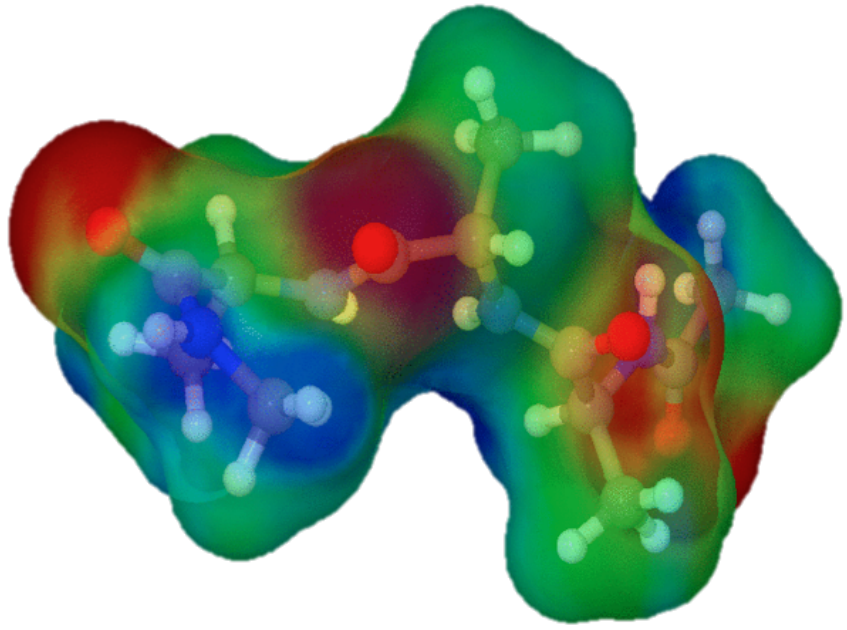
TMS 143rd Annual Meeting, San Diego, Feb. 16-20, 2014

# Smart use of Density Functional Theory calculations to drive Newtonian dynamics

SAND2014-1079C

**Reese Jones** and Mickey Shaugnessy

*Sandia National Laboratories and Synopsis Corporation*



Sandia National Laboratories is a multiprogram laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

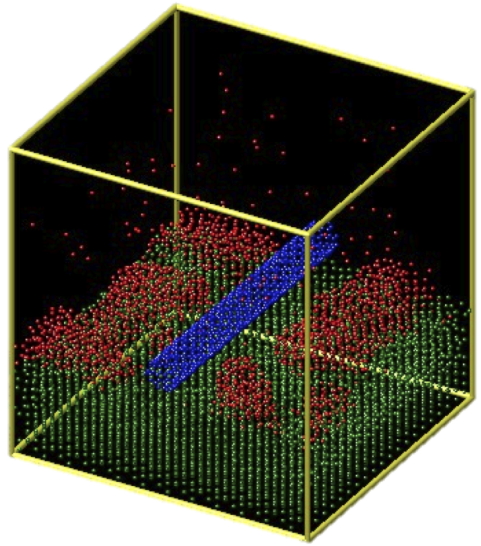
# Outline

**Motivation:** *compositionally* and *structurally complex* materials and processes: e.g. deposition and radiation damage, require the dynamic simulation of *large* atomic systems with *multiple* interspecies interactions.

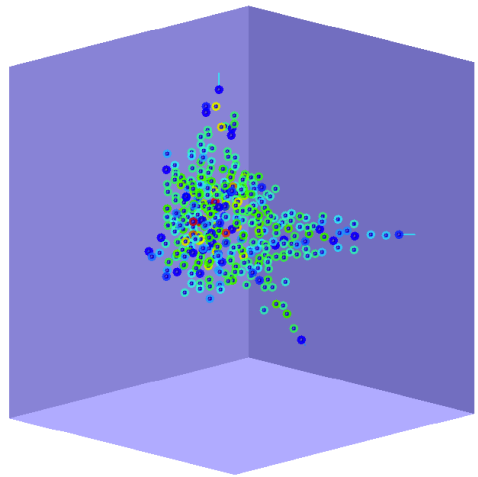
Outline:
- Basic idea
- Distance metric
- Metric search
- Force interpolation
- Results
- Conclusion

please ask questions

atomic deposition



radiation damage



unfortunately the last pretty picture for today

# Basic idea

**Premise:** we can decompose the system into clusters and if the clusters are sufficiently large the forces on the central atoms are sufficiently accurate.

