Coarse-graining potential-energy surfaces from ab initio data using artificial neural networks

Jörg Behler

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

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

Introduction

Paul Dirac (1902 – 1984),
Nobel prize in physics 1933

"The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble."

"It therefore becomes desirable that approximate practical methods of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation."

P.A.M. Dirac, Proc. Roy. Soc. Lond. Ser. A 123 (1929) 714.

Potentials for Different Length and Time Scales

RUB

The accuracy of theoretical simulations depends critically on the reliability of the employed interatomic potentials.

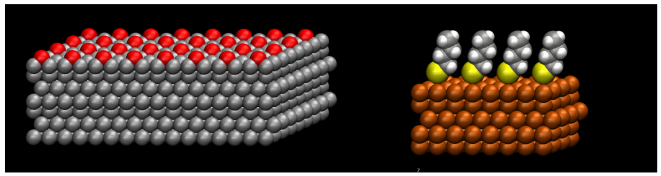
Method	Atoms	Simulation Time
CI, CC, MP2, MP4, …	10 - 100	0
Density Functional Theory	100 - 1000	100 ps
Semiempirical Methods Tight Binding	1000 - 10 000	1 ns
„Reactive Potentials“ EAM, Tersoff, …	1000 - 100 000	10 ns
Classical Force Fields (LJ, harmonic, Coulomb)	1 000 000	1 ms

Examples: Large, Reactive, Complex Systems

RUB

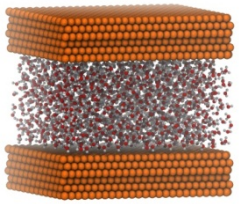
Surface Science:

- heterogeneous catalysis
- corrosion
- self-assembled monolayers



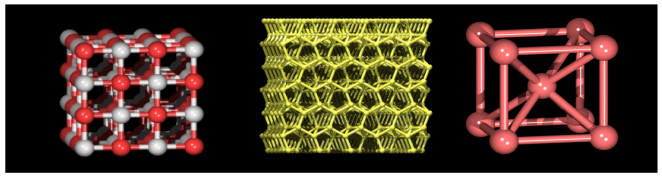
Solid-Liquid Interface:

- heterogeneous catalysis
- electrochemistry



Materials Science:

- crystal structure prediction
- phase diagrams
- properties of materials



Coordination:

- solvation of ions
- metal organic frameworks
- active sites in enzymes

⇒ **Reliable potentials for large-scale simulations often not available!**

Definition:

- large (>> 1000 atoms)
  ⇒ electronic structure calculations too costly

- various types of bonding
  ⇒ necessary to describe simultaneously
    - covalent bonds
    - non-bonded interactions (vdW)
    - metals, delocalized electrons
    - polarization, charge transfer, electrostatics

  } very difficult to find a universal functional form

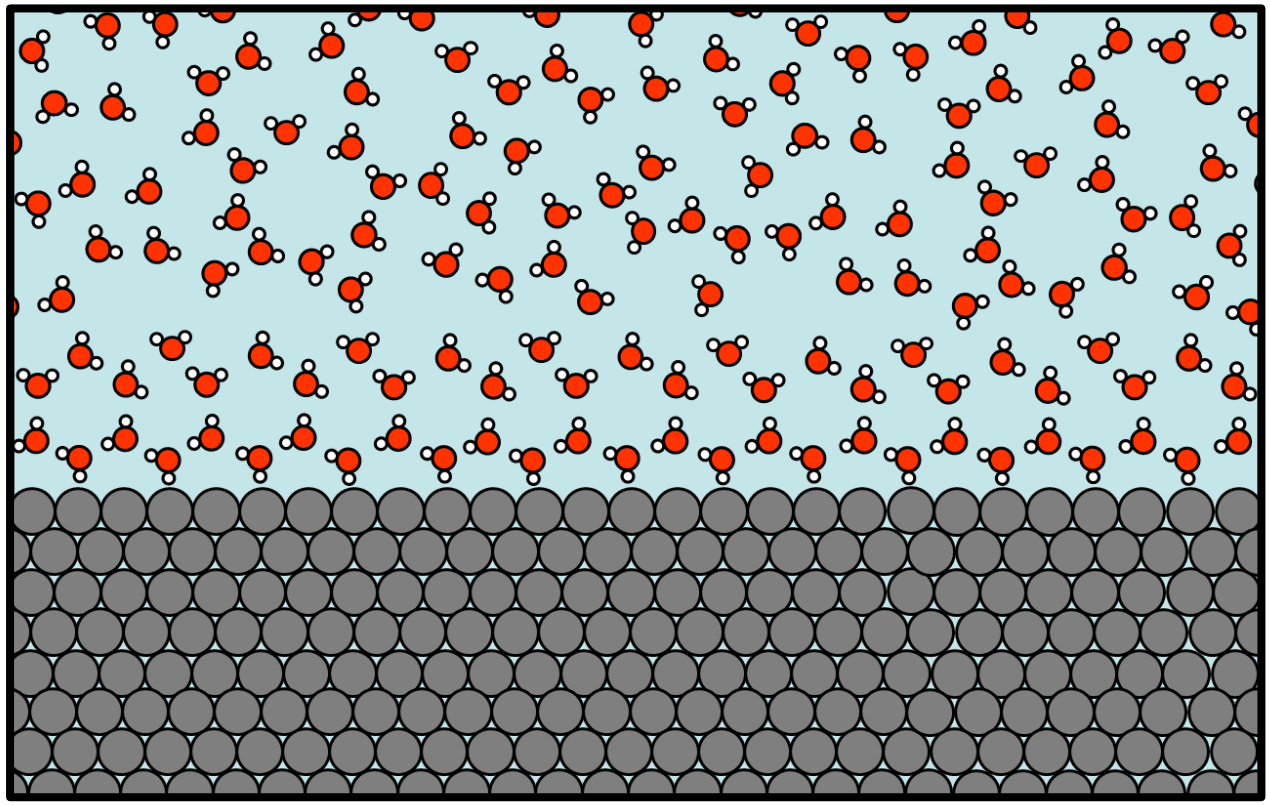
- dynamics, statistics, reactivity
    - bond making and breaking
    - complex structural rearrangements
    - sampling, free energies

Potentials of *ab initio* quality are often not available!

System Size: Solid-Liquid Interface Example

RUB

Example: Solid-Liquid Interface

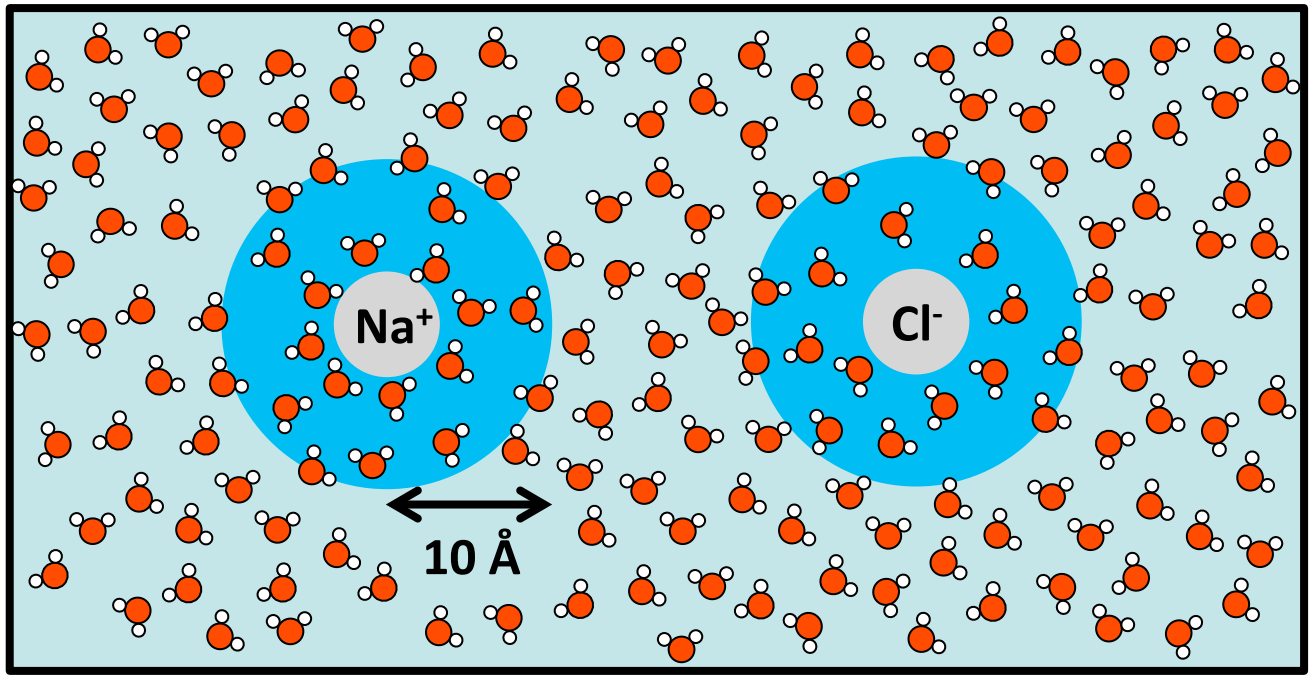


Which system size is required to describe a *liquid bulk electrolyte* in contact with a metal electrode?

System Size: Solid-Liquid Interface Example

Starting point: Low concentration ⇒ e.g. solvated Na^+Cl^- ion pair

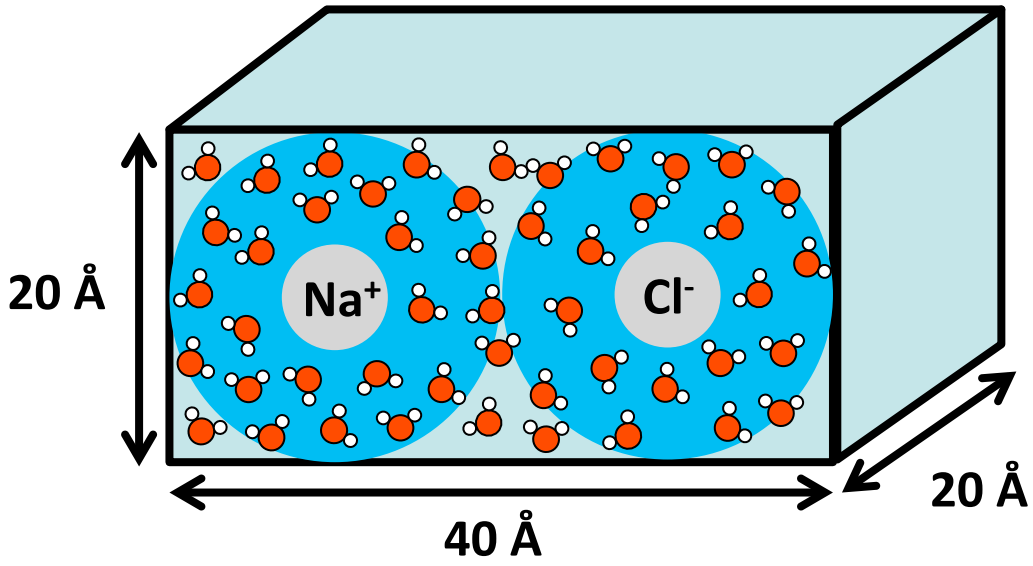
Ions modify the water structure up to a distance of about 10 Å.



System Size: Solid-Liquid Interface Example

Starting point: Low concentration ⇒ e.g. solvated Na^+Cl^- ion pair

⇒ smallest possible system for "non-overlapping", uncorrelated solvation spheres:



⇒ volume = 40 x 20 x 20 $Å^3$ = 16 000 $Å^3$

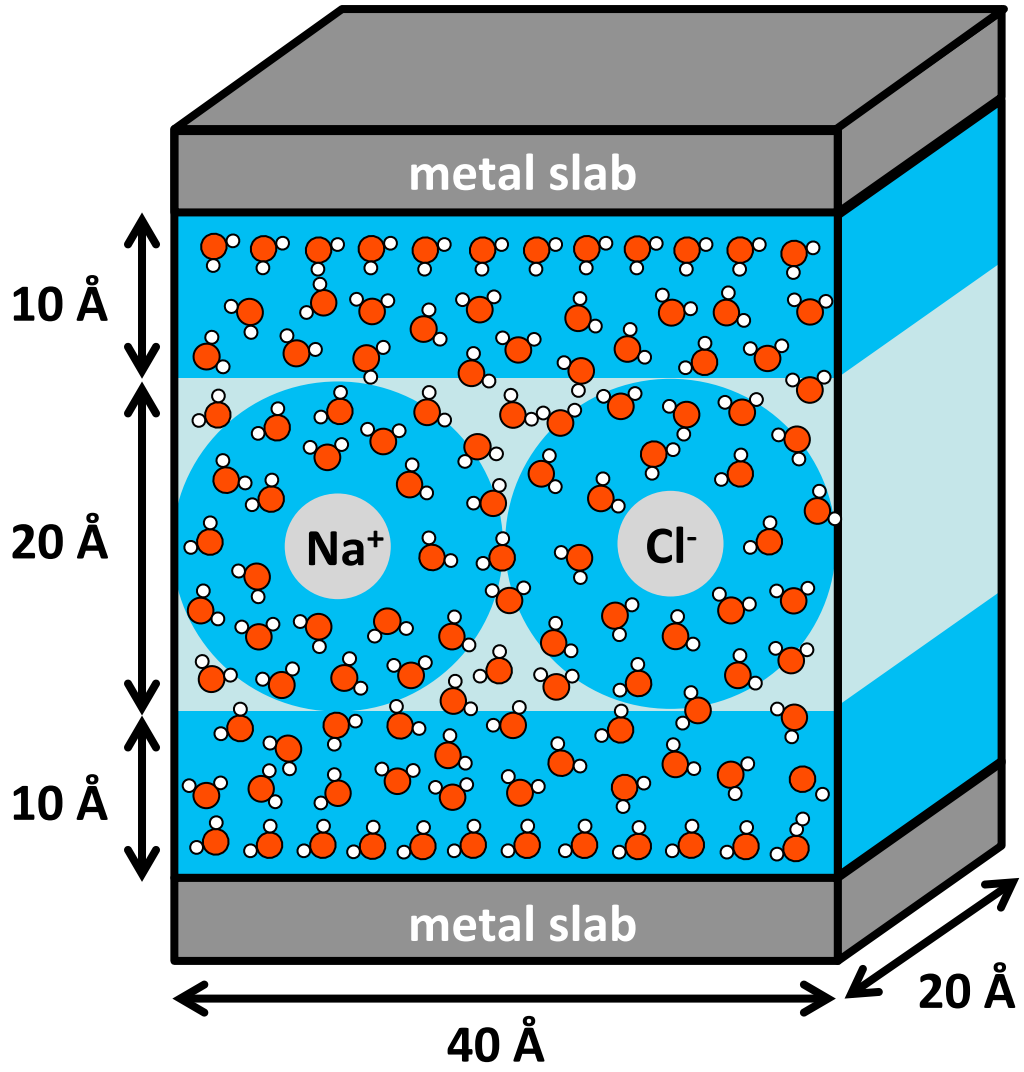
volume of a single water molecule: ≈ 30 $Å^3$

⇒ ≈ 530 water molecules (concentration 0.1 M)

System Size: Solid-Liquid Interface Example

RUB

Slab model: Surfaces modify water structure up to 10 Å.



⇒ volume = 32 000 Å^3

⇒ ≈ 1060 water molecules
+ ≈ 700 metal atoms
(7 layer slab)

⇒ ≈ 3900 atoms

System size cannot be reduced for higher concentrations!

⇒ beyond *ab initio* MD

Holds for any concentration!!!

Interatomic Potentials: Motivation

Schrödinger equation:

$$H\Psi = E\Psi$$

Born Oppenheimer approximation ⇒ Electronic Schrödinger equation and Hamiltonian:

$$H_{el} = -\tfrac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}}$$

Total Energy:

$$E_{tot} = E_{el} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}}$$

Hamiltonian depends on

- **nuclear positions R**
- **nuclear charges *Z* (chemical elements)**
- **system total charge *Q* (usually neutral)**

⇒ Structure contains all information required to fully define the total energy

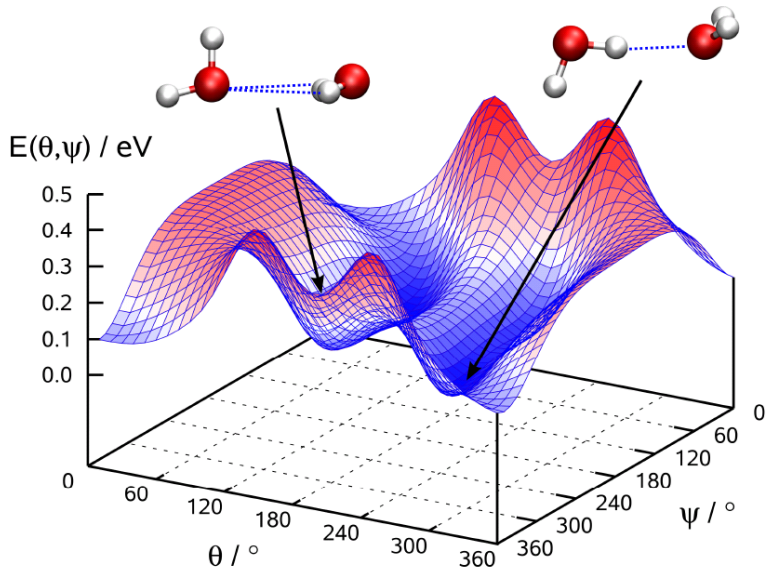
⇒ If we knew the function *E*(*R*,*Z*,*Q*) we could avoid solving the electronic structure problem explicitly again and again in MD simulations

Potential Energy Surfaces

The *Potential Energy Surface* (PES) is a high-dimensional real-valued function, which yields the total energy of the system if the atomic positions are provided in form of a suitable set of coordinates.

3*N*-6-dimensional ⇒ very complicated

Example: 2D cut of the 12-dimensional PES of the water dimer



Criteria for the “Perfect Potential”

RUB

1) **High Accuracy**
   The potential should be close to *ab initio* quality.

2) **Systematic Improvements**
   If the potential fails in some situation, it should be easy to improve the accuracy systematically.

3) **General Functional Form**
   The potential should be applicable to all types of systems and bonding.

4) **“Reactive”**
   The potential should be able to describe the breaking and making of bonds.

5) **High-Dimensional**
   The potential should depend on all atomic positions in the system and include all many-body effects.

6) **Effort**
   The construction of the potential should not require much human work.

Criteria for the “Perfect Potential”

RUB

7) **Transferability**
The potential should be applicable to a wide range of bonding situations.

8) **Efficiency**
The potential should be fast.

9) **Costs**
The construction of the potential should not require much CPU time.

10) **Derivatives**
Analytic derivatives must be available for the calculation of forces.

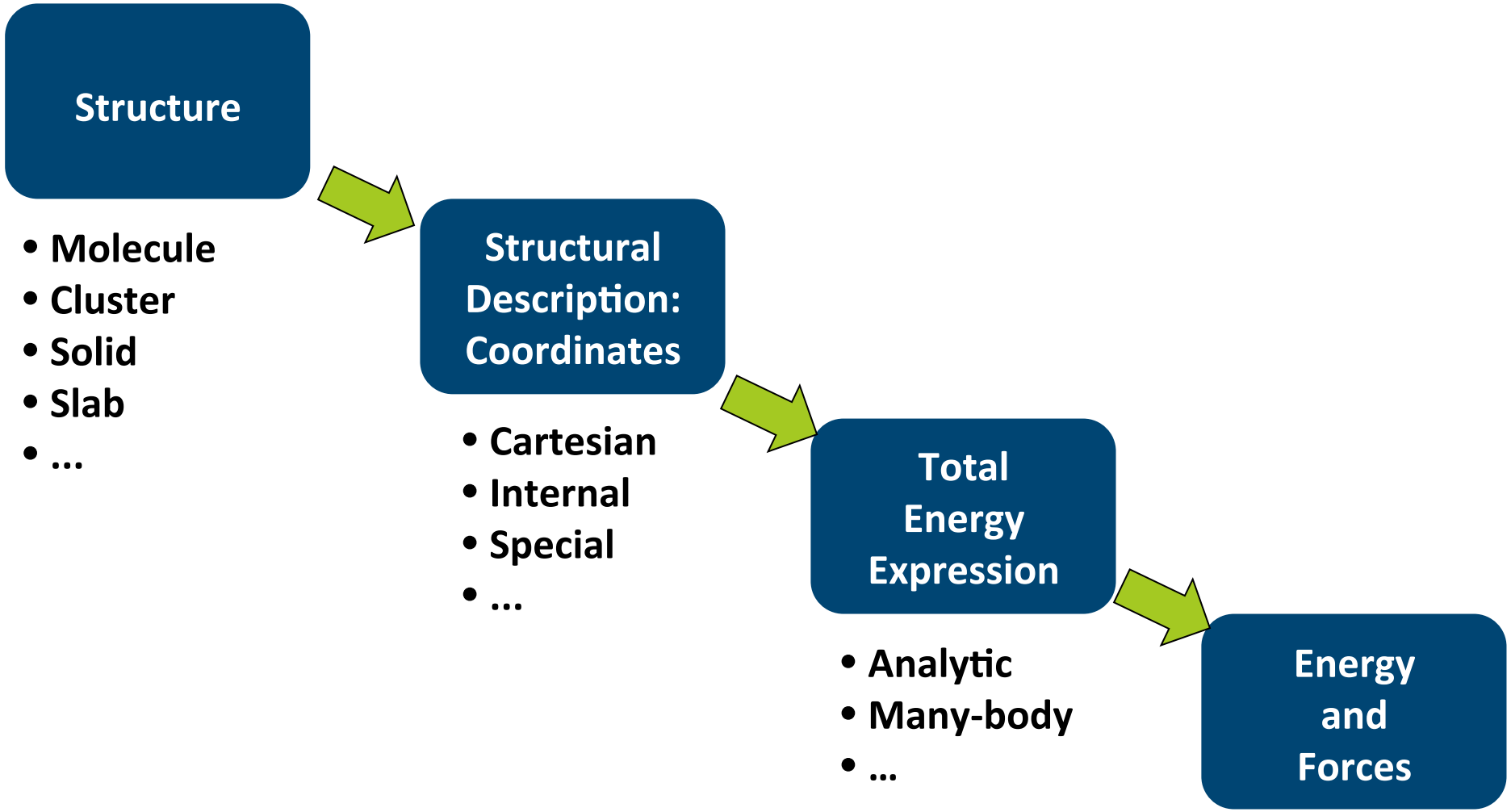
None of the presently available potentials fulfills all requirements!

⇒ Very active field of research

Definition of the Problem

RUB

Construction of a functional relation between the structure and its energy



Classification of Interatomic Potentials

Physical Approach: „Conventional“ Potentials

Typically fixed functional form ⇒ intrinsically limited accuracy

Examples:

- **Lennard-Jones potential (classical force fields)**

$$E_{ij} = 4\varepsilon\left[\left(\frac{\sigma}{R_{ij}}\right)^{12} - \left(\frac{\sigma}{R_{ij}}\right)^{6}\right]$$

- **Tersoff, Brenner, Stillinger Weber**

Mathematical Approach:

Use of extremely flexible functions adapting to high-level reference data.

