INTERATOMIC POTENTIALS AND MACHINE LEARNING

Logan Ward
Asst. Computational Scientist
Argonne National Laboratory

3 February 2022

But first: what are interatomic potentials

We can't solve everything with quantum mechanics

Main Issue: Steep computational costs

Large scaling with electron count (N) ->
Only small systems are possible

Abbrev.	Runtime
CISDTQ	$O(N^{10})$
CC	$O(N^7)$
CISD	$O(N^6)$
MP2	$O(N^5)$
QMC	$O(N^3) - O(N^4)$
HF	$O(N^3) - O(N^4)$
DFT	$O(N^3)$
TB	$O(N^3)$
MM	$O(N^2)$

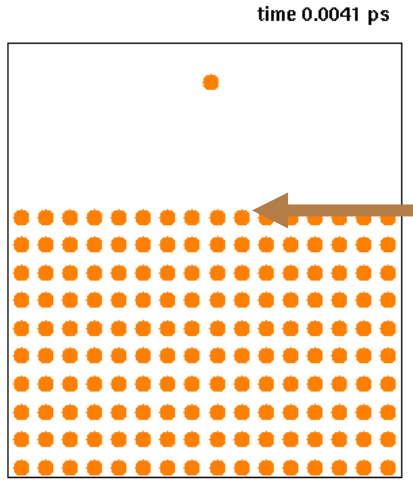
Table: Rupp. IJQC. (2015)

Observation: Do not *need* electrons for…

- Molecular dynamics
- Gibbs sampling
- …

Benefit: Cheap and linear scaling

Key need: Approximate models of forces



Only **energies** and **forces** needed for simulating molecular dynamics

Wide range of approaches

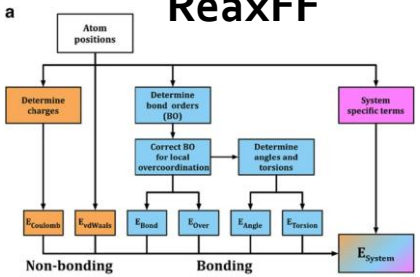
Physics-Motivated

Machine Learned

Lennard-Jones

$$E(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]$$

ReaxFF



Neural Network Potentials

Finnis-Sinclair

$$E(r) = -\sqrt{\rho} + V(r)$$

"Only include terms I understand"

"Good predictions imply good physics"

ReaxFF Figure: Senftle et al. npj Comp Mat (2016)

A grounding: Semi-empirical potentials

Original Formulation
Mostly physics

Lennard-Jones Potentials (1924)

Pauli repulsion: $\frac{1}{r^{12}}$

$$V(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]$$

Dispersive forces: $\frac{1}{r^6}$

Newer version: Wang et al. (2019)

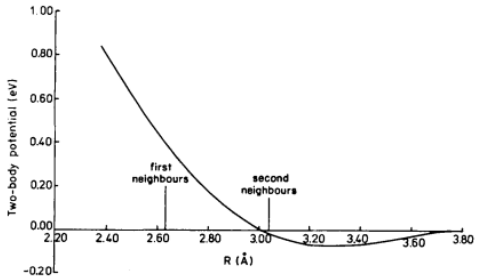
Newer Incarnations
Deviating slightly

$$V(r) = \epsilon\alpha\left(\left(\frac{\sigma}{r}\right)^{2\mu} - 1\right)\left(\left(\frac{r_c}{r}\right)^{2\mu} - 1\right)^{2v}$$

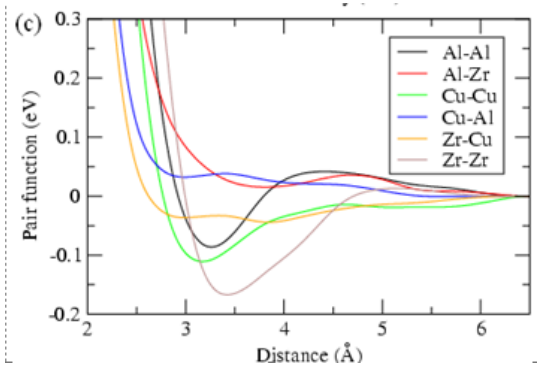
Finnis-Sinclair/EAM (1983)

$$E(r) = -\sqrt{\rho} + V(r)$$

"[pair potential] is purely repulsive"



Newer uses: Cheng et al. PRL (2009)



Semi-empirical potentials can be quite complex

MMFF94: Halgren 1994

The MMFF94 energy expression can be written as:

$$E_{MMFF} = \sum EB_{ij} + \sum EA_{ijk} + \sum EBA_{ijk} + \sum EOOP_{ijk;l} + \sum ET_{ijkl} + \sum EvdW_{ij} + \sum EQ_{ij} \quad (1)$$

$$E_{vdW_{ij}} = \varepsilon_{IJ}\left(\frac{1.07R^*_{IJ}}{R_{ij} + 0.07R^*_{IJ}}\right)^7\left(\frac{1.12R^{*7}_{IJ}}{R^7_{ij} + 0.12R^{*7}_{IJ}} - 2\right) \quad (8)$$

This form is used in conjunction with an expression that relates the minimum-energy separation R^*_{II} to the atomic polarizability α_I [eq. (9)], with specially formulated combination rules [eqs. (10) and (11)], and with a Slater–Kirkwood expression for the well depth ε_{IJ} [eq. (12)]:

$$R^*_{II} = A_I\alpha_I^{1/4} \quad (9)$$

$$R^*_{IJ} = 0.5(R^*_{II} + R^*_{JJ})\left(1 + 0.2\left(1 - \exp(-12\gamma^2_{IJ})\right)\right) \quad (10)$$

$$\gamma_{IJ} = (R^*_{II} - R^*_{JJ})/(R^*_{II} + R^*_{JJ}) \quad (11)$$

$$\varepsilon_{IJ} = \frac{181.16G_IG_J\alpha_I\alpha_J}{(\alpha_I/N_I)^{1/2} + (\alpha_J/N_J)^{1/2}}\frac{1}{R^{*6}_{IJ}} \quad (12)$$

ReaxFF: van Duin (2001)

