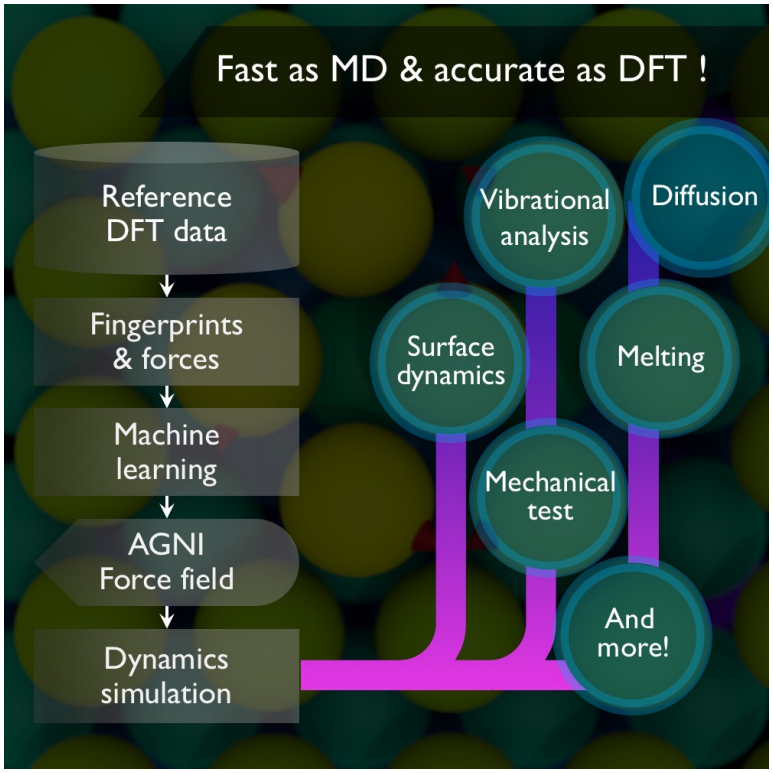
Machine Learning Based Atomistic Force Fields

Rohit Batra, Huan Tran, Venkatesh Botu, James Chapman and **Rampi Ramprasad**

School Of Materials Science And Engineering, Georgia Institute Of Technology, GA



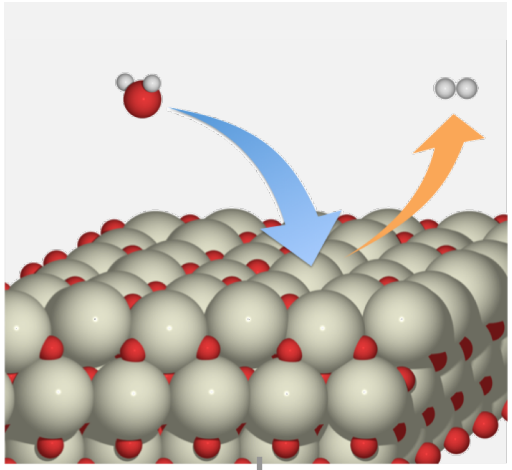
NIST Workshop on Atomistic Simulations for Industrial Needs

1st August, 2018

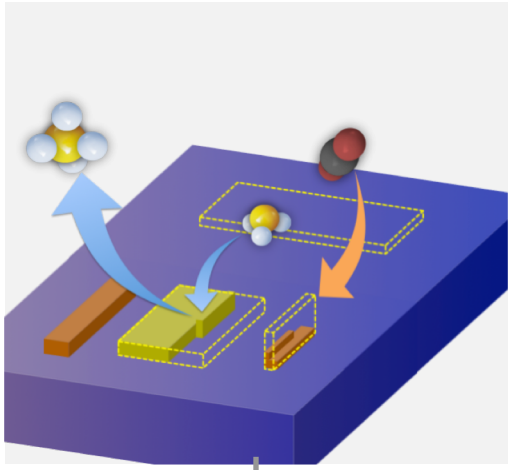
MOLECULAR DYNAMICS

Used for *in silico* study of several chemical and materials science phenomena

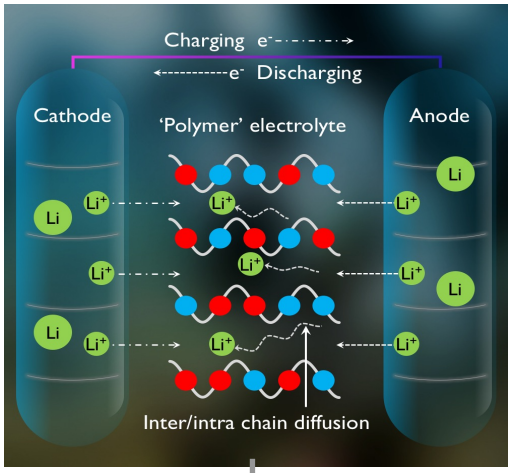
Catalysis



Materials growth



Electrochemistry



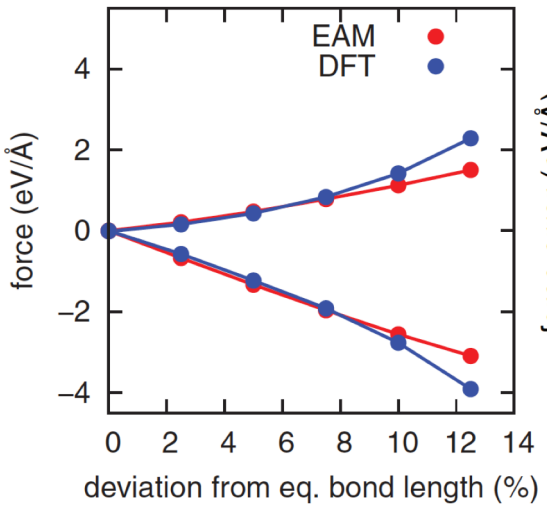
Input: Atomic forces $\{ F_i(R_{ij}) \}$ and time-step

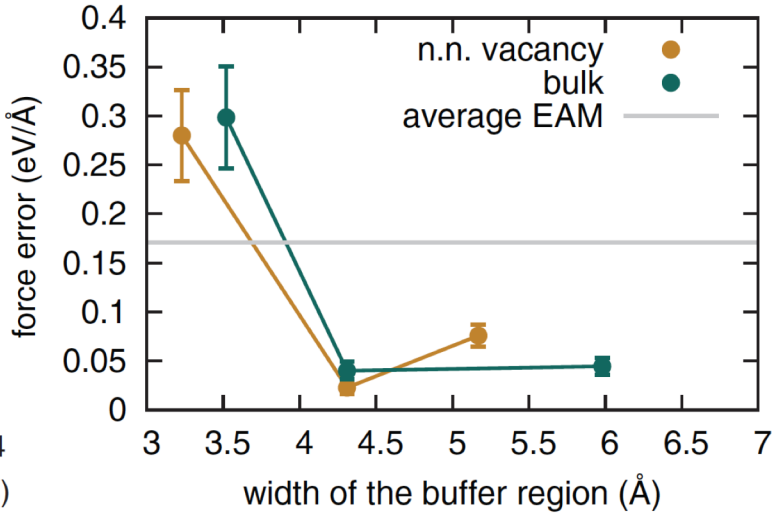
Quantum mechanics | Semi-empirical | Machine learning

LIMITATIONS OF CLASSICAL POTENTIALS

Non-transferable

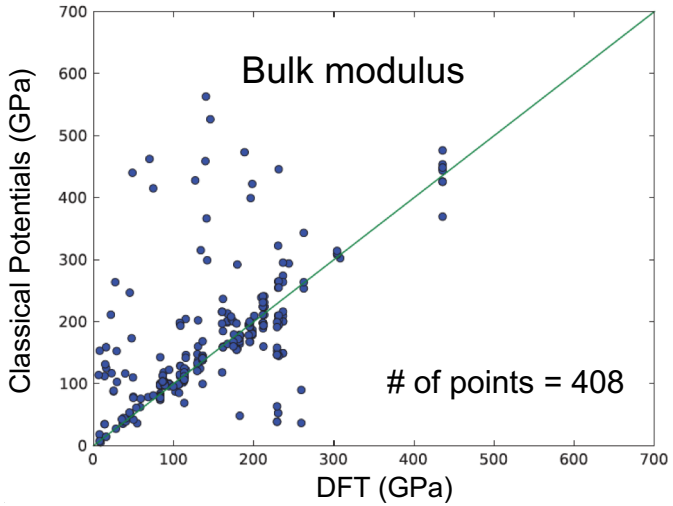
Inaccurate forces in regimes far from equilibrium





Non-generalizable

No common potential functional form across materials



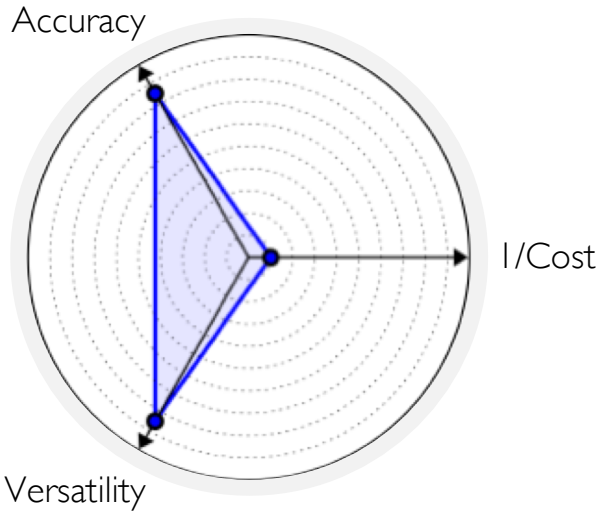
Potential fit for binary Al-Cu system might not be accurate for elemental Al, Cu.

