Neural Network Potentials for Atomistic Simulations

Jörg Behler

Lehrstuhl für Theoretische Chemie II
Ruhr-Universität Bochum, Germany
and
Research Center
Chemical Sciences and Sustainability
Research Alliance Ruhr

1

Overview

Topic of this talk

Machine Learning as a tool to extend the time and length scales of ab initio molecular dynamics

Goal 1: General method for all types of systems

Goal 2: Quality of reference method (here DFT)

Goal 3: Answer questions that cannot be answered with AIMD directly

2

Introduction

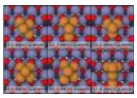
RUB

Goal:

Atomic-level understanding of complex systems in chemistry and materials science

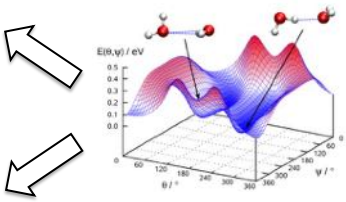
⇒ Predictive computer simulations with first-principles quality

Energy
global and local minima

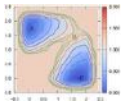


Forces
dynamics, free energies

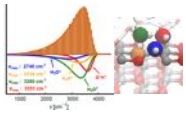
Central Role:
Potential Energy Surface



Reactions
barriers / transition states



Vibrations
properties, analysis



The accuracy of the obtained results depends on the quality of the atomic interactions

3

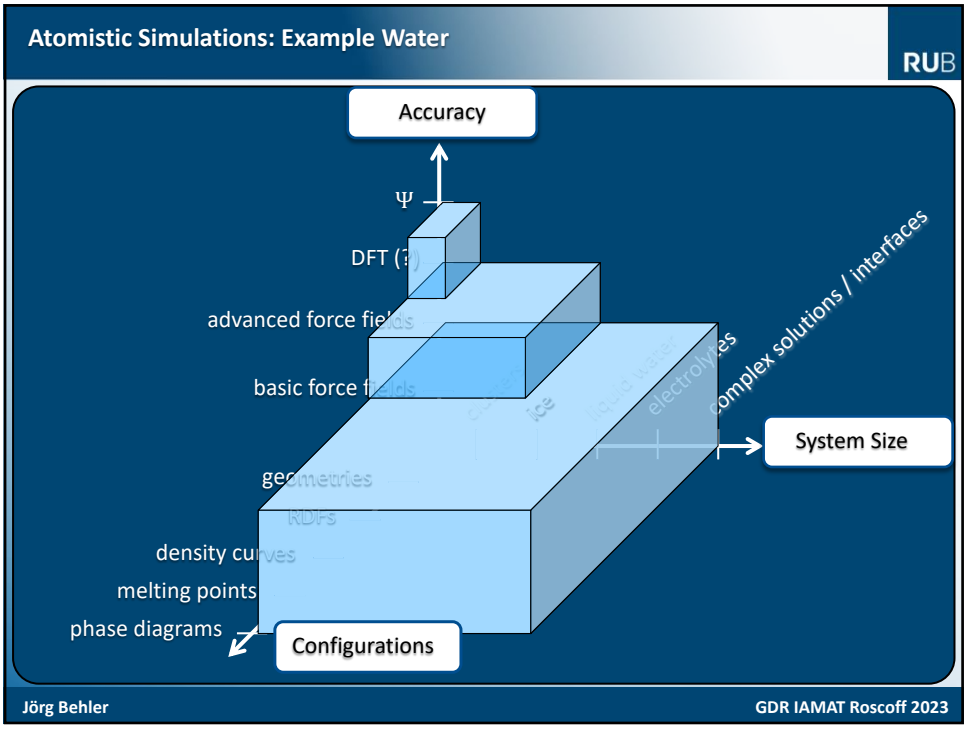
Potentials for Different Length and Time Scales

RUB

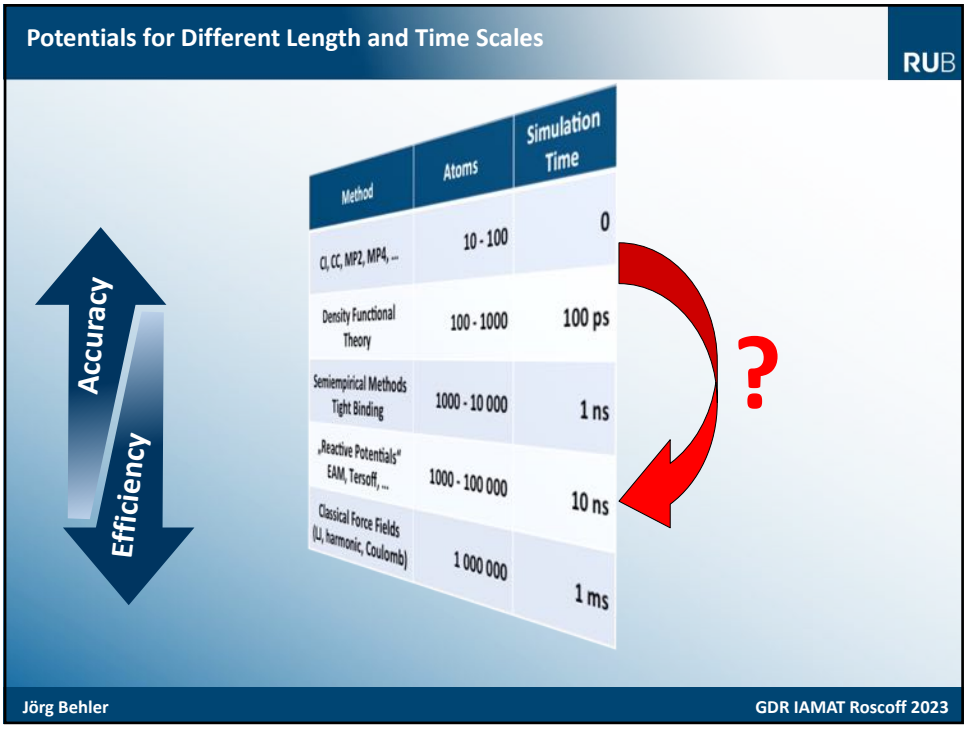
⇒ PES is the central quantity for atomistic simulations

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

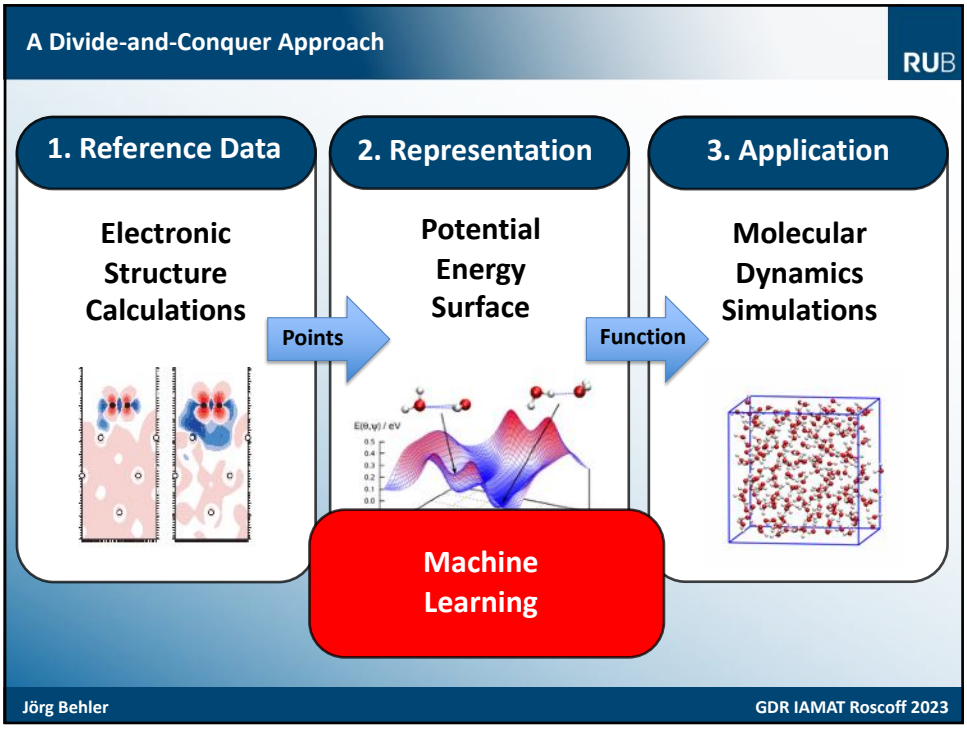
4


Atomistic Simulations: Example Water
RUB
Accuracy
Ψ
DFT (?)
advanced force fields
basic force fields
geometries
RDFs
density curves
melting points
phase diagrams
Configurations
complex solutions / interfaces
electrolytes
liquid water
ice
System Size
Jörg Behler
GDR IAMAT Roscoff 2023

