Atomistic simulations with applications to Si and Ge systems

J. Hickman

NRC Postdoctoral fellow
MML
NIST: Gaithersburg MD

2018 Atomistic Simulations for Industrial Needs Workshop
August 01-03, Rockville, MD

National Institute of Standards and Technology
U.S. Department of Commerce

Presentation Overview

- Physically informed neural network (PINN) interatomic potentials (Mishin 2017)
  - Motivation and Introduction
  - Overview of potential model
  - Training set generation
  - Potential Development/training process
  - Si PINN results
  - Al PINN results
  - Future work (SiGe)

Introduction

Motivation

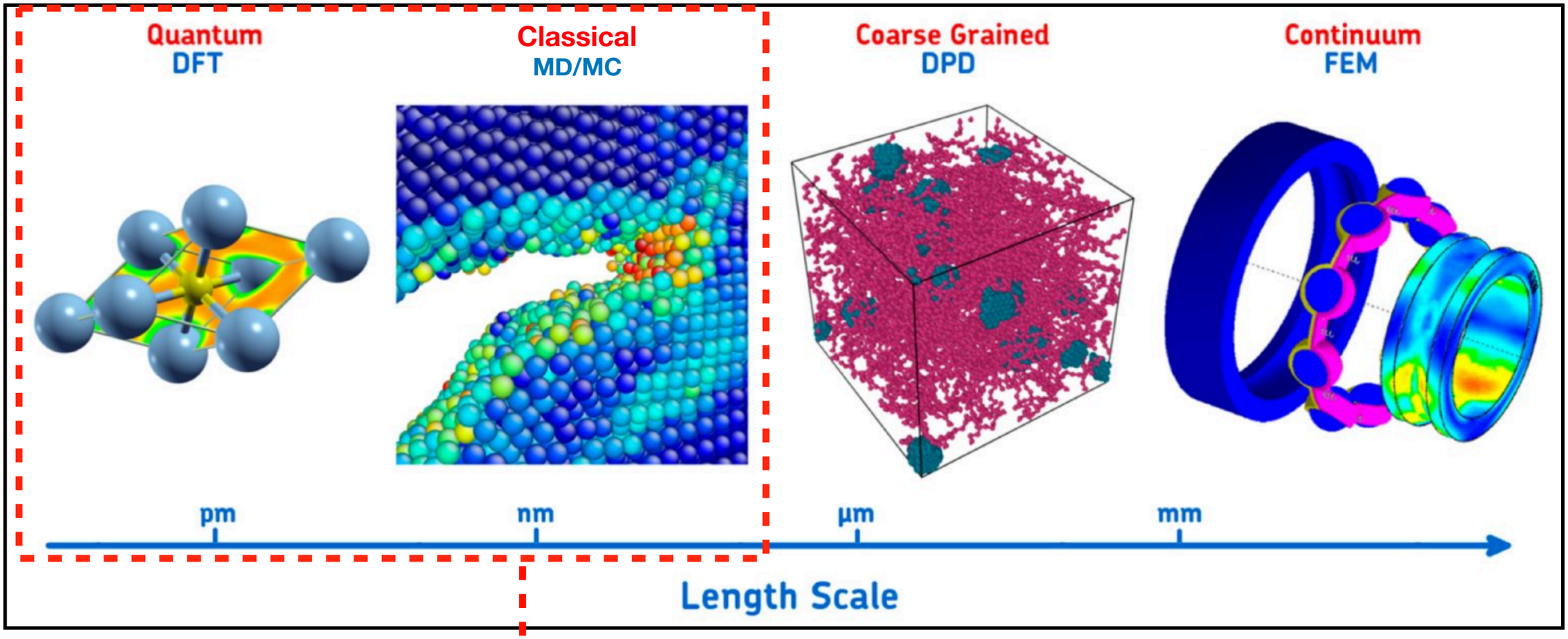


image source: http://evolution.skf.com/us/bearing-research-going-to-the-atomic-scale

Ab-initio/DFT

- Slow
- Size limited (N~ 10^2)
- 0K or short timescales
- Very accurate

Compromises

Speed
Accuracy
Scalability
Transferability

Classical (MD/MC)

- Fast
- Larger systems (N~ 10^7)
- Kinetic phenomena/long simulations
- ★ Accuracy depends on *approximation* of the potential energy surface (PES)

$$E = E(\mathbf{r}_1, \ldots, \mathbf{r}_N)$$

Better potential models

Types of interatomic potentials

Traditional interatomic potential

$$E = E(\mathbf{r}_1, \ldots, \mathbf{r}_N)$$

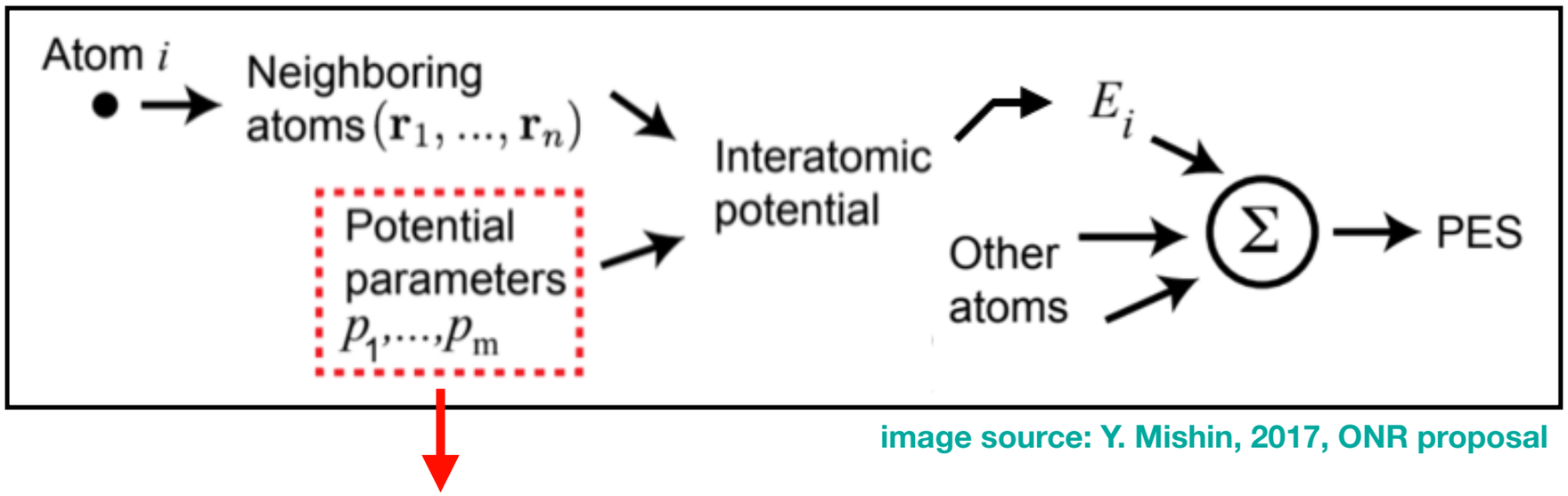


image source: Y. Mishin, 2017, ONR proposal

-EAM, ADP, REAX, COMB, REBO, REAXFF … etc

-PES approximated via physically derived analytic functions

~10-20 parameters

Fit or "train" the potential parameters using experimental and DFT data

Machine learning potentials →

Approximate PES via DFT energy or force interpolation

- **Gaussian process regression**
- **Interpolating moving least squares**
- **Kernel ridge regression**
- **Compressed sensing**
- ★ **Artificial neural network (ANN) potentials**

Types of machine learning applications

Machine Learning

only inputs

Unsupervised Learning

- Dimensionality Reduction
  - Meaningful Compression
  - Structure Discovery
  - Big data Visualistaion
  - Feature Elicitation
- Clustering
  - Recommender Systems
  - Targetted Marketing
  - Customer Segmentation

inputs and known outputs

Supervised Learning

- Classification
  - Image Classification
  - Customer Retention
  - Idenity Fraud Detection
  - Diagnostics
- Regression
  - Advertising Popularity Prediction
  - Weather Forecasting
  - Population Growth Prediction
  - Market Forecasting

inputs+guidlines

Reinforcement Learning

- Real-time decisions
- Game AI
- Robot Navigation
- Skill Acquisition
- Learning Tasks

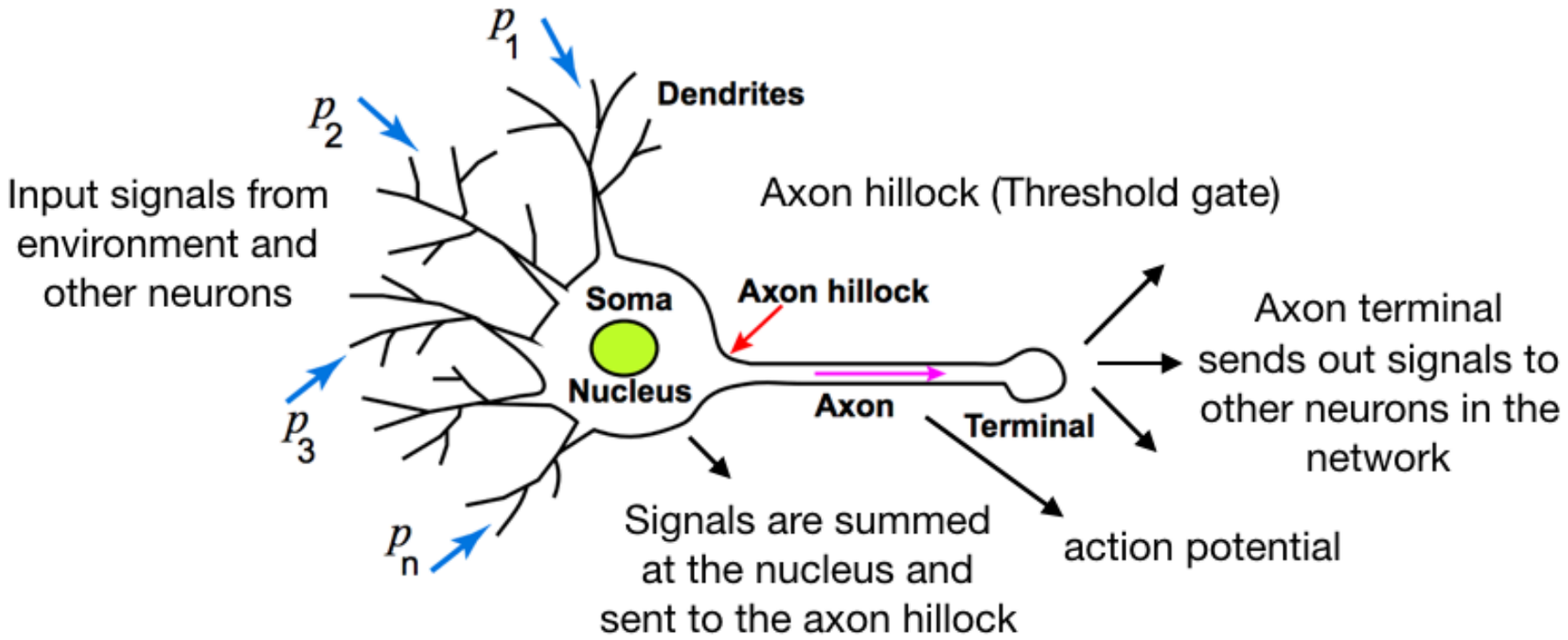
High dimensional function interpolation via artificial neural networks

Current work

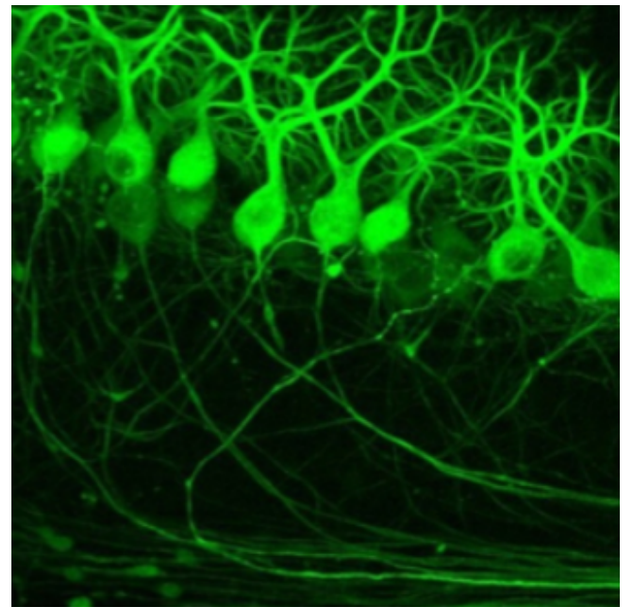
https://www.wordstream.com/blog/ws/2017/07/28/machine-learning-applications

Neural networks-1

Biological neuron

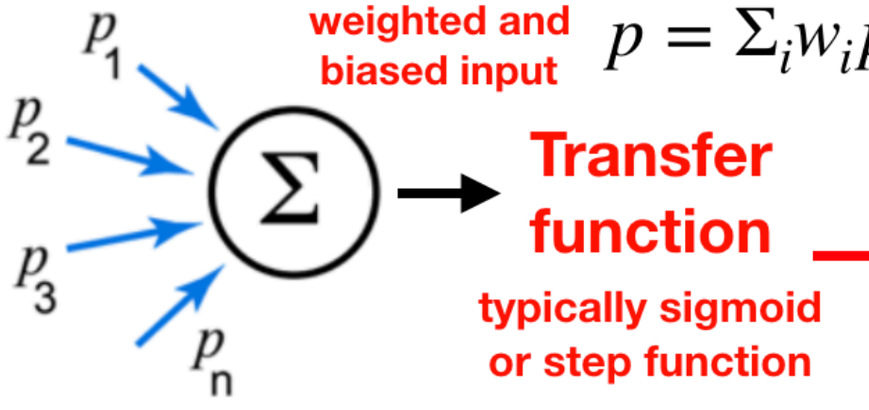


Biological neural network



Artificial neuron

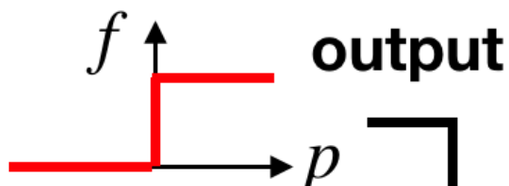
AKA: nodes or perceptrons



weighted and biased input $p = \Sigma_i w_i p_i + b$

Transfer function

typically sigmoid or step function



output

$$q = \begin{cases} 1, & \Sigma_i w_i p_i + b > 0 \\ 0, & \Sigma_i w_i p_i + b < 0 \end{cases}$$

Each note has **one** bias and as many weights as it has inputs

Artificial neural network

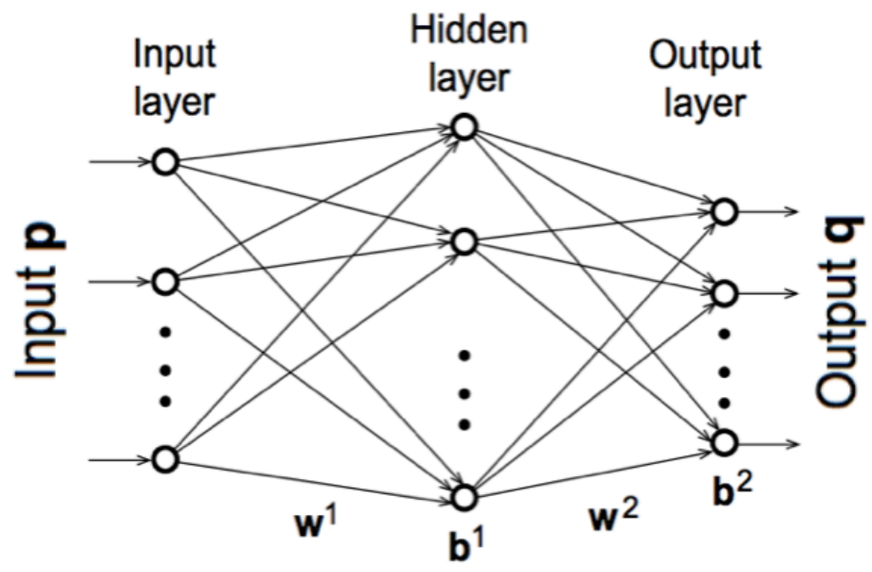
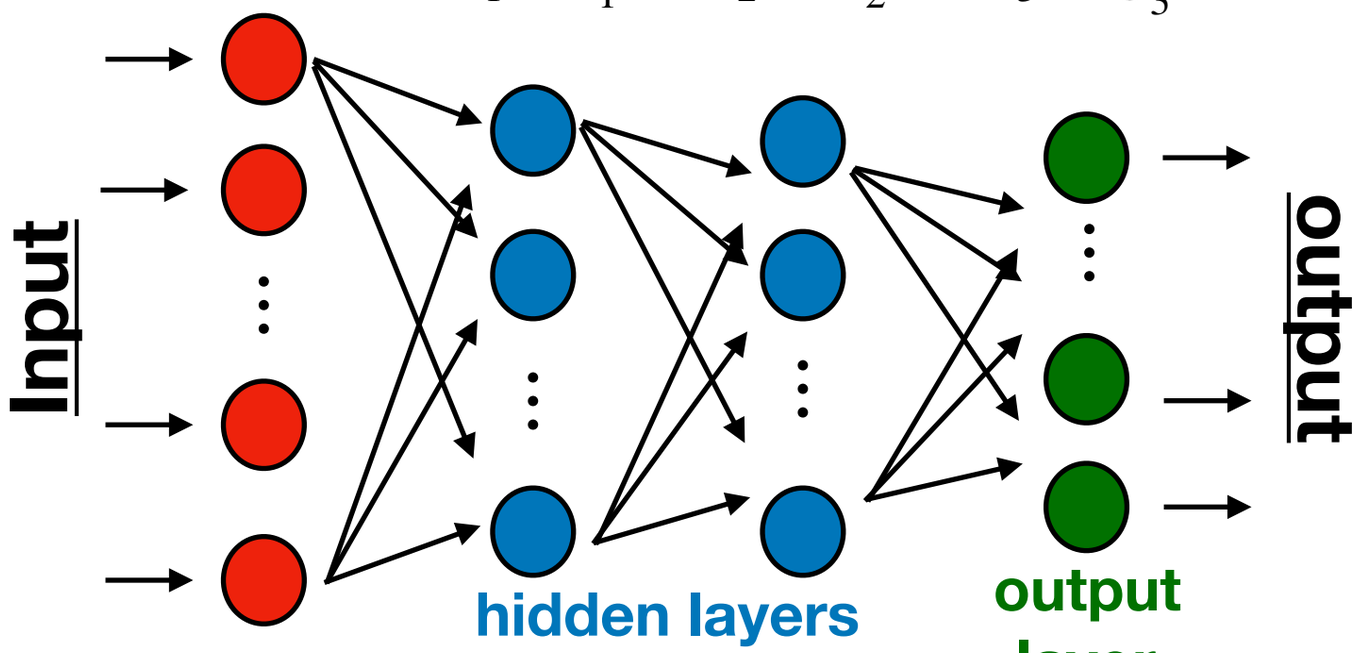


image source: Y. Mishin, 2017, ONR proposal

Neural networks-2

Architecture:

fitting parameters: $\underline{w}_1$ $\overline{b}_1$ $\underline{w}_2$ $\overline{b}_2$ $\underline{w}_3$ $\overline{b}_3$

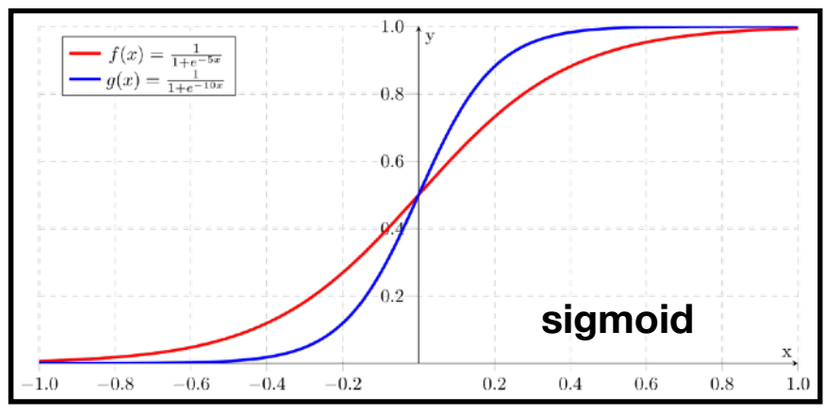


60 16 16 8

Architecture: 16x16x8

Network description/architecture
layers+transfer functions+weights and biases

Transfer function: $$f(p) = \frac{1}{1 + e^{-p}}$$



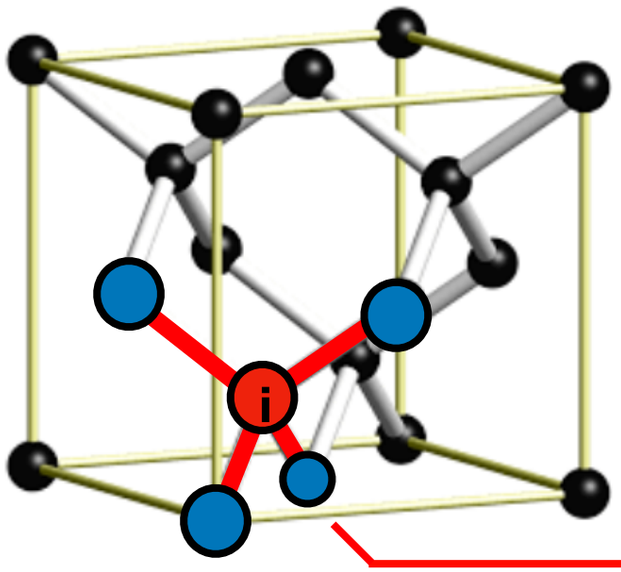
$$N_{fit} = (60 \times 16 + 16) + (16 \times 16 + 16) + (16 \times 8 + 8) = \boxed{1384}$$ **Number of NN fitting parameters**

Training

known inputs $\mathbb{R}^N$ → NN → known outputs $\mathbb{R}^M$ → **Once trained the NN can map between arbitrary inputs and output**