Examples:

- **Splines**
- **Polynomial Fits**
- **Gaussian Approximation Potentials**
- **Artificial Neural Networks**

Example: The AMBER force field

W.D. Cornell et al., *J. Am. Chem. Soc.* 117 (1995) 5179.

$$E = \sum_{\text{bonds}} k_b \left(r - r_0\right)^2 + \sum_{\text{angles}} k_a \left(\alpha - \alpha_0\right)^2 + \sum_{\text{torsions}} \tfrac{1}{2} V_n \left[1 + \cos\left(n\omega - \gamma\right)\right]$$

$$+ \sum_{j=1}^{N-1} \sum_{i=j+1}^{N} \left\{ \varepsilon_{ij} \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - 2 \left(\frac{\sigma}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \right\}$$

Advantages:
- **very fast**
- **easy to implement**

Disadvantages:
- **not very accurate**
- **no reactivity (there are exceptions, e.g. ReaxFF)**

Other Force fields:

CHARMM: B.R. Brooks et al., *J. Comp. Chem.* 4 (1983) 187.

GROMOS: W.F. van Gunsteren, H.J.C. Berendsen , GROMOS Library Manual, Groningen (1987).

Standard Potential for Materials Science: Tersoff Potential

J. Tersoff, *Phys. Rev. Lett.* 56 (1986) 632.

J. Tersoff, *Phys. Rev. B* 37 (1988) 6991.

J. Tersoff, *Phys. Rev. B* 38 (1988) 9902.

Idea: Morse potential-like expression with environment-dependent bond strength parameter (bond order dependence)

$$E = \sum_i E_i = \frac{1}{2} \sum_i \sum_{j \neq i} E_{ij}$$

equivalent: atomic or pair energy formulation

$$E_{ij} = f_{cut}\left(R_{ij}\right)\left[A \cdot e^{-\lambda_1 R_{ij}} - B_{ij} e^{-\lambda_2 R_{ij}}\right]$$

(Morse potential: $\lambda_1 = 2\lambda_2$)

$$B_{ij} = B_0 e^{-z_{ij}/b}$$

bond order term, z_{ij} is complicated function of bonding environment, *b* determines decay of bond strength

reactive, but difficult to apply to multicomponent systems

One of the standard potentials in materials science!
Difficult to parameterize and not very accurate!

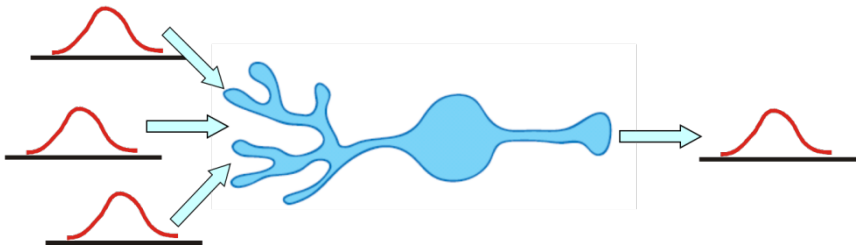
Neural Network Potentials

Neural Networks (NN)

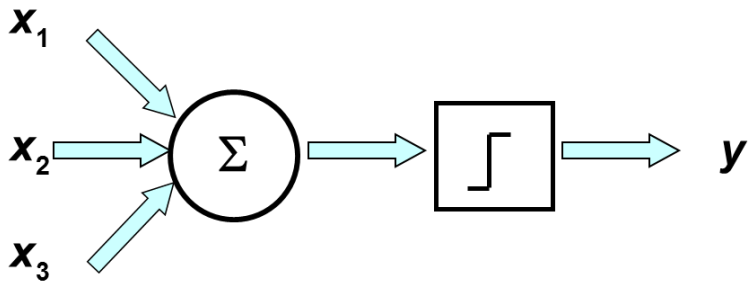
*"**Artificial neural networks** are massively parallel interconnected networks of simple (usually adaptive) elements and their hierarchical organizations, which are intended to interact with the objects of the real world in the same way as **biological nervous systems** do."*

T. Kohonen, *Neural Networks* 1, 3 (1988).

"Bio-Neurons"



Artificial Neuron

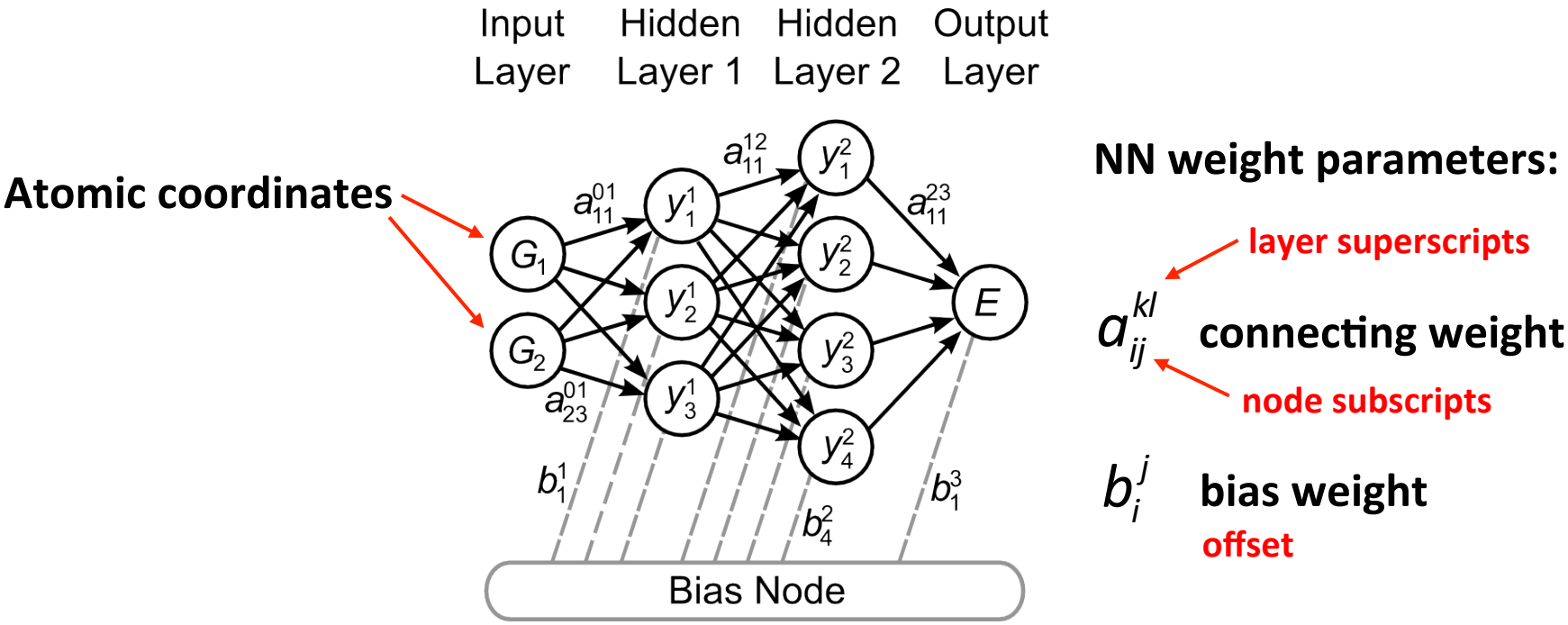


W. McCulloch and W. Pitts, *Bull. Math. Biophys.* 5 (1943) 115.

⇒ **artificial NNs mimic the signal processing in the nervous system**

Feed-Forward Neural Networks: Analytic Form

Small example for a 2D feed-forward Neural Network:



NN weight parameters:

layer superscripts

a_{ij}^{kl} **connecting weight**

node subscripts

b_i^j **bias weight**

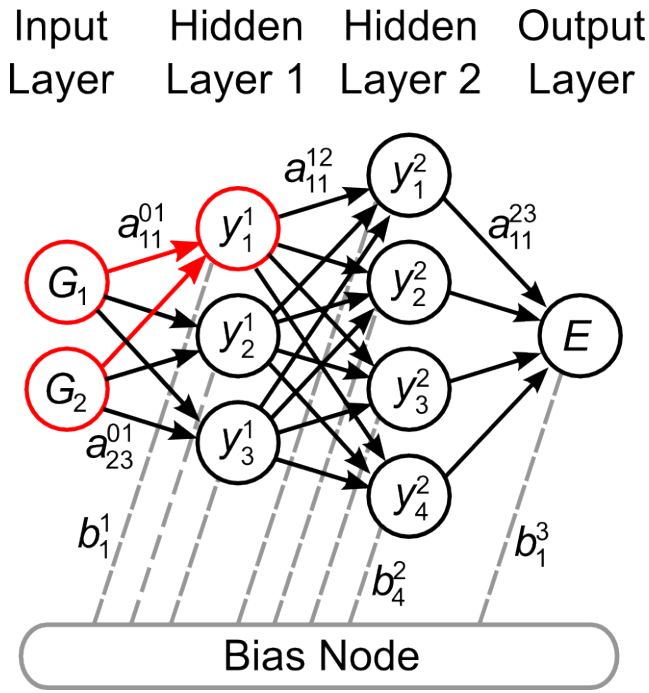
offset

Functional form:

$$E = f_1^3\left(b_1^3 + \sum_{k=1}^{4} a_{k1}^{23} \cdot f_k^2\left(b_k^2 + \sum_{j=1}^{3} a_{jk}^{12} \cdot f_j^1\left(b_j^1 + \sum_{i=1}^{2} a_{ij}^{01} \cdot G_i \right) \right) \right)$$

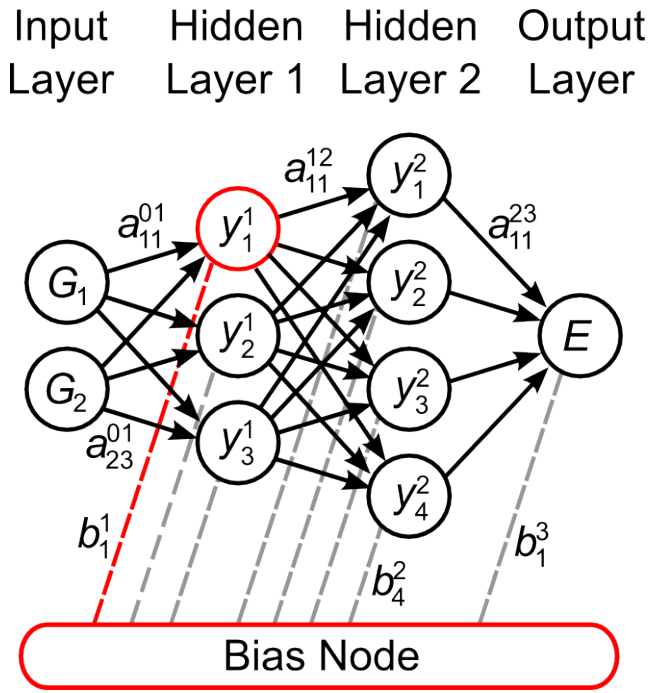
Feed-Forward Neural Networks: Analytic Form

Step 1: Linear combination of input values



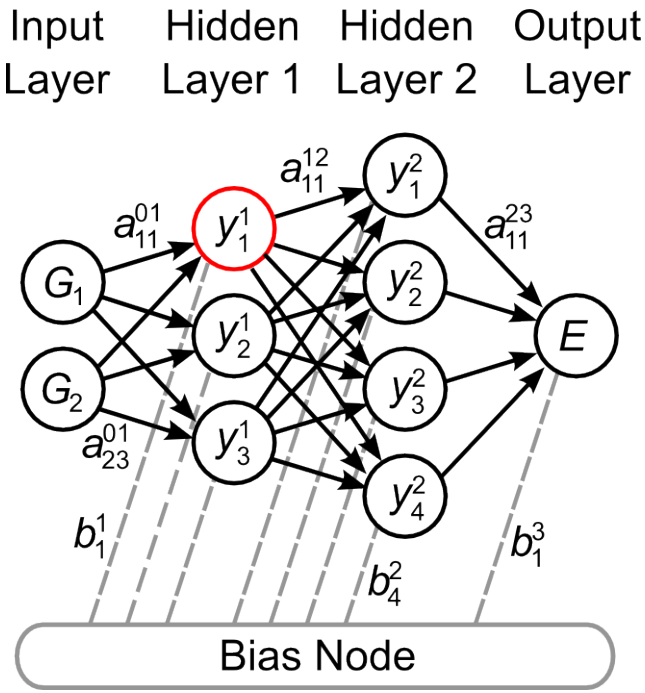
$$E = f_1^3\left(b_1^3 + \sum_{k=1}^{4} a_{k1}^{23} \cdot f_k^2\left(b_k^2 + \sum_{j=1}^{3} a_{jk}^{12} \cdot f_j^1\left(b_j^1 + \sum_{i=1}^{2} a_{ij}^{01} \cdot G_i\right)\right)\right)$$

Step 2: Adding bias weight



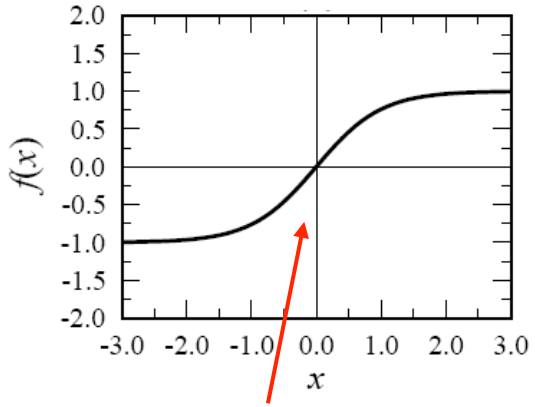
$$E = f_1^3\left(b_1^3 + \sum_{k=1}^{4} a_{k1}^{23} \cdot f_k^2\left(b_k^2 + \sum_{j=1}^{3} a_{jk}^{12} \cdot f_j^1\left(b_j^1 + \sum_{i=1}^{2} a_{ij}^{01} \cdot G_i\right)\right)\right)$$

Step 3: Apply Activation Function



Activation Function

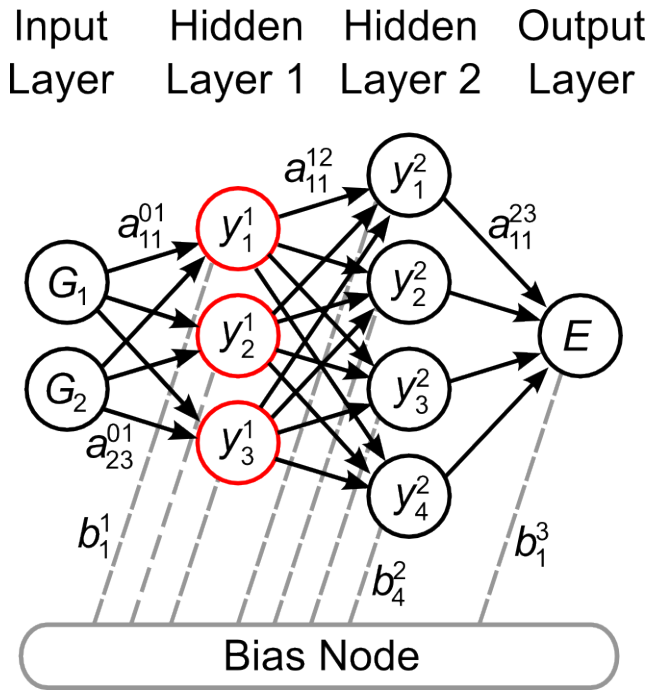
$$y = \tanh(x)$$



nonlinear

$$E = f_1^3\left(b_1^3 + \sum_{k=1}^{4} a_{k1}^{23} \cdot f_k^2\left(b_k^2 + \sum_{j=1}^{3} a_{jk}^{12} \cdot f_j^1\left(b_j^1 + \sum_{i=1}^{2} a_{ij}^{01} \cdot G_i\right)\right)\right)$$

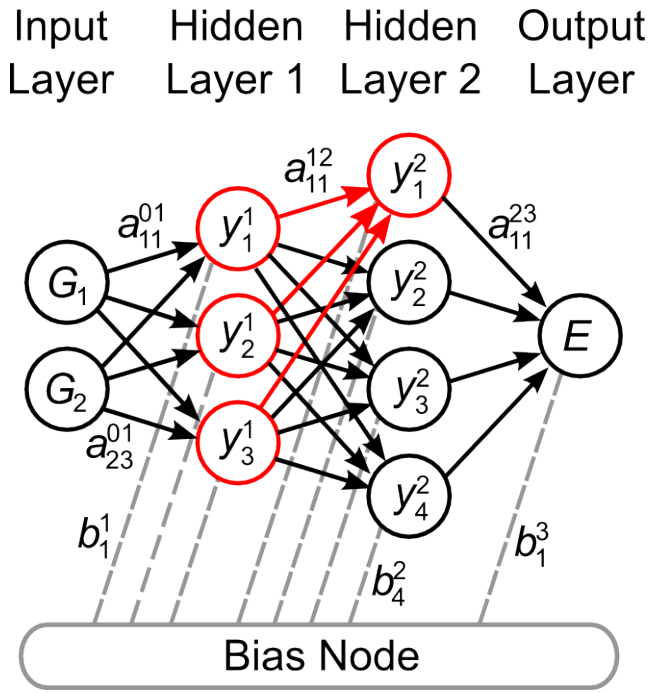
Step 4: Calculate all nodes in hidden layer 1 in the same way



$$E = f_1^3\left(b_1^3 + \sum_{k=1}^{4} a_{k1}^{23} \cdot f_k^2\left(b_k^2 + \sum_{j=1}^{3} a_{jk}^{12} \cdot f_j^1\left(b_j^1 + \sum_{i=1}^{2} a_{ij}^{01} \cdot G_i\right)\right)\right)$$

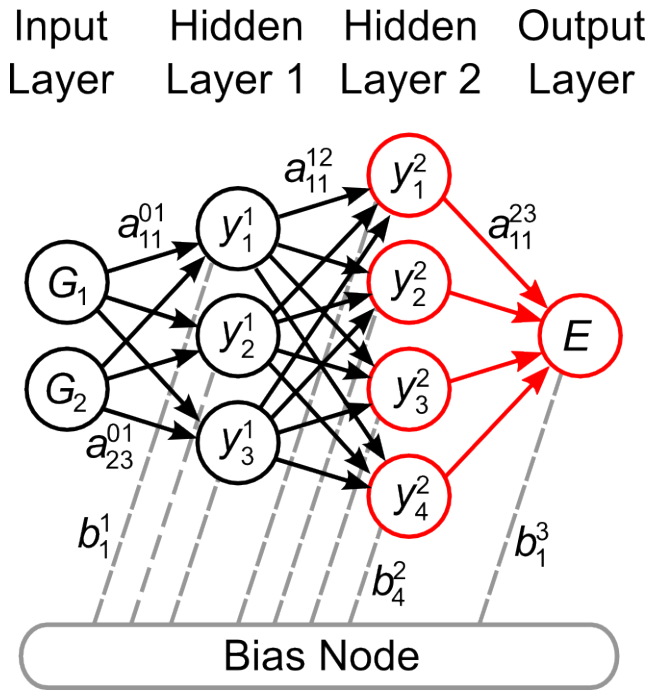
Step 5: Calculate all nodes in hidden layer 2 in the same way

- **linear combination**
- **add bias weight**
- **activation functions**



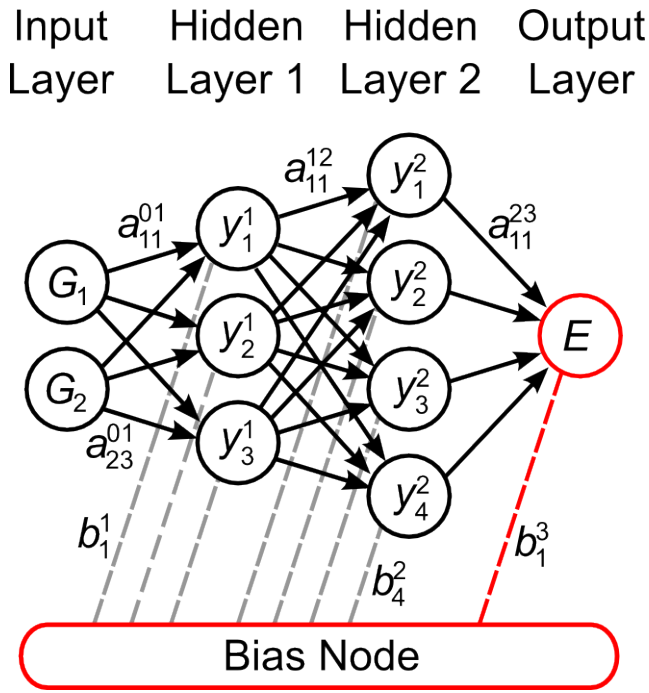
$$E = f_1^3\left(b_1^3 + \sum_{k=1}^{4} a_{k1}^{23} \cdot f_k^2\left(b_k^2 + \sum_{j=1}^{3} a_{jk}^{12} \cdot f_j^1\left(b_j^1 + \sum_{i=1}^{2} a_{ij}^{01} \cdot G_i\right)\right)\right)$$

Step 6: Calculate output node



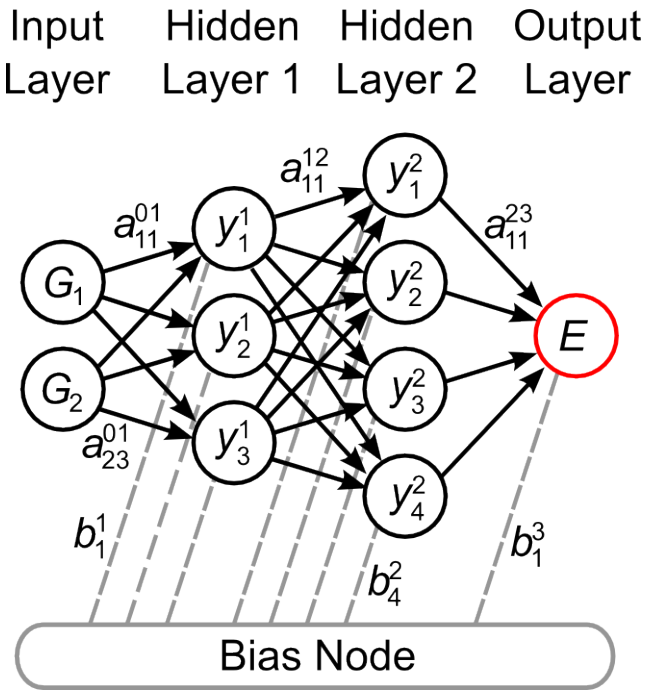
$$E = f_1^3\left(b_1^3 + \sum_{k=1}^{4} a_{k1}^{23} \cdot f_k^2\left(b_k^2 + \sum_{j=1}^{3} a_{jk}^{12} \cdot f_j^1\left(b_j^1 + \sum_{i=1}^{2} a_{ij}^{01} \cdot G_i\right)\right)\right)$$