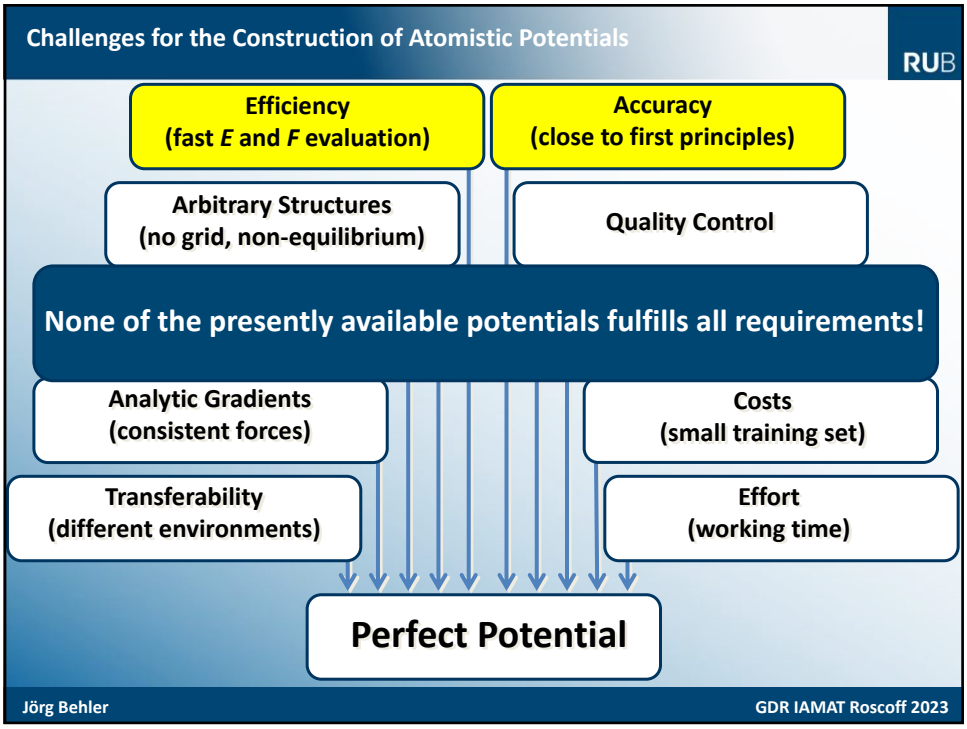
5


Potentials for Different Length and Time Scales
RUB
Accuracy
Efficiency
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
?
Jörg Behler
GDR IAMAT Roscoff 2023

6


A Divide-and-Conquer Approach
RUB
1. Reference Data
2. Representation
3. Application
Electronic Structure Calculations
Potential Energy Surface
Molecular Dynamics Simulations
Points
Function
Machine Learning
Jörg Behler
GDR IAMAT Roscoff 2023

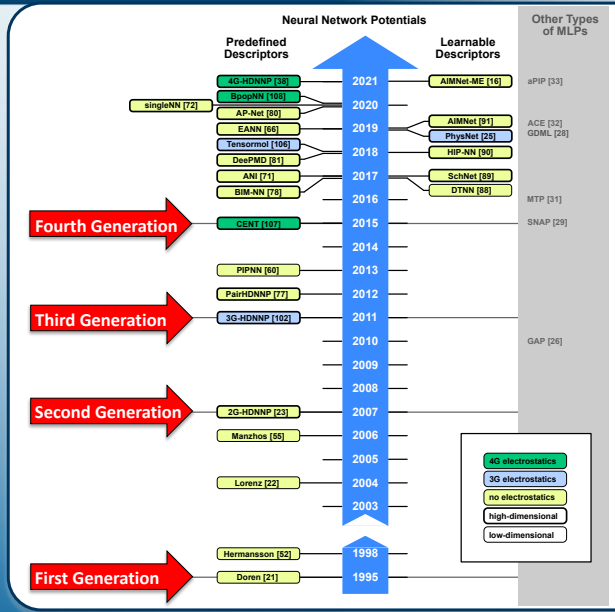
7


Challenges for the Construction of Atomistic Potentials
RUB
Efficiency (fast E and F evaluation)
Accuracy (close to first principles)
Arbitrary Structures (no grid, non-equilibrium)
Quality Control
None of the presently available potentials fulfills all requirements!
Analytic Gradients (consistent forces)
Costs (small training set)
Transferability (different environments)
Effort (working time)
Perfect Potential
Jörg Behler
GDR IAMAT Roscoff 2023

8

Neural Network Potentials

RUB



J. Behler, *Chem. Rev.* 121 (2021) 10037.

E. Kocer, T. W. Ko, J. Behler, *Ann. Rev. Phys. Chem.*, in press (2022).

9

First-Generation Neural Network Potentials 1995 – 2007

10

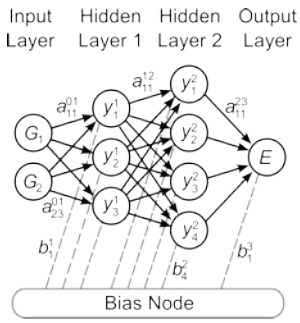
Machine Learning Potentials

RUB

1995 2000 2005 2010 2015 2020

First Machine Learning Potential = First Neural Network Potential

T.B. Blank, S.D. Brown, A.W. Calhoun, and D.J. Doren, *J. Chem. Phys.* **103** (1995) 4129.



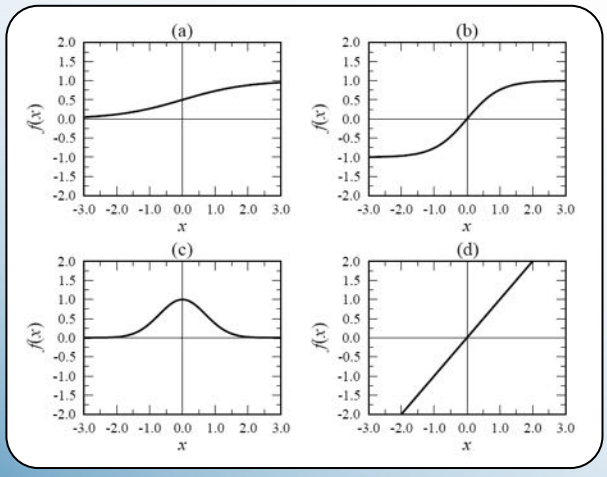
Feed-forward neural network

11

Activation Functions: Examples

RUB

Activation functions enable the fitting of general nonlinear functions.



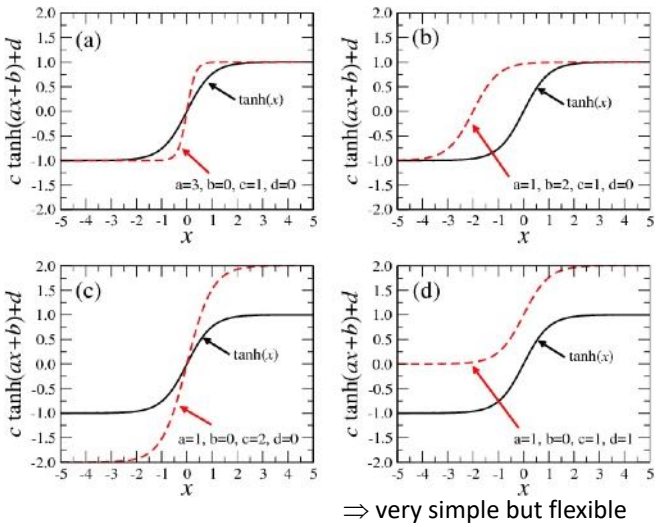
Activation functions
- converge for very small and very large arguments
- have a nonlinear shape for intermediate values

12

Neural Network Potentials: Activation Functions

RUB

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



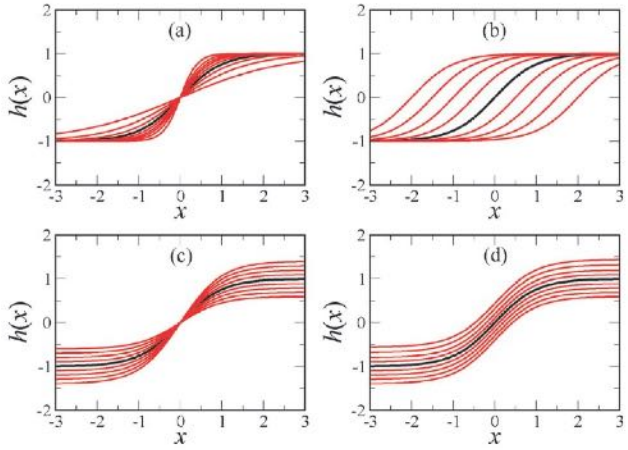
⇒ very simple but flexible

13

Neural Network Potentials: Activation Functions

RUB

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



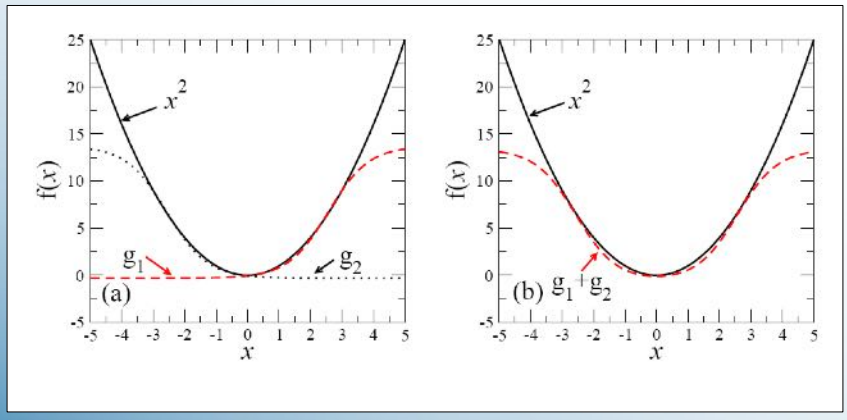
⇒ very simple but flexible

14

Model Potential: Harmonic Oscillator

RUB

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

15

Machine Learning Potentials

RUB

1995 2000 2005 2010 2015 2020

First-Generation MLPs

- about 30 papers (1995 – 2007) from about 10 groups
- all methods in the first decade are based on neural networks