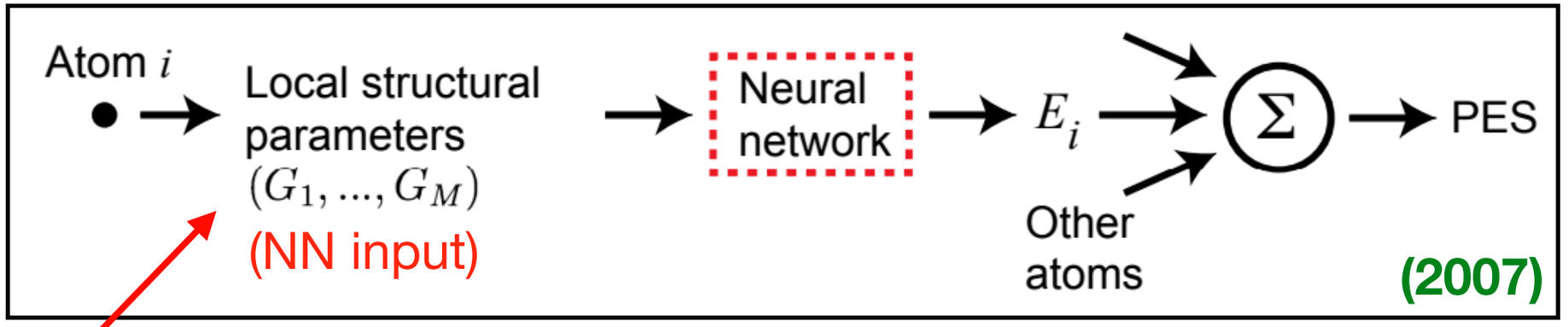
fit/optimize weights and biases

Artificial neural network potentials



Quantify local structural "fingerprint"

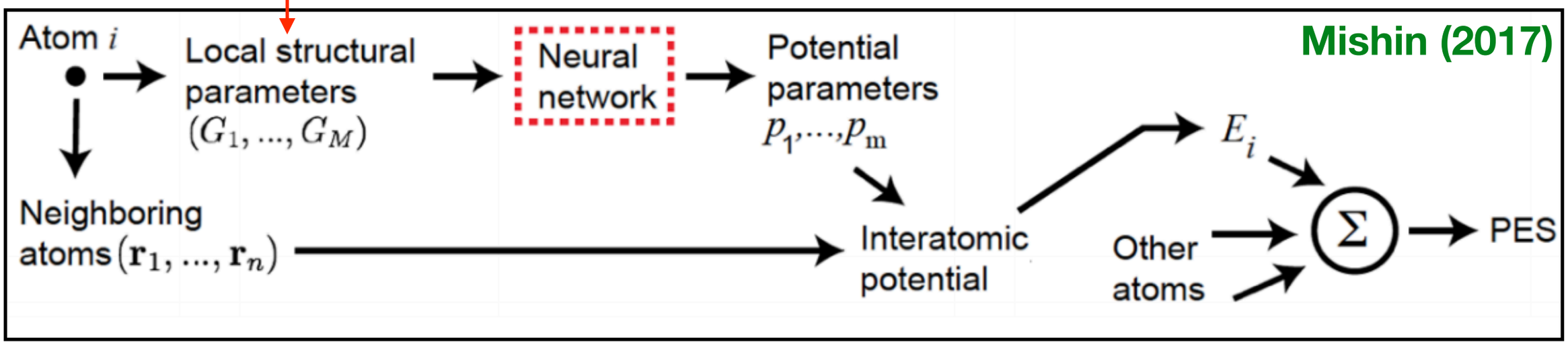
"Mathematical" or "straight NN potential model:



[33] J. Behler, M. Parrinello, Generalized neural-network representation of high-dimensional potential-energy surfaces, Phys. Rev. Lett. 98 (2007) 146401.

[34] A. Bholoa, S. D. Kenny, R. Smith, A new approach to potential fitting using neural networks, Nucl. Instrum. Methods Phys. Res. 255 (2007) 1–7.

Physically informed neural network (PINN) potential model:



Training process:

"Training set" — Structures | DFT energies

Model error → Model energies

$$RMSE = \left(\frac{\Sigma_i (E_i - \tilde{E}_i)^2}{N} \right)^{\frac{1}{2}}$$ —— Fitting

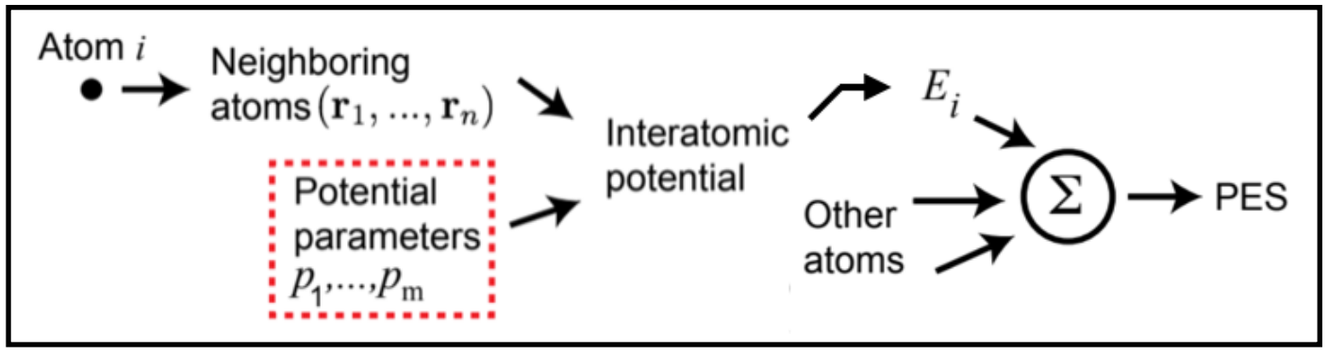
config-1	E_1	$\tilde{E}_1$
config-2	E_2	$\tilde{E}_2$
⋮	⋮	⋮
config-N	E_N	$\tilde{E}_N$

Find the NN weights and bias's which minimize the RMSE

image source: Y. Mishin, 2017, ONR proposal

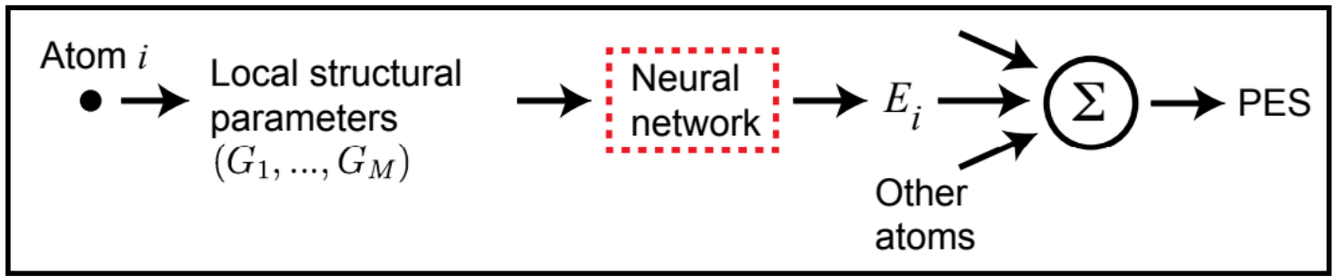
Potential model comparison

Traditional interatomic potential:

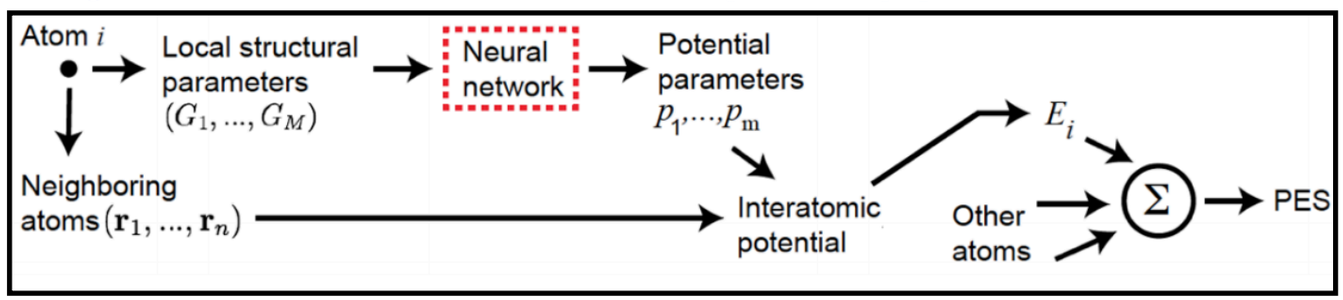


- extrapolation=predictions outside training region

"Mathematical" or "straight NN potential:



Physical neural network (PINN) potential:

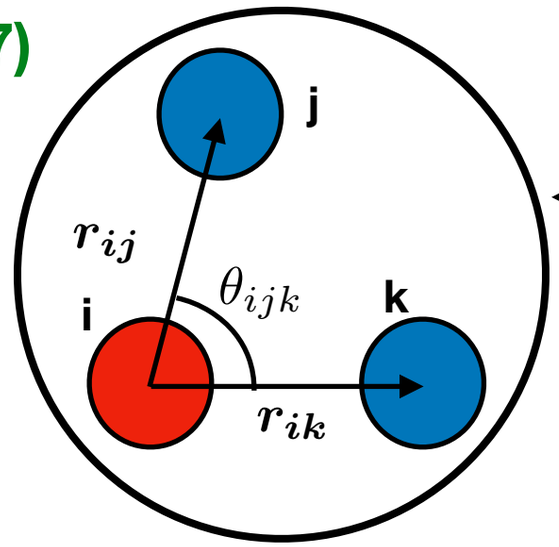


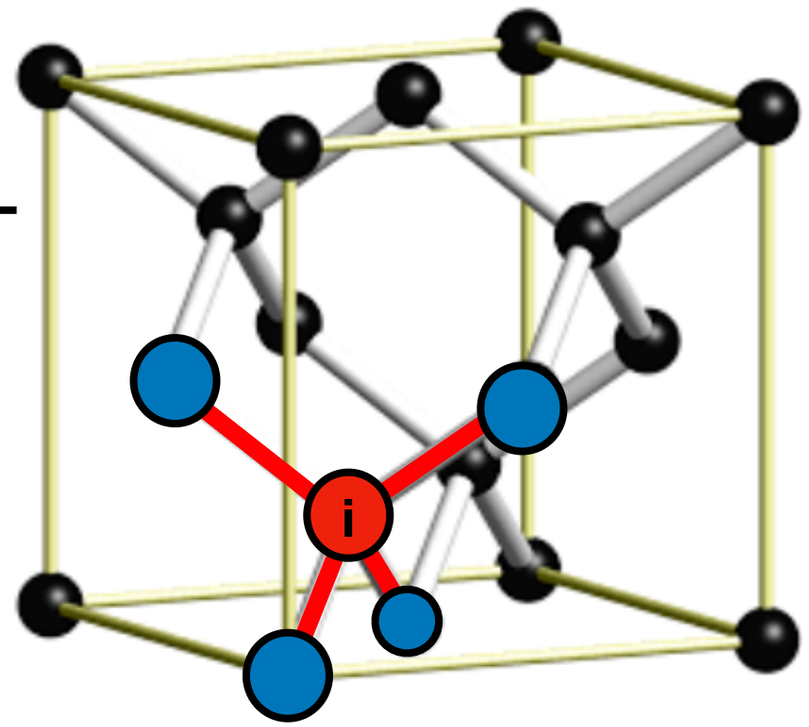
Pros	Cons
• Very fast • Decent extrapolation • Potential models derived from physics	• Difficult to train/fit • requires intuition • Hard to improve upon once finalized • Accuracy limitations
• Fast relative to DFT • DFT level accuracy (~1-5 mEv) within training set • Relatively straight forward/routine to train/fit • Systematic improvement (add more data)	• Slow relative to traditional potentials • Bad extrapolation
• Same as straight NN • Decent extrapolation • Potential models derived from physics	• Slow relative to traditional potentials

PINN potential: Structure parameters

Structure parameters: (Mishin: 2017)

$$G_i^l = \sum_{j,k} P_l(cos(\theta_{ijk}))f(r_{ij})f(r_{ik})$$





Angular term:

$$P_l(cos(\theta)) \qquad l = 0,1,2,4,6$$

(Legendre polynomials)

Radial term:

$$f(r) = \frac{1}{\sigma^3} e^{-\frac{(r-r_o)^2}{\sigma^2}} f_c(r)$$

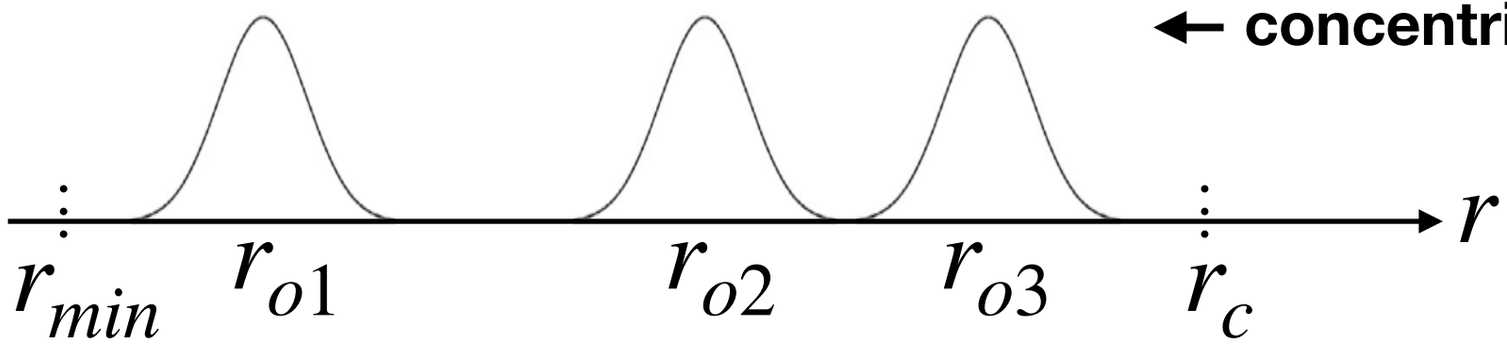
$f_c(r)$ = Cutoff function

choose multiple r_o values

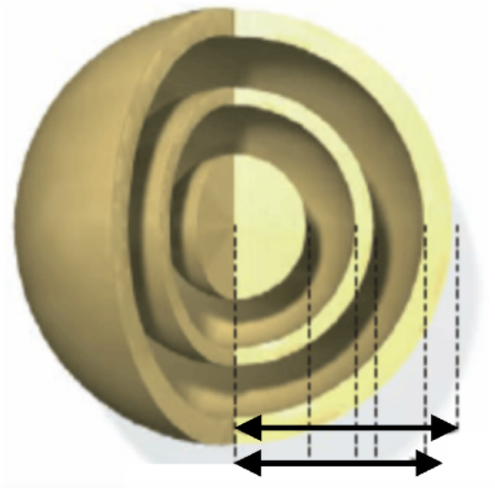
(~8 to 12)

Gi's act as "fingerprints" of the local atomic environment

$12\ r_o's \rightarrow 60\ G_i's$ per atom



← concentric shells →



PINN potential: BOP format (Mishin: 2017)

$$f_c(r) = \begin{cases} \dfrac{(r - r_c)^4}{d^4 + (r - r_c)^4} & r \le r_c \\ 0, & r \ge r_c \end{cases}$$

cutoff function

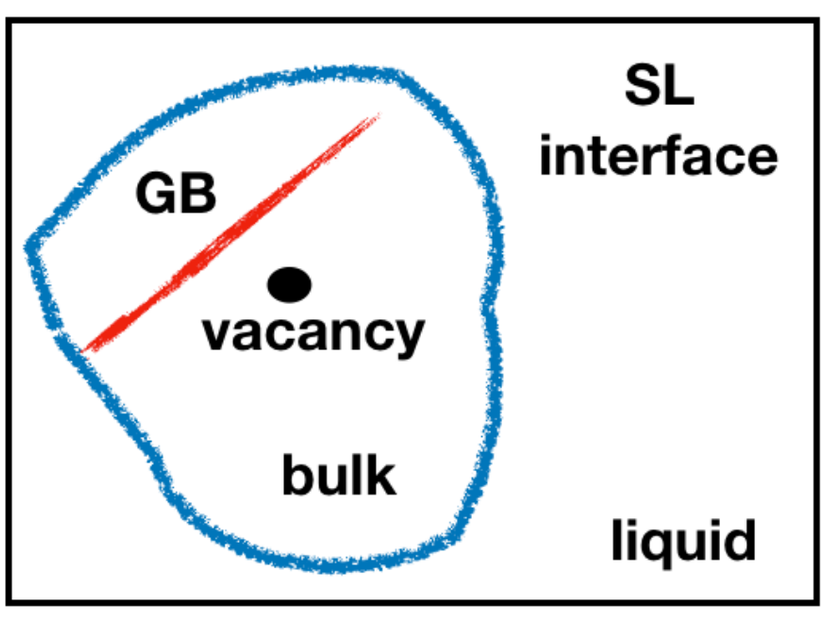
Note: r_c *and* d *are fixed*

$$E_i = \frac{1}{2} \sum_{j \neq i} \left[A_i e^{-\alpha_i r_{ij}} - B_i S_{ij} b_{ij} e^{-\beta_i r_{ij}} \right] f_c(r_{ij}) + E_i^{(p)}$$

(atom specific)

bond order parameter

promotion energy



$$b_{ij} = (1 + z_{ij})^{-1/2}$$

$$E_i^{(p)} = -\sigma_i \left(\sum_{j \neq i} S_{ij} b_{ij} f_c(r_{ij}) \right)^{1/2}$$

$$z_{ij} = \sum_{k \neq i,j} a_i S_{ik} (\cos\theta_{ijk} + h_i)^2 f_c(r_{ij}) \longleftrightarrow S_{ij} = \prod_{k \neq i,j} S_{ijk}$$

$$S_{ijk} = 1 - f_c(r_{ik} + r_{jk} - r_{ij}) e^{-\lambda_i (r_{ik} + r_{jk} - r_{ij})}$$

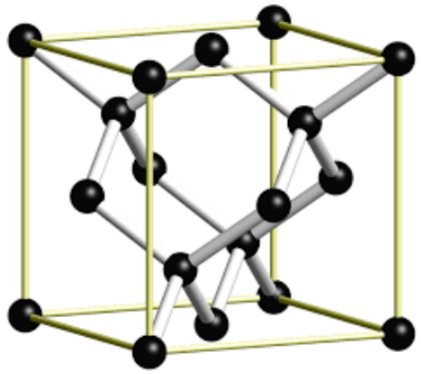
8 Adjustable parameter which are controlled by the outputs of the NN

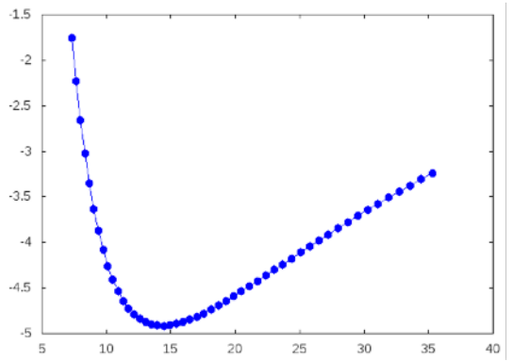
A_i, B_i, α_i, β_i, a_i, h_i, λ_i and σ_i

Silicon PINN potential
(preliminary results)

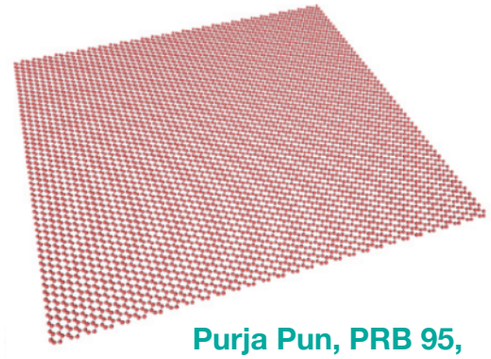
Training/test set generation

Stable structure: (Diamond)

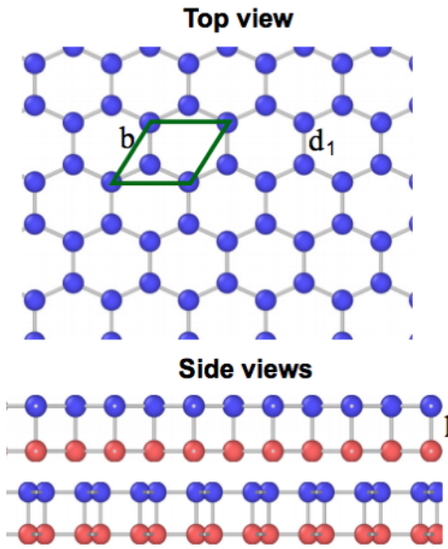




Two dimensional structures:

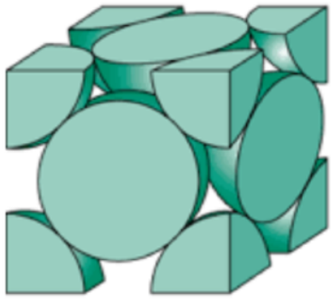


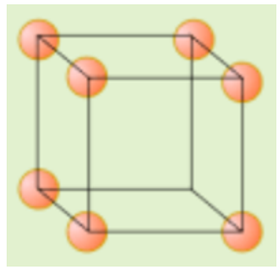
Purja Pun, PRB 95, 224103 (2017)

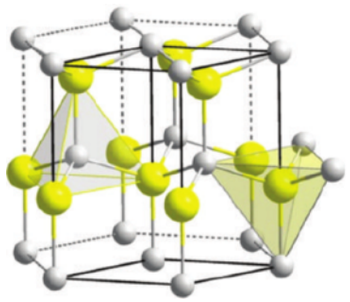


6 silicene allotropes

Alternative structures:







~14 alternate structures

- FCC, BCC, HEX, HCP, SC, Liquid, Amorphous …etc

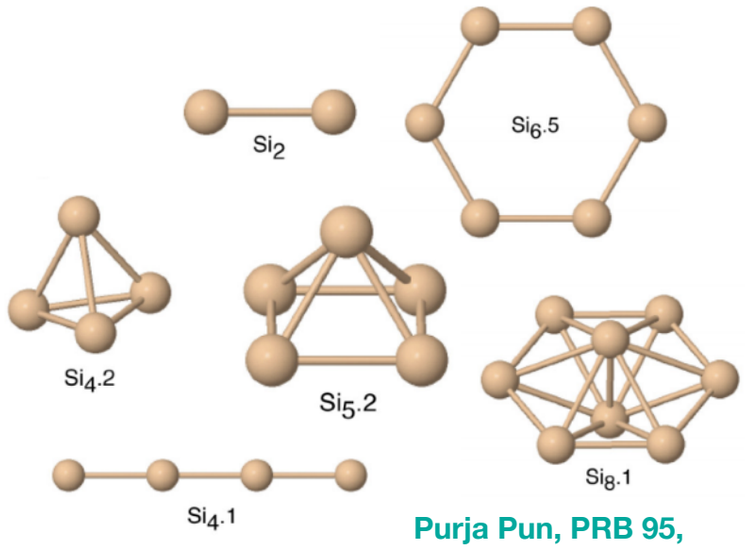
DFT calculation details

- Functional/PP: PBE/PAW
- ENCUT=600
- ~4300 structures
- block size 1-96 atoms
- k-point convergence tests for each group

Non-equilibrium sampling

- Isotropic expansions/compressions
- Random local atomic perturbations
- Anisotropic box variations

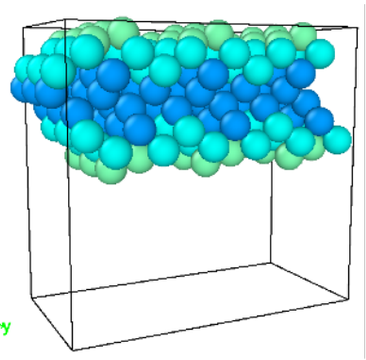
Clusters:

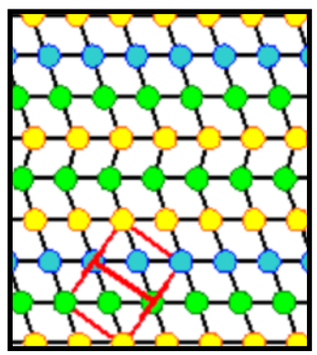


Purja Pun, PRB 95, 224103 (2017)

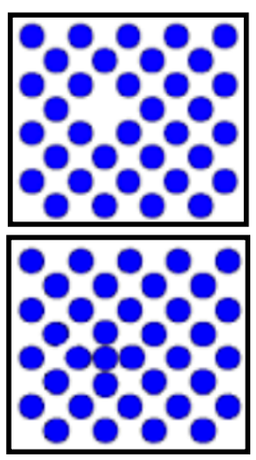
~18 atomic clusters

Defects:





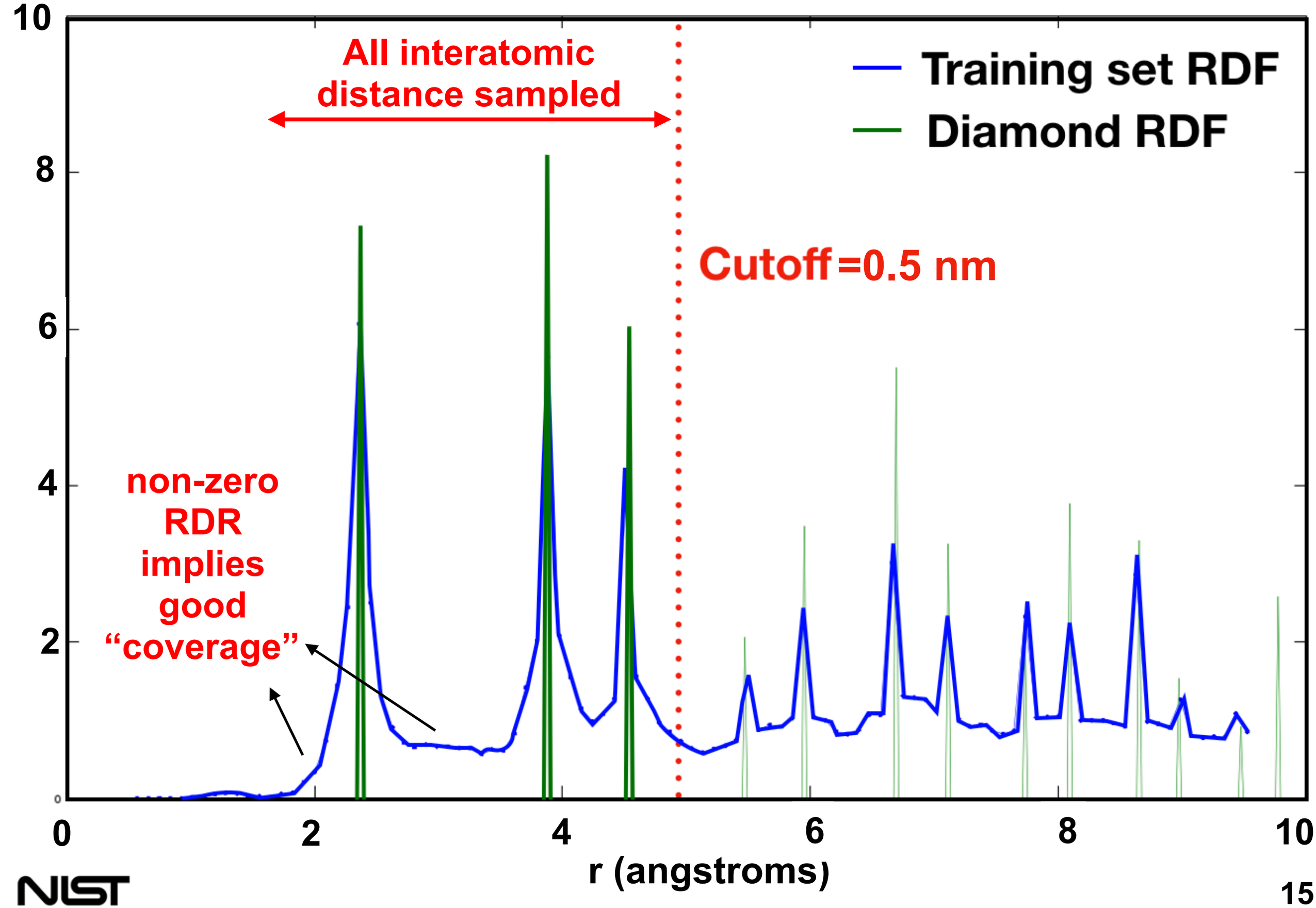
www.tf.uni-kiel.de



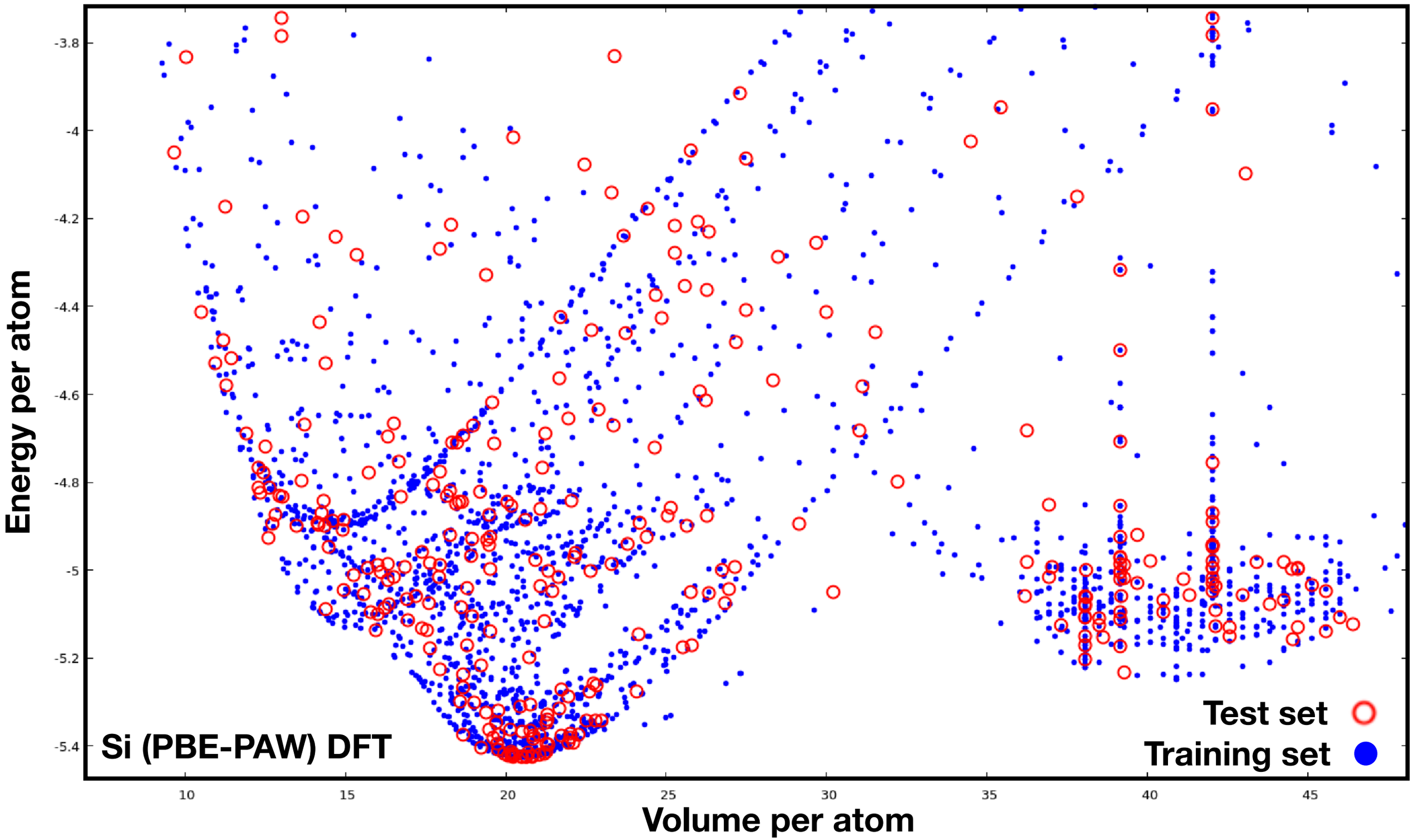
~27 different defects

- Vacancies, Various self interstitials, Surfaces, Stacking faults

Train set coverage



Training and test set



Full set=Test+Training=110,000 atoms (4279 structures)

(random subsets) ↓ **15,000 atoms (575 structures)**

↳ **95,000 atoms (3706 structures)**

NIST

Si PINN potential details

Methodological details:

PINN:

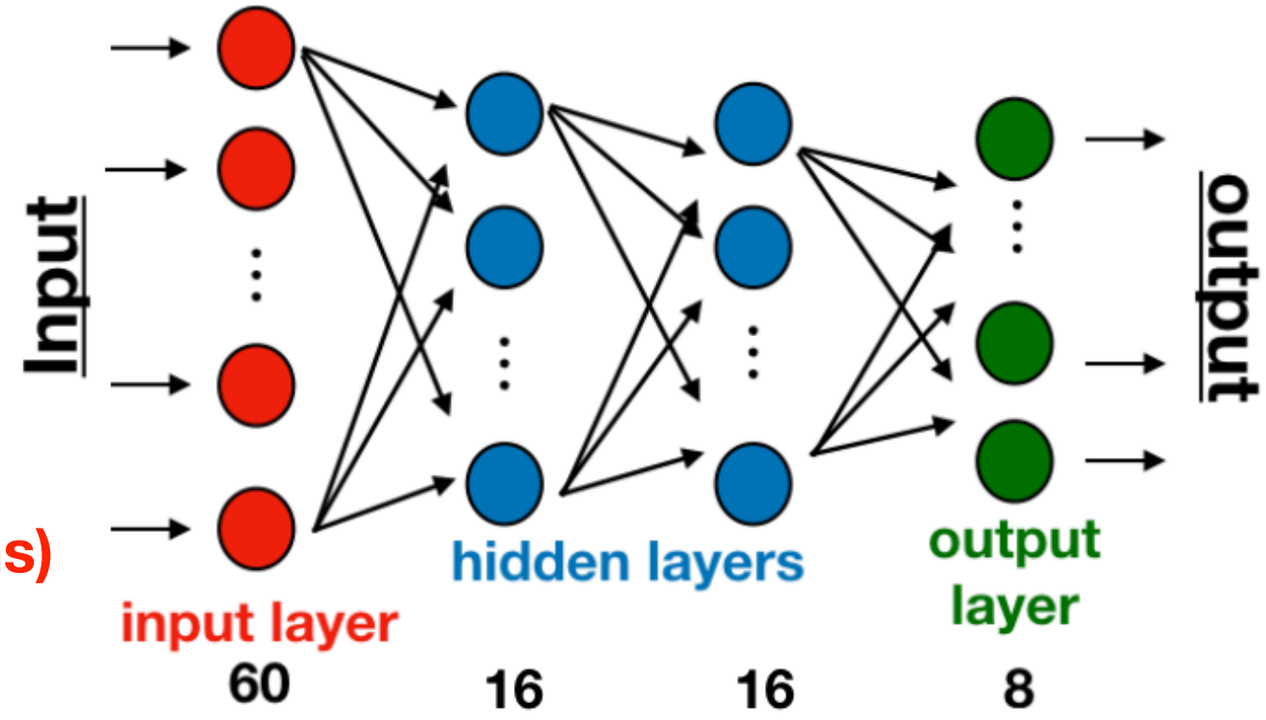
$16x16x8$

12 uniformly spaced ro choices:

ro=1.25 1.59 1.92 2.26 2.60 2.93
3.27 3.60 3.94 4.28 4.61 4.95

→ **60 Gi's total**

$\sigma = 0.1\ nm$



Training set: (0.5 full set) (exploratory runs)

Number of structures: 1640
Total number of atoms: 42832

Test set:

Number of structures: 501
Total number of atoms: 12812

NN fitting parameters: 1384

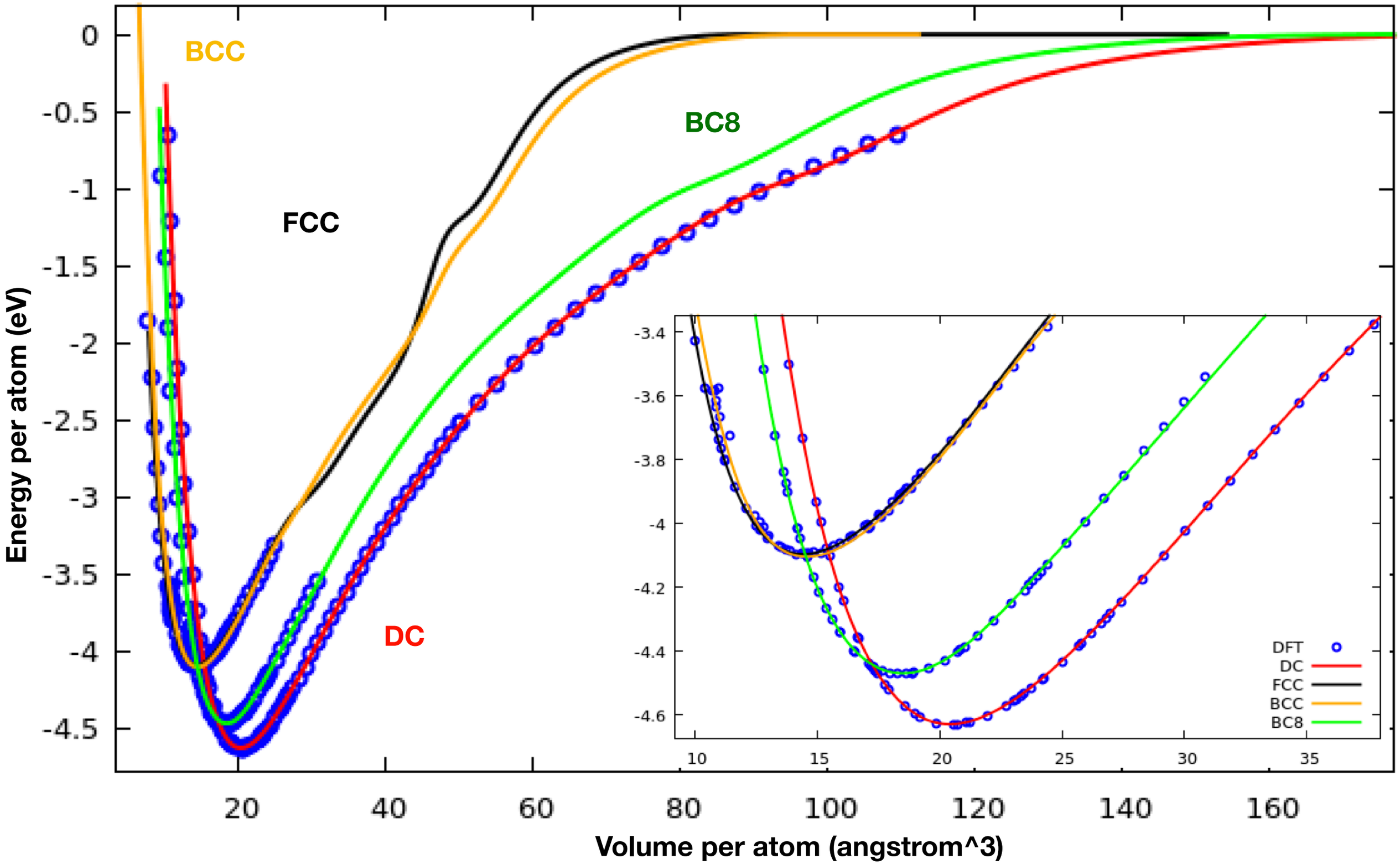
Cutoff function parameters (global) $r_c = 0.5\ nm$, $d = 0.15\ nm$

- **DFT energy shifted by 0.79502 eV/atom → DC=4.63 (eV/atom)**

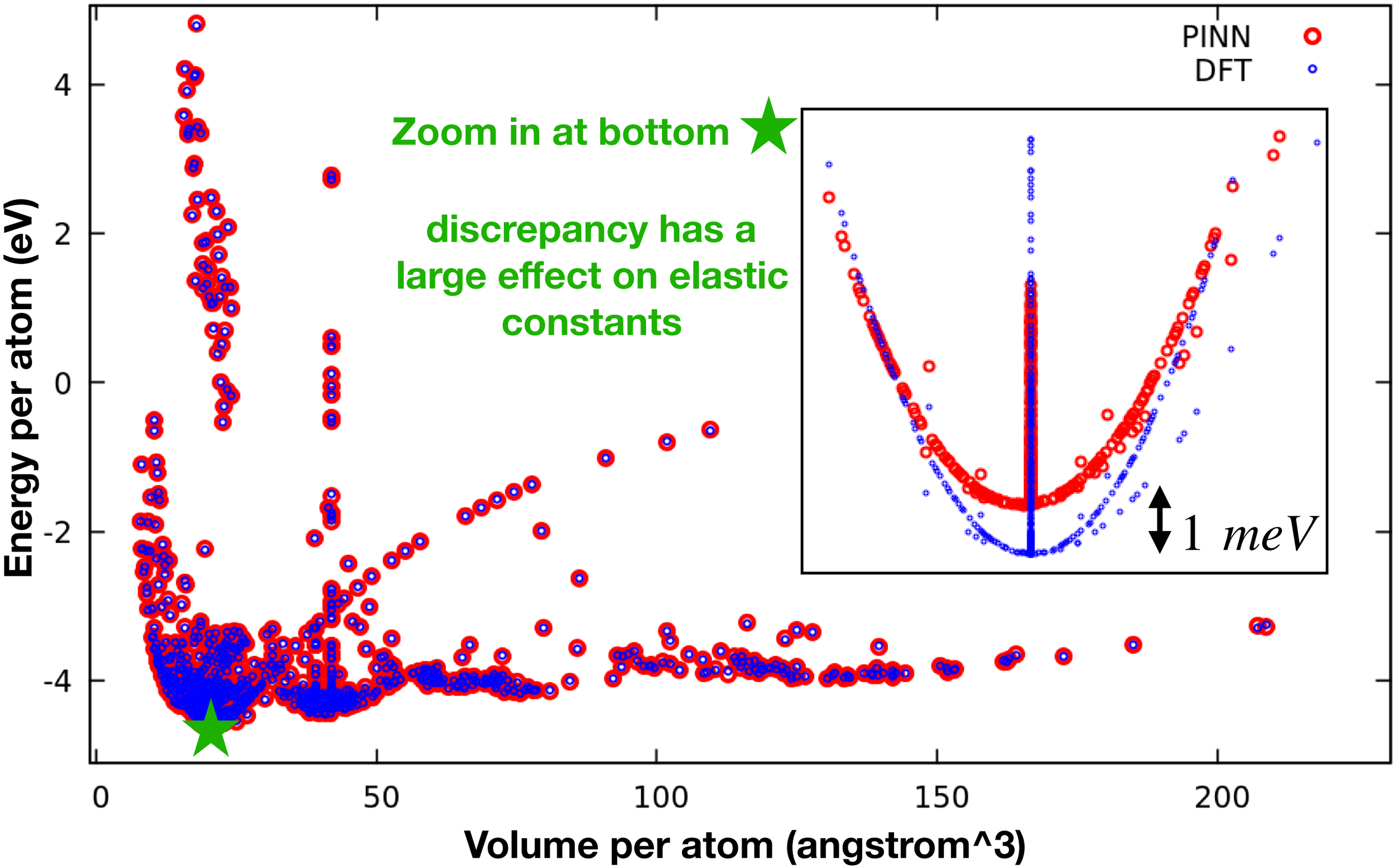
Start many NN's from different randomized IC, optimize and choose best

Fitting code reference: ("To be published") G. P. Purja Pun(1), R. Batra(2), R. Ramprasad (3) and Y. Mishin(1)
(1) George Mason Univ., (2) Univ. Connecticut, (3) Georgia Tech