**Procedure:**

1. as dynamics ensues new/query configurations are generated and compared to a cluster-force database.
2. if sufficient stored clusters are close the query cluster the stored forces are interpolated at the new/query cluster
3. otherwise, *ab initio* forces are calculated for the new cluster and stored

As opposed to the *globally*-tuned *empirical* potential, we explicitly propogate the system in time using locally interpolated DFT forces.

# Cluster distance

A *cluster* $\mathcal{C}_A$ is a set $\mathcal{C}_A = \{\Delta\mathbf{x}_{1a}, \Delta\mathbf{x}_{2a}, \Delta\mathbf{x}_{3a}, \ldots, \Delta\mathbf{x}_{N_A a}\}$ of distance vectors $\Delta\mathbf{x}_{\alpha a} \equiv \mathbf{x}_\alpha - \mathbf{x}_a$ relative to the central atom $\mathbf{x}_a$,

A distance $d(A, B)$ between clusters $\mathcal{C}_A$ & $\mathcal{C}_B$, needs (a) basic metric properties:

M1 **coincidence**: $d(A, B) = 0$ iff $A = B$

M2 **positivity**: $d(A, B) > 0$

M3 **symmetry**: $d(A, B) = d(B, A)$

M4 **triangle inequality**: $d(A, C) \geq d(A, B) + d(B, C)$

M5 **reverse triangle inequality**: $d(A, C) \leq |d(A, B) - d(B, C)|$

and (b) physical invariances $d(A, B) = d(A, B')$:

I1 **translation** $\mathcal{C}_B \to \mathcal{C}_{B'} = \mathcal{C}_B + \mathbf{a}$

I2 **rotation** $\mathcal{C}_B \to \mathcal{C}_{B'} = \mathbf{R}\mathcal{C}_B$

I3 **permutation** $\mathcal{C}_B \to \mathcal{C}_{B'} = \mathbf{P}\mathcal{C}_B$

# Root Mean Square Distance

Assuming $N_A = N_B$, the root mean square deviation (RMSD) comparison metric is

$$\begin{aligned} d_{\mathrm{RSMD}}(A,B) &= \min_{\mathbf{R},\mathbf{P}} \|\mathrm{X}_A - \mathbf{R}\mathbf{P}\mathrm{X}_B\|_w \\ &= \min_{\mathbf{R},\mathbf{P}} \sqrt{(\mathrm{X}_A - \mathbf{R}\mathbf{P}\mathrm{X}_B)\cdot \mathrm{W}_{AB}\,(\mathrm{X}_A - \mathbf{R}\mathbf{P}\mathrm{X}_B)} \\ &\equiv \min_{\mathbf{R},\mathbf{P}} \sqrt{\sum_{\alpha,\beta=1}^{N_b} \|\Delta\mathbf{x}_{\alpha a} - \mathbf{R}P_{\alpha\beta}\Delta\mathbf{x}_{\beta b}\|^2 \; w_{\alpha\beta}} \end{aligned} \tag{1}$$

where:

- $\mathrm{X}_A$ and $\mathrm{X}_B$ are matrices of the cluster vectors,
- $\mathbf{R} \in \mathrm{Orth}^+$ is a rotation of the cluster and
- $\mathbf{P}$ is a permutation of the cluster ordering,

i.e. a binary orthogonal matrix which is simply the rearrangment of the rows or columns of the identity matrix.

