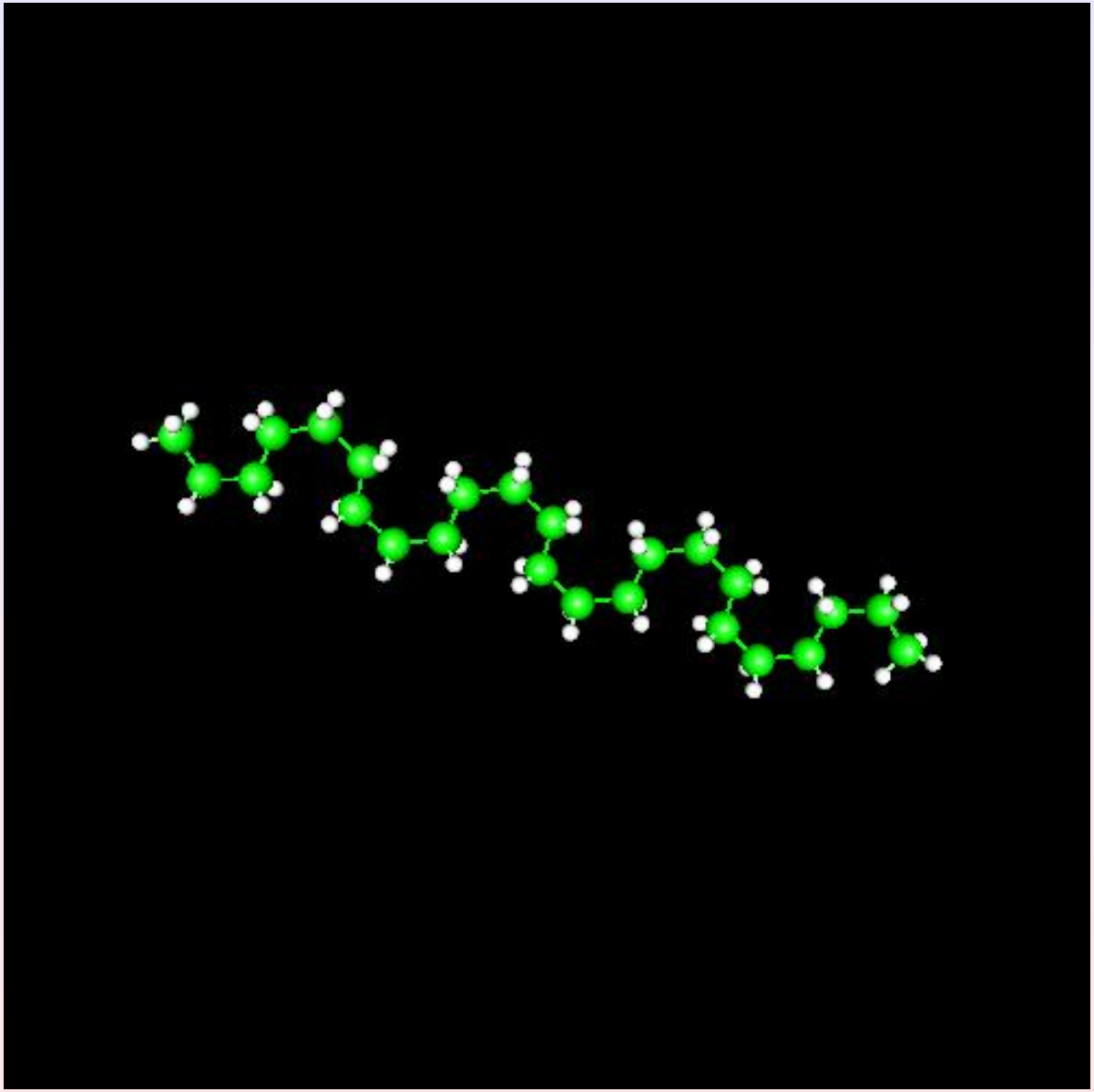
Linear Scaling Algorithm for Kohn-Sham[1]

1: Given (localized) wave functions Ψ_0.
2: **repeat** {(*Self-Consistency Loop (SCF)*)}
3: Compute electronic density: ρ
4: Compute effective potential: $V_{\text{eff}}[\rho]$.
5: **repeat** {(*Localized Subspace Iteration*)}
6: Estimate Fermi energy.
7: Filtering Step: $\Phi = T_n(H)\Psi$.
8: Localization Step: Localize ψ_r for $r = 1, \cdots, k$.
9: Truncation beyond cut-off radius.
10: **until** Convergence of LINEAR iteration
11: Update electronic density (mixing).
12: **until** $\|\rho_{k+1} - \rho_k\|_1 \leq Tol$.

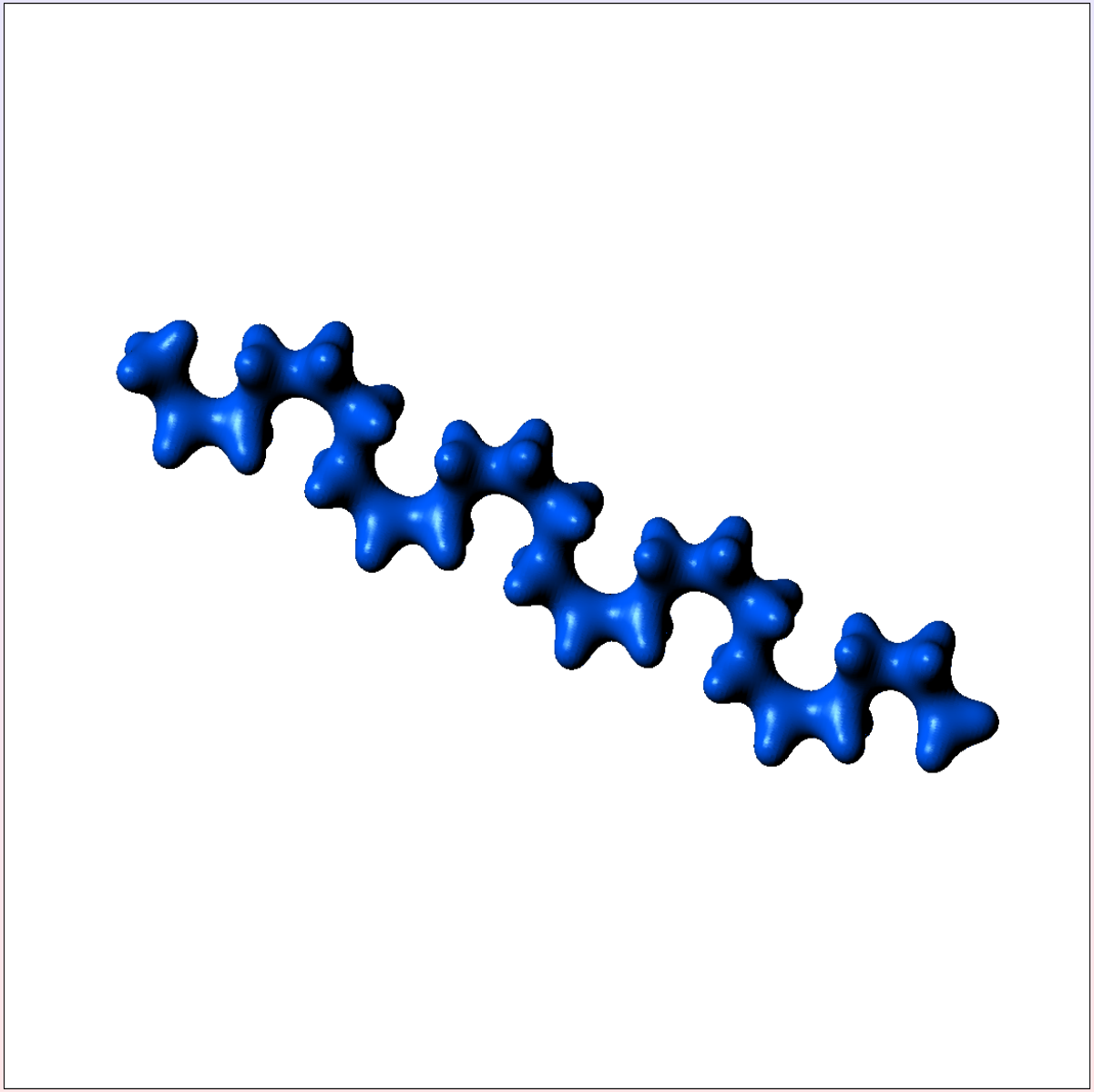
- Similar to the subspace iteration method of Zhou, Saad, Tiago, and Chelikowsky '06, **but**, we avoid diagonalization and orthogonalization (which is $O(N^3)$).
- In practice, only one filtering step is performed.

