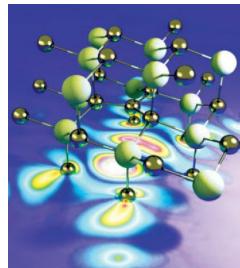


Learning Many-body Potential Energy Functions for Condensed Matter Systems

Ming Chen
University of California, Berkeley

Multiscale Modeling in Condensed Matter

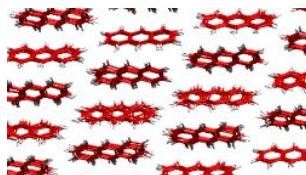
Timescale



ab initio

A. Boretti, *Nature Photonics.* **8**, 88 (2014).

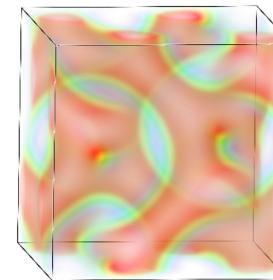
nm



All Atom Molecular Mechanics



μm



Coarse-Graining

F. J. Van Eerden *et al*, *Nature Communications.* **8**, 15214 (2017).

mm

System Size

Continues Models

K. T. Delaney *et al*, *Comp Phys Comm.* **184**, 2102 (2013).

From *ab initio* to Molecular Mechanics

Ab initio: Solving the Schrödinger Equation (Time-dependent/independent)

“Accurate”

“Transferable”

Slow (Scaling,
Prefactor)

From *ab initio* to Molecular Mechanics

Ab initio: Solving the Schrödinger Equation (Time-dependent/independent)



Many problems only require potential energy function, i.e.

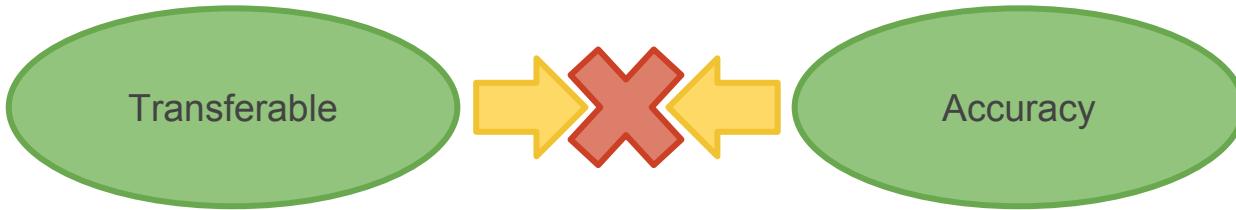
$$\mathbf{x} = (\mathbf{r}_1, \dots, \mathbf{r}_{N_{atom}}) \in \mathbb{R}^{3N_{atom}} \rightarrow \begin{matrix} \text{Energy (Force, Hessian, etc)} \\ E(\mathbf{x}) \quad \mathbf{f}(\mathbf{x}) \quad \mathbf{H}(\mathbf{x}) \end{matrix}$$

System Cartesian Coordinates

Potential Energy Function (PEF)

A black arrow points from the text "Potential Energy Function (PEF)" to the symbol $E(\mathbf{x})$ in the matrix above.

Transferable Force Field



It looks like this...

$$\begin{aligned} U = & \sum_{bonds} \kappa_b (r - r^0)^2 + \sum_{UB} \kappa_{UB} (r_{13} - r_{13}^0)^2 + \\ & \sum_{angles} \kappa_\theta (\theta - \theta^0)^2 + \sum_{dihedrals} \kappa_\phi (1 + \cos(n\phi - \phi^0)) \\ & + \sum_{impropers} \kappa_\omega (\omega - \omega^0)^2 + \\ & \sum_{nonbound} \epsilon_{ij} \left[\left(\frac{R_{ij}^0}{r_{ij}} \right)^{12} - \left(\frac{R_{ij}^0}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{\varepsilon r_{ij}} \end{aligned}$$

