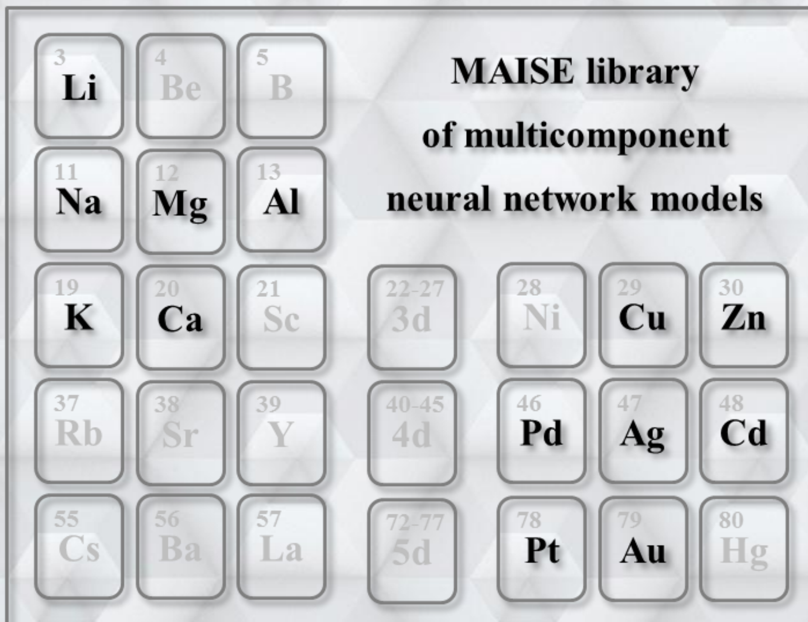
BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK

Construction of standardized neural network-based interatomic models for structure prediction acceleration

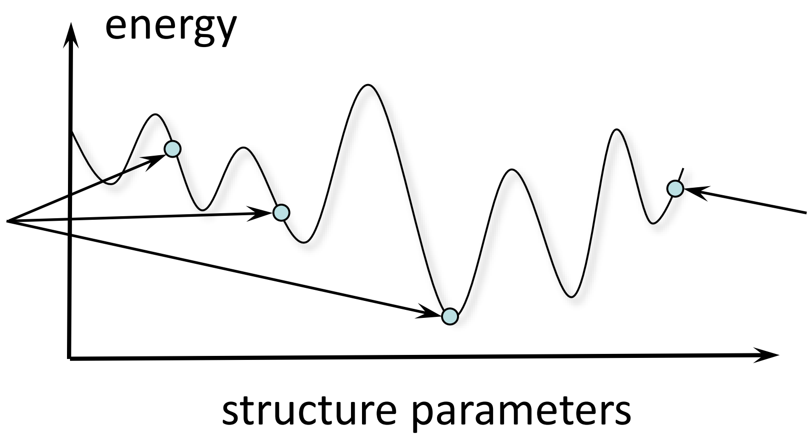
Alexey N. Kolmogorov

Binghamton University
State University of New York



Acceleration of materials prediction: outline

Efficient sampling
of parameter space



Accurate evaluation
of energy, force, stress

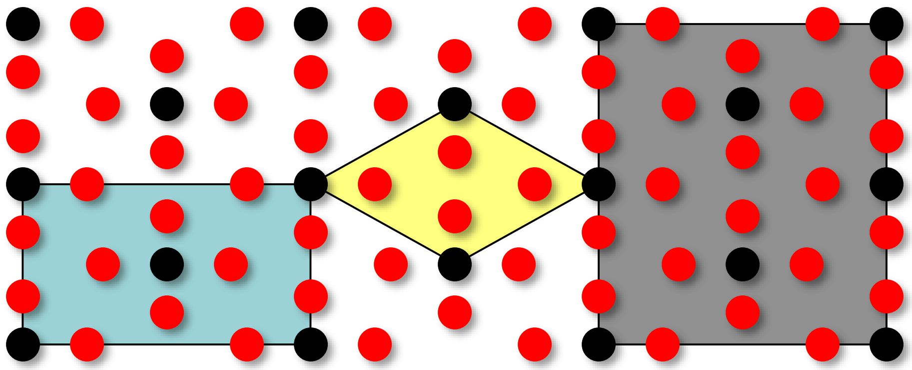
Structure prediction

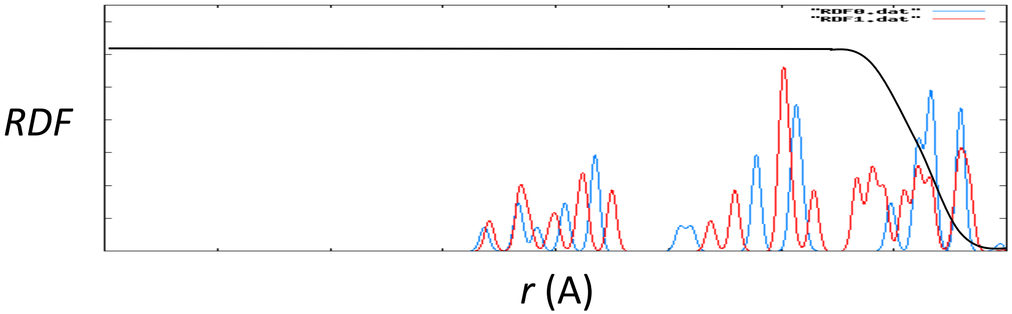
search strategies
confirmed predictions
need for acceleration

Interaction description with neural networks

advantages and challenges
hierarchical (stratified) construction
performance and application

Structure similarity via radial distribution function (RDF)





$$C_1 \times C_2 = \int_0^{R_{cut}} \sum_{ij} RDF_{1,ij}(r) RDF_{2,ij}(r) dr,$$

[Duplicate elimination in ES]

Kolmogorov *et al*. PRL 105, 217003 (2010)

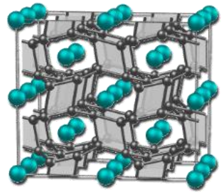
[Atomic environment analysis]

Choi *et al*. PRL 108, 127204 (2012)

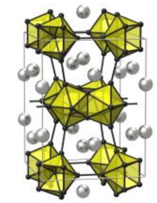
[Distortion analysis]

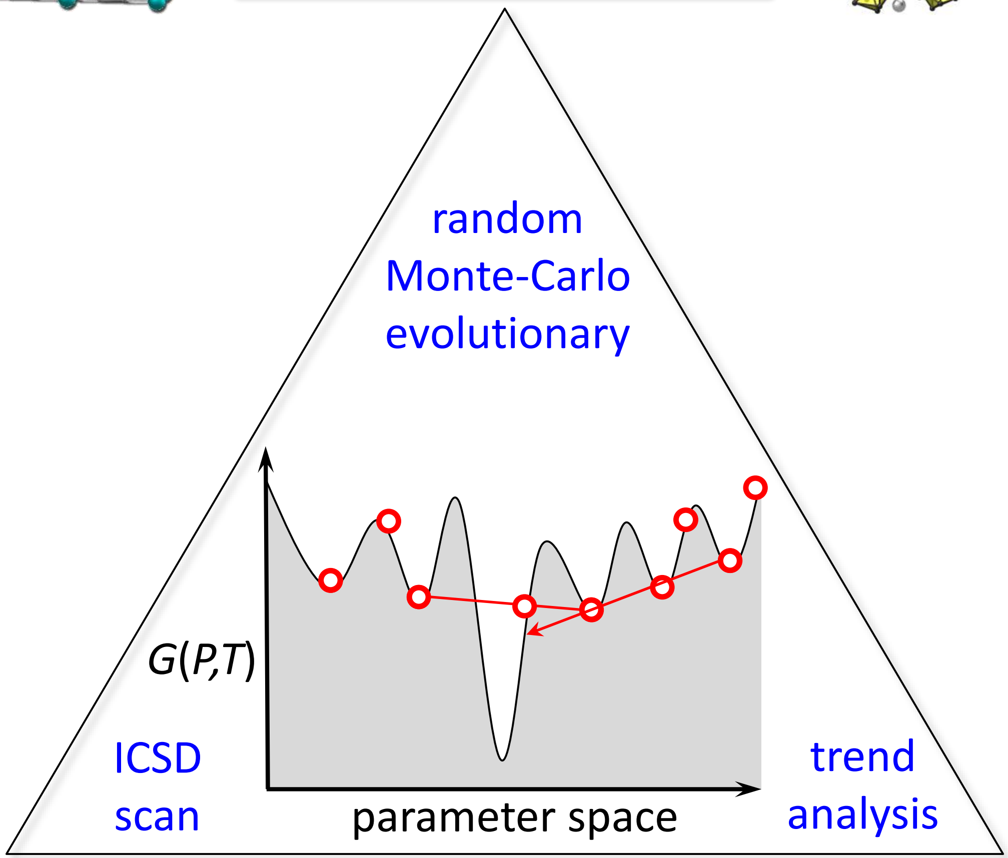
Kolmogorov *et al*. PRL 109, 075501 (2012)