Step 7: Add bias weight at output node



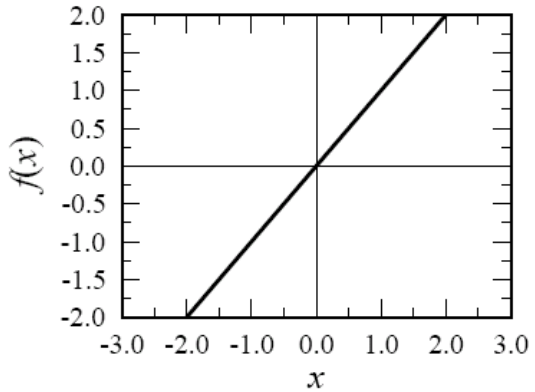
$$E = f_1^3\left(b_1^3 + \sum_{k=1}^{4} a_{k1}^{23} \cdot f_k^2\left(b_k^2 + \sum_{j=1}^{3} a_{jk}^{12} \cdot f_j^1\left(b_j^1 + \sum_{i=1}^{2} a_{ij}^{01} \cdot G_i \right) \right) \right)$$

Step 8: Apply linear activation function at output node



Output node: Activation function

$$y = x$$

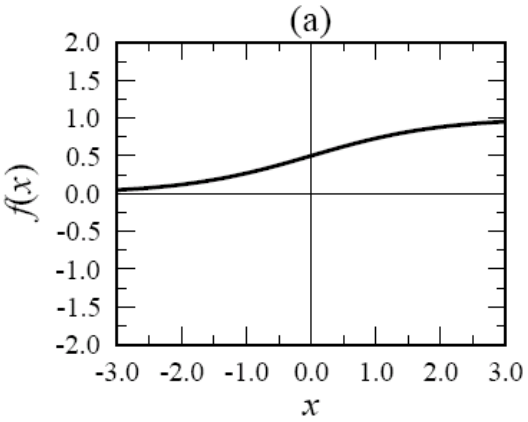


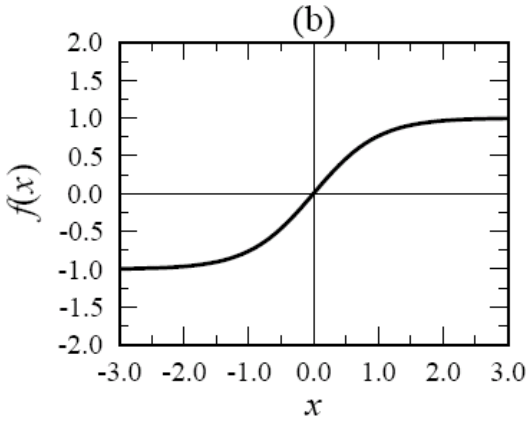
$$E = f_1^3\left(b_1^3 + \sum_{k=1}^{4} a_{k1}^{23} \cdot f_k^2\left(b_k^2 + \sum_{j=1}^{3} a_{jk}^{12} \cdot f_j^1\left(b_j^1 + \sum_{i=1}^{2} a_{ij}^{01} \cdot G_i\right)\right)\right)$$

Activation functions enable the representation of general nonlinear functions.

sigmoid function

$$f(x) = \frac{1}{1+e^{-x}}$$



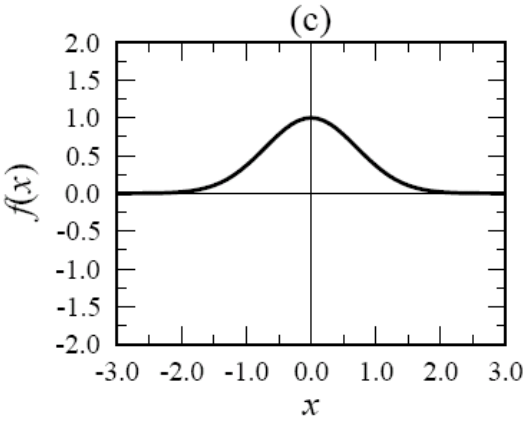


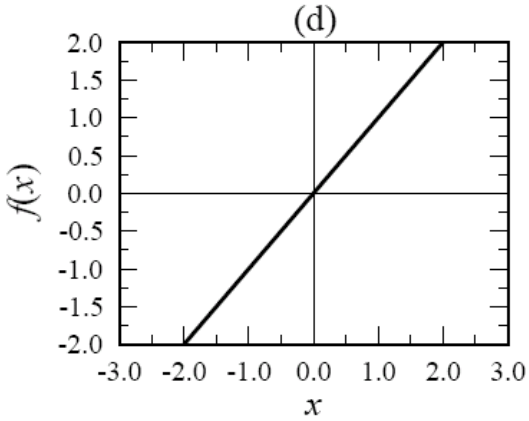
hyperbolic tangent

$$f(x) = \tanh(x)$$

Gaussian function

$$f(x) = e^{-\eta x^2}$$



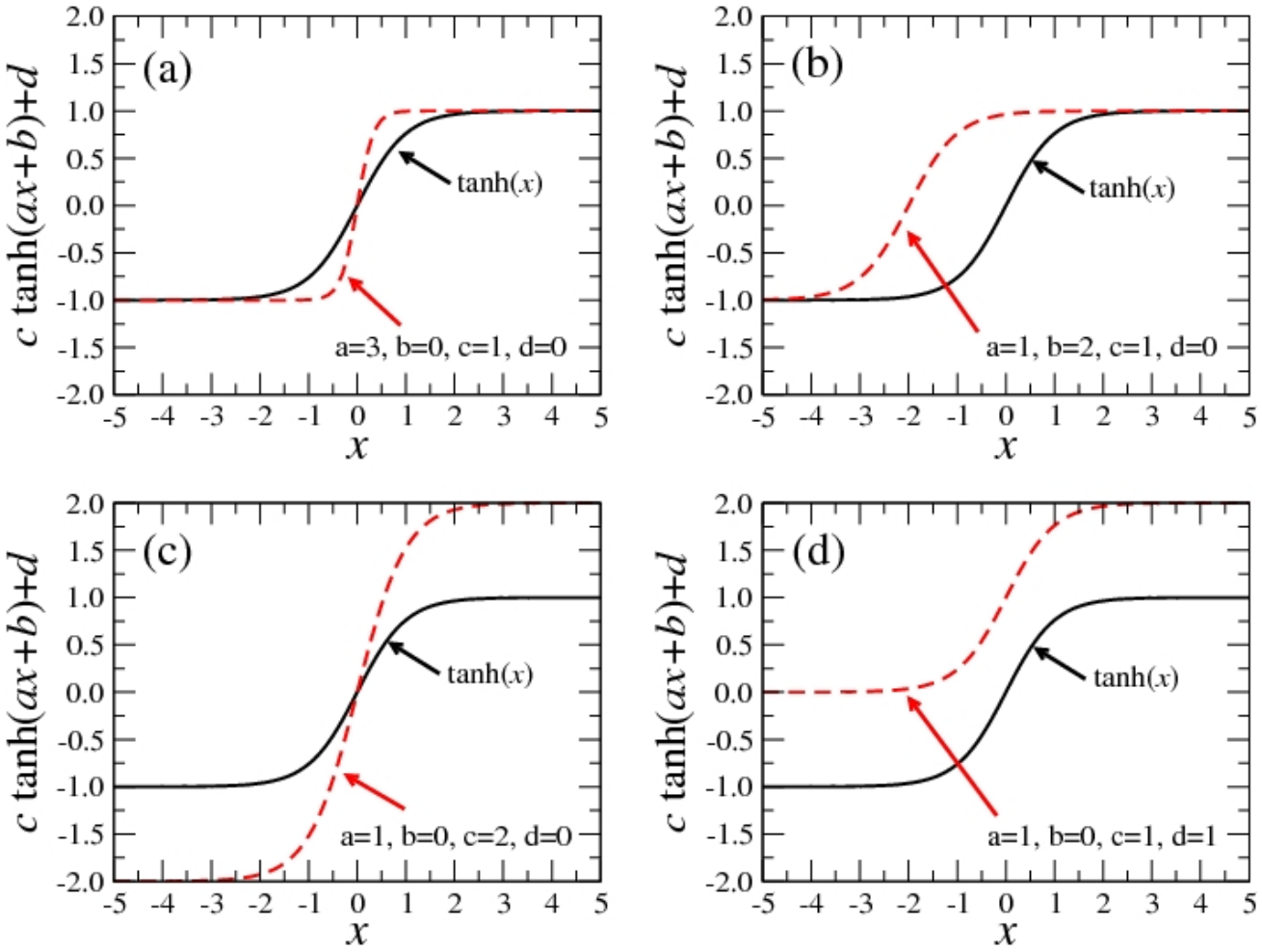


linear function

$$f(x) = x$$

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

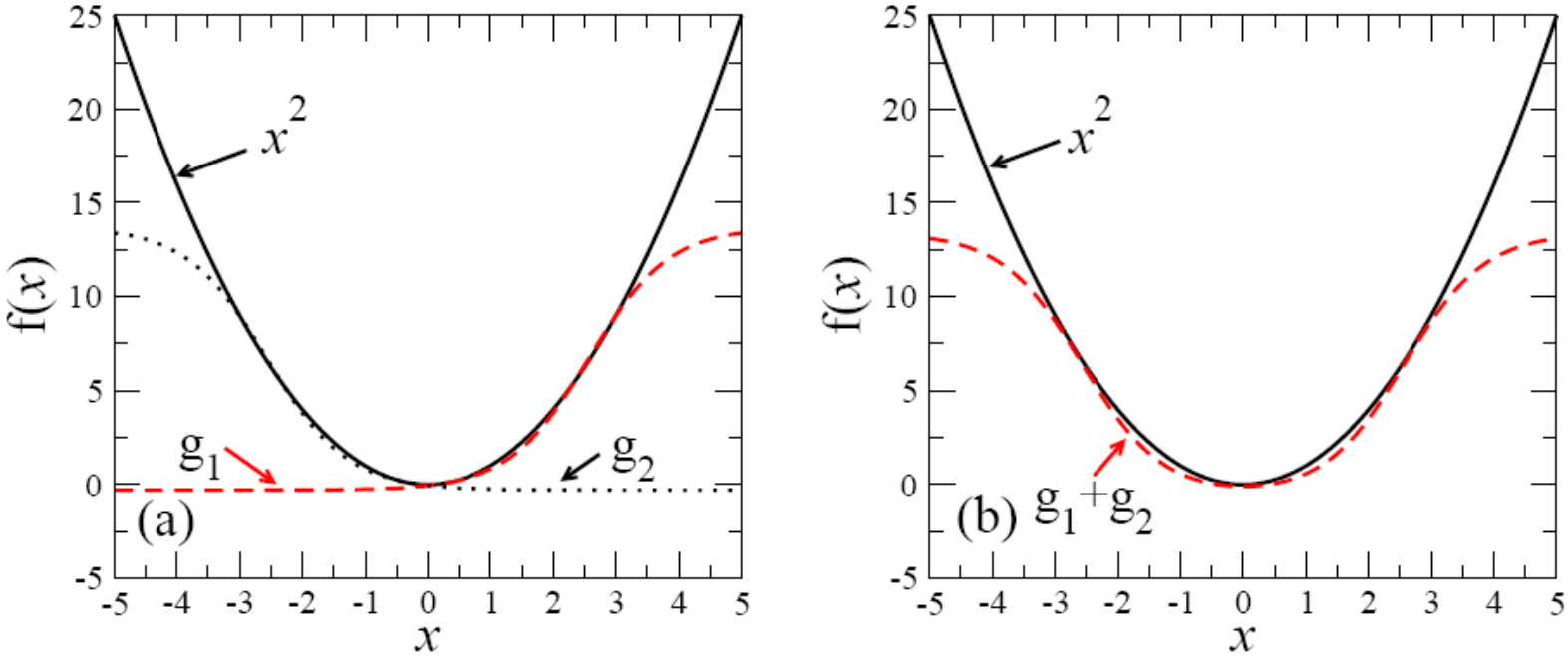
Basic functional element of the NN: $f(x) = c \cdot \tanh(a \cdot x + b) + d$



⇒ very simple but flexible

Model Potential: Harmonic Oscillator

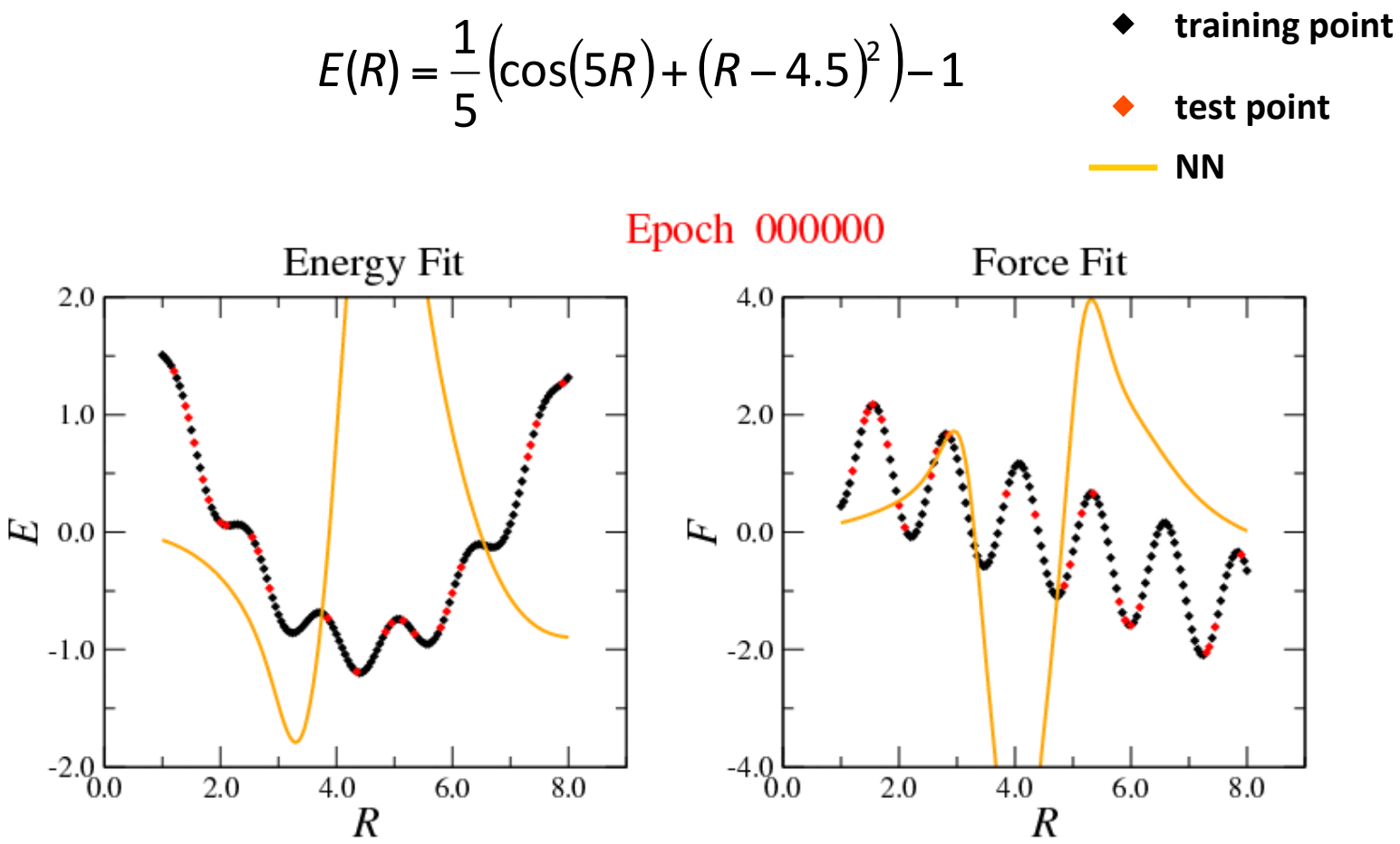
Goal: Represent potential of the 1D harmonic oscillator in the interval [-3 , 3]



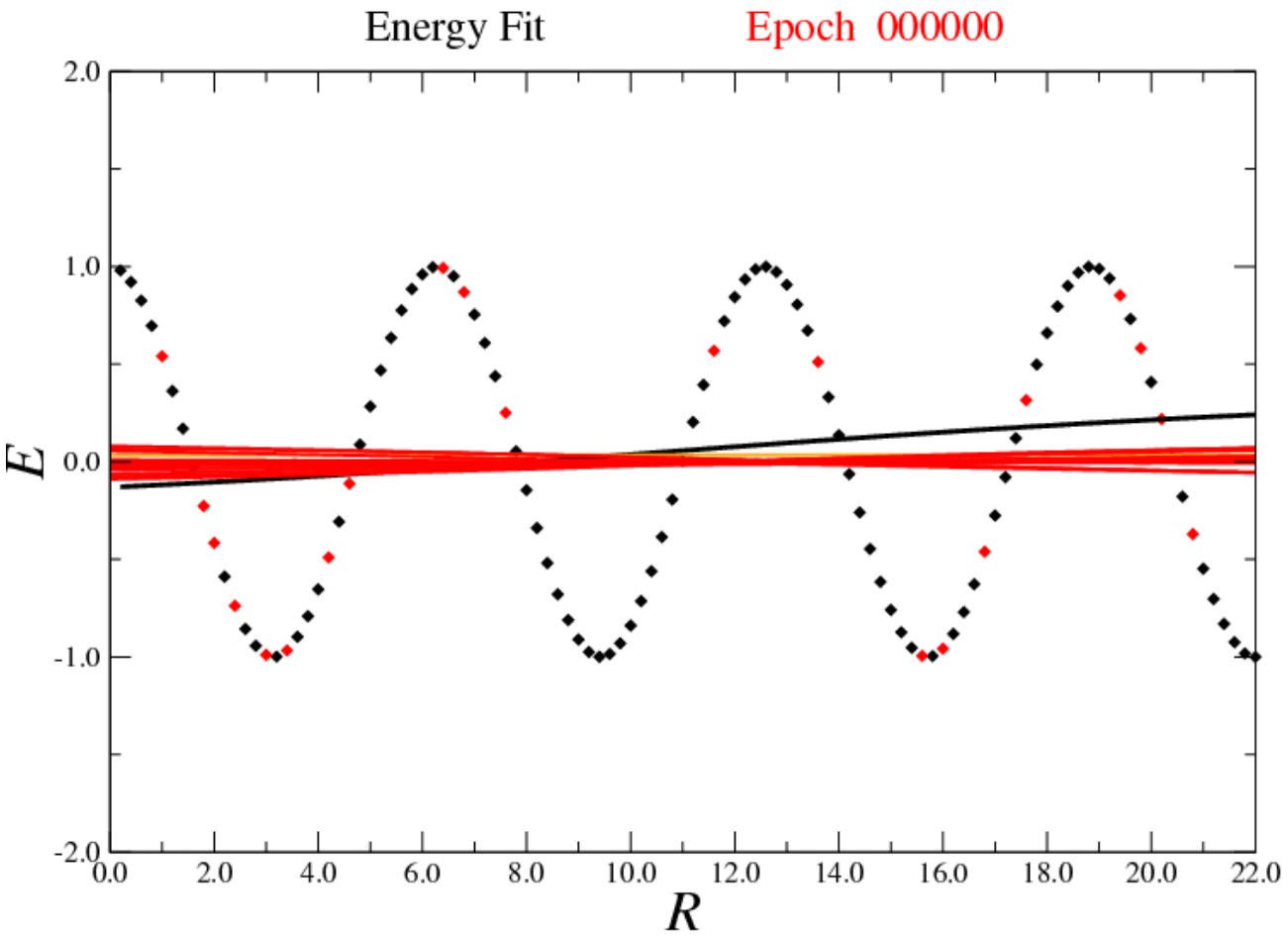
- just two activation functions provide a good approximation
- further improvement possible by adding more functions

Representation of a 1-dimensional model potential function:

$$E(R) = \frac{1}{5}\left(\cos(5R) + (R - 4.5)^2\right) - 1$$



NN based on initial random weight parameters!



bias weight

individual nodes

NN fit

training point

test point

NN architecture: 1-14-1 tl

Overview: Neural Network Potentials (incomplete)

RUB

Neural Network Potentials for Small Molecules:

F.V. Prudente, and J.J.S. Neto, *Chem. Phys. Lett.* **287**, 585 (1998).	H_3^+
H. Gassner et al., *J. Phys. Chem. A* **102**, 4596 (1998).	$H_2O\cdots Al^{3+}\cdots H_2O$
L.M. Raff et al., *J. Chem. Phys.* **122**, 084104 (2005).	CH_2=CHBr
S. Manzhos, and T. Carrington, Jr., *J. Chem. Phys.* **125**, 084109 (2006).	H_2O, HOOH

Molecule – (Frozen) Surface Interactions:

T.B. Blank et al., *J. Chem. Phys.* **103**, 4129 (1995).	H_2@Si, CO@Ni
S. Lorenz, A. Groß, and M. Scheffler, *Chem. Phys. Lett.* **395**, 210 (2004).	H_2@Pd
J. Behler, S. Lorenz, and K. Reuter, *J. Chem. Phys.* **127**, 014705 (2007).	O_2@Al
J. Ludwig, and D.G. Vlachos, *J. Chem. Phys.* **127**, 154716 (2007).	H_2@Pt
D.A.R.S. Latino et al., *J. Electroanal. Chem.* **624**, 109 (2008).	ethanol@Au

Reviews:

C.M. Handley, and P.L.A. Popelier, *J. Phys. Chem. A* **114**, 3371 (2011).