Examples:

T.B. Blank, S.D. Brown, A.W. Calhoun, and D.J. Doren, *J. Chem. Phys.* **103** (1995) 4129.
H. Gassner, M. Probst, A. Lauenstein, K. Hermansson, J. Phys. Chem. A **102** (1998) 4596.
S. Lorenz, A. Groß, M. Scheffler, Chem. Phys. Lett. **395** (2004) 210.
S. Manzhos, T. Carrington, Jr., J. Chem. Phys. **125** (2006) 194105.
J. Behler, S. Lorenz, K. Reuter, J. Chem. Phys. **127** (2007) 014705.

Focus:

Training
Symmetry
Surfaces
Spectroscopy
Surface Symmetry

⇒ Basic ideas and key concepts are well established

16

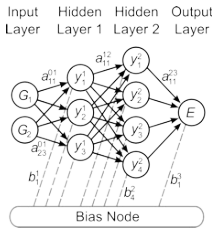
Machine Learning Potentials

RUB

1995 2000 2005 2010 2015 2020

First-Generation MLPs

Limitation: Applicable to low-dimensional systems only



Challenges:

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

⇒ No generally applicable solution for all systems

17

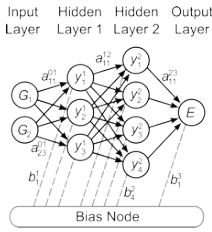
Machine Learning Potentials

RUB

1995 2000 2005 2010 2015 2020

First-Generation MLPs

Limitation: Applicable to low-dimensional systems only



Challenges:

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

⇒ Another approach is required for high-dimensional systems

Jörg Behler GDR IAMAT Roscoff 2023

18

Second-Generation Neural Network Potentials: Locality

19

High-Dimensional Neural Network Potentials

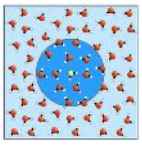
RUB

3 Steps:

1. Total energy is the sum of atomic energies

$$E = \sum_i E_i$$

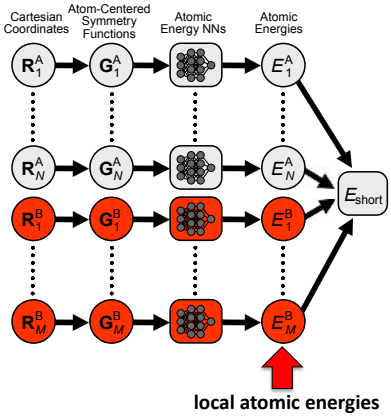
2. Atomic energies depend on local environments
   ⇒ cutoff radius



3. Description of local atomic environments by many-body atom-centered symmetry functions
   ⇒ structural fingerprints
   (invariances: rotation, translation, permutation)

J. Behler, M. Parrinello, *Phys. Rev. Lett.* 98 (2007) 146401.
J. Behler, *J. Chem. Phys.* 134 (2011) 074106.
J. Behler, *Angew. Chem. Int. Ed.* 56 (2017) 12828.

Structure for a binary system A_NB_M



⇒ **applicable to thousands of atoms**

20

How do the symmetry functions work?

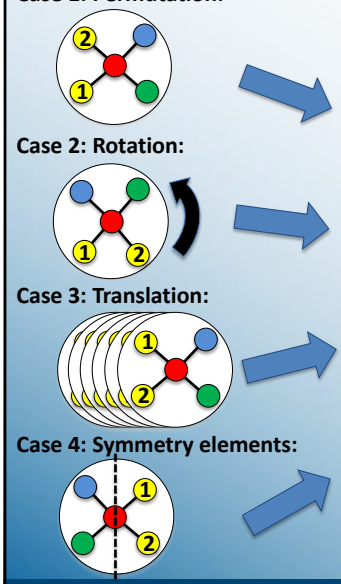
RUB

Case 1: Permutation:

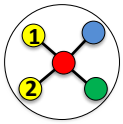
Case 2: Rotation:

Case 3: Translation:

Case 4: Symmetry elements:



Example environment:

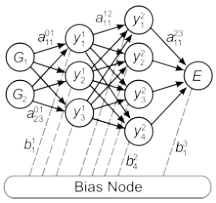


Symmetry function vector of central atom

$$\begin{pmatrix} 1.223 \\ 1.442 \\ \vdots \\ 2.332 \end{pmatrix}$$

Numerically identical

Atomic NN



⇒ Same property value (e.g. energy)

21

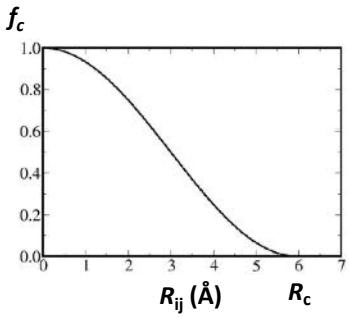
Symmetry Functions: Cutoff Function

RUB

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

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



Typical cutoff radius: R_c = 6 - 8 Å

22

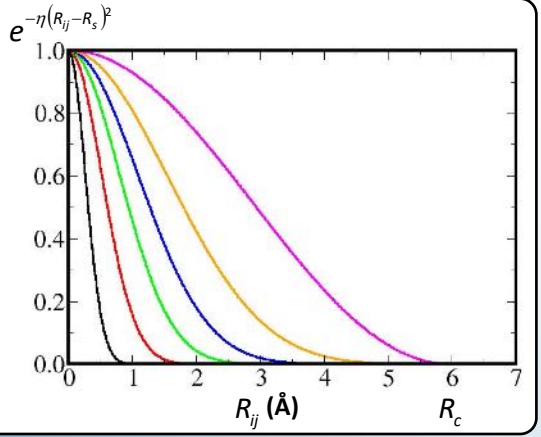
Symmetry Functions: Radial Functions

RUB

- 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.

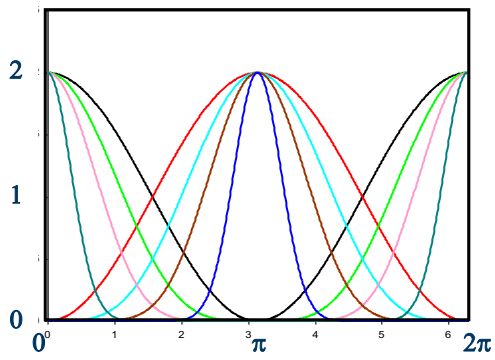
23

Symmetry Functions

RUB

Angular symmetry functions

$$G_i^{\mu} = 2^{1-\xi} \sum_{j,k\neq i} \left(1+\lambda\cos\theta_{ijk}\right)^{\xi} \cdot e^{-\eta\left(R_{ij}^2+R_{ik}^2+R_{jk}^2\right)} f_c\left(R_{ij}\right) f_c\left(R_{ik}\right) f_c\left(R_{jk}\right)$$



⇒ **Several angular functions are used**

- **all symmetry functions are many-body terms**
- **typical number: about 10 radial and 40 angular functions**

24

Symmetry Functions: Combinatorial Growth

RUB

For an *N* element system, there are *N* separate atomic NN types
⇒ no significant increase in complexity

For each radial function there are *N* functions for the possible neighboring elements
⇒ increase in complexity

For each angular function there are *N(N+1)/2* functions for the neighboring elements
⇒ strong increase in complexity