[1] CJGC, Jianfeng Lu, Weinan E '07; CJGC, Jianfeng Lu, Yulin Xuan, Weinan E, '09

Example: *Alkane* - $CH_3(CH_2)_{10}CH_3$ (74 atoms)



Example: *Alkane* - $CH_3(CH_2)_{10}CH_3$ - Density



LSI Timings

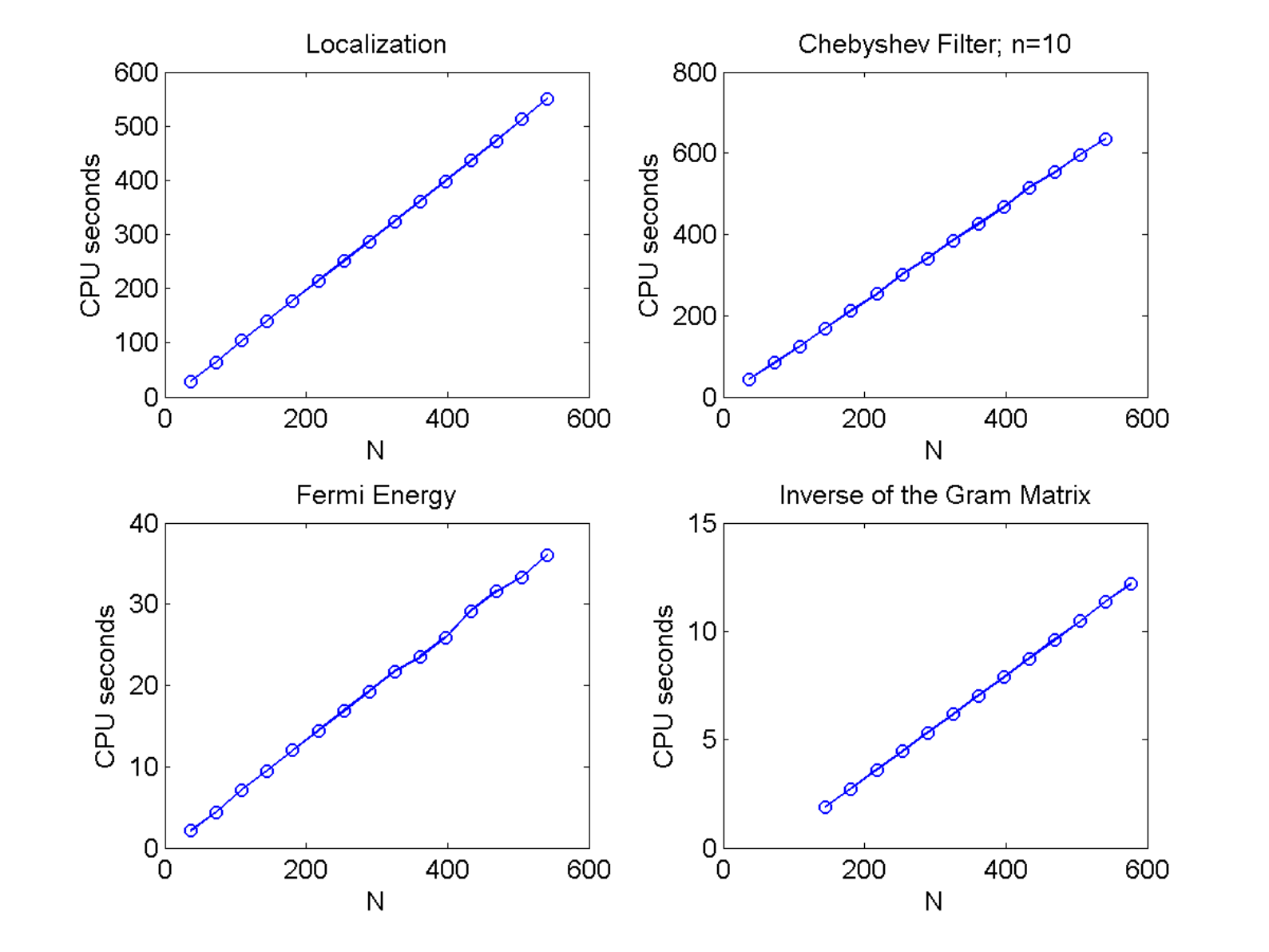


Figure : Timings obtained with the LSI code. Linear scaling is observed.

Precomputing W_{CB}[3]

- Concurrent computation of the electronic density is impractical.
- Stress is a function of the strain: This is a function of six variables.
- This can be precomputed and stored in a table from where we can interpolate, in the context of the Heterogeneous Multiscale Method (HMM)[1].
- An effective way of doing this is using a **sparse representation**.
- Can be used to assess accuracy of empirical potentials[2].

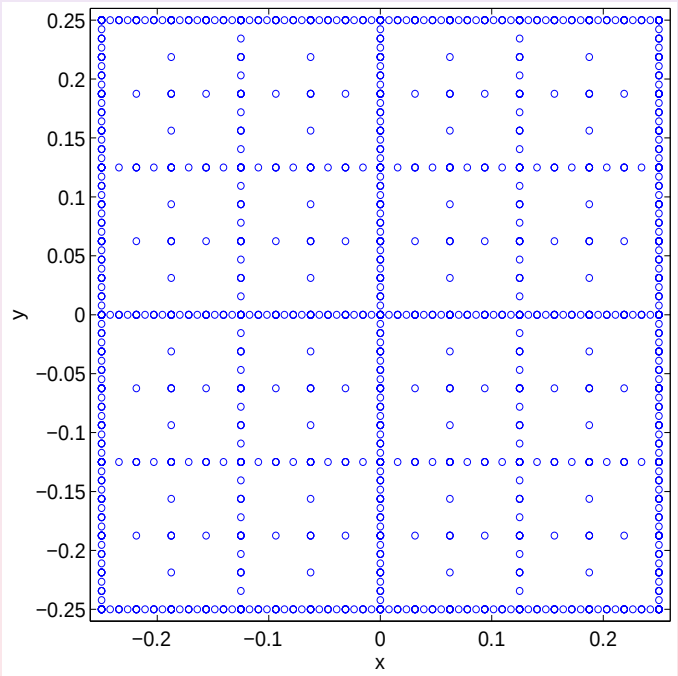
[1] E, Engquist '03

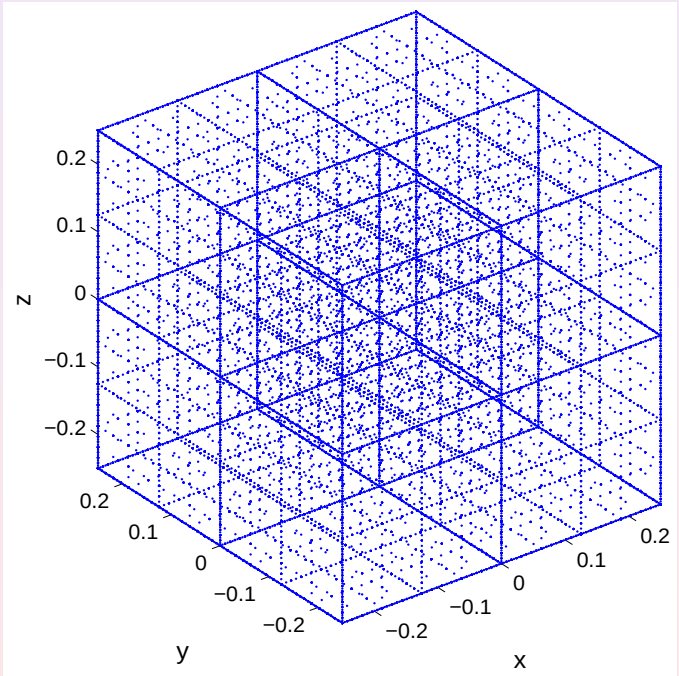
[2] G. Wu, G. Lu, CJGC, and W. E, *Phys. Rev. B*, 79 (2009)

[3] CJGC, Ren, Lu, and E, *Comm. Comp. Phys.*, 4 (2008)

Sparse representation

- If we attempt to represent a function of d variables using a uniform grid, we would need $O(2^{nd})$ grid points.
- Using sparse grids, we need only $O(n^{d-1}2^n)$ points[1].





[1] Bungartz and Griebel, *Acta Numerica (2004)*

Elastic deformation of a solid

Given a deformation $u : \Omega \to \mathbb{R}^3$, we denote by $F = \nabla u$ the deformation tensor, and define the elastic energy density as

$$\omega[F] = \frac{1}{|V_0|} W_{\mathrm{CB}}[F], \tag{8}$$

where the Cauchy-Born energy, $W_{\mathrm{CB}}[F]$, is the energy of the deformed unit cell, obtained by transforming the basis vectors as $b_i = Fa_i$, $i = 1, 2, 3$. Due to frame indifference, the elastic energy depends only on the right Cauchy-Green strain tensor (RCGST), $C = F^T F$.

Computational Cost

- $2,572,288$ grid points were used, which provided up to four digits of accuracy.
- To achieve the same accuracy with a regular grid, $128^6 \geq 4 \times 10^{12}$ grid points would be required.