Bianchini et al., Modelling Simul. Mater. Sci. Eng. 24 045012 (2016), Choudhary et al., Scientific Data, 4 (2017)

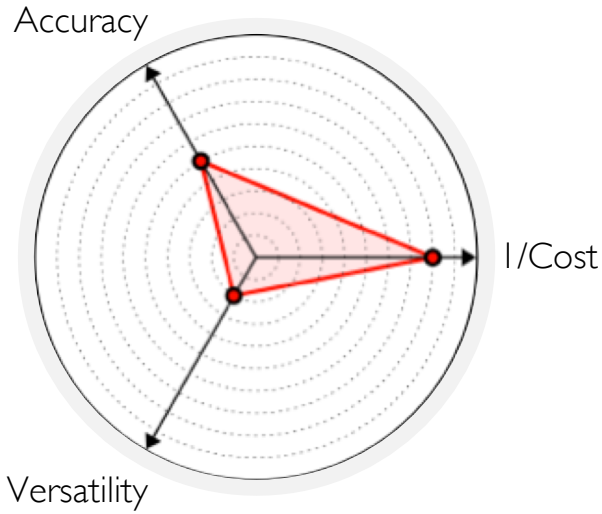
MOTIVATION

ML force fields provide balance between cost, accuracy and versatility

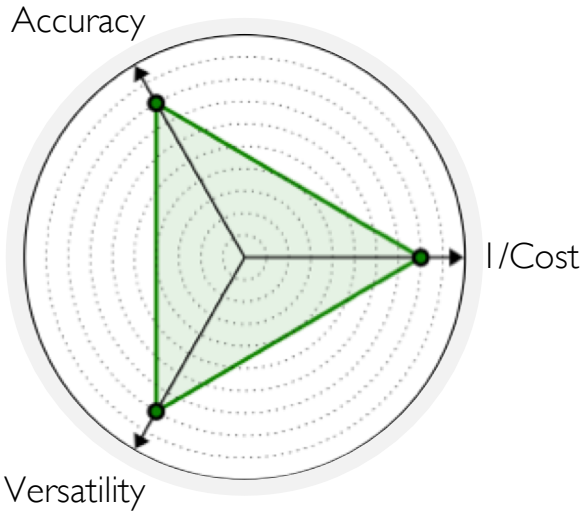
Quantum mechanics



Semi-empirical



Machine learning



V. Botu, "Surface Chemistry with Machine Learning and Quantum Mechanics", Doctoral Dissertation, University of Connecticut (2016)

MACHINE LEARNING: AN EXAMPLE

“ What should be the credit limit ? ”

Index	Feature	Value
1	Age	23
2	Gender	Male
3	Salary	$30,000
4	Years in job	1 year
…	…	…
n	Current debt	$15,000

$y = F(X)$

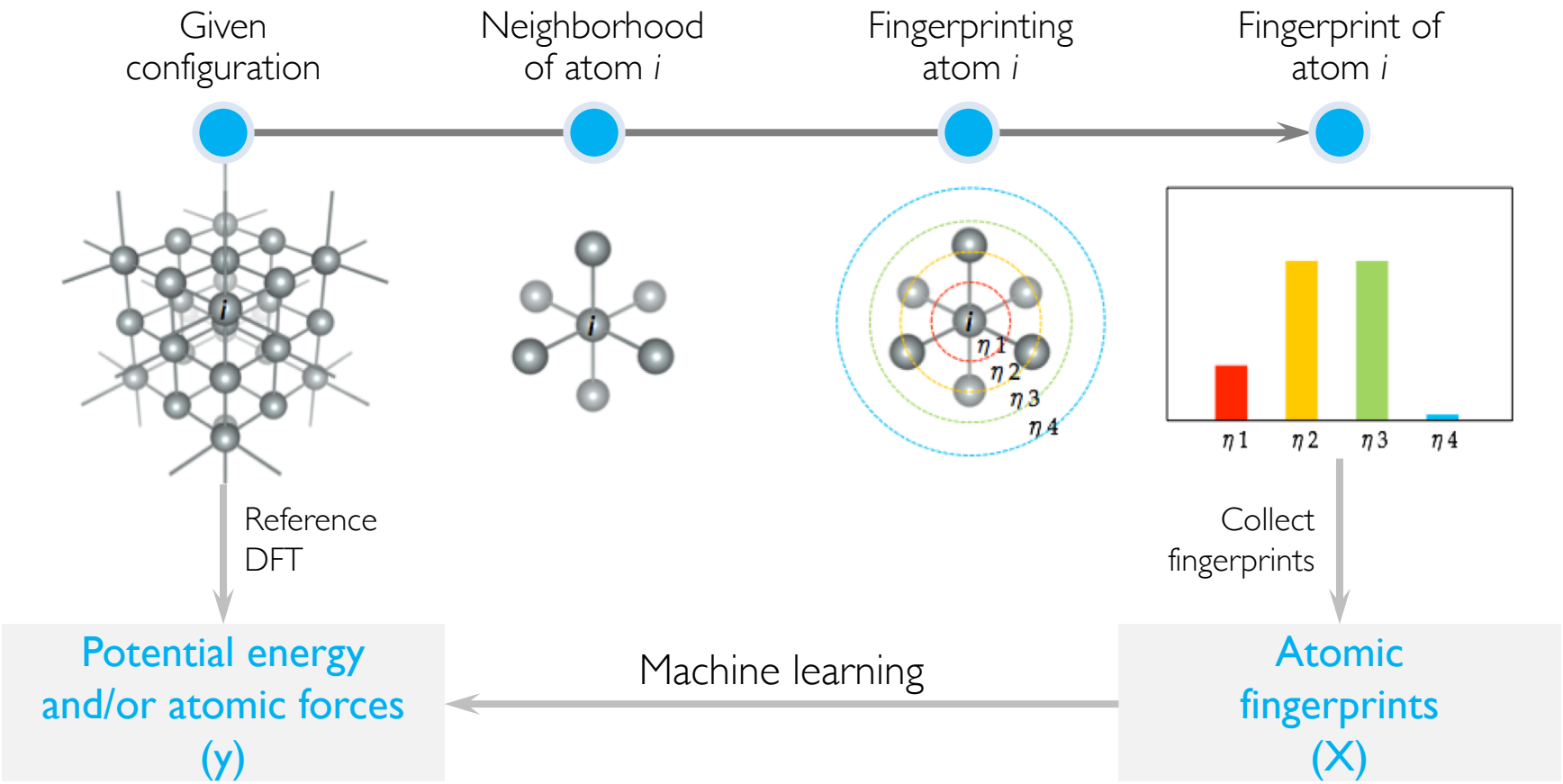
Machine learning finds approximate function

X				y
Age	Gender	…	Current debt	Credit limit
25	M	…	$10,000	$5,000
19	F	…	$50,000	$1,000
…	…	…	…	…

$y_{predict} = F(X_{new})$

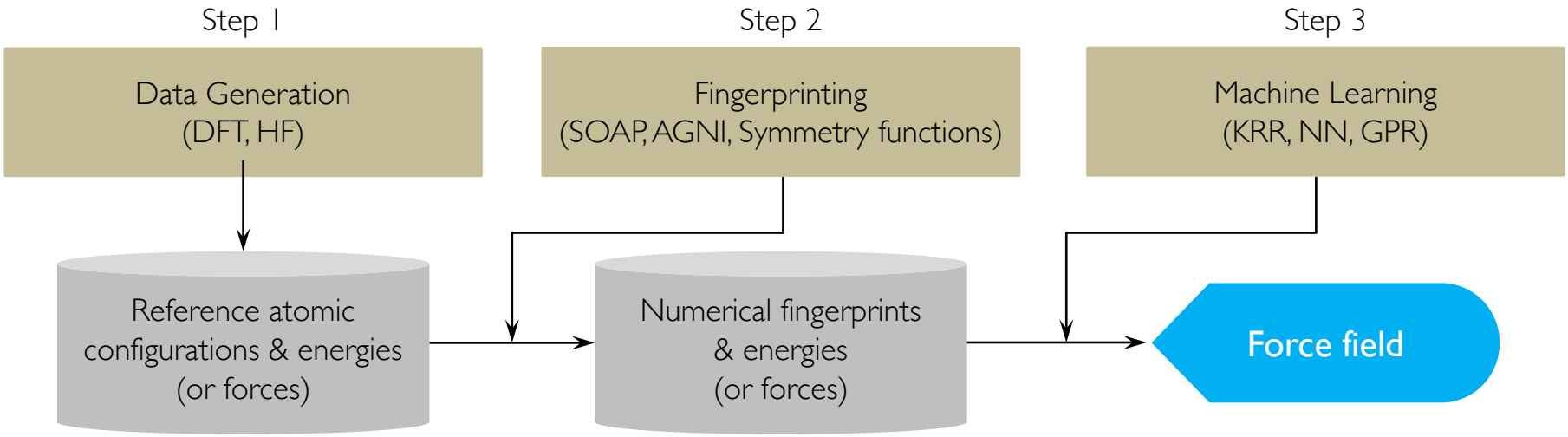
30	M	…	$20,000	?
25	F	…	$15,000	?

