RUB

Representing High-Dimensional Potential-Energy Surfaces by Artificial Neural Networks

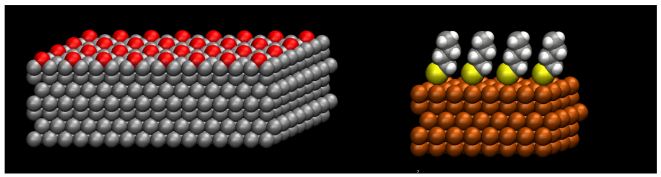
Jörg Behler

Lehrstuhl für Theoretische Chemie
Ruhr-Universität Bochum
D-44780 Bochum, Germany

joerg.behler@theochem.ruhr-uni-bochum.de

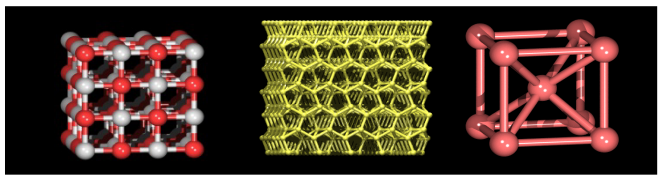
Surface Science:

- heterogeneous catalysis
- corrosion
- self-assembled monolayers



Materials Science:

- crystal structure prediction
- phase diagrams
- properties of materials



Challenges:

- complicated bonding situations
- unusual coordination
- different types of bonding: covalent, metallic, vdW etc.
- large systems
- long simulation times

⇒ Reliable potentials for large-scale simulations are needed!

The reliability of the results obtained in theoretical simulations depends on the accuracy of the underlying potentials.

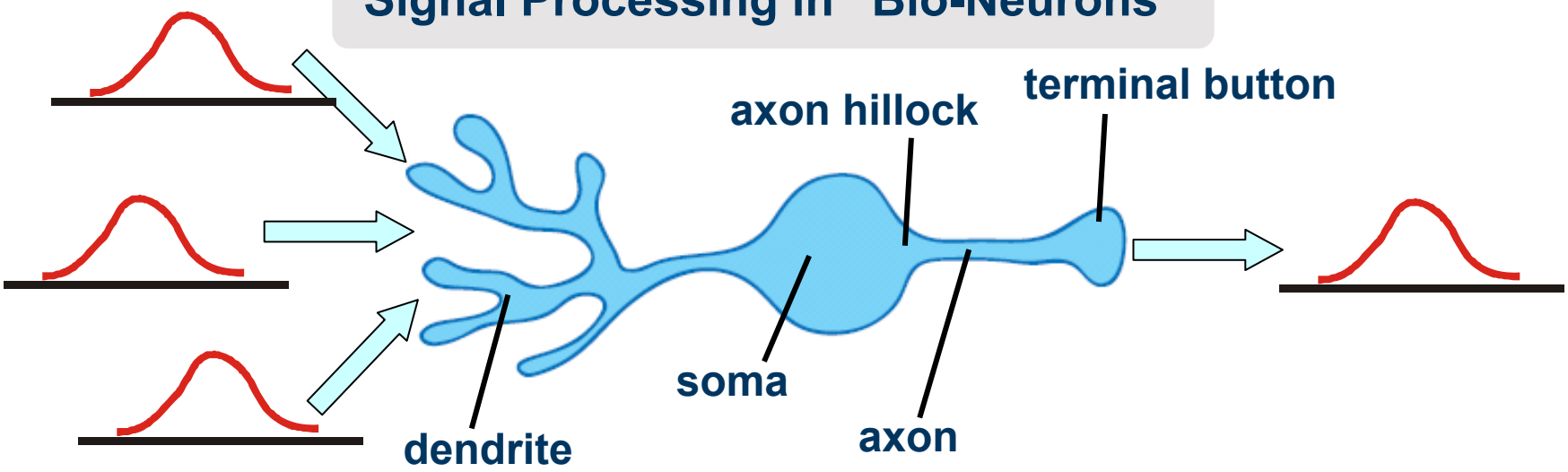
	atoms	time
ab initio (MP2, CC, CI)	≈ 10	0
Hartree Fock / density functional theory	≈ 100	100 ps
semiempirical / tight binding	≈ 1000	1 ns
empirical potentials	≈ 10,000	10 ns
classical force fields	≈ 100,000	100 ns

RUB

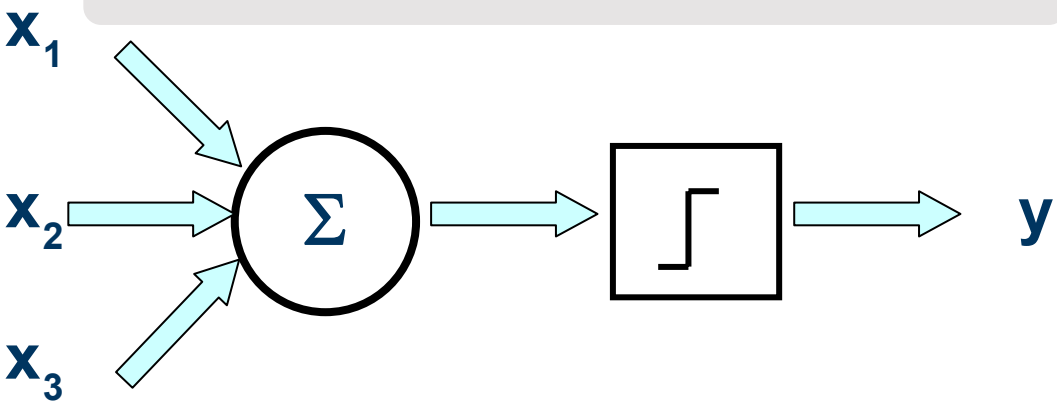
“List of wishes”

- should be very accurate
- should describe making and breaking of bonds
- should be applicable to all types of systems (metals and insulators, solids and molecules)
- should be fast to evaluate
- systematic improvement should be possible
- no “manual” work for construction
- should be transferable (“predictive”)

Signal Processing in "Bio-Neurons"

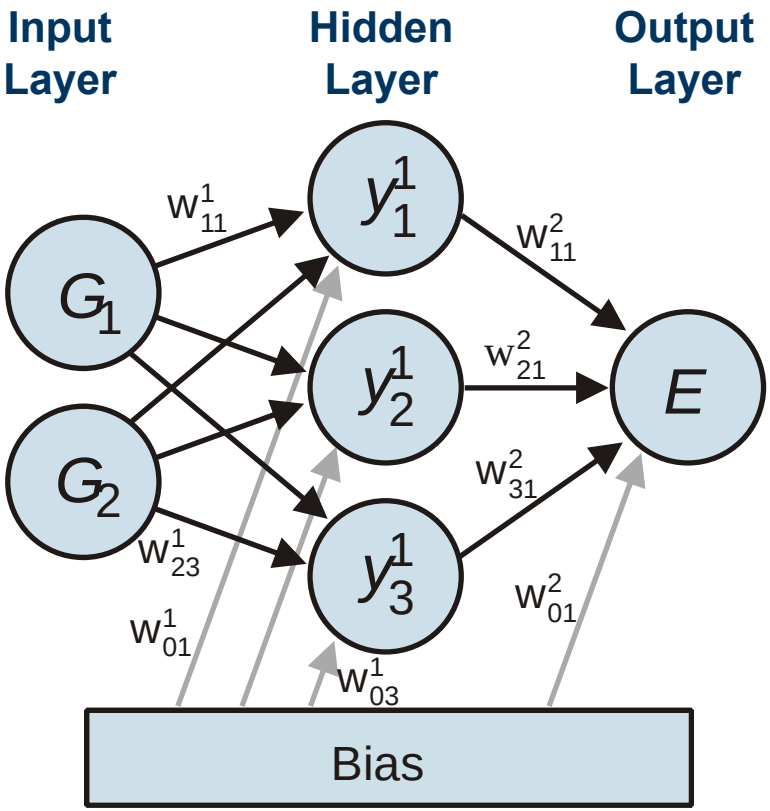


Artificial Neuron



W. McCulloch, and W. Pitts, *Bull. Math. Biophys.* 5, 115 (1943).

Example:



Analytic Total Energy Expression:

$$E = f_a^2\left(w_{01}^2 + \sum_{j=1}^{3} w_{j1}^2 \cdot f_a^1\left(w_{0j}^1 + \sum_{i=1}^{2} w_{ij}^1 G_i \right) \right)$$

Properties of a NN Potential:

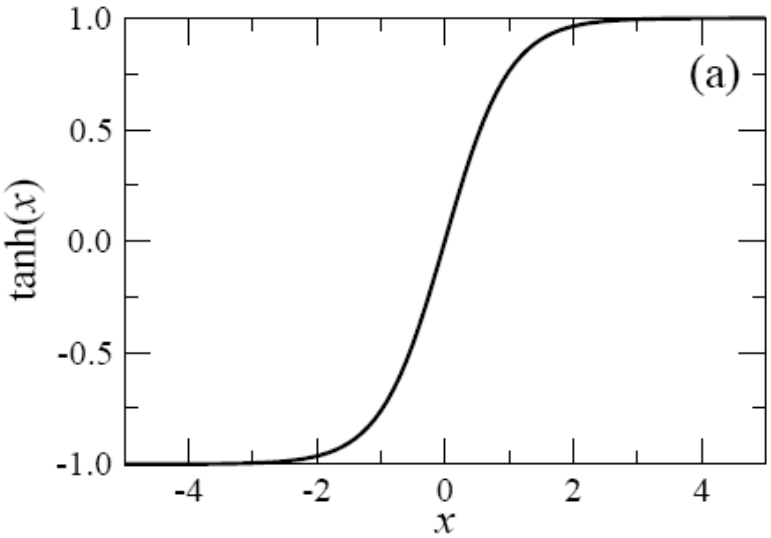
- very flexible functional form
- very accurate
- no knowledge of functional form required
- properties of the system can be included via mapping on symmetry functions
- bad for extrapolation
- many training points needed
- analytic forces

T.B. Blank, S.D. Brown, A.W. Calhoun, and D.J. Doren, *J. Chem. Phys.* 103 (1995) 4129.
S. Lorenz, A. Groß, and M. Scheffler, *Chem. Phys. Lett.* 395 (2004) 210.
J. Behler, S. Lorenz, and K. Reuter, *J. Chem. Phys.* 127 (2007) 014705.