# Optimal rotation

$$d(A,B) = \min_{\mathbf{R},\mathbf{P}} \sqrt{(\mathrm{X}_A - \mathbf{RP}\mathrm{X}_B) \cdot \mathrm{W}_{AB} (\mathrm{X}_A - \mathbf{RP}\mathrm{X}_B)}$$

$$= \min_{\mathbf{R},\mathbf{P}} \sqrt{\|\mathrm{X}_A\|_w^2 + \|\mathrm{X}_B\|_w^2 - 2\mathrm{X}_A \cdot \mathrm{W}_{AB}\mathbf{RP}\mathrm{X}_B}$$

To determine the rotation $\mathbf{R}$, Kabsch (1976) noticed the last term

$$2\mathrm{X}_A \cdot \mathrm{W}\mathbf{RP}\mathrm{X}_B = 2\,\mathrm{tr}\left(\mathbf{RP}\mathrm{X}_B\mathrm{X}_A^T\mathrm{W}_{AB}\right)$$

is the only one dependent on $\mathbf{R}$.

A singular value decomposition $\mathrm{SVD}[\mathbf{P}\mathrm{X}_B\mathrm{X}_A^T\mathrm{W}_{AB}] = \mathrm{UDV}^T$ gives the solution $\mathbf{R} = \mathbf{VU}^T$ and

$$d(A,B) = \min_{\mathbf{P}} \sqrt{\|\mathrm{X}_A\|_w^2 + \|\mathrm{X}_B\|_w^2 - 2\,\mathrm{tr}\,\mathbf{D}(\mathbf{P})}$$

# Distance from Gaussian densities

Finding the optimal *permutation* **P** is considerably harder,

As in Bartok (2013), a permutation invariant metric can be formed from atomic densities

$$\rho(\mathbf{x}) = \Delta(\mathbf{0}) + \sum_{\alpha} \Delta(\mathbf{x} - \mathbf{x}_\alpha) = \sum_{\alpha} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} c_\alpha^{lm}(r) Y_{lm}(\mathbf{r})$$

represented by a spherical harmonic expansion, where

$$\Delta(\mathbf{x}) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}_\alpha\|^2}{2\sigma^2}\right)$$

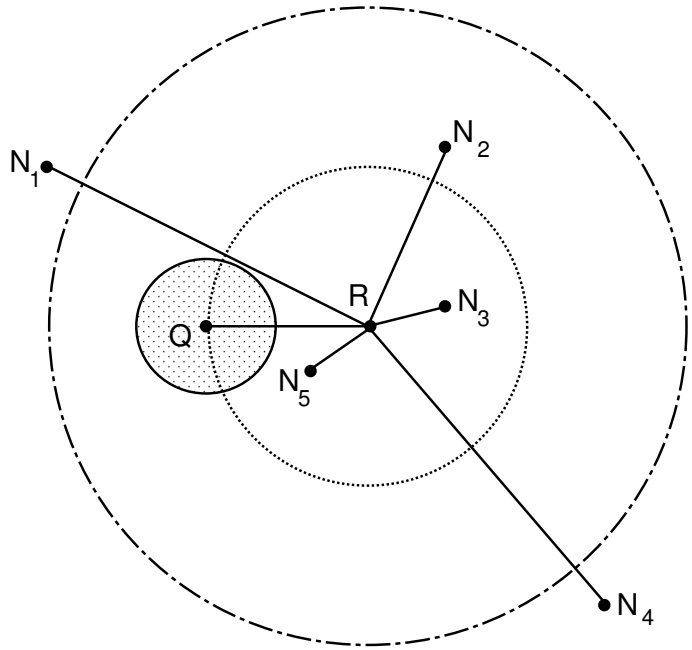
And so

$$d_{\mathrm{SOAP}}(\rho_A, \rho_B) = \sqrt{\frac{S(\rho_A, \rho_A)S(\rho_B, \rho_B) - S(\rho_A, \rho_B)^2}{S(\rho_A, \rho_B)^2}}$$

where $S(\rho_A, \rho_B) = \int d\mathbf{x}\, \rho_A(\mathbf{x})\rho_B(\mathbf{Rx})$

# Metric database search

- Select a set of points $\{R_i\}$ in the database randomly or from previous time step.
- Loop
  1. Find closest in set: $\text{argmin}_R\, d(Q, R_i)$, where Q is the query configuration
  2. *Retrieve* neighboring set $\{N\}$ of points about the point $R$ for which $2d(N_i, R) < d(Q, R)$
  3. Stop if the neighborhood is the desired radius or size.