Select Si PINN potential equations of state

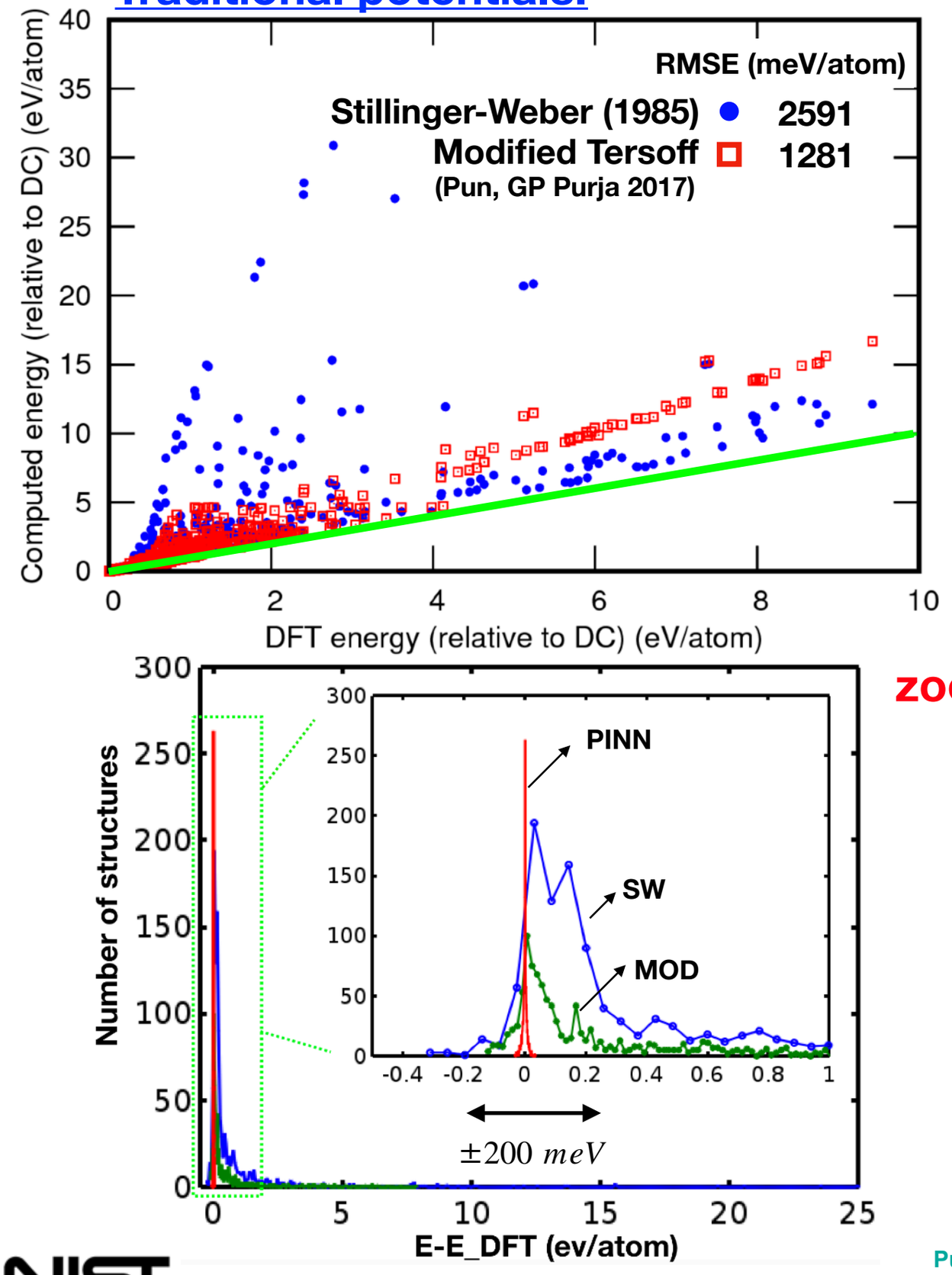


Results: Silicon PINN Potential

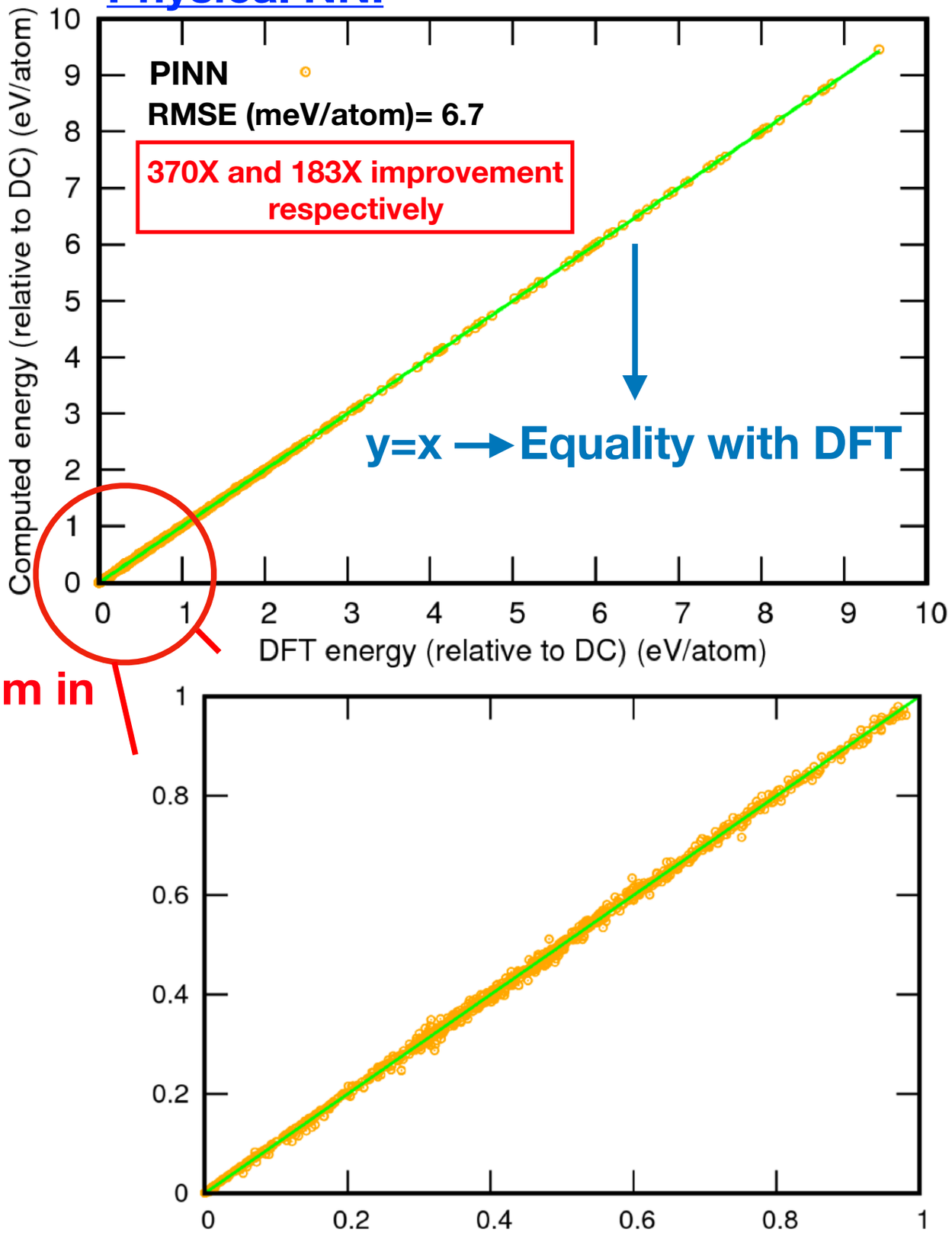


Si PINN comparison with traditional potential

Traditional potentials:

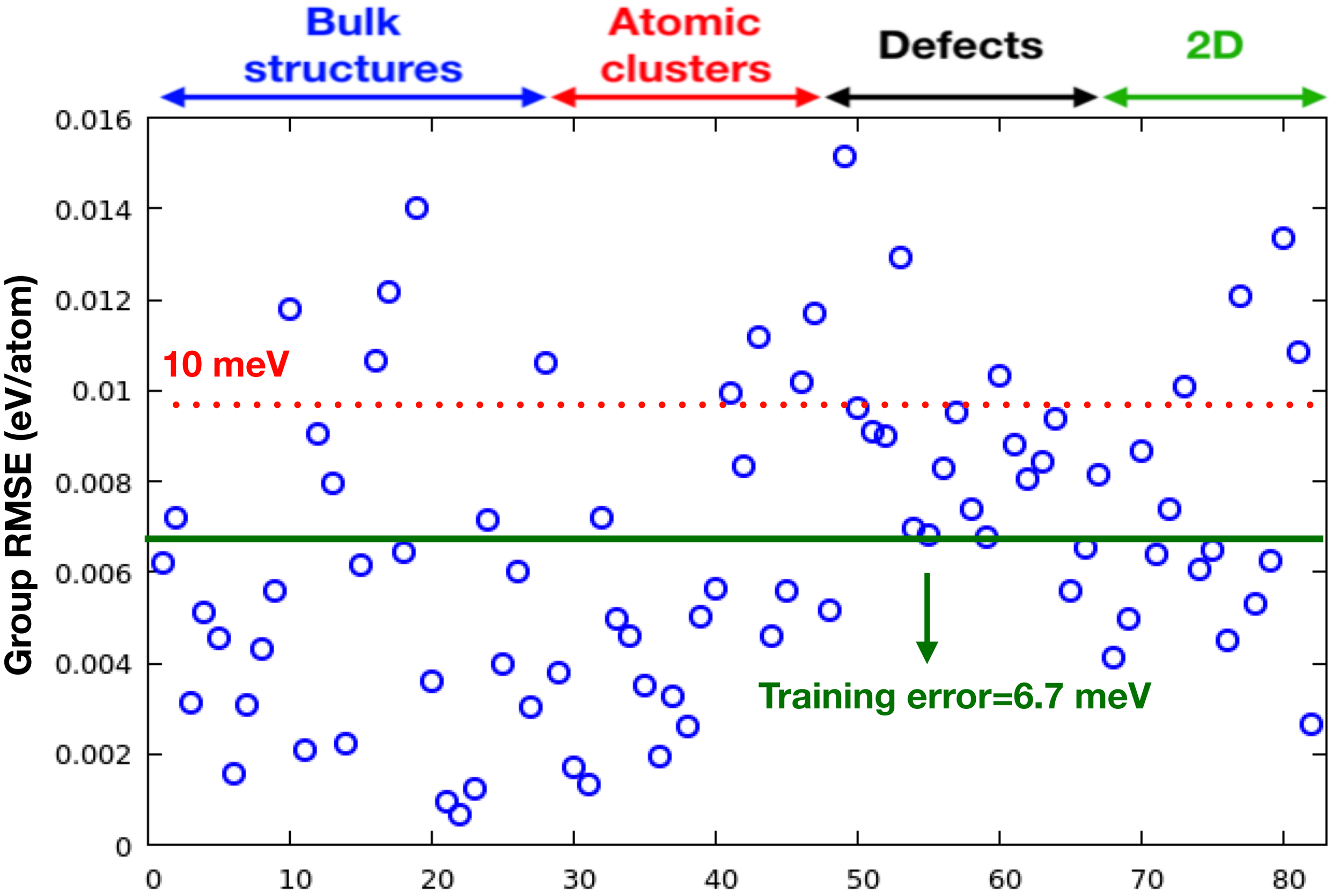


Physical NN:



Pun, GP Purja, and Y. Mishin. "Optimized interatomic potential for silicon and its application to thermal stability of silicene." *Physical Review B* 95.22 (2017): 224103.

Si PINN subgroup RMSE



Aluminum

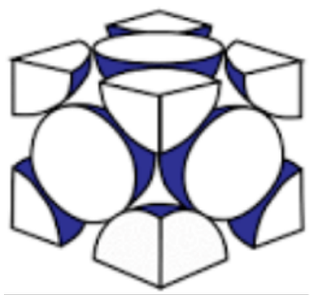
("To be published") G. P. Purja Pun(1), R. Batra(2), R. Ramprasad (3) and Y. Mishin(1)
(1) George Mason Univ., (2) Univ. Connecticut, (3) Georgia Tech

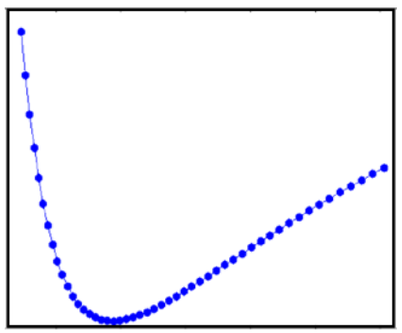
Aluminum NN and PINN Potential

Training/test set:

Stable Phase (ab-initio MD)

FCC (a=3.75 Å)
FCC (a=3.96 Å)
FCC (a=4.0 Å)
FCC (a=4.10 Å)
FCC (a=4.15 Å)
FCC (a=4.35 Å)





↓ T=700 K, 1200 K, 2500 K

expansion/compression

Database source:

V. Botu, R. Batra, J. Chapman, and R. Ramprasad. Machine learning force fields: Construction, validation, and outlook. *The Journal of Physical Chemistry C*, 121(1):511–522, 01 2017. URL: https://doi.org/10.1021/acs.jpcc.6b10908, doi:10.1021/acs.jpcc.6b10908.

Alternate structures (0K)

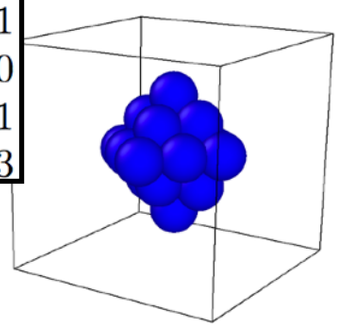
FCC	isotropic strain at 0 K
BCC	isotropic strain at 0 K
HCP	isotropic strain at 0 K
SC	isotropic strain at 0 K
DC	isotropic strain at 0 K
FCC	uniaxial ⟨100⟩ at 0 K
A15*	isotropic strain at 0 K
HEX*	isotropic strain at 0 K
FCC*	uniaxial ⟨100⟩ at 0 K
FCC*	uniaxial ⟨111⟩ at 0 K

Clusters+Defects (ab-initio MD)

GB 111
GB 210
GB 310
GB 320
GB 510
SF⟨211⟩,(111)

Surface 100
Surface 110
Surface 111
Surface 200
Surface 311
Surface 333

1 Vac
1 adatom on (100)
2 adatoms on (111)
Dimer adatom on (111)
Trimer adatom on (111)



Full set=Test+Training=127,592 atoms (3649 structures)

(random subsets) ↓ ↳ 108,052 atoms (3154 structures)

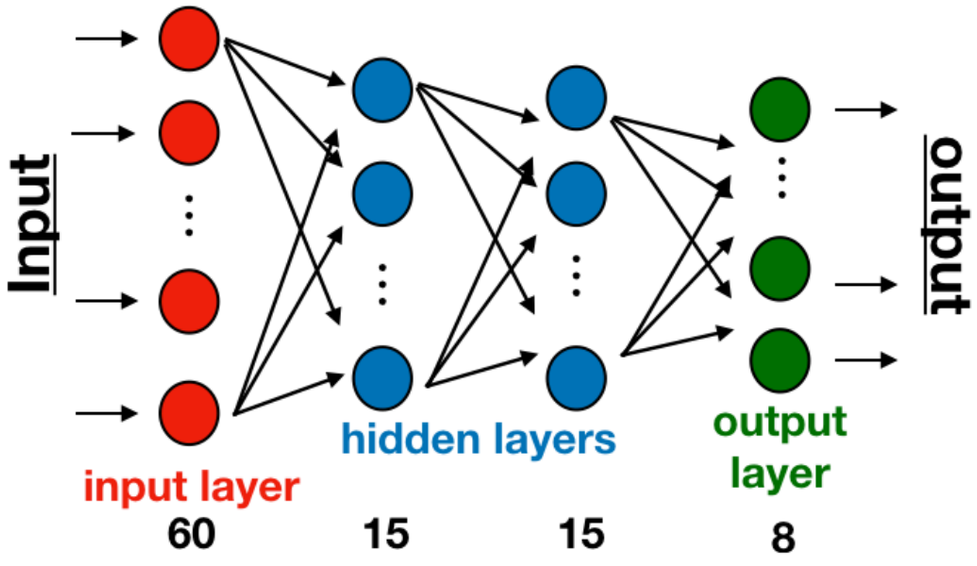
19,540 atoms (495 structures)

Methodological details:

- Both straight and physical NN
- DFT energy shifted by 0.38446 eV/atom

PINN:

15×15×8



Fit	Size	N_P	N_{Gi}	N_{r_0}
NN	16×16×1	1265	5	12
Physical NN	15×15×8	1283	5	12

NN fitting parameters

60 Gi's total

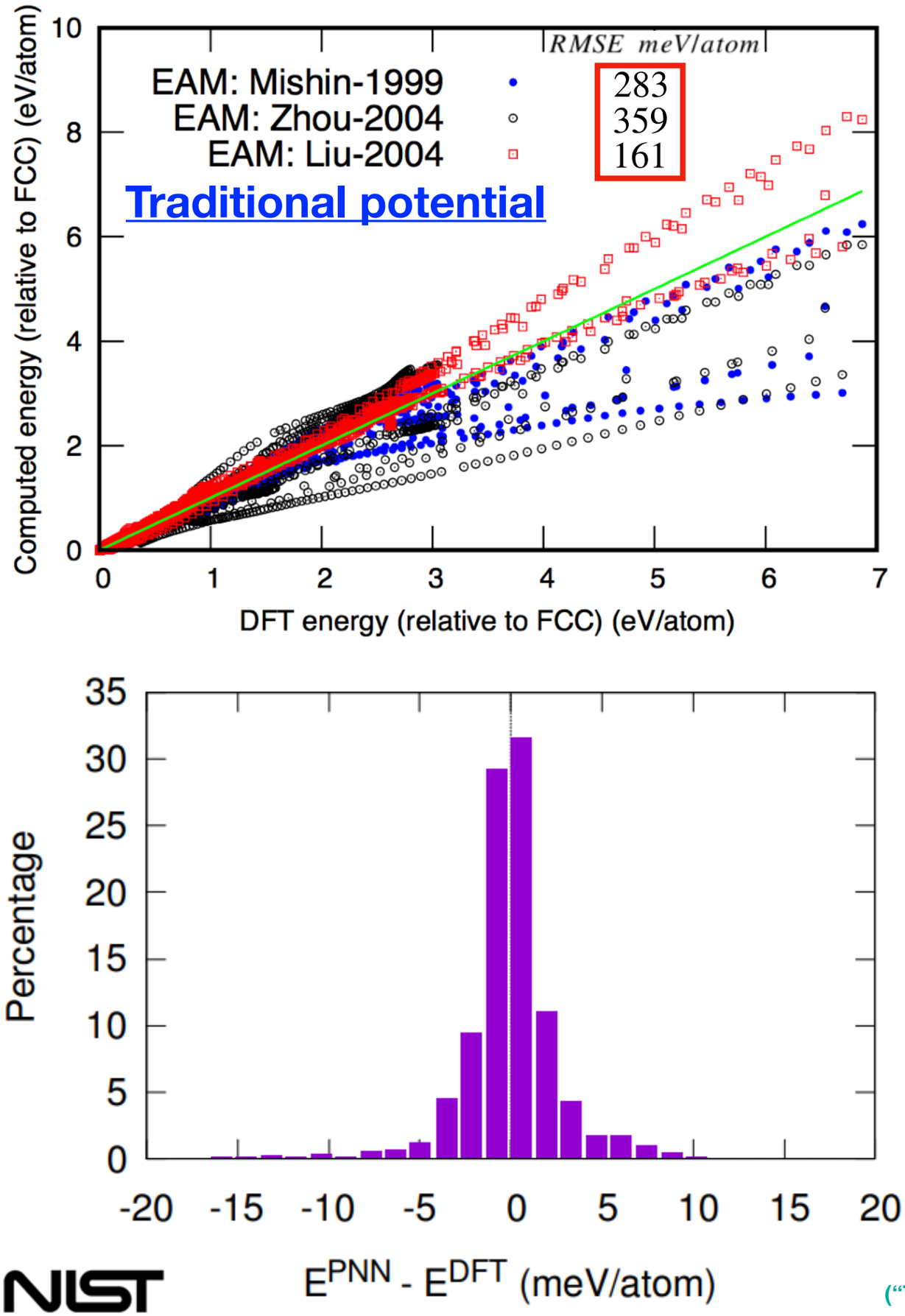
Cutoff function parameters (global)