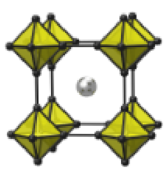
Prediction strategies



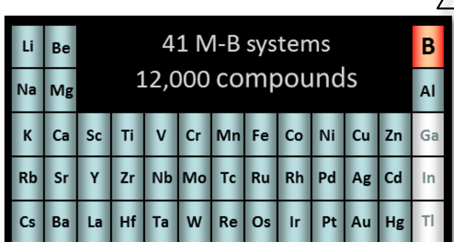
Unconstrained search

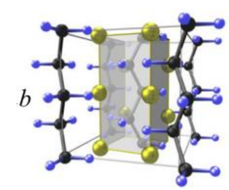




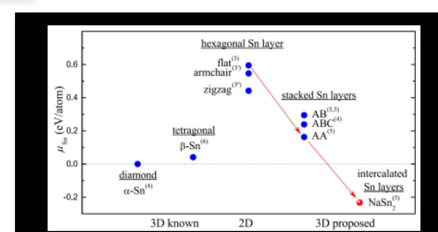


Constrained search





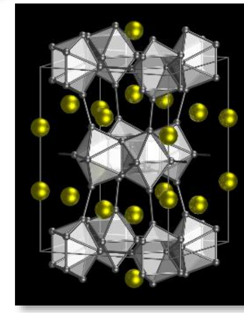
'Rational' design



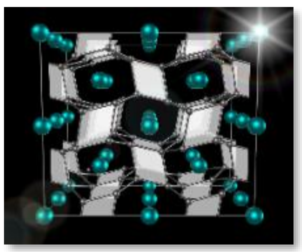
Group's standing and confirmed predictions