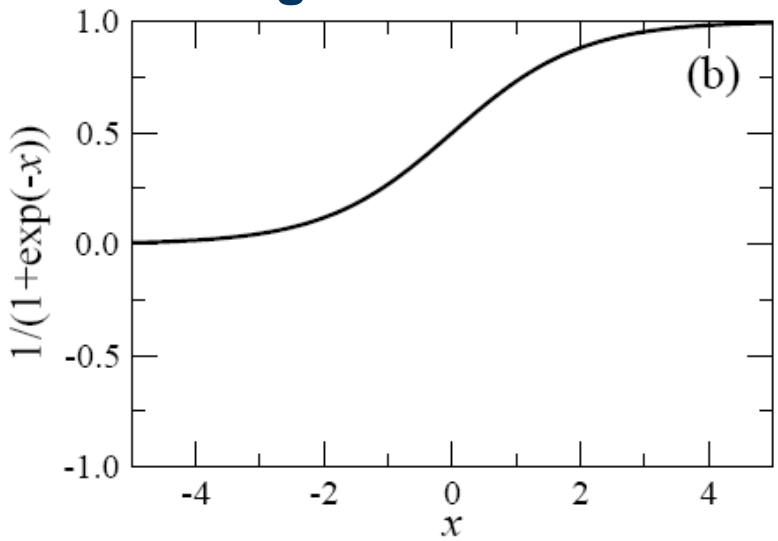
Activation functions enable the fitting of general nonlinear functions.

Examples:

hyperbolic tangent



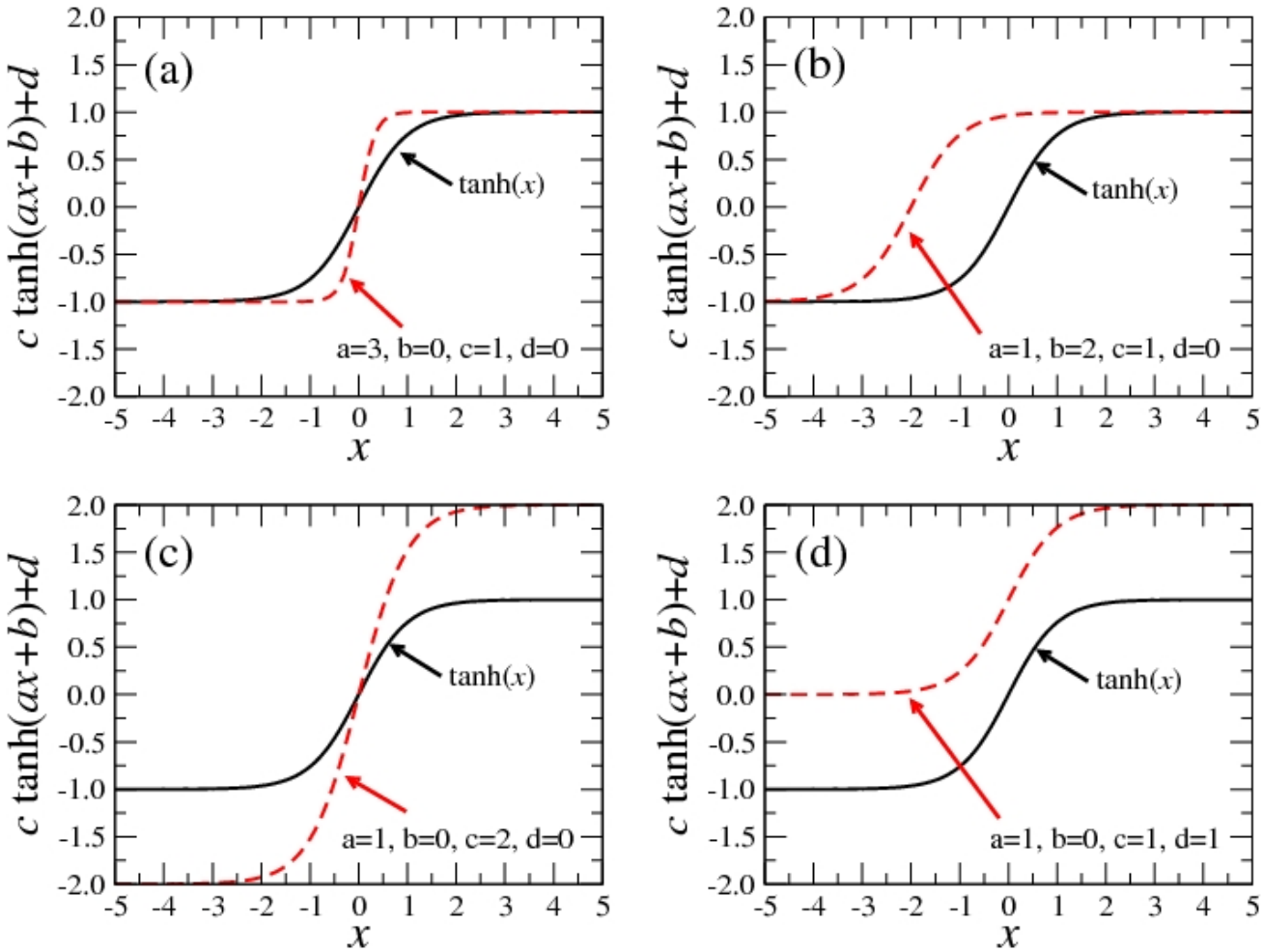
sigmoid function



Activation functions

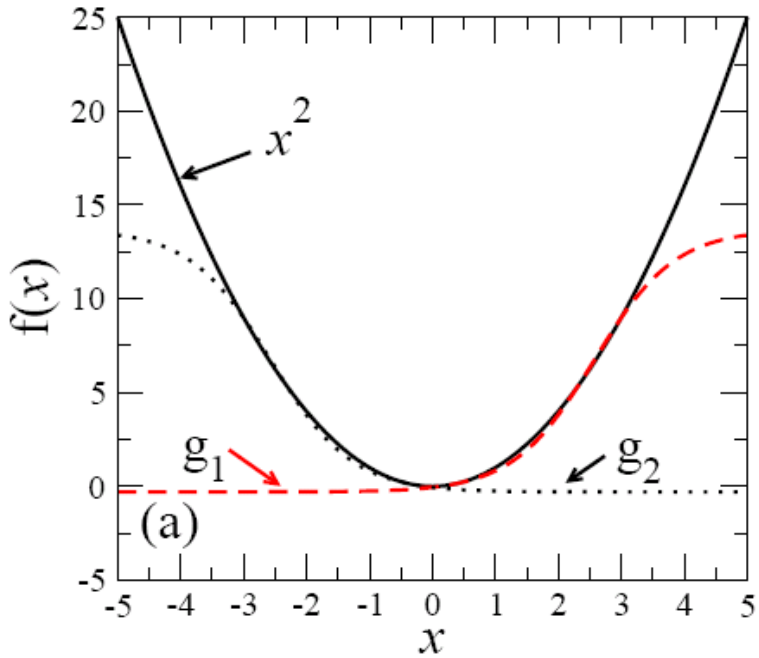
- converge for very small and very large arguments
- have a nonlinear shape for intermediate values

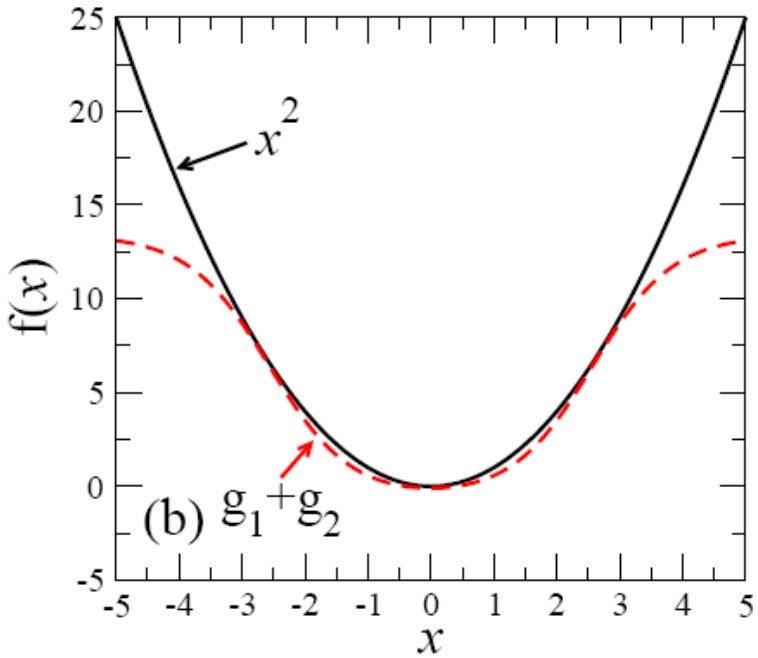
Basic functional element of the NN: $f(x) = c\tanh(ax+b)+d$



⇒ very flexible but simple

Fitting of the potential of a 1D harmonic oscillator for -3 < x < 3:

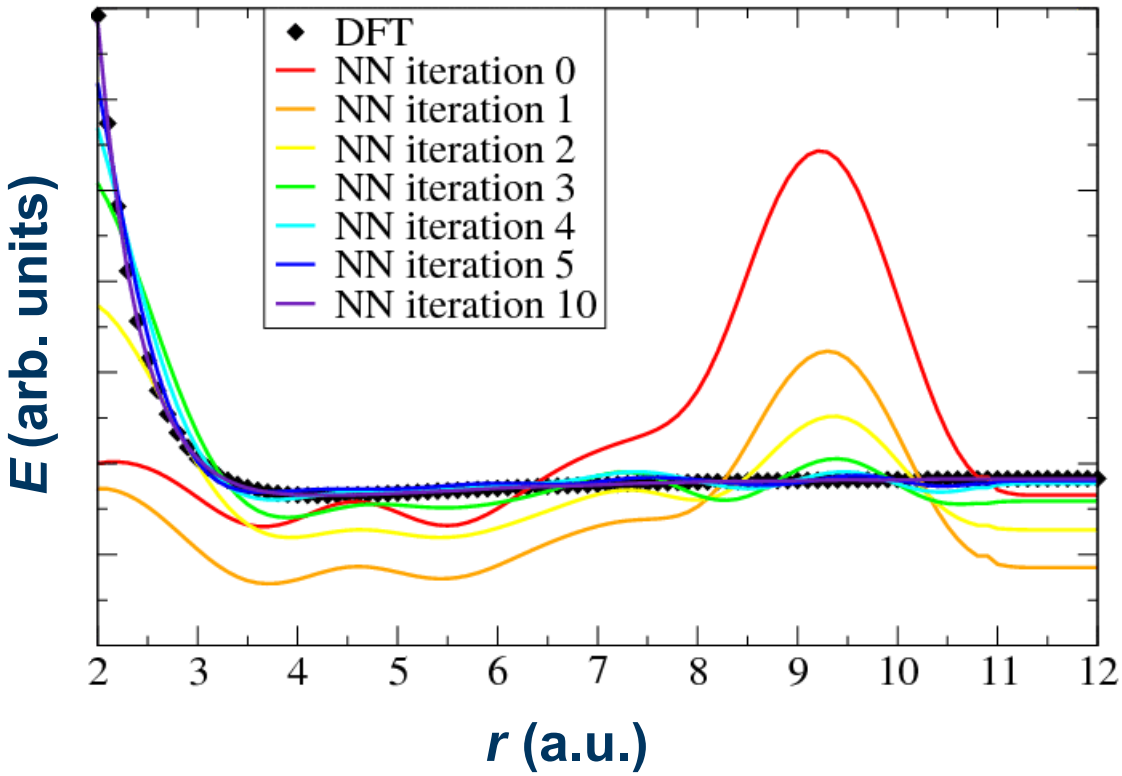




⇒ Just two functions provide a rather good approximation

⇒ Fit is further improved by adding more functions

More realistic example: Pair potential



⇒ NN has a non-physical functional form, but can learn physics

Two Challenges:

Find the relevant structures

- experiment
- chemical intuition
- molecular dynamics (e.g. Parrinello Rahman)
- genetic algorithms
- metadynamics
- …

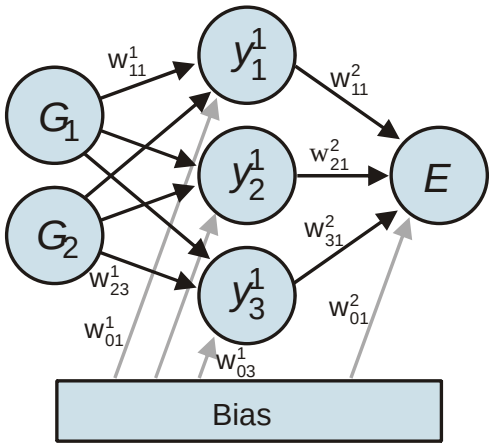
Determine the stability of many structures

- fast empirical potentials for first evaluation ⇒ often unreliable
- DFT ⇒ too expensive

⇒ We employ a Neural Network potential

RUB

Is it possible to use Neural Network potentials to construct high-dimensional potential energy surfaces?

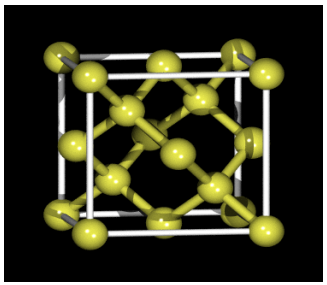
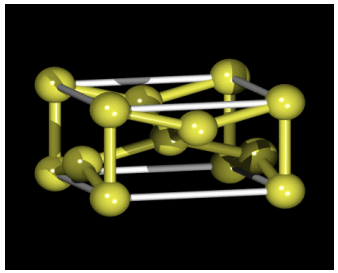
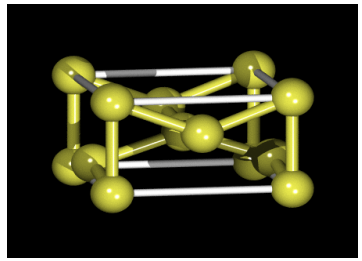
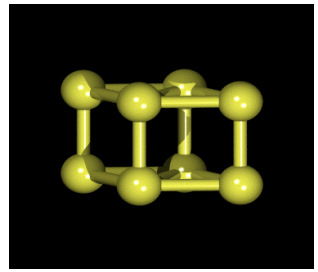


Conceptual problems for a direct application of Neural Networks to condensed systems

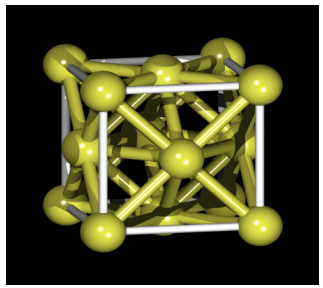
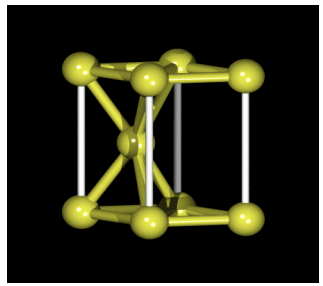
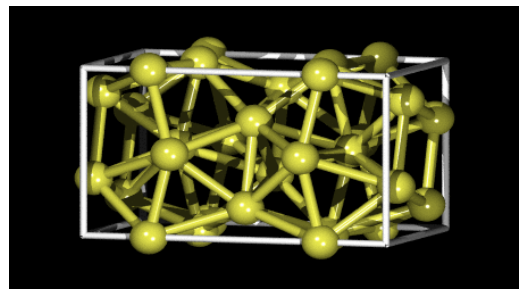
- limited number of dimensions (up to 12)
- symmetry of the system is not included (exchange of atoms changes the energy output)
- energy possibly depends on rotation and translation
- fit is valid only for a given system size

⇒ A new Neural Network scheme is required to deal with high-dimensional systems

The High-Pressure Crystal Structures of Silicon

cubic diamond —11.7 GPa→ β-tin —13.2 GPa→ Imma —15.4 GPa→ simple hexagonal —38 GPa→ Cmca —42 GPa→ hcp —79 GPa→ fcc

⇒ **Challenging model system to study pressure-induced phase transitions**

Review: A. Mujica, A. Rubio, A. Muñoz, and R.J. Needs, *Rev. Mod. Phys.* 75 (2003) 863.

Code: PWSCF

- plane waves
- pseudo potentials
- stress tensor for k-points

PWSCF:
S. Baroni, A. Dal Corso, S. de Gironcoli, P. Giannozzi, C. Cavazzoni, G. Ballabio, S. Scandolo, G. Chiarotti, P. Focher, A. Pasquarello, K. Laasonen, A. Trave, R. Car, N. Marzari, A. Kokalj, http://www.pwscf.org/.

Calculational Details:

- 20 Ry cutoff
- 3x3x3 k-points

System Details:

- 64 Si atoms
- Pressure range 0 – 100 GPa, Temperature 0 – 3000 K

Timing: 4-10 h on a 2.8 GHz Opteron (serial run)

A metadynamics run takes about 100 metasteps, and each step includes a MD run of about 2 ps.

⇒ 200 000 energy and force evaluations per metadynamics run

⇒ 1 600 000 CPU hours = 67 000 CPU days = 183 CPU years

⇒ We cannot use DFT directly!

We need *many* metadynamics runs!

⇒ We need a significant speed-up in the potential evaluation, but without a significant loss in accuracy

Can we combine metadynamics with neural networks?

Idea:

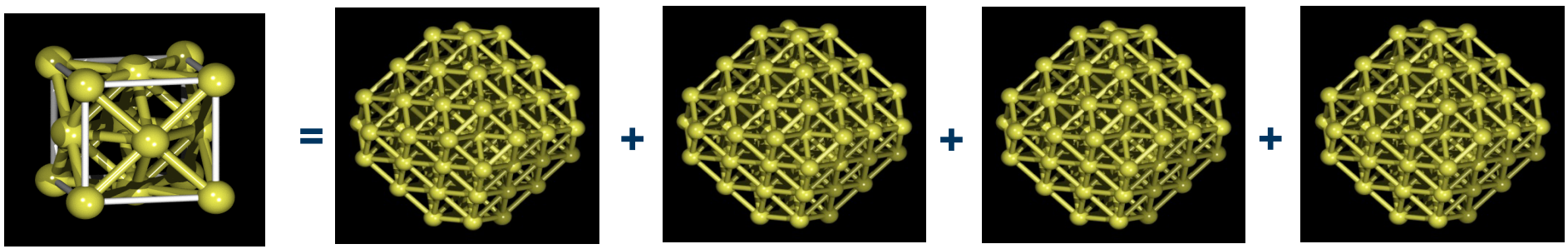
Represent the total energy as a sum of atomic energies:

$$E = \sum_i E_i$$

Each E_i is given by an individual NN as a function of the chemical environment

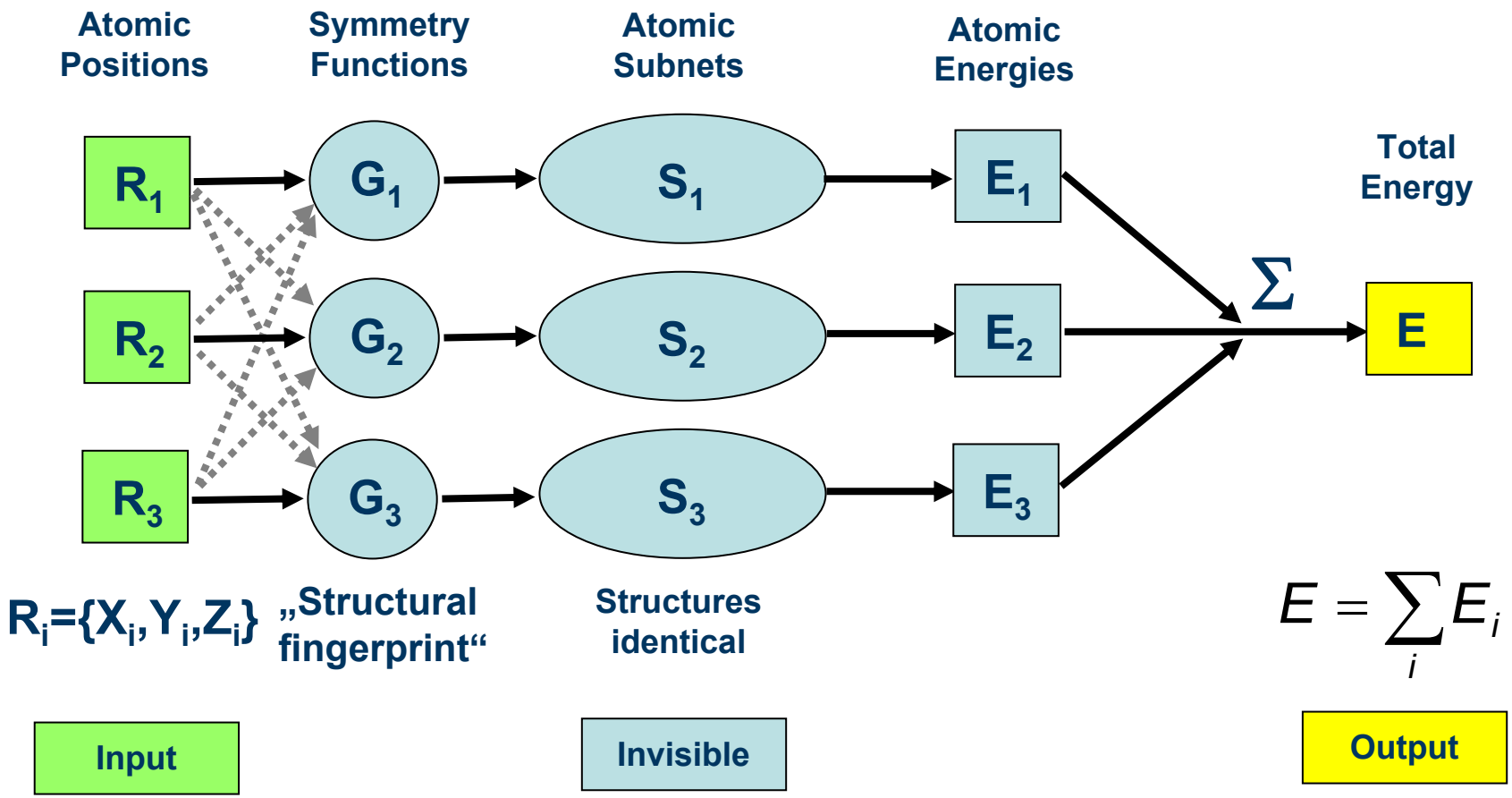
$\Rightarrow$ requires a new Neural Network architecture and suitable symmetry functions

Example: fcc structure (4 atom unit cell)



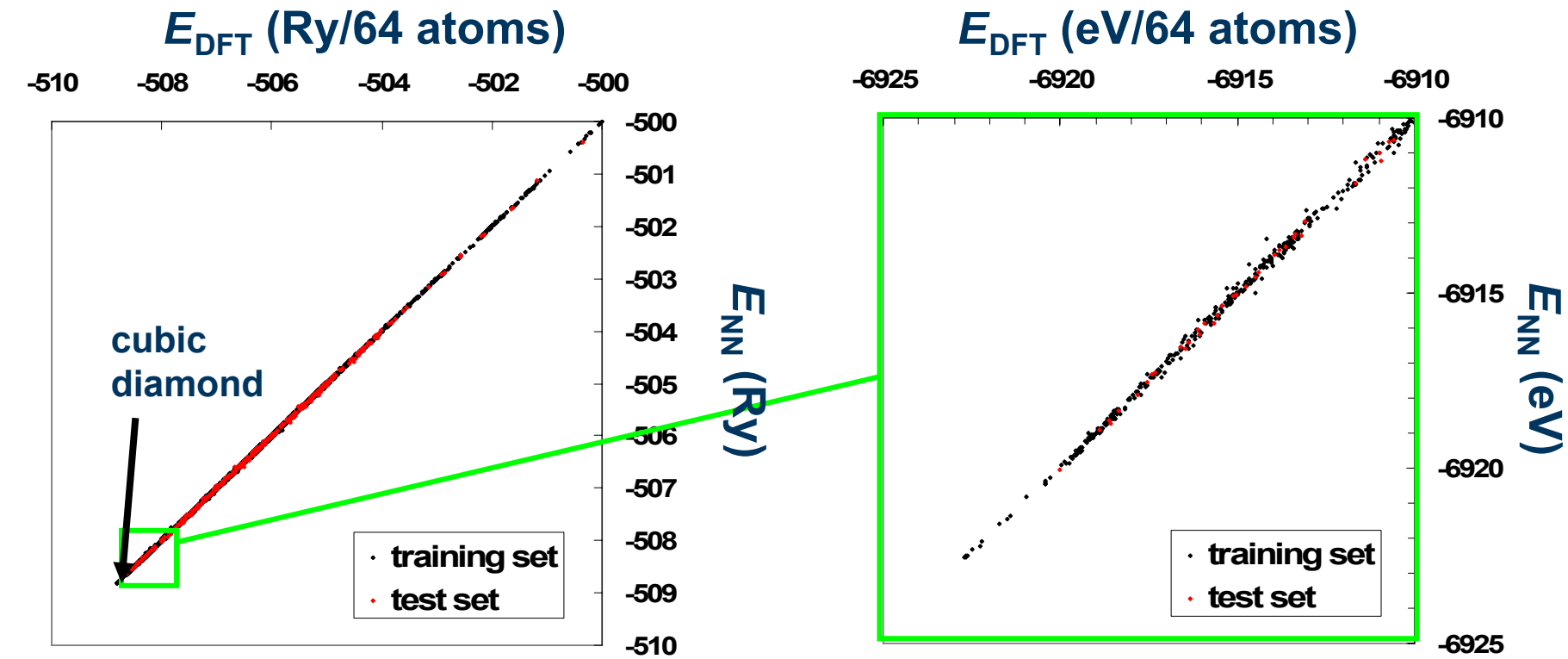
J. Behler, and M. Parrinello, *Phys. Rev. Lett.* 98, 146401 (2007).

Example: 3-atom system



J. Behler, and M. Parrinello, *Phys. Rev. Lett.* 98, 146401 (2007).
J. Behler, R. Martoňák, D. Donadio, and M. Parrinello, *phys. stat. sol. (b)* 245, 2618 (2008).