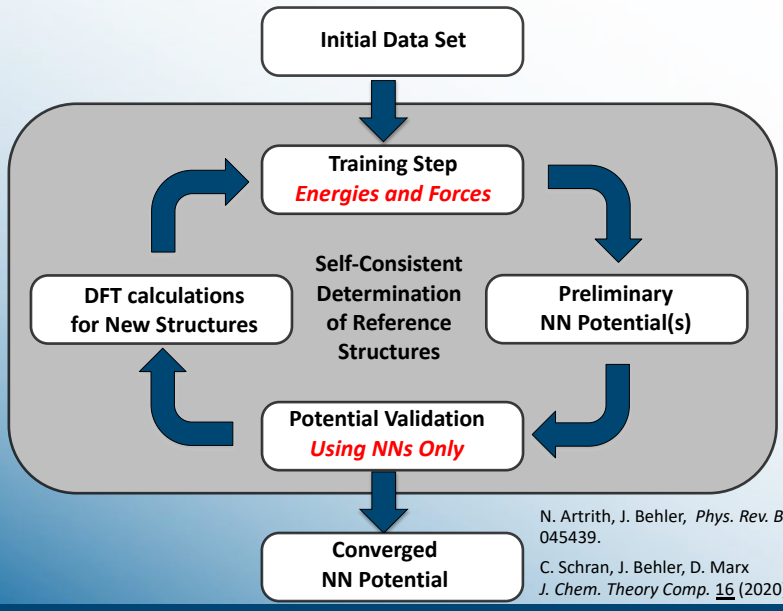
Elements	atomic NNs	multiplier radial/angular symmetry functions	typical total number of symmetry functions
1	1	1 / 1	6 + 25
2	2	2 / 3	12 + 75
3	3	3 / 6	18 + 150
4	4	4 / 10	24 + 250

Often, a reduction is possible, if only some compositions or structures are relevant

25

Construction of the Reference Set: Active Learning

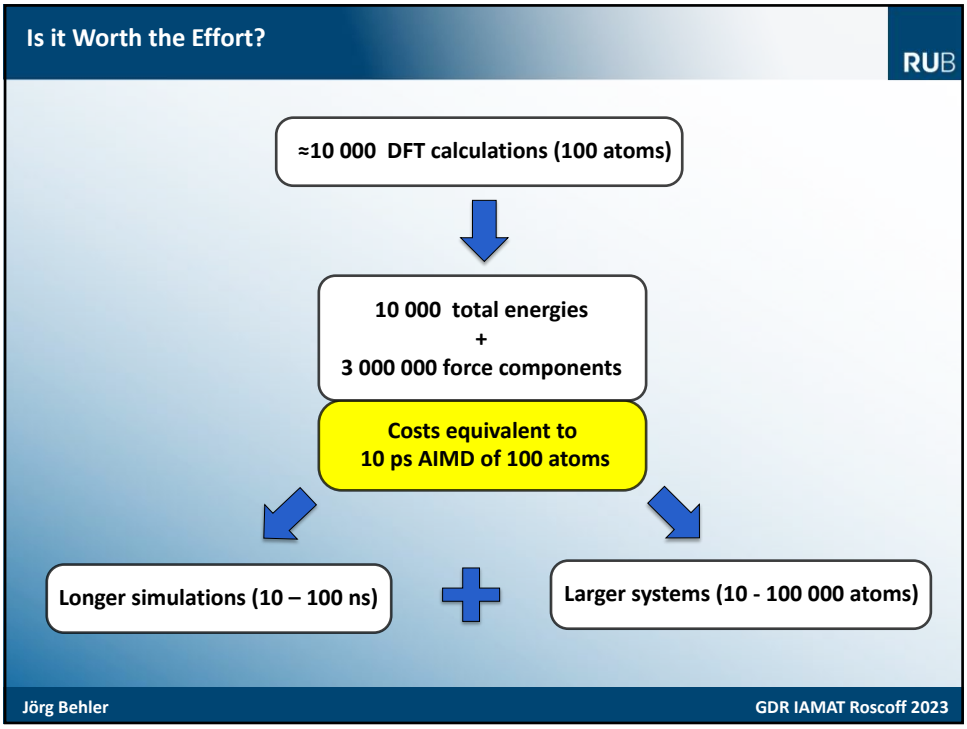
RUB



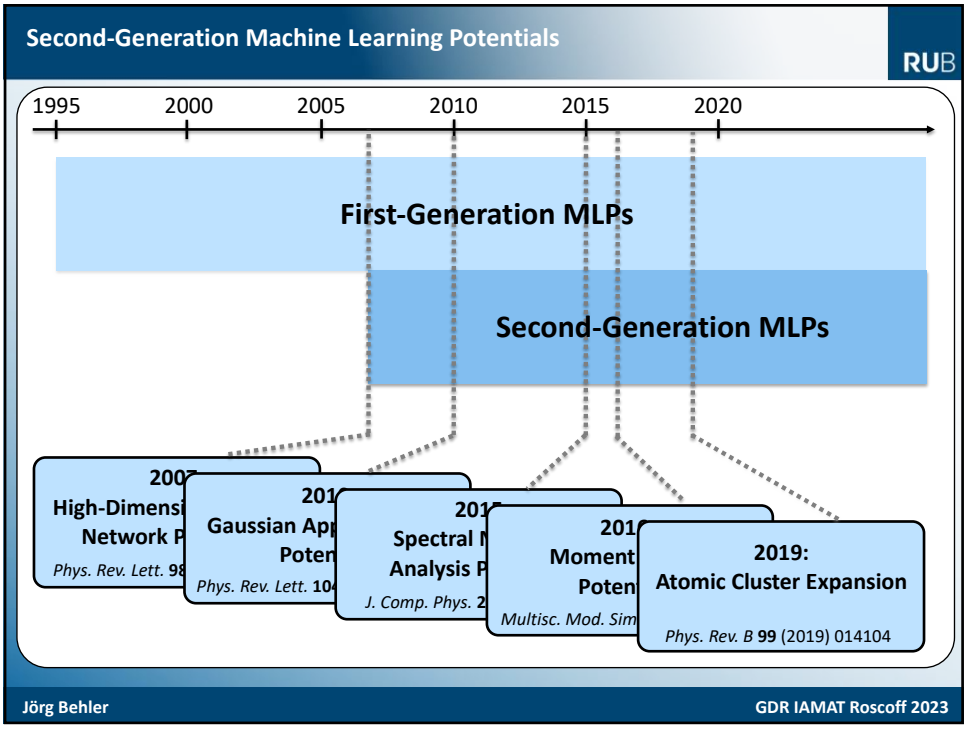
N. Artrith, J. Behler, *Phys. Rev. B.* 85 (2012) 045439.

C. Schran, J. Behler, D. Marx
J. Chem. Theory Comp. 16 (2020) 88.

26


Is it Worth the Effort?
RUB
≈10 000 DFT calculations (100 atoms)
10 000 total energies
+
3 000 000 force components
Costs equivalent to
10 ps AIMD of 100 atoms
Longer simulations (10 – 100 ns)
Larger systems (10 - 100 000 atoms)
Jörg Behler
GDR IAMAT Roscoff 2023

27

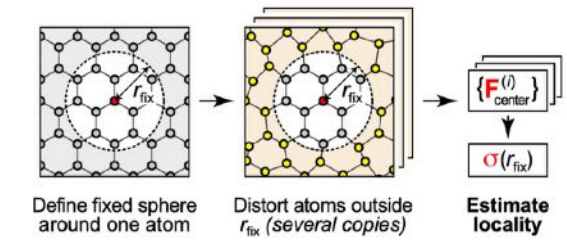

Second-Generation Machine Learning Potentials
RUB
1995
2000
2005
2010
2015
2020
First-Generation MLPs
Second-Generation MLPs
2019:
Atomic Cluster Expansion
Phys. Rev. B 99 (2019) 014104
Jörg Behler
GDR IAMAT Roscoff 2023

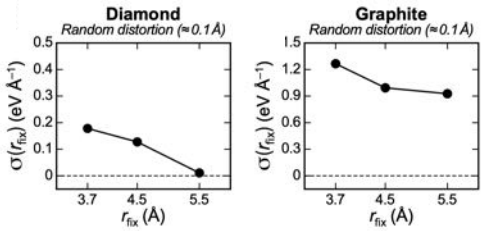
28

Interaction Range

RUB

Locality Test





V. L. Deringer and G. Csányi, Phys. Rev. B 95 (2017) 094203.

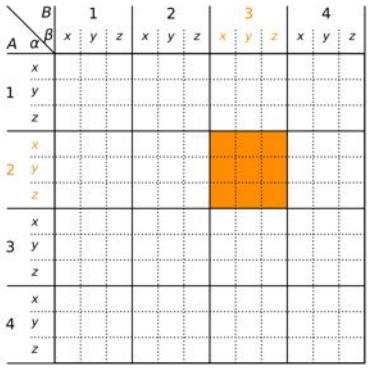
29

Interaction Range

RUB

A Hessian-Based Analytic Locality Test

Hessian $H_{A_\alpha B_\beta} = \frac{\partial^2 E}{\partial A_\alpha \partial B_\beta} = -\frac{\partial f_{B_\beta}}{\partial A_\alpha} = -\frac{\partial f_{A_\alpha}}{\partial B_\beta} \qquad \alpha, \beta = \{x, y, z\}$



Hessian submatrix norm

$$||\mathbf{h}_{AB}|| = \sqrt{\sum_{\alpha=x,y,z} \sum_{\beta=x,y,z} h^2_{A_\alpha B_\beta}}$$