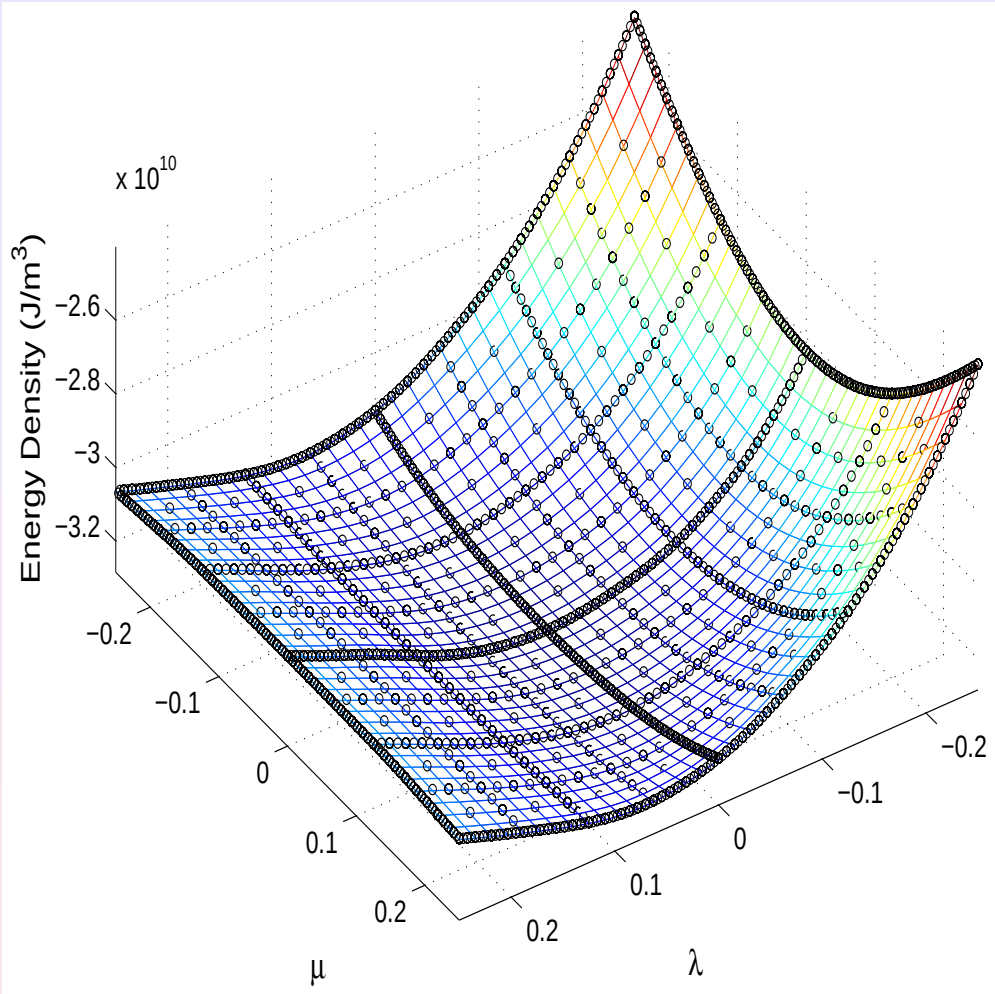


Figure : Elastic energy for the shear and expansion deformation. We plot the energy interpolated to a uniform grid. The sparse grid nodes are superimposed.

Cauchy-Born: Limitations

- No explicit general formulas, but can be precomputed.
- Unable to capture defects.
- Accuracy for finite systems? Finite size effects?

(Domain) Decomposition Approaches

- Domain decomposition: Quasicontinuum[1]
  - Domain decomposed into a local and nonlocal region.
  - Local region: Cauchy-Born elasticity.
  - Nonlocal region: Fully atomistic.
  - Interface matching: Can produce *ghost forces*[2].
  - Extended to Quantum-Mechanical models[3].
- Force decomposition[4]
  - Short wavelength components: atomistic.
  - Long wavelength components: Based on linear elasticity theory.
  - Atomistic forces are transferred to a mesh, and displacement are obtained via linear elasticity, and transferred back to the atomic locations.

---

[1] Tadmor, Ortiz, Phillips, *Philosophical Magazine A*, **73** (1996)

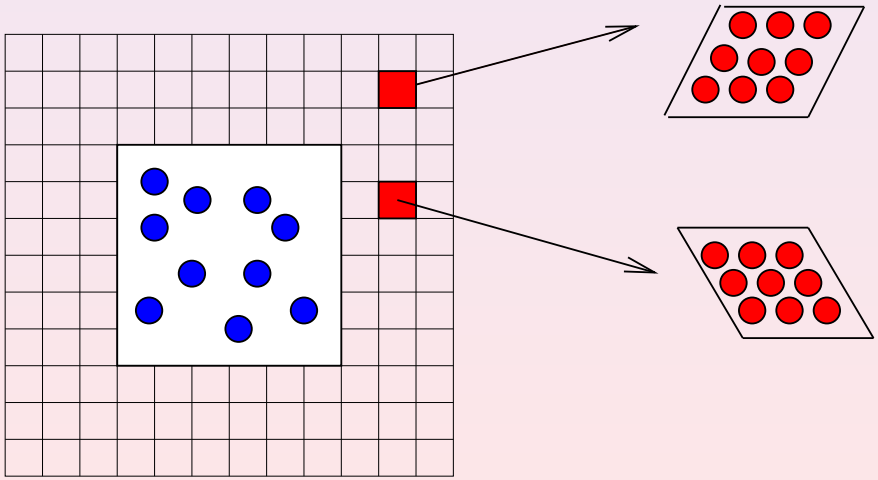
[2] Tadmor, Phillips, Ortiz ('96), E, Ming ('05), E, Lu, Yang ('06), Dobson and Luskin ('07)

[3] E, Lu, Kaxiras ('06), Hayes, Ho, Ortiz, and Carter ('06), G-C, Lu, E ('07), Gavini, Bhattacharya and Ortiz ('07), Peng, Zhang, Hung, Carter, Lu ('08)

[4] S. Goedecker, F. Lançon, and T. Deutsch, Phys. Rev. B, 64 (2001), 161102

The Quasicontinuum Method[1]

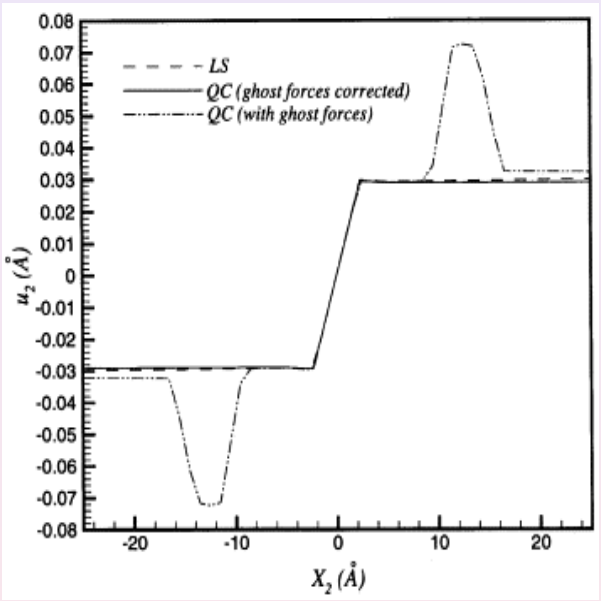
- Consider a material sample with a defect, e.g., a crack, vacancy, dislocation, etc.
- Decompose the domain into a **nonlocal** region containing the defect, and a **local** region, containing the rest.
- In the nonlocal region, atoms are treated directly.
- In the local region, representative atoms (rep-atoms) and Cauchy-Born elasticity is used.



[1] Tadmor, Ortiz, Phillips, *Philosophical Magazine A*, **73** (1996)

The Quasicontinuum Method (cont.)

- Overcomes some limitations of Cauchy-Born.
- Matching at the atomistic/continuum interface is difficult: Ghost forces[1] (Nonzero force for equilibrium configuration).



- Energy-based correction[2], Force-based correction [3].
- No general method to remove ghost forces is available to date.

[1] Shenoy, Miller, Tadmor, Rodney, Phillips, Ortiz, *J. Mech. Phys. Solids*, 47 (1999)

[2] E, Lu, Yang *Phys. Rev. B* 74 (2006).

[3] Luskin, Dobson, *Math. Modelling Num. Anal.* 42 (2008).

Force Decomposition[1]

- No domain decomposition: No matching conditions.
- Information is transferred between a regular mesh and atomistic locations:
  - Atomistic forces $\mathbf{f}_a$ are transferred to a regular grid: $\mathbf{f}_G$.
  - Solve for displacement on regular grid: $-\mathbf{A}\mathbf{u}_G = \mathbf{f}_G$ (Linear Elasticity).
  - Transfer mesh displacement to atomic displacements: $\mathbf{u}_a$.
- The matrix $\mathbf{A}$ contains the material elastic constants.
- Difficulty with defects due to plastic deformation.

[1] S. Goedecker, F. Lançon, and T. Deutsch, Phys. Rev. B, 64, 161102 (2001)

Multilevel Methods

- Multigrid philosophy: A. Brandt[1]
  - Proposes Multigrid as multiscale approach, although no concrete examples are provided.
- Multigrid + Cauchy-Born[2].
  - Based on cascadic or one-way multigrid.
  - Cauchy-Born elasticity used at coarse levels.
  - Fully atomistic at fine level.
- Multigrid + Constrained Cauchy-Born[3]
  - Not a systematic approach.
  - Requires knowledge of the defect structure.

---

[1] Multiscale Scientific Computation: Review (2001)

[2] J. Chen and P.B. Ming, Comm. Comput. Phys., 10 (2011), 70

[3] J. Chen, P.B. Ming, and J. Z. Yang, preprint

Multigrid Method

- Introduced by Achi Brandt[1] and Wolfgang Hackbusch[2] for the solution of the algebraic equations that result from discretizing partial differential equations.
- Typically based on:
  1. Smoothing operator: Removes high frequencies in the error.
  2. Coarse Grid Correction: Captures low frequencies.