predictions
confirmations

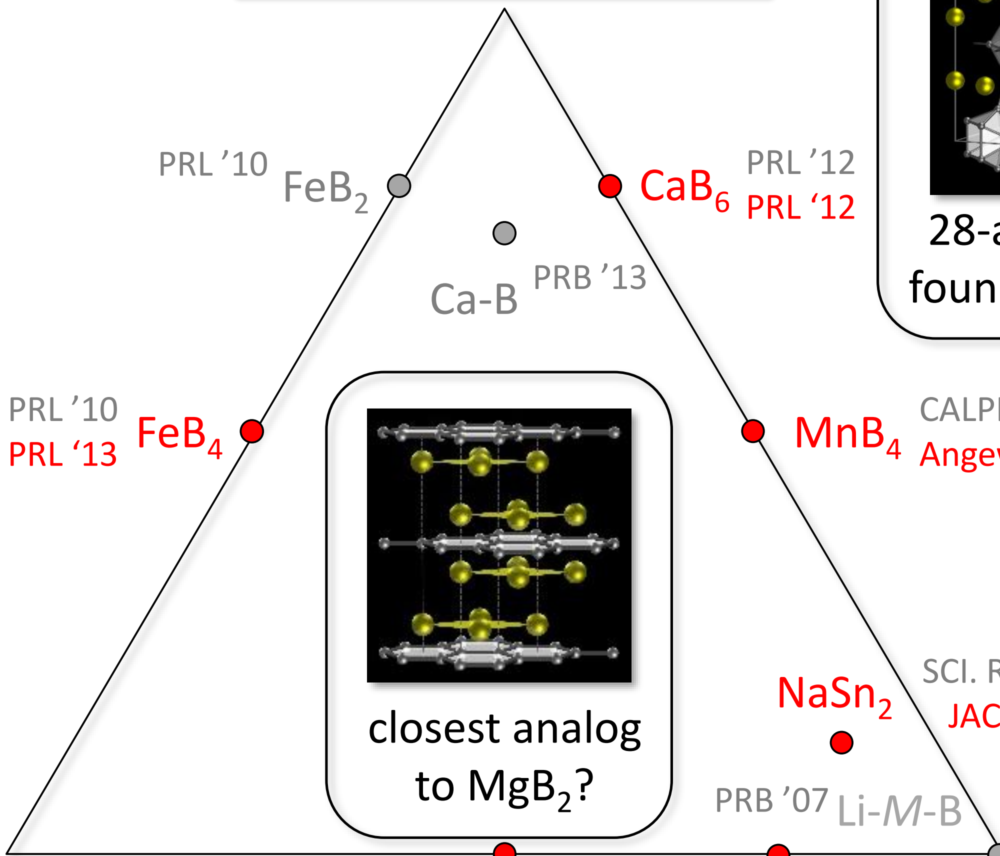
Unconstrained search



28-atom u.c. found *ab initio*



confirmed BCS superconductor

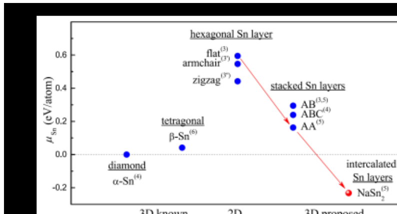


ICSD screening

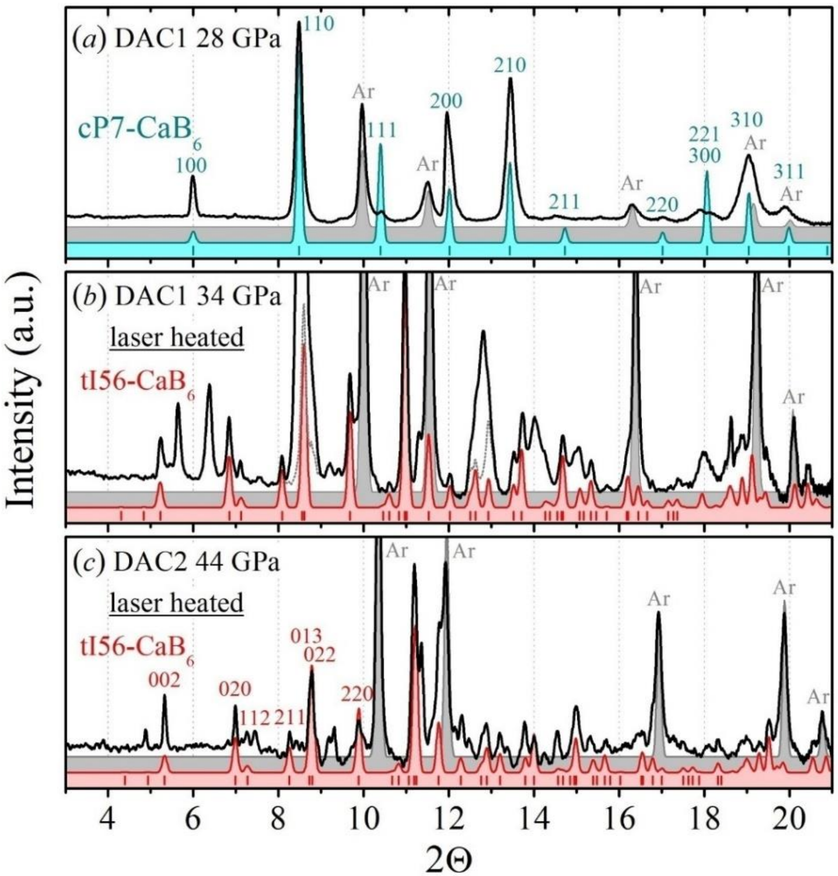
41 M-B systems
12,000 compounds

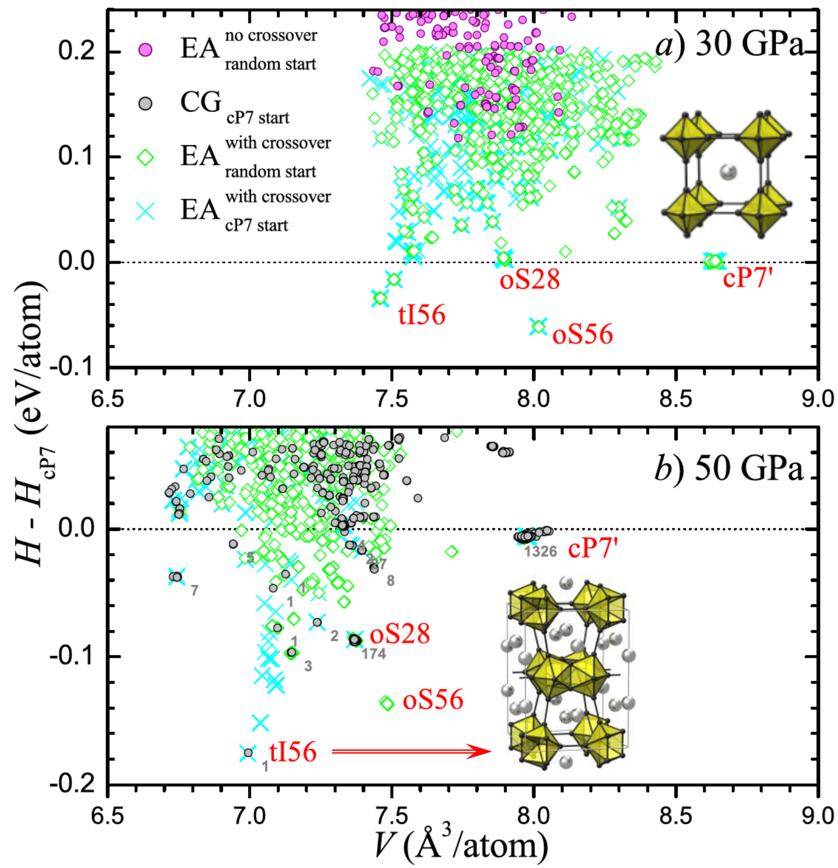
Li	Be											B
Na	Mg											Al
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In
Cs	Ba	La	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl

'Rational' design



Benchmarking optimization methods on 28-atom CaB_6



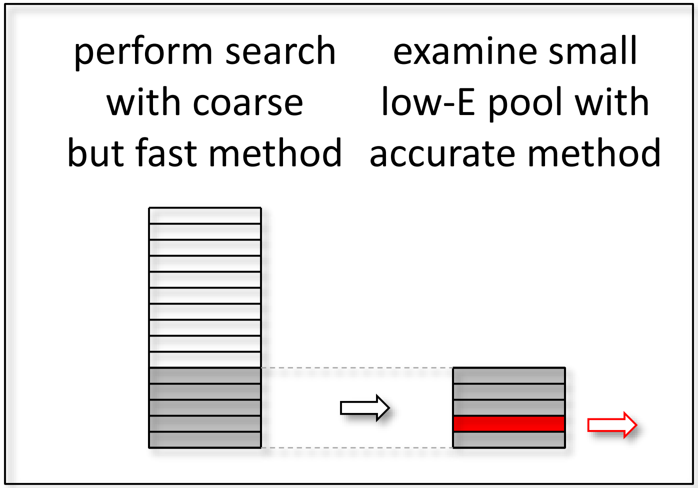


		From scratch	Supercell initialization	Known lattice parameters
28 CaB_6 [1]	With crossover (evolutionary)	~10,000 runs	~ 500 runs	~ 200 runs
	Without crossover (Monte-Carlo)	failed		~ 200 runs
	Conjugate gradient (deterministic)		~ 10,000 runs	
28 γ-B [2]	Conjugate gradient (deterministic)			~ 200 runs
44 Li [3]	Conjugate gradient (deterministic)			~ 200 runs

[1] Kolmogorov et al., PRL 105, 217003 (2012) [2] Oganov et al., Nature 457, 863 (2009) [3] Marques et al., PRL 106, 095502 (2011)

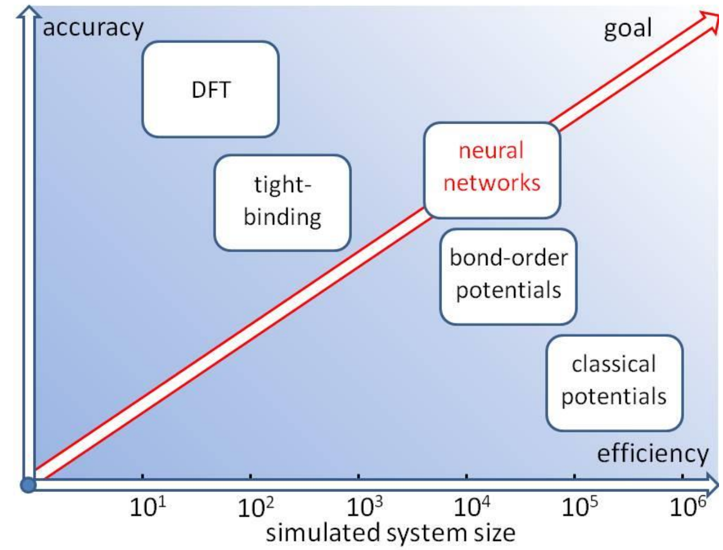
Prediction acceleration with classical potentials

Basic acceleration strategy

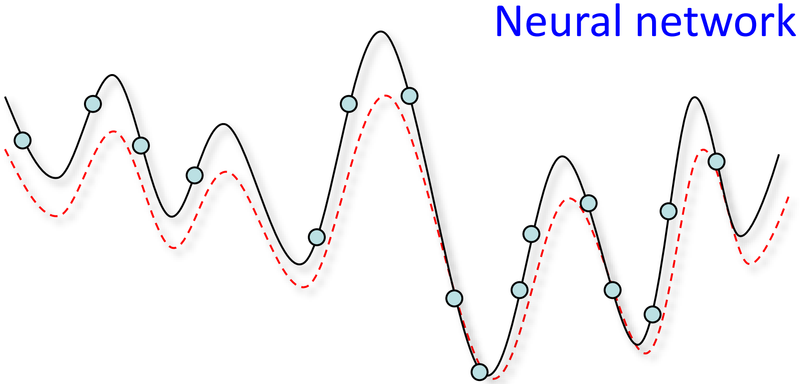


Farrow, Chow, Woodley, PCCP 16, 21119 (2014)
Wu, *et al.*, K.M. Ho: JPCM 26, 035402 (2014)

Interaction description methods



Neural network interpolator of DFT data



generate DFT database	50,000 CPU h / system
train neural network	1,000 CPU h / system
expected/observed speed-up	$\times\ 10^3$-10^4

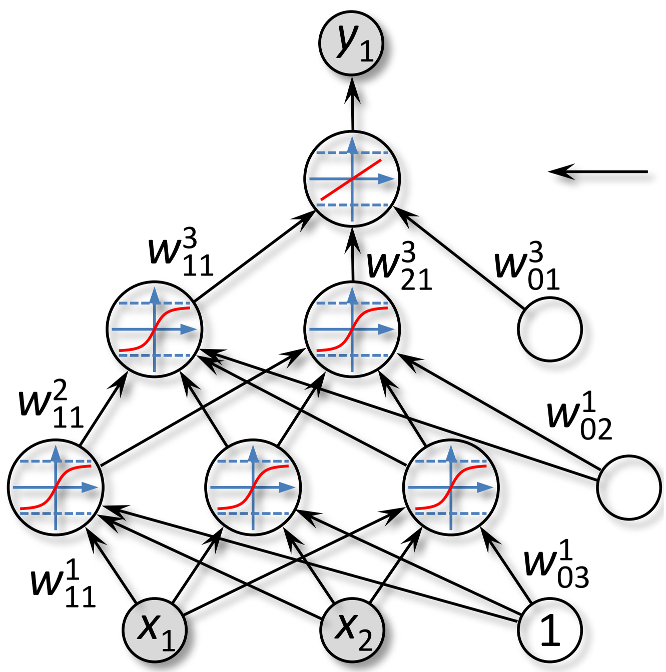
DFT reliability for predicting ground states (PBE/LDA, T = 0 K)