CHARMM, AMBER, OPLS_AA,
GROMOS and many others

- Reasonably good for modeling some biomolecules like proteins.
- Need careful benchmark in many condensed matter systems, e.g. molecular crystals

Challenges in Developing Transferable Force Fields for Condensed Matter Systems

- High quality data for different systems
 - Limited experimental data
 - Limited accurate *ab initio* calculations
- Complicated short range interactions
 - Many-body effect
 - Breaking/forming covalent bonds
- Polarization in long range interactions (polarizable models)

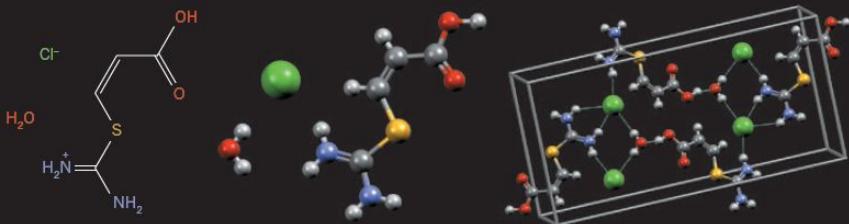


New ways to construct many-body potential energy functions.

System-specific Potential Energy Functions

CRYSTAL CHALLENGE

The 3D structure that a molecule adopts in a crystal is very difficult to predict — but defines what properties the molecule has.



The structural formula of a molecule reveals which atoms are connected at a 2D level.

Chemists are making progress at predicting how complex molecules will assemble in 3D space — there are millions of possibilities.

The 3D orientation repeats in a crystalline lattice with a structure that dictates the molecule's mechanical, chemical and physical properties.

CHEMISTRY

Software predicts crystal structures

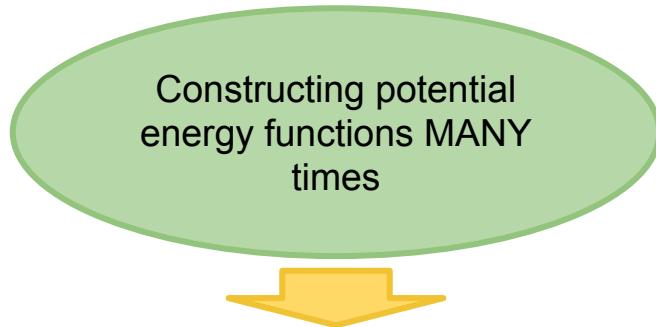
Chemists have succeeded at a fiendish task — forecasting how complex molecules will assemble in 3D.

Target	Chemical diagram	Crystallization conditions, remarks and clarifications
(XXII)		Crystallized from an acetone/water mixture; chiral-like character due to potential flexibility of the six-membered ring, but no chiral precursors used in synthesis.
(XXIII)		Five known polymorphs (A–E); three $Z' = 1$ (A, B, D), two $Z' = 2$ (C and E). The most stable polymorphs at 257 and 293 K are both $Z' = 1$. Crystallization conditions include slow evaporation of acetone solution and of ethyl acetate: water mixture.
(XXIV)		Crystallized from 1 M HCl solution. The substituents of the C=C double bond are in the cis configuration.
(XXV)		Slow evaporation of a methanol solution, which contained a racemic mixture of the enantiomers of Tröger's base.
(XXVI)		Slow evaporation from 1:1 mixture of hexane and dichloromethane. No chiral precursors used in synthesis.

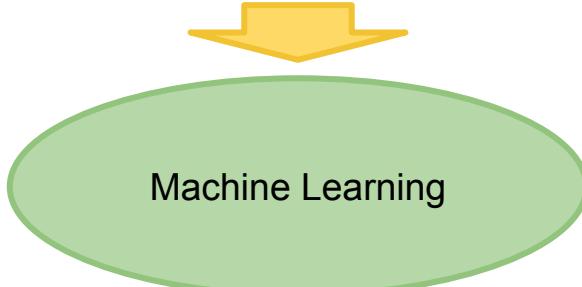
System-specific Potential Energy Functions

Team	Members	(XXII)		(XXIII)								(XXIV)		(XXV)		(XXVI)	
		L1	L2	A		B		C		D		E		L1	L2	L1	L2
1	Chadha & Singh	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
2	Cole <i>et al.</i>	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
3	Day <i>et al.</i>	3	1	23	—	—	75	—	—	75	—	—	—	—	—	—	—
4	Dzyabchenko	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
5	van Eijck	4	—	83	—	20	—	—	—	—	—	—	—	1	—	—	—
6	Elking & Fusti-Molnar	—	—	—	—	78	—	—	(73)	—	—	—	—	—	8	—	1
7	van den Ende, Cuppen <i>et al.</i>	9	90	—	—	—	—	—	—	—	—	—	—	—	—	—	—
8	Facelli <i>et al.</i>	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
9	Obata & Goto	2	—	—	—	13	—	—	(66)	—	—	—	—	—	—	—	—
10	Hofmann & Kuleshova	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
11	Lv, Wang, Ma	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
12	Marom <i>et al.</i>	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
13	Mohamed	1	—	—	88	—	—	—	—	—	—	—	—	—	—	—	—
14	Neumann, Kendrick, Leusen	2	—	26	85	2	4	—	6	11	39	—	—	2	6	1	1
15	Pantelides, Adjiman <i>et al.</i>	6	—	70	—	13	—	—	—	—	—	—	—	1	—	—	—
16	Pickard <i>et al.</i>	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
17	Podeszwa <i>et al.</i>	8	3	—	—	—	—	—	—	—	—	—	—	—	—	—	—
18	Price <i>et al.</i>	6	2	—	—	1	2	—	—	85	44	—	—	1	1	2	1
19	Szalewicz <i>et al.</i>	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
20	Tuckerman, Szalewicz <i>et al.</i>	4	—	—	—	—	—	—	—	—	—	—	—	2	—	—	—
21	Zhu, Oganov, Masunov	3	—	—	—	—	—	—	—	—	—	—	—	2	—	—	—
22	Boese	*	—	*	*	*	*	*	*	*	*	*	*	*	*	*	*
23	Brandenburg & Grimme	—	—	—	—	11	1	—	—	—	—	—	—	2	—	—	—
24	Szalewicz <i>et al.</i>	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
25	Tkatchenko <i>et al.</i>	3	1	—	—	2	5	—	—	14	2	—	—	1	—	—	—

System-specific Potential Energy Functions



- General function forms that can be used in many different systems.
- Many parameters: flexible models
- Systematic and efficient ways to fit and validate



Learning Potential Energy Functions

- Explicit, deterministic, smooth almost sure

$$\mathbf{f}(\mathbf{x}) = -\nabla E(\mathbf{x}) \quad \mathbf{H}(\mathbf{x}) = \nabla \nabla E(\mathbf{x})$$

- Relatively easy to evaluate
- Efficient and easy to train
- Transferable with respect to different system size
- Satisfy basic symmetry requirements
 - Permutational Symmetry
 - Translational/Rotational Symmetry

Learning Potential Energy Functions

- Explicit, deterministic, smooth almost sure

$$\mathbf{f}(\mathbf{x}) = -\nabla E(\mathbf{x}) \quad \mathbf{H}(\mathbf{x}) = \nabla \nabla E(\mathbf{x})$$

- Relatively easy to evaluate
- Efficient and easy to train
- Transferable with respect to different system size
- Satisfy basic symmetry requirements
 - Permutational Symmetry
 - Translational/Rotational Symmetry

Learning Potential Energy Functions

- Explicit, deterministic, smooth almost sure

$$\mathbf{f}(\mathbf{x}) = -\nabla E(\mathbf{x}) \quad \mathbf{H}(\mathbf{x}) = \nabla \nabla E(\mathbf{x})$$

- Relatively easy to evaluate
- Efficient and easy to train
- Transferable with respect to different system size
- Satisfy basic symmetry requirements
 - Permutational Symmetry
 - Translational/Rotational Symmetry

Learning Potential Energy Functions

- Explicit, deterministic, smooth almost sure

$$\mathbf{f}(\mathbf{x}) = -\nabla E(\mathbf{x}) \quad \mathbf{H}(\mathbf{x}) = \nabla \nabla E(\mathbf{x})$$

- Relatively easy to evaluate
- Efficient and easy to train
- Transferable with respect to different system size
- Satisfy basic symmetry requirements
 - Permutational Symmetry
 - Translational/Rotational Symmetry

Learning Potential Energy Functions

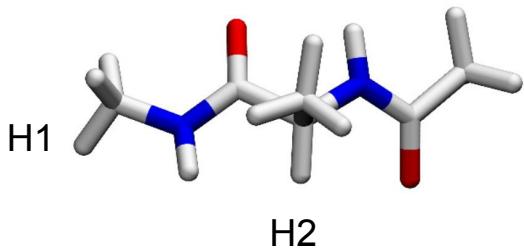
- Explicit, Deterministic, Smooth almost sure

$$\mathbf{f}(\mathbf{x}) = -\nabla E(\mathbf{x}) \quad \mathbf{H}(\mathbf{x}) = \nabla \nabla E(\mathbf{x})$$

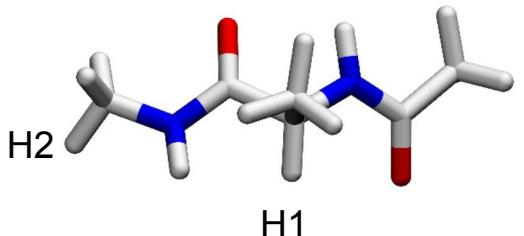
- Relatively easy to evaluate
- Efficient and easy to train
- Transferable with respect to different system size
- **Satisfy basic symmetry requirements**
 - Permutational Symmetry
 - Translational/Rotational Symmetry

Symmetry requirements

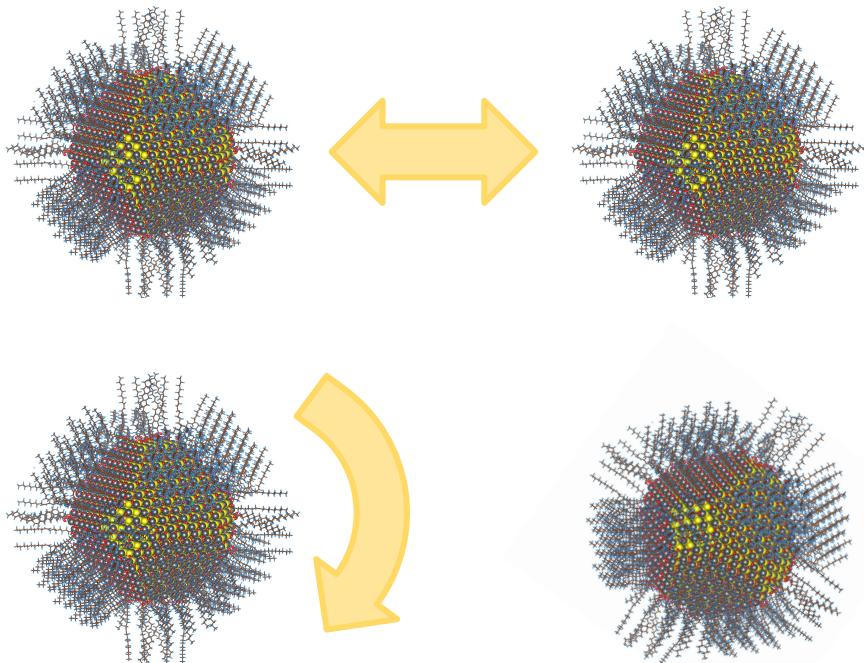
Permutational Symmetry



Exchanging labels doesn't affect energy (from quantum mechanics)



Translational/Rotational Symmetry

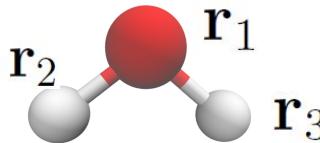


Weak Correlated System: Many-body Expansion

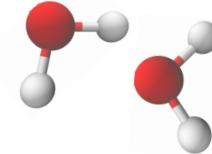
$$U = \sum_i^N U_{1B}(\mathbf{R}_i) + \sum_{i < j} U_{2B}(\mathbf{R}_i, \mathbf{R}_j) + \sum_{i < j < k} U_{3B}(\mathbf{R}_i, \mathbf{R}_j, \mathbf{R}_k) + \dots + U_{NB}(\mathbf{R}_1, \dots, \mathbf{R}_N)$$

$$\mathbf{R}_1 = (\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$$

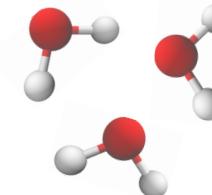
1 Body (1B)



2 Bodies (2B)



3 Bodies (3B)



Y. Wang, X. Huang, B. C. Shepler, B. J. Braams, J. M. Bowman, *J. Chem. Phys.* **134**, 094509 (2011)

V. Babin,,C. Leforestier, F. Paesani, *J. Chem. Theory Comput.* **9**, 5395 (2013)

V. Babin, G. R. Medders, F. Paesani, *J. Chem. Theory Comput.* **10**, 1599 (2014)

G. R. Medders, V. Babin, F. Paesani, *J. Chem. Theory Comput.* **10**, 2906

Weak Correlated System: Many-body Expansion

Already good enough

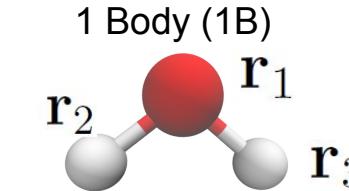
$$U = \sum_i^N U_{1B}(\mathbf{R}_i) + \sum_{i < j} U_{2B}(\mathbf{R}_i, \mathbf{R}_j)$$

$$+ \sum_{i < j < k} U_{3B}(\mathbf{R}_i, \mathbf{R}_j, \mathbf{R}_k) + \dots + U_{NB}(\mathbf{R}_1, \dots, \mathbf{R}_N)$$

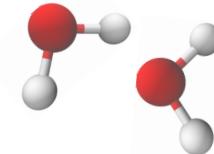
Complicated: High dimensional/Many-body Correlation

Y. Wang, X. Huang, B. C. Shepler, B. J. Braams, J. M. Bowman, *J. Chem. Phys.* **134**, 094509 (2011)

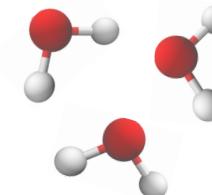
$$\mathbf{R}_1 = (\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$$



2 Bodies (2B)



3 Bodies (3B)



Emperical

V. Babin,,C. Leforestier, F. Paesani, *J. Chem. Theory Comput.* **9**, 5395 (2013)

V. Babin, G. R. Medders, F. Paesani, *J. Chem. Theory Comput.* **10**, 1599 (2014)

G. R. Medders, V. Babin, F. Paesani, *J. Chem. Theory Comput.* **10**, 2906

Polynomial Regression in Many-body Water Potential

$$U_{2B}(\mathbf{R}_i, \mathbf{R}_j) = s(\mathbf{R}_i, \mathbf{R}_j) \sum_{l=1}^{N_{poly}} C_l \phi_l(\mathbf{R}_i, \mathbf{R}_j) + \text{long range}$$

$$U_{3B}(\mathbf{R}_i, \mathbf{R}_j, \mathbf{R}_k) = s(\mathbf{R}_i, \mathbf{R}_j, \mathbf{R}_k) \sum_{l=1}^{N_{poly}} C_l \phi_l(\mathbf{R}_i, \mathbf{R}_j, \mathbf{R}_k)$$

Polynomial Regression in Many-body Water Potential

$$U_{2B}(\mathbf{R}_i, \mathbf{R}_j) = s(\mathbf{R}_i, \mathbf{R}_j) \sum_{l=1}^{N_{poly}} C_l \phi_l(\mathbf{R}_i, \mathbf{R}_j) + \text{long range}$$

$$U_{3B}(\mathbf{R}_i, \mathbf{R}_j, \mathbf{R}_k) = s(\mathbf{R}_i, \mathbf{R}_j, \mathbf{R}_k) \sum_{l=1}^{N_{poly}} C_l \phi_l(\mathbf{R}_i, \mathbf{R}_j, \mathbf{R}_k) + \text{????}$$

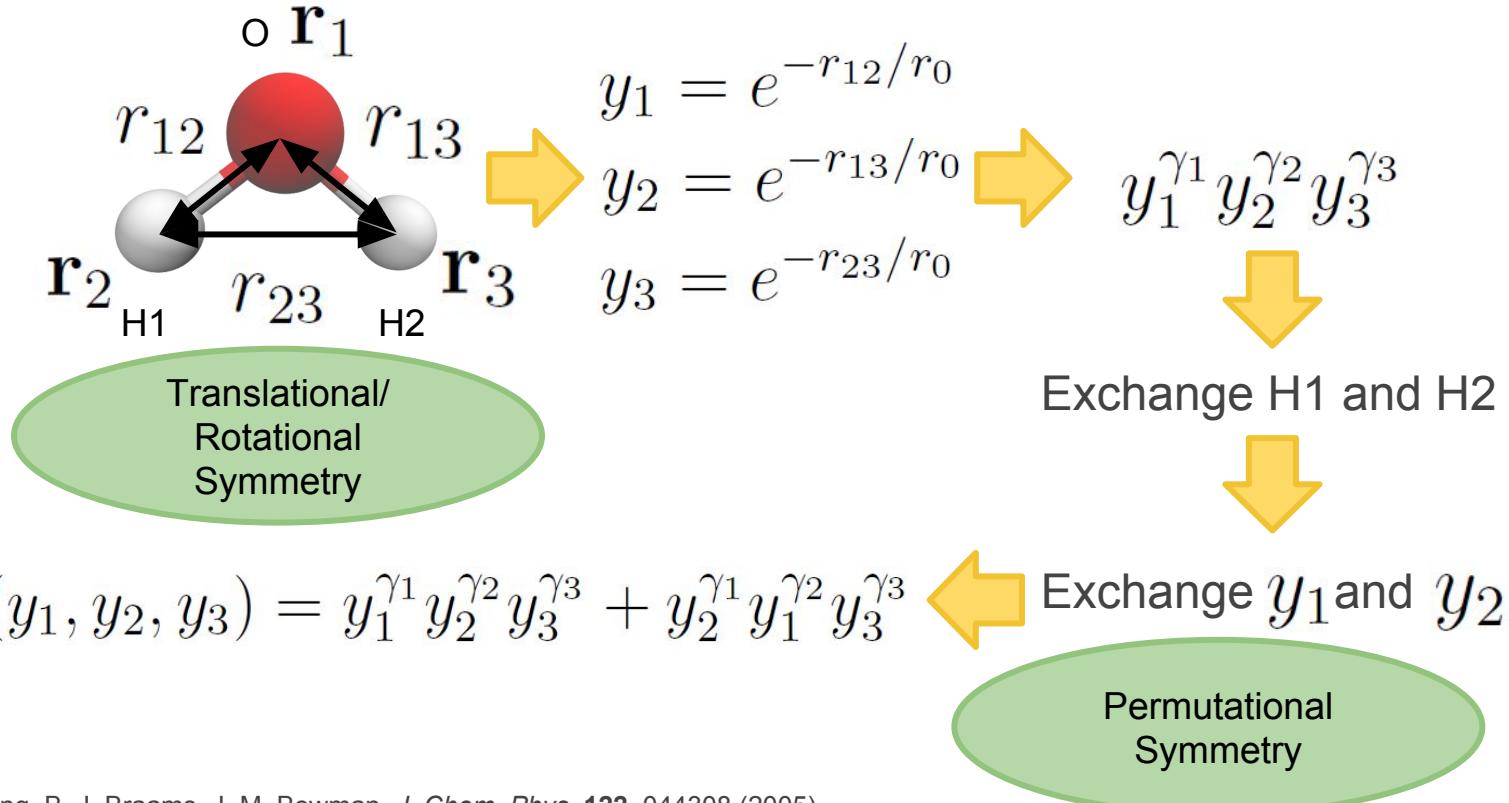
$$\chi = \sum_{n=1}^{N_{sample}} w(E)(U_{nB}(\mathbf{x}_n) - E_{nB}^{qm}(\mathbf{x}_n))^2 + \lambda \sum_{l=1}^{N_{poly}} C_l^2$$

V. Babin,,C. Leforestier, F. Paesani, *J. Chem. Theory Comput.* **9**, 5395 (2013)

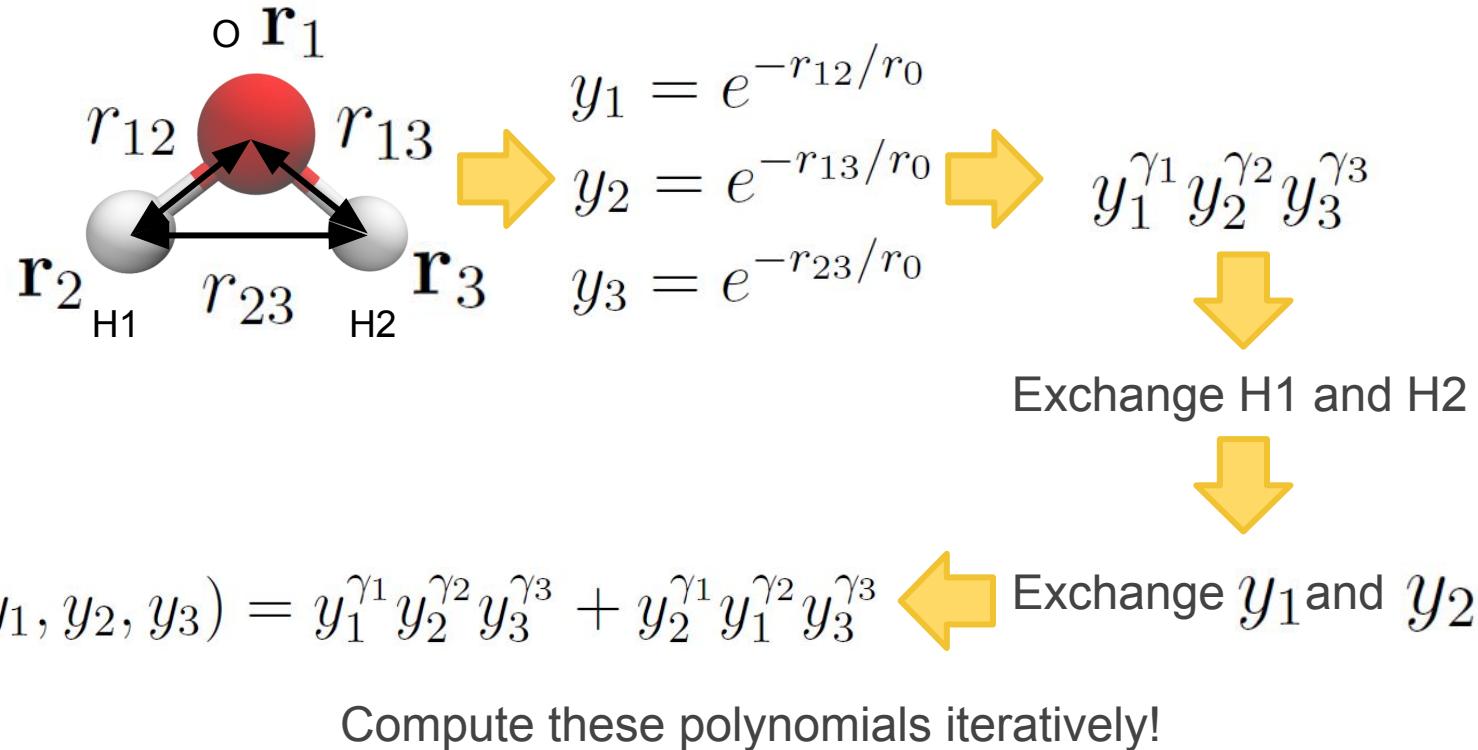
V. Babin, G. R. Medders, F. Paesani, *J. Chem. Theory Comput.* **10**, 1599 (2014)

G. R. Medders, V. Babin, F. Paesani, *J. Chem. Theory Comput.* **10**, 2906

Permutationally Invariant Polynomial



Permutationally Invariant Polynomial



Water: From Cluster to Bulk

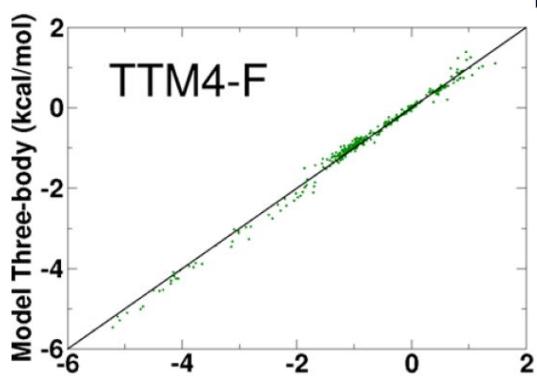
MB-pol potential: polynomial+polarizable model (TTM4-F)

C. J. Burnham, et al, *J. Chem. Phys.* **128**, 154519 (2008)

$$U = \sum_i^N U_{1B}(\mathbf{r}_i) + \sum_{i < j} U_{2B}^{short}(\mathbf{r}_i, \mathbf{r}_j) + \sum_{i < j < k} U_{3B}^{short}(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k) + U_{TTM} + U_{disp}$$

2, 3, many body contributions

Long range dispersion



V. Babin,,C. Leforestier, F. Paesani, *J. Chem. Theory Comput.* **9**, 5395 (2013)

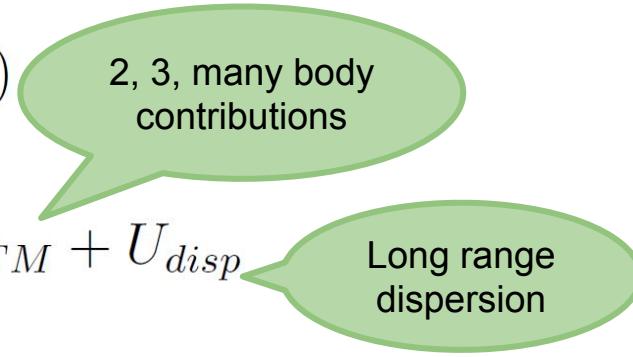
V. Babin, G. R. Medders, F. Paesani, *J. Chem. Theory Comput.* **10**, 1599 (2014)

G. R. Medders, V. Babin, F. Paesani, *J. Chem. Theory Comput.* **10**, 2906

Water: From Cluster to Bulk

MB-pol potential: polynomial+polarizable model (TTM4-F)

Learn the
“Correction”

$$U = \sum_i^N U_{1B}(\mathbf{r}_i) + \sum_{i < j} U_{2B}^{short}(\mathbf{r}_i, \mathbf{r}_j) + \sum_{i < j < k} U_{3B}