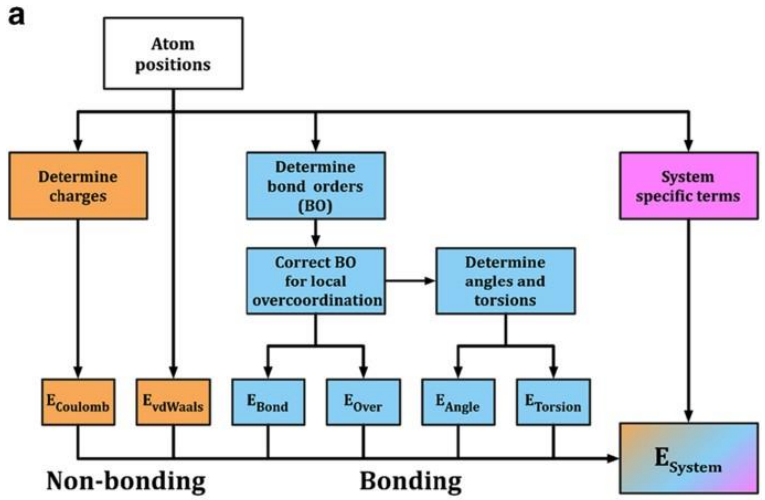


Figure: Senftle (2016)

MEAM: Baskes (~1987)

Based on same many-body equation as Finnis-Sinclair

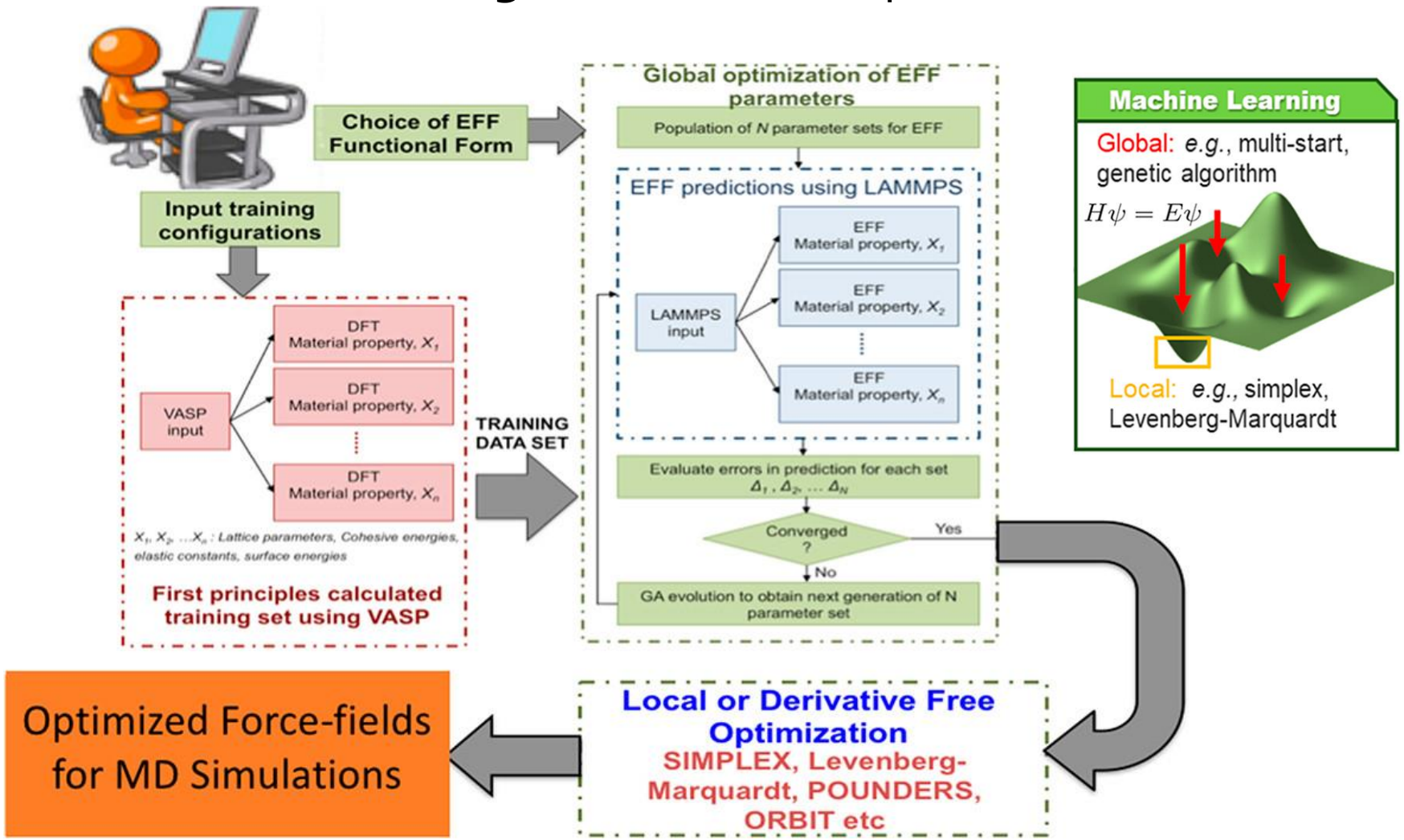
$$E_i = F(\rho_{ij}) + \sum_j \phi_{ij}$$

… but with angular terms to fit

$$(\rho^{(3)})^2 = \sum_{\alpha,\beta,\gamma}\left[\sum_i \rho^{a(3)}(r^i)\frac{r^i_\alpha r^i_\beta r^i_\gamma}{r^{i\,3}}\right]^2$$

Fitting semi-empirical potential

Secret Ingredient: Global optimization



Figures: Chan et al. JPCC. (2019)

Great resource: NIST Interatomic Potential Repository

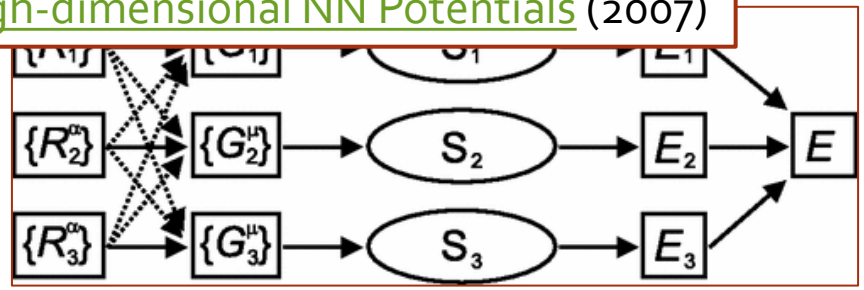
A reflection about empiricism from 1983

"Since the present approach is empirical [...]; the consequences of the form of the model which we now derive, and the fitting to experimental data which follow, do not depend on the physical interpretation"

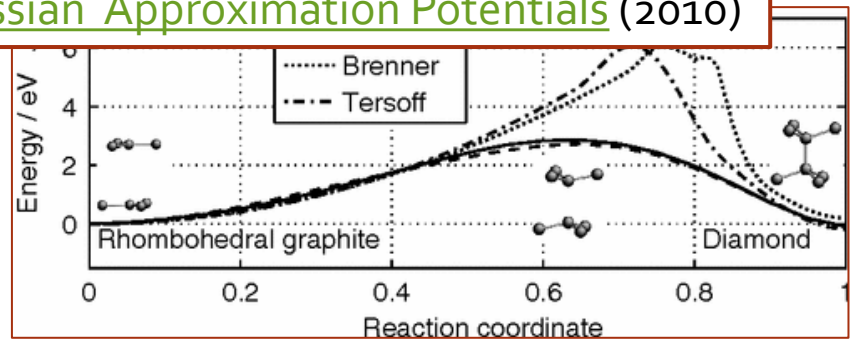
– M.W. Finnis and J.E. Sinclair (1983)