Dependence of force on each individual atom in the system can be quantified

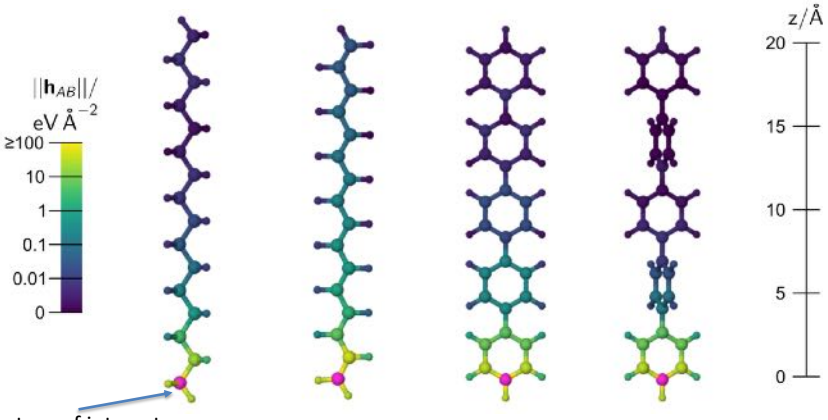
M. Herbold and J. Behler, J. Chem. Phys. **156** (2022) 114106.

30

Interaction Range

RUB

A Hessian-Based Analytic Locality Test



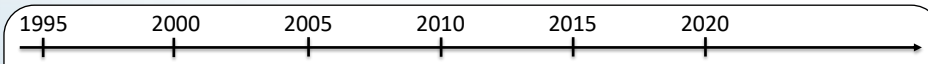
- contribution of each atom in the system can be quantified
- no cancellation of contributions

M. Herbold and J. Behler, J. Chem. Phys. **156** (2022) 114106.

31

Second-Generation Machine Learning Potentials

RUB



First-Generation MLPs

Second-Generation MLPs

Limitation: Truncation of the interactions at the cutoff radius

32

Third-Generation Neural Network Potentials: Long-Range Interactions

33

Third-Generation Neural Network Potentials: Method

RUB

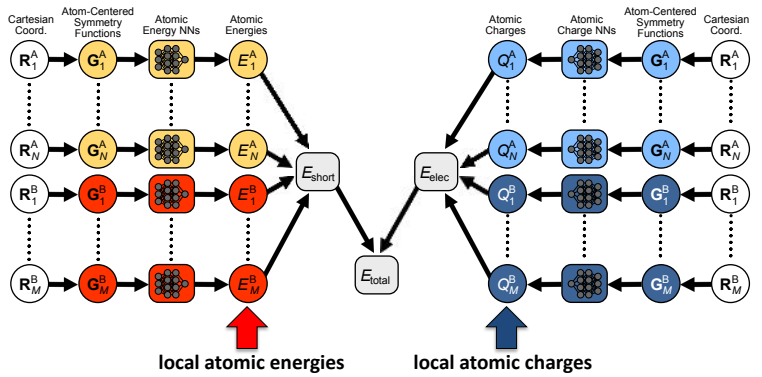
Short-Range Part:

local bonding

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

Long-Range Part:

electrostatics (and vdW)



N. Artrith, T. Morawietz and J. Behler, *Phys. Rev. B* 83 (2011) 153101.
T. Morawietz and J. Behler, *J. Chem. Phys.* 136 (2012) 064103.

34

Machine Learning Potentials

RUB

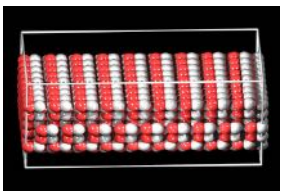
1995 2000 2005 2010 2015 2020

First-generation MLPs

Second-generation MLPs

Third-generation MLPs

Zinc Oxide Surface



Examples:

N. Artrith, T. Morawietz and J. Behler, *Phys. Rev. B* 83 (2011) 153101.
T. Morawietz and J. Behler, *J. Chem. Phys.* 136 (2012) 064103.
K Yao et al., *Chem. Sci.* 9 (2018) 2261.
O. T. Unke and M. Meuwly, *J. Chem. Theory Comput.* 15 (2019) 3678.

35

Forces

RUB

Short-range part:

Force with respect to some coordinate α:

$$F_{\alpha,s} = -\frac{\partial E_s}{\partial \alpha} = -\sum_{j=1}^{N_{atom}} \frac{\partial E_j}{\partial \alpha}$$

$$= \underset{\text{chain rule}}{-} \sum_{j=1}^{N_{atom}} \sum_{\mu=1}^{N_{sym,j}} \frac{\partial E_j}{\partial G_{j\mu}} \cdot \frac{\partial G_{j\mu}}{\partial \alpha}$$

from atomic NN architecture

from symmetry function definition

$\Rightarrow$ NNP provides exact analytic derivatives/forces
$\Rightarrow$ Important for energy conservation in MD

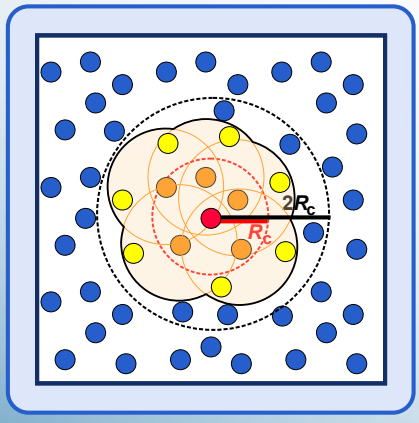
J. Behler, J. Chem. Phys. **134** (2011) 074106

36

Environment-Dependence of the Forces

RUB

An interesting consequence:
The environment-dependence of the forces is twice as large at the environment-dependence of the atomic energies!



J. Behler, Angew. Chem. Int. Ed. **56** (2017) 12828.

37

Forces including electrostatics

RUB

$$F_\alpha = F_{\alpha,s} + F_{\alpha,elec} = -\frac{\partial E_s}{\partial \alpha} - \frac{\partial E_{elec}}{\partial \alpha}$$

$$F_{\alpha,s} = -\frac{\partial E_s}{\partial \alpha} = -\sum_{j=1}^{N_{atom}} \frac{\partial E_j}{\partial \alpha} = -\sum_{j=1}^{N_{atom}} \sum_{\mu=1}^{N_{sym,j}} \frac{\partial E_j}{\partial G_{j\mu}} \cdot \frac{\partial G_{j\mu}}{\partial \alpha}$$

$$F_{\alpha,elec} = -\frac{\partial}{\partial \alpha} \frac{1}{2} \sum_{i=1}^{N_{atom}} \sum_{j=1, j\neq i}^{N_{atom}} \frac{Q_i Q_j}{R_{ij}}$$

$$= -\sum_{i=1}^{N_{atom}} \sum_{j=1, j\neq i}^{N_{atom}} \frac{1}{2R_{ij}^2} \left[\underbrace{\frac{\partial Q_i}{\partial \alpha} Q_j R_{ij} + Q_i \frac{\partial Q_j}{\partial \alpha} R_{ij}}_{\text{not included in standard force fields}} - Q_i Q_j \frac{\partial R_{ij}}{\partial \alpha} \right]$$

38

Fourth-Generation Neural Network Potentials: Global Electronic Structure

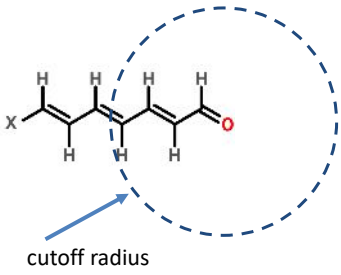
39

Fourth-Generation Neural Network Potentials

RUB

The Challenge:

Example molecule:

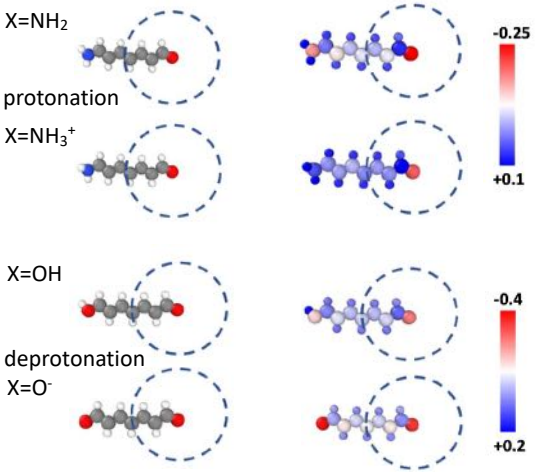


cutoff radius

X = functional group

⇒ oxygen does not see the functional group X

Structure **DFT charges**



T. W. Ko, J. A. Finkler, S. Goedecker, J. Behler, Nat. Commun. 12 (2021) 398.

40

Overview: Fourth-Generation Methods

RUB

2015: Charge Equilibration Neural Network Technique - CENT

Idea: Global charge equilibration + Charge-dependent total energy expression
Applicability: ionic inorganic systems