[1] A. Brandt, *Math. Comp.*, **31** (1977)

[2] W. Hackbusch, *Computing*, **20** (1978)

Multigrid Method (cont.)

Basic two-grid method for the linear case:

- Fine-grid problem (discretization of original problem): $A_h u_h = f_h$. Solve iteratively.
- **Defect correction**: Given an approximate solution $u_h^{(k)}$: $u_h^{(k+1)} = u_h^{(k)} + e$, where $A_h e = f_h - A_h u_h^{(k)}$.
- Iterative methods such as Gauss-Seidel, Jacobi, etc, stagnate: **Smoothing** (efficient with high frequency components, but not for low frequency components).
- **Coarse Grid correction**:
  - Low frequencies in fine grid become high frequencies in coarse grid.
  - Use coarse grid to provide initial guess in fine grid.
- How should we correct on mesh Ω_h? What problem should we solve in mesh Ω_H?

Multigrid Method (cont.)

- How should we correct on mesh Ω_h? What problem should we solve in mesh Ω_H?
  - In the symmetric, positive definite case:

$$u_h = \arg\min_{u \in \Omega_h} \phi(u),$$

where

$$\phi(u) = \frac{1}{2} u^T A u - u^T f.$$

  - Define

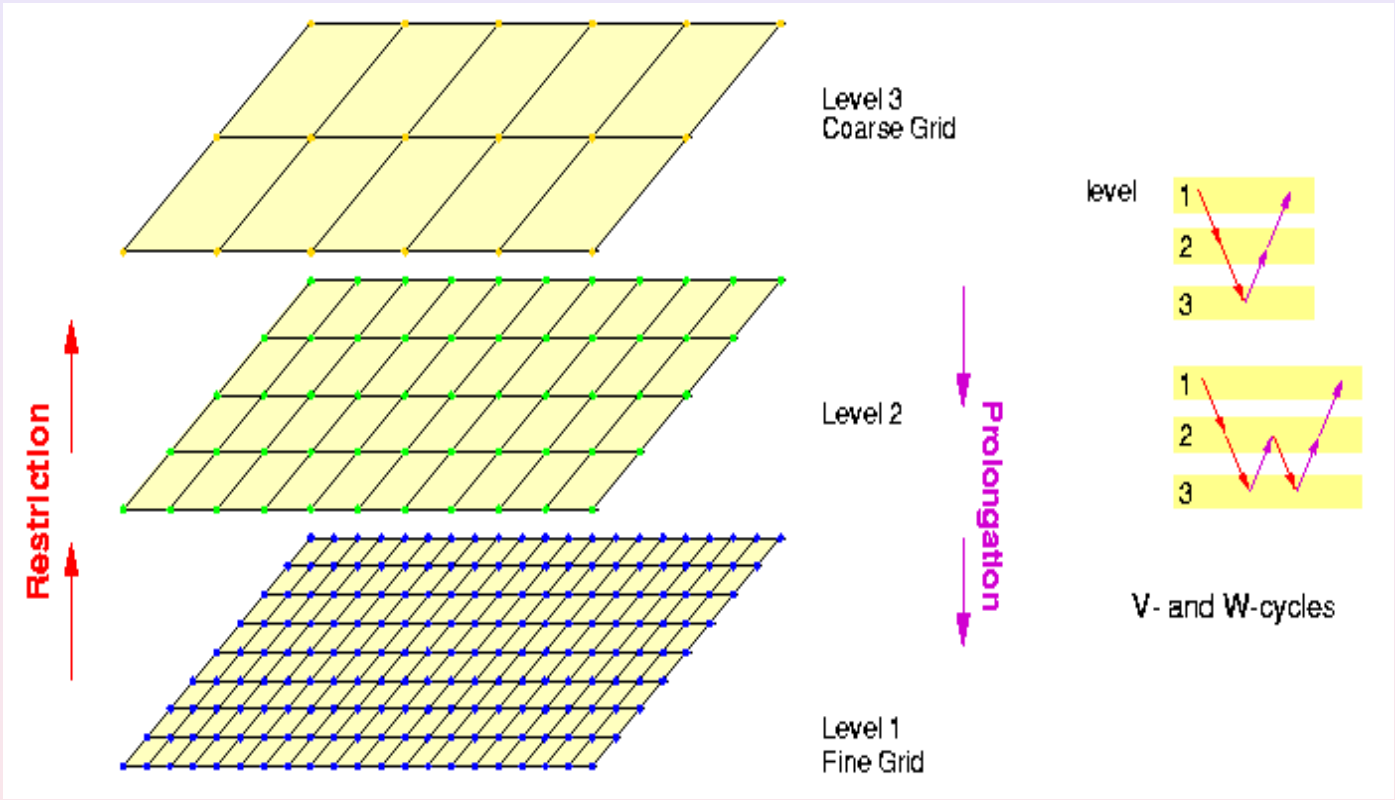
$$e_H = \arg\min_{v \in \Omega_H} \phi(u^{(k)} + I_H^h v),$$

where $I_H^h : \Omega_H \to \Omega_h$ is an *interpolation operator*.

- Coarse-grid problem: $A_H v_H^k = r_H^k$
  - $r_H^k = I_h^H (f_h - A_h u_h^k)$, with $I_h^H = c(I_H^h)^T$.
  - $A_H = I_h^H A_h I_H^h$.
- Correction: $u_h^{k+1} = u_h^k + I_H^h e_H^k$.

Multigrid Method is obtained by recursive application of two-grid method.

Multigrid Method (cont.)



Multigrid for Nonlinear Problems

- Consider $A(u) = f$.
- Newton Multigrid:
  - Linear system solved using *linear* Multigrid.
  - Global linearization: *Local* convergence.
- Full Approximation Scheme (FAS): Takes advantage of multigrid idea.
  - The nonlinear residual equation is $A(u + e) - A(u) = r(= f - A(u))$.
  - **FAS idea**: $A(u_h + e) = A(u_H) + r_H$, with
    - $u_H = I_h^H u_h$,
    - $r_H = I_h^H (f_h - A(u_h))$.
  - Reduces to linear multigrid for $A(u)$ linear.
  - The actual solution is a *fixed point*.

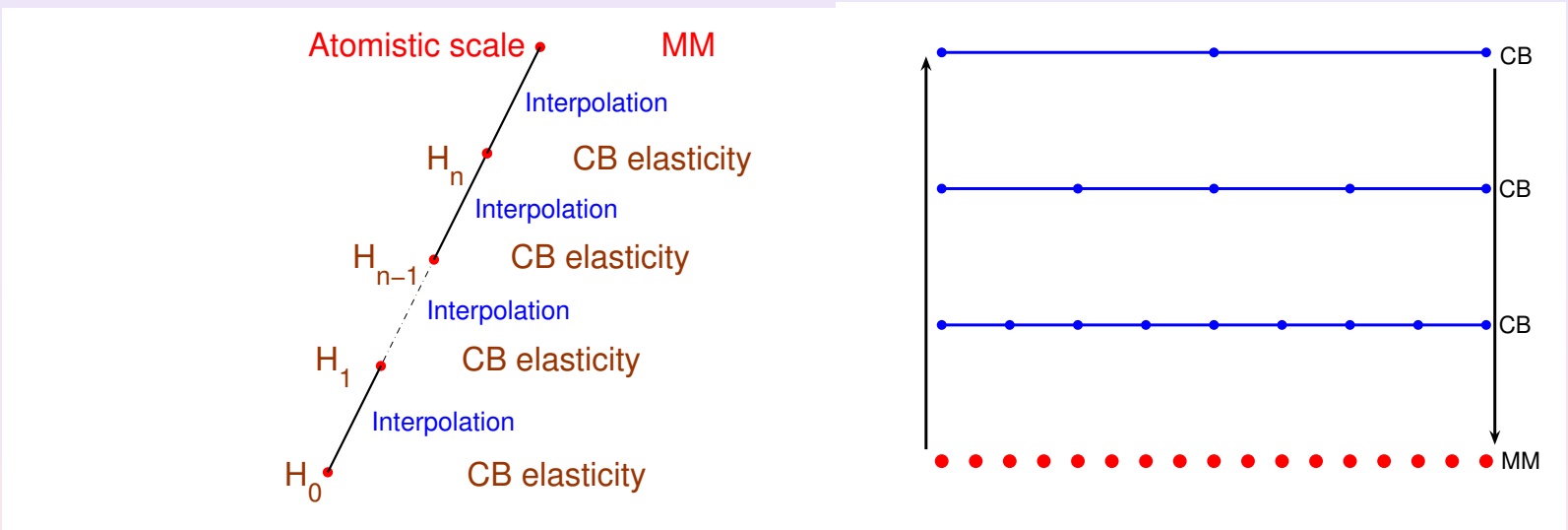
Cascadic Multigrid for Molecular Mechanics[1]

- Consider a nested sequence of triangulations $\mathcal{T}_0 \subset \mathcal{T}_1 \subset \ldots \mathcal{T}_L \subset \Omega$ and associated finite element spaces $X_0 \subset X_1 \subset \cdots \subset X_L$.
- Initialization: Let $\boldsymbol{v}_0 = 0$ be the initial guess. Minimize the CB elasticity problem discretized on $\mathcal{T}_0$ to obtain $\boldsymbol{u}_0$.
- For $i = 1, \ldots, L$,
  - Interpolate $\boldsymbol{v}_i = I_{i-1}^i \boldsymbol{u}_{i-1}$, where $I_{i-1}^i : X_{i-1} \to X_i$ is the interpolation operator.
  - Use $\boldsymbol{v}_i$ as initial guess to minimize the CB problem discretized on $\mathcal{T}_i$.
- At the finest level L, construct the initial atomic locations by $\mathbf{y}_{CB} = \mathbf{x} + \mathbf{v}_L(\mathbf{x})$, and solve the molecular mechanics problem using $\mathbf{y}_{CB}$ as initial guess.