ML FORCE FIELDS: KEY CONCEPT



R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakkithodi, C. Kim,
Machine Learning and Materials Informatics: Recent Applications and Prospects, npj Computational Materials 3, 54 (2017)

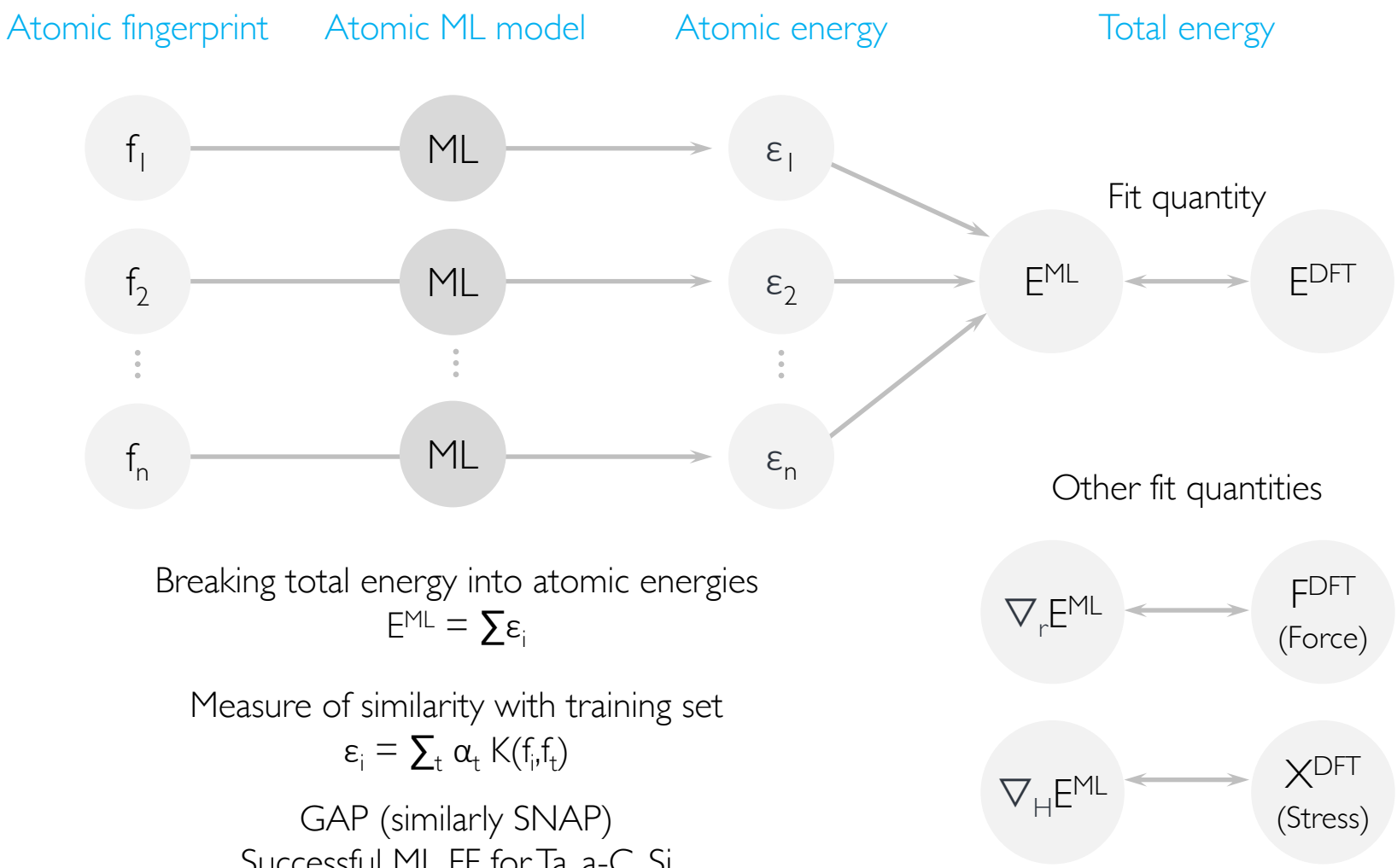
ML FORCE FIELDS: METHODOLOGY



1. Reference data should be exhaustive (as ML is interpolative)
2. Atomic Fingerprint is the key
3. Choose ML method based on fingerprint and amount of reference data

R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakkithodi, C. Kim,
Machine Learning and Materials Informatics: Recent Applications and Prospects, npj Computational Materials 3, 54 (2017)

ML FORCE FIELDS: LEARNING ENERGY



Breaking total energy into atomic energies

$E^{ML} = \sum \varepsilon_i$

Measure of similarity with training set

$\varepsilon_i = \sum_t \alpha_t K(f_i, f_t)$

GAP (similarly SNAP)

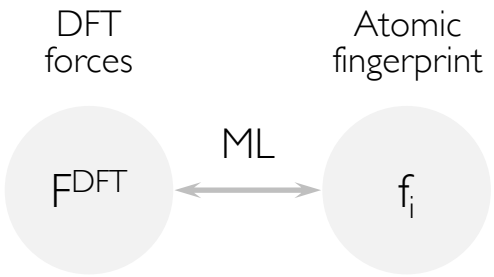
Successful ML FF for Ta, a-C, Si

Behler, J. Chem. Phys. 145, 170901 (2016), Thompson et al., J. Comput. Phys. 285, 316 (2015), Deringer PRB 95, 094203 (2017)

ML FORCE FIELDS: LEARNING FORCES

Learn atomic forces directly

Successful ML FF for Al and Si



	Pros	Cons
Energy-based ML FF	• Well-understood (available codes) • Momentum conserved for isolated system	• Based on ill-defined atomic energies • Indirect access to forces (slow) • Difficult to fit forces accurately (e.g. planar defects)
Force-based ML FF	• Direct access to forces (fast) • Based on well-defined atomic forces • Easy to fit forces for complex structures	• Indirect access to total energy • Noise in ML predictions (non-zero net force) • Difficult to access stresses / pressures

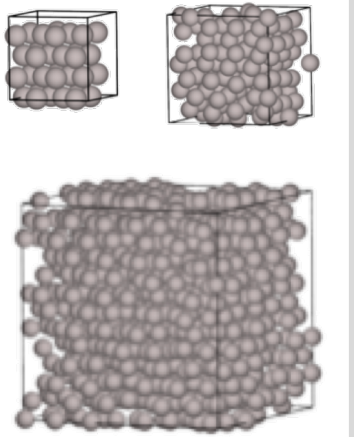
Our method: AGNI

Botu et al., J. Phys. Chem. C, 121 (1), 511 (2017), Li et al., PRL 114, 096405 (2015), Botu et al., Phys. Rev. B, 92, 094306 (2015)

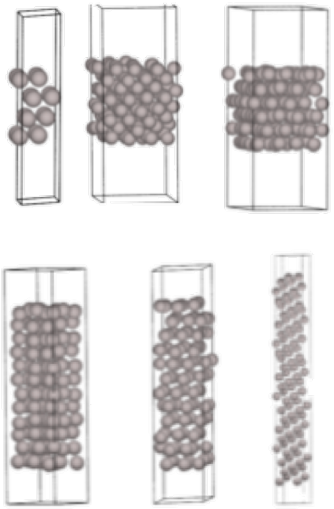
STEP 1. REFERENCE DATA

Periodic & non-periodic configurations of Al

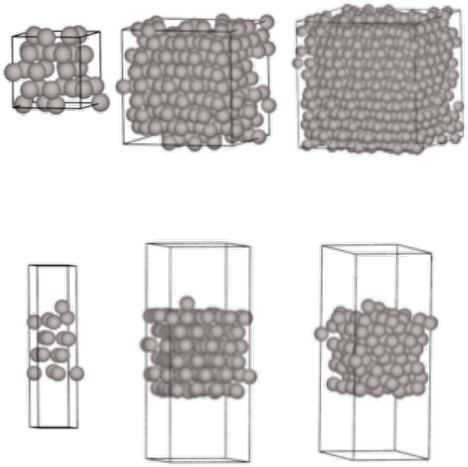
Bulk



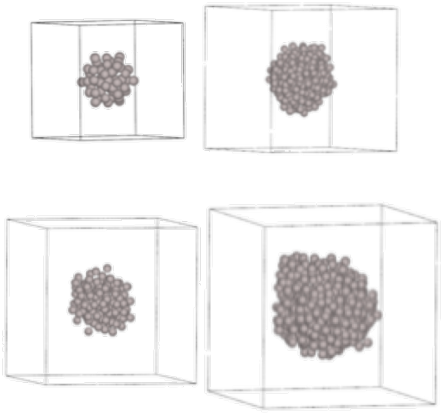
Surface



Defect & surface features

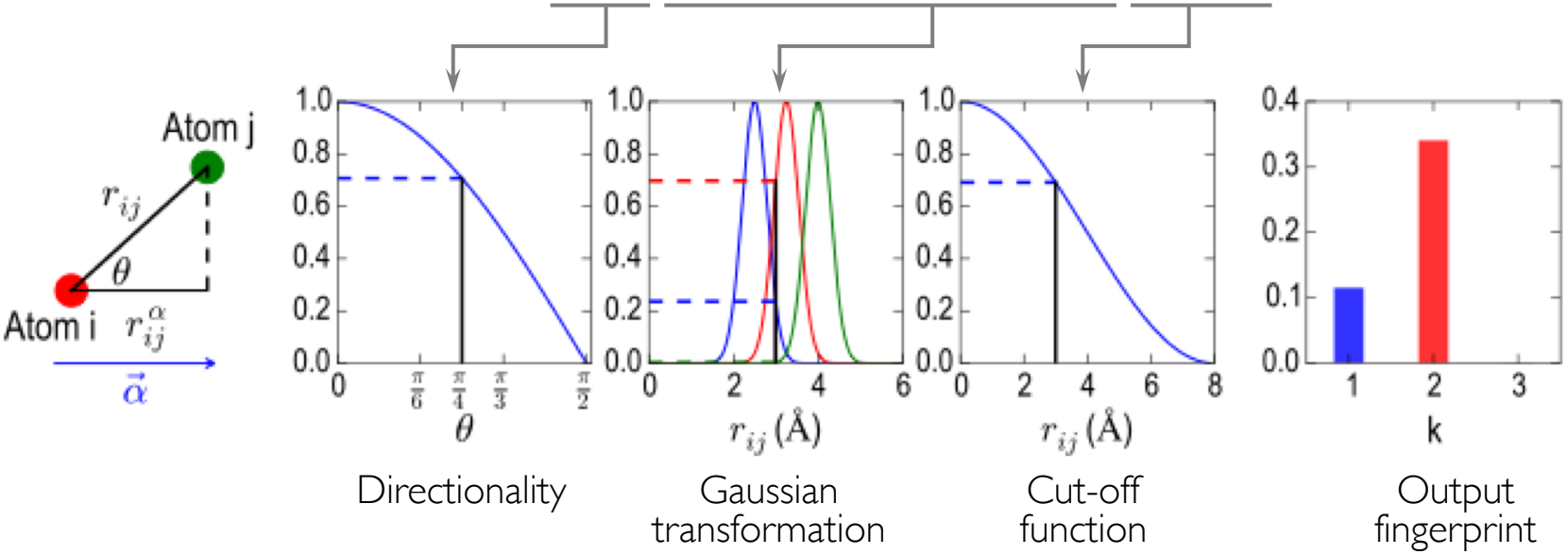


Cluster



STEP 2. AGNI FINGERPRINT

$$V_{i,\alpha;k} = \sum_{i \neq j} \frac{r_{ij}^{\alpha}}{r_{ij}} \exp\left[-\frac{1}{2}\left(\frac{r_{ij} - a_k}{w}\right)^2 \right] f_c(r_{ij})$$