Abandoning model forms: Approaches for ML potentials

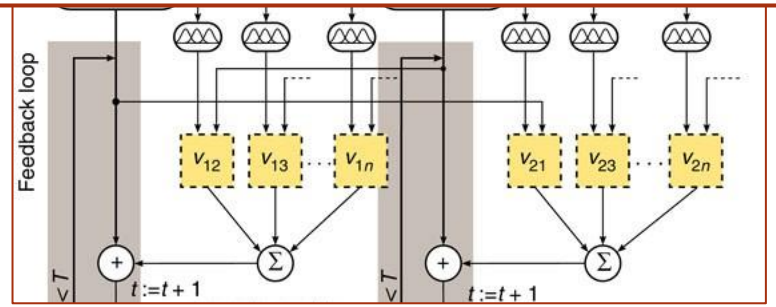
High-dimensional NN Potentials (2007)



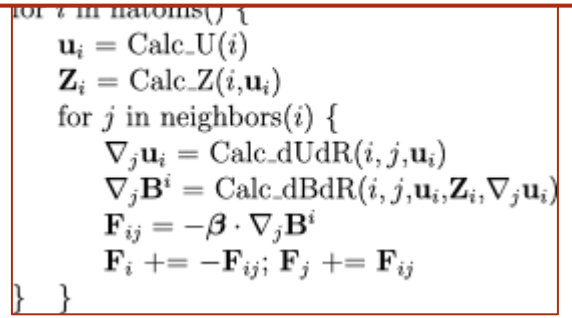
Gaussian Approximation Potentials (2010)



Deep Tensor Neural Networks / MPNNs (2016)

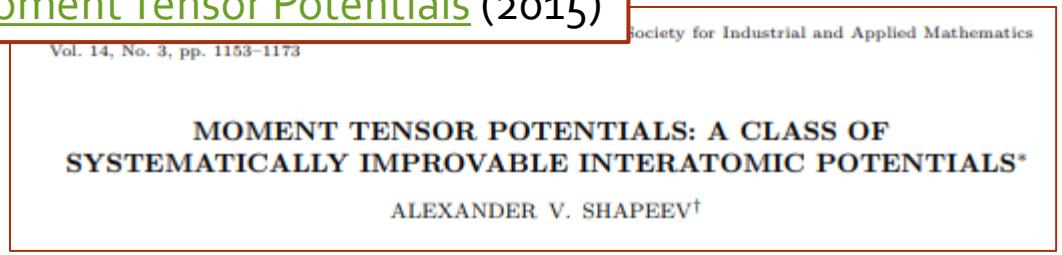


Spectral Neighbor Analysis Potentials (2014)



$\mathbf{u}_i = \text{Calc\_U}(i)$
$\mathbf{Z}_i = \text{Calc\_Z}(i, \mathbf{u}_i)$
for j in neighbors(i) {
$\nabla_j \mathbf{u}_i = \text{Calc\_dUdR}(i, j, \mathbf{u}_i)$
$\nabla_j \mathbf{B}^i = \text{Calc\_dBdR}(i, j, \mathbf{u}_i, \mathbf{Z}_i, \nabla_j \mathbf{u}_i)$
$\mathbf{F}_{ij} = -\boldsymbol{\beta} \cdot \nabla_j \mathbf{B}^i$
$\mathbf{F}_i \mathrel{+}= -\mathbf{F}_{ij}$; $\mathbf{F}_j \mathrel{+}= \mathbf{F}_{ij}$
} }

Moment Tensor Potentials (2015)



Society for Industrial and Applied Mathematics
Vol. 14, No. 3, pp. 1153–1173

MOMENT TENSOR POTENTIALS: A CLASS OF SYSTEMATICALLY IMPROVABLE INTERATOMIC POTENTIALS*

ALEXANDER V. SHAPEEV†

And many other works:
ANI-1 (2017), AGNI (2015), MAISE (2009), Sparse Interatomic Potentials (2014), MBTR (2017), Genetic Programming (2020),
and many I am likely missing…

Watch for how many body effects are captured

Two-body interactions (r_{ij}) are easy to compute, but pair-potentials are inaccurate

Importance of bond angles

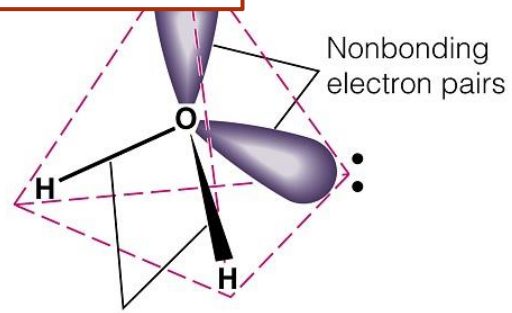


Figure: Stack Exchange

Non-linearity of energy vs coordination

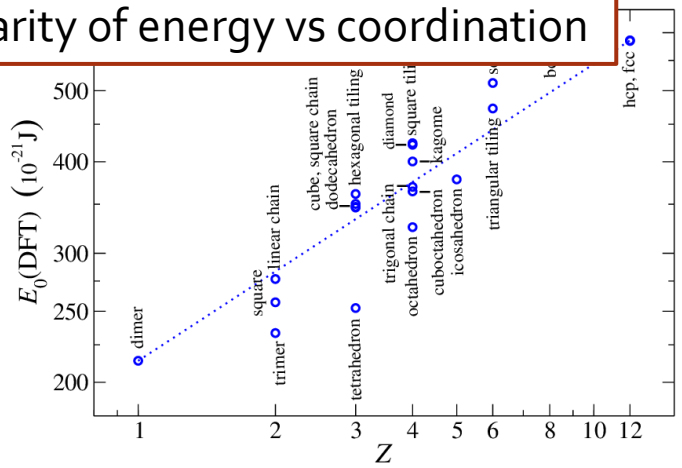


Figure: Jalkanen and Müser. (2015)