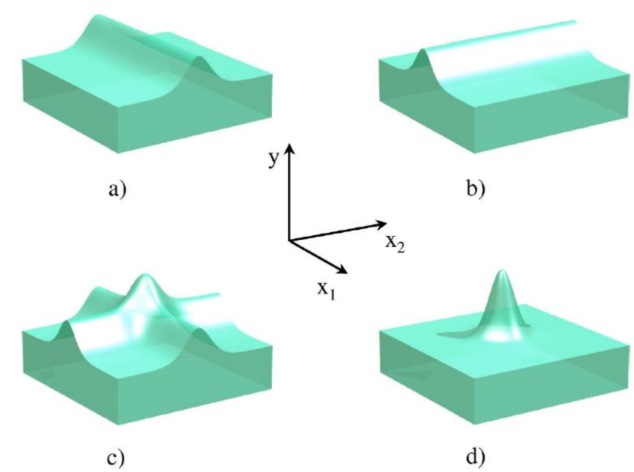
metal-metal compound stability	97% agreement with experiment, Curtarolo 2005
metal-boron compound stability	83% agreement with experiment, Kolmogorov 2014

Neural networks: multidimensional approximation functions

example of a peak approximation with a multi-layer neural network

the position, shape, and height controlled by weights





possible to represent a continuous multivariable function
by superposition of functions of one variable

A. N. Kolmogorov, Dokladi Akademii Nauk USSR, **114**, 953 (1957).

Description of interatomic interactions with Neural Networks

E, $\boldsymbol{F}_i$, $\boldsymbol{s}_{\alpha\beta}$

$\Uparrow$

output

weights

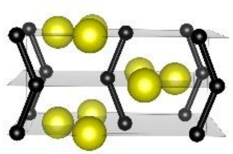
input

descriptor

$\Uparrow$

$\{ \boldsymbol{R}_i \} = \{ \boldsymbol{R}_{ij} , \boldsymbol{\theta}_{ijk} \}$

structure



development focus

design & train NN for multi-element systems

test available functions

generate diverse datasets

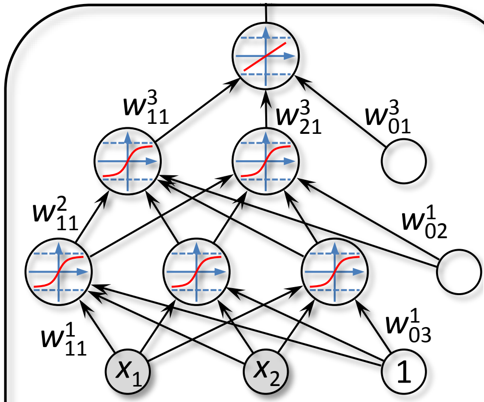
Description of interatomic interactions with Neural Networks

E, $\boldsymbol{F}_i$, $s_{\alpha\beta}$

output

weights

input

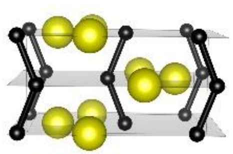


descriptor

symmetry functions

$\{ \boldsymbol{R}_i \} = \{ \boldsymbol{R}_{ij} , \boldsymbol{\theta}_{ijk} \}$

structure



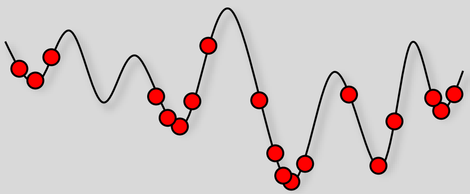
development focus

design & train NN for multi-element systems

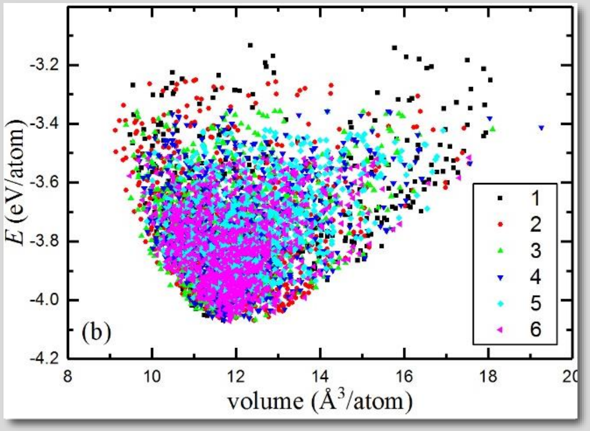
test available functions

generate diverse datasets

data generation protocols



traditional MD-based vs proposed evolutionary



sampling of relevant space

automated generation

built-in data diversification

Description of interatomic interactions with Neural Networks

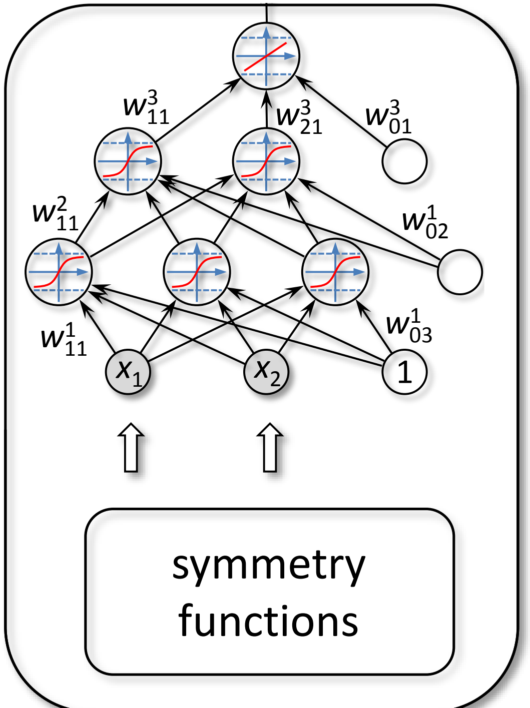
E, $\boldsymbol{F}_i$, $\boldsymbol{s}_{\alpha\beta}$

output

weights

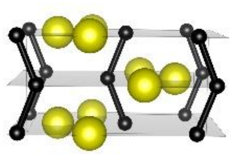
input

descriptor



$\{ \boldsymbol{R}_i \} = \{ \boldsymbol{R}_{ij} , \boldsymbol{\theta}_{ijk} \}$

structure



development focus

design & train NN for multi-element systems

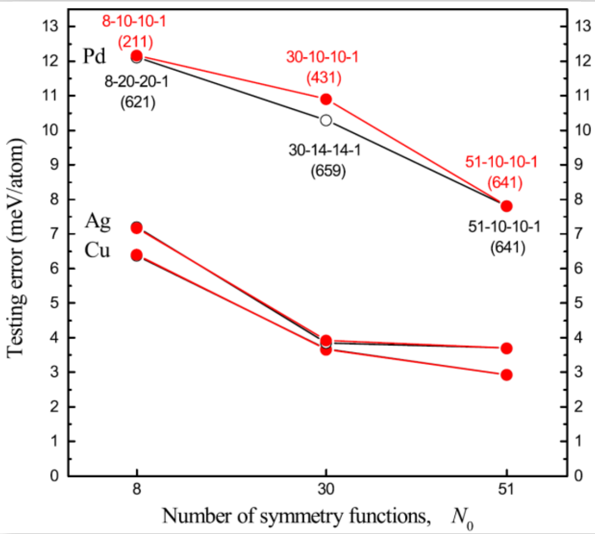
test available functions

generate diverse datasets

descriptor choice

Bartok et al. PRB 87, 184115 (2013)
Behler Parrinello PRL 98, 146401 (2007)

test & use of PB descriptor



good for metallic bonding
up to 3-body terms
30 or 51-function sets

Description of interatomic interactions with Neural Networks

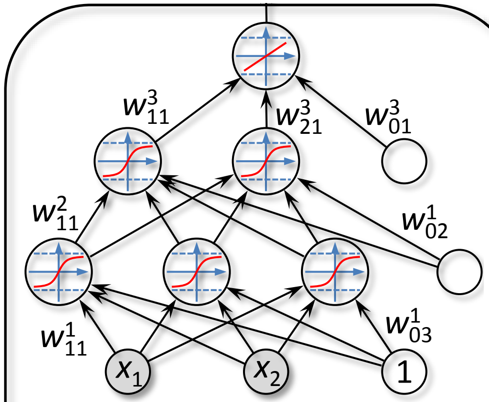
E, $\boldsymbol{F}_i$, $\boldsymbol{s}_{\alpha\beta}$