J. Behler, *Phys. Chem. Chem. Phys* **13**, 17930 (2011).

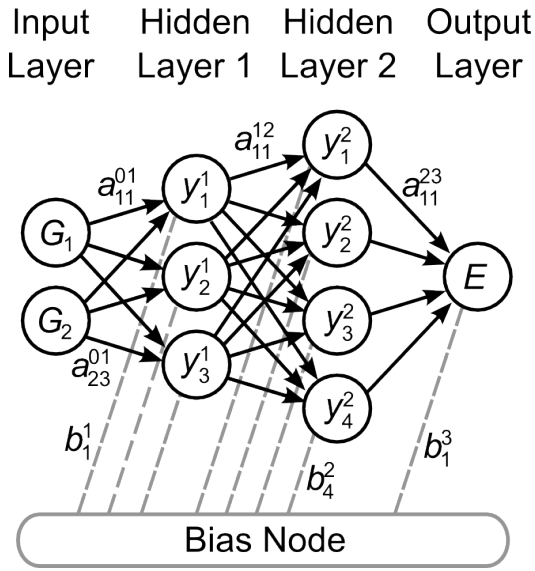
High-Dimensional Neural Network Potentials

High-Dimensional Neural Network Potentials

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 large systems:



- **limited number of dimensions (up to ≈ 12)**
- **permutation symmetry of the system is not included (exchange of atoms changes the energy)**
- **energy generally depends on rotation and translation**
- **potential is valid only for a given system size**

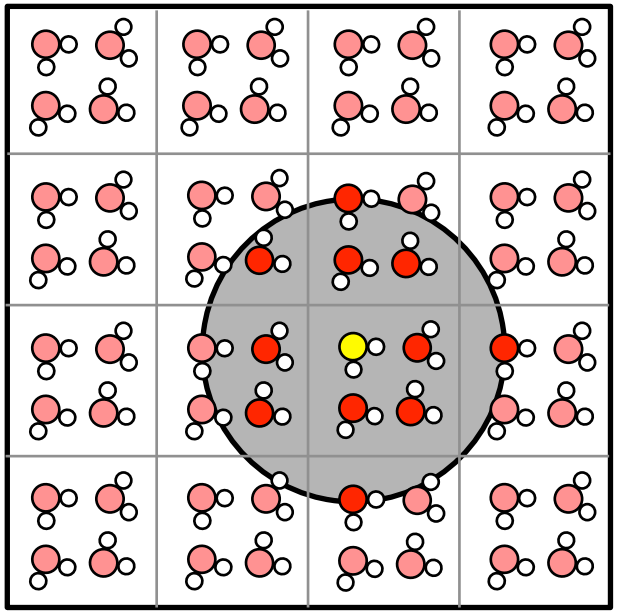
⇒ A new Neural Network scheme is required to deal with high-dimensional systems

High-Dimensional Neural Network Potentials

RUB

Total energy represented as a sum of atomic energies E_i:

$$E = \sum_i E_i$$



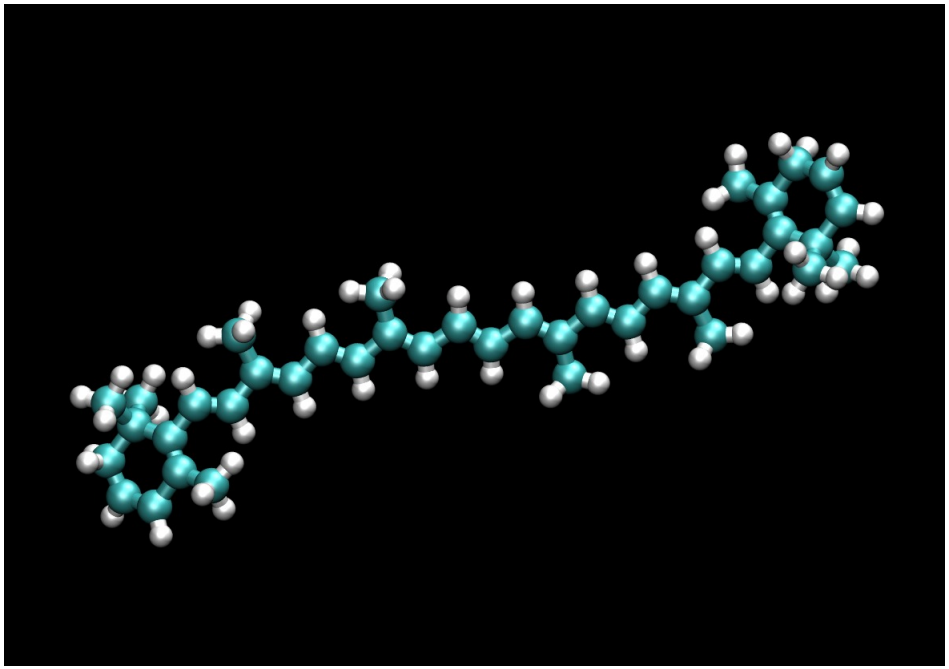
- **separate NN for each atom**
- **E_i depend on chemical environment**
- **environments defined by sufficiently large cutoff $R_c \approx 6$ Å – 10 Å**
- **chemical environment is described by many-body „symmetry functions“**

J. Behler and M. Parrinello, *Phys. Rev. Lett.* 98 (2007) 146401.
J. Behler, *J. Chem. Phys.* 134 (2011) 074106.

Local Atomic Environments

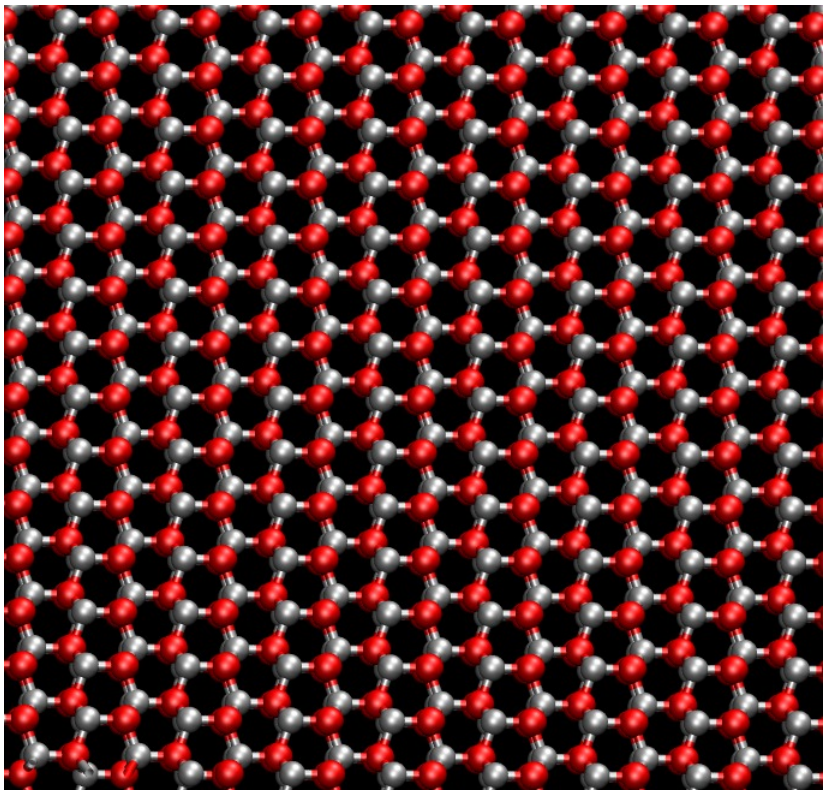
RUB

Molecular Case:



β-carotene

Periodic Case:

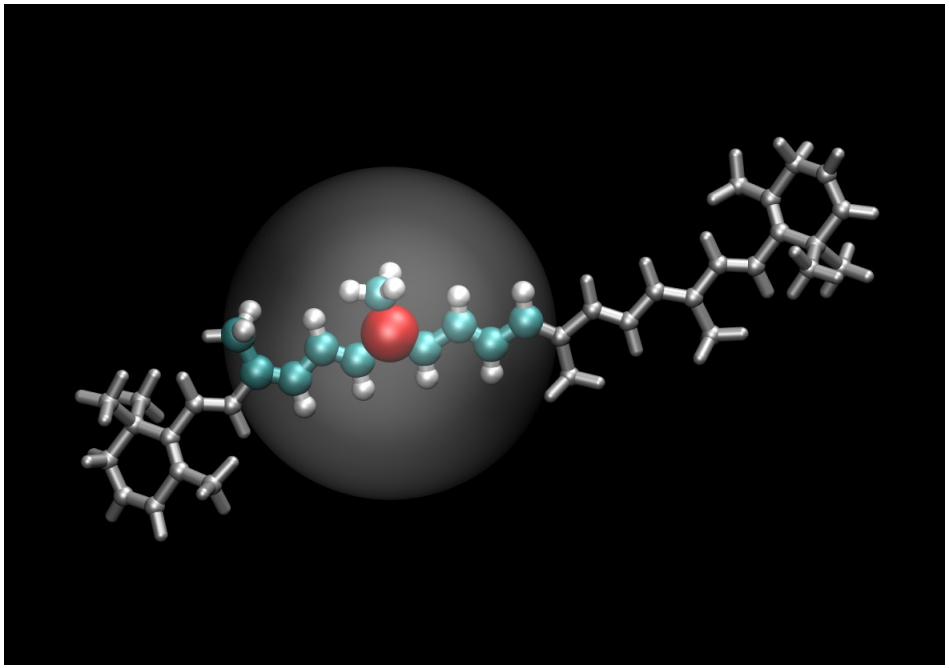


zinc oxide (wurtzite structure)

Local Atomic Environments

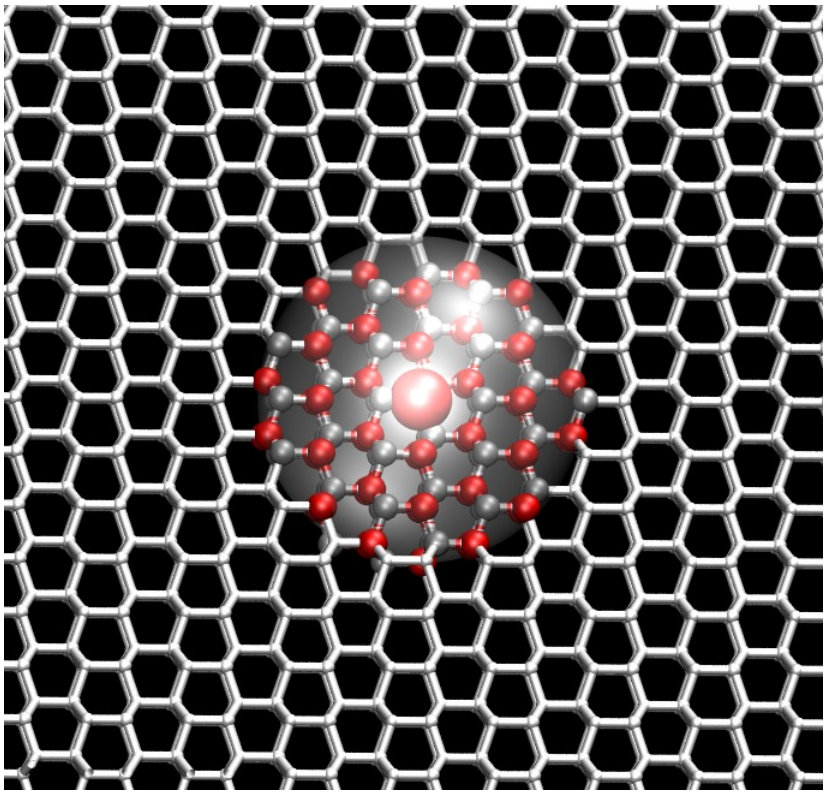
RUB

Molecular Case:



β-carotene

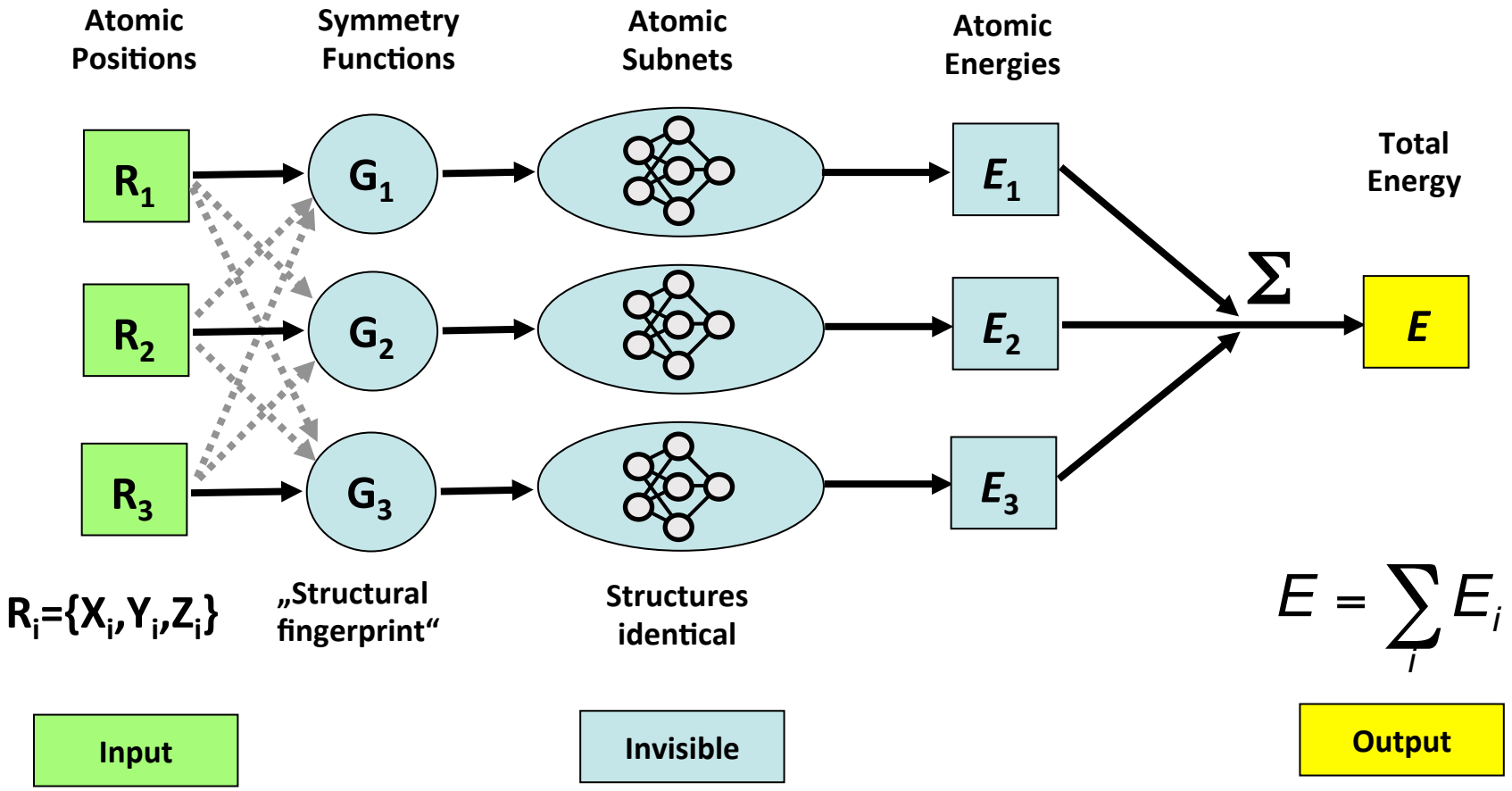
Periodic Case:



zinc oxide (wurtzite structure)

High-Dimensional Neural Network Potentials

Example: 3-Atom System



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

Neural Network Potentials for Multicomponent Systems

RUB

Problem: Interactions are not necessarily short-ranged

Example: Multicomponent-systems: oxides, water

- **significant charge-transfer**
- **long-ranged electrostatic interactions must be included**
- **modification of the Neural Network scheme is needed**

Generalized total energy expression:

$$E_{tot} = E_{short} + E_{elec}$$

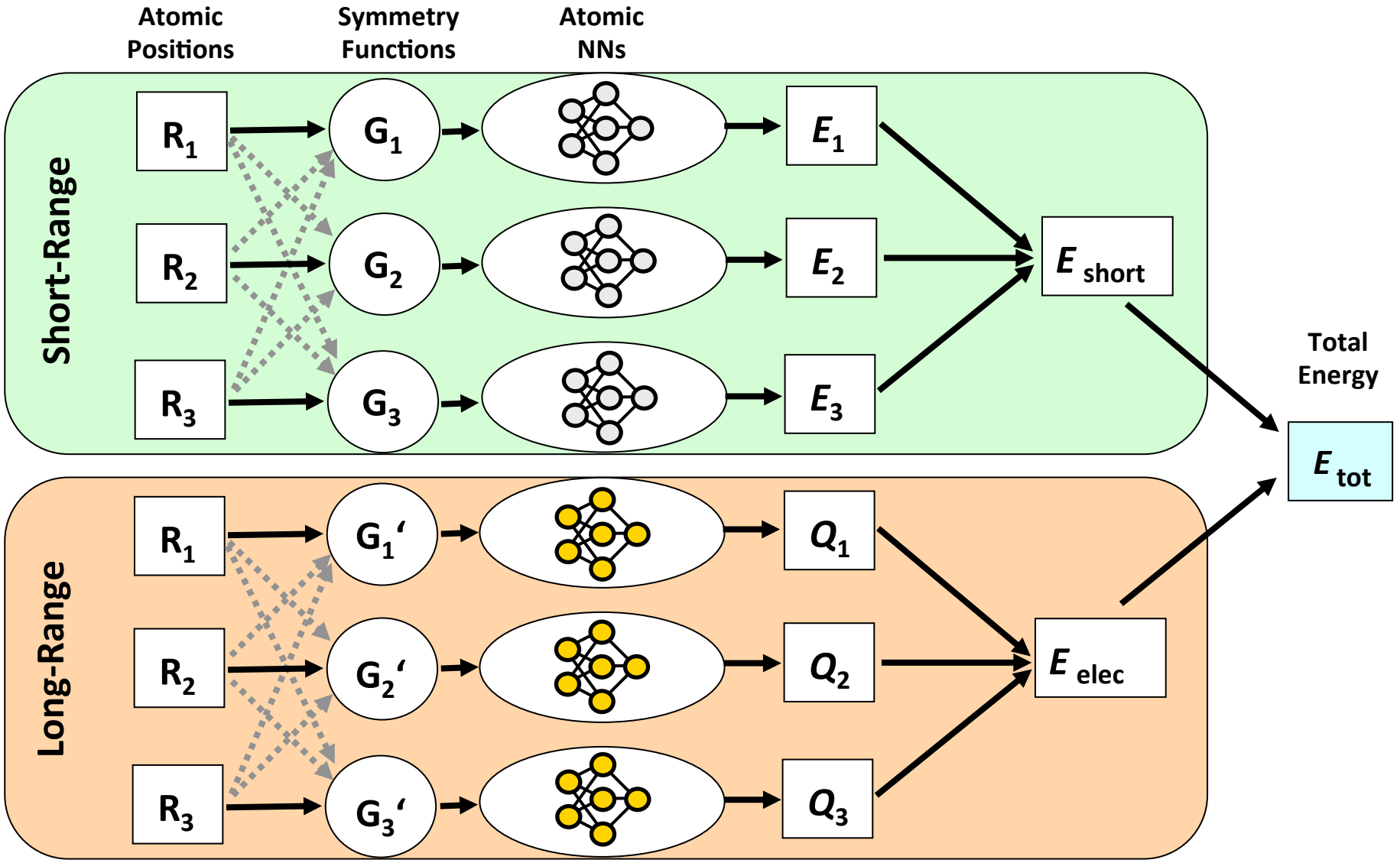
"Short-range" part: atomic energy contributions depending on local environment

Long-range part: atomic charges depending on local environment

- **electrostatic interactions are not truncated**
- **standard methods (Ewald sum) can be used**
- **Neural Network potential applicable to general systems**

N. Artrith, T. Morawietz, and J. Behler, *Phys. Rev. B* 83 (2011) 153101.

Neural Network Potentials for Multicomponent Systems



Symmetry Functions

Symmetry Functions: Basic Properties

RUB

Starting Point: Atomic Structure (Cartesian Coordinates)

Goal: Vector of coordinates $\{G_i\}$ (symmetry functions) with the properties:

- structural description
- translational and rotational invariance
- atomic permutation symmetry included
- number independent of coordination
- zero value and derivative at cutoff radius
- continuous in value and slope

Two-Step Approach

$$\{R_i\} \xrightarrow[\text{to symmetry functions}]{\text{transformation}} \{G_i\} \xrightarrow[\text{functional relation by NN}]{\text{construction of a}} E$$

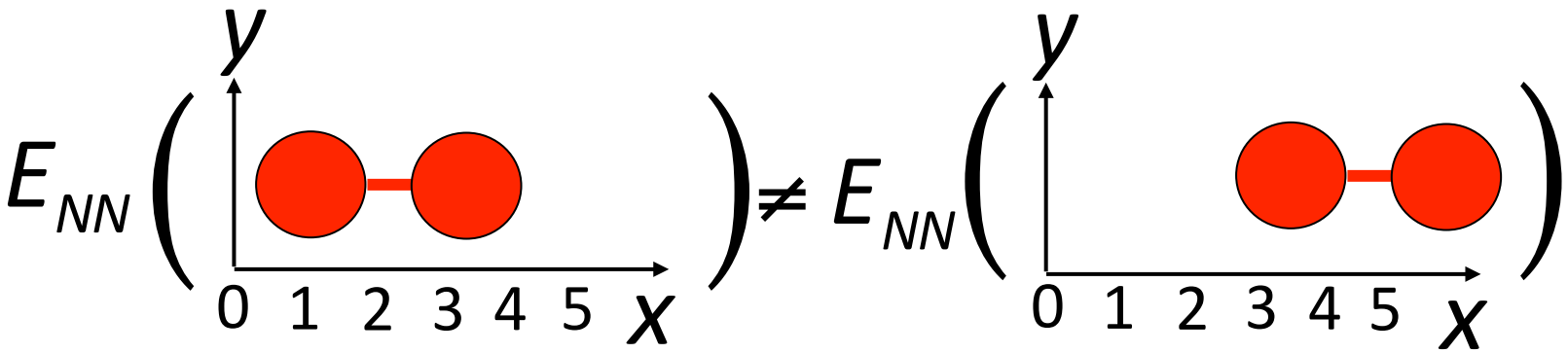
Symmetry Functions: Translational and Rotational Invariance

Problem 1: Cartesian Coordinates

- Neural Networks just process numbers
- absolute values of Cartesian coordinate have no meaning
- only relative atomic positions are important

⇒ Cartesian coordinates cannot be used

Example:



Solution: Use internal coordinates („bond lengths“, angles)

Problem 2: Permutation Symmetry

Example: **Water monomer at finite temperature**

$$E_{NN}\left(\; 1 \;\; 2 \;\; R_{OH,1} \;\; R_{OH,2} \;\right) \neq E_{NN}\left(\; 1 \;\; 2 \;\; R_{OH,1} \;\; R_{OH,2} \;\right)$$

⇒ **Both molecules have a different set of coordinate values**

Solution: Symmetrization step

⇒ **invariant with respect to order of H atoms**

$$G_1 = \left(R_{OH,1} + R_{OH,2}\right)^2 \qquad G_3 = R_{HH}$$

$$G_2 = \left(R_{OH,1} - R_{OH,2}\right)^2$$

⇒ **Now exchanging both H-atoms does not change the NN input vector**

Problem 3: Internal Coordinates

- symmetrization is limited to very small molecules (≈ 4-5 atoms)
- internal coordinates are not unique
- number of coordinates depends on system size