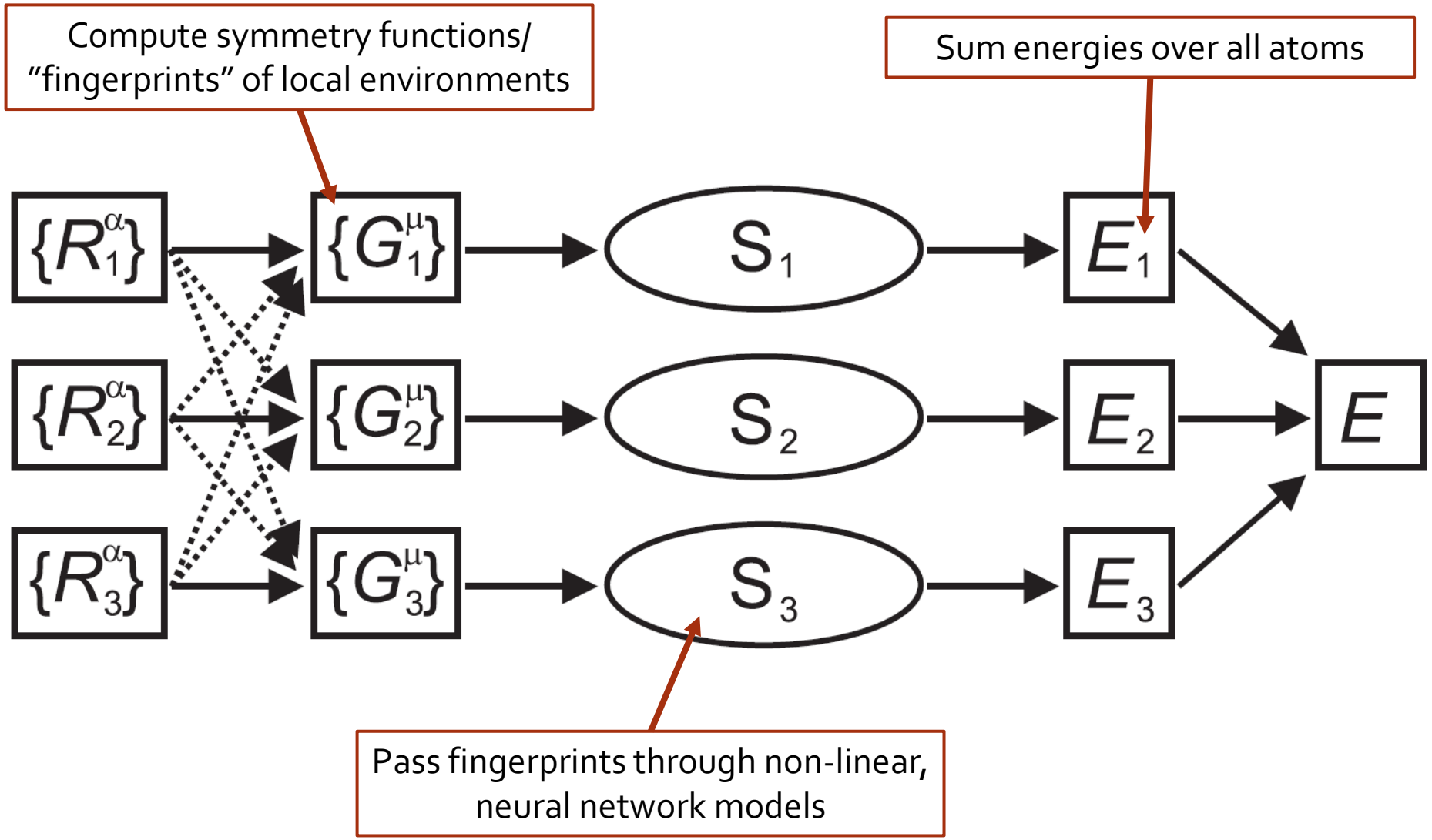
Simple example: Non-linear functions over sums of pairs

Example: Finnis-Sinclair potentials

$$E_i = \sqrt{\sum_j \rho(r_{ij})} + \sum_j \phi(r_{ij})$$

Accounts for coordination, but not angles

Neural Network Potentials: High Dimensional

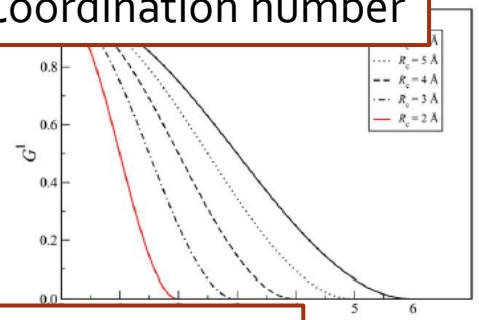


Reference: Behler and Parrinello. PRL (2007)

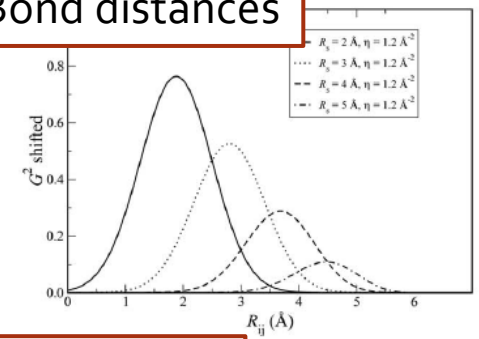
High Dimensional NNs: Subtle Details

Symmetry Functions

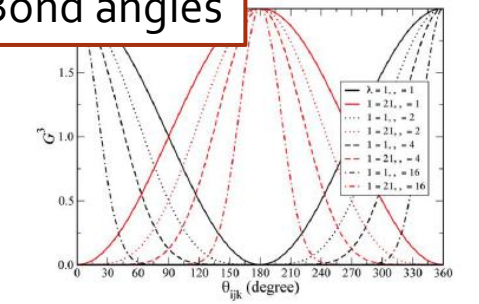
Coordination number



Bond distances

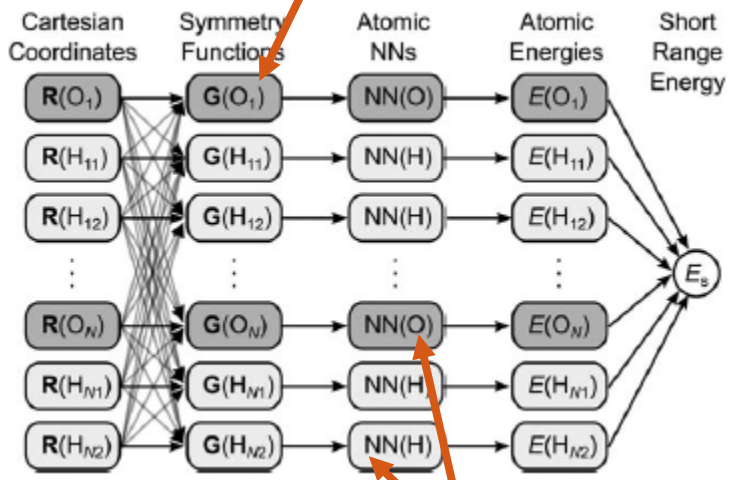


Bond angles



>1 Element in System

Different symmetry function for each atom type



Separate NN per element

Long-range Interactions

Problem: Electrostatic repulsion acts over long distance, but symmetry functions are local