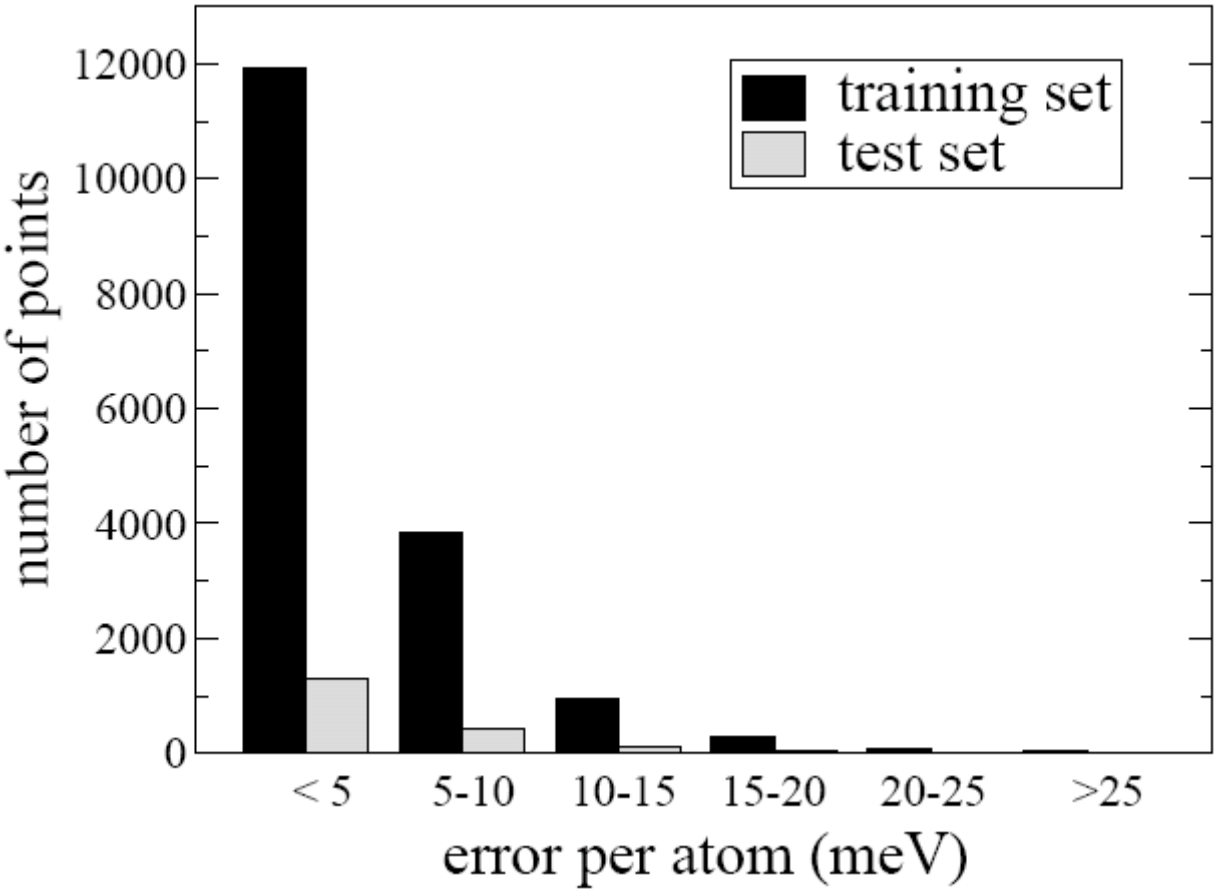
Neural Network Accuracy: Total Energies



Accuracy

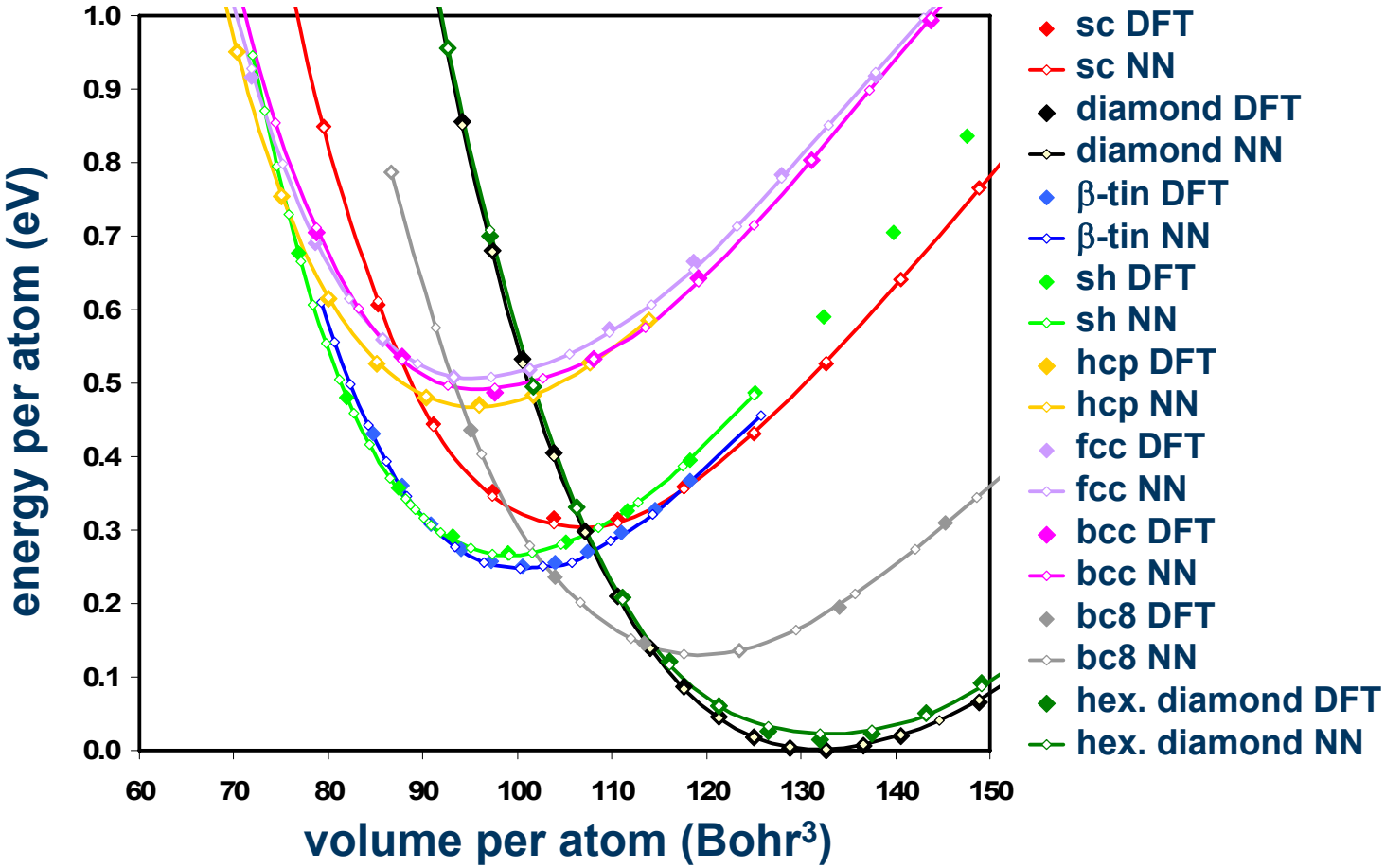
	Points:	RMSE (meV/atom):	MAD (meV/atom):
Training set	17144	5	4
Test set	1907	6	5

Energy error distribution:



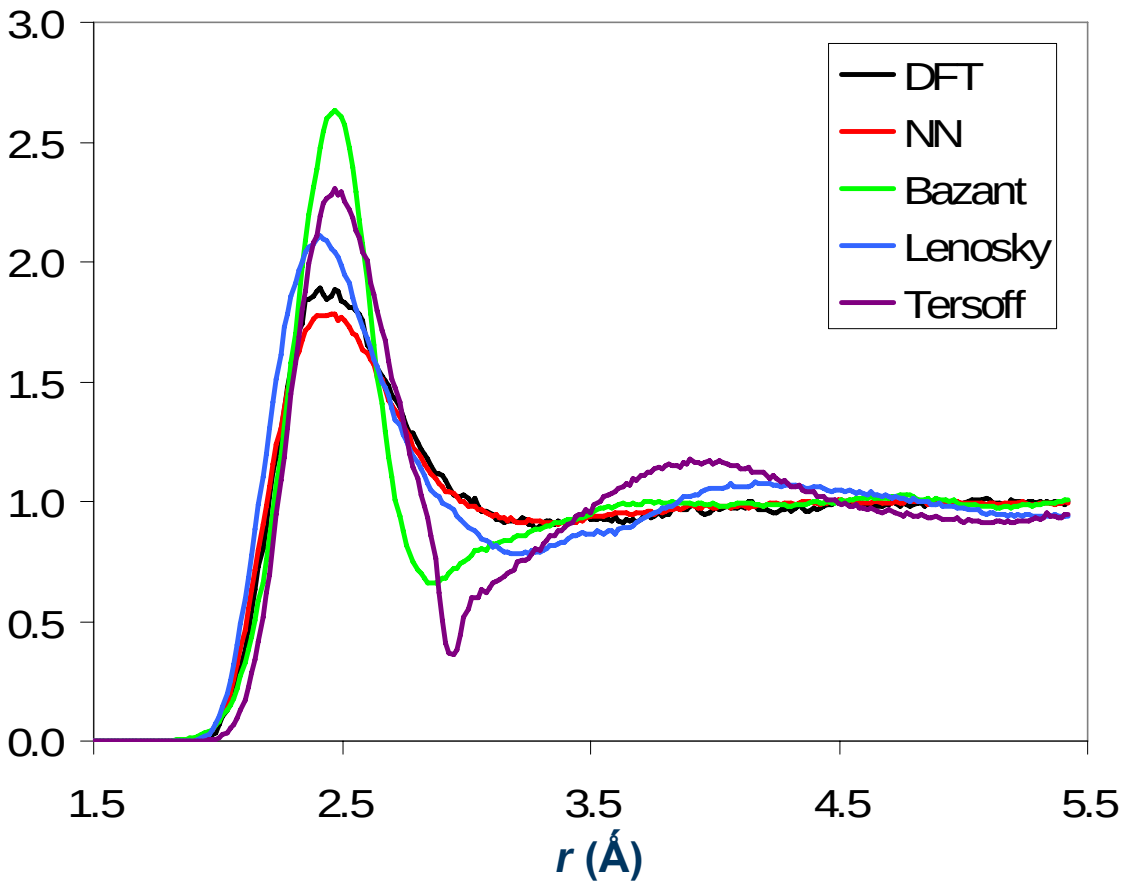
RUB

Ideal Crystal Structures



Disordered Structures:

Liquid Si, 64 atom cell, 3000 K, a = 10.862 Å



DFT data: T. Kühne, *priv. comm.*

RUB

64 atom silicon cell

Timing for the calculation of energy, forces and stress (single Opteron CPU) :

DFT (scf): about 5-10 hours (k-points!)

NN: ≈ 1 second (depending on symmetry functions)

⇒ Speed up factor: 20000

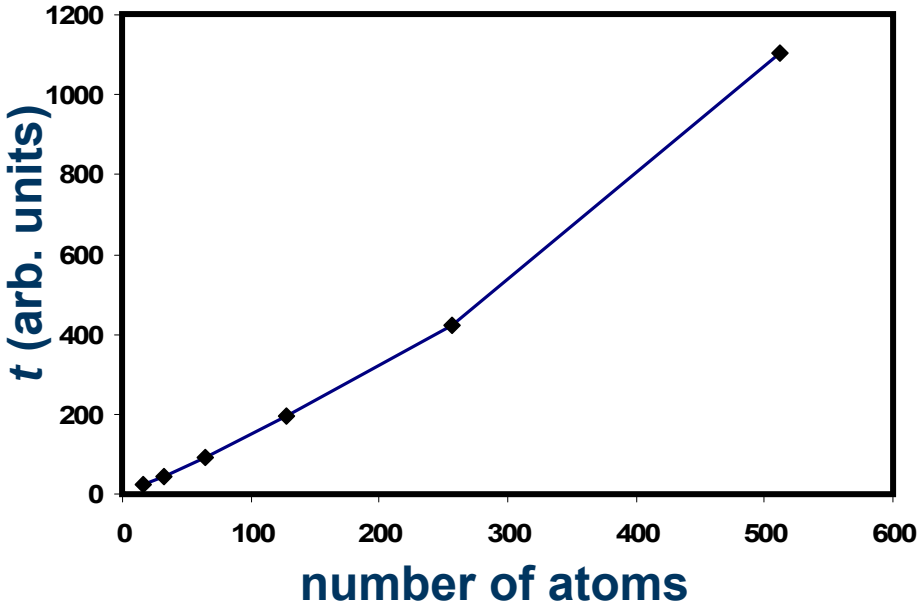
⇒ Now 100 ps MD on a single CPU per day with about DFT accuracy

⇒ Time for a metadynamics run is reduced from 200 CPU years to about 2 days on a standard PC.

