TECHNISCHE
UNIVERSITÄT
WIEN

165.144 Simulation of condensed matter

2025 Ass.Prof. Esther Heid esther.heid@tuwien.ac.at

Summary chapter 7

- Transferability of force fields is achieved through a careful selection of atom types balancing specificity and generality

Summary chapter 7

- Transferability of force fields is achieved through a careful selection of atom types balancing specificity and generality
- Periodic boundary conditions require a special treatment for non-bonded interactions:

Summary chapter 7

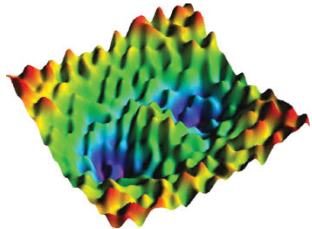
- Transferability of force fields is achieved through a careful selection of atom types balancing specificity and generality
- Periodic boundary conditions require a special treatment for non-bonded interactions:
  - **Lennard-Jones interactions** quickly decay, and can be cut-off after $1 - 3$ nm. A switching functions is used to smoothly transition from the full LJ-interaction to zero.

Summary chapter 7

- Transferability of force fields is achieved through a careful selection of atom types balancing specificity and generality
- Periodic boundary conditions require a special treatment for non-bonded interactions:
  - **Lennard-Jones interactions** quickly decay, and can be cut-off after $1 - 3$ nm. A switching functions is used to smoothly transition from the full LJ-interaction to zero.
  - **Coulomb interactions** do not decay quickly, and have to be taken into account using Ewald summation (or particle-mesh Ewald). The Ewald sum only converges if the net charge in the whole box is zero (!)

Summary chapter 7

- Transferability of force fields is achieved through a careful selection of atom types balancing specificity and generality
- Periodic boundary conditions require a special treatment for non-bonded interactions:
  - **Lennard-Jones interactions** quickly decay, and can be cut-off after $1 - 3$ nm. A switching functions is used to smoothly transition from the full LJ-interaction to zero.
  - **Coulomb interactions** do not decay quickly, and have to be taken into account using Ewald summation (or particle-mesh Ewald). The Ewald sum only converges if the net charge in the whole box is zero (!)
- Quantum mechanics provides an "ab-initio/first principles" (no prior knowledge or measurements required) means to study small systems
- Force fields can be parametrized using quantum mechanics

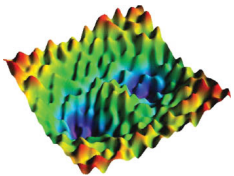


Chapter 8: Machine-learned force fields

Why settle on a functional form?

The potential energy $E_{\mathrm{pot}}(\mathbf{x})$ has many complex features:

- Number of minima exponential on N
- Multiple transition paths
- Features at different length and energy scales



https://doi.org/10.1515/pac-2014-0212

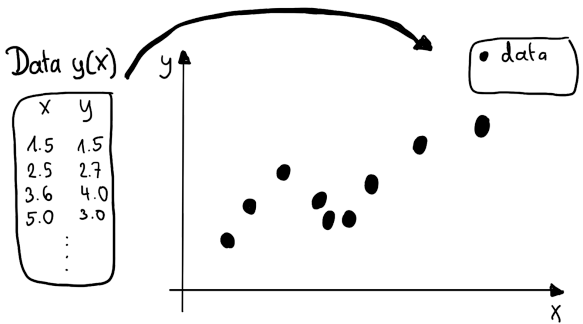
Approaches to approximate the true potential energy function:

- Quantum-mechanical calculations would be very accurate, but are way too expensive
- The classical forcefields from the previous lectures are cheap but only a crude approximation
- Machine learning can approximate a region of the function close to QM accuracy but with much less expensive function evaluations

Machine learning

Regression: Predict the continuous outcome/function value given some new unseen input

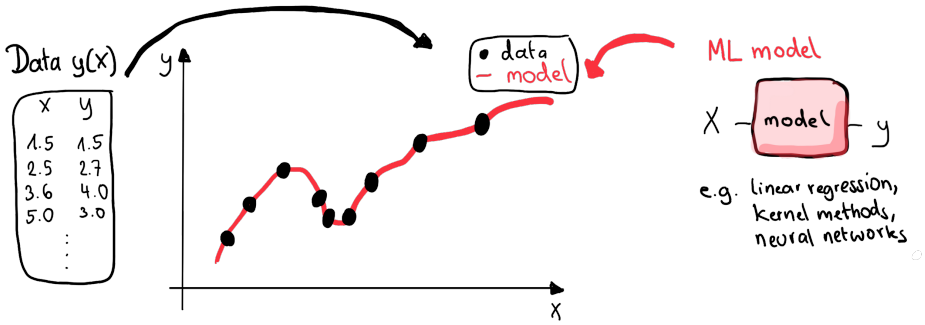
Step 1: Collect data (measurements, calculations)