Mathematically complete fingerprint to numerically and uniquely represent any configuration

Huan et al., npj comput. mater., 3, 27 (2017)

STEP 3. MACHINE LEARNING

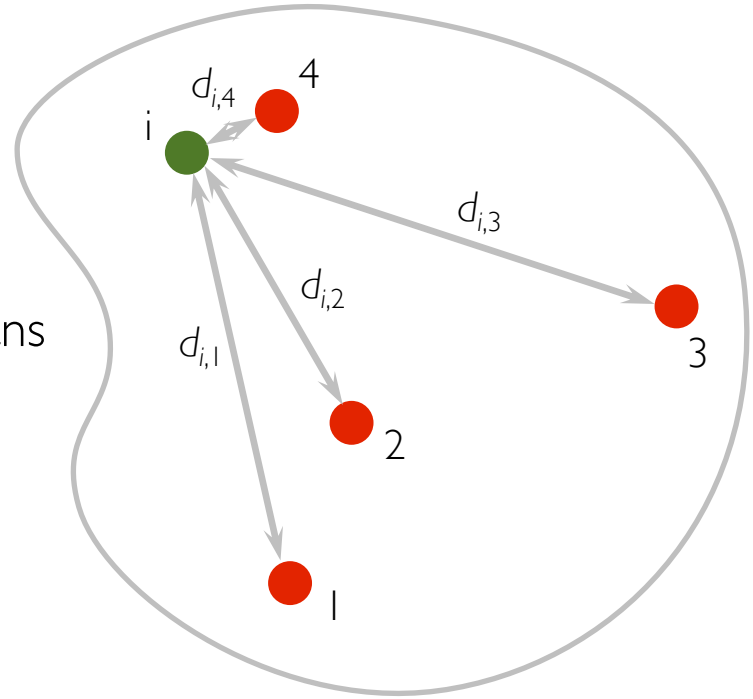
Kernel ridge regression → Measures (dis)similarity amongst data

Measure of similarity: Euclidean distance

$$d_{i,t}^{u} = ||V_i^u(\eta) - V_t^u(\eta)||$$

Property Estimation: Sum of weighted Gaussians

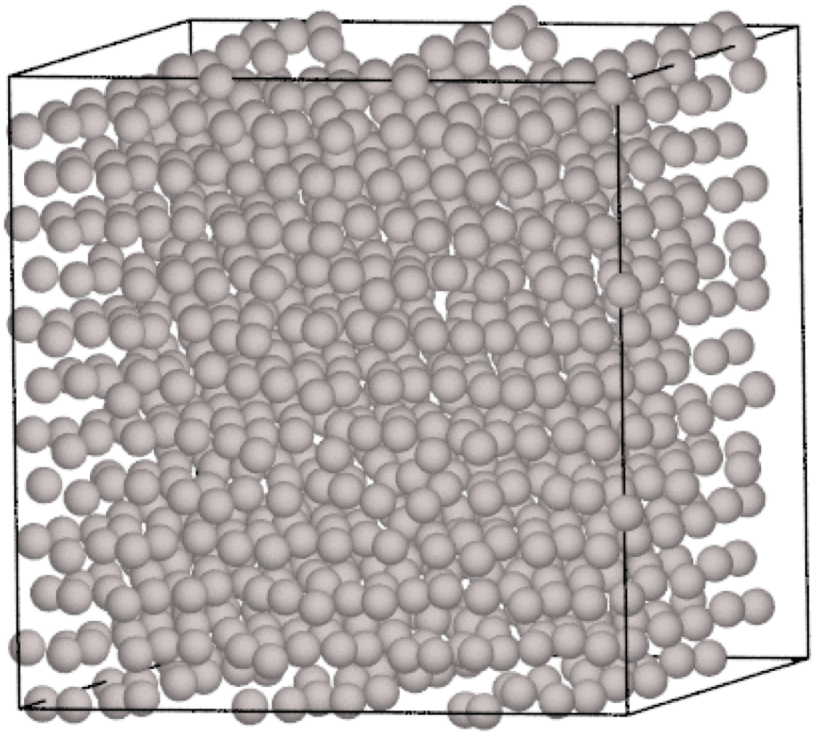
$$F_i^u = \sum_t^{N_t} \alpha_t \cdot \exp\left[-\frac{\left(d_{i,t}^u\right)^2}{2l^2}\right]$$



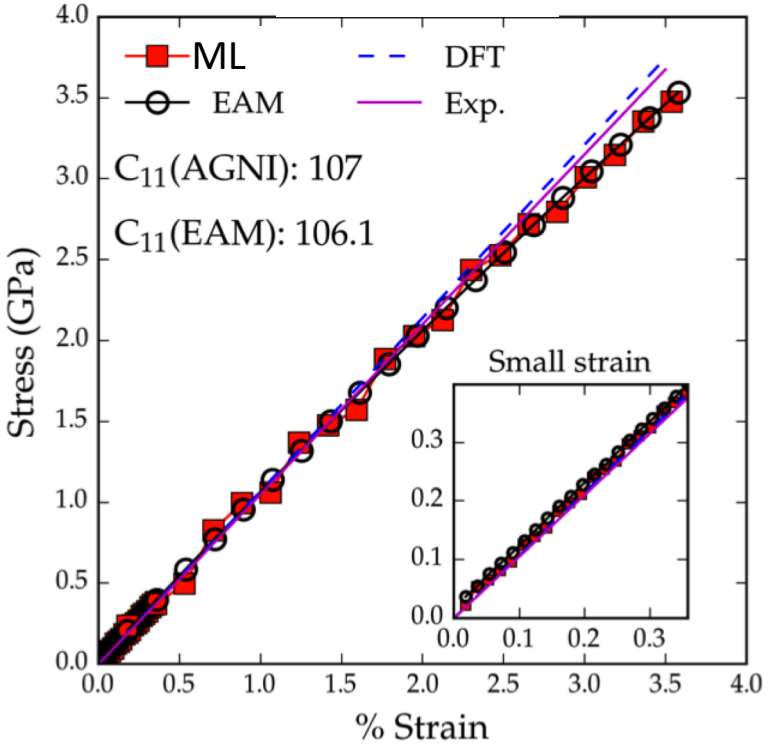
AGNI scheme compatible with other ML methods as well

AGNI: STATIC TESTS

Geometry optimization



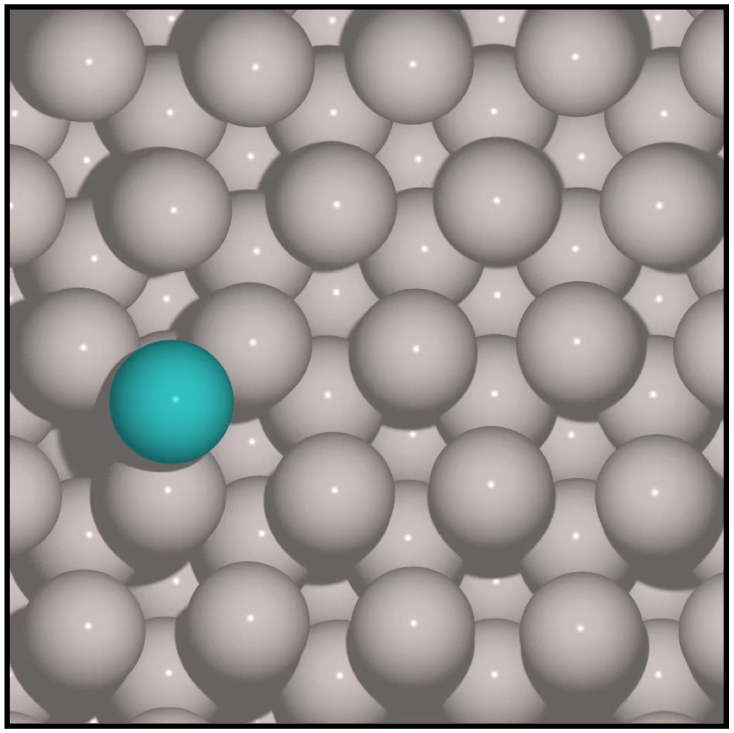
Mechanical behavior



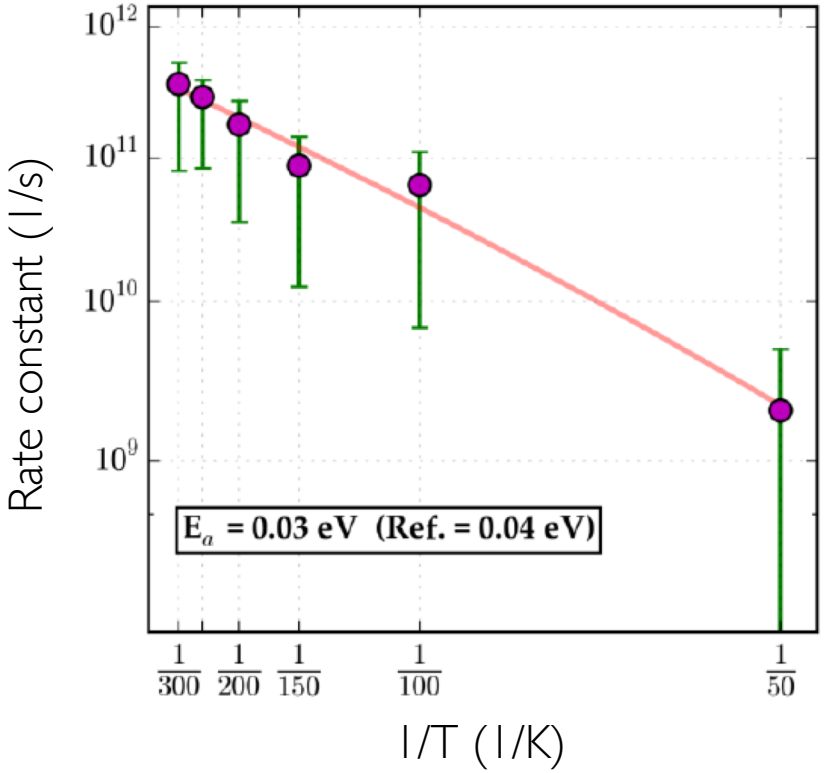
Botu et al., J. Phys. Chem. C, 121 (1), 511 (2017), Botu et al., IJQC 115, 1074 (2015)