Query point Q, reference point R and its neighbors $N_i$. The subset { $N_2$, $N_3$, $N_5$ } are candidates for configurations in the interpolation ball around Q. After computation of distances of the subset to Q, $N_5$ will be selected as the best candidate for the new R.

# Interpolation of forces

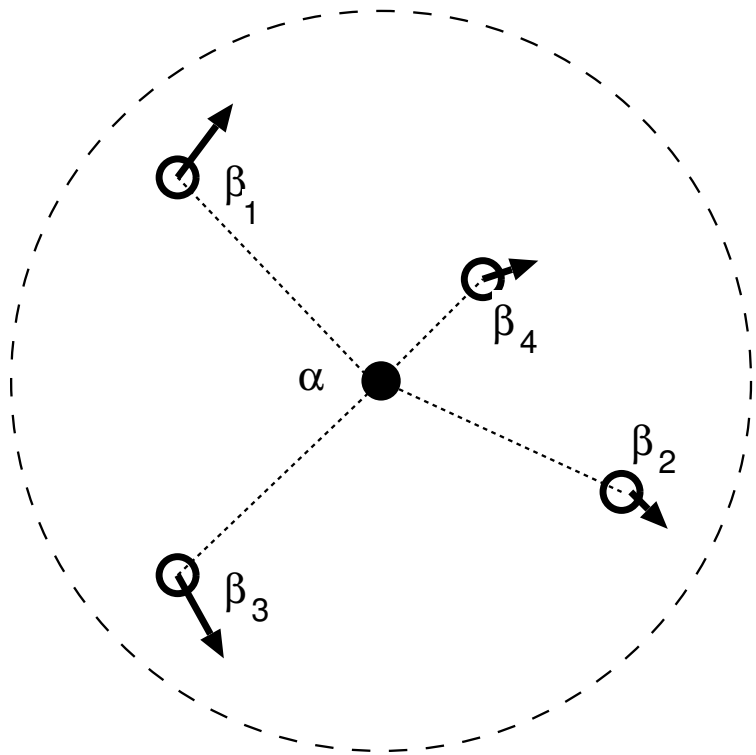
Given $\{d_{\alpha\beta}\} \mid \beta \in \mathcal{B}_\alpha$ for $\alpha$ an accurate force $\mathbf{f}_\alpha$ can be obtained from simple interpolation

$$\mathbf{f}_\alpha = \sum_{\beta \in \mathcal{B}_\alpha} \mathbf{R}_\beta \mathbf{f}_\beta \phi(d_{\alpha\beta})$$

such that $\sum_\beta \phi(d_{\alpha\beta}) = 1$ and $\phi(r) \sim \frac{1}{r}$.

Here $\mathbf{R}_\beta$ is an optimal rotation from the reference cluster $\beta_i$ to the query cluster $\alpha$, obtained by