⇒ not applicable to large systems

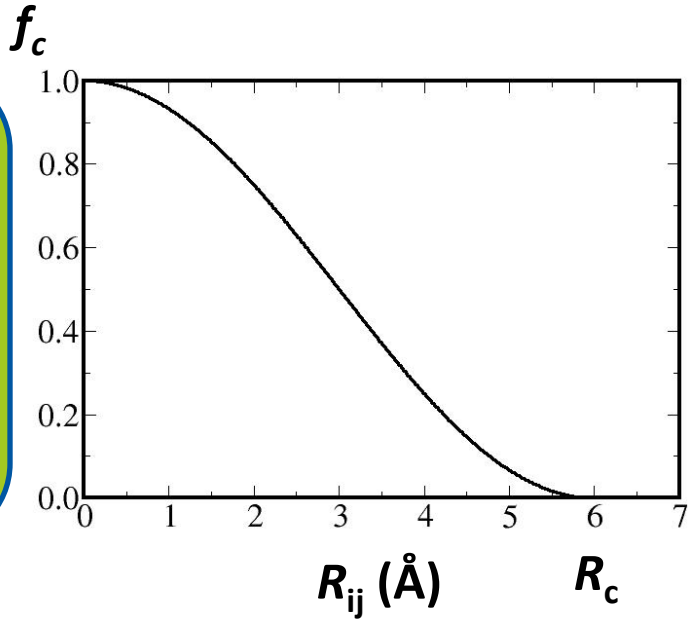
Solution: Local Motifs and Many-Body Symmetry Functions

- describe only close (chemically interacting) atoms
  ⇒ cutoff
- use many-body functions depending on all neighbors
  ⇒ independent of number of neighbors
- translationally and rotationally invariant like internal coordinates
- continuous in value and slope
  ⇒ forces

Cutoff Function

- decays to zero in value and slope at R_c
- reflects decreasing chemical interaction
- central component of all symmetry functions
- R_c is increased until potential converges

$$f_c\left(R_{ij}\right) = \begin{cases} \frac{1}{2}\left[\cos\left(\frac{\pi R_{ij}}{R_c}\right)+1\right] & \text{for} \quad R_{ij} < R_c \\ 0 & \text{for} \quad R_{ij} > R_c \end{cases}$$



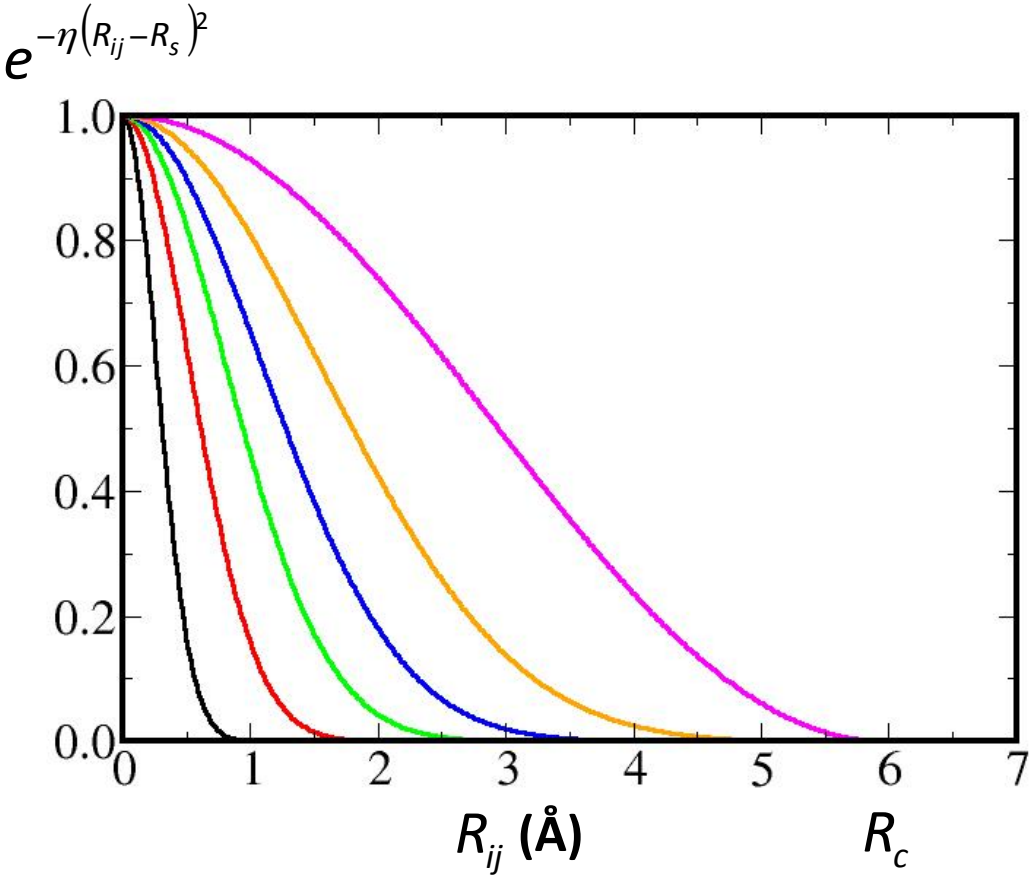
Typical cutoff radius: R_c = 6 - 8 Å

Symmetry Functions: Radial Functions

- decay with increasing distance ⇒ Gaussians
- summation over all neighbors
- many-body term, interpretation as coordination number
- one-to-one correspondence between function value and R_{ij}

$$G_i^{rad} = \sum_j e^{-\eta\left(R_{ij}-R_s\right)^2} f_c\left(R_{ij}\right)$$

Set of radial functions: „Radial Fingerprint“



J. Behler, *J. Chem. Phys.* 134 (2011) 074106.

Crystal Structure Prediction

Crystal Structure Prediction

RUB

„One of the continuing scandals in the physical sciences is that it remains in general impossible to predict the structure of even the simplest crystalline solids from a knowledge of their chemical composition.“

J. Maddox, *Nature* 335 (1988) 201. (editor of *Nature* 1966-73 and 1980-1995)

Challenges in Crystal Structure Prediction:

Identification of Candidate Structures

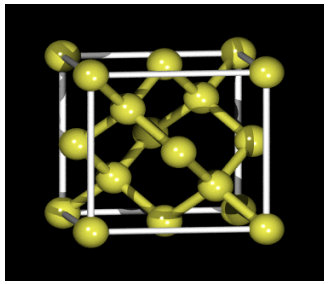
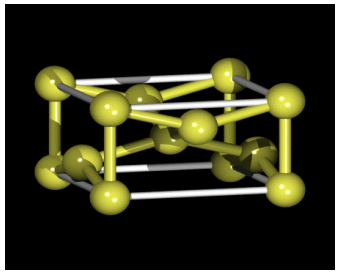
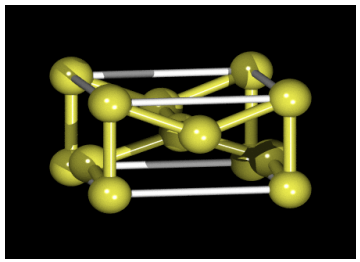
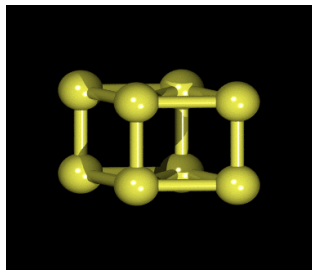
- experiment
- chemical intuition
- molecular dynamics
- genetic algorithms
- metadynamics

Investigation of the Stability

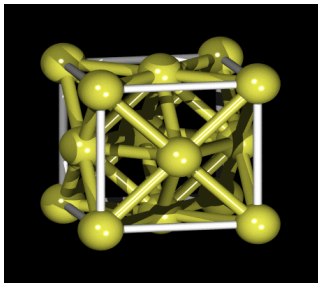
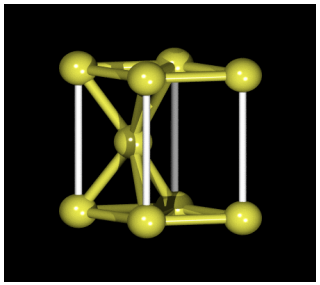
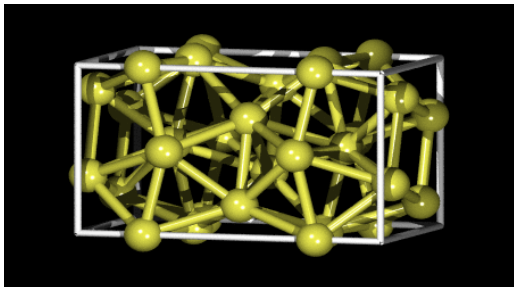
- fast empirical potentials for first evaluation ⇒ often unreliable
- DFT ⇒ too expensive

⇒ Can we use a Neural Network potential?

High-Pressure Phase Diagram of Silicon

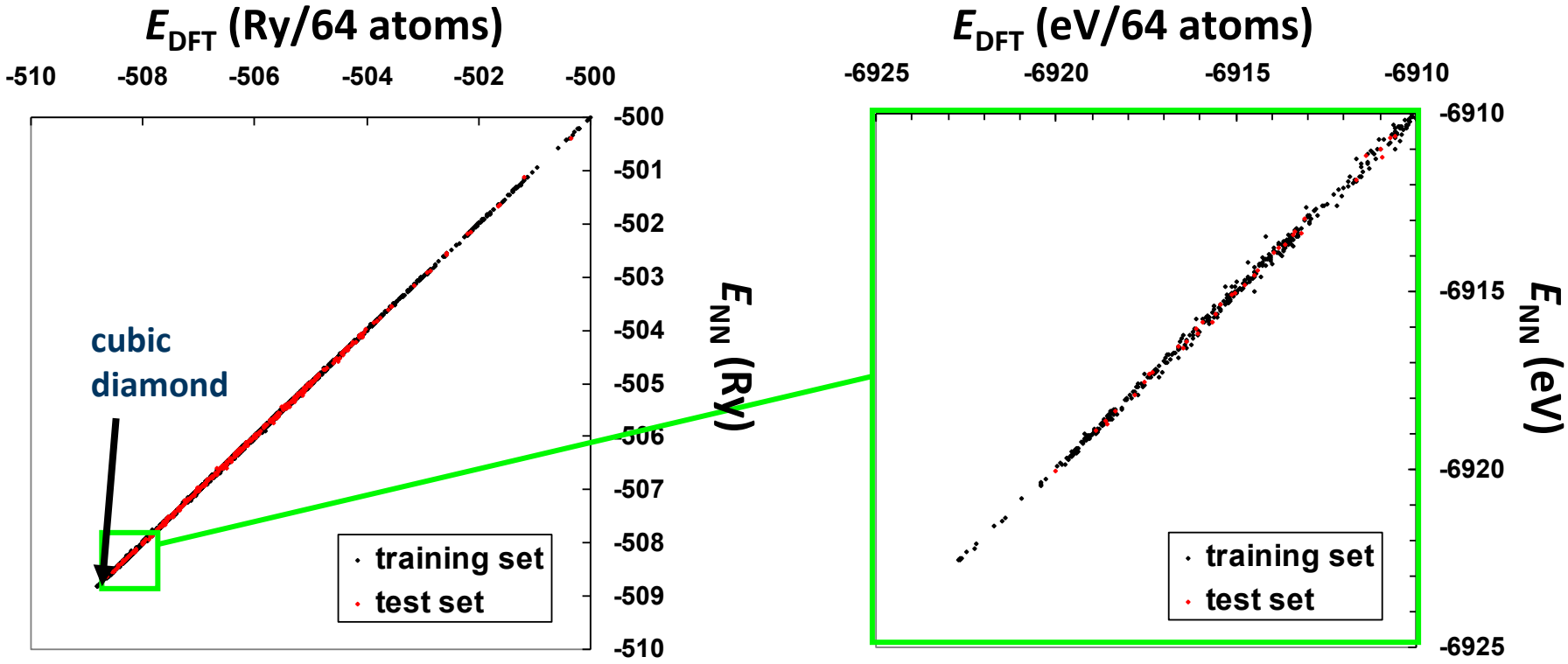
cubic diamond —11.7 GPa→ β-tin —13.2 GPa→ Imma —15.4 GPa→ simple hexagonal —38 GPa→ Cmca —42 GPa→ hcp —79 GPa→ fcc

⇒ Challenging model system to test potentials

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

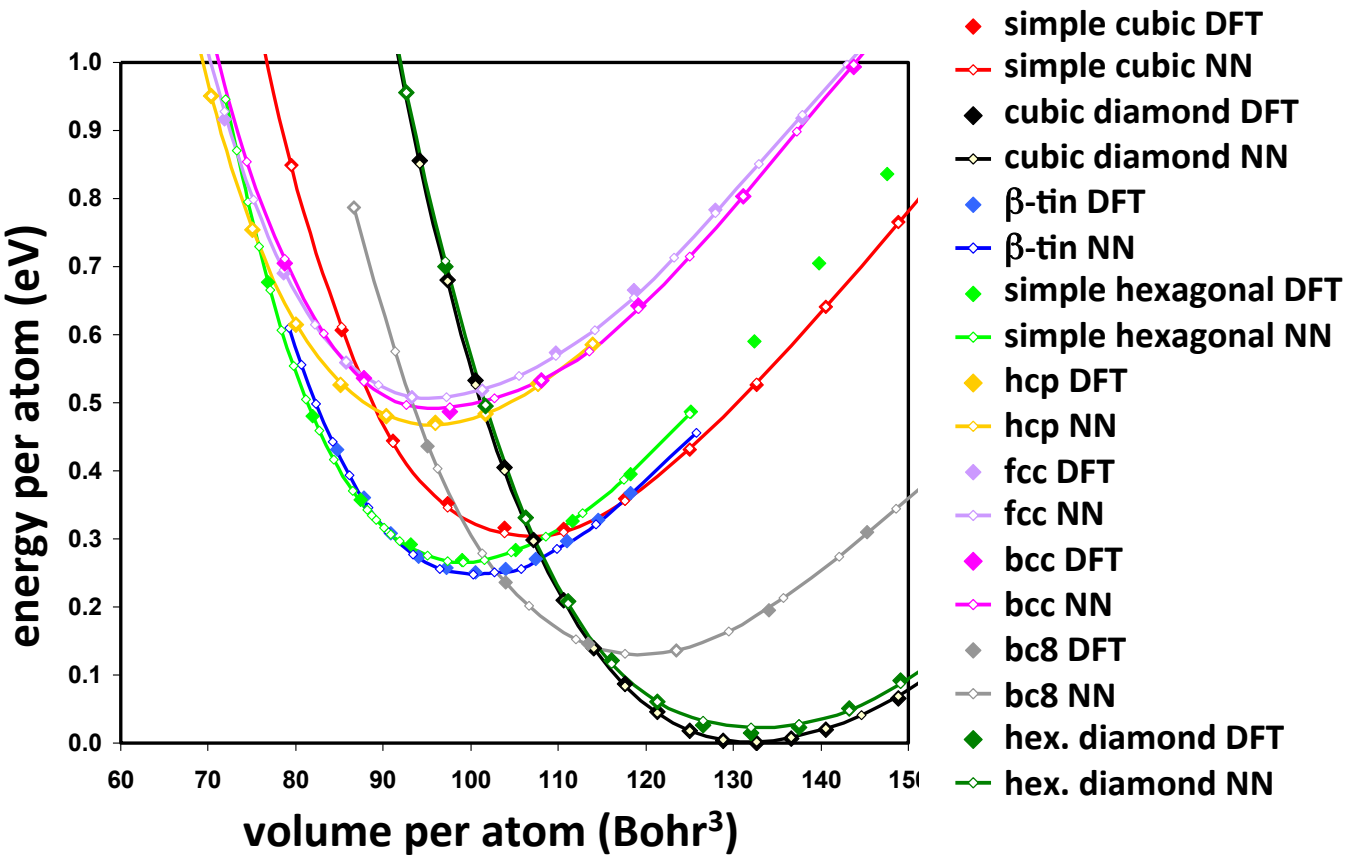
Metadynamics Simulations of Silicon: CPU Time



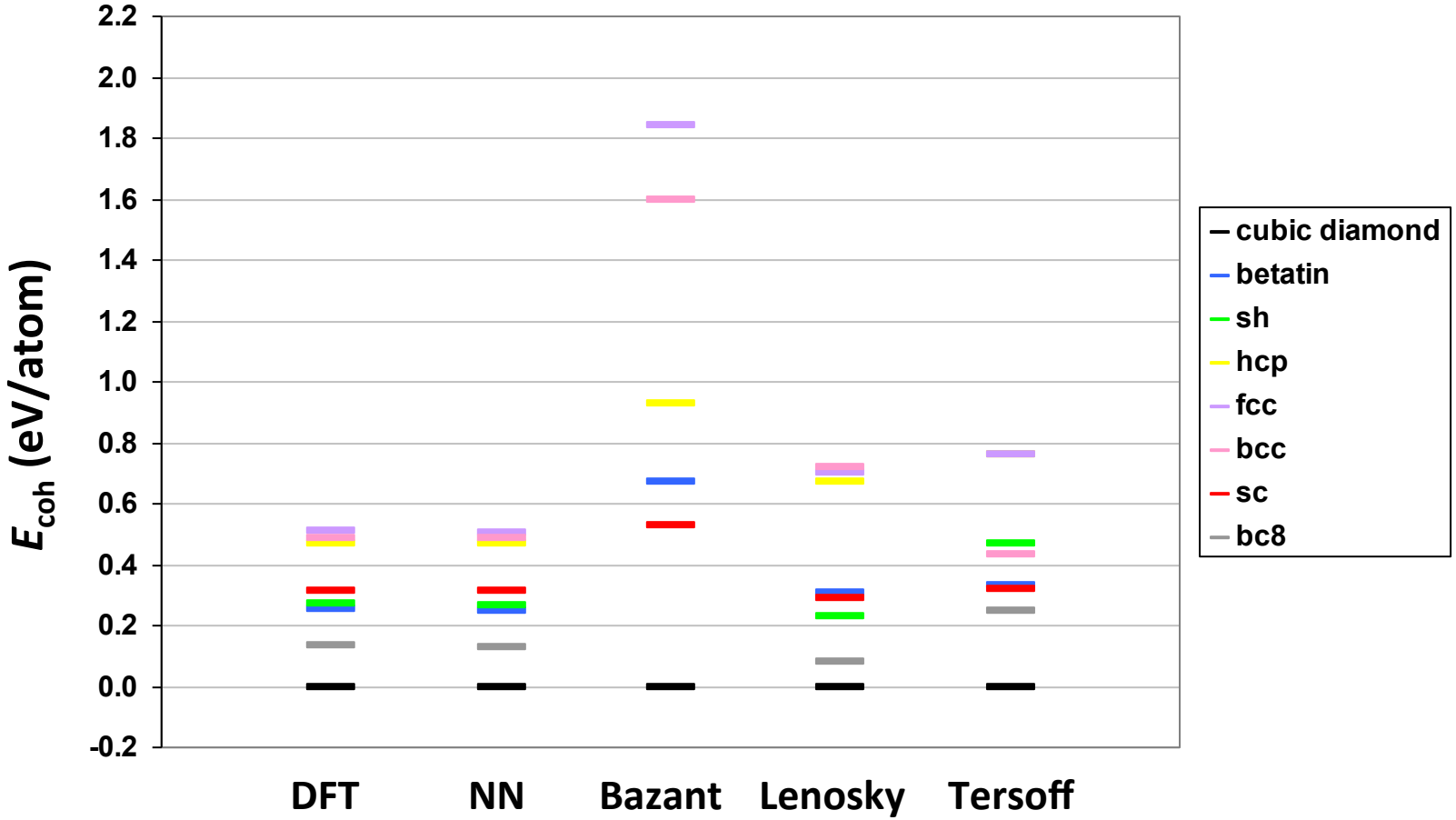
Accuracy

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

Neural Network Potential for Silicon: Equation of State



Silicon: Stability of Crystal Structures



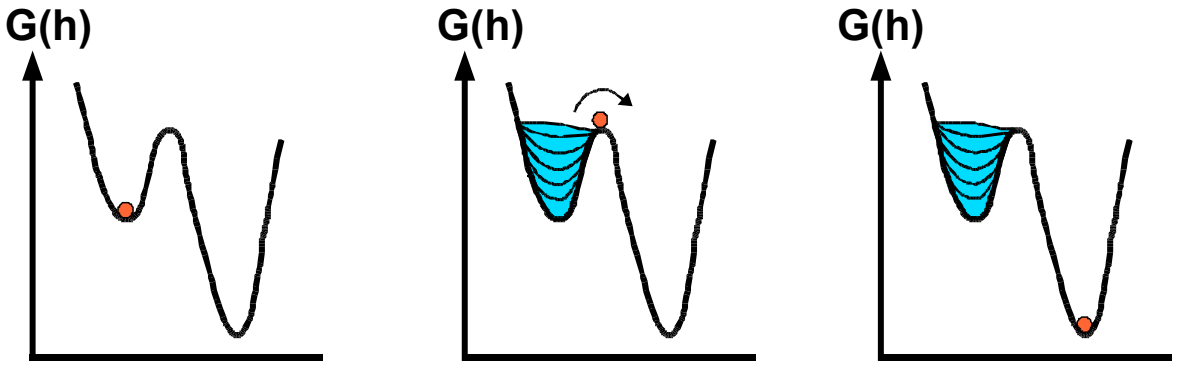
M.Z. Bazant, and E. Kaxiras, *Phys. Rev. Lett.* 77 (1996) 4370.

T.J. Lenosky et al., *Model. Simul. Mater. Sci. Eng.* 8 (2000) 825.

Implementation: S. Goedecker, *Comp. Phys. Comm.* 148 (2002) 124.