AGNI: MODELING DIFFUSION

Captures underlying dynamics accurately



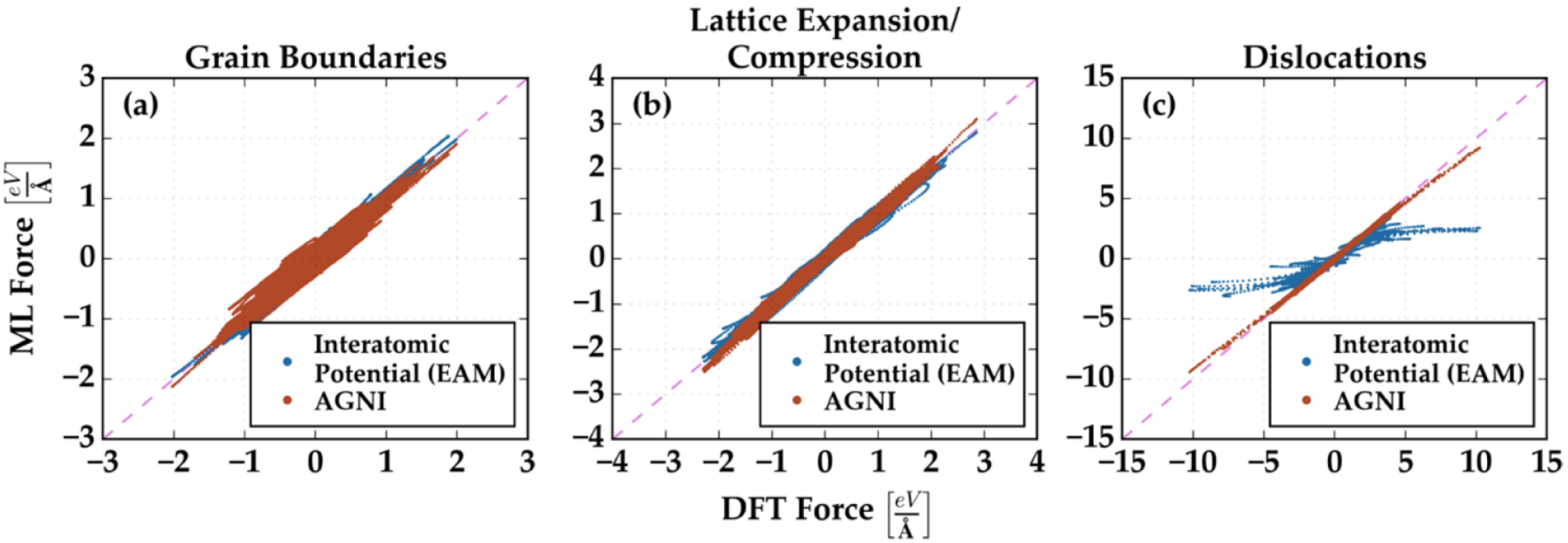
Self-diffusion on Al (111) surface



Botu et al., Phys. Rev. B, 92, 094306 (2015)

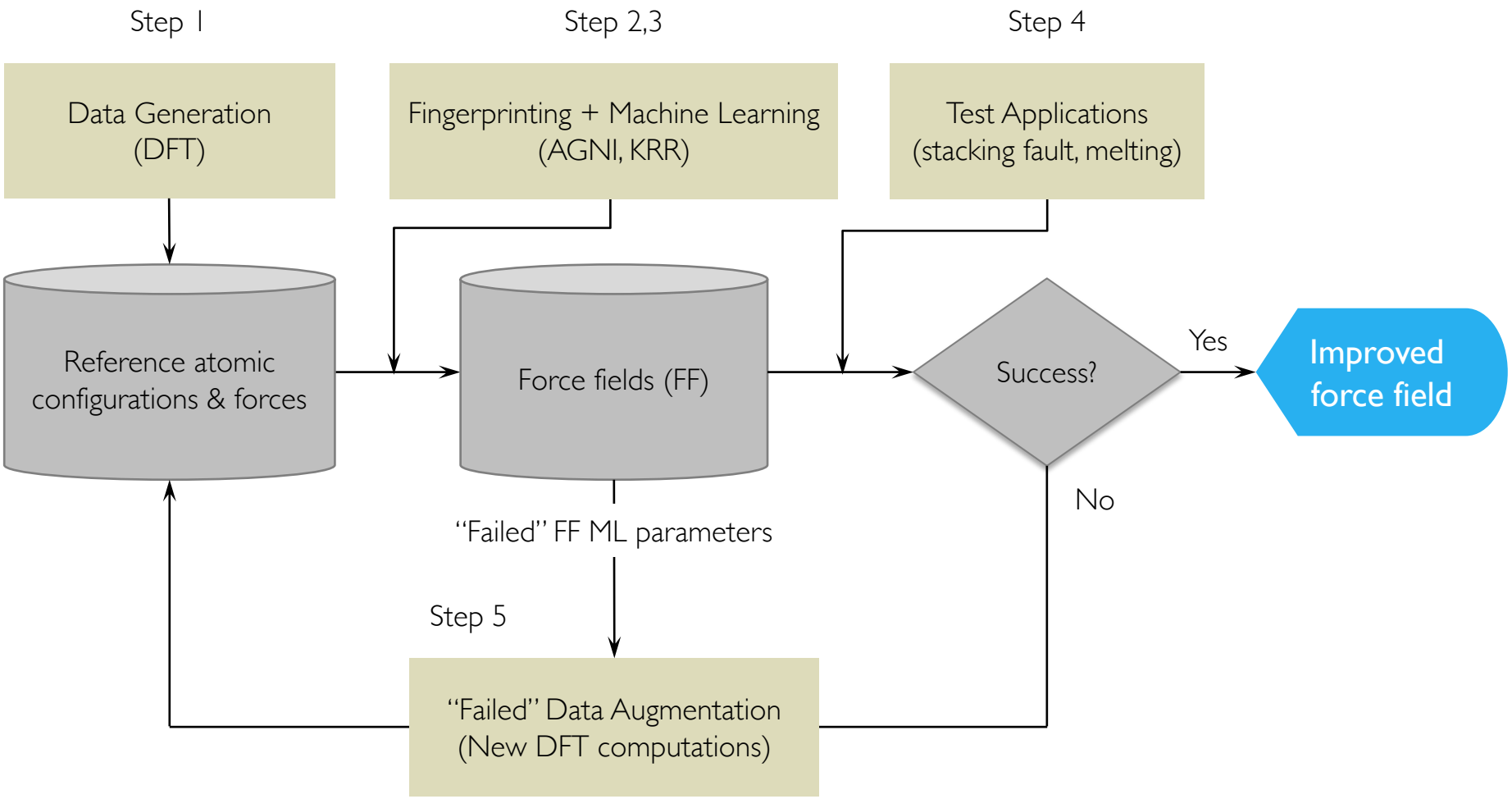
AGNI: BEYOND "TRAINING" DATA

Learning in fingerprint space



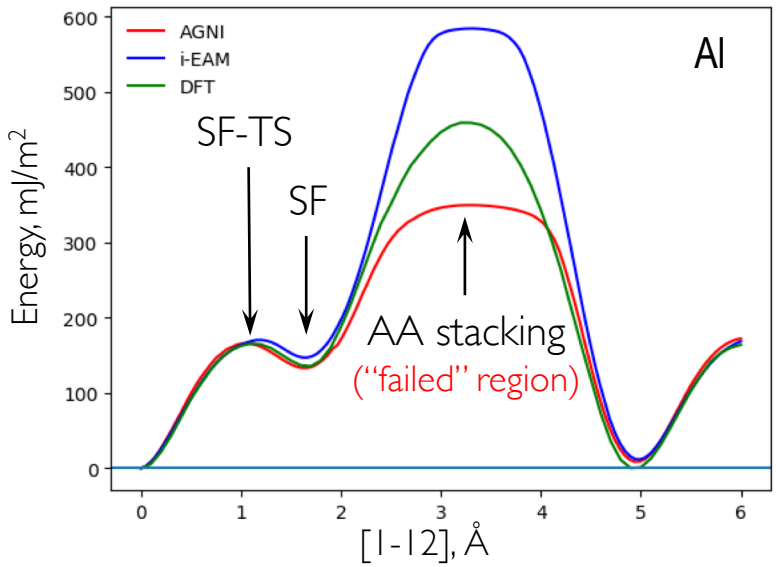
Botu et al., J. Phys. Chem. C, 121 (1), 511 (2017)