[1] J. Chen, P.B. Ming, Comm. Comput. Phys., 10 (2011), 70

Cascadic Multigrid[1] (II)

Cascadic Multigrid:



- Cauchy-Born solution not a good initial guess near defects.
- Only sees defect at atomistic level.

[1] J. Chen, P.B. Ming, Comm. Comput. Phys., 10 (2011), 70

Our Approach: Goals[1]

- We are interested in a general, robust method.
- Propose more general coarse-grid models, beyond Cauchy-Born related continuum models.
- Extend to more general situations such as dislocations in a robust way.
- We have extended the work of Chen and Ming in two directions:
  - Quasiatomistic approach: CB elasticity is replaced by the introduction of *quasiatoms*.
  - Multiscale Multigrid: Atomistic locations are kept at every level.

[1] C. García-Cervera and J. Chen, *in preparation*

Coarse-grid Model: Quasiatomistic Description

- Select grid points or rep-atoms (robust adaptive technique required)

$$\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{N_{\text{rep}}}\}$$

- Reconstruct atomic configuration (interpolation)

$$\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N\} := \Pi\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{N_{\text{rep}}}\}$$

- Compute system energy by summing site energies ($\mathcal{O}(N)$ cost)

$$\begin{aligned} E_{\text{QA}}^{\text{tot}}(\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{N_{\text{rep}}}\}) &= E^{\text{tot}}(\Pi\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{N_{\text{rep}}}\}) \\ &= V(\Pi\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{N_{\text{rep}}}\}) - \sum_{i=1}^{N} \boldsymbol{f}_i \cdot \Pi_i\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{N_{\text{rep}}}\}. \end{aligned}$$

Algorithm

1. Solve the quasiatomistic model on the coarsest level $l = 1$
2. For $m = 1, \ldots,$ Nvcycle
3. For $l = 2, \ldots, L$ (number of levels, including atomistic level)
4. Interpolate $\{\boldsymbol{u}\}_{l-1}$ to $\{\boldsymbol{u}\}_l$
5. Relax system energy in the quasiatomistic model on level l with Nrelax relaxation steps in the manner of a V-cycle
6. End
7. End
8. Solve on level L (fully atomistic) with the displacement obtained from level $L - 1$.

Continued...

- Minimize $E_{\mathrm{QA}}^{\mathrm{tot}}$ and find equilibrium solution $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{N_{\mathrm{rep}}}\}$ ($\mathcal{O}(N_{\mathrm{rep}})$ cost)

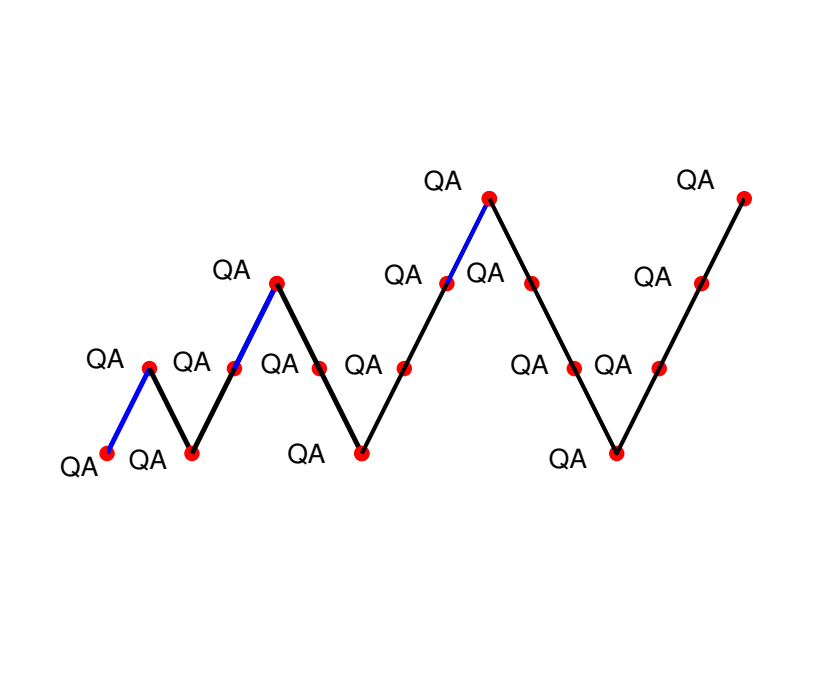
$$\mathrm{argmin}\, E_{\mathrm{QA}}^{\mathrm{tot}}(\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{N_{\mathrm{rep}}}\})$$

- Reconstruct atomic configuration again ($\mathcal{O}(N)$ cost)

Continued...

Remarks

1. Optimal cost in the framework of multigrid method : $\mathcal{O}(N \log N)$.
2. Uniform mesh for the time being.
3. Preliminary results indicate this approach is robust.
4. However, as in Cauchy-Born elasticity, it fails to capture defects effectively at coarse levels.



1D Tension

One-dimensional chain with Lennard-Jones potential up to nearest neighbor without body-force

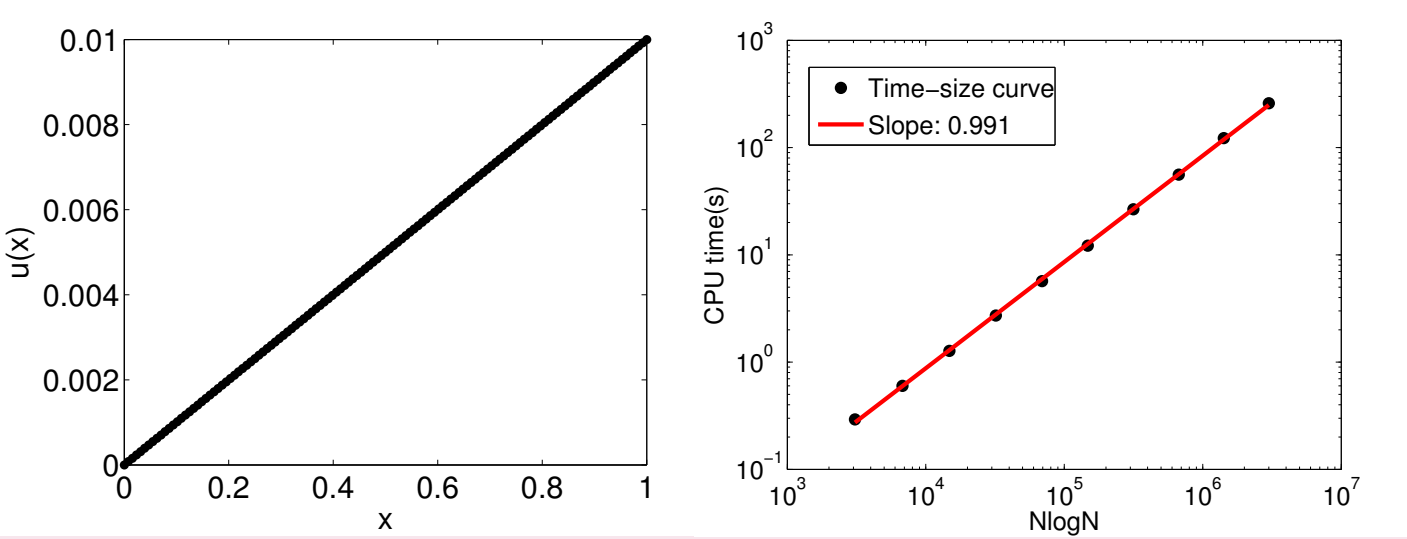
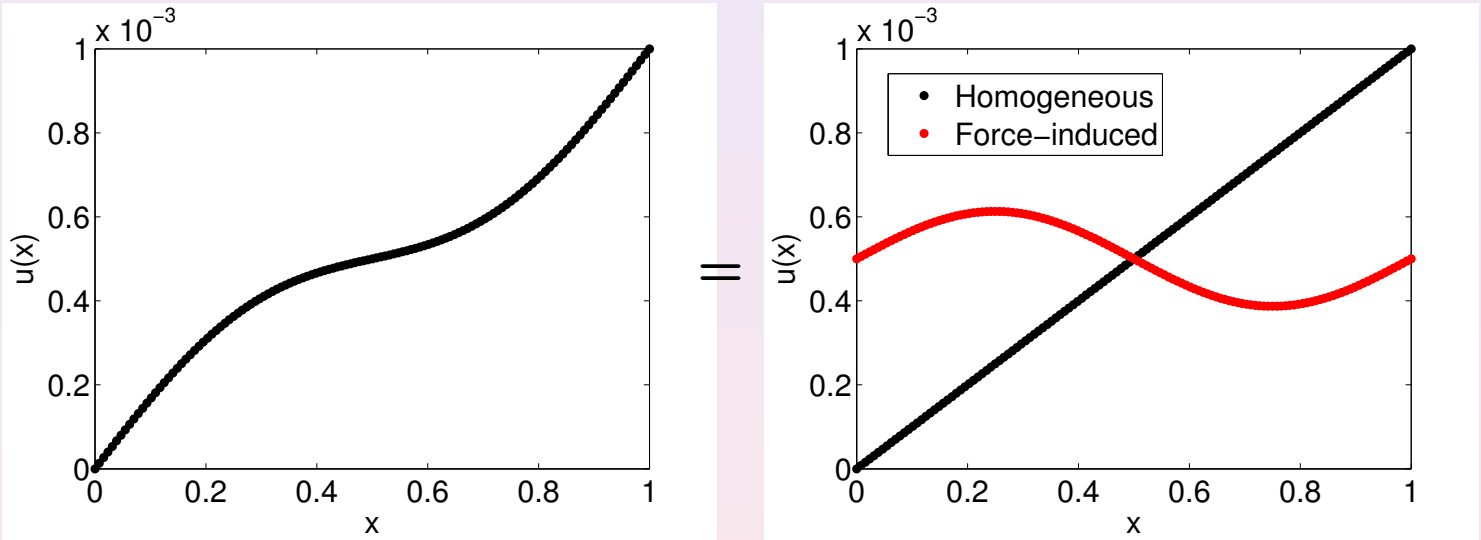