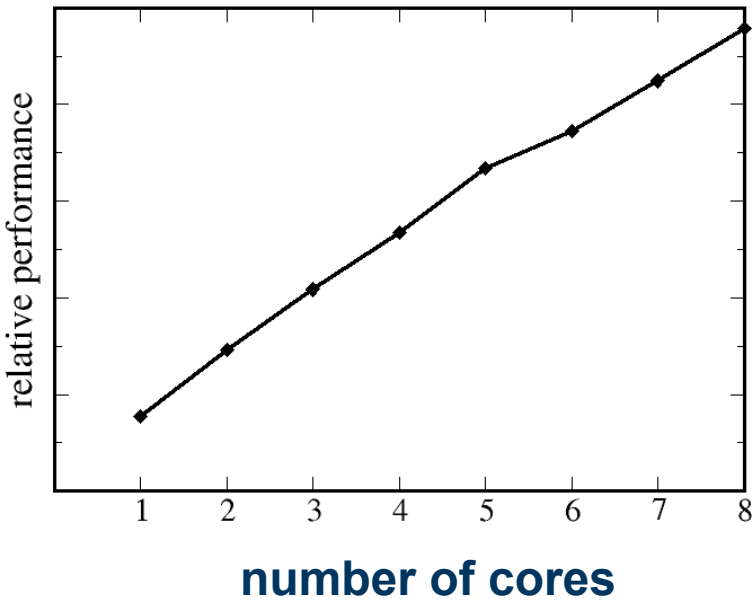
RUB

Linear Scaling with System Size:

Benchmark: MD Simulation



Parallelization:



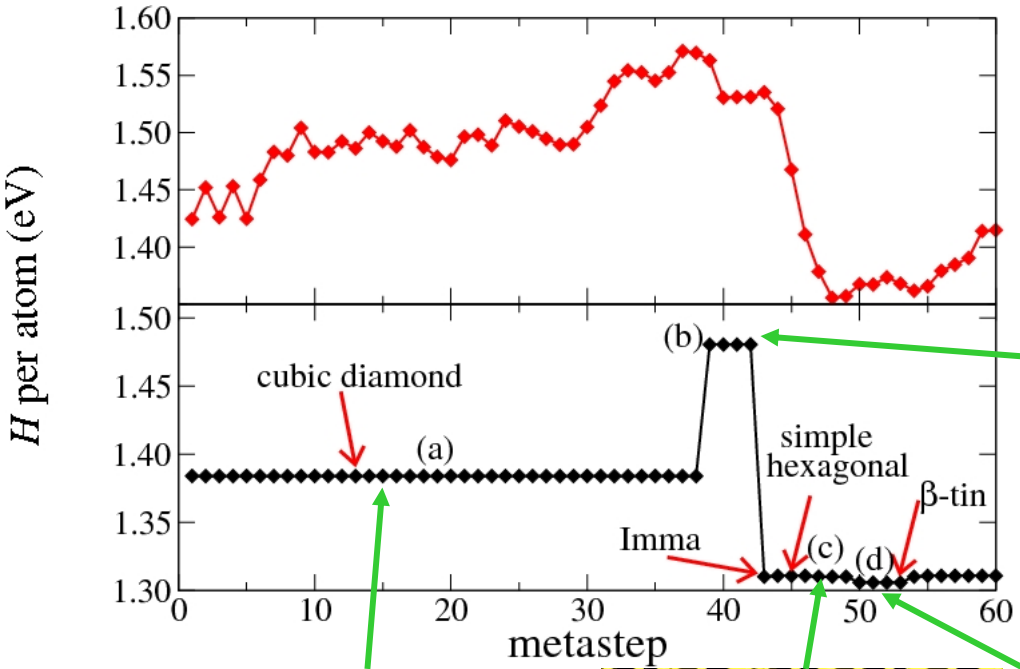
⇒ Neural Network is about 20000 – 100000 times faster than DFT

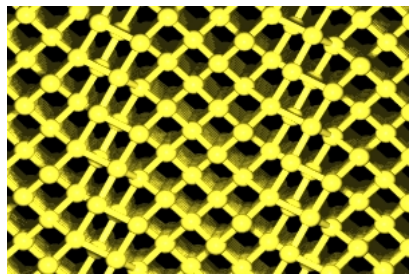
RUB

diamond → β-Sn → Imma → sh → Cmca → hcp → fcc

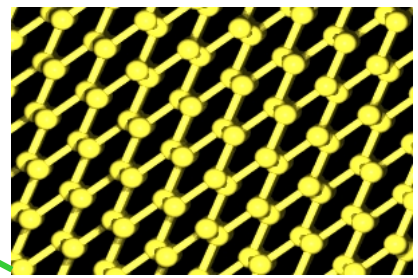
T = 300 K, *p* = 12 GPa

Enthalpy barrier ≈ 0.1 eV/atom



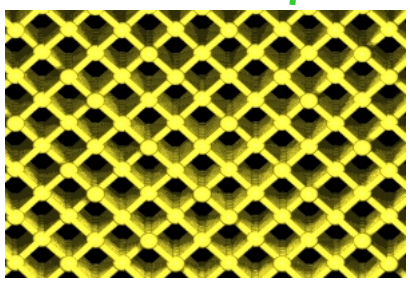


(b)

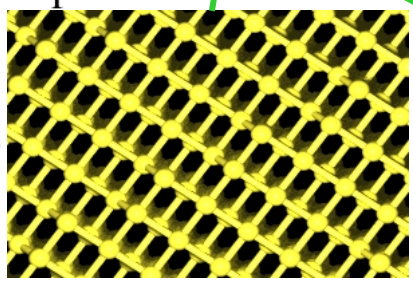


(d)

(a)



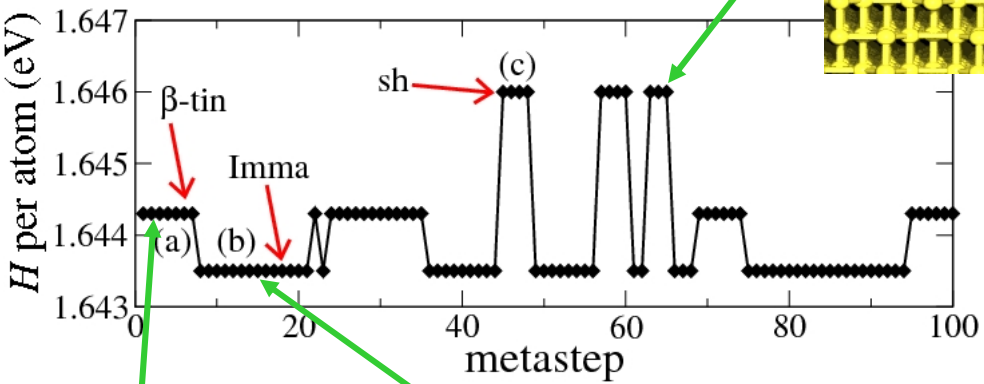
(c)

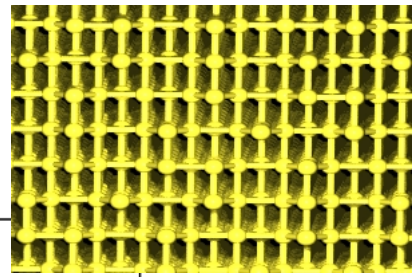


J. Behler, R. Martoňák,
D. Donadio, and M. Parrinello,
Phys. Rev. Lett. 100, 185501 (2008).

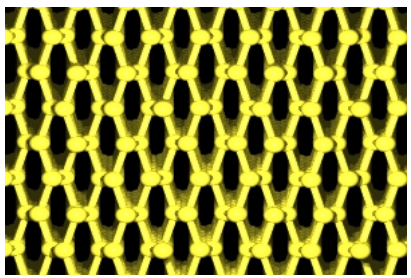
diamond → β-Sn → Imma → sh → Cmca → hcp → fcc

T = 300 K, p = 16 GPa

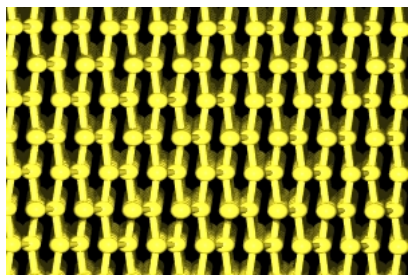




(c)



(a)



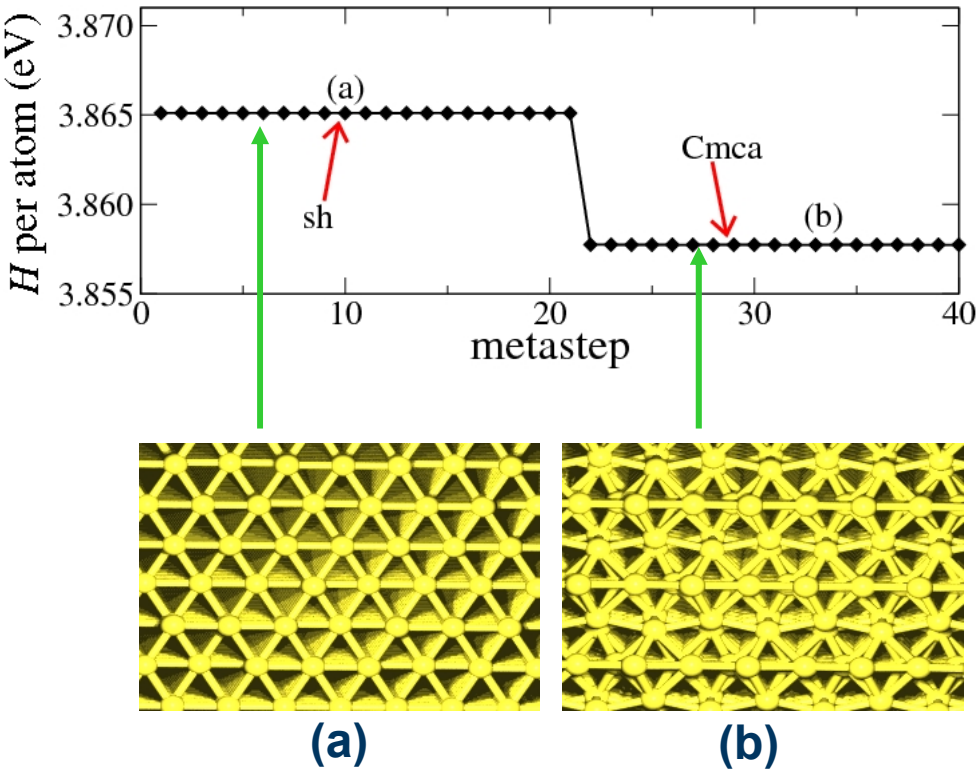
(b)