⇧

output

weights

input



⇧ ⇧

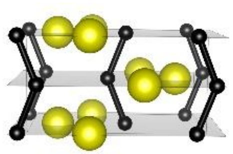
descriptor

symmetry functions

⇧

$\{ \boldsymbol{R}_i \} = \{ \boldsymbol{R}_{ij} , \boldsymbol{\theta}_{ijk} \}$

structure



development focus

design & train NN for multi-element systems

test available functions

generate diverse datasets

multi-element NNs

Artrith, Hiller, Behler, CuZnO PSS B 250 1191 (2013)

general problem

Pd data	CuPd data	Cu data	CuAg data	Ag data

⇩ ⇩ ⇩

NN_{CuPd} NN_{Cu} NN_{CuAg}

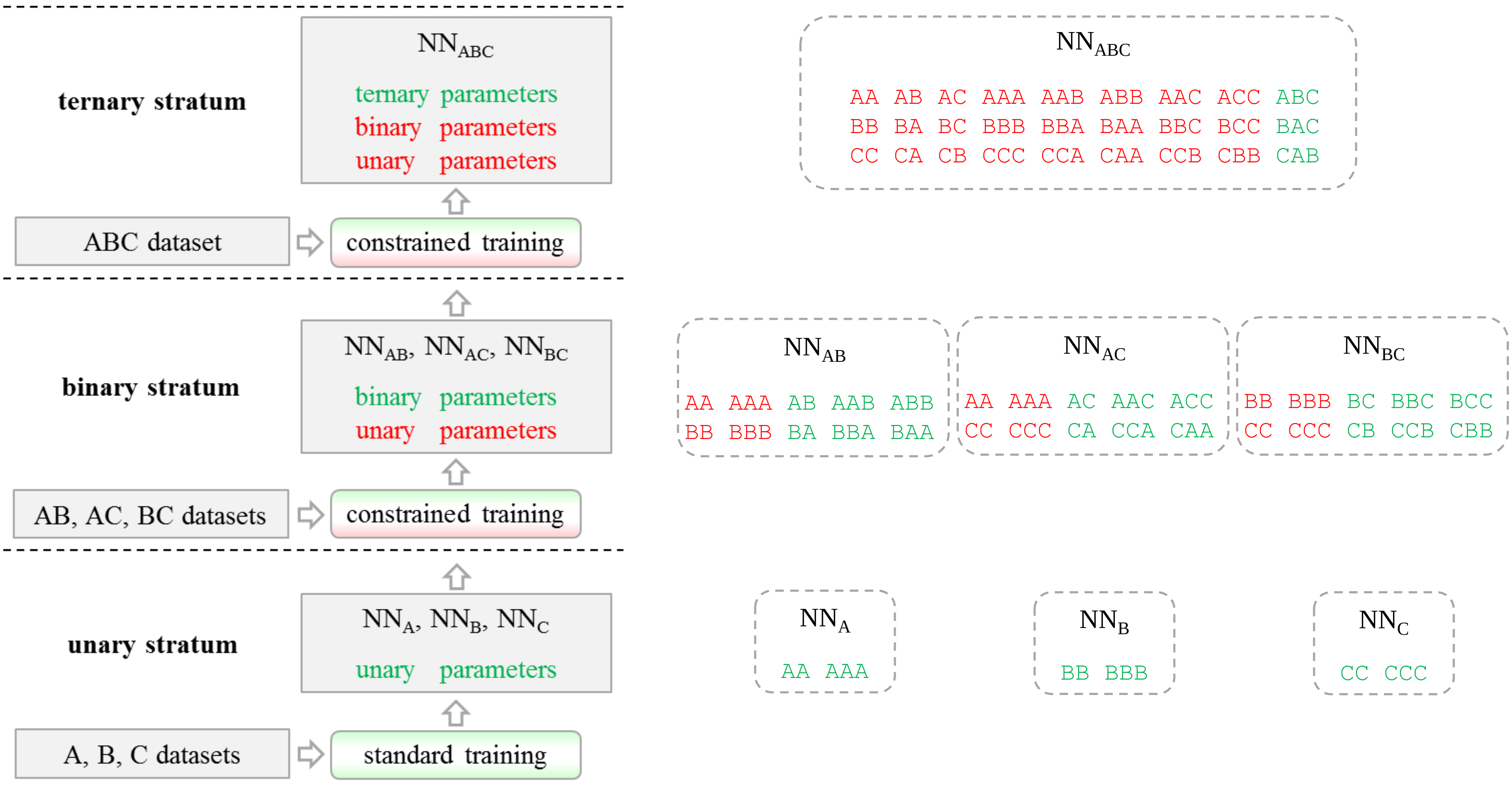
⇩ ⇩ ⇩

E(Cu) $\neq$ E(Cu) $\neq$ E(Cu)

proposed procedure

hierarchical (stratified) training

Stratified construction of multi-element NNs



Training stratification: key points

has been used previously for classical and tight-binding models

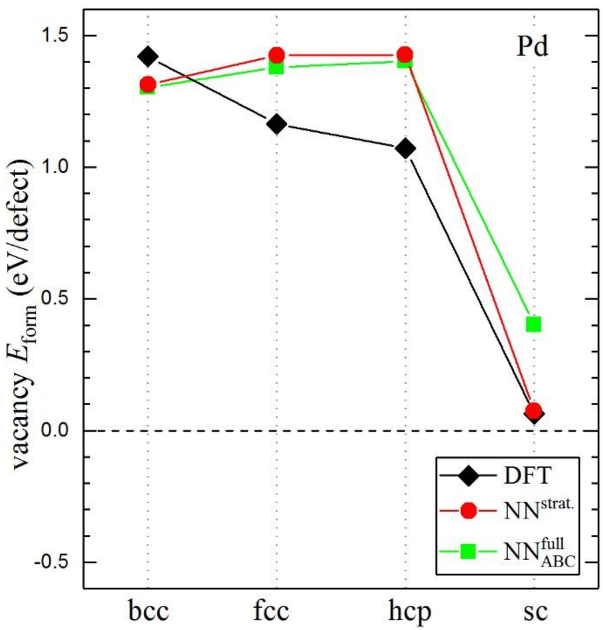
offers training speed-up and standardization of NN models

could lead to drop in accuracy due to constraints?