Figure : 1% strain. Left: Displacement; Right: CPU time

1D Tension with Smooth External Force

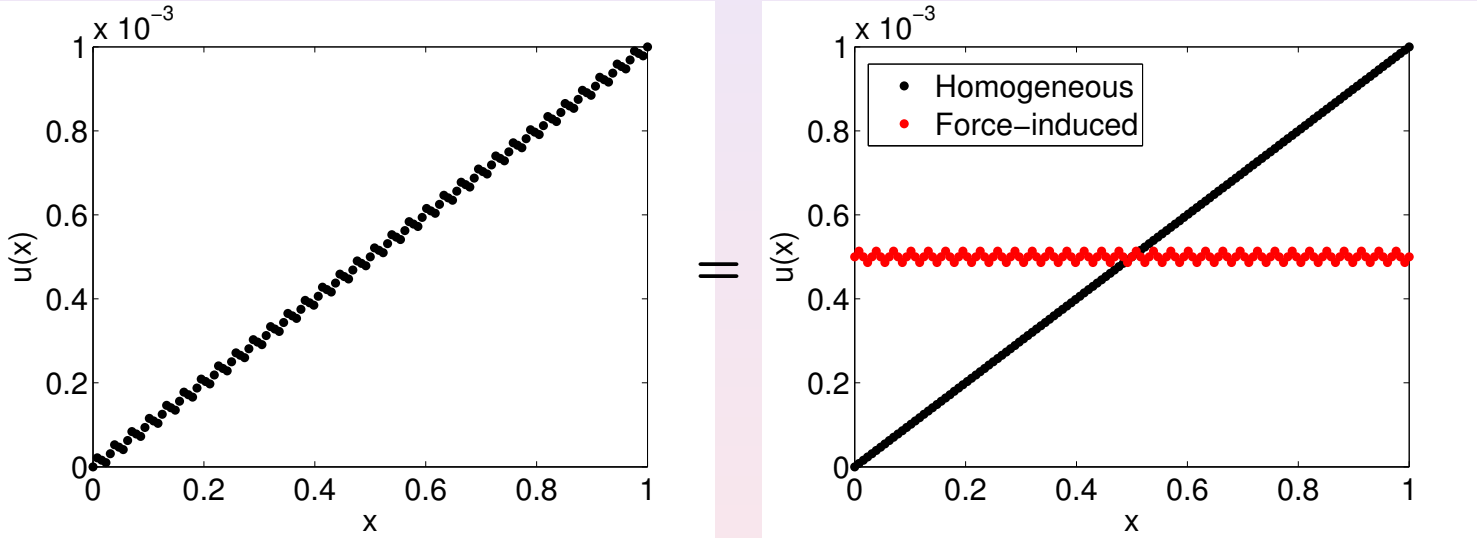
Figure : 1% strain with force $f(x) = \sin(2\pi x)$. Decomposition of displacement field (shifted).



- CPU time scales $\mathcal{O}(N \log N)$.

1D Tension with Oscillatory External Force

Figure : 1% strain with force $f(x) = \sin(2\pi x/\epsilon)$ with ϵ lattice constant. Decomposition of displacement field (shifted).



- CPU time scales $\mathcal{O}(N \log N)$.

Shear

Embedded Atom Method (EAM)[1] $\uparrow$ [001] $\rightarrow$ [110]

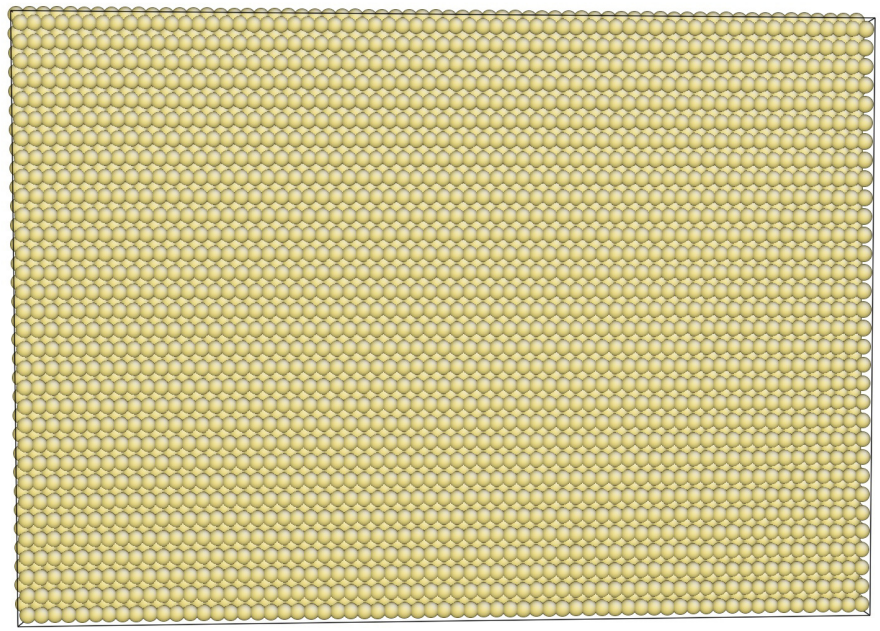


Figure : Atomic structure: 0.1% strain.

[1] Daw and Baskes *Phys. Rev. B*, **29** ('84), Ercolessi and Adams *Europhys. Lett.* **26** ('94)