J. Tersoff, *Phys. Rev. B* 38 (1988) 9902.

Metadynamics Simulations: The Idea



Evolution of the system:

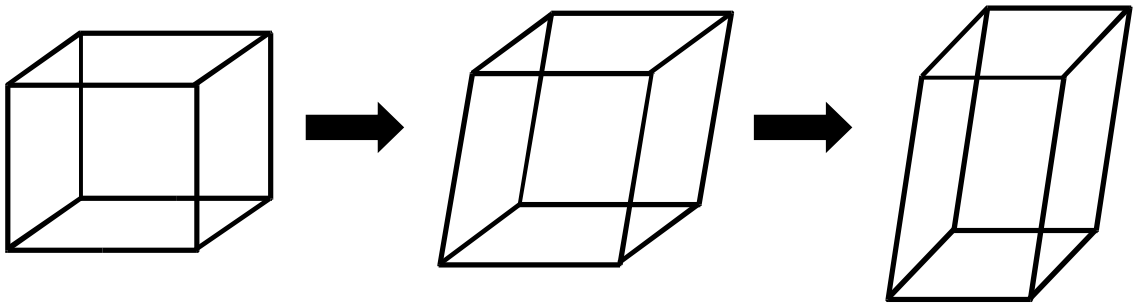
$$\mathbf{h}^{t+1} = \mathbf{h}^{t} + \delta h \frac{\mathbf{F}}{|\mathbf{F}|}$$

Thermodynamic force:

$$F_i = -\frac{\partial G}{\partial h_i} \quad \text{with} \quad -\frac{\partial G}{\partial h_{ij}} = V\left[\mathbf{h}^{-1}\left(-\sigma - p\mathbf{I}\right)\right]_{ji}$$

Force of time-dependent potential:

$$F_i \rightarrow F_i - \frac{\partial}{\partial h_i} \sum_{t'<t} We^{-\frac{\left|\mathbf{h} - \mathbf{h}^{t'}\right|^2}{2\delta h^2}}$$



Collective Variables:

$$\mathbf{h} = \begin{pmatrix} a_x & b_x & c_x \\ 0 & b_y & c_y \\ 0 & 0 & c_z \end{pmatrix}$$

A. Laio, and M. Parrinello, *Proc. Nat. Acad. Sci*. 99 (2002) 12562.

R. Martoňák, A. Laio, and M. Parrinello, *Phys. Rev. Lett*. 90 (2003) 75503.

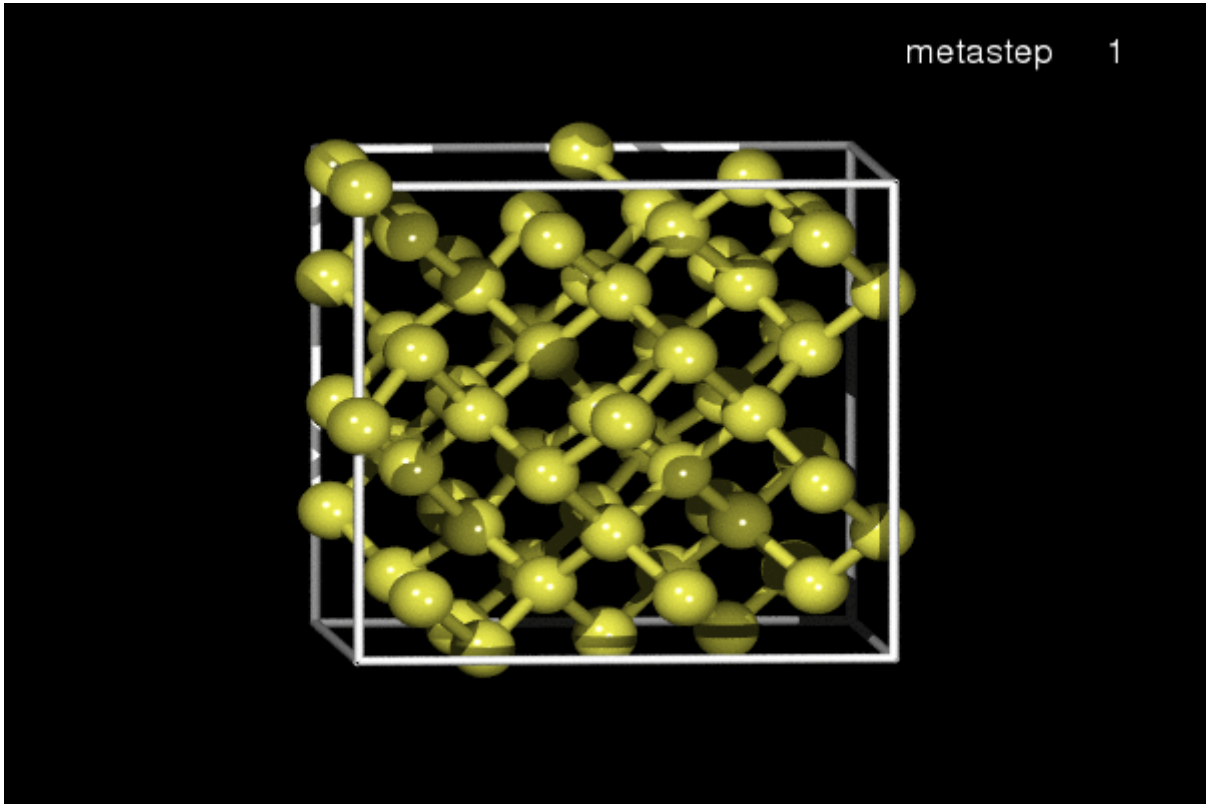
Metadynamics Simulations of Silicon

RUB

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

T **= 300 K,** ***p*** **= 12 GPa**

Movie contains the final frame of each metastep



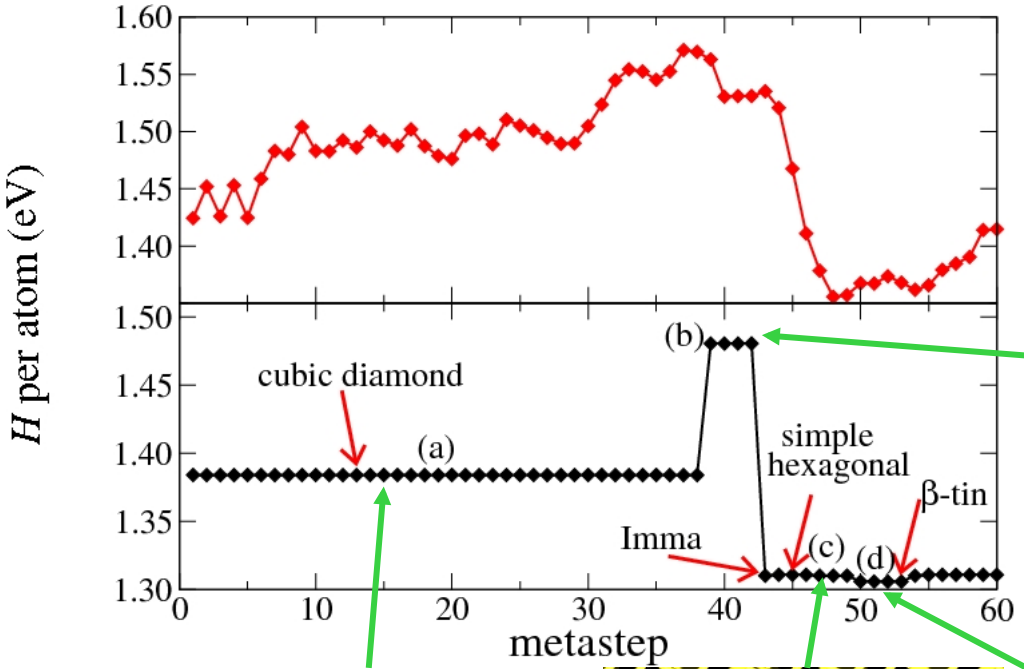
Metadynamics Simulations of Silicon

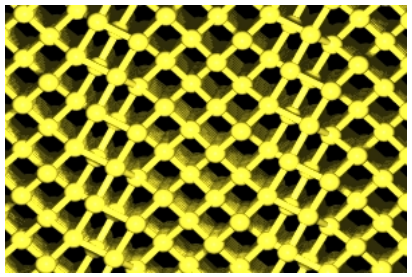
RUB

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

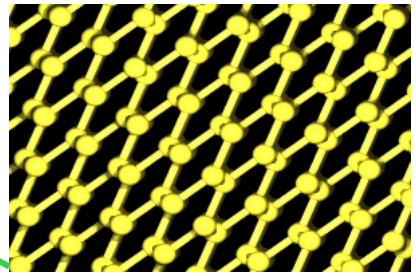
T **= 300 K,** ***p*** **= 12 GPa**

Enthalpy barrier ≈ 0.1 eV/atom



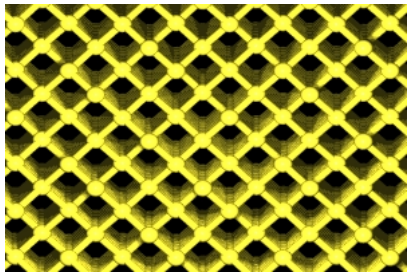


(b)

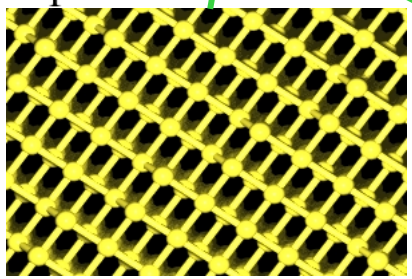


(d)

(a)



(c)



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

Metadynamics Simulations of Silicon

RUB

- NN potential describes phase stabilities with DFT quality
- NN efficiency allows predictive simulations

Computational costs DFT:

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

⇒ ≈ 200 CPU years per simulation

⇒ ≈ 20 000 CPU years per phase diagram

Computational costs NN:

- Construction of the NN PES: ≈ 20 CPU years
- Metadynamics simulations: ≈ 1 CPU day per simulation

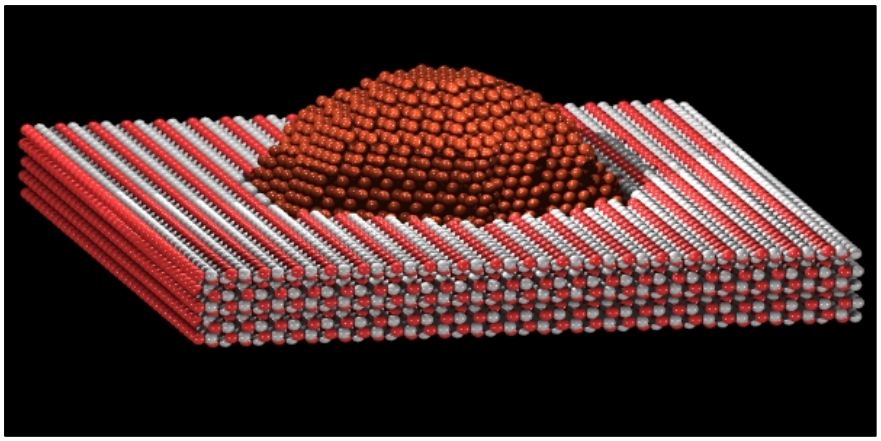
⇒ ≈ 20 CPU years for full phase diagram

Heterogeneous Catalysis

Copper Clusters at Zinc Oxide: Methanol Synthesis

Theory: Required System Size

> 20 000 atoms



Copper clusters contain 5 000 - 50 000 atoms (+ oxide support needed)
⇒ too large for DFT

Key questions:

- **shape of the cluster**
- **chemical composition of the subsystems**
- **surface structure, reactive sites**
- **processes at the Cu/ZnO interface, alloy formation?**

A Neural Network Potential for Copper Surfaces

RUB

Are Neural Network Potentials applicable to metal surfaces?

Data Set for Copper:

	Training	Test
Bulk	13,903	1,545
Slabs	12,483	1,413
Clusters	7,577	842
Total	**33,963**	**3,800**

⇒ 617,475 atomic environments

⇒ 1,852,425 pieces of information

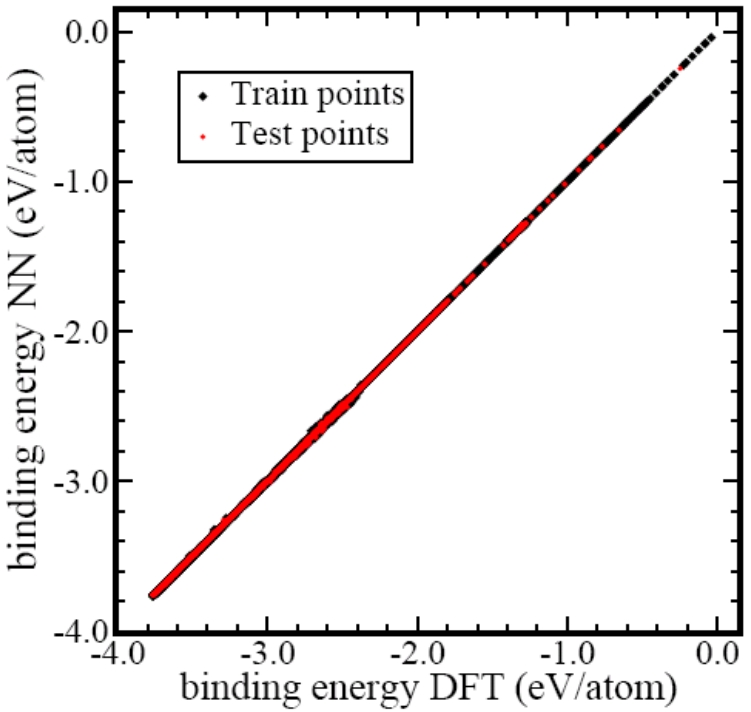
RMSEs:

	E (eV/ atom)	*F* (eV/ Bohr)
Training	0.0036	0.0428
Test	0.0039	0.0420

DFT Code: FHI-aims, PBE XC

V. Blum *et al.*, *Comp. Phys. Comm*. 180 (2009) 2175.

***E*(NN) vs. *E*(DFT)**



NN architecture: 51-30-30-1 *ttl*

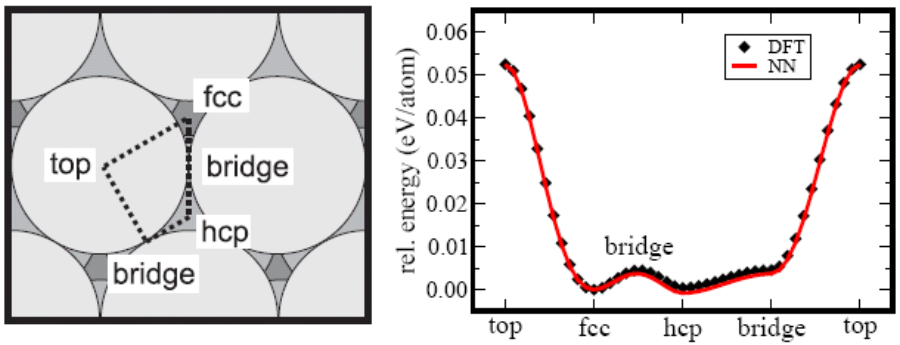
A Neural Network Potential for Copper Surfaces

RUB

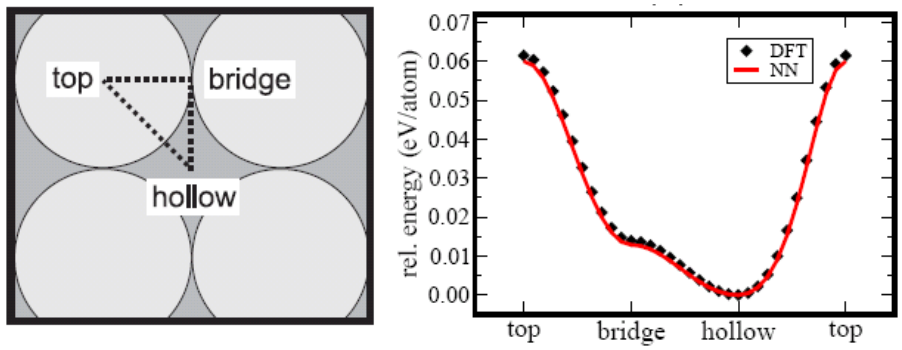
Bulk Copper:

	E_{coh} eV/atom		Lattice parameters Å		B GPa	
Structure	**DFT**	**NN**	**DFT**	**NN**	**DFT**	**NN**
fcc	3.763	3.756	a=3.630	a=3.630	140	138
bcc	3.719	3.716	a=2.885	a=2.887	137	135
sc	3.281	3.282	a=2.407	a=2.407	103	108
hcp	3.740	3.740	a=4.862 c/a=1.63	a=4.856 c/a=1.63	-	-

Cu(111):



Cu(100):

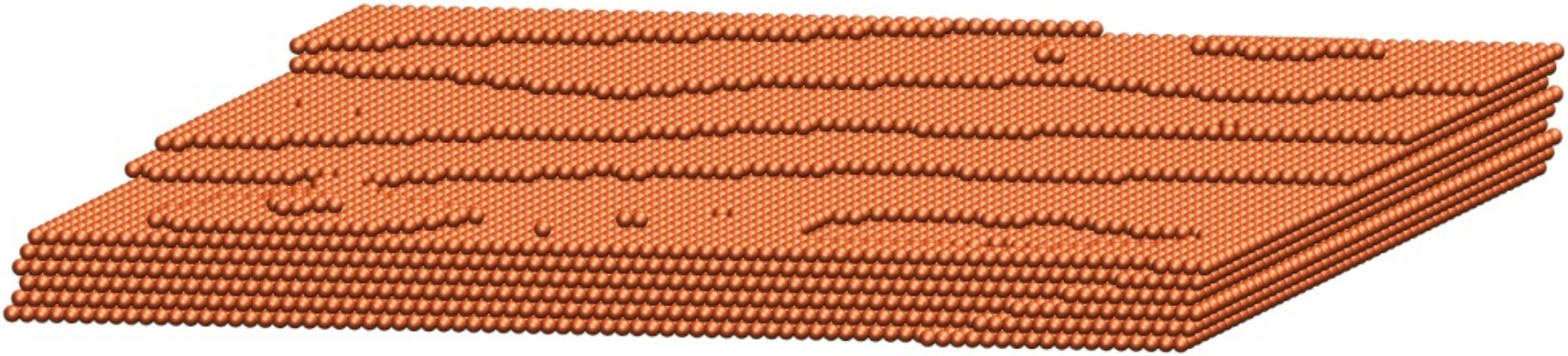


N. Artrith and J. Behler, *Phys. Rev. B* 85 (2012) 045439.

A Neural Network Potential for Copper Surfaces

RUB

Model of a "real" surface with steps, kinks, and adatoms (≈ 29 000 atoms).



DFT ⇒ impossible
NN ⇒ ≈ 2 mins (single core)

⇒ NN potentials can be used to study systems of this size.

How do we know if the potential is correct?

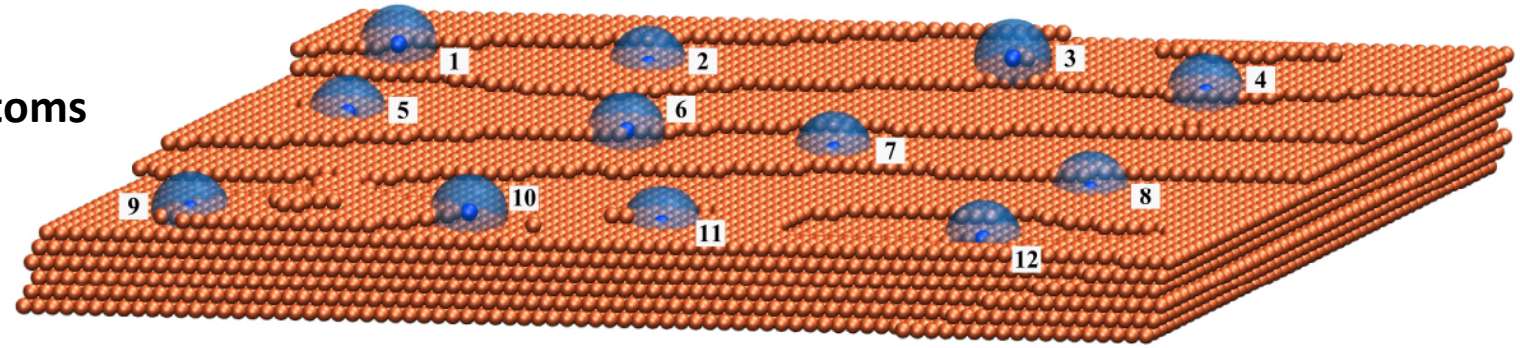
N. Artrith and J. Behler, *Phys. Rev. B* 85 (2012) 045439.

A Neural Network Potential for Copper Surfaces

RUB

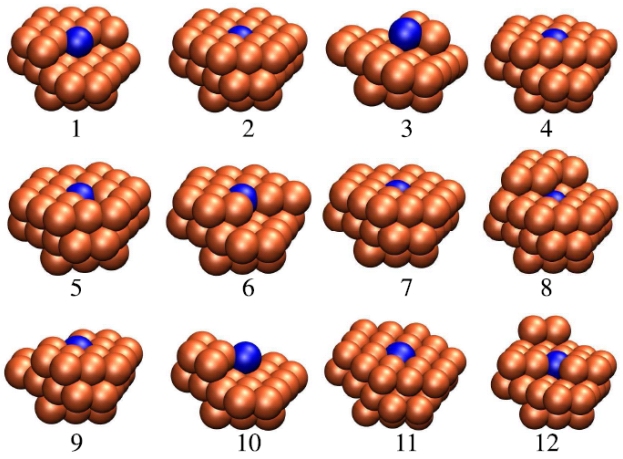
A "Real" Copper Surface:

29 000 atoms

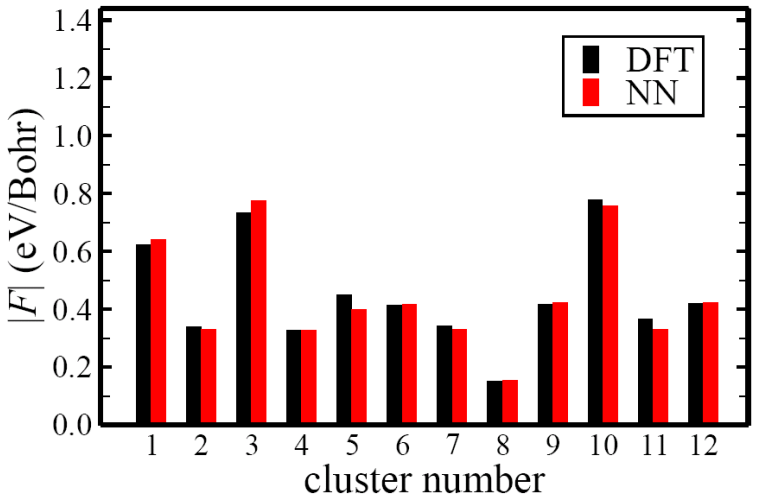


We need local properties related to the potential energy surface ⇒ forces

Clusters:



Forces at central atoms in clusters:



N. Artrith and J. Behler, *Phys. Rev. B* 85 (2012) 045439.

Neural Network Potential for Zinc Oxide

Crystal Structures of Zinc Oxide

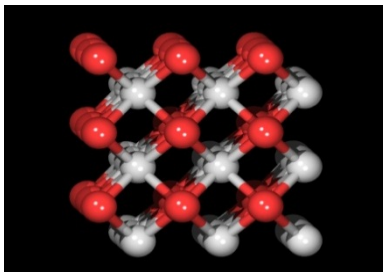
Training points: 35 000

RMSEs Energy = 0.002 eV/atom
Charges = 0.009 *e* /atom
Forces = 0.162 eV/Bohr

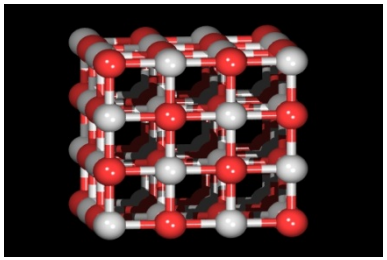
Testing points: 3 800

RMSEs Energy = 0.003 eV/atom
Charges = 0.009 *e* /atom
Forces = 0.161 eV/Bohr

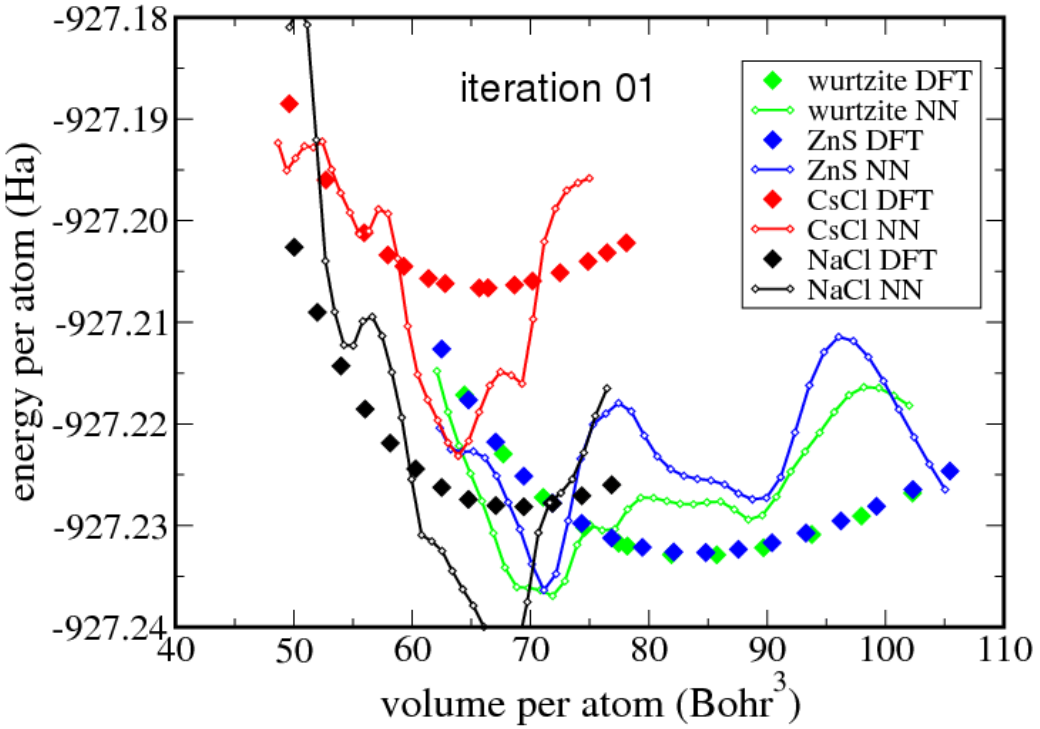
CsCl



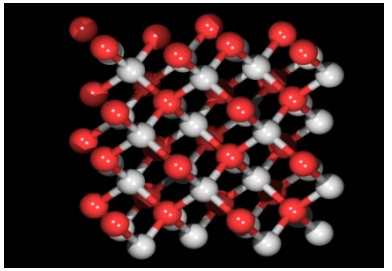
NaCl



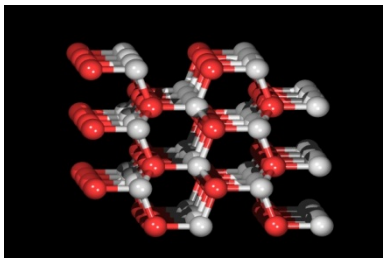
Iterative Optimization:



zinc blende

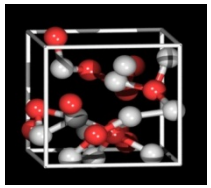


wurtzite

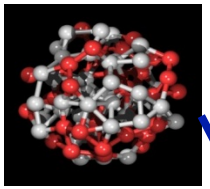


N. Artrith, T. Morawietz, and J. Behler, *Phys. Rev. B* 83 (2011) 153101.

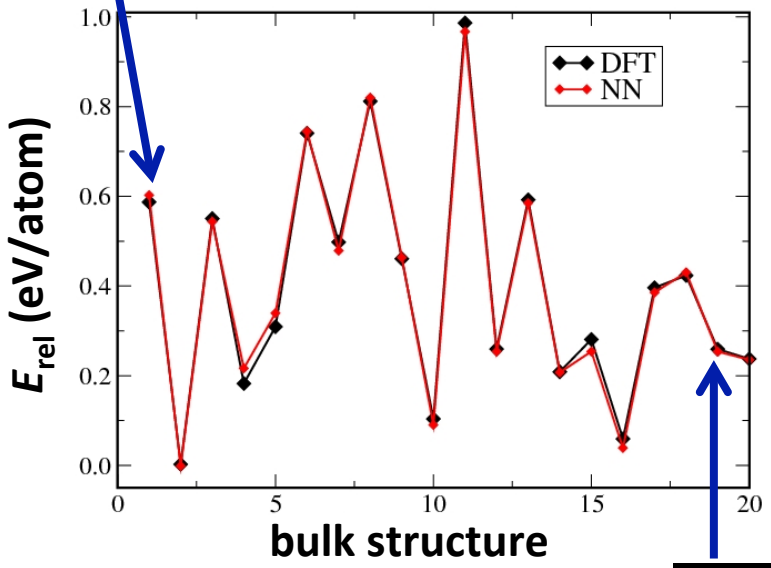
Neural Network Potential for Zinc Oxide

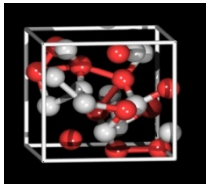


Comparison DFT - NN:

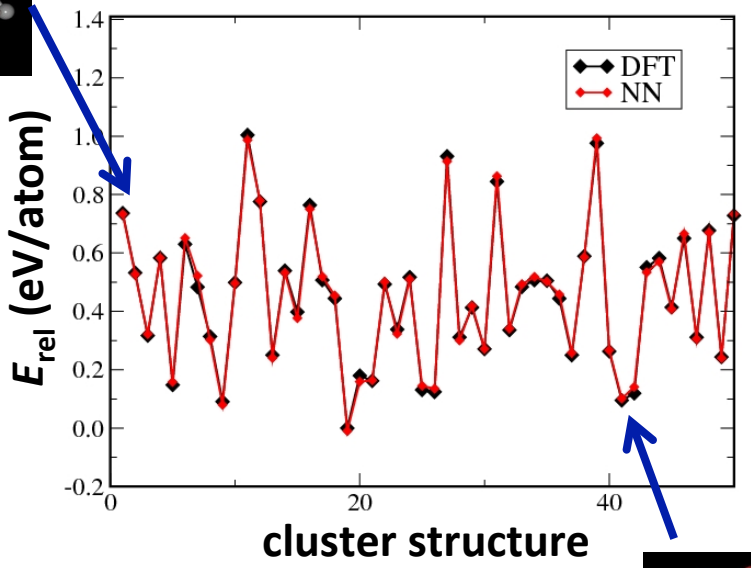


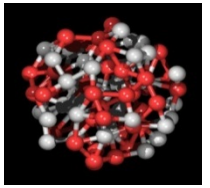
Bulk: $Zn_{12}O_{12}$





Clusters: $Zn_{40}O_{40}$





Neural Network Performance:

Example: 1000 atom system: 10 s CPU time on standard-PC (energy + forces)

Accuracy of the NN Potential for the Ternary System

RUB

From monocomponent and binary systems to CuZnO:

		Energy eV / atom		Force eV / Bohr	
System	Structures	train	test	train	test
copper	40 000	0.004	0.004	0.043	0.042
ZnO	40 000	0.002	0.003	0.161	0.162
CuZnO	100 000	0.005	0.005	0.094	0.089

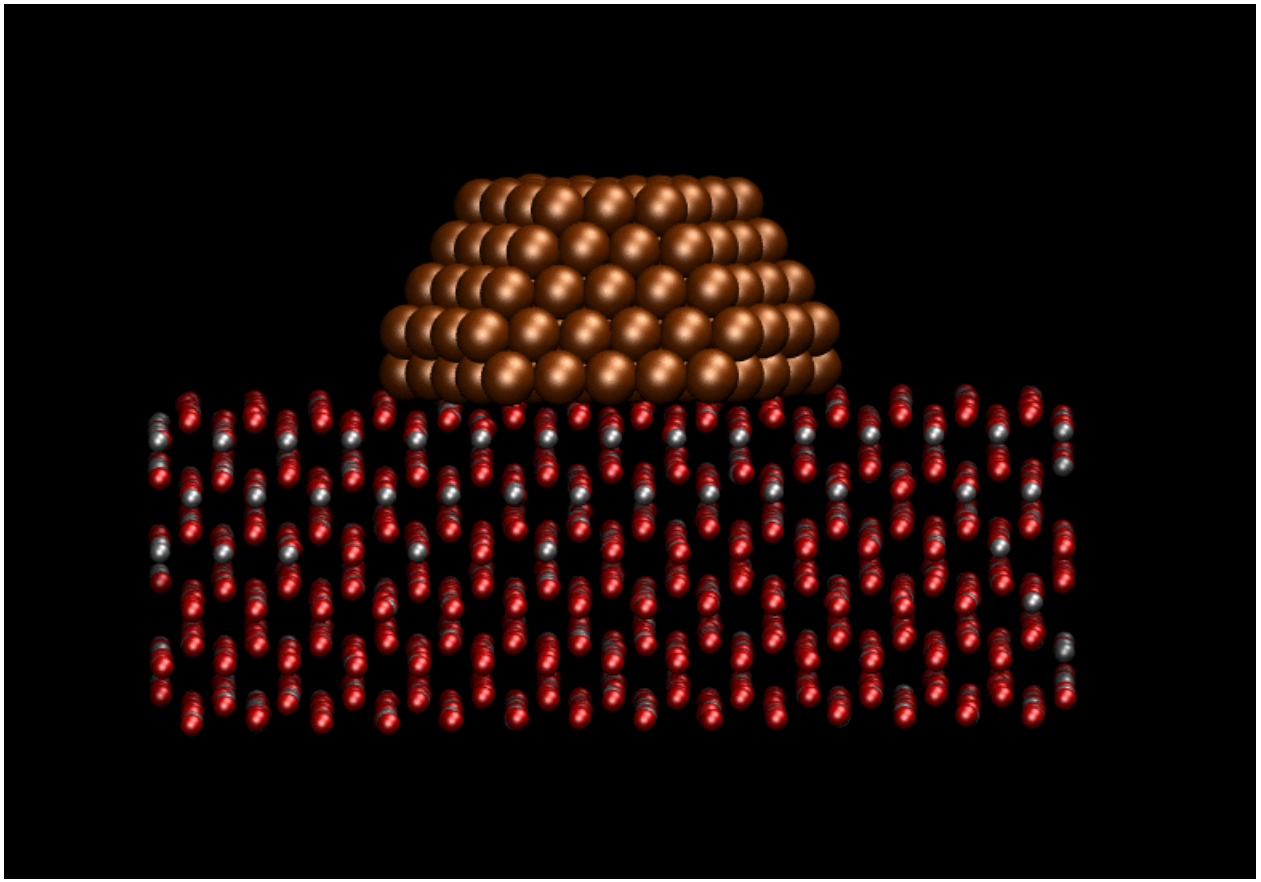
Ternary NN architecture: 156-15-15-15-1

⇒ NN potential for ternary system has the same quality as for pure Cu or ZnO!
⇒ Limit of NNs has not yet been reached

Copper Clusters at Zinc Oxide

RUB

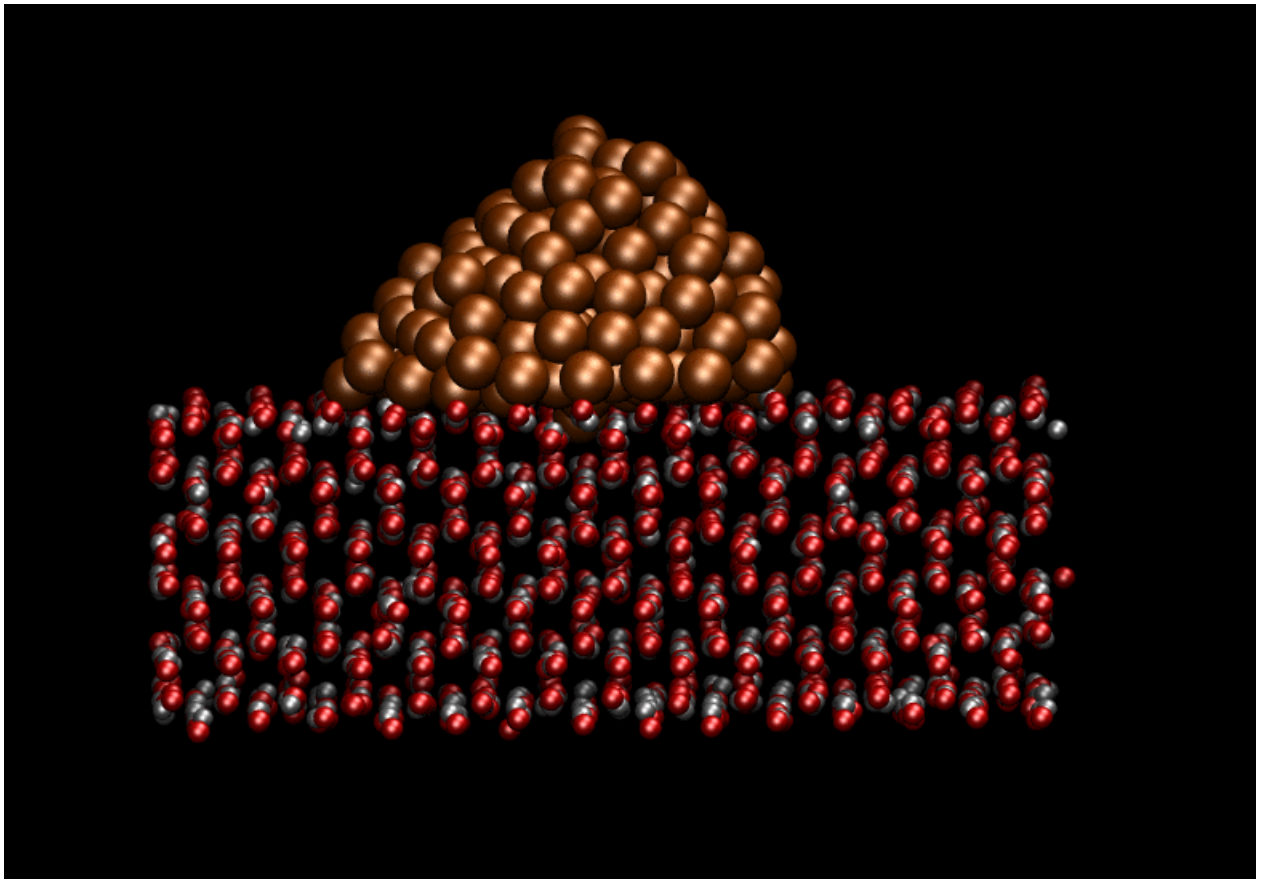
Cu(111)/ZnO(10-10)



233 copper atoms
1344 ZnO formula units

⇒ 2921 atoms

Copper Clusters at Zinc Oxide



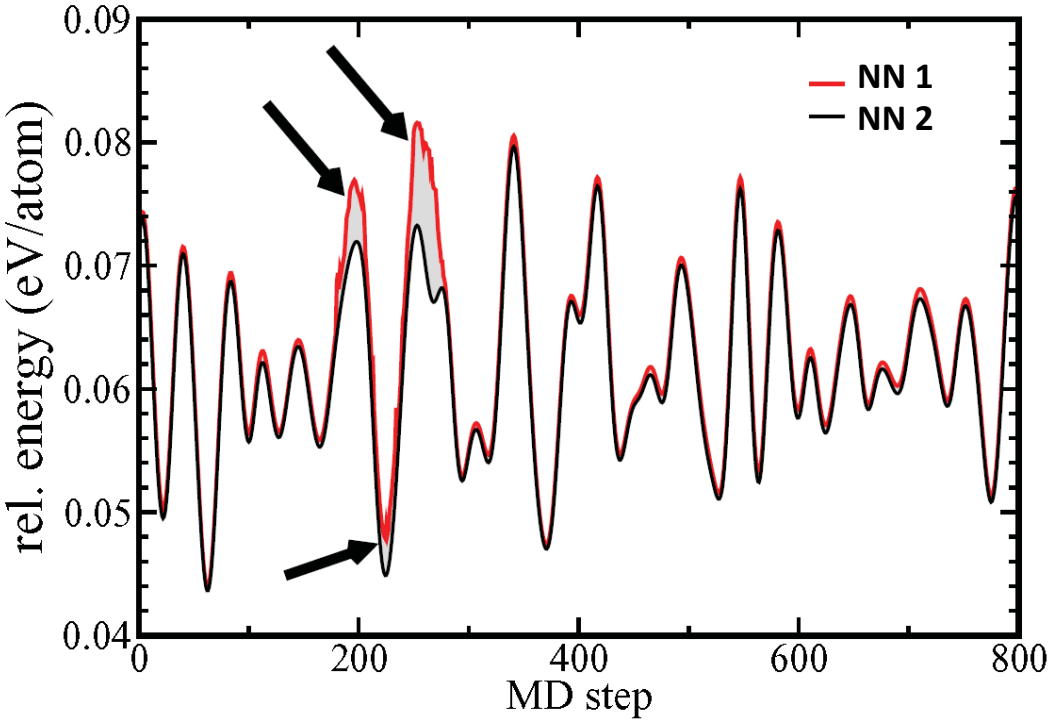
⇒ MD simulations allow to improve the potential systematically

Systematic Improvement

Neural Network Potentials must be reliable for structures NOT included in the training set.

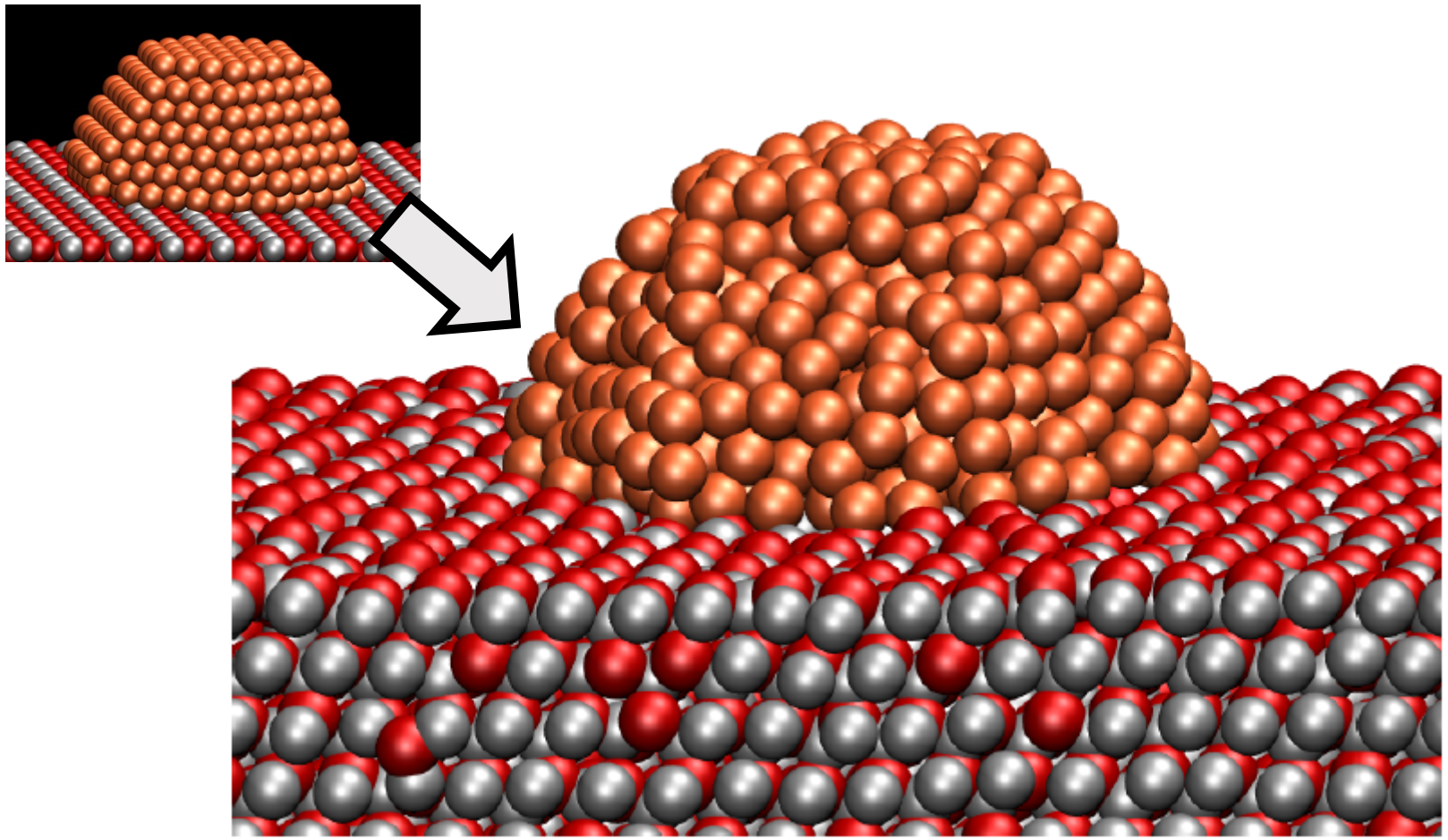
1. Independent test set not used for training

2. Systematic identification of missing parts of configuration space



N. Artrith and J. Behler, *Phys. Rev. B* 85 (2012) 045439.

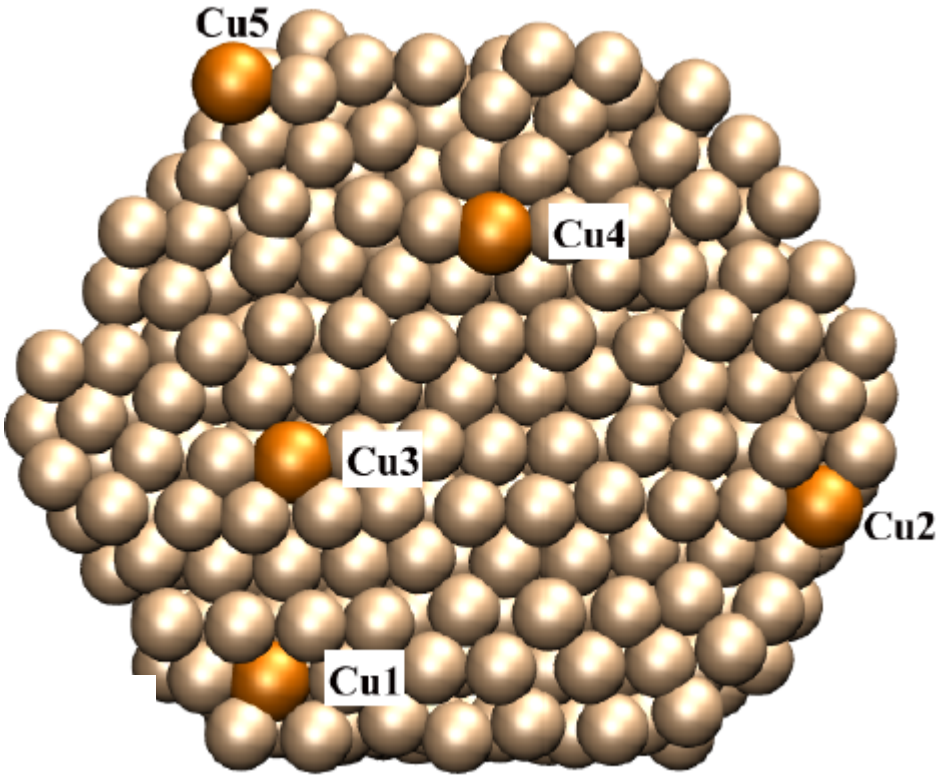
Random Snapshot from MD Simulation at 1000 K