Solution: Add in repulsion as a separate term, learn "short-range" energies

Learn with NN

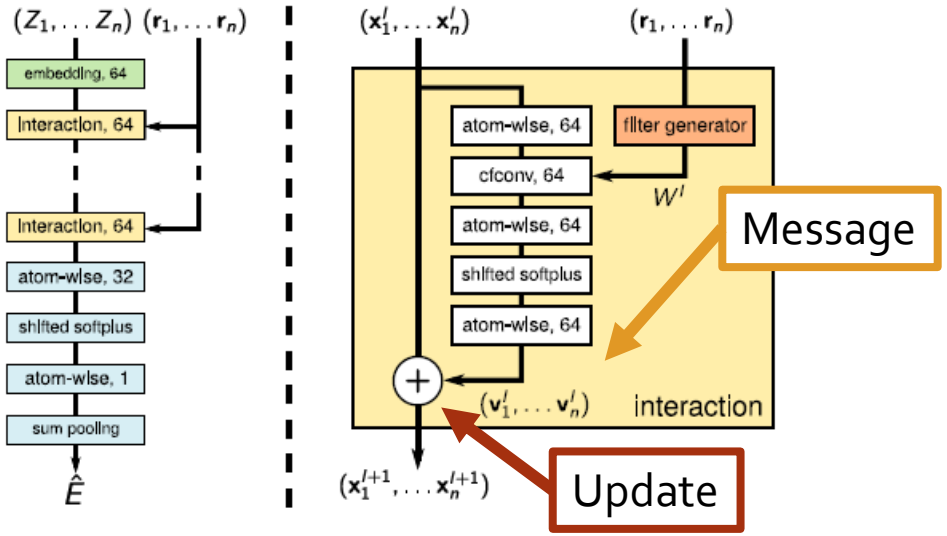
$$E = E_{short} + E_{elec}$$

Explicitly model

Figures from Behler IJQC (2015)

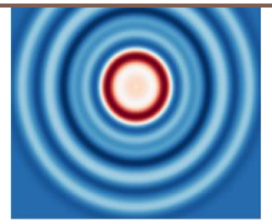
Other approaches to HDNN: ANI-1 (2017), MAISE (2017), Artrith and Urban (2016)

Neural Network Potentials: Message-Passing



$$\text{cfconv}(x_i) = \sum_j^{n_{atoms}} x_j \circ W(r_j - r_i)$$

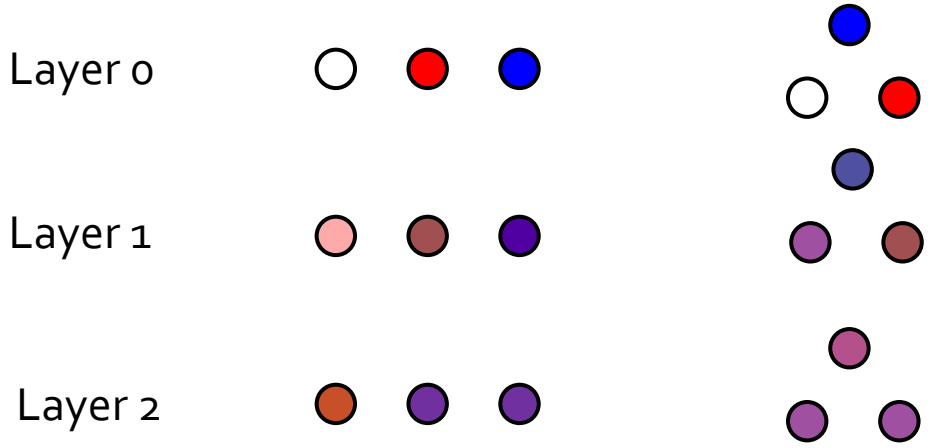
Continuous Filter



Ref: Schütt et al. JCP, (2018)

Subtle question: How do we account for angles if we only deal with pairwise interaction?

Solution: Message passing



Changes in bond angles are communicated on subsequent passes because they yield different environments

Gaussian Approximation Potentials: Two Key Innovations

Sparse Gaussian Process Potentials

Simple Concept: Kernel methods to express energy of a system

$$\mathrm{E}(\boldsymbol{b}) = \sum_n \alpha_n G(\boldsymbol{b}, \boldsymbol{b_n})$$

$\boldsymbol{b}$: Basis sets that describe local environment

$\boldsymbol{b_n}$: … that describe reference environments

$\boldsymbol{G}$: Kernel function

$\boldsymbol{\alpha_n}$: Coefficients for reference environment

Local Environment-Based Descriptors

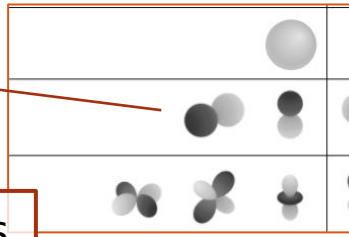
Easy Idea, Unfamiliar Mathematics

Step 1: Describe environments as neighbor densities

Step 2: Decompose into many-body spherical harmonics

$$p_i(\boldsymbol{R}) = \sum_j f_c(r_{ij}) \exp\left(-\frac{|\boldsymbol{R} - \boldsymbol{R_{ij}}|^2}{2\sigma^2}\right)$$

$$= \sum_{nlm} c_{nlm} g_n(R) Y_{lm}(\boldsymbol{R})$$



Step 3: Sum over direction-dependent terms

$$b_{n_1 n_2 l} = \sum_{m=l}^{l} c^*_{n_1 lm}\, c_{n_2 lm}$$

Refs: Bartók et al. PRL. (2010), Bartók et al. PRB. (2013)