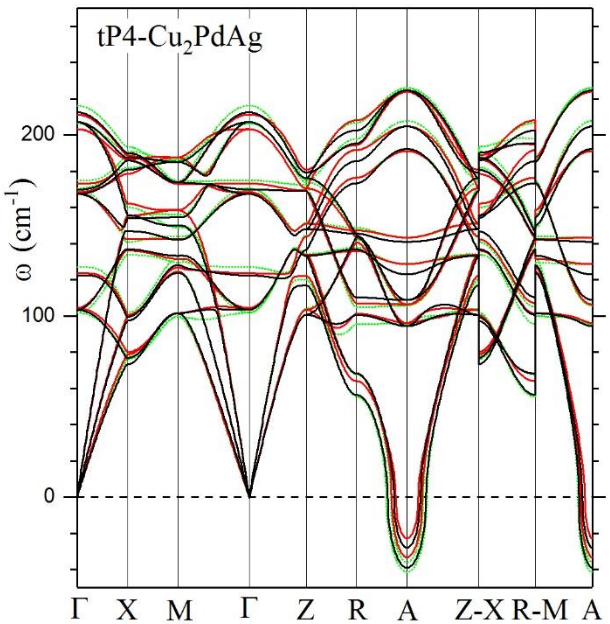
NN accuracy for Cu-Pd-Ag: DFT vs NN$^{strat.}$ vs NN^{full}

DFT:	PAW-PBE, E_{cut} = 500 eV	57,000 energy data
NNs:	30/82/145-10-10 MLP	5,073 parameters

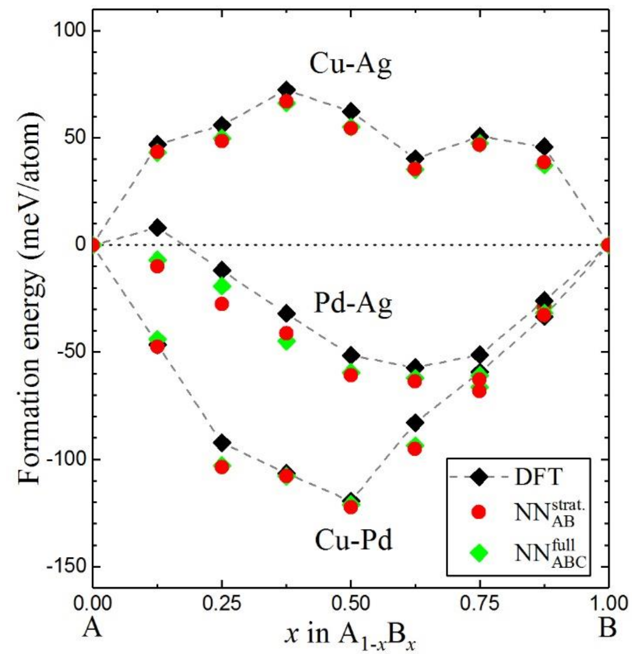
vacancy formation



phonon dispersion



formation energy



NNs vs DFT

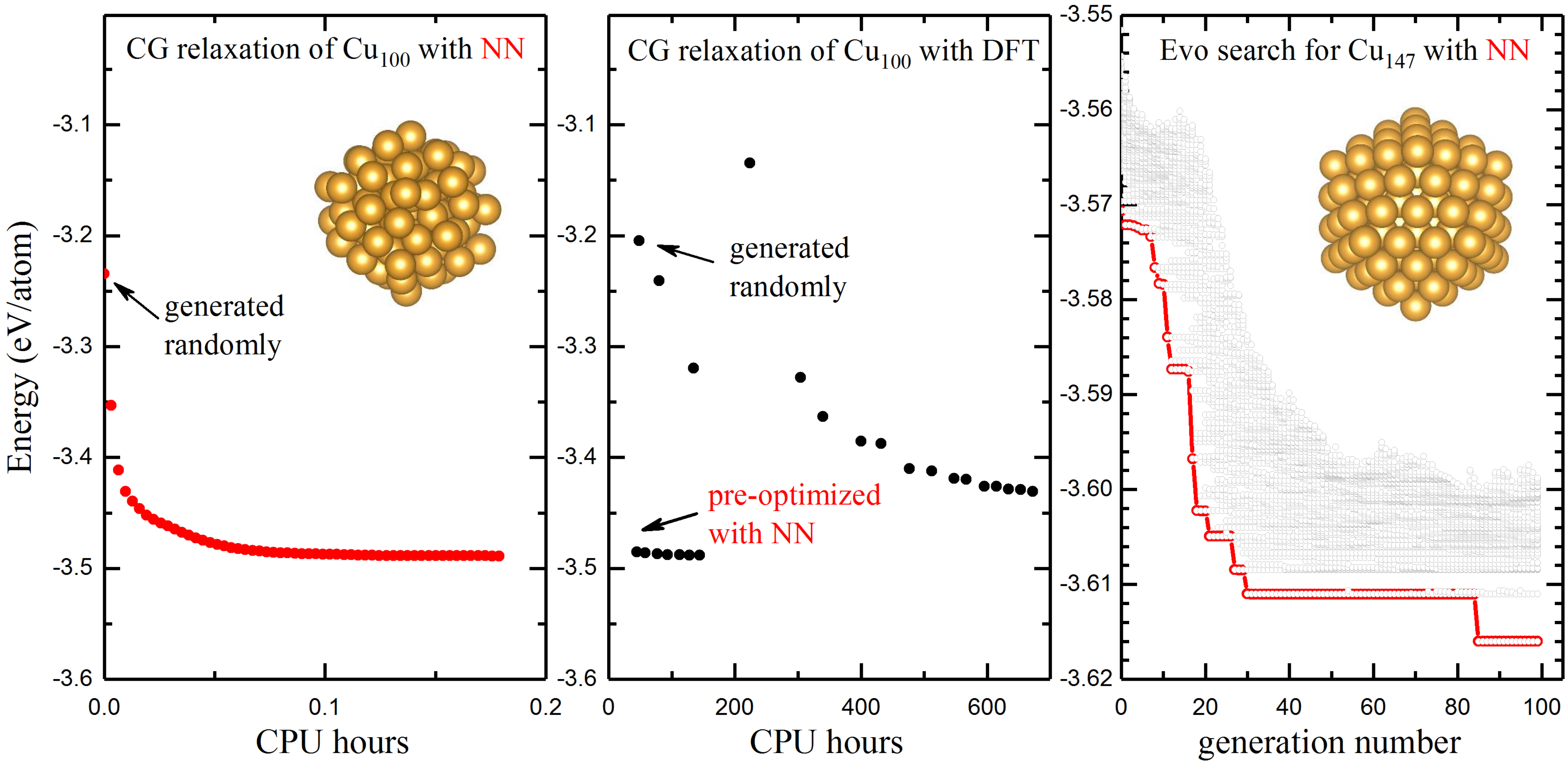
within ~0.3 eV/defect	captures softening	within 10 meV/atom

stratified vs full NN training:

no decrease in accuracy due to constraints

S. Hajinazar, J. Shao, and A.N. Kolmogorov, PRB 95, 014114 (2017)

NN benchmarking for nanoparticles



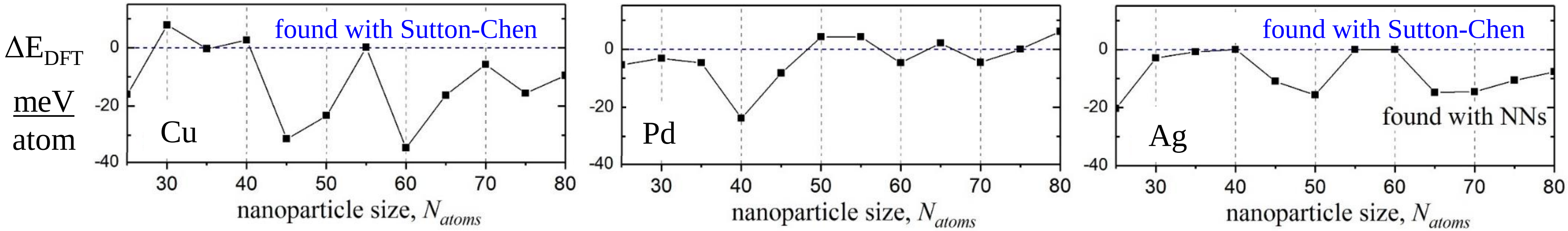
Cu_{100} relaxation with NN takes 10 mins on a single core

NPs pre-optimized with NN are a better start for DFT optimization

Evolutionary search with NNs converges to correct g.s.