Shear

$\uparrow [001] \rightarrow [110]$

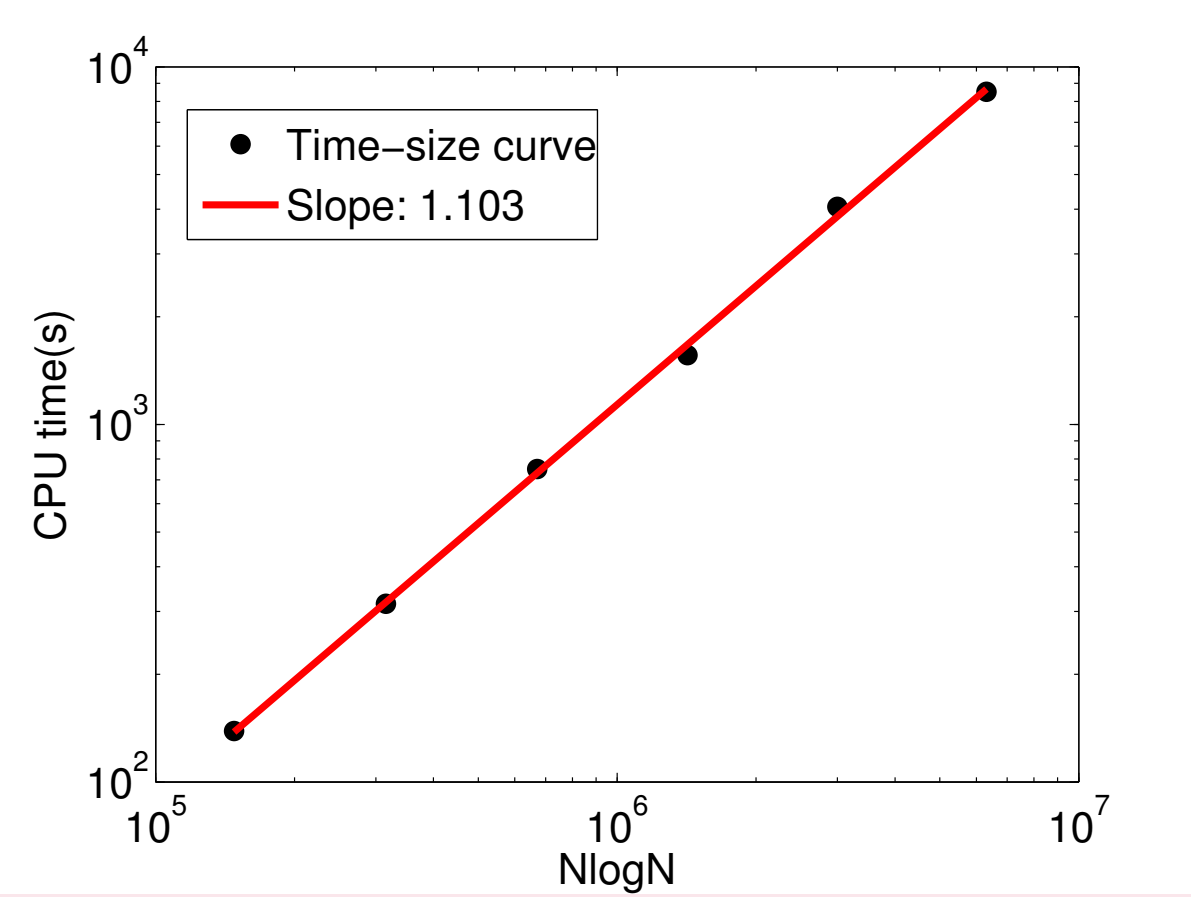
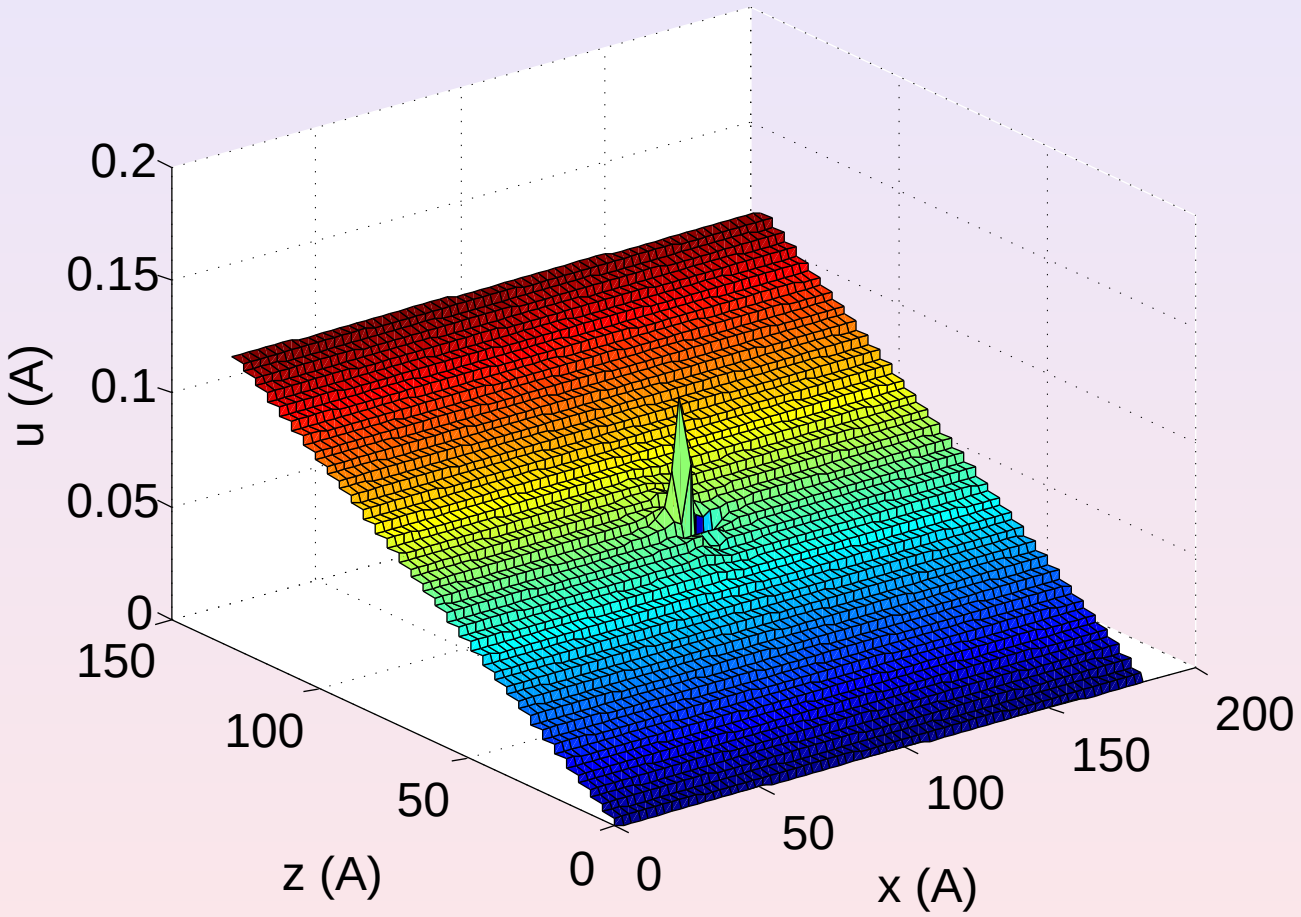


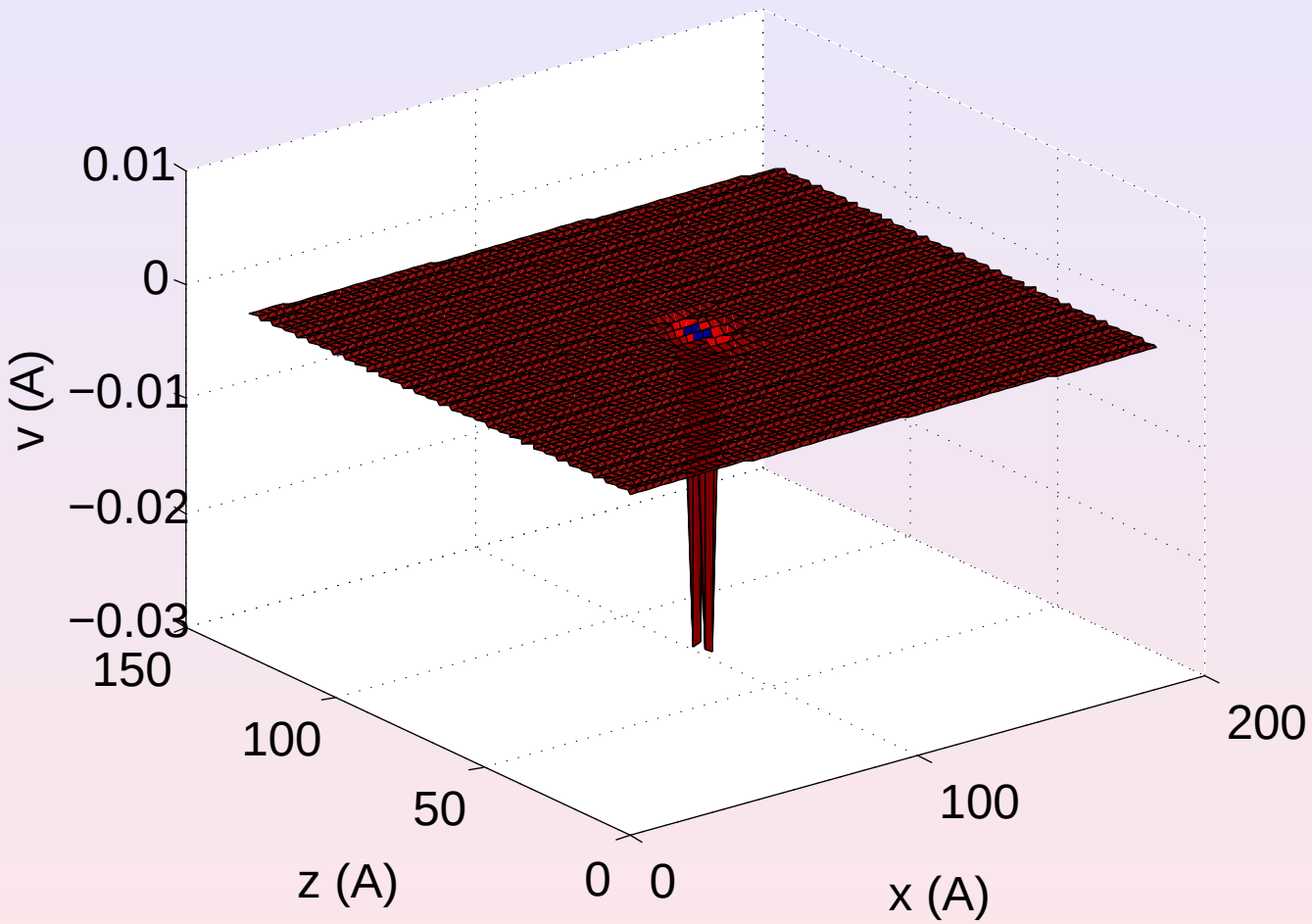
Figure : CPU time

Saves 94% CPU time for a system of 32768 atoms.