Machine learning (regression task)

Regression: Predict the continuous outcome/function value given some new unseen input

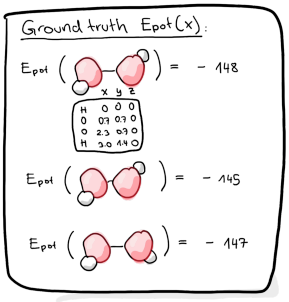
Step 2: Fit model (find best parameters of function)



Step 3: Apply model, e.g. y(2.0)=2.1

Machine-learning (ML) potentials

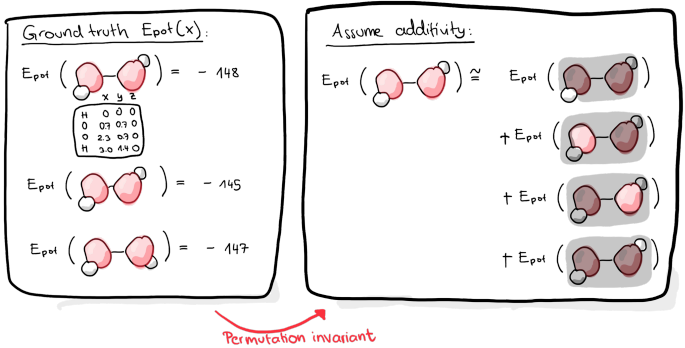
Data: $E_{\mathrm{Pot}}(\mathbf{x})$ ... potential energy as a function of the atomic coordinates from QM calculations for many relevant conformations (+ optionally atomic forces $\mathbf{f}$)



Problem: Learned potential energy would not be permuation, translation, or rotation invariant!

Machine-learning (ML) potentials

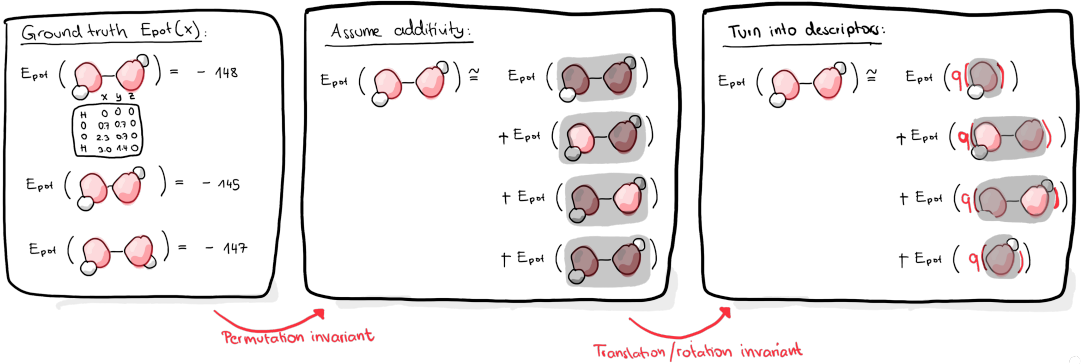
Assumption 1: $E_{\mathrm{Pot}}(\mathbf{x})$ is approximately additive with respect to the atomic contributions $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$



Problem: Learned potential energy would not be translation or rotation invariant!

Machine-learning (ML) potentials

Assumption 2: The atomic contribution $E_{\mathrm{Pot},N}(\mathbf{x})$ can also be learned from $E_{\mathrm{Pot},N}(\mathbf{q_N})$ with $\mathbf{q_N}$ being a descriptor representation of the local environment around atom N

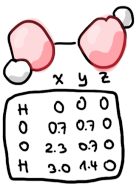


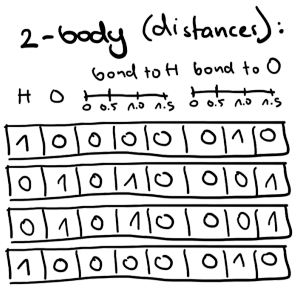
Problem: We must choose/design suitable descriptors