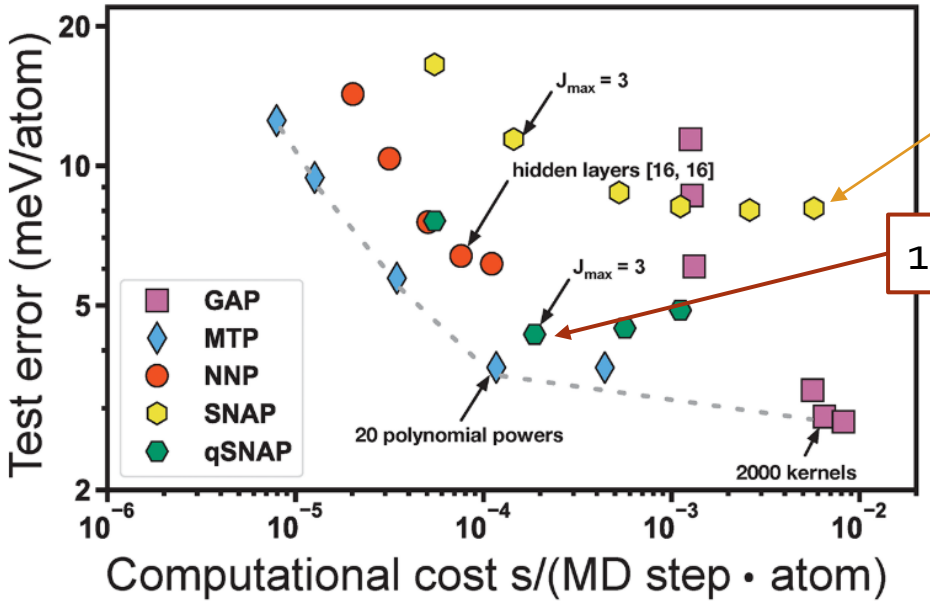
Linear-regression-based potentials

Problem with GAP: Execution time scales with number of reference points

Alternative Solution: Use linear or polynomial regression

Key method: Spectral Neighbor Analysis Potentials (SNAP)

- Similar features (4D instead of 3D harmonics), but faster computation



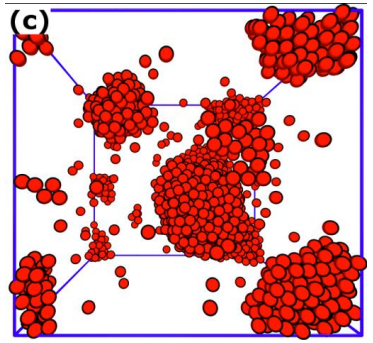
Control cost by adding more harmonics

100x speedup, but loss in accuracy

Reference: Zuo et al. JPCA (2020) <- Highly recommended read!

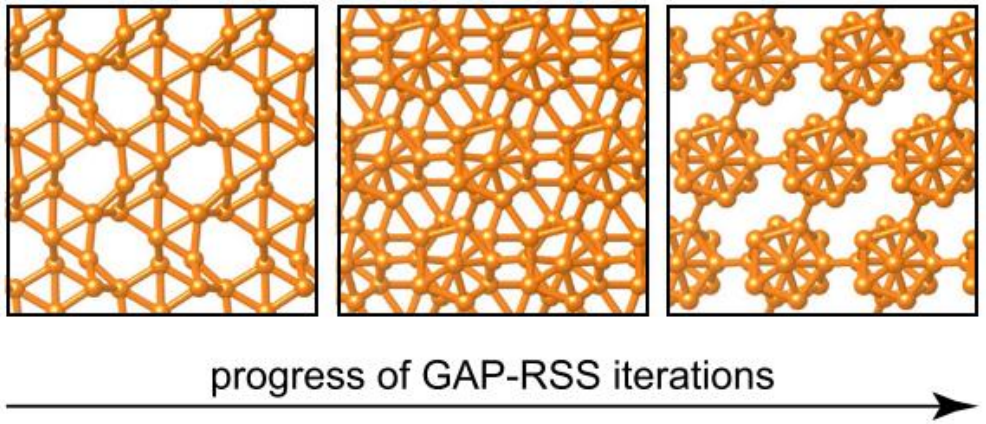
Fitting Techniques: Forced Sampling

Problem: Potentials have "blind spots"



Example: Potential trained on solids does incorrect things in liquids.
Ref: Ward et al. ArXiv (2012)

Solution: Run dynamics to find them, use to get new training data



Ref: Bernstein et al. npj Comp Mat (2019)