POWER OF ML: ACTIVE LEARNING



Development of application "targeted" ML force field

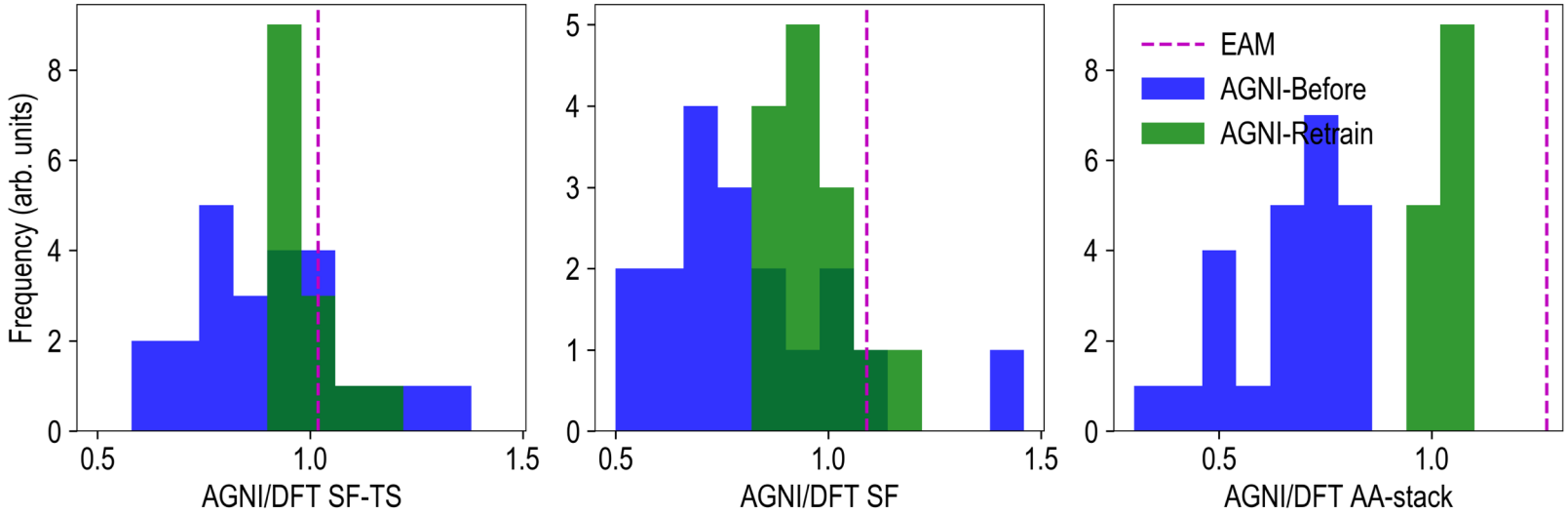
ACTIVE LEARNING: STACKING FAULT



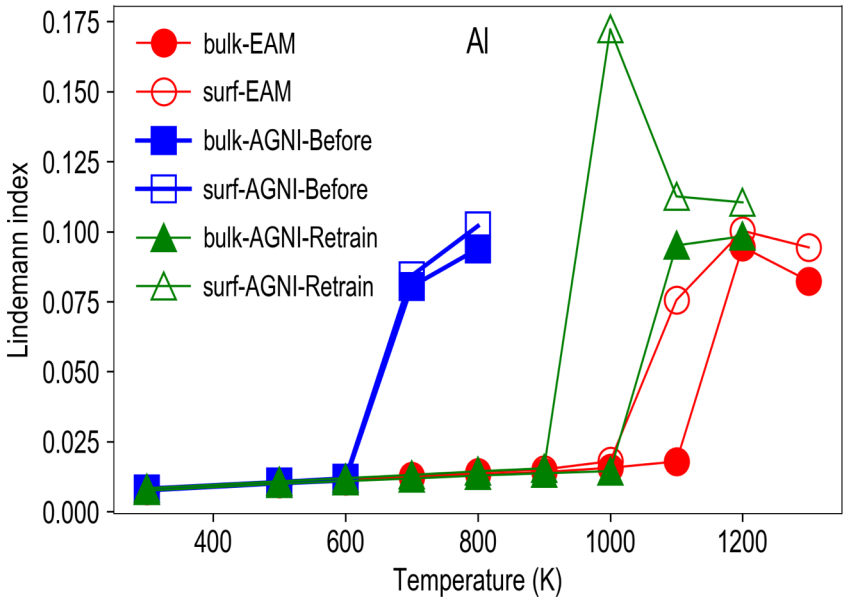
- Force integration to capture underlying PES
- Unable to accurately capture AA stacking

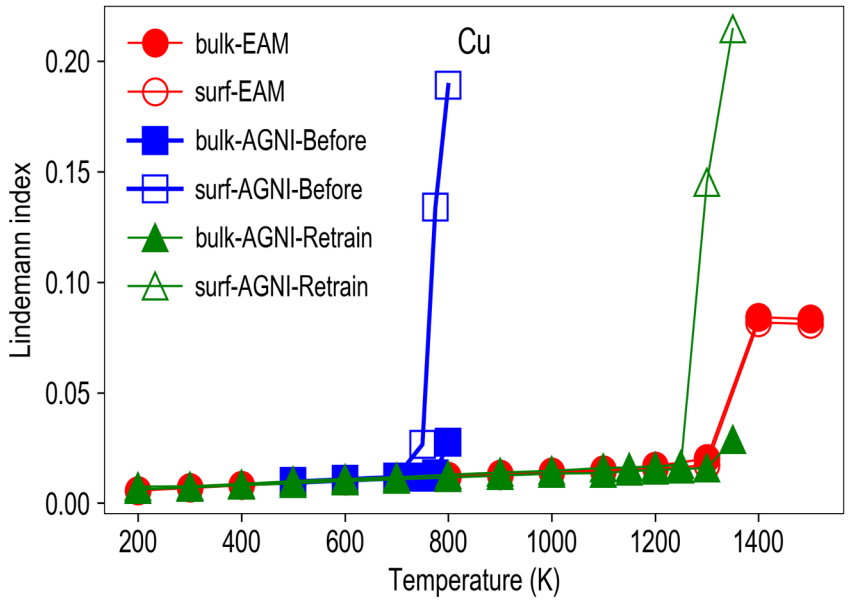
Relatively easy retraining process

- Statistical improvement with active learning



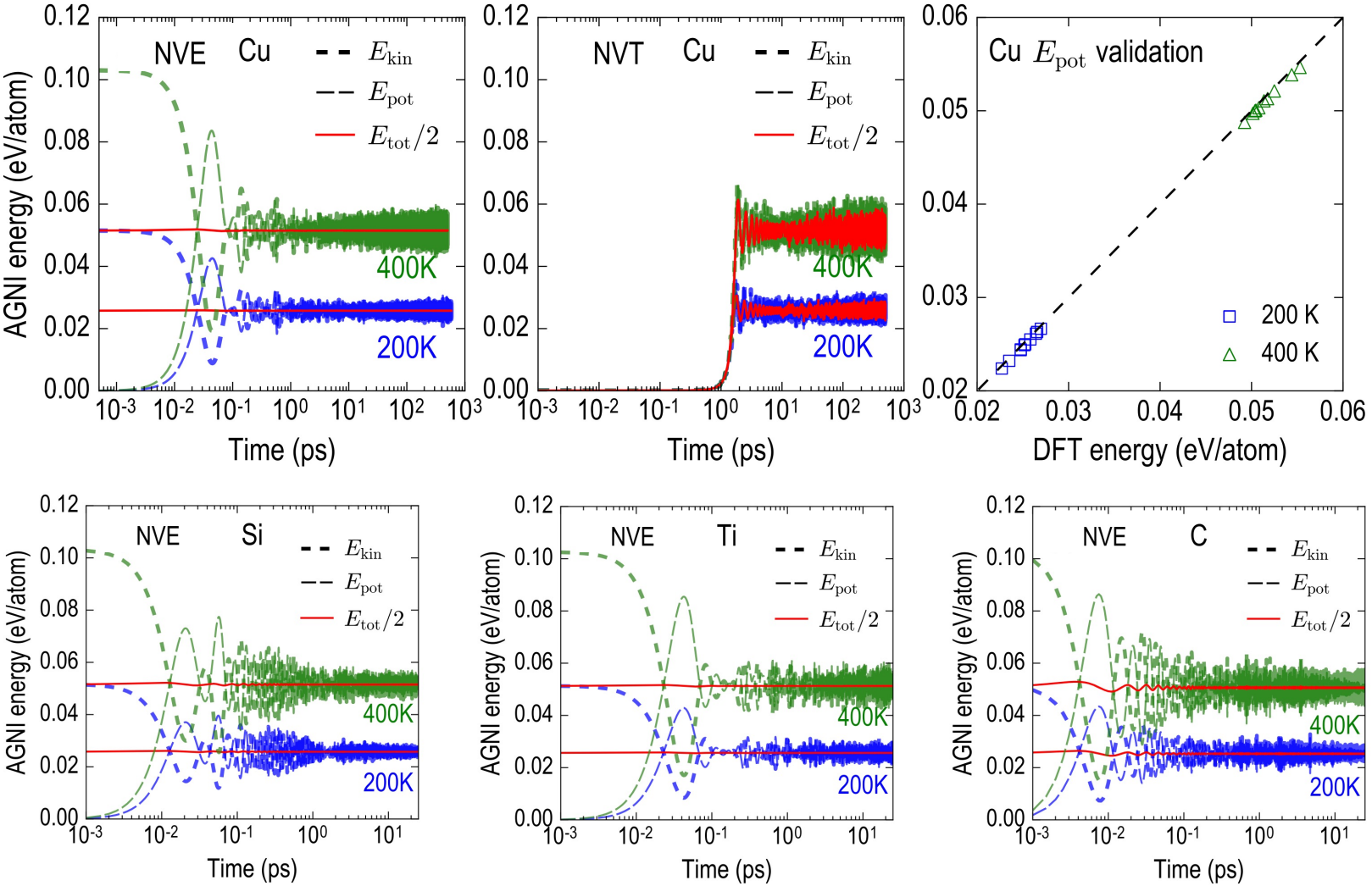
ACTIVE LEARNING: MELTING





Active learning improves AGNI FFs to capture melting behavior

AGNI: UNIVERSAL FRAMEWORK

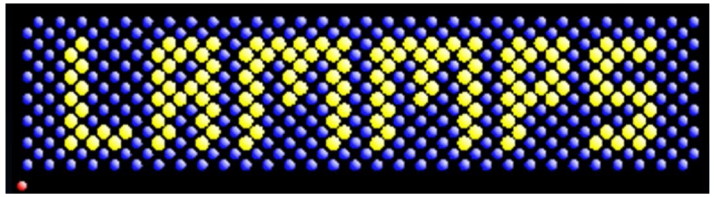


Huan et al., npj comput. mater., 3, 27 (2017)