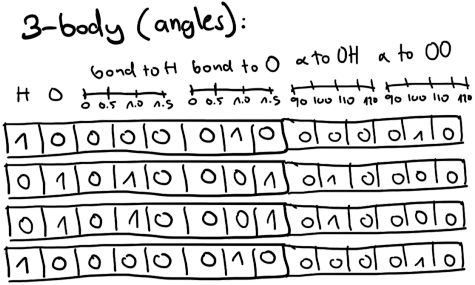
How to design a descriptor

Descriptors must be invariant with respect to translation, and invariant or equivariant with respect to rotation.

Simplest choice: Two-body (atom + distance information), or three-body (atom + distance + angle information)

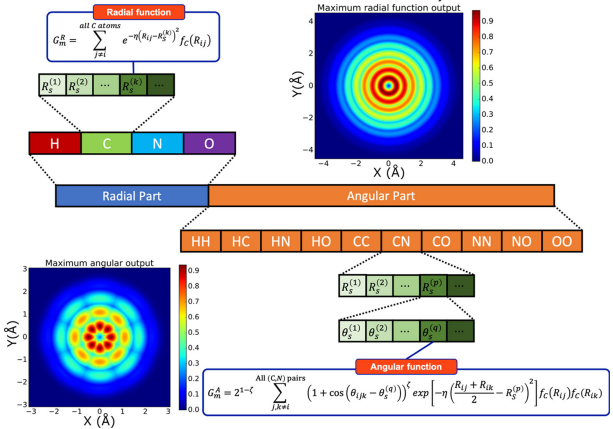






How to design a descriptor

Usually: Many-body representation (all neighboring atoms up to a cutoff) with a radial part and an angular part (usually based on spherical harmonics):



doi.org/10.1021/acs.accounts.0c00868

Descriptor-based architectures

Design of descriptors is still a work in progress, many variants published recently:

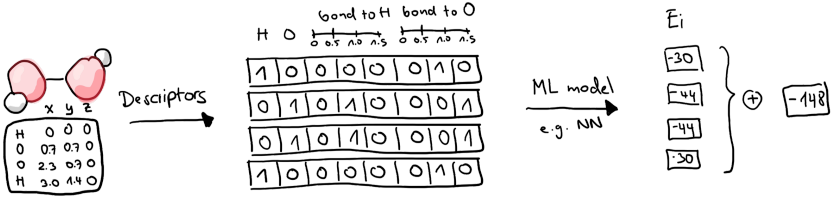
- Atom-centered symmetry functions (ACSFs)
- Smooth overlap of atomic positions (SOAP)
- Eigenvalues of the Coulomb matrix

Exemplar architectures:

- GAP: Gaussian process regression on SOAP descriptors
- NeuralIL: Residual neural network on Bessel descriptors
- ANI: Neural network on ACSFs

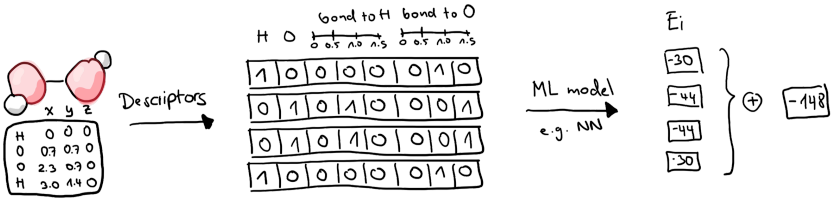
Learned descriptors

So far, we have custom-made the descriptors, relying on expert knowledge which information is important:

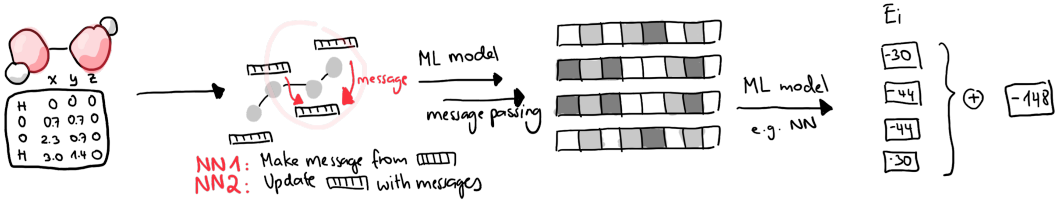


Learned descriptors

So far, we have custom-made the descriptors, relying on expert knowledge which information is important:



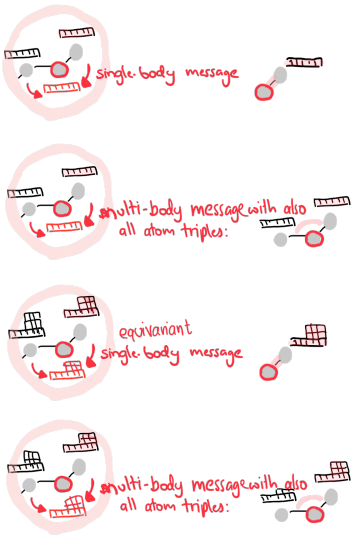
Instead, we can learn the descriptor from the data using message passing (= another model):



Learned descriptors

Evolution of learned descriptors:

1. 2-Body (atoms I, J) invariant message passing with scalar features: Very basic learned features
2. Multi-body messages including also triples (atoms I, J, K) or more: Can capture more information
3. 2-Body (atoms I, J) equivariant message passing with scalar and vectorial features: Can also capture directional information
4. Combination of equivariant message-passing with multi-body messages: Currently one of the best models

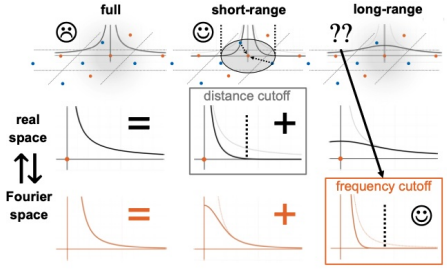


What about the long-range part?

Previous methods only use local descriptors. What about long-range forces though?

No commonly accepted approach yet. Some possibilities:

- Fixed atomic multipoles (charges, dipoles, etc)
- Machine-learning method for the charge density and the derived potential energy
- Project long-range part onto local descriptors
- Ewald-type ML



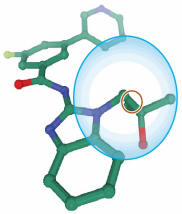
proceedings.mlr.press/v202/kosmala23a.html

Overview of machine-learned potentials

Neural network models

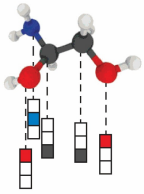
Predefined descriptor

ANI
DeepMD
EANN
HDNN
PIP-NN, FI-NN
TensorMol
wACSF, SingleNN, …



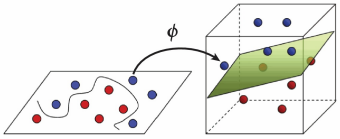
Learnable descriptor

AIMNet, AIMNet-NSE
DTNN, SchNet
HIP-NN
NequIP
PhysNet
PaiNN, SpookyNet…



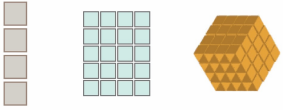
Kernel-based models

GPR/KRR
GAP
sGDML
SNAP
…



Others (linear and polynomial models)

ACE
ChIMEs
MTP
PIP
…



Yang Y, et al. 2024
Annu. Rev. Phys. Chem. 75:371–95

Useful software

- 

(scikit-learn.org): Comprehensive, well documented ML kit for Python, covering whole workflow.
- 

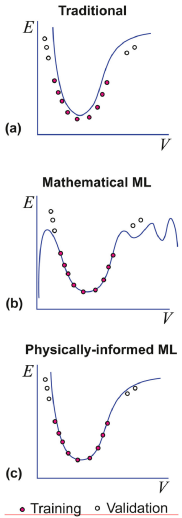
(tensorflow.org): Google's open soure software library for numerical computation using data flow graphs
- 

(pytorch.org): Meta's competitor to TensorFlow.
- 
(github.com/google/jax): Automatic differentiation and JIT compilation for Python and Numpy with the goal of high-performance machine-learning research.

Disadvantages/Challenges

The development of machine-learned force fields is far from finished, and many challenges remain. Maybe you want to contribute?

- No functional form means no safety net - predictions can get arbitrarily different from the ground truth. Predictions can thus even have a wrong sign, i.e. be attractive instead of repulsive
- Generalization ability often poor
- High dependence on training set
- Currently promising results after some finetuning, but not off-the-bench



doi.org/10.1016/j.actamat.2021.116980