Codes becoming widely available

README.md

SchNetPack - Deep Neural Networks for Atomistic Systems

build passing | codecov 76% | code style black

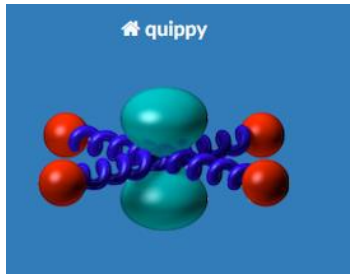
potfit

potfit wiki

open source force-matching

AMP

quippy



$\hat{H}(\vec{R})|\Psi\rangle = E|\Psi\rangle$

TORCHANI

MAML

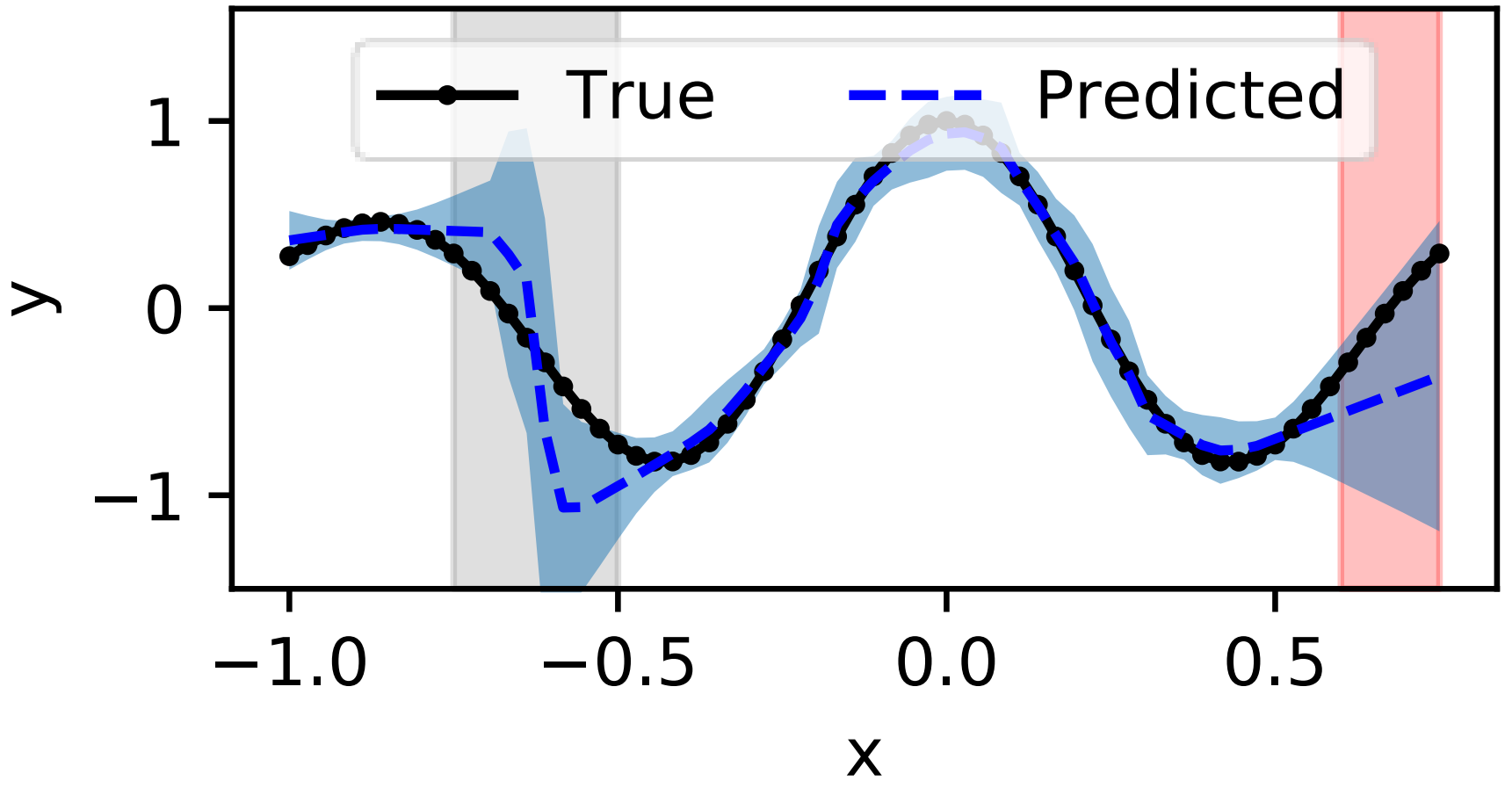
README.md

n2p2 - A neural network potential package

DOI 10.5281/zenodo.1344446 | release v2.1.0 | build passing | codecov 72% | License GPLv3

SIDE LESSON:
DOMAIN OF APPLICABILITY

It is hard to tell when a prediction is “untrustworthy”



Many routes to assessing uncertainty

Disagreement between ensembles

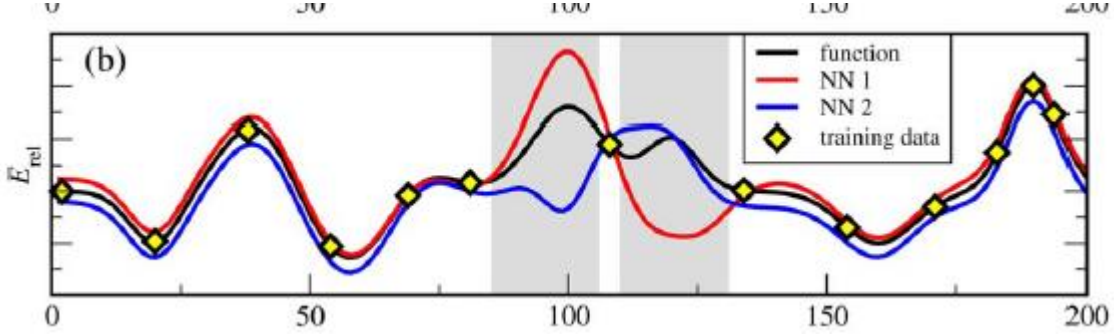


Figure: Behler. IJQC. (2015)

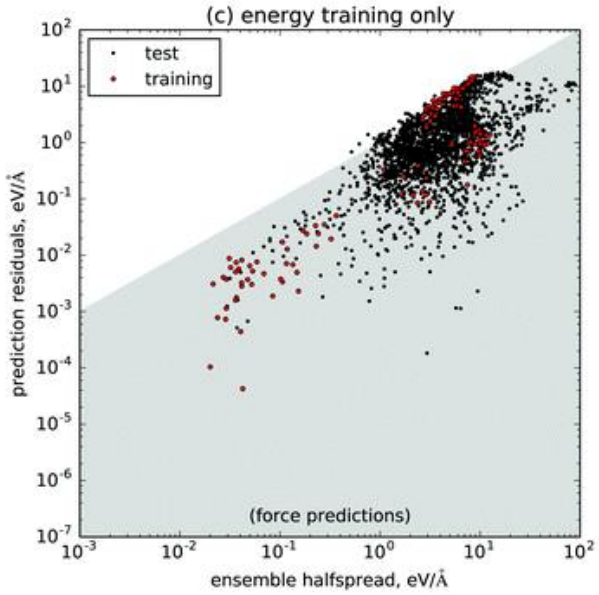


Figure: Peterson et al. PCCP. (2017)

Distance from training points

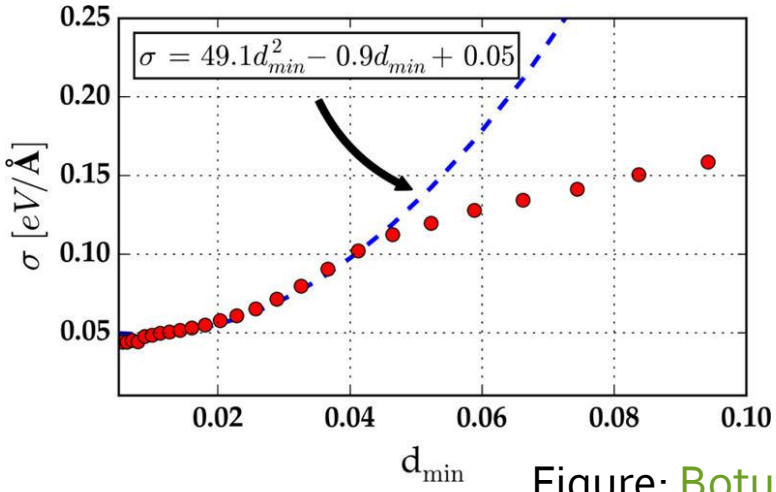


Figure: Botu et al. JPCC (2017)

"Extrapolation Index"