S. A. Ghasemi, A. Hofstetter, S. Saha, S. Goedecker, Phys. Rev. B 92 (2015) 045131.

2020: Becke Population Neural Networks - BpopNN

Idea: SCF charge distribution, populations as additional inputs in modified SOAP
Applicability: molecular systems

X. Xie, K. A. Persson, D. W. Small, J. Chem. Theory Comput. 16 (2020) 4256.

2021: Fourth-Generation High-Dimensional NNPs – 4G-HDNNP

Idea: Global charge equilibration + accurate short-range energies
Applicability: very general

T. W. Ko, J. A. Finkler, S. Goedecker, J. Behler, Nat. Commun. 12 (2021) 398.

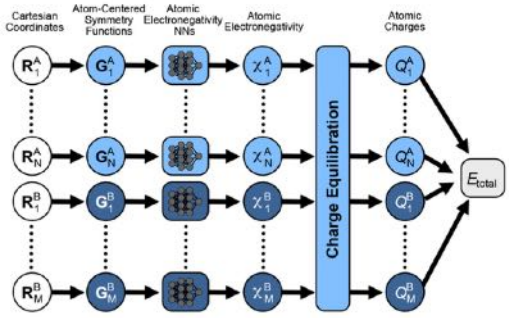
41

CENT – Charge Equilibration Neural Network Technique

RUB

$$E_{\mathrm{tot}}(\{Q_i\}) = \sum_{i=1}^{N_{\mathrm{atom}}} \left(E_i^0 + \chi_i Q_i + \frac{1}{2} J_{ii} Q_i^2 \right) + \frac{1}{2} \iint \frac{\rho(\mathbf{R})\,\rho(\mathbf{R}')}{|\mathbf{R} - \mathbf{R}'|} \mathrm{d}\mathbf{R}\,\mathrm{d}\mathbf{R}'$$

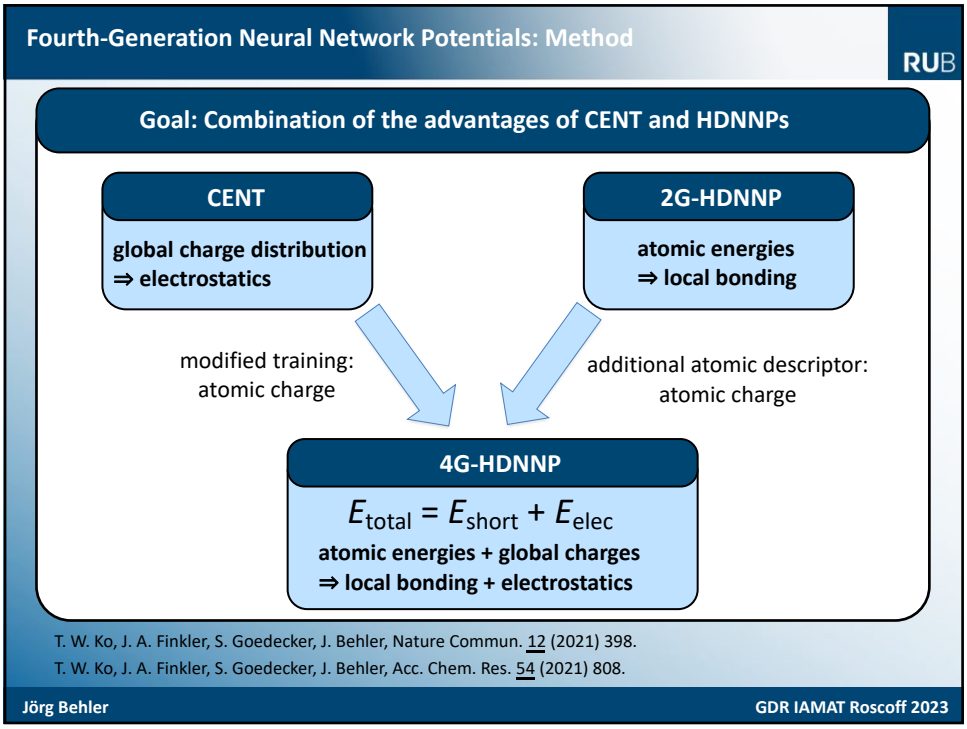
electronegativities — hardness — Coulomb energy



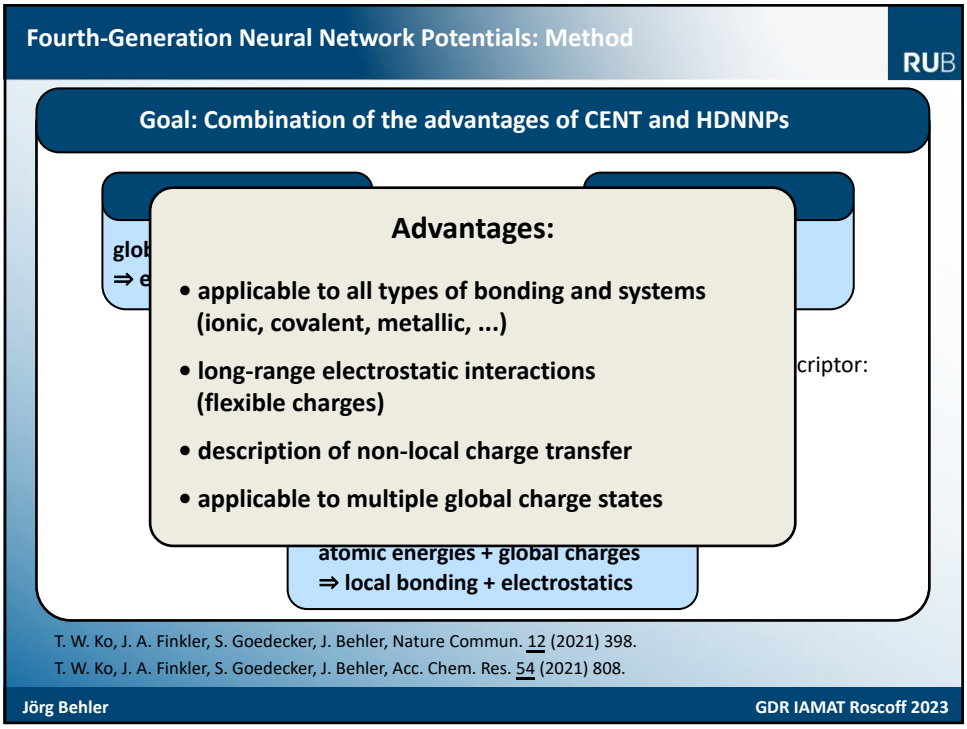
⇒ global electronic structure included (non-local charge transfer)
⇒ applications: systems with primarily ionic bonding

S. A. Ghasemi, A. Hofstetter, S. Saha, S. Goedecker, Phys. Rev. B **92** (2015) 045131.

42


Fourth-Generation Neural Network Potentials: Method
RUB
Goal: Combination of the advantages of CENT and HDNNPs
CENT
global charge distribution
⇒ electrostatics
2G-HDNNP
atomic energies
⇒ local bonding
modified training:
atomic charge
additional atomic descriptor:
atomic charge
4G-HDNNP
$E_{total} = E_{short} + E_{elec}$
atomic energies + global charges
⇒ local bonding + electrostatics
T. W. Ko, J. A. Finkler, S. Goedecker, J. Behler, Nature Commun. 12 (2021) 398.
T. W. Ko, J. A. Finkler, S. Goedecker, J. Behler, Acc. Chem. Res. 54 (2021) 808.
Jörg Behler
GDR IAMAT Roscoff 2023

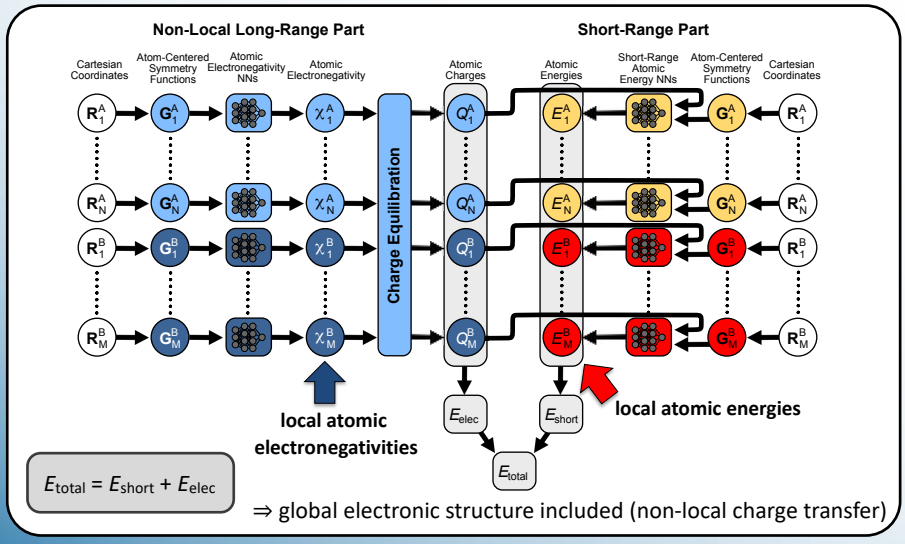
43


Fourth-Generation Neural Network Potentials: Method
RUB
Goal: Combination of the advantages of CENT and HDNNPs
Advantages:
• applicable to all types of bonding and systems (ionic, covalent, metallic, ...)
• long-range electrostatic interactions (flexible charges)
• description of non-local charge transfer
• applicable to multiple global charge states
criptor:
atomic energies + global charges
⇒ local bonding + electrostatics
T. W. Ko, J. A. Finkler, S. Goedecker, J. Behler, Nature Commun. 12 (2021) 398.
T. W. Ko, J. A. Finkler, S. Goedecker, J. Behler, Acc. Chem. Res. 54 (2021) 808.
Jörg Behler
GDR IAMAT Roscoff 2023