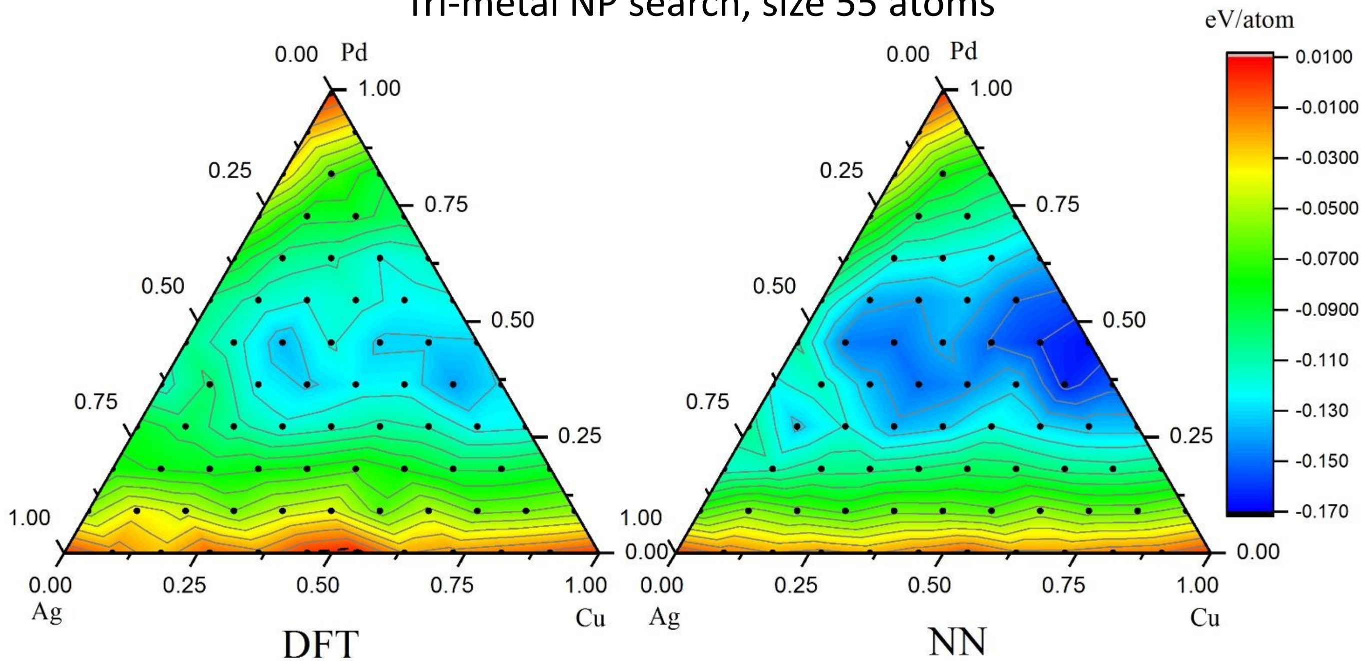
NN benchmarking for nanoparticles (NPs)

Single-element NP search, sizes 25-80 atoms
compared to Doye & Wales, New J. Chem. 733 (1998)



more stable configurations found, confirmed with DFT

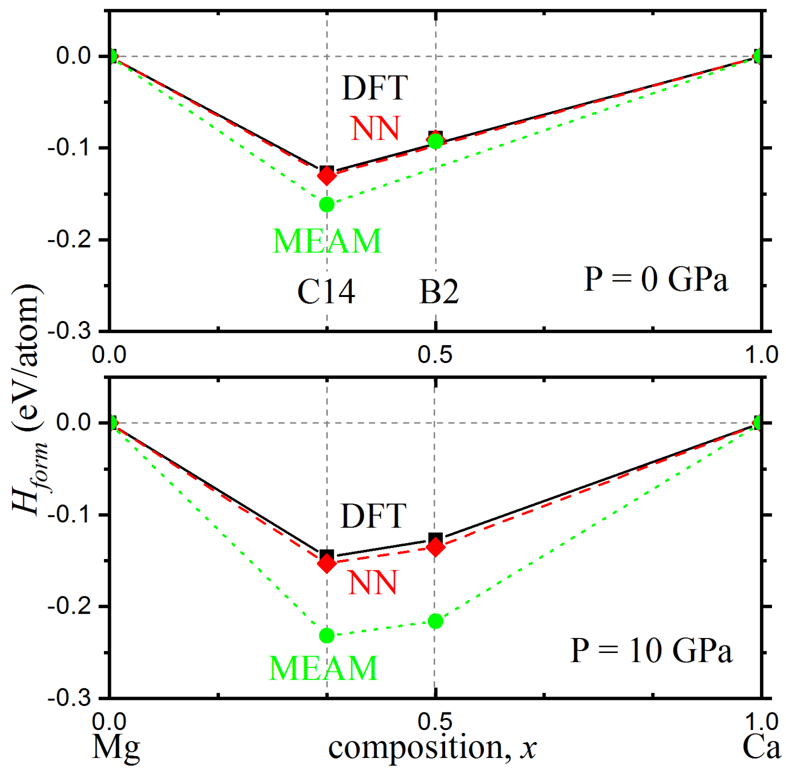
Tri-metal NP search, size 55 atoms



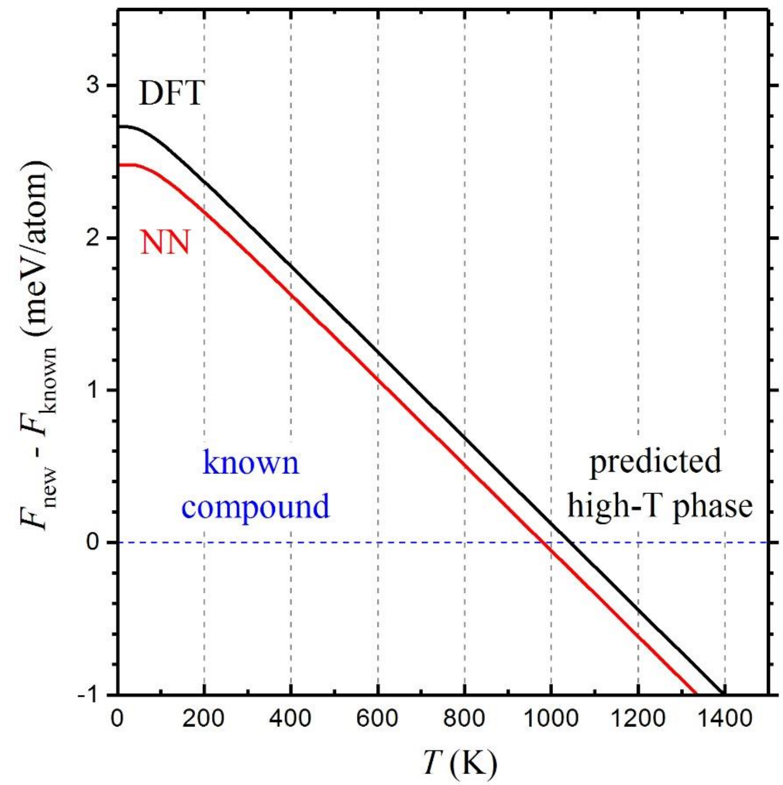
full composition range explored with ES+NN
stability islands agree in DFT and NN treatments

First predictions accelerated with NNs

NN vs DFT and MEAM
for formation enthalpy H(P) = E + PV



NN vs DFT
for free energy at high T $F(T) = E - TS_{phon}$



NN reproduces convex hull within 10 meV/atom

MEAM overbinds compounds by 100 meV/atom

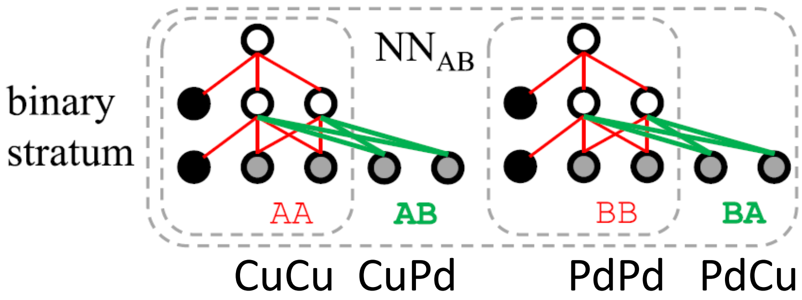
~10^4-fold acceleration of high-T analysis

new high-T Mg-Ca phases predicted

Interpretation of NN parameters?

architecture of a

simplified 16-10-1 NN



are there correlations between

CuCu and CuPd weights

connecting to same neurons?