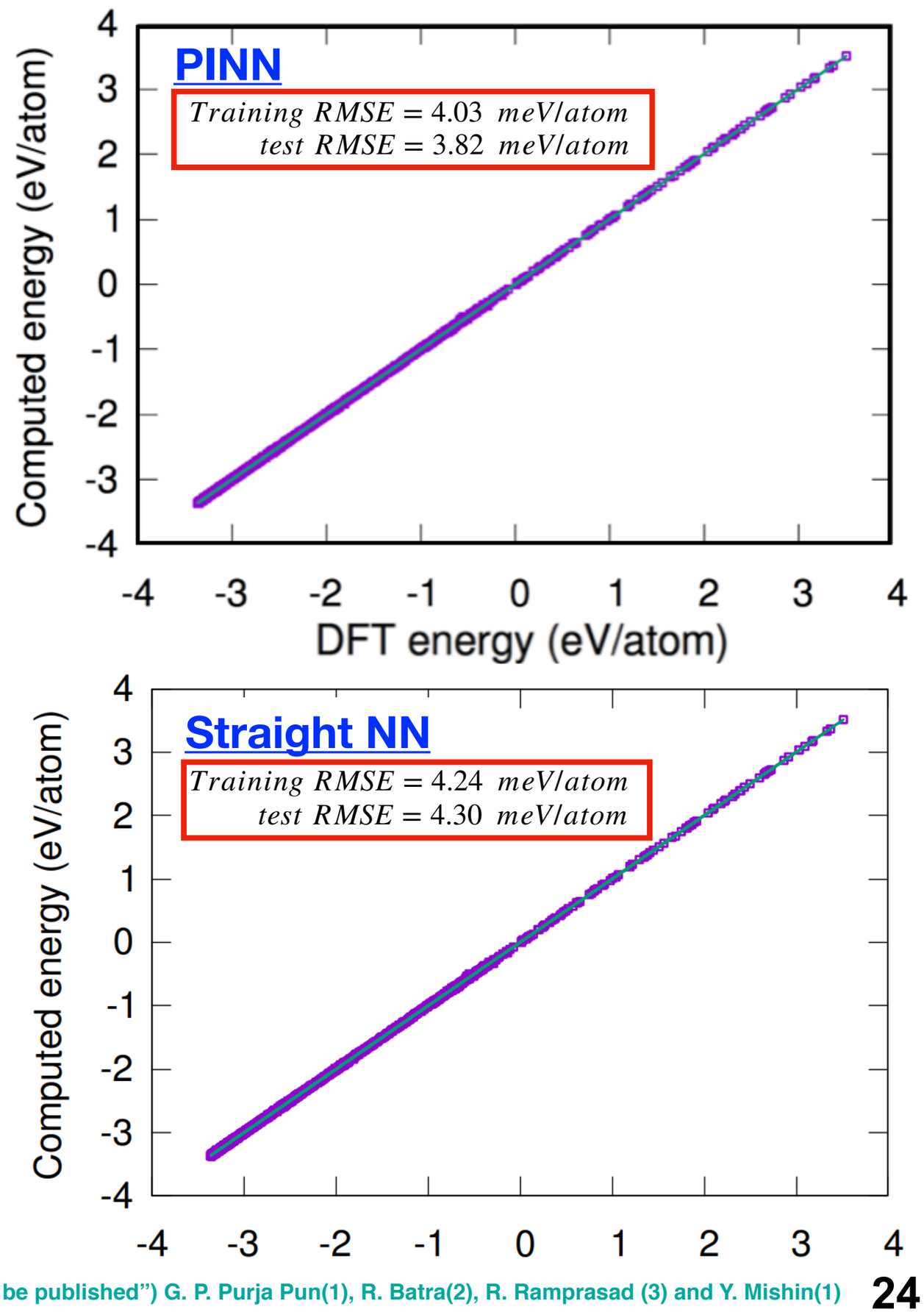
$r_c = 0.6\ nm$

$d = 0.15\ nm$

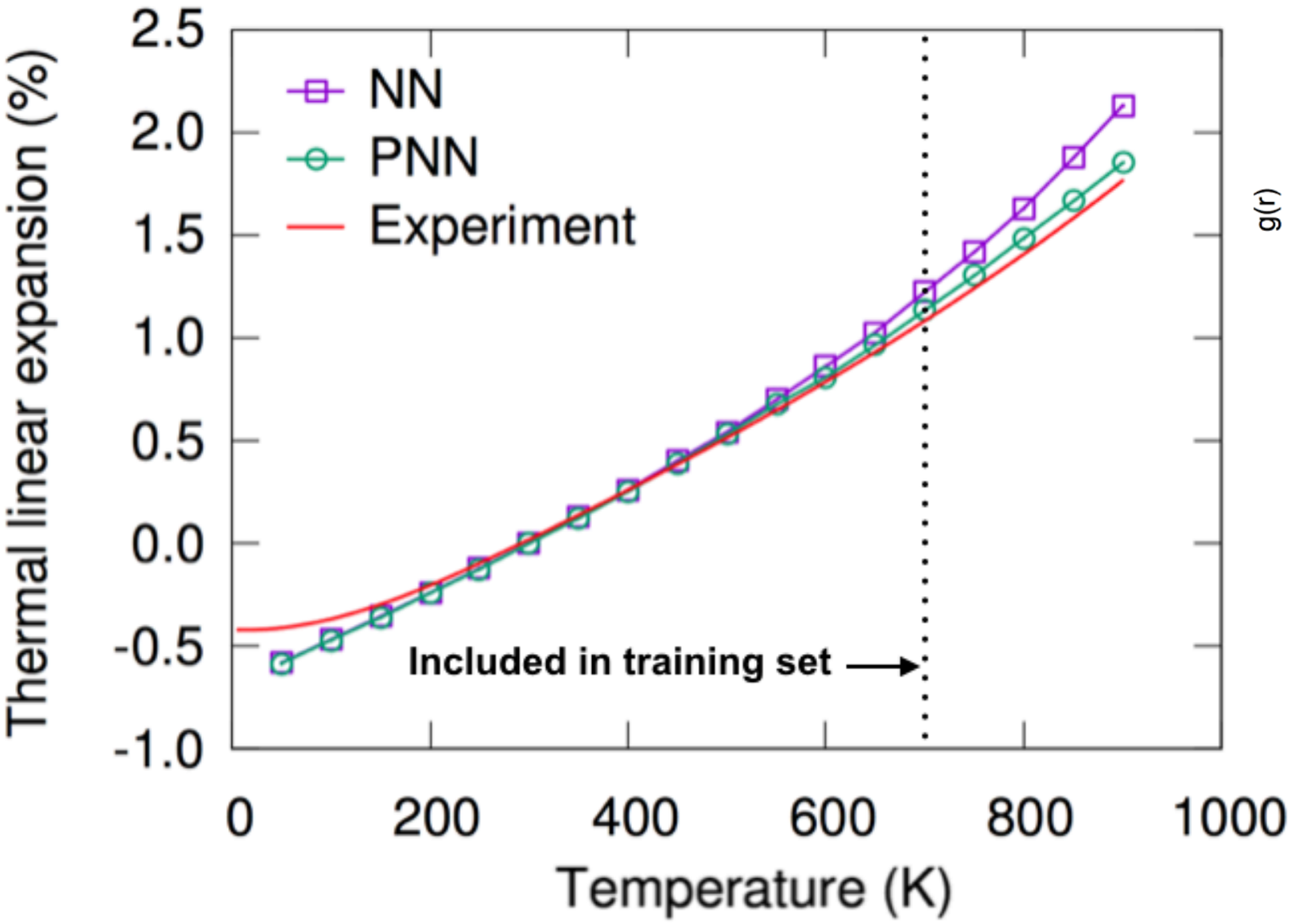
NIST

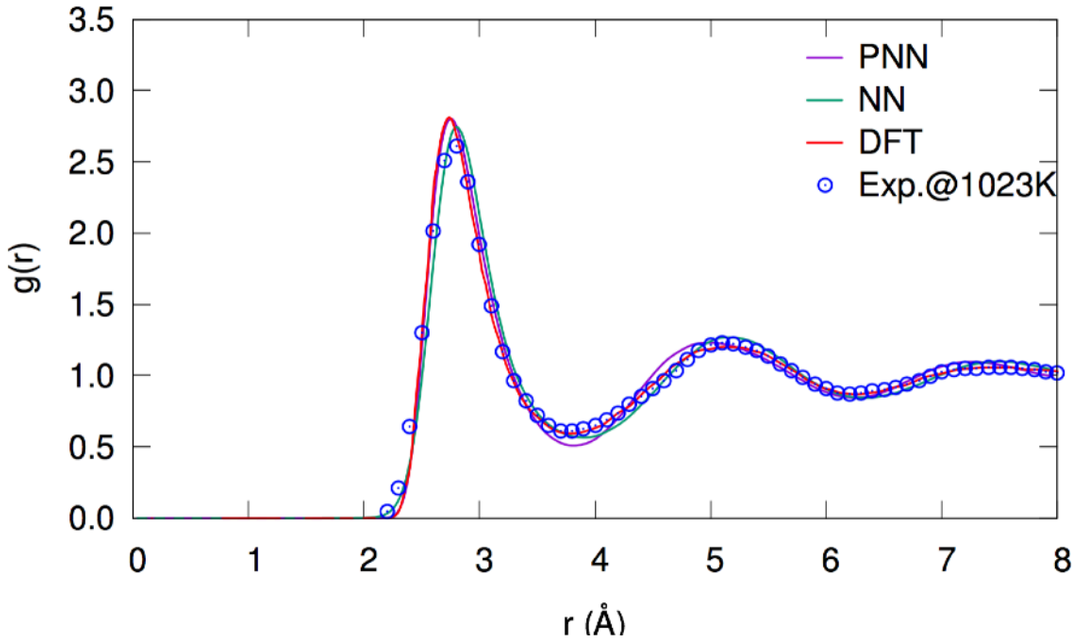
Results: Aluminum NN and PINN Potential

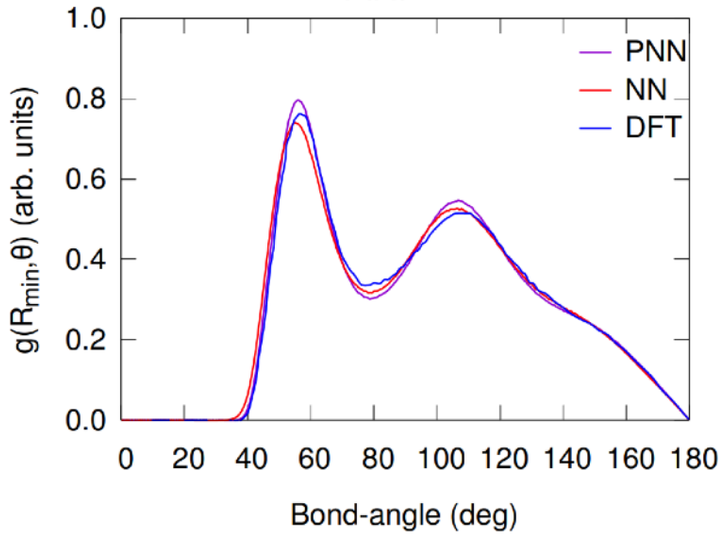




Results: Aluminum NN and PINN Potential







Additional Properties:

Property	Ab initio	NN	Physical NN
E_0 (eV/atom)	3.7480[a]	-3.3609	-3.3611
a_0 (Å)	4.039[a,d]; 3.9725–4.0676[c]	4.0431	4.0398
B (GPa)	83[a]; 81[f]	79	81
c_{11} (GPa)	104[a]; 103–106[d]	109	118
c_{12} (GPa)	73[a]; 57–66[d]	65	62
c_{44} (GPa)	32[a]; 28–33[d]	26	30
$\gamma_s(100)$ (Jm^{-2})	0.92[b]	0.8974	0.9047
$\gamma_s(110)$ (Jm^{-2})	0.98[b]	1.0089	0.9644
$\gamma_s(111)$ (Jm^{-2})	0.80[b]	0.8471	0.8238

Property	Ab initio	NN	Physical NN
E_v^f (eV)	0.6646–1.3458[c]; 0.7[e]	0.6558	0.7120
E_v^f (eV) unrelaxed	0.78[e]	0.7412	0.7830
E_I^f (T_d) (eV)	2.2001–3.2941[c]	2.7105	2.9133
E_I^f (O_h) (eV)	2.5313–2.9485[c]	2.1573	2.5480
E_I^f ⟨100⟩ (eV)	2.2953–2.6073[c]	1.8189	2.0558
E_I^f ⟨110⟩ (eV)	2.5432–2.9809[c]	2.7924	2.6725
E_I^f ⟨111⟩ (eV)	2.6793–3.1821[c]	2.6073	2.8375
γ_{SF} (mJ/m^2)	145.67[g]	122	130
γ_{us} (mJ/m^2)		130	140

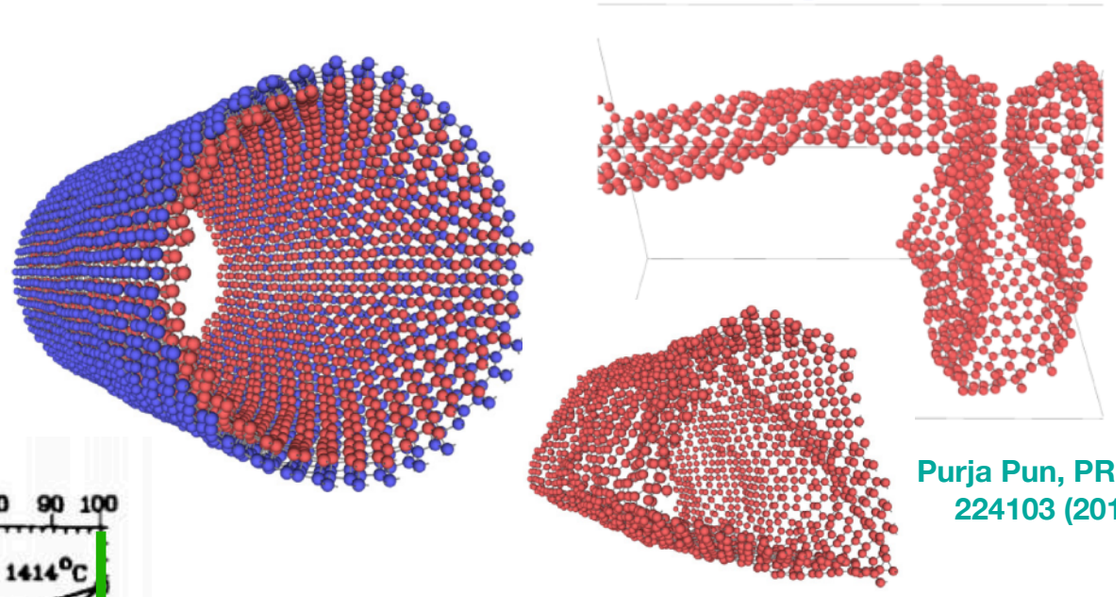
Future work

Future work-2

Germanium PINN potential:

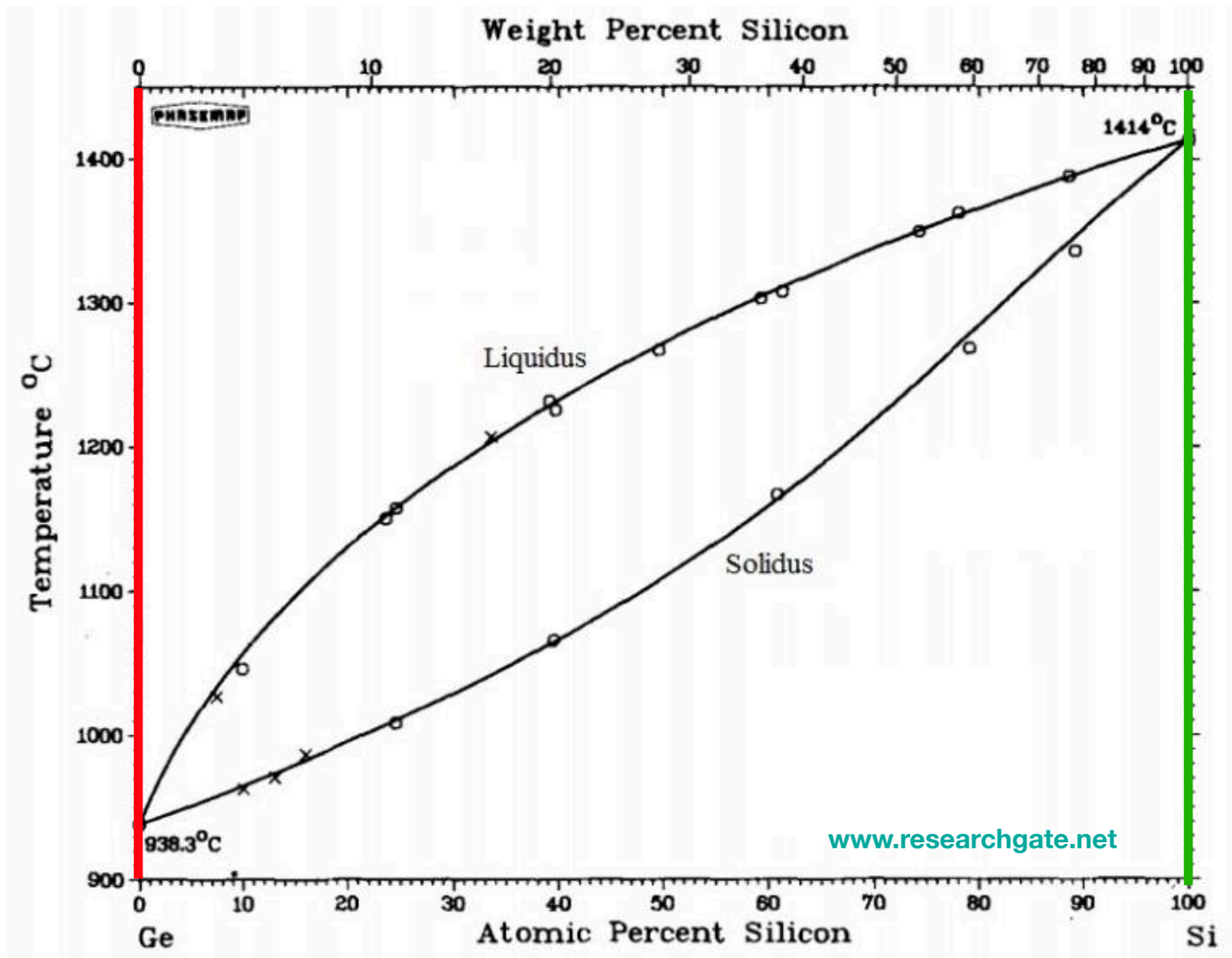
→ 2D Germanene stability

↓ Germanene thermal conductivity



Purja Pun, PRB 95, 224103 (2017)

Binary SiGe PINN potential:



- **Methodological:**
  - **leap frog**
  - **forward thinking**
  - **NN architecture**
- **Binary training set**
- **inheritance**
- **scalability**

Ge PINN ⟷ **Si PINN**

Si-Ge PINN

Conclusions

- Developed a new silicon interatomic potential using the new PINN potential format
- Even in preliminary stage we are obtaining excellent agreement with the DFT energies
- Current potential reproduces DFT data around 300x better than current traditional potentials
- Investigating methodological considerations to streamlining the fitting procedure for faster future development

Acknowledgements

- Ganga Purja Pun
- Kamal Choudhary
- Vesselin Yamakov
- Francesca Tavazza
- Yuri Mishin
- GMU,NRC,NIST

Image sources:

- https://www.tf.uni-kiel.de/matwis/amat/iss/kap_5/backbone/r5_3_3.html
- https://www.wordstream.com/blog/ws/2017/07/28/machine-learning-applications
- https://www.researchgate.net/figure/268158499_fig1_Figure-1-Phase-diagram-of-SiGe-alloys-showing-separation-of-the-solidus-and-liquidus
- http://evolution.skf.com/us/bearing-research-going-to-the-atomic-scale
- /https://www.chegg.com/homework-help/questions-and-answers/consider-concentric-metal-sphere-spherical-shells-shown-figure--innermost-solid-sphere-rad-q4808250

Stillinger, Frank H., and Thomas A. Weber. "Computer simulation of local order in condensed phases of silicon." *Physical review B* 31.8 (1985): 5262.

("To be published") G. P. Purja Pun(1), R. Batra(2), R. Ramprasad (3) and Y. Mishin(1): (1) George Mason Univ., (2) Univ. Connecticut, (3) Georgia Tech

Pun, GP Purja, and Y. Mishin. "Optimized interatomic potential for silicon and its application to thermal stability of silicene." *Physical Review B* 95.22 (2017): 224103.

Extra Slides

Si PINN comparison with of properties

Properties:

```
Total number of clusters: 210
Check total number of configurations: 1640
Check total RMSError: 0.007201
deformation = 0.99976546
equil. energy of diam structure = -4.62903632 (eV/atom)
equil. volume of diam structure = 20.429451 (A^3/atom)
energy of (100) surface = 2.183343 (J/m^2)
energy of (110) surface = 1.776560 (J/m^2)
energy of (111) surface = 1.314244 (J/m^2)
vacancy formation energy = 3.20324940 (eV) (-4.62120403e+03)
Td interstitial formation energy = 4.013617 (eV)
<100> dumbbell formation energy = 8.435771 (eV)
<110> dumbbell formation energy = 23.287577 (eV)
HEX interstitial formation energy = 4.287206 (eV)
bond-center interstitial formation energy = 72.593833 (eV)
bulk modulus = 97.181035 (GPa)
c11 = 125.226780 (GPa)
shear modulus = 42.065657 (GPa)
c12 = 83.161123 (GPa)
c44 = 105.045461 (GPa)
EXECUTION TIME ON 32 CPU(S): 191.889600 s
```

Property	Experiment	DFT
E_c (eV/atom)	4.63[c]	4.84[r]
a (Å)	5.430[a]	5.451[r]
c_{11}(GPa)	165[a]; 167.40[b]	
c_{12} (GPa)	64[a]; 65.23[b]	
c_{44} (GPa)	79.2[a]; 79.57[b]	
ν_{max} (THz)	15.7[o]	
Vacancy:		
E_v^f (T_d) (eV)	3.6[j]	3.17[m]; 3.69[t]
		3.29–4.3[h]; 3.70 – 3.84[s]
E_v^f (D_{3d}) (eV)		3.97[t]; 4.29[v]; 4.37[n]
		3.67 – 3.70[s]; 5.023[i]
Interstitials:		
E_i^f (hex) (eV)		3.31–5[h]; 2.87 – 3.80[s]
E_i^f (T_d) (eV)		3.43–6[h]; 3.43 – 5.10[s]
E_i^f (B) (eV)		4–5[h]
E_i^f ⟨110⟩(eV)		3.31–3.84[h]; 2.87 – 3.84[s]
Surface energy γ_s (J m^{-2}):		
{111}	1.24[q]; 1.23[p]	1.57[l]; 1.74[f]
{100}		2.14[l]; 2.39[f]; 2.36[k]
$\{100\}_{2\times1}$	1.36[p]	1.71[g]; 1.45[f]; 1.51[k]
{110}	1.43[p]	1.75[k]
Melting:		
T_m(K)	1687	
$\Delta V_m / V_{solid}$ (%)	−5.1[a]	
L (kJ/mol)	50.6[a]	