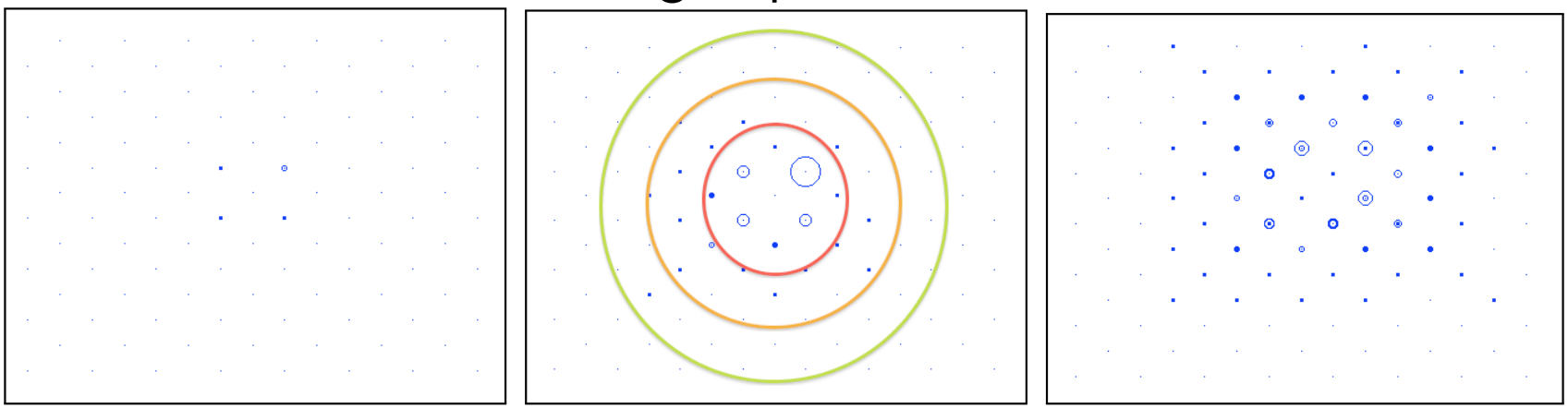
$$\text{argmin}_{\mathbf{R}}\, d(C_\alpha, \mathbf{R}C_\beta)$$



# Convergence of force with cluster size

Magnitude of the Hellmann-Feynman force for a 512 atom Si supercell

Increasing displacement $\Longrightarrow$



Error is $\|\mathbf{f}_{\mathrm{Ball}(R)} - \mathbf{f}^*\|$ where $\mathbf{f}^*$ is the force in the full system.

# Correlation of cluster distance and forces

The difference in forces on the central atom for two clusters

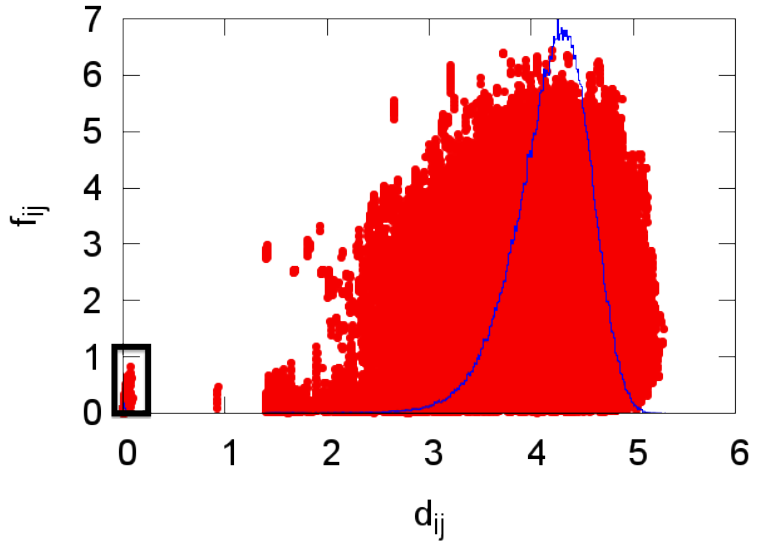
$$\|\mathbf{f}_\alpha - \mathbf{f}_\beta\|$$

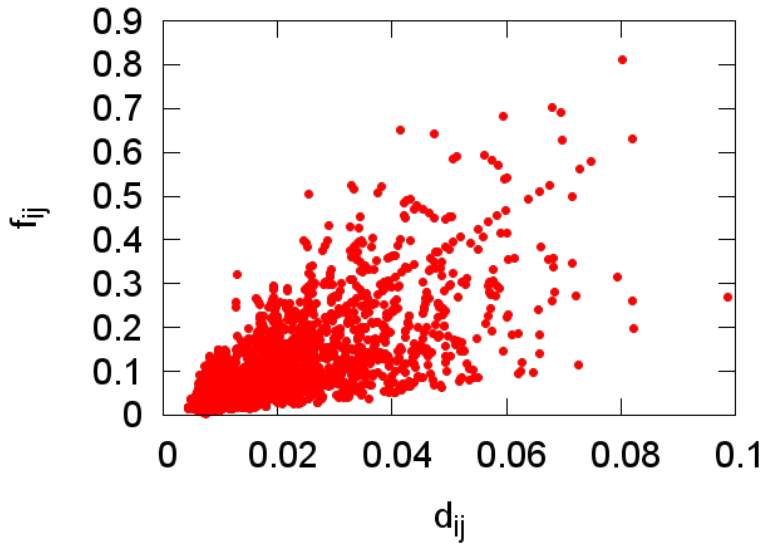
is highly correlated with with the inter-cluster distance $d_{\alpha\beta}$.