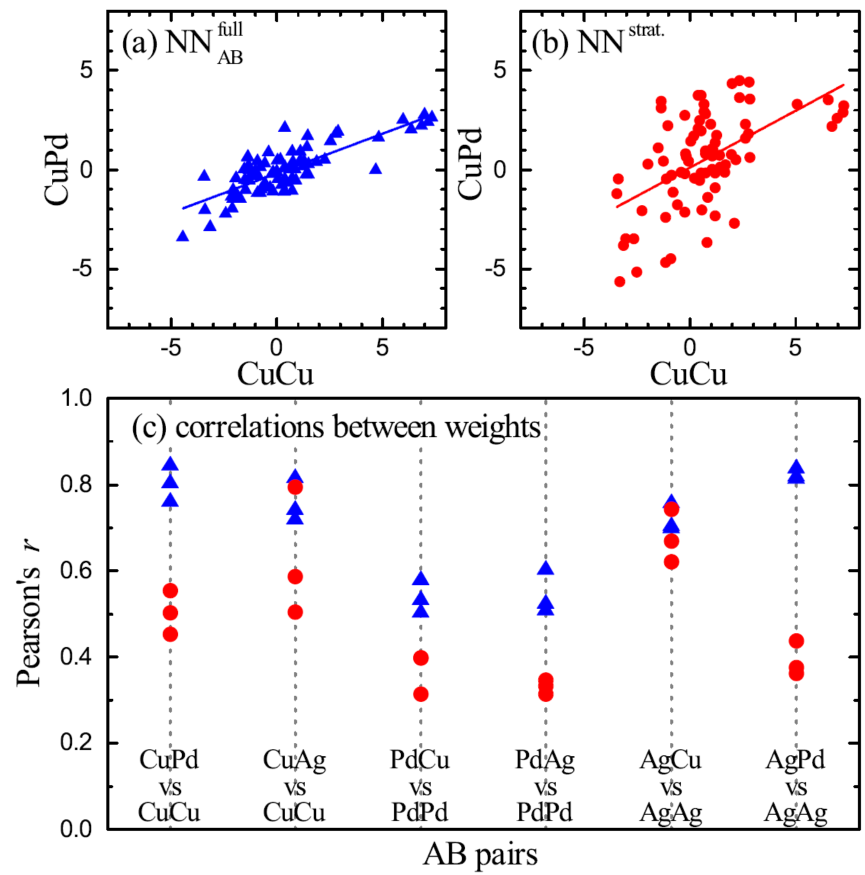
more correlation for 'full' NN

still little insight into 'inner workings'

weight correlations

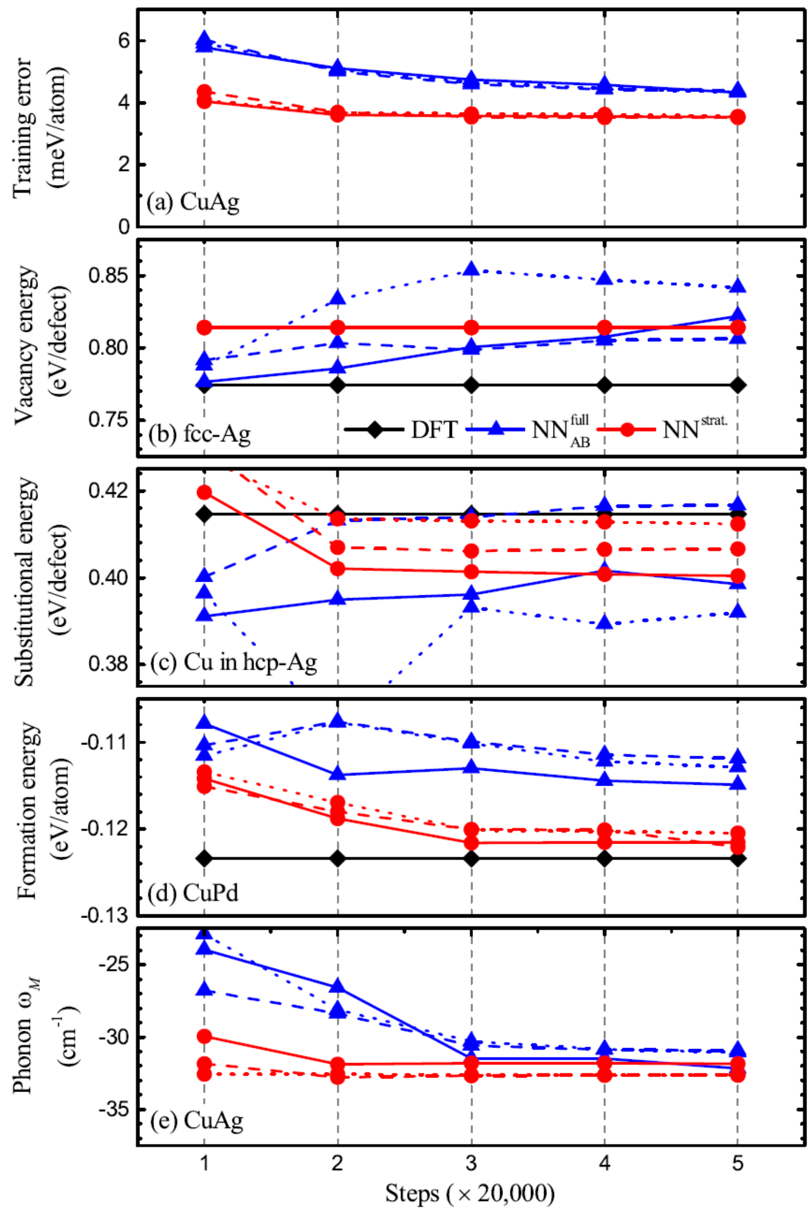
'full' NN

stratified NN



NN optimization details

Data set	Stratified NN ΔE eV/atom	Stratified NN No. of data	Stratified NN No. of weights	Full NN ΔE eV/atom	Full NN No. of data	Full NN No. of weights
Cu	0.5172	8551	431	–	–	431
Pd	0.6262	8487	431	–	–	431
Ag	0.4481	8533	431	–	–	431
CuPd	0.5975	7623	1040	0.5829	24661	1902
CuAg	0.5391	7617	1040	0.6601	24701	1902
PdAg	0.5989	7601	1040	0.6833	24621	1902
CuPdAg	0.2170	8917	660	0.6227	57329	5073



total energy as target only; forces/stresses possible

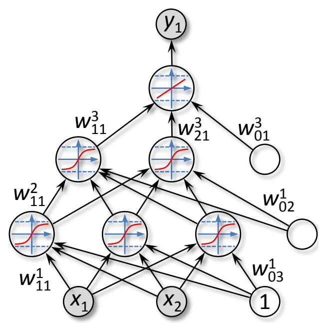
100,000 optimization steps with BFGS/CG

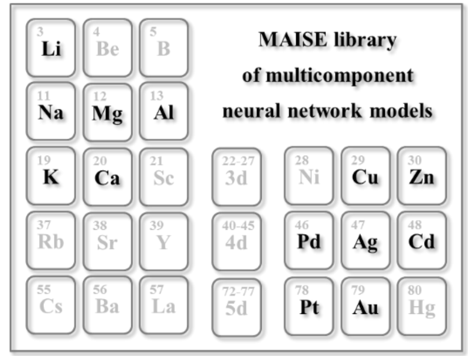
performance insensitivity to weight initialization

Summary

Stratified construction of neural network based interatomic models for multicomponent materials

S. Hajinazar, J. Shao, and A.N. Kolmogorov, PRB 95, 014114 (2017)





Method development	evolutionary data generation stratified construction of models	sampling of relevant space no loss of accuracy
NN model development	metallic: 13 unaries, 7 binaries, 2 ternaries covalent: in progress	3-10 meV/atom noticeably lower accuracy
NN model application	robust for evolutionary optimization accurate for high-T stability analysis	new low-energy NPs predicted new high-T phases predicted
MAISE development	evolutionary optimization, MD to be released as open source	crystals surfaces nanoparticles collaborations welcome
Acknowledgments	S. Hajinazar, E. Sandoval (BU) J. Shao (BU), A. Romero (WVU)	A Jephcoat (Okayama U.) N & L Dubrovinskie (Bayreuth)
Support	TAE (Binghamton)	NSF-DMR