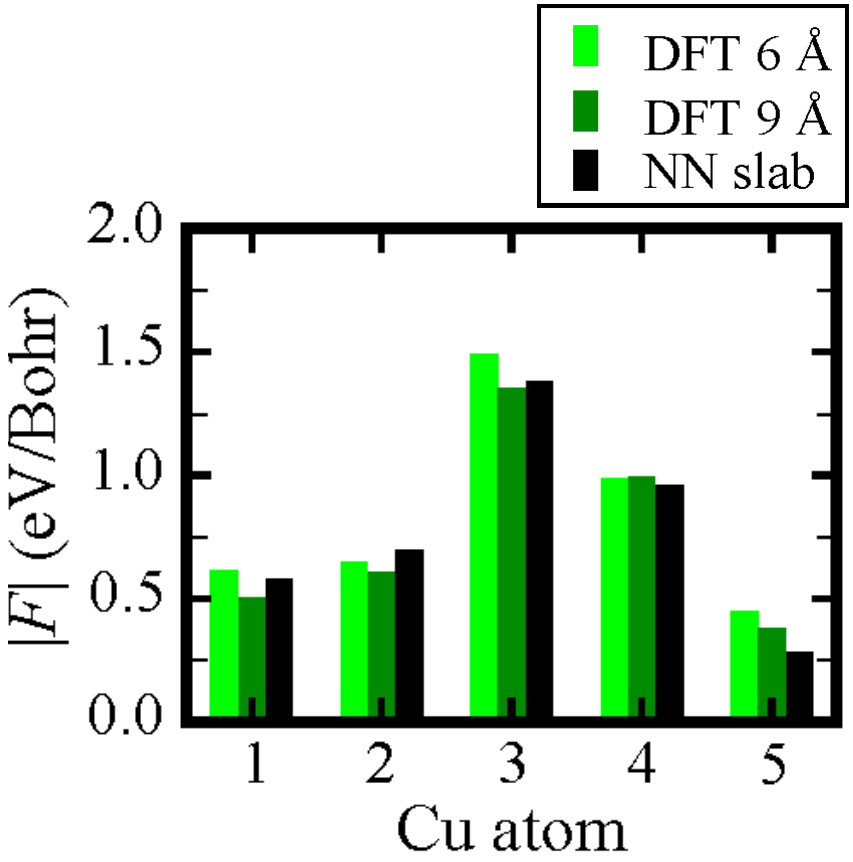


N. Artrith, B. Hiller, and J. Behler, *Phys. Status Solidi B,* 250, 1191 (2012).

Forces at Selected Copper Atoms of the Interface

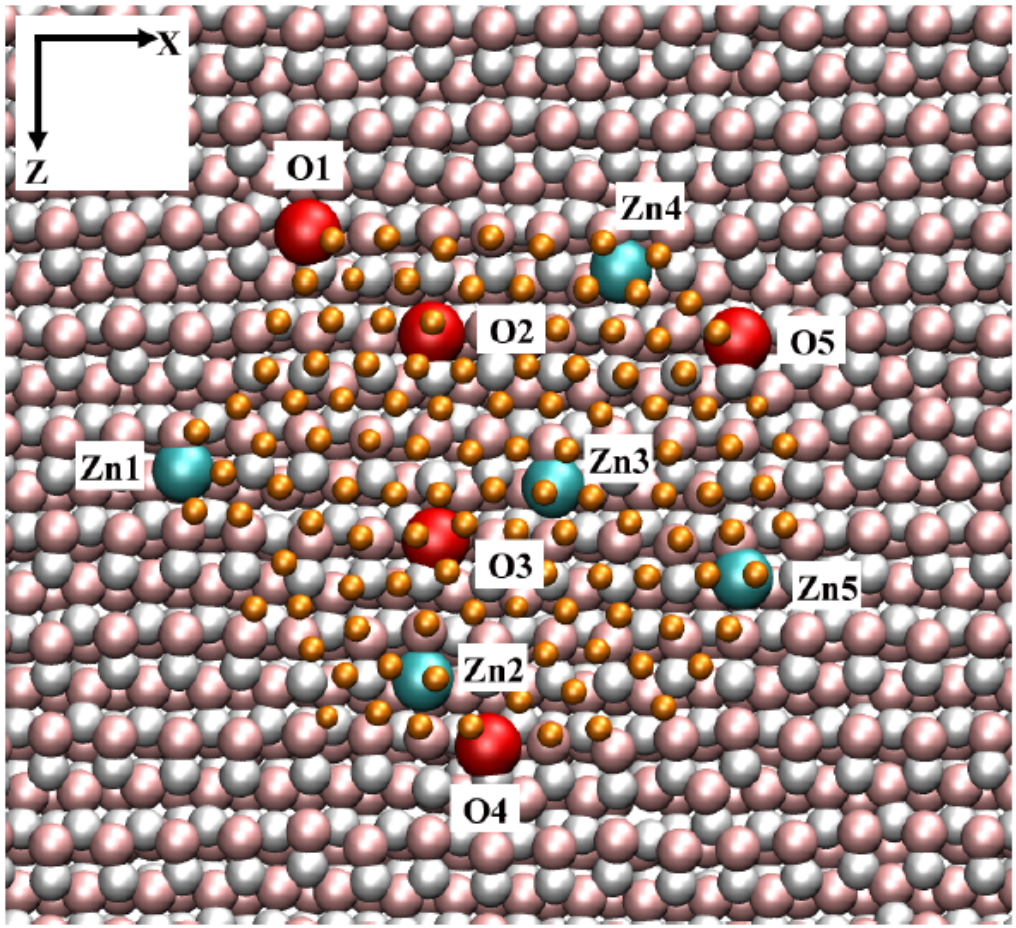
Bottom View:

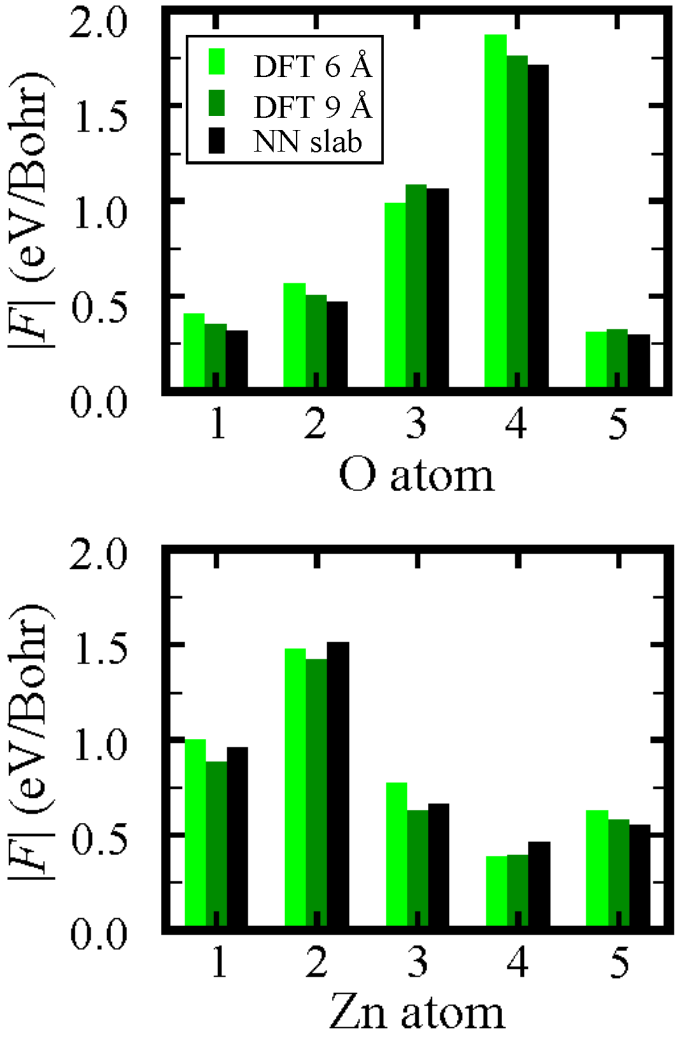




N. Artrith, B. Hiller, and J. Behler, *Phys. Status Solidi B,* 250, 1191 (2012).

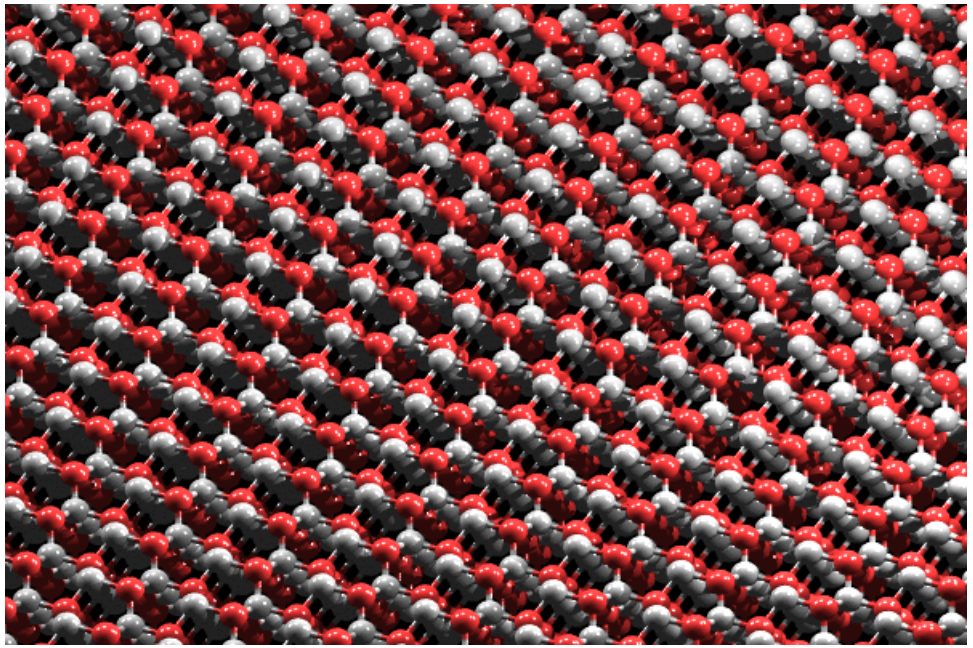
Forces at Selected Zinc anc Oxygen Atoms at the Interface





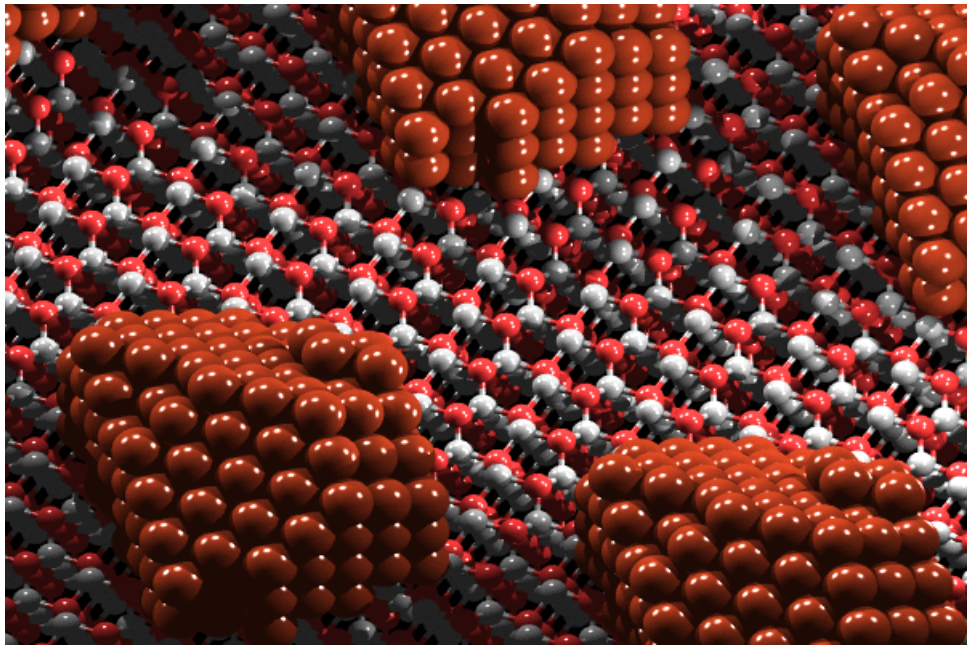
Copper Growth at ZnO(1 0 -1 0)

Preliminary Monte Carlo results:



- (5 x 10) ZnO(1 0 -1 0) supercell, 6 layers ⇒ 600 atoms
- growth of 200 Cu atoms
- 1000 Metropolis equilibration steps per added Cu atom

Preliminary Monte Carlo results:



⇒ cluster shape close to Wulff constructions

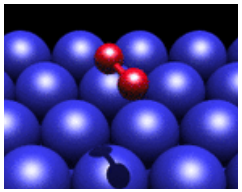
Next steps: lattice-free simulations

Summary
Neural Networks

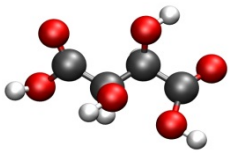
Further Applications of Neural Network Potentials

RUB

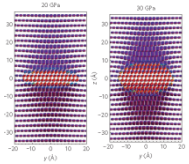
- **explanation of oxygen sticking at Al(111)**
  with K. Reuter, S. Lorenz, M. Scheffler, A. Groß, C. Carbogno
  PRL 94 (2005) 036104; *PRB* 75 (2007) 115409; *PRB* 77 (2008) 115421; *PRL* 101 (2008) 096104; *PRB* 81 (2010) 035410.



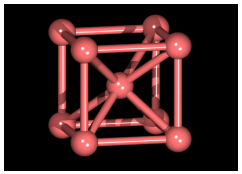
- **neural network potential for tartaric acid**
  S. Klees and J. Behler, *in preparation* (2013).



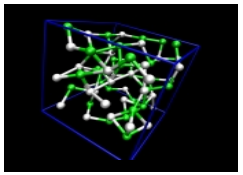
- **nucleation mechanism for the graphite-to-diamond phase transition**
  with R. Khaliullin, H. Eshet, T.D. Kühne and M. Parrinello
  PRB 81 (2010) 100103; *Nature Materials* 10 (2011) 693.



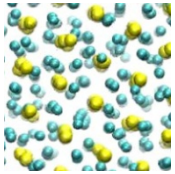
- **melting of sodium**
  with H. Eshet, R. Khaliullin, T.D. Kühne and M. Parrinello
  PRB 81 (2010) 184107; *PRL* 108 (2012) 115701.



- **working mechanism of phase change materials (GeTe)**
  with G. Sosso and M. Bernasconi
  PRB 85 (2012) 174103; *PRB* 86 (2012) 104301; *Phys. Stat. Sol. B* 249 (2012) 1880.



- **solid-liquid hybrid phase of high-chalcocite (Cu_2S)**
  with A. Singraber and C. Dellago
  In preparation (2013).



Summary: Take Home Messages

Advantages of Neural Network Potentials

- provide very accurate energies (error ≈ 5 meV/atom)
- only atomic positions needed ⇒ "reactive"
- fast (100 atoms per second per core)
- can be combined with any electronic structure method
- universal functional form, no system-specific terms
  ⇒ applicable to solids, molecules and surfaces
- systematic improvement possible

PCCP Vol. 13, Issue 40, 2011

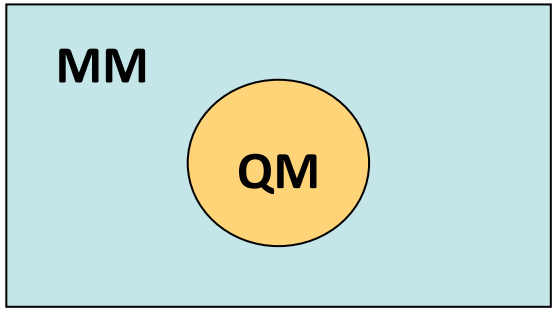
Disadvantages of Neural Network Potentials

- large training sets needed ⇒ costly
- currently restricted to 3-4 chemical elements
- "non-physical" (unbiased) functional form,
  ⇒ limited extrapolation capabilities
  ⇒ construction and application needs to be done with care

Important Applications of Neural Network Potentials

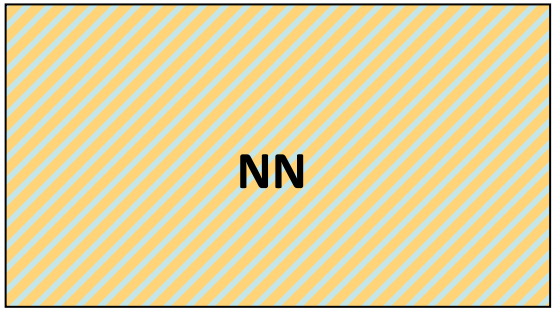
If reactive part of the system is not localized:

Localized reactions: QM/MM



Example:
- enzymes

Delocalized reactions: NN



Example:
- structural phase transitions
- chemistry at surfaces and interfaces

If very different types of bonding are present and system is large:

Example:
- self-assembled monolayers of organic molecules at metal surfaces

Other Machine Learning Techniques

Gaussian Approximation Potentials

A.P. Bartok, M.C. Payne, R. Kondor, G. Csanyi, *Phys. Rev. Lett.* 104 (2010) 136403.

	HDNNs (Behler and Parrinello 2007)	GAPs (Bartok et al. 2010)
Similarities:		
Accuracy:	very high	very high
Cutoff:	yes	yes
Fixed Form:	no	no
Reactions:	yes	yes
Symmetry:	both: rotational, translational and permutational invariance	
Speed:	the same (about 100 atoms per second and core in 2010)	
Differences:		
Structural Description:	symmetry functions	4D spherical harmonics
Energy Expression:	atomic NNs	Gaussians

Further Machine Learning Approaches

RUB

Some examples:

Polynomials

A. Brown, B.J. Braams, K. Christoffel, Z. Jin, J.M. Bowman, *J. Chem. Phys.* 119 (2003) 8790.

Systematic Neural Network method: truncated many-body expansion

S. Manzhos, T. Carrington, Jr., *J. Chem. Phys.* 125 (2006) 84109.

M. Malshe et al., *J. Chem. Phys.* 130 (2009) 184102.

Prediction of atomization energies for a wide range of organic molecules

M. Rupp, A. Tkatchenko, K.-R. Müller, and O.A. von Lilienfeld, *Phys. Rev. Lett.* 108 (2012) 058301.

and many others ...

Representing energies and potential-energy surfaces by machine learning techniques is a rapidly advancing field.

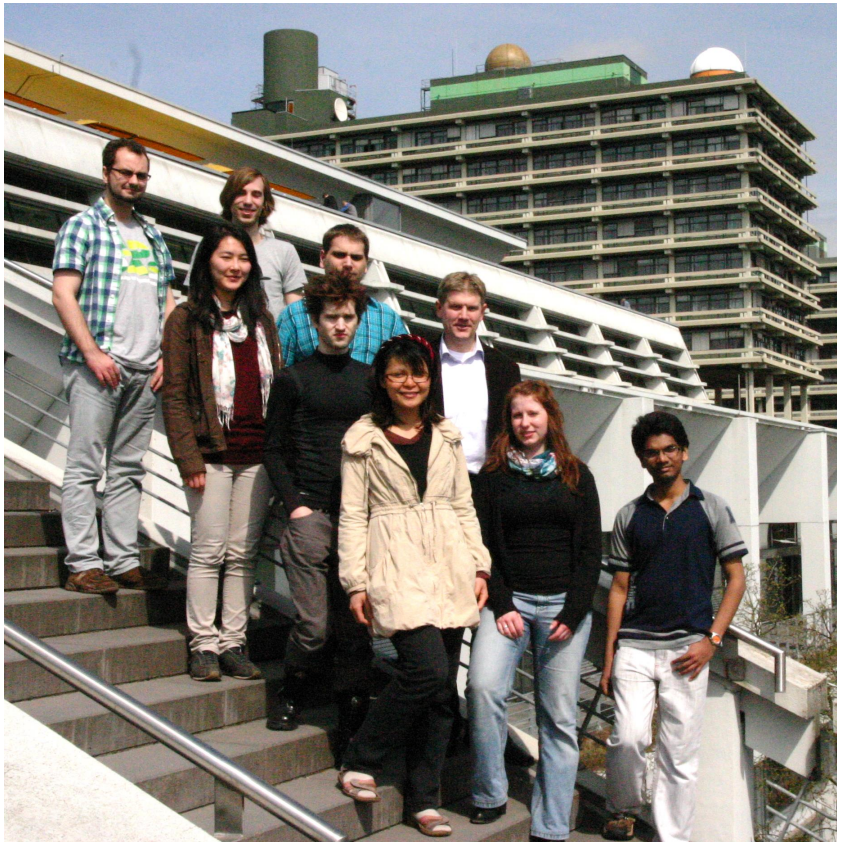
Acknowledgements

The Group@RUB

Christopher Handley
Björn Hiller
Felix Hoffmann
Eriko Hojo
Philipp Schienbein
Sinja Klees
Suresh Kondati Natarajan
Tobias Morawietz
Porntip Seema

Alumni:

Nongnuch Artrith (MIT)
Jovan Jose K.V. (Indiana University)
Marcus Maschke (RUB)
Vikas Sharma (University of Delhi)



Acknowledgements

Fritz-Haber Institut Berlin
Sönke Lorenz
Matthias Scheffler
Volker Blum + FHI-aims team
Karsten Reuter (now TU München)

University of Ulm
Christian Carbogno (now FHI)
Axel Groß

ETH Zürich
Michele Parrinello
Roman Martoňàk (now Bratislava)
Thomas Kühne (now University of Mainz)
Hagai Eshet (now Tel Aviv University)
Rustam Khaliullin (now University of Mainz)

Ruhr-Universität Bochum
Dominik Marx + Theoretische Chemie
Christof Wöll (now KIT), Martin Muhler + SFB 558
Martina Havenith + RESOLV

Universität Wien
Andreas Singraber
Christoph Dellago

Università di Milano Bicocca
Gabriele Sosso
Marco Bernasconi

Funding:

Emmy Noether-Programm
Deutsche Forschungsgemeinschaft DFG

Sonderforschungsbereich 558
Metall-Substrat-Wechselwirkungen in der heterogenen Katalyse
der Ruhr-Universität Bochum

FCI FONDS DER CHEMISCHEN INDUSTRIE

RESOLV
RUHR EXPLORES SOLVATION
CLUSTER OF EXCELLENCE - EXC 1069

Nordrhein-Westfälische Akademie der Wissenschaften

Stiftung Mercator
IDEEN BEFLÜGELN.

Studienstiftung des deutschen Volkes

RUB
RESEARCH DEPARTMENT
Interfacial Systems Chemistry

Acknowledgements