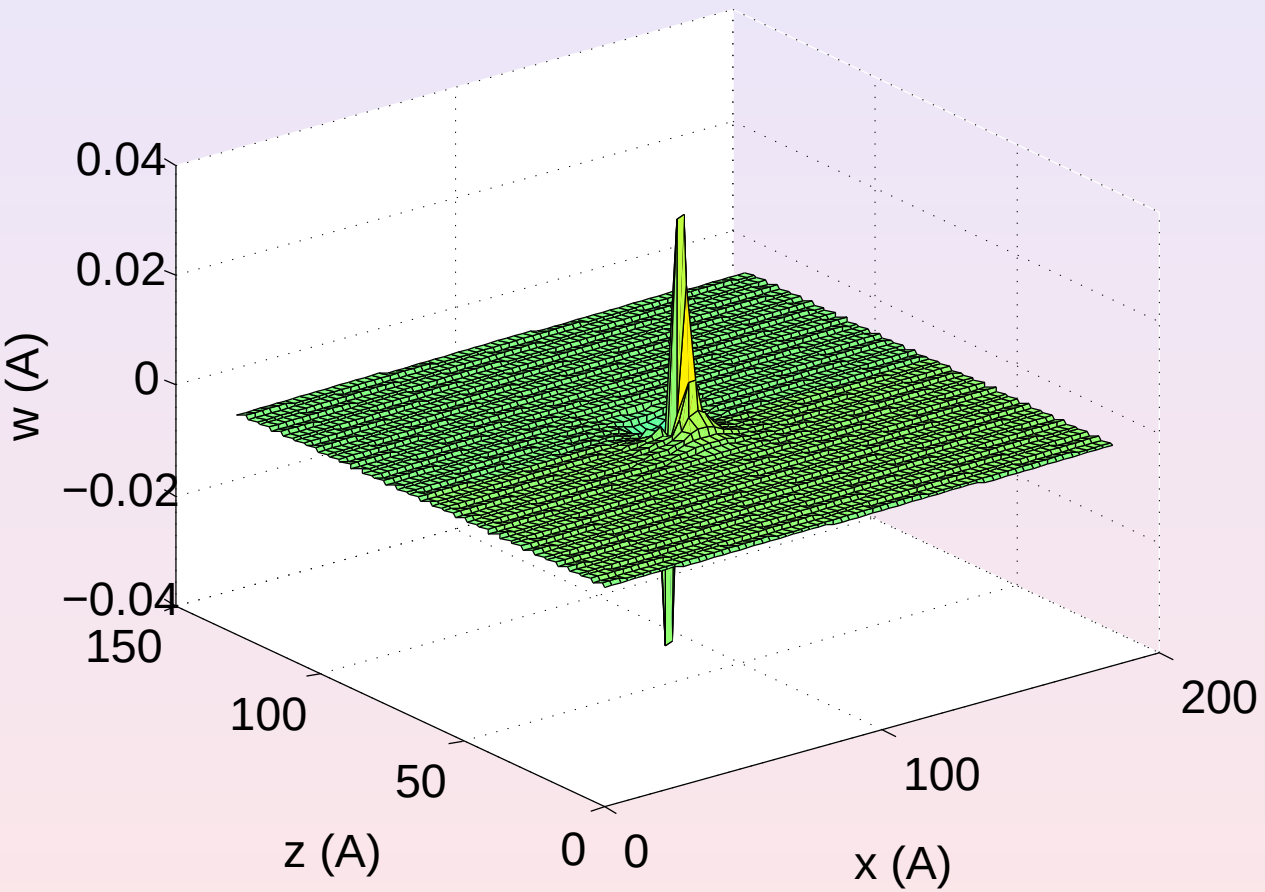
Aluminum: Vacancy



Aluminum: Vacancy



Aluminum: Vacancy



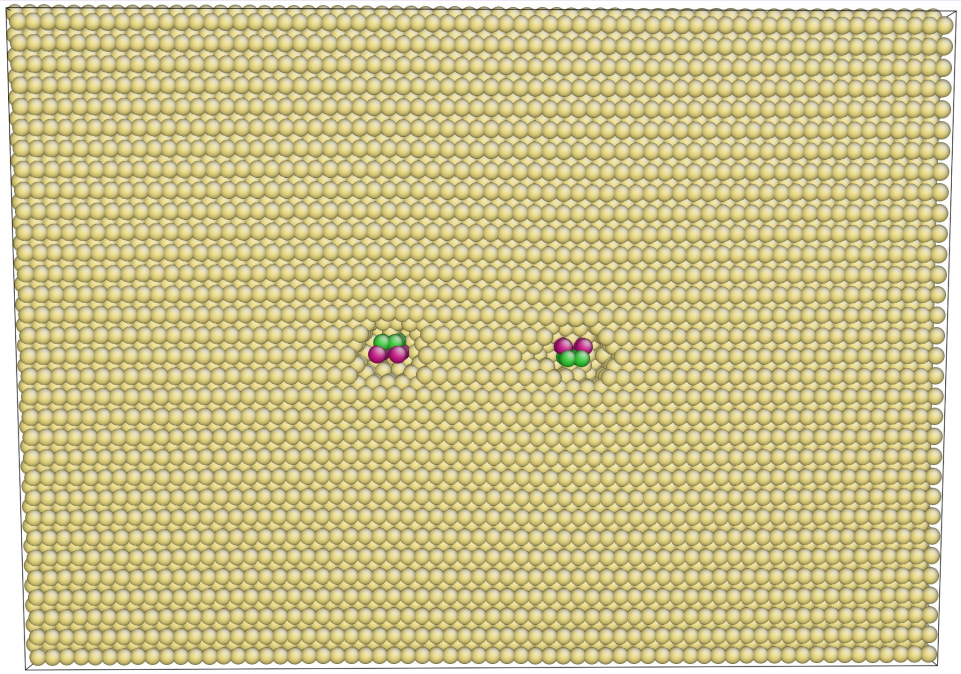
Lomer Dislocations[1]

$\uparrow$ [001] 387.1 Å Dirichlet B.C.

$\rightarrow$ [110] 410.6 Å Dirichlet B.C.

$\otimes[\bar{1}10]$ 11.4 Å Periodic B.C. (110592 atoms)

Shear $\longrightarrow$



$\longleftarrow$ Shear

[1] Miller and Tadmor, Modelling Simul. Mater. Sci. Eng. 17(2009), 053001.

Continued...

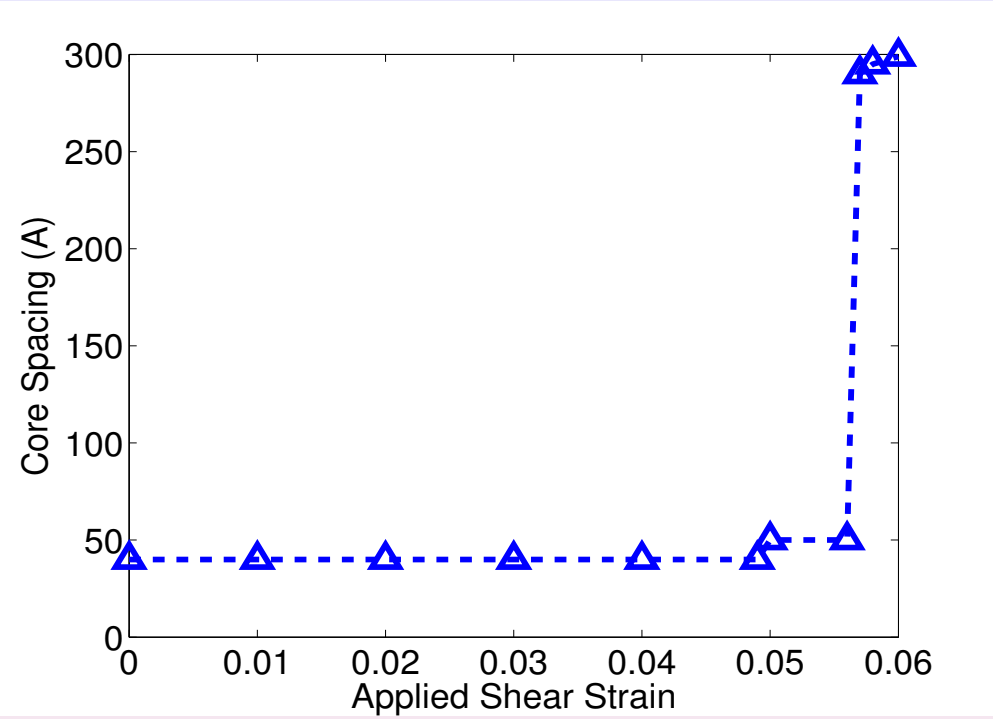


Fig. : Distance between dislocation cores vs. shear strain (γ)

- Dislocations move slightly until $\gamma \approx 0.05$;
- Spacing increases from 40 Å to 50 Å until $\gamma \approx 0.056$;
- Dislocations move a long distance at $\gamma = 0.057$;
- Beyond $\gamma = 0.057$ spacing is more or less constant due to boundary conditions.

Summary and Perspectives

- We have developed an effective strategy for multiscale modelling based on full multigrid.
- Both methodologies provide a general description of the atomistic model at coarse levels (*systematic coarsening*).
- Preliminary results seem encouraging.

  - Needs to be tested in more complex situations.
  - Combine with coarse-grained elasticity model (as in Quasicontinuum).
  - Can this be extended to noncrystalline systems? Polymers?
  - Finite temperature extension? Could use ideas from Renormalization Multigrid or Monte Carlo Multigrid.