The blue line depicts the probability density - showing that the mostly likely difference is quite peaked.

Also the candidates for interpolation in the **black** box are separated from the bulk of the clusters.

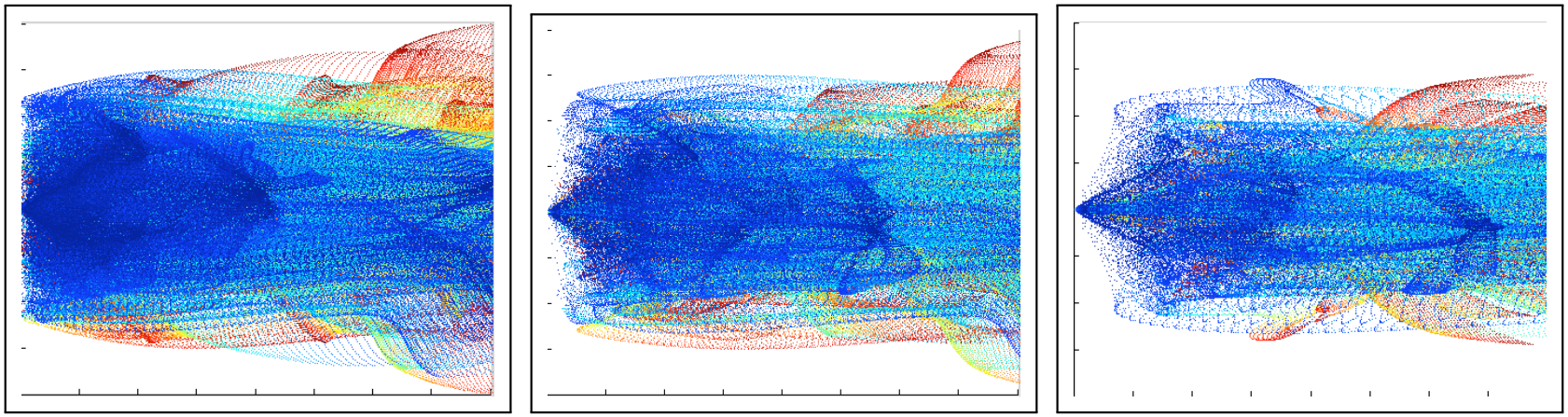
The close-up shows that the correlation is quite linear



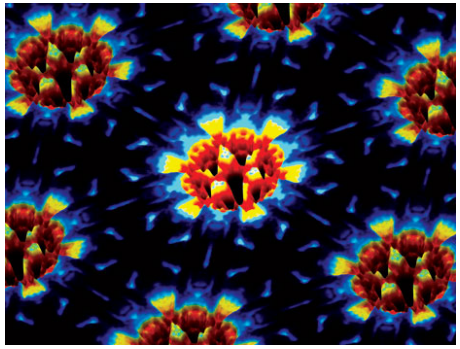
# Sensitivity to the cluster size

Force-distance correlation for increasing cluster size $\Longrightarrow$



# Current work

- We have results for the dynamics of 1- and 3-D test systems but they are aren't that interesting since we have only been testing consistency of the method and full *ab initio* and classical molecular dynamics.
- Application to atomic deposition, radiation damage, and structurally interesting materials like topological insulators is upcoming



- Since the method reduces to *ab initio* molecular dynamics in the limit to full system clusters and no interpolation, we are trying to make stronger connections to it
- We are also exploring using machine learning techniques to enhance the search and interpolation methodology.

Reese Jones **rjones@sandia.gov**