44

Fourth-Generation Neural Network Potentials: Method

RUB



⇒ global electronic structure included (non-local charge transfer)

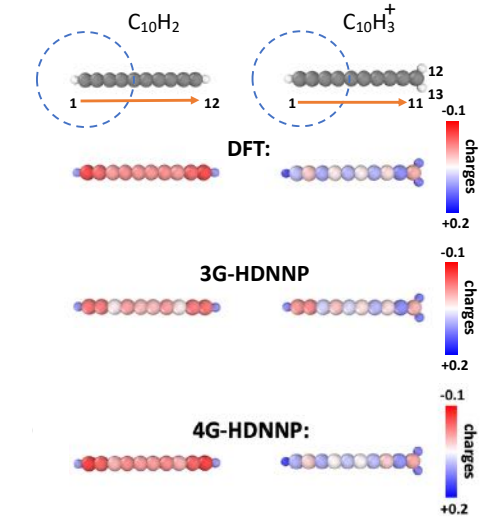
T. W. Ko, J. A. Finkler, S. Goedecker, J. Behler, Nature Commun. 12 (2021) 398.
T. W. Ko, J. A. Finkler, S. Goedecker, J. Behler, Acc. Chem. Res. 54 (2021) 808.

45

Fourth-Generation Neural Network Potentials: Examples

RUB

Test Case: Organic Molecules



Root Mean Squared Errors

Method		Q (e)	E (meV/atom)	F (eV/Å)
2G	train	---	1.6	0.131
	test	---	1.6	0.130
3G	train	0.027	3.2	0.653
	test	0.027	3.2	0.658
4G	train	0.006	1.1	0.078
	test	0.007	1.2	0.078

46

Fourth-Generation Neural Network Potentials: Examples

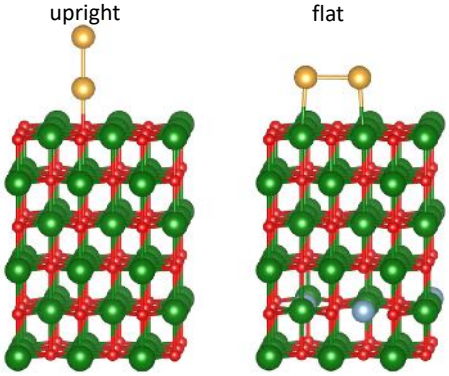
RUB

Au_2 cluster adsorbed on MgO(001)

DFT Optimization:

upright

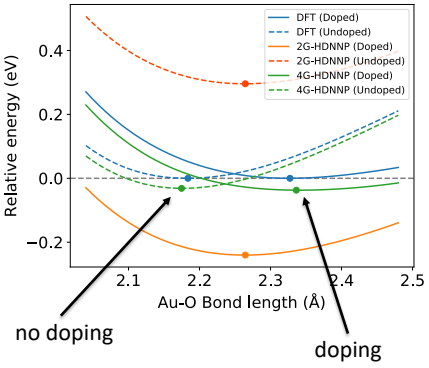
flat



no doping

Al doping

Energy for upright geometry



⇒ 4G-HDNNP can describe the effect of distant doping!

47

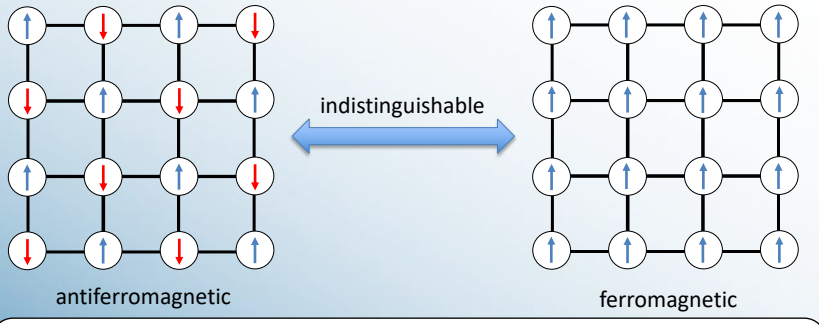
What About Magnetism?

48

Magnetic High-Dimensional Neural Network Potentials

RUB

Most descriptors do not depend on atomic spins



⇒ spin-related energy changes are treated as noise
⇒ unreliable energy surfaces
⇒ improved descriptors are needed

M. Eckhoff and J. Behler, *npj Comput. Mater.* **7** (2021) 170.

49

Magnetic High-Dimensional Neural Network Potentials

RUB

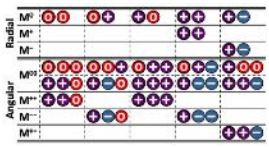
Spin-Dependent Atom-Centered Symmetry Functions (sACSF)

Step 1: Atomic spin coordinates (collinear spin)

$$s_i = \begin{cases} 0 & \text{for } |M_S| < M_S^{\text{thres}} \\ \text{sgn}(M_S) & \text{otherwise} \end{cases} \quad \text{with} \quad M_S = \tfrac{1}{2}(n_\uparrow - n_\downarrow)$$

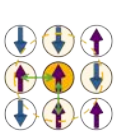
Step 2: Spin-augmentation functions (SAF), radial case

$$M^0(s_i, s_j) = 1\ ,$$
$$M^+(s_i, s_j) = \tfrac{1}{2}\,|s_i s_j| \cdot |s_i + s_j| \quad \Rightarrow \text{spin-sensitive filter}$$
$$M^-(s_i, s_j) = \tfrac{1}{2}\,|s_i s_j| \cdot |s_i - s_j|$$



Step 3: sACSF, radial case

$$G_i^{\text{rad}} = \sum_j M^x(s_i, s_j) \cdot \mathrm{e}^{-\eta R_{ij}^2} \cdot f_c(R_{ij})$$



⇒ Second-generation (environment-dependent) extension

⇒ Magnetic High-Dimensional Neural Network Potentials

M. Eckhoff and J. Behler, *npj Comput. Mater.* **7** (2021) 170.

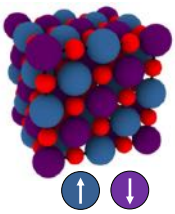
50

Magnetic High-Dimensional Neural Network Potentials

RUB

Application: Manganese Oxide (MnO)

- antiferromagnetic ground state (AFM-II) with rhombohedral distortion
- magnetic cell is 2x2x2 supercell of geometric unit cell



Training RMSE energy (3000 2x2x2 supercells, HSE06 functional):

HDNNP: 11 meV/atom (AFM-II = FM)

mHDNNP: 1 meV/atom (AFM II ≠ FM)

Results ΔE(AFM-II /FM) mHDNNP: 46.3 meV/atom (HSE06: 45.9 meV/atom)

Lattice parameters AFM-II
mHDNNP: a = 4.433 Å, α = 90.77°
HSE06: a = 4.434 Å, α = 90.89°
Exp.: a = 4.430 Å, α = 90.62°

M. Eckhoff and J. Behler, *npj Comput. Mater.* **7** (2021) 170.

51

How do we know if a Neural Network Potential is reliable?

What can we do with Neural Network Potentials?

52

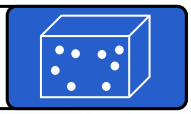
Quality Control

53

Multistep Quality Control

RUB

1. Generate reasonable initial training set
⇒ preliminary NN potential



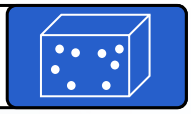
2. Check for extrapolation

54

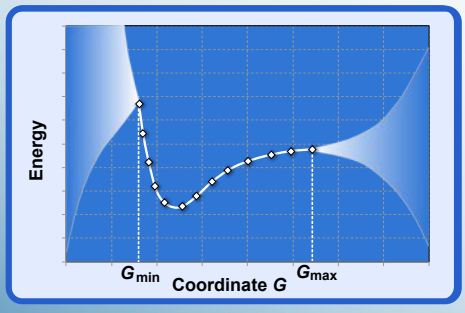
Multistep Quality Control

RUB

1. Generate reasonable initial training set
⇒ preliminary NN potential



2. Check for extrapolation



⇒ easy to detect

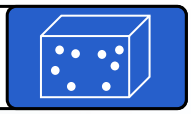
J. Behler, Angew. Chem. Int. Ed. 56 (2017) 12828.