Training and test set

1	A1	AMORPHOUS	UNPERTURBED+ISO	25	64	1600	1600
2	A2	AMORPHOUS	PERTURBED+ISO	65	64	4160	5760
3	B1	DC-P	UNPERTURBED+ISO	90	2	180	5940
4	B10	HCP-P	UNPERTURBED+ISO	70	2	140	6080
5	B11	DC-P	UNPERTURBED+UNIAXIAL+100	35	2	70	6150
6	B14	DC-C	UNPERTURBED+UNIAXIAL+111	25	48	1200	7350
7	B15	DC-C	PERTURBED+ISO	170	64	10880	18230
8	B17	SH-P-(HEX)	UNPERTURBED+ISO	70	1	70	18300
9	B2	DC-P	UNPERTURBED+FINE+ISO	80	2	160	18460
10	B20	Wurtzite-P	UNPERTURBED+ANISO	90	4	360	18820
11	B21	ST12-P	UNPERTURBED+ISO	70	12	840	19660
12	B23	FCC-P	PERTURBED+ISO	20	8	160	19820
13	B24	HCP-P	PERTURBED+ANISO	20	16	320	20140
14	B25	BCC-P	PERTURBED+ISO	20	8	160	20300
15	B26	SH-P-(HEX)	PERTURBED+ANISO	20	8	160	20460
16	B27	BSN-P	PERTURBED+ANISO	20	16	320	20780
17	B28	BC8-C	PERTURBED+ISO	5	128	640	21420
18	B29	Wurtzite-P	PERTURBED+ANISO	20	32	640	22060
19	B3	BCC-P	UNPERTURBED+ISO	90	1	90	22150
20	B30	ST12-P	PERTURBED+ANISO	10	96	960	23110
21	B31	DC-P	FINE-UNPERTURBED+SHEAR-2	250	2	500	23610
22	B32	DC-P	FINE-UNPERTURBED+100	250	2	500	24110
23	B33	DC-P	FINE-UNPERTURBED+SHEAR-1	250	2	500	24610
24	B4	SC-C	UNPERTURBED+ISO	90	1	90	24700
25	B5	BC8-C	UNPERTURBED+ISO	90	16	1440	26140
26	B6	BSN-P	UNPERTURBED+ISO	70	2	140	26280
27	B7	Wurtzite-P	UNPERTURBED+ISO	70	4	280	26560
28	B8	cP46-C	UNPERTURBED+ISO	50	46	2300	28860
29	B9	FCC-P	UNPERTURBED+ISO	90	1	90	28950
30	C11	Si5_2-PENTAMER-PYRAMID	UNPERTURBED+ISO	25	5	125	29075
31	C12	Si5_3-PENTAMER-BRIDGE	UNPERTURBED+ISO	25	5	125	29200
32	C13	Si5_4-PENTAMER-BIPYRAMID	UNPERTURBED+ISO	25	5	125	29325
33	C15	Si6_1-HEXAMER-CHAIR	UNPERTURBED+ISO	25	6	150	29475
34	C16	Si6_2-HEXAMER-OCTAHEDRON	UNPERTURBED+ISO	25	6	150	29625
35	C18	Si6_4-HEXAMER-FACECAP	UNPERTURBED+ISO	25	6	150	29775
36	C21	Si7_1-HEPTAMER-BI-PYRAMID	UNPERTURBED+ISO	25	7	175	29950
37	C22	Si8_1-OCTAMER-BI-PYRAMID	UNPERTURBED+ISO	25	8	200	30150
38	C6	Si4_2-TETRAMER-TD	UNPERTURBED+ISO	25	4	100	30250

Training and test set

39	C7	Si4_3-D4h-TETRAMER-SQUARE	UNPERTURBED+ISO	25	4	100	30350
40	C9	Si4_5-TETRAMER	UNPERTURBED+ISO	25	4	100	30450
41	D1	DC-C	1-VACANCY+PERTURBED+ISO	50	7	350	30800
42	D10	DC-C	110-INTERSTITIAL+PERTURBED+ISO	50	9	450	31250
43	D11	DC-C	2-VACANCY+PERTURBED+ISO	35	62	2170	33420
44	D12	DC-C	HEX-INTERSTITIAL+PERTURBED+ISO	35	65	2275	35695
45	D13	DC-C	Td-INTERSTITIAL+PERTURBED+ISO	35	65	2275	37970
46	D14	DC-C	B-INTERSTITIAL+PERTURBED+ISO	35	65	2275	40245
47	D15	DC-C	110-INTERSTITIAL+PERTURBED+ISO	35	65	2275	42520
48	D2	DC-C	1-VACANCY+PERTURBED+ISO	25	63	1575	44095
49	D3	DC-C	<-110><001>(110)+MESH	82	48	3936	48031
50	D4	DC-C	<010><001>(100)+MESH	116	32	3712	51743
51	D5	DC-C	<10-1><-12-1>(111)+MESH-GLIDE	106	96	10176	61919
52	D6	DC-C	<101><121>(111)+MESH-SHUFFLE	106	96	10176	72095
53	D7	DC-C	B-INTERSTITIAL+PERTURBED+ISO	50	9	450	72545
54	D8	DC-C	HEX-INTERSTITIAL+PERTURBED+ISO	50	9	450	72995
55	D9	DC-C	Td-INTERSTITIAL+PERTURBED+ISO	50	9	450	73445
56	L2	LIQUID	PERTURBED+ISO	65	64	4160	77605
57	S1	SI-SURFACE-(100)-R	PERTURBED+ISO	35	16	560	78165
58	S10	SI-SURFACE-(320)	PERTURBED+ISO	35	40	1400	79565
59	S11	SI-SURFACE-(321)	PERTURBED+ISO	35	36	1260	80825
60	S12	SI-SURFACE-(322)	PERTURBED+ISO	35	36	1260	82085
61	S2	SI-SURFACE-(110)-R	PERTURBED+ISO	35	16	560	82645
62	S3	SI-SURFACE-(111)-R	PERTURBED+ISO	35	24	840	83485
63	S4	SI-SURFACE-(111)	PERTURBED+ISO	35	24	840	84325
64	S5	SI-SURFACE-(210)	PERTURBED+ISO	35	20	700	85025
65	S6	SI-SURFACE-(211)	PERTURBED+ISO	35	24	840	85865
66	S8	SI-SURFACE-(310)	PERTURBED+ISO	35	24	840	86705
67	S9	SI-SURFACE-(311)	PERTURBED+ISO	35	18	630	87335
68	T1	SILICENE-1-LAYER-HONEYCOMB	UNPERTURBED+ISO	25	18	450	87785
69	T10	SILICENE-PLANAR-BILAYER-AA_p	PERTURBED+ISO	45	36	1620	89405
70	T11	SILICENE-PLANAR-BILAYER-AA_p	PERTURBED+ANISO	25	36	900	90305
71	T13	SILICENE-BUCKLED-BILAYER-AA^p	PERTURBED+ISO	45	36	1620	91925
72	T14	SILICENE-BUCKLED-BILAYER-AA^p	PERTURBED+ANISO	25	36	900	92825
73	T15	SILICENE-BUCKLED-BILAYER-AB	UNPERTURBED+ISO	25	36	900	93725
74	T16	SILICENE-BUCKLED-BILAYER-AB	PERTURBED+ISO	45	36	1620	95345
75	T17	SILICENE-BUCKLED-BILAYER-AB	PERTURBED+ANISO	25	36	900	96245
76	T2	SILICENE-1-LAYER-HONEYCOMB	PERTURBED+ISO	45	18	810	97055
77	T3	SILICENE-1-LAYER-BUCKLED	UNPERTURBED+ISO	25	18	450	97505
78	T4	SILICENE-1-LAYER-BUCKLED	PERTURBED+ISO	45	18	810	98315
79	T5	SILICENE-1-LAYER-BUCKLED	PERTURBED+ANISO	25	18	450	98765
80	T7	SILICENE-1-LAYER-DUMBELL	PERTURBED+ISO	44	63	2772	101537
81	T8	SILICENE-1-LAYER-DUMBELL	PERTURBED+ANISO	25	63	1575	103112
82	T9	SILICENE-PLANAR-BILAYER-AA_p	UNPERTURBED+ISO	25	36	900	104012

Minimization algorithm

Davidon–Fletcher–Powell formula

From Wikipedia, the free encyclopedia

The **Davidon–Fletcher–Powell formula** (or **DFP**; named after William C. Davidon, Roger Fletcher, and Michael J. D. Powell) finds the solution to the secant equation that is closest to the current estimate and satisfies the curvature condition (see below). It was the first quasi-Newton method to generalize the secant method to a multidimensional problem. This update maintains the symmetry and positive definiteness of the Hessian matrix.

Given a function $f(x)$, its gradient (∇f), and positive-definite Hessian matrix B, the Taylor series is

$$f(x_k + s_k) = f(x_k) + \nabla f(x_k)^T s_k + \frac{1}{2} s_k^T B s_k + \dots,$$

and the Taylor series of the gradient itself (secant equation)

$$\nabla f(x_k + s_k) = \nabla f(x_k) + B s_k + \dots$$

is used to update B.

The DFP formula finds a solution that is symmetric, positive-definite and closest to the current approximate value of B_k:

$$B_{k+1} = (I - \gamma_k y_k s_k^T) B_k (I - \gamma_k s_k y_k^T) + \gamma_k y_k y_k^T,$$

where

$$y_k = \nabla f(x_k + s_k) - \nabla f(x_k),$$
$$\gamma_k = \frac{1}{y_k^T s_k},$$

and B_k is a symmetric and positive-definite matrix.

The corresponding update to the inverse Hessian approximation $H_k = B_k^{-1}$ is given by

$$H_{k+1} = H_k - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} + \frac{s_k s_k^T}{y_k^T s_k}.$$

B is assumed to be positive-definite, and the vectors s_k^T and y must satisfy the curvature condition

$$s_k^T y_k = s_k^T B s_k > 0.$$

The DFP formula is quite effective, but it was soon superseded by the BFGS formula, which is its dual (interchanging the roles of *y* and *s*).

Elastic Constants

DFT PBE

Jarvis (ionic relaxation)

$$c_{11} = 152.7\ GPa$$
$$c_{12} = 57.0\ GPa$$
$$c_{44} = 74.4\ GPa$$

https://www.ctcms.nist.gov/~knc6/jsmol/JVASP-1001.html

Material Project (ionic relaxation)

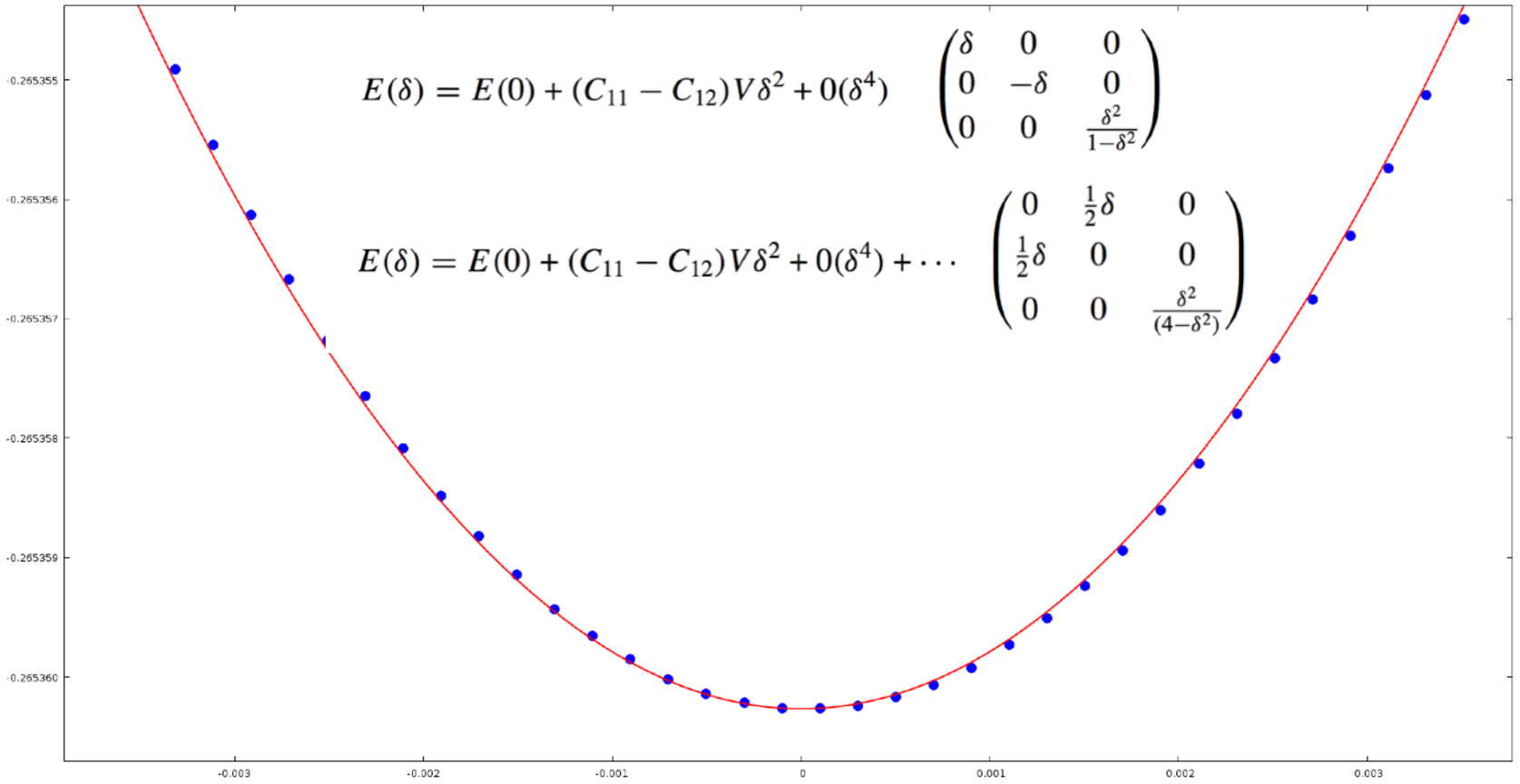
$$c_{11} = 144.0\ GPa$$
$$c_{12} = 53.0\ GPa$$
$$c_{44} = 75\ GPa$$

https://materialsproject.org/materials/mp-149/

Current (no ionic relaxation)

$$c_{11} = 153.3\ GPa$$
$$c_{12} = 56.7\ GPa$$
$$c_{44} = 97.2\ GPa$$

https://www.ctcms.nist.gov/~knc6/jsmol/JVASP-1001.html



Classical potential (mod/c)

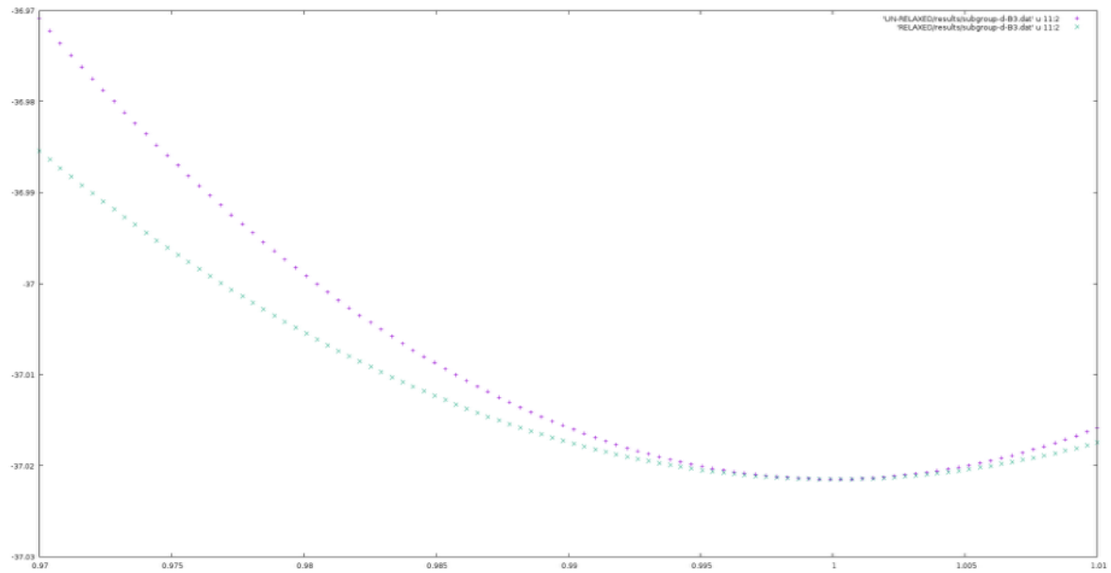
c_{11}(GPa)	172.6	165[a]; 167.40[b]
c_{12} (GPa)	64.6	64[a]; 65.23[b]
c_{44} (GPa)	81.3	79.2[a]; 79.57[b]
		experiment

(no ionic relaxation)

```
C11=171.554032204297 GPA
C12=66.5676160791408 GPA
C44=109.991181148439 GPA
```

(ionic relaxation)

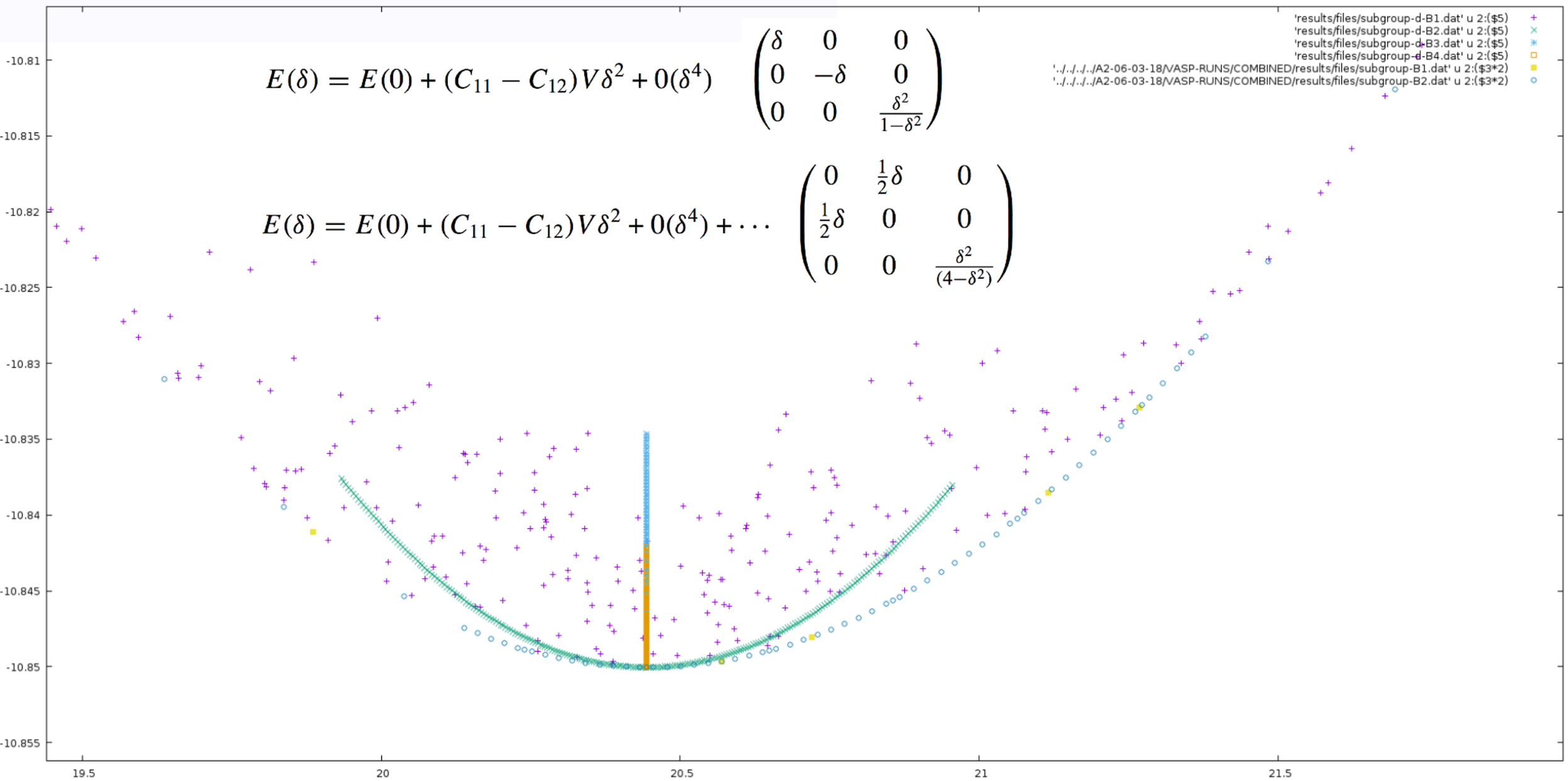
```
C11=171.554032204157 GPA
C12=66.567616082438 GPA
C44=78.4100485501238 GPA
```



Elastic Constants

$$E(\delta) = E(0) + (C_{11} - C_{12})V\delta^2 + 0(\delta^4) \quad \begin{pmatrix} \delta & 0 & 0 \\ 0 & -\delta & 0 \\ 0 & 0 & \frac{\delta^2}{1-\delta^2} \end{pmatrix}$$

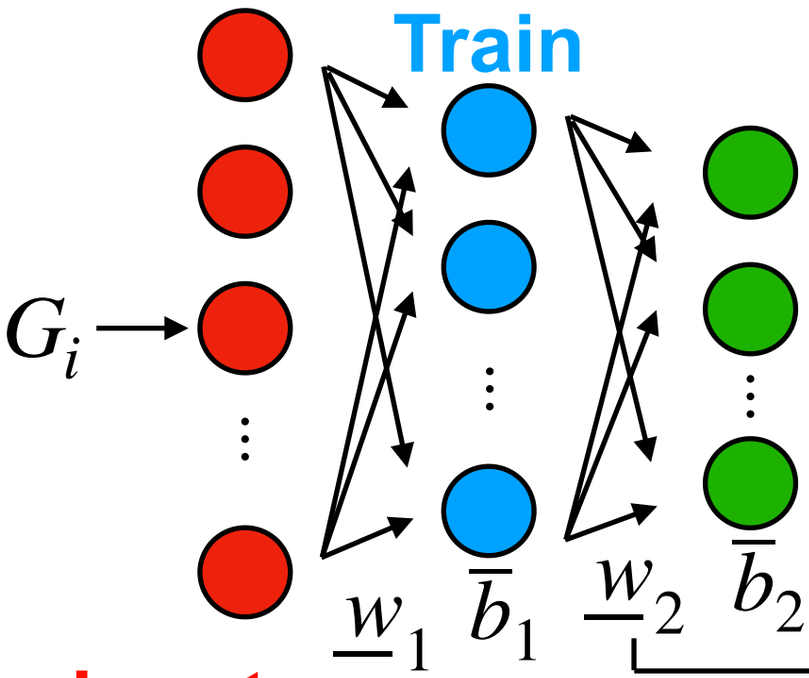
$$E(\delta) = E(0) + (C_{11} - C_{12})V\delta^2 + 0(\delta^4) + \cdots \quad \begin{pmatrix} 0 & \frac{1}{2}\delta & 0 \\ \frac{1}{2}\delta & 0 & 0 \\ 0 & 0 & \frac{\delta^2}{(4-\delta^2)} \end{pmatrix}$$