CRITICAL STEPS FORWARD

- Use ML FF (along with active learning) in regions where classical potentials are know to have limitations (far from equilibrium)
- Extension of ML FF to multi-elemental systems (> 2 elements) and advanced ML methods (DNN)
- Uncertainty quantification and regions of applicability

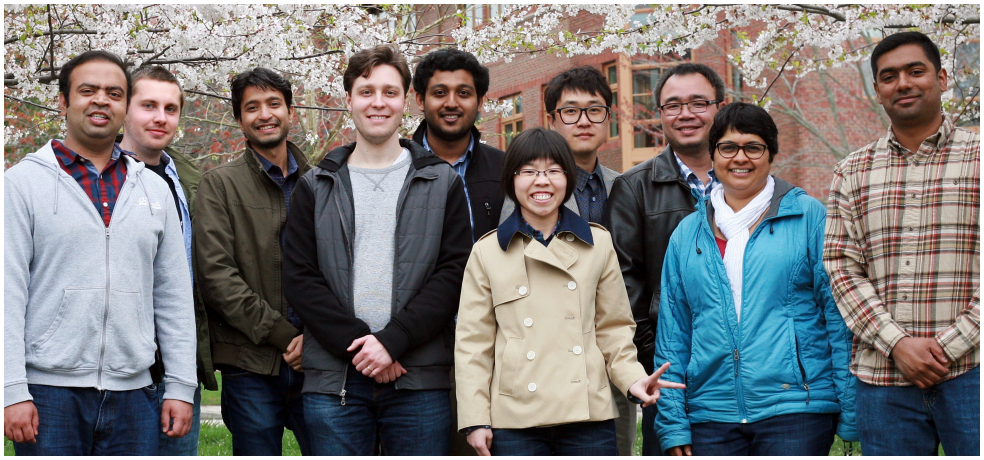
AGNI: Available in LAMMPS



`pair_style agni command`

Part of USER-MISC package

ACKNOWLEDGEMENTS



Past members
Dr. Sridevi Krishnan

Current members
Dr. Chiho Kim, Dr. Lihua Chen and
Dr. Anand Chandrasekharan

Special thanks to
Dr. Blas Uberuaga and Dr. Ghanshyam Pilania (Los Almos National Lab, NM, USA)

Funding agency

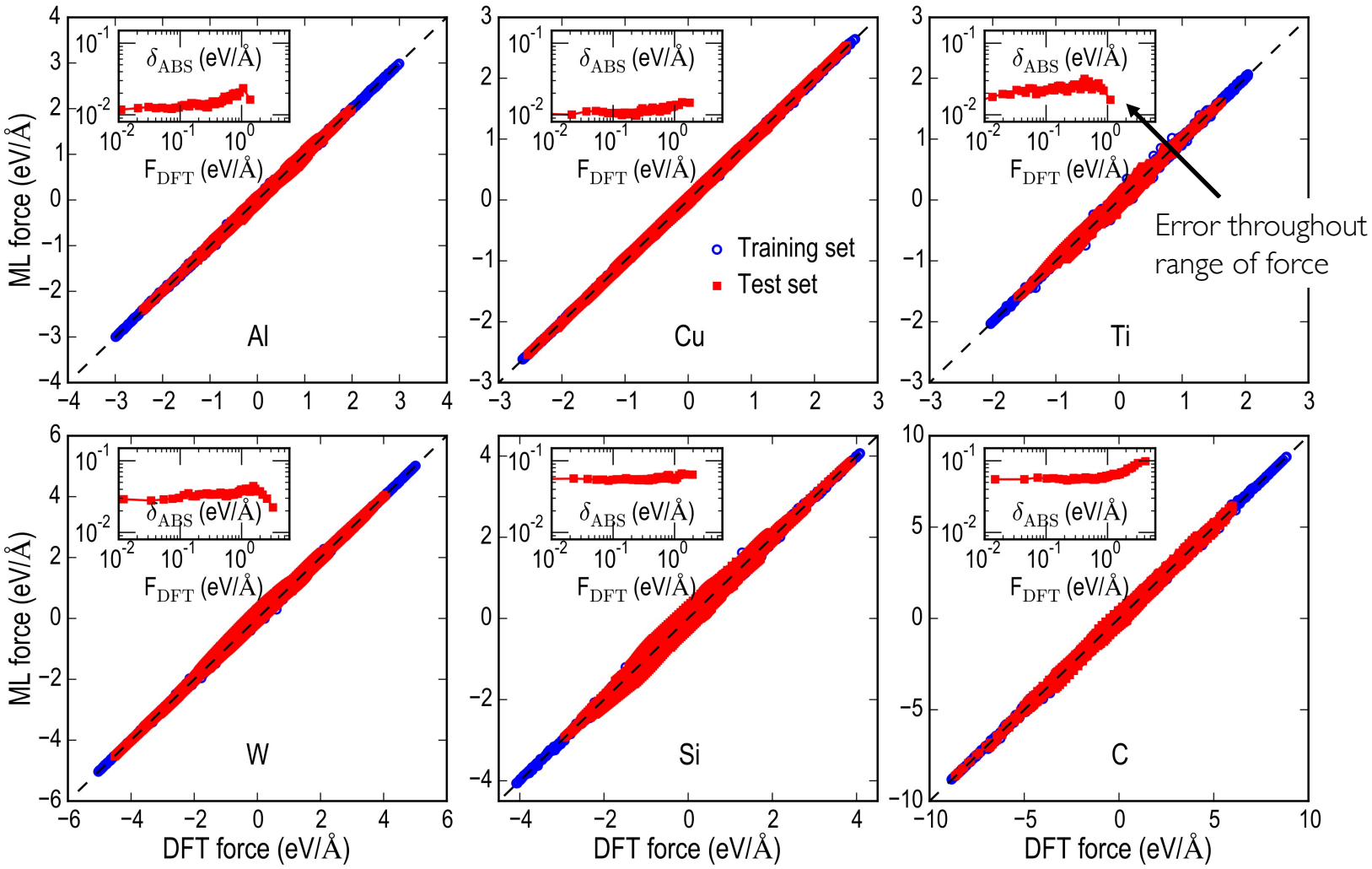
Office of Naval Research
ONR
USN
Science & Technology