55

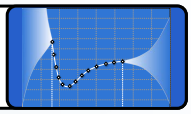
Multistep Quality Control

RUB

1. Generate reasonable initial training set
⇒ preliminary NN potential



2. Check for extrapolation
⇒ extend sampled "volume" in configuration space



3. Use a validation set ("Early Stopping Method")

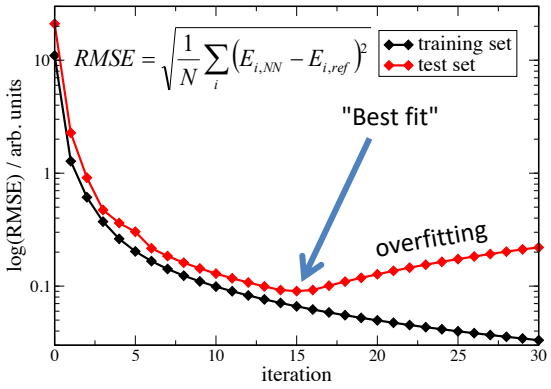
56

Neural Network Potentials: Early Stopping Method

RUB

Reference set → Training set: 90 %
Reference set → Test set: 10 %

Generalization → Cross Validation Methods



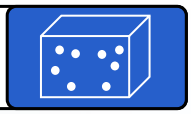
⇒ **overfitting cannot be detected by monitoring the error of the training set**

57

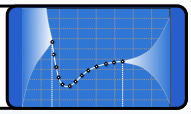
Multistep Quality Control

RUB

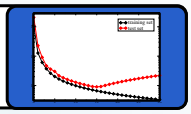
1. Generate reasonable initial training set
⇒ preliminary NN potential



2. Check for extrapolation
⇒ extend sampled "volume" in configuration space



3. Use a validation set ("Early Stopping Method")
⇒ estimate accuracy for <u>similar</u> structures not included in the training set

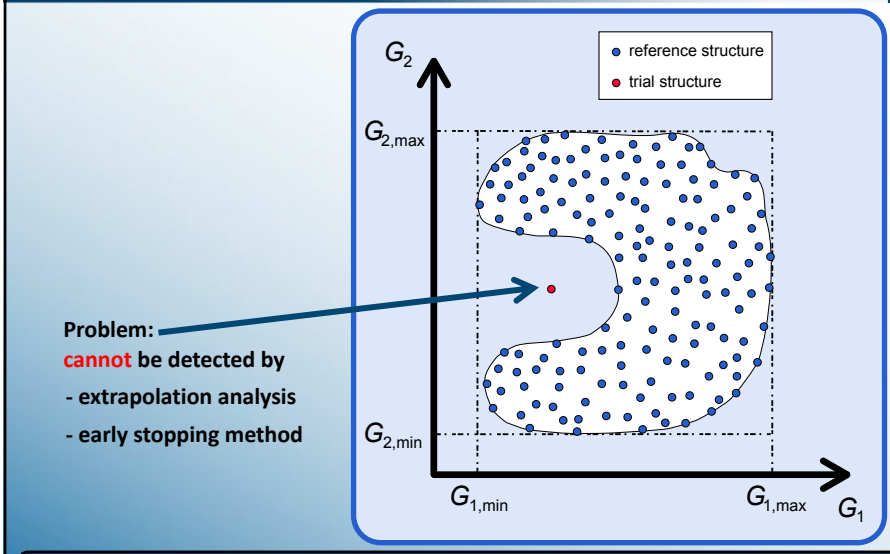


4. Check for "holes in the training set"

58

Detection of Problematic Configurations

RUB



Problem:
cannot be detected by
- extrapolation analysis
- early stopping method

⇒ **Early stopping method (training and test set) is not sufficient!**

59

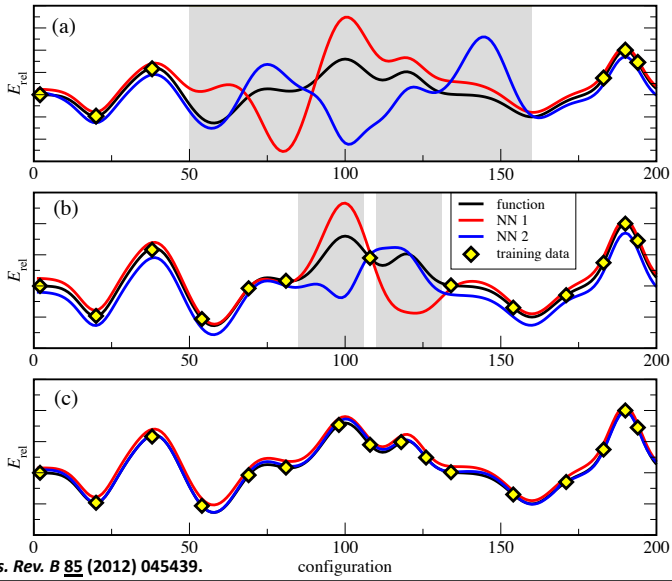
Iterative Improvement of the Reference Set

RUB

initial NN

improved NN

final NN



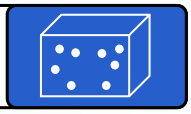
N. Artrith and J. Behler, *Phys. Rev. B* 85 (2012) 045439.

60

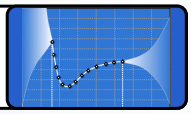
Multistep Quality Control

RUB

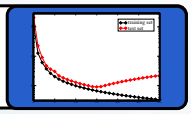
1. Generate reasonable initial training set
⇒ preliminary NN potential



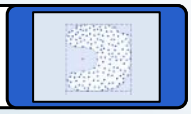
2. Check for extrapolation
⇒ extend sampled "volume" in configuration space



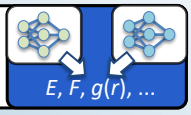
3. Use a validation set ("Early Stopping Method")
⇒ estimate accuracy for similar structures not included in the training set



4. Check for "holes in the training set"
⇒ add structures in poorly sampled regions (iteratively)



5. Check robustness of the results
⇒ repeat simulations with independent NN potentials



61

Summary: Four Generations of Neural Network Potentials

RUB

First-Generation Neural Network Potentials

Global description ⇒ low-dimensional systems

Second-Generation Neural Network Potentials

Short-range: yes Long-range: no Non-local: no Spin: yes

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

Third-Generation Neural Network Potentials

Short-range: yes Long-range: yes Non-local: no Spin: yes

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

Fourth-Generation Neural Network Potentials

Short-range: yes Long-range: yes Non-local: yes Spin: yes

T. W. Ko, J. A. Finkler, S. Goedecker, J. Behler, Nature Commun. 12 (2021) 398.

62

Summary: Four Generations of Neural Network Potentials

RUB

First-Generation Neural Network Potentials

Global description ⇒ low-dimensional systems

Second-Generation Neural Network Potentials

Short-range: yes Long-range: no

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

Neural Network Potentials fill a gap in the toolbox of chemistry, physics and materials science

Third-Generation Neural Network Potentials

Short-range: yes Long-range: yes Non-local: no Spin: yes

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

Fourth-Generation Neural Network Potentials

Short-range: yes Long-range: yes Non-local: yes Spin: yes

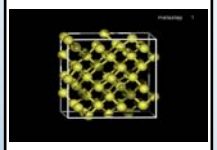
T. W. Ko, J. A. Finkler, S. Goedecker, J. Behler, Nature Commun. 12 (2021) 398.

63

Summary: Applications of Neural Network Potentials

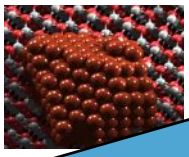
RUB

Bulk Materials



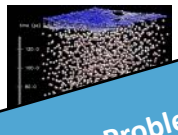
PRL 98 (2007) 146401.
PRL 100 (2008) 185501

Nanostructures

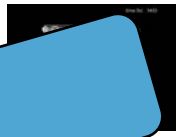


PRB 100103.
Nature Materials 10 (2011) 693.

Nucleation

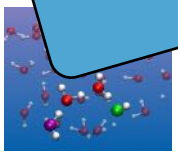


Phase Transitions



Neural Networks Help Solving Problems

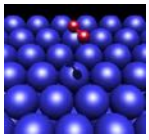
P



PNAS 113 (2016) 8368.

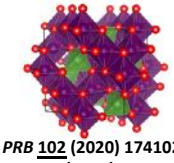
PCCP 18 (2016) 28704.
JPCC 121 (2017) 4368.

Gas-Surface Dynamics



JPCL 10 (2019) 2957.
JPCL 10 (2019) 1763.

Li-Intercalation Compounds



PRB 102 (2020) 174102.
JCP 153 (2020) 164107.

and many more …

64

Acknowledgements: Funding and Support

RUB