Future work-1: Forward thinking method

Stage-1

Train

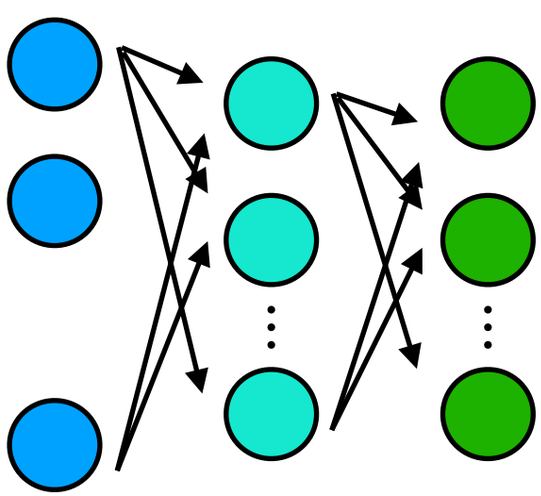


BOP Param

At the end of training dump output of hidden layer "new training set"

Stage-2

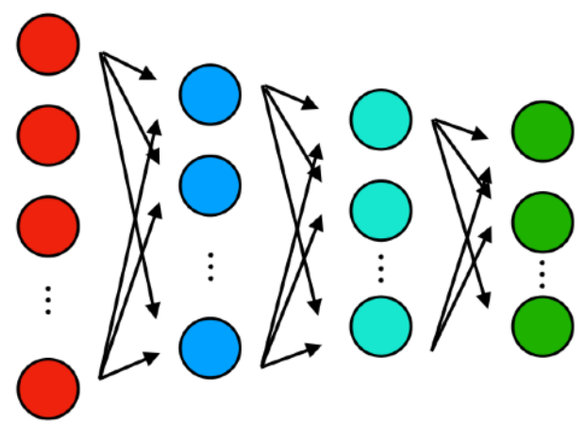
Discard: w_2 b_2

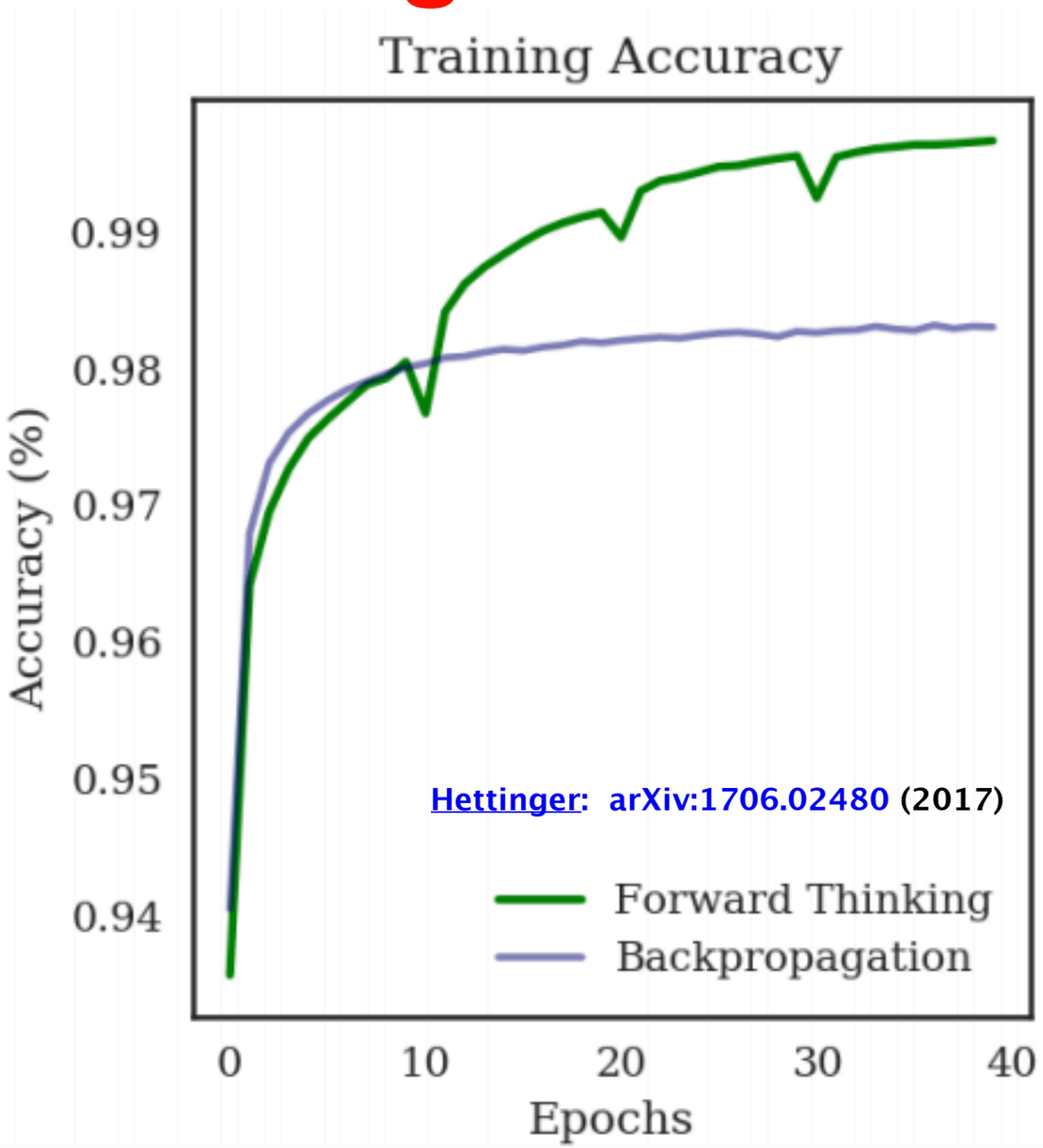


Train

BOP Param

Repeat and combine at the end



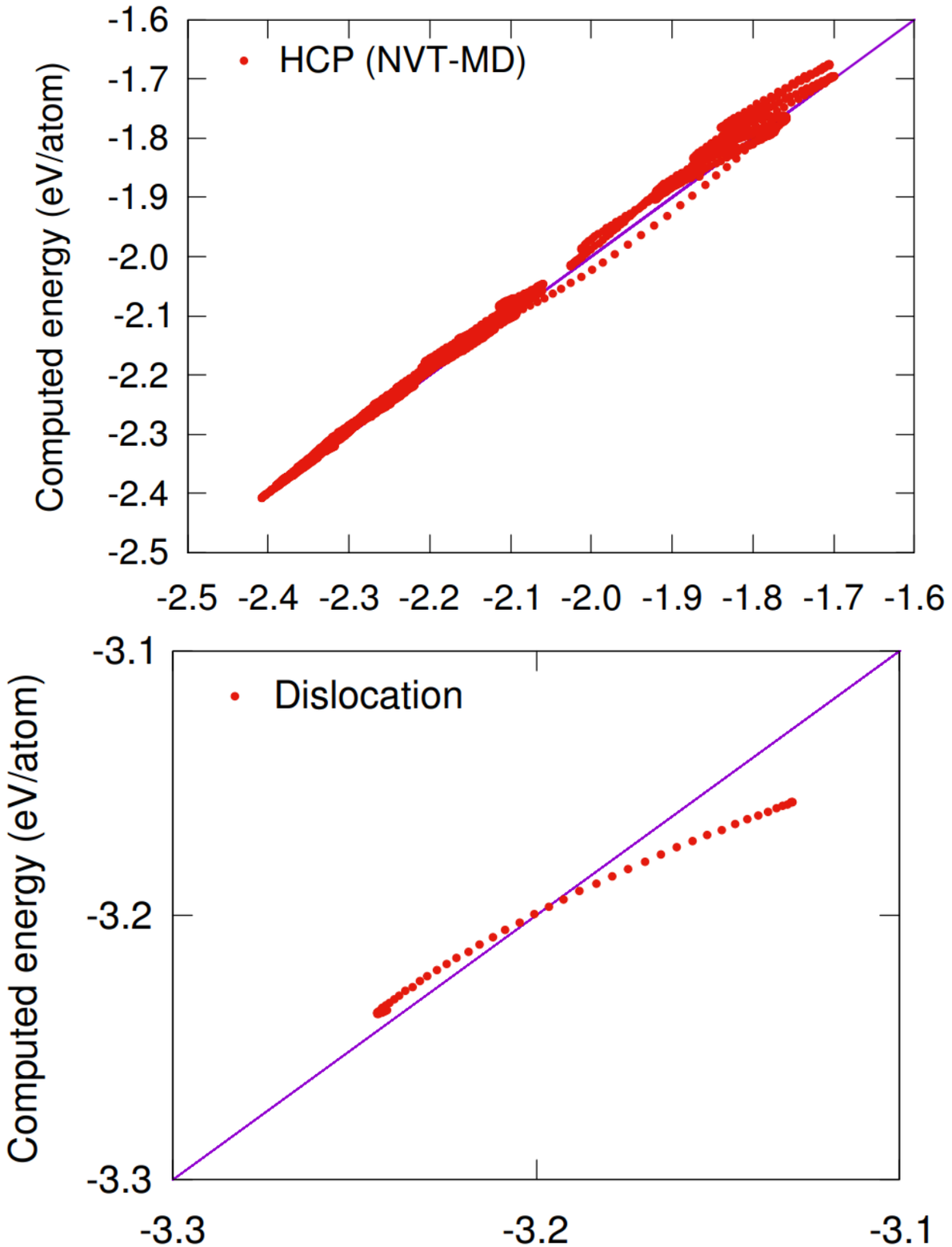


Hettinger: arXiv:1706.02480 (2017)

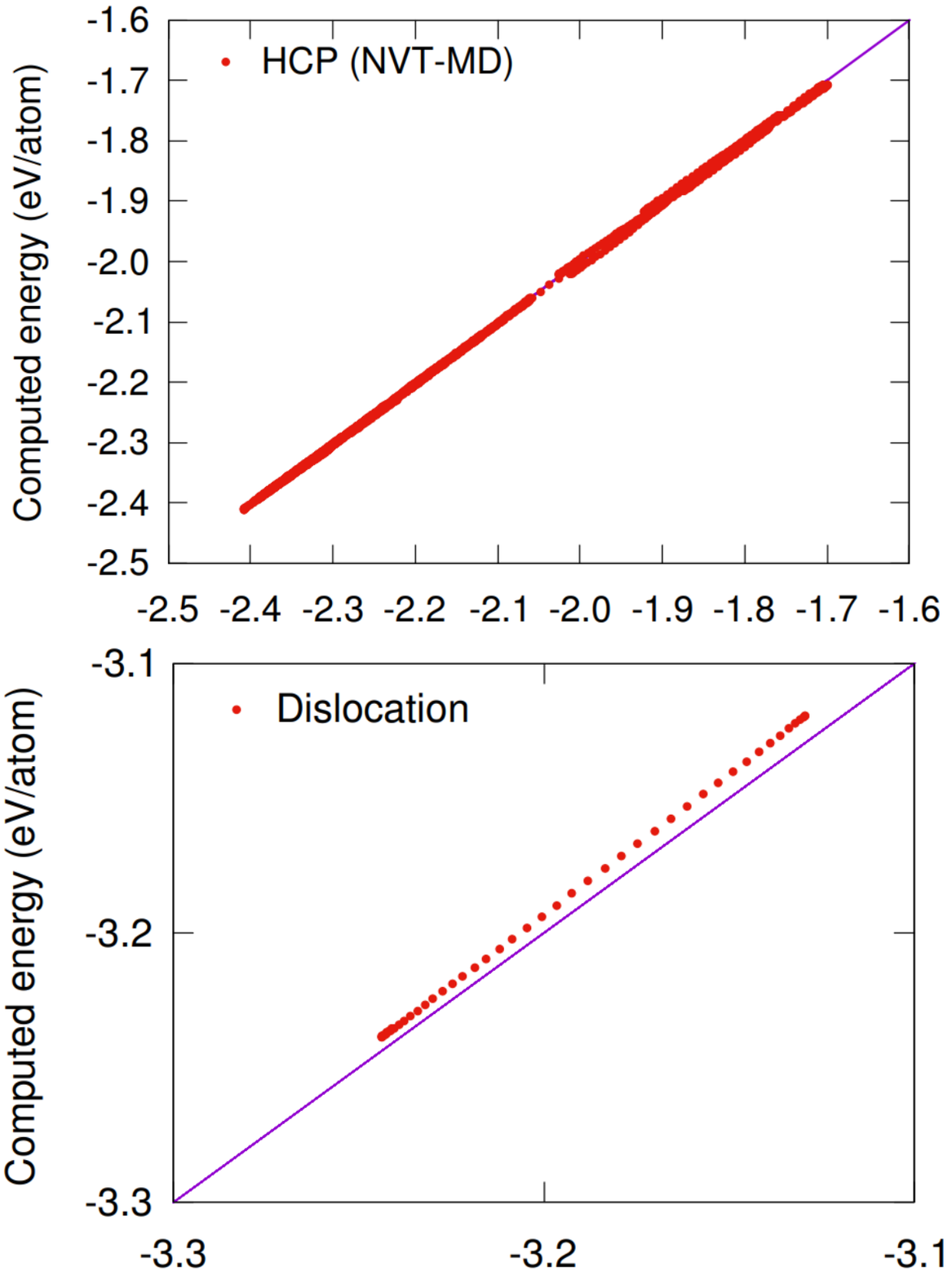
Dynamically build NN by training one layer at a time

Results: Aluminum NN and PINN Potential

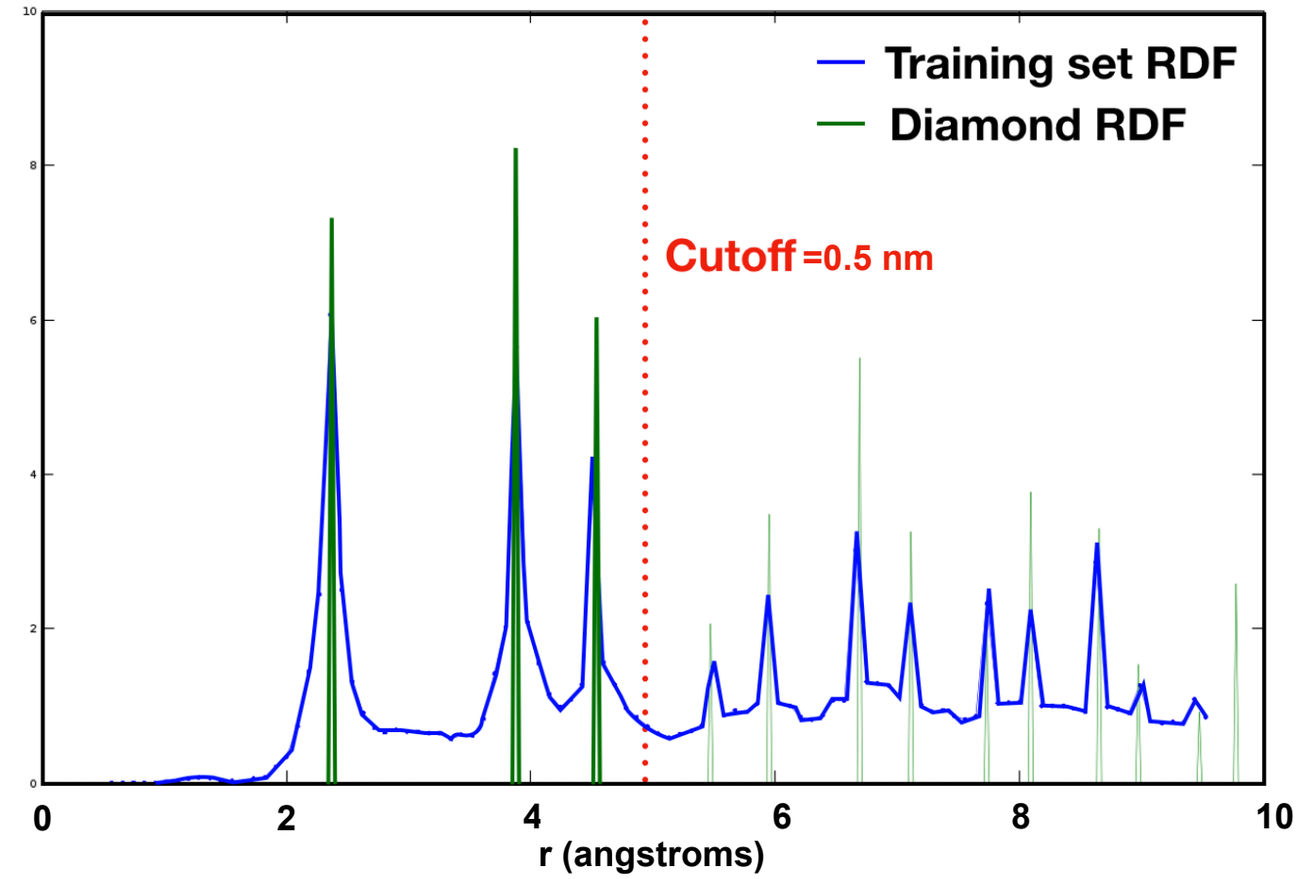
DFT energy (eV/atoms) (NN)

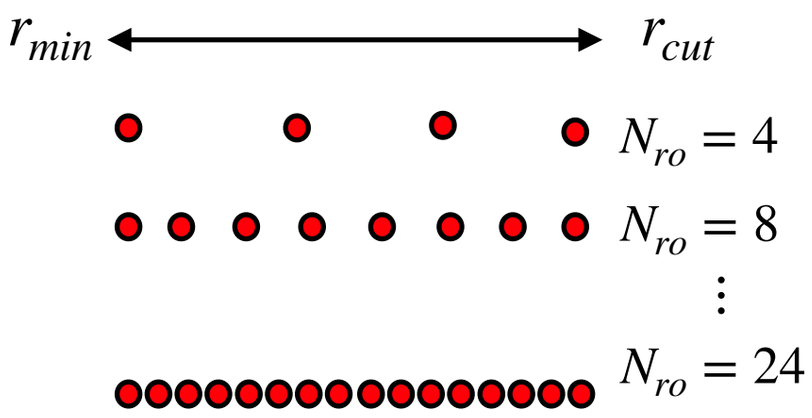


DFT energy (eV/atoms) (PINN)



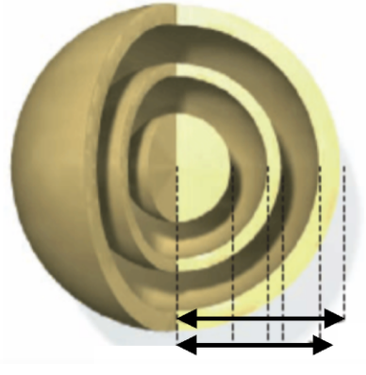
Local structure parameter choice



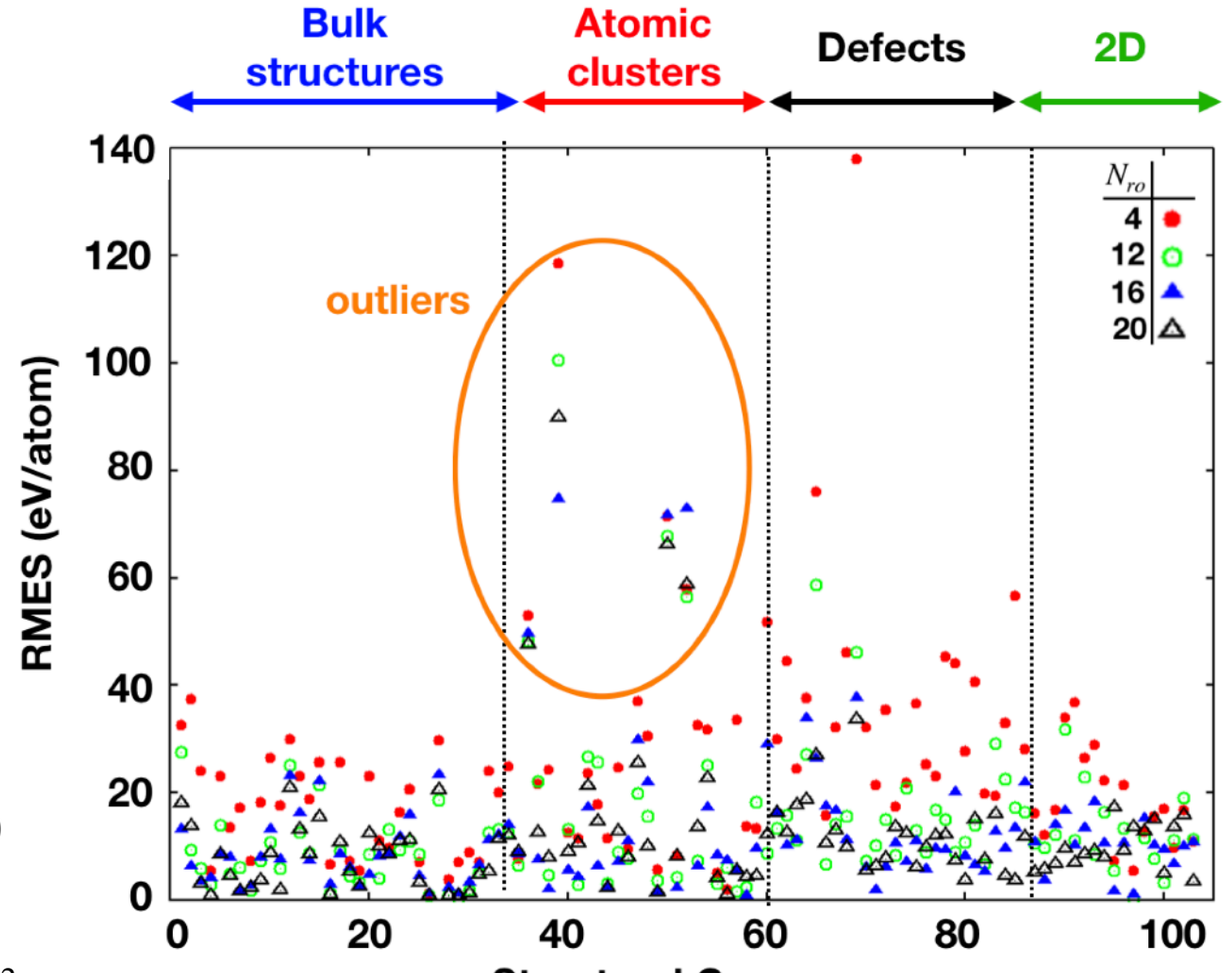


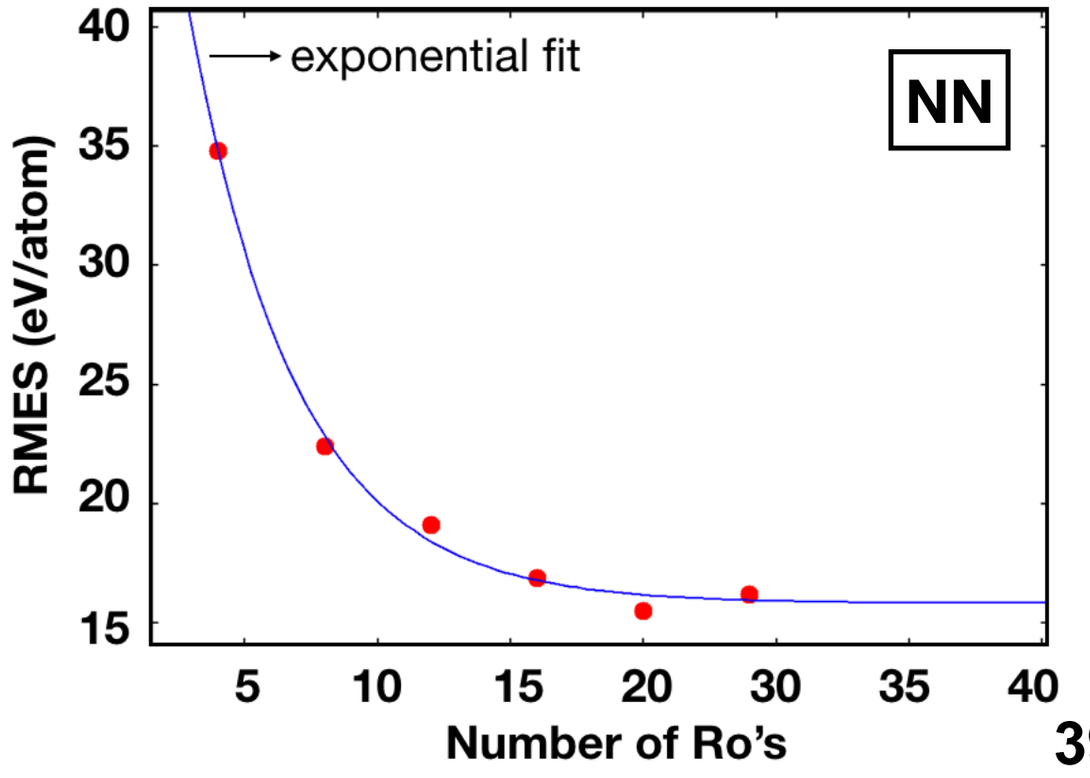
$$f(r) = \frac{1}{\sigma^3} e^{-\frac{(r - r_O)^2}{\sigma^2}} f_c(r)$$