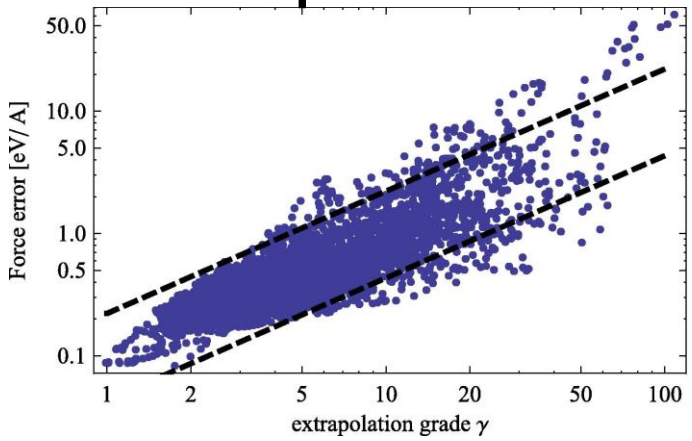
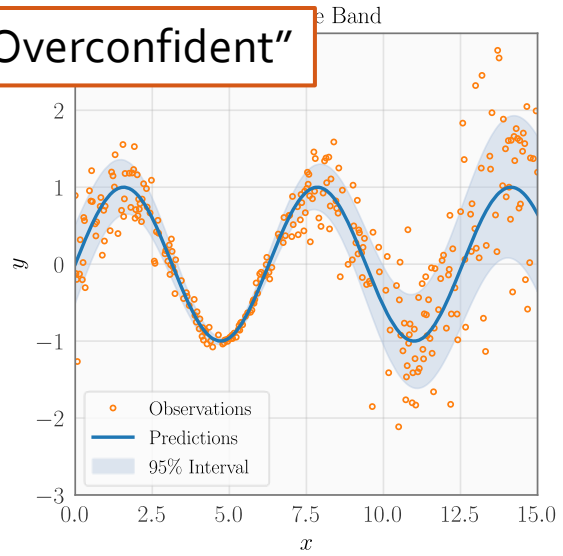
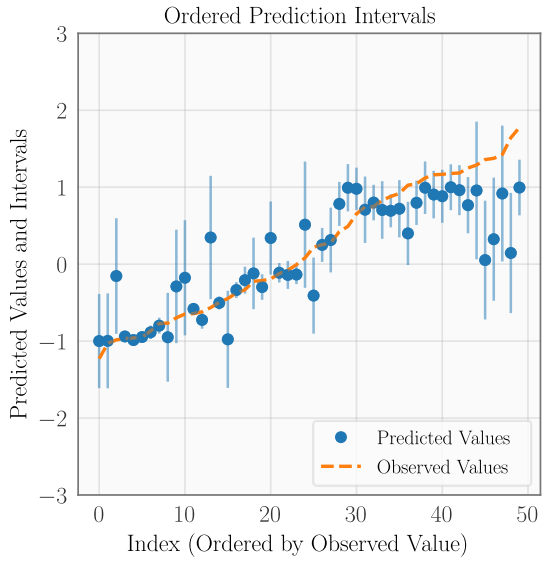


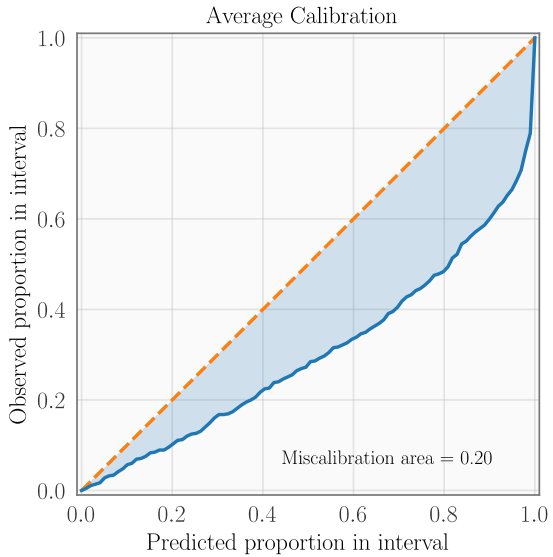
Figure: Podryabinkin and Shapeev. CMS (2017)

You can measure uncertainty metric quality

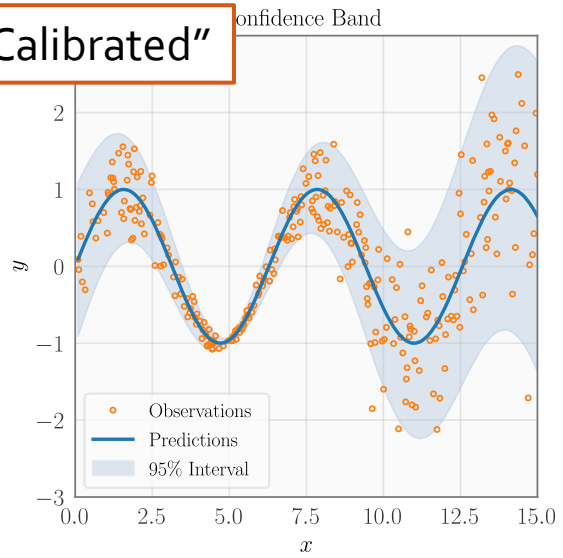
"Overconfident"

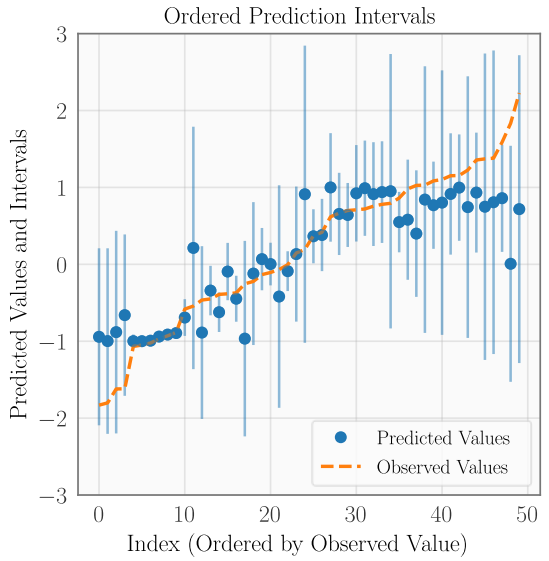


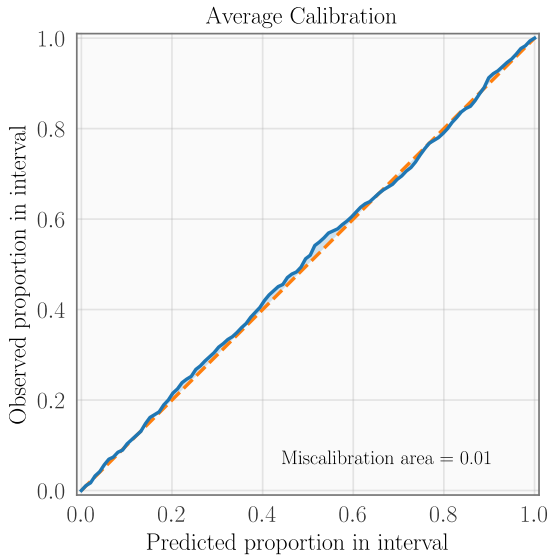




"Calibrated"







https://github.com/uncertainty-toolbox/uncertainty-toolbox

Take-away points

ML vs "Semi-empirical"

"Semi-empirical"

- Fixed model form
- Parameters fit using global optimization

"ML Potentials"

- Learned model form
- Parameters fit using supervised learning

Types of ML Potentials

Neural Networks:

- High-dimensional
- Message passing

Kernel Learning

- GAP
- AGNI

Polynomial Regression

- SNAP
- Moment-tensor potentials

Domain of Applicability

Many things to consider:

- Distance to training set
- Model uncertainty
- Would it update model?