NSF

XSEDE
Extreme Science and Engineering Discovery Environment

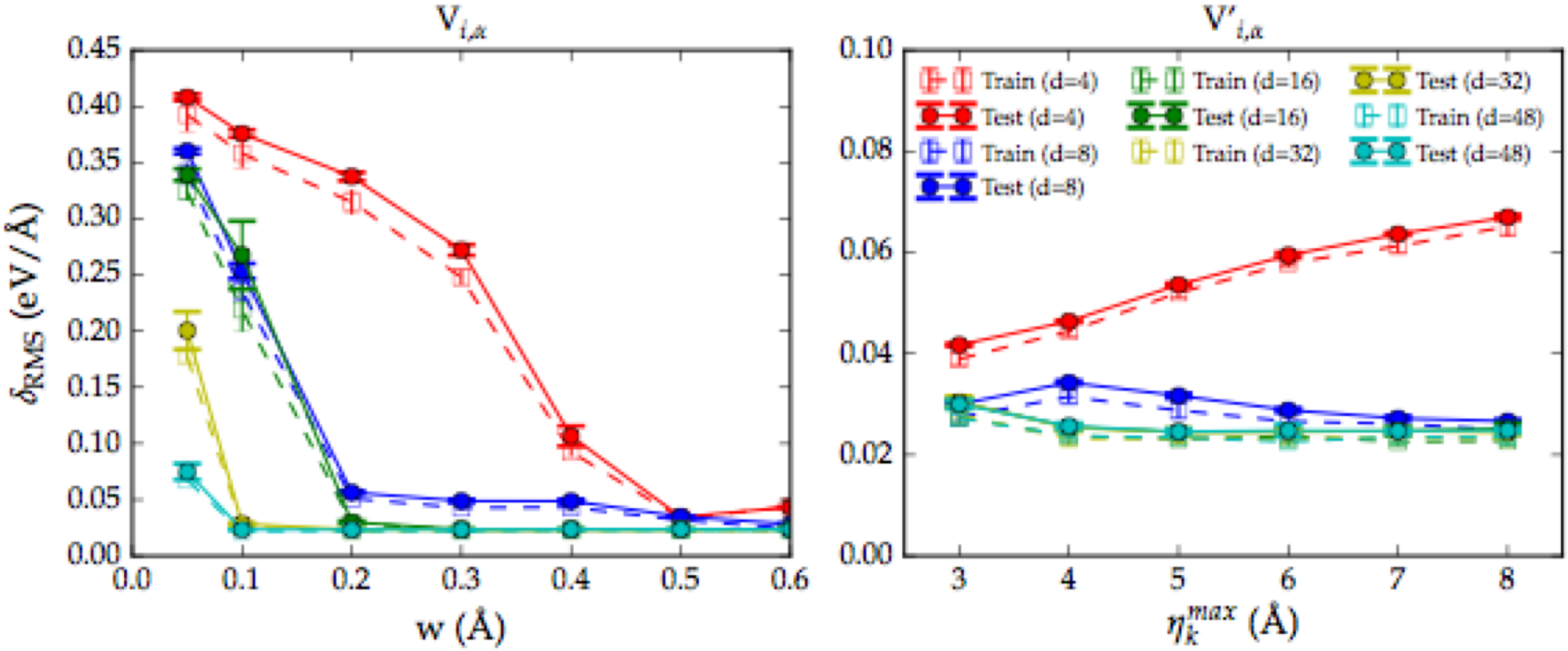
Backup slides

RESULTS: FORCE ACCURACY I

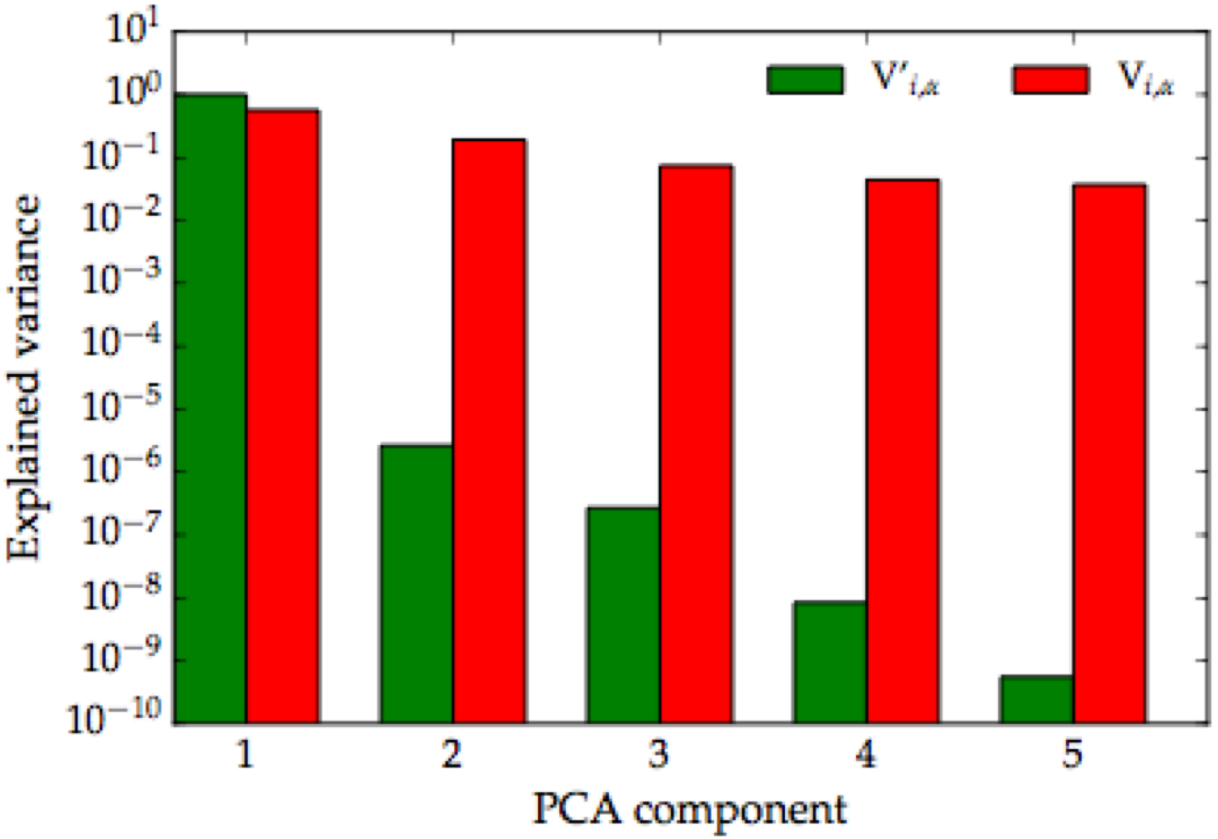


High force accuracy for all elements and for all range of forces

HYPERPARAMETERS DETERMINATION



FINGERPRINT COMPARISION



OVERALL ERRORS

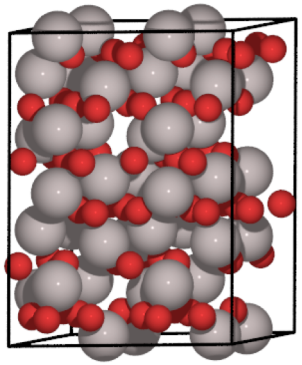
Table 1. Force evaluation errors of the force fields developed, obtained for a training set size $N_t = 1000$

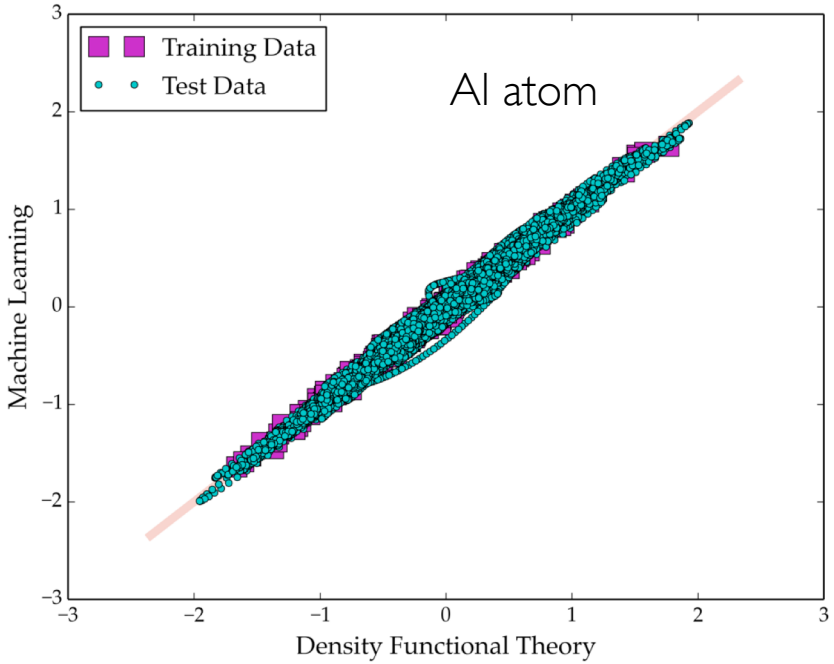
FF	Al			Cu		
	δ_{RMS}	δ_{MAX}	δ_{STD}	δ_{RMS}	δ_{MAX}	δ_{STD}
(I)	0.025	0.100	0.025	0.024	0.093	0.024
(II)	0.023	0.096	0.023	0.017	0.071	0.017
(III)	0.023	0.092	0.023	0.017	0.074	0.017
(IV)	0.025	0.097	0.025	0.018	0.076	0.017
(V)	0.021	0.082	0.021	0.016	0.056	0.016
FF	Ti			W		
	δ_{RMS}	δ_{MAX}	δ_{STD}	δ_{RMS}	δ_{MAX}	δ_{STD}
(I)	0.065	0.290	0.065	0.094	0.398	0.094
(II)	0.054	0.306	0.054	0.063	0.268	0.063
(III)	0.045	0.173	0.045	0.065	0.244	0.065
(IV)	0.047	0.162	0.047	0.068	0.253	0.068
(V)	0.035	0.149	0.034	0.049	0.200	0.049
FF	Si			C		
	δ_{RMS}	δ_{MAX}	δ_{STD}	δ_{RMS}	δ_{MAX}	δ_{STD}
(I)	0.081	0.296	0.081	0.161	0.778	0.161
(II)	0.074	0.251	0.074	0.088	0.373	0.088
(III)	0.074	0.260	0.074	0.083	0.322	0.083
(IV)	0.074	0.263	0.074	0.087	0.339	0.087
(V)	0.074	0.253	0.074	0.085	0.326	0.085

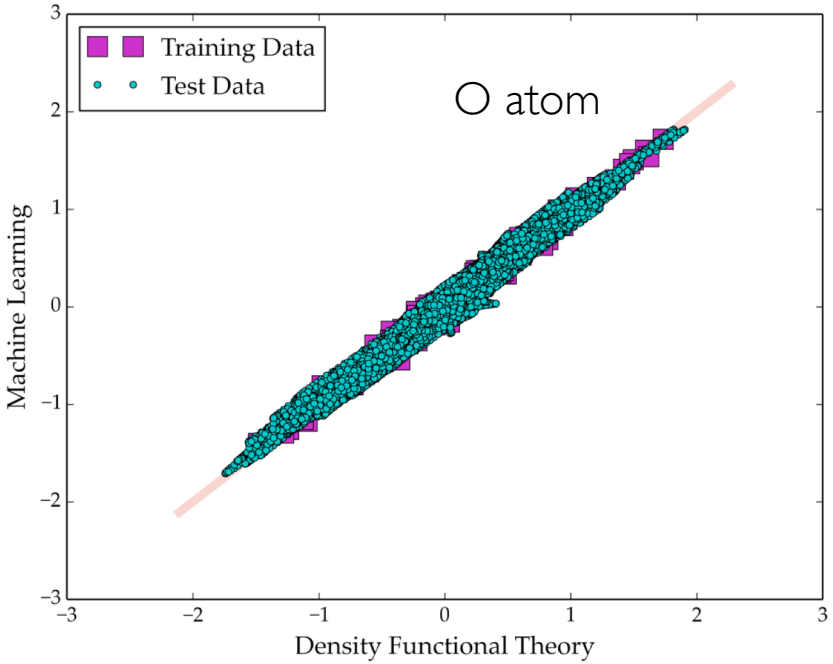
For each material, three error measures, i.e., δ_{RMS}, δ_{MAX}, and δ_{STD} are reported in eV/Å for (I), (II), (III), (IV), and (V), five recipes of force field creation (described in the text)

PERFORMANCE ON MULTI-ELEMENT CASES

Alumina



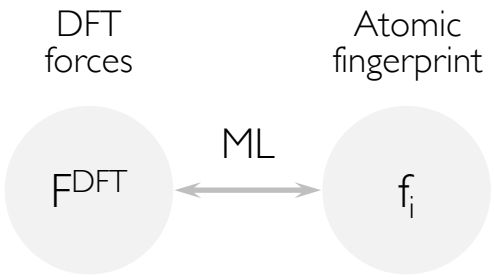




ML FORCE FIELDS: LEARNING FORCES

Learn atomic forces directly

Successful FF for Al and Si



Extracting energy from forces

1. Static energy calculations: $E = E_o - \sum F_i^u \Delta r_i^u, \; u \subset (x, y, z)$

2. Dynamic energy calculations: $E_t = E_{t-\Delta t} - \Delta t(\sum_{i,u} F_i^u v_i^u)$

Botu et al., J. Phys. Chem. C, 121 (1), 511 (2017), Li et al., PRL 114, 096405 (2015), Botu et al., Phys. Rev. B, 92, 094306 (2015)