Metadynamics Simulations of Phase Transitions in Silicon

RUB

diamond → β-Sn → Imma → **sh** → **Cmca** → hcp → fcc

***T* = 800 K, *p* = 45 GPa**



J. Behler, R. Martoňàk, D. Donadio, and M. Parrinello, *phys. stat. sol. (b)* 245, 2618 (2008).

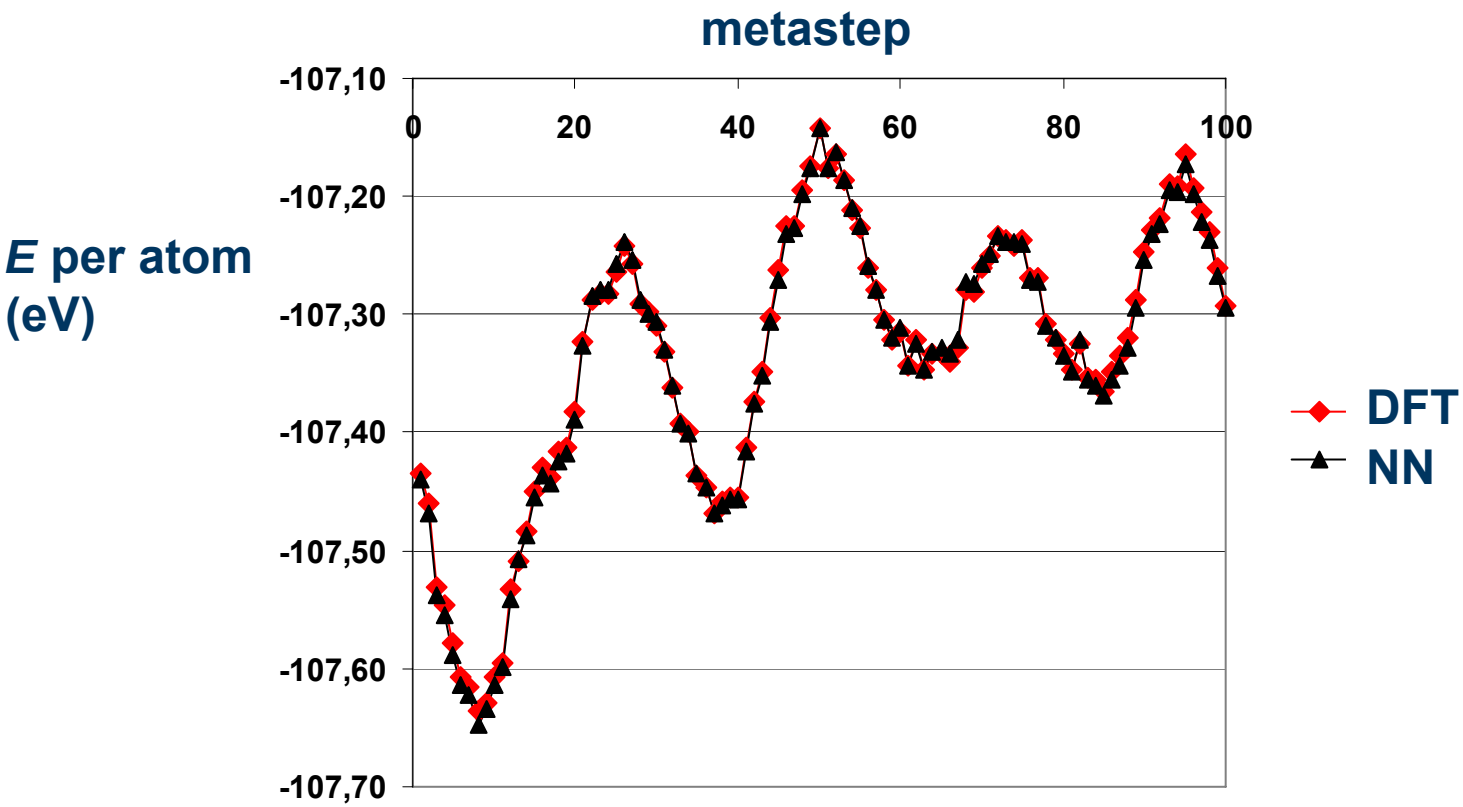
Metadynamics Simulations: Check of the Potential

RUB

diamond → β-Sn → Imma → **sh** → **Cmca** → hcp → fcc

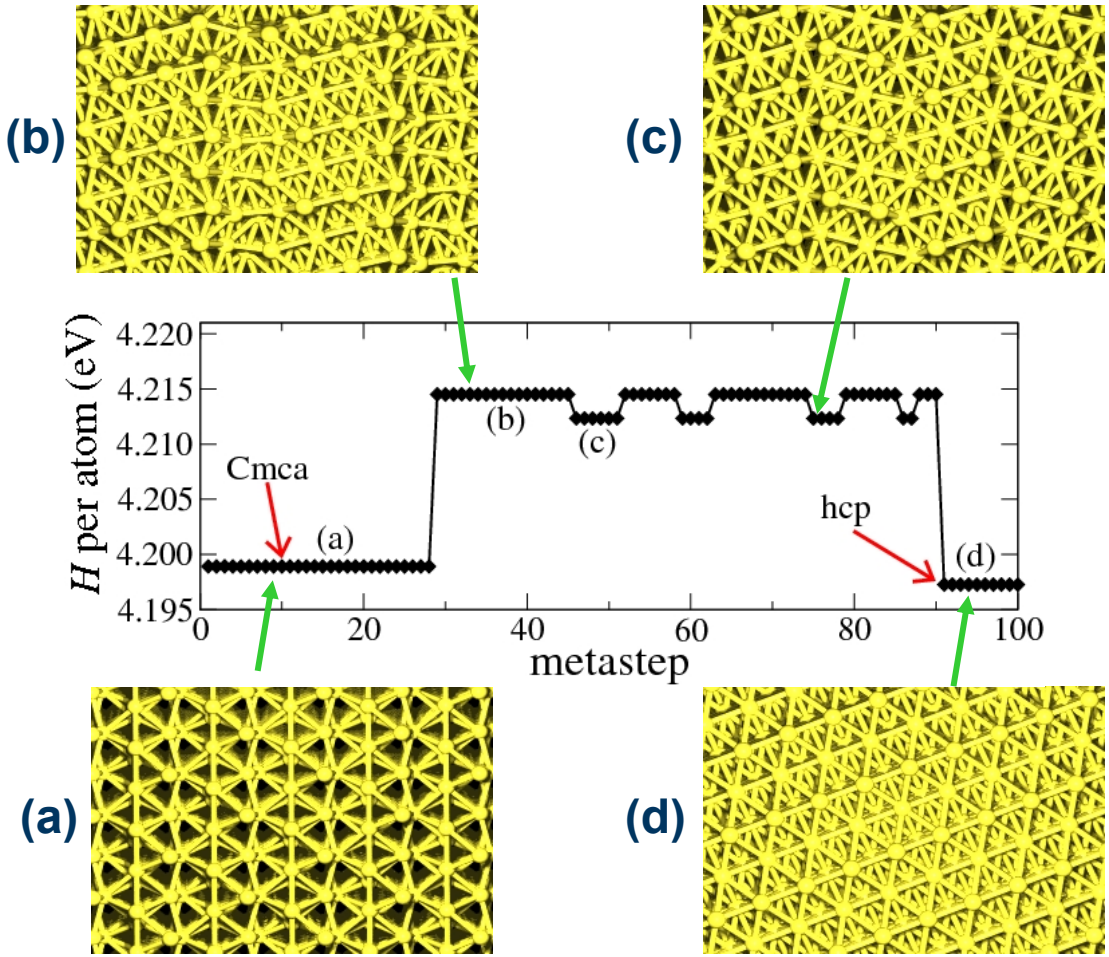
***T* = 800 K, *p* = 45 GPa**

Recalculation of the final structure of each metastep with DFT:



diamond → β-Sn → Imma → sh → **Cmca** → **hcp** → fcc

T = 300 K, p = 50 GPa

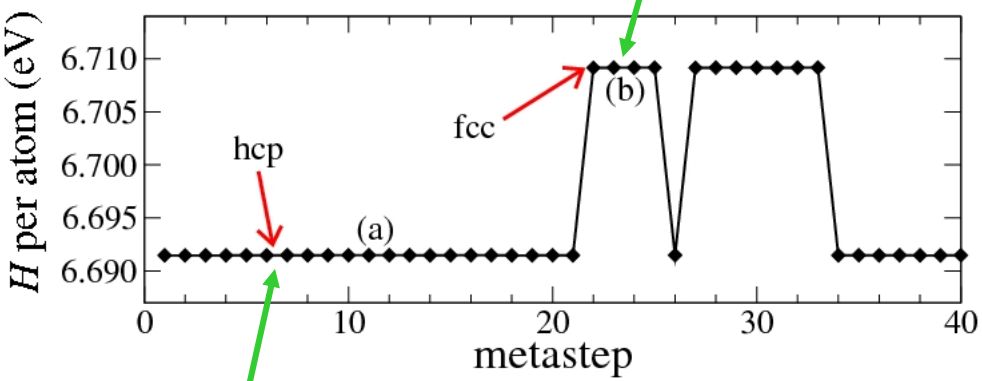


RUB

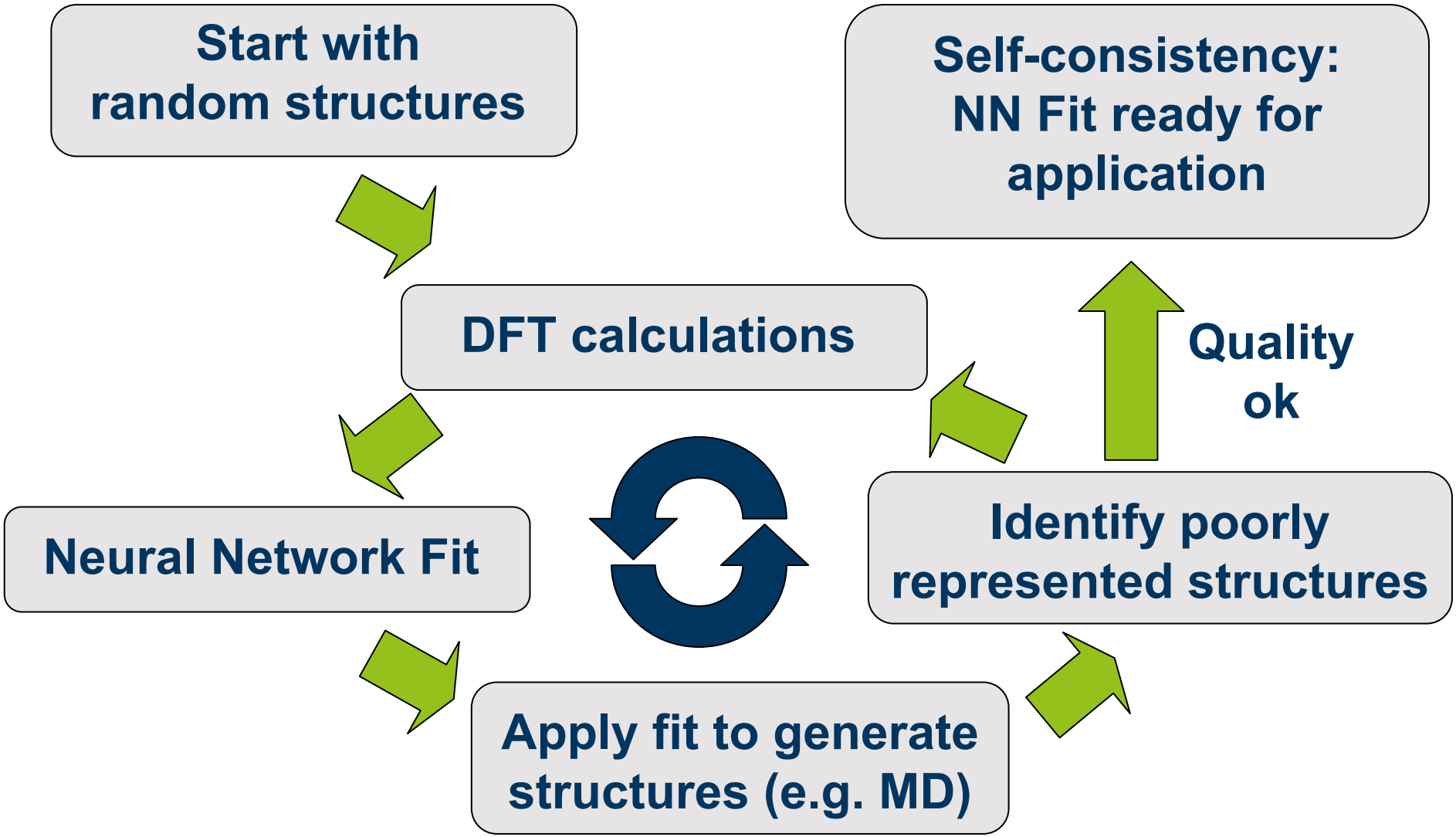
diamond → β-Sn → Imma → sh → Cmca → **hcp** → **fcc**

***T* = 300 K, *p* = 90 GPa**

(b)



(a)


Start with random structures
DFT calculations
Neural Network Fit
Apply fit to generate structures (e.g. MD)
Identify poorly represented structures
Quality ok
Self-consistency: NN Fit ready for application

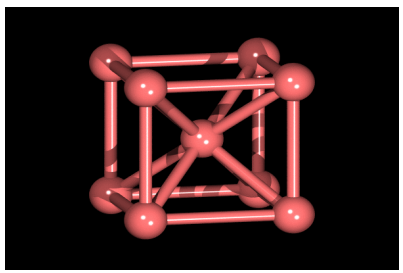
RUB

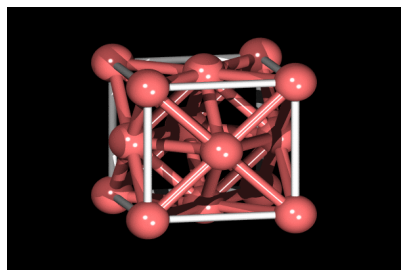
DFT Code: PWSC, PBE functional

PWSCF: S. Baroni et al., http://www.pwscf.org/.

	Structures:	RMSE (meV/atom):
Training Set:	16500	0.72
Test Set:	1500	0.91

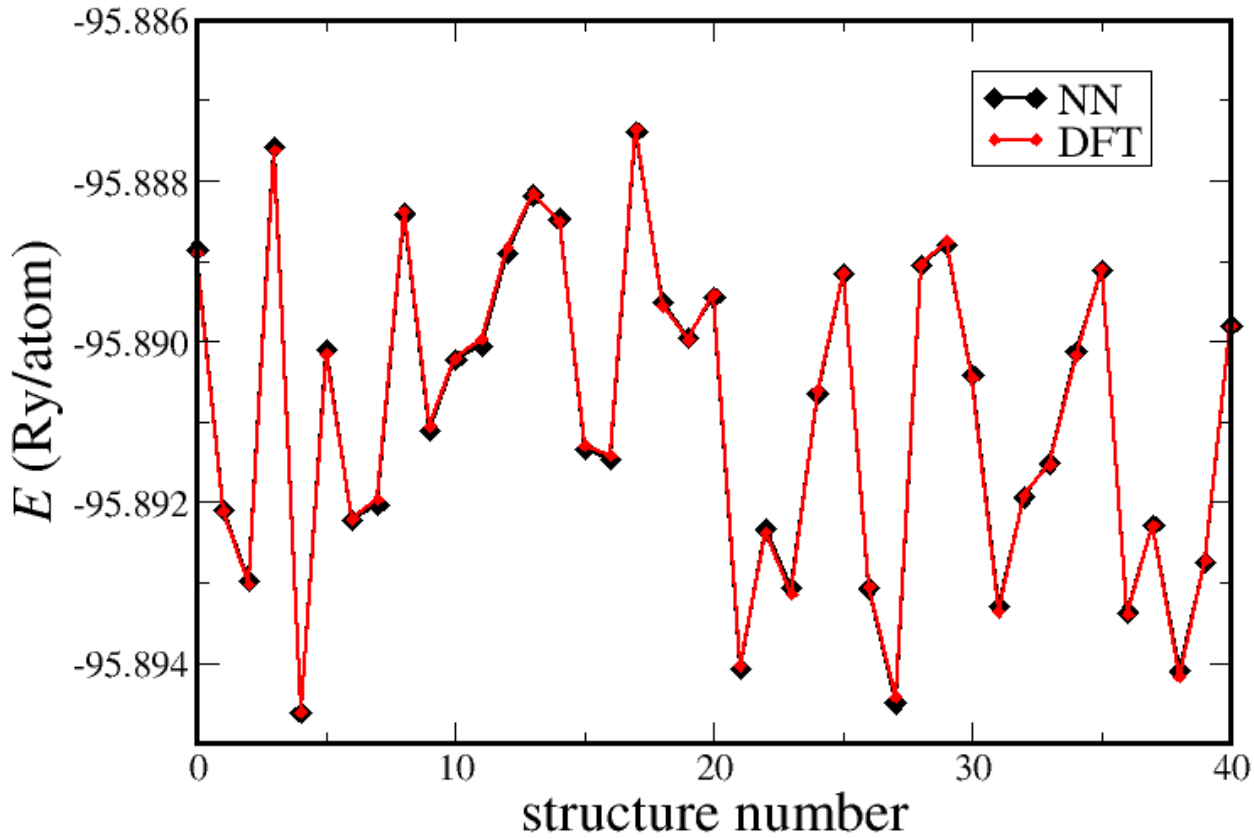
	bcc		fcc	
	DFT	NN	DFT	NN
a_0 (Å)	4.201	4.202	5.297	5.297
B_0 (GPa)	7.63	7.59	7.624	7.613





H. Eshet, R.Z. Khaliullin, T.D. Kühne, J. Behler, and M. Parrinello, *Phys. Rev. B, in press* (2010).

NPT simulation of liquid sodium at 600 K, 100 GPa (64 atom cell)



H. Eshet, R.Z. Khaliullin, T.D. Kühne, J. Behler, and M. Parrinello, *Phys. Rev. B, in press* (2010).

Summary

Neural Network Potentials:

- now applicable to condensed systems (solids, large clusters, liquids)
- NN reproduces total energies very accurately (also in value!)
- provide analytic derivatives (forces, stress)
- fast and linear scaling with system size
- can be constructed using any electronic structure method

Limitations of NN Potentials:

- need many reference calculations
- limited transferability, no extrapolation
- limited number of chemical elements

Outlook:

- useful tool to speed up extended simulations