$$\sigma = 0.1\ nm$$



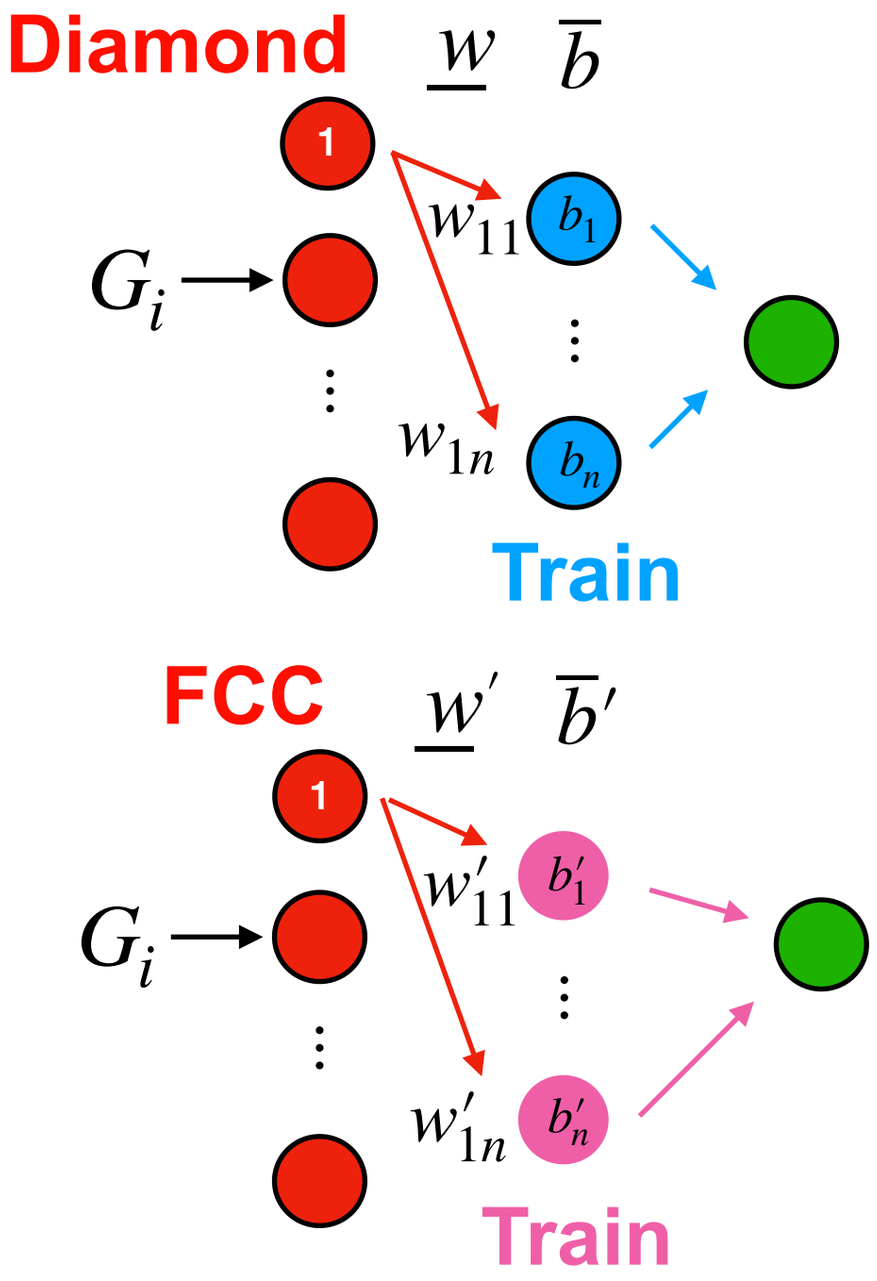
Observed a similar effect for PINN





Future work-1: Forward thinking method

Stage-1

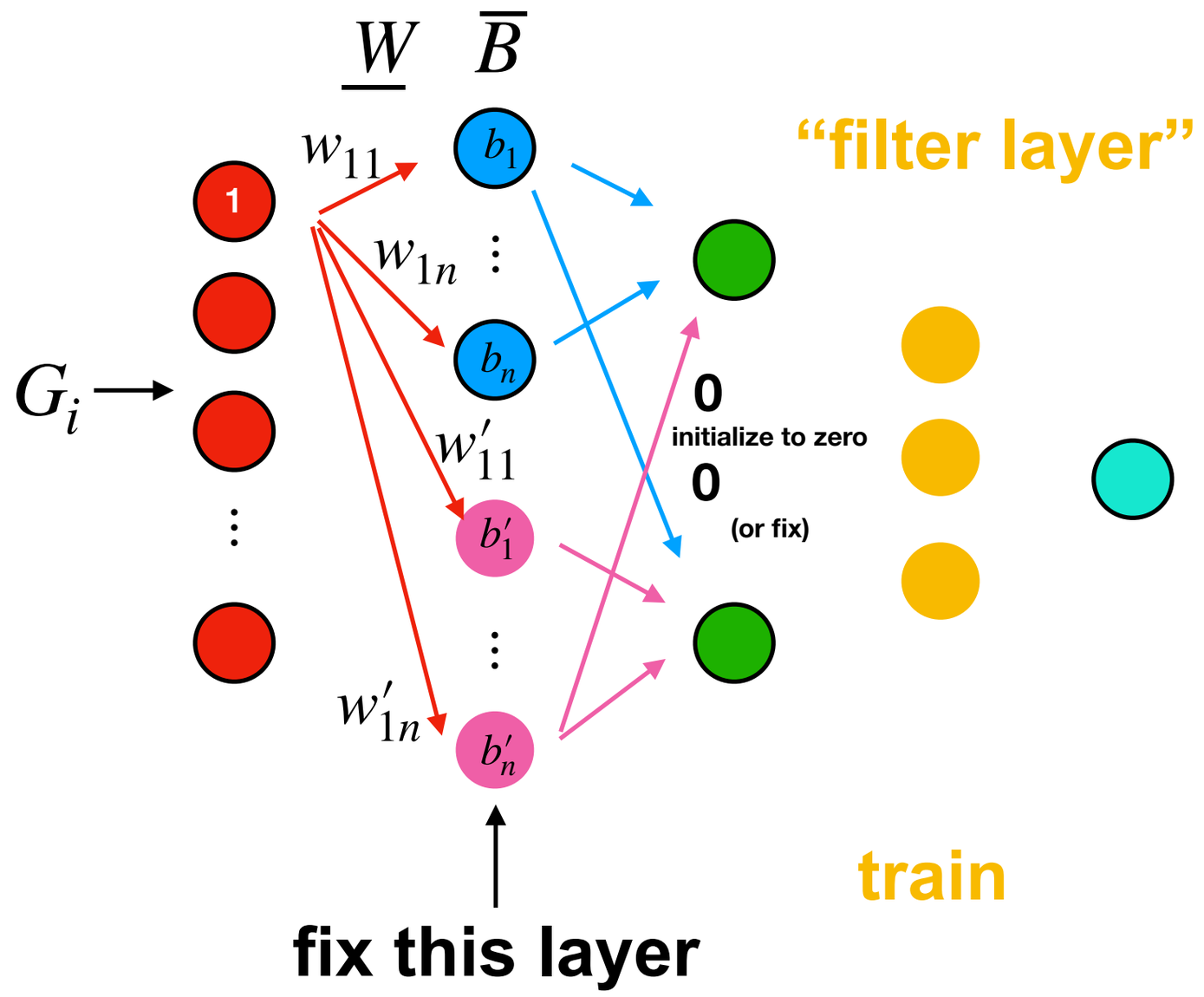


40-2-1

(40*2+2+2+1)=85 —> (~8500 fitting)

(40*1+1+1+1)=43 —> (~4300 fitting)

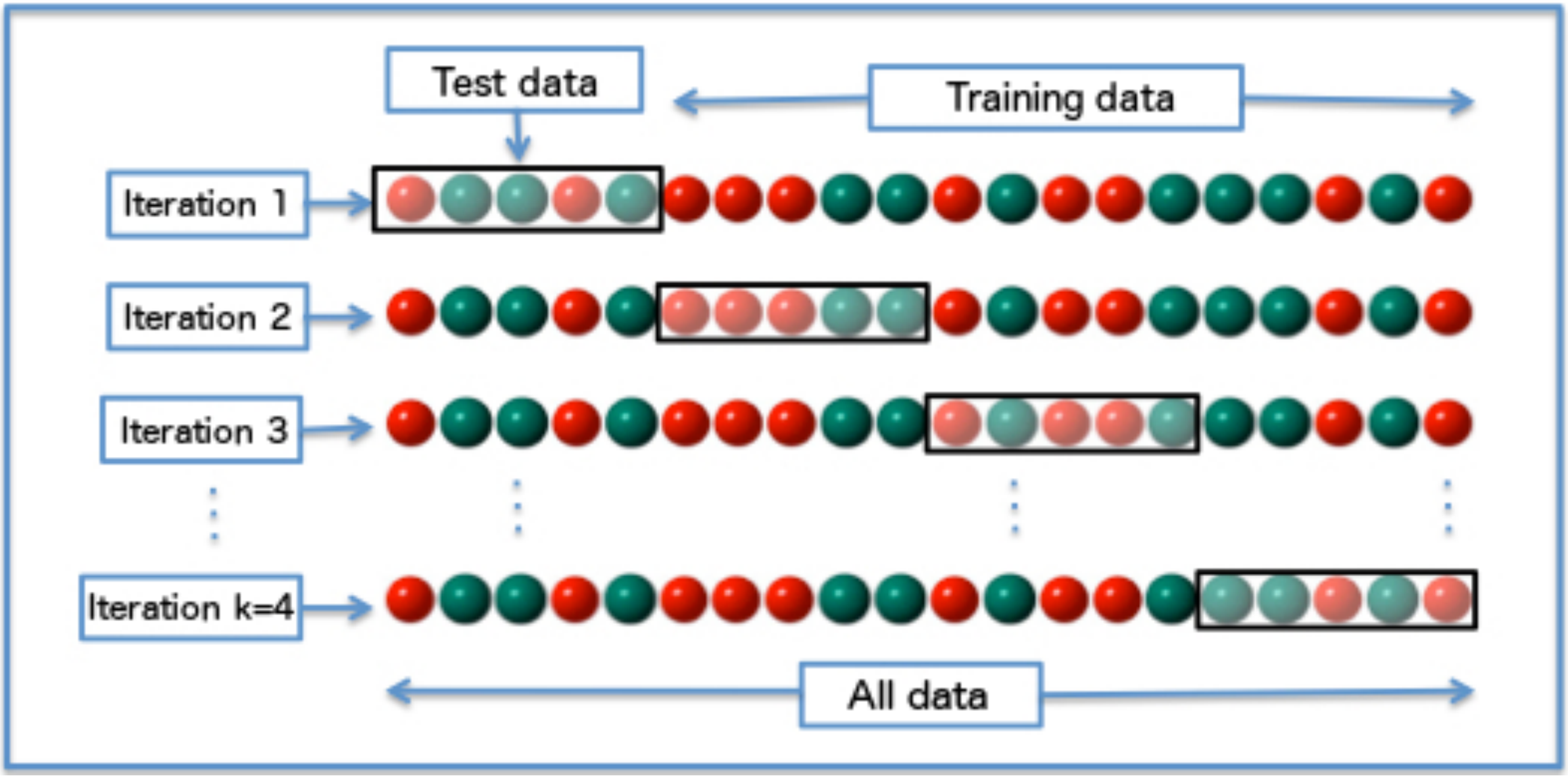
Stage-2



40 200 100

(40*2+2+2+1)=85 —> (~8500 fitting)

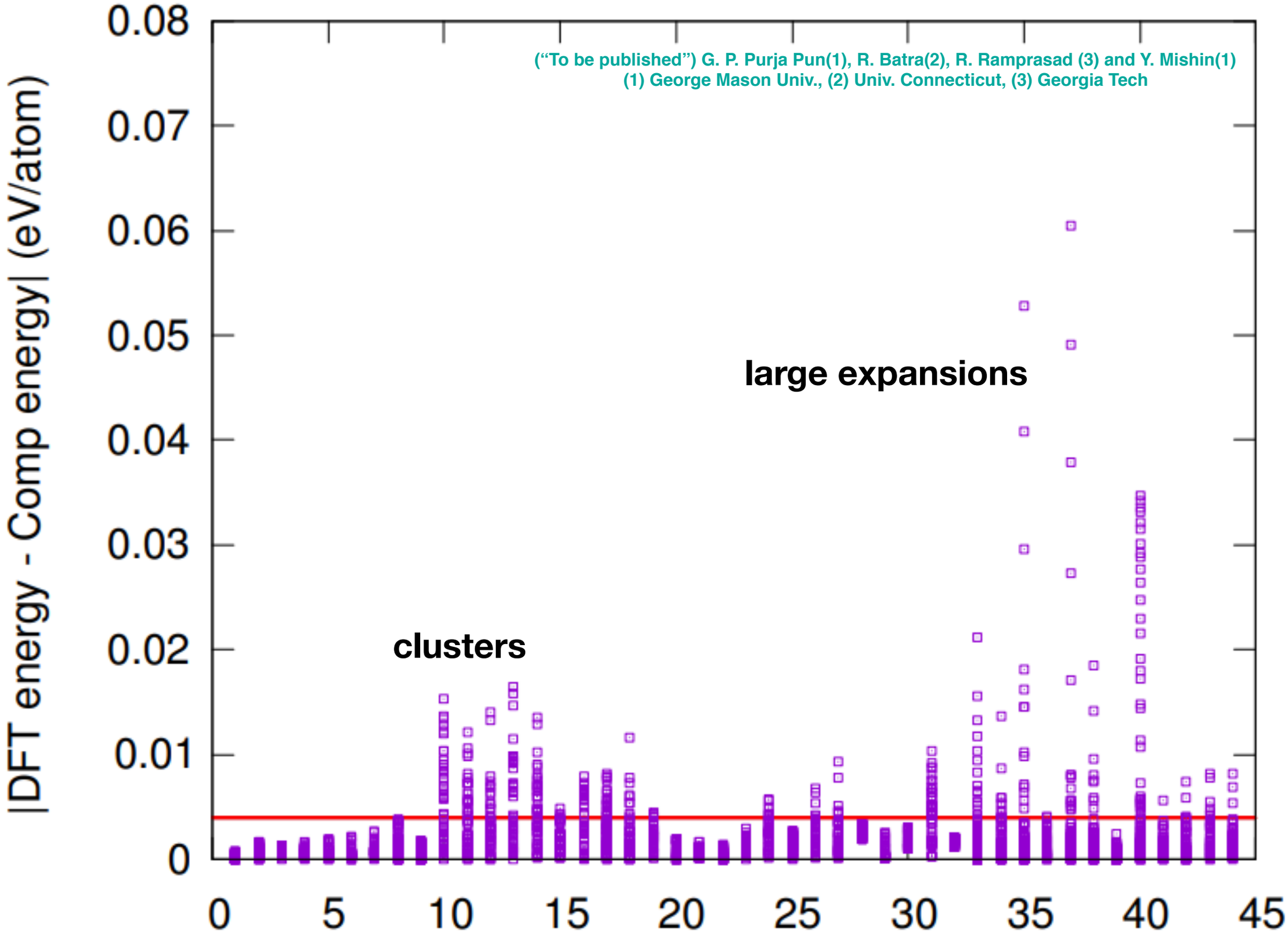
K-fold validation



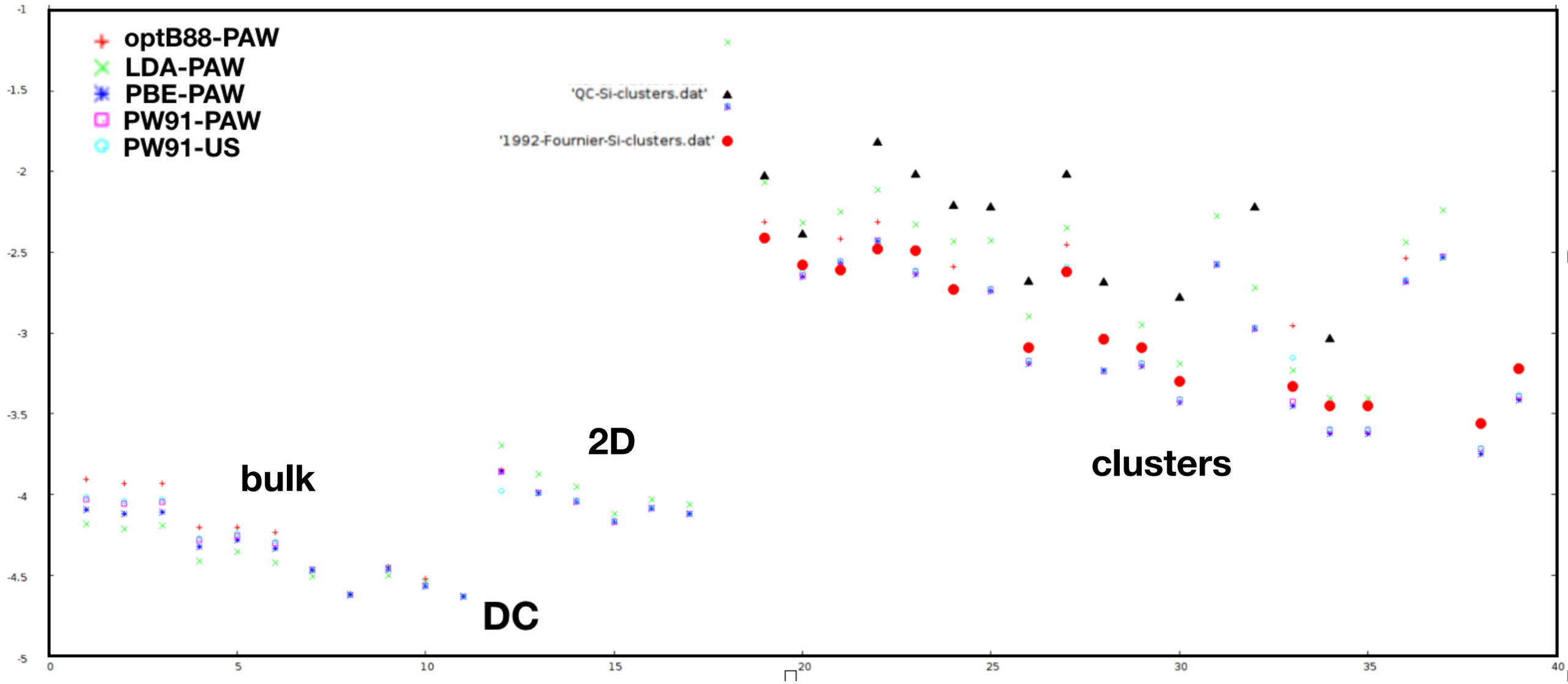
https://en.wikipedia.org/wiki/Cross-validation_(statistics)

Results: Aluminum PINN Potential

("To be published") G. P. Purja Pun(1), R. Batra(2), R. Ramprasad (3) and Y. Mishin(1)
(1) George Mason Univ., (2) Univ. Connecticut, (3) Georgia Tech



DFT variation



DFT variation

- **Points**
  - **Don’t focus so much on details (wiggles,NN size)**
  - **Say humans and animals rather than just animals**
  - **pairwise repulsion, angular dependence, longer distance interactions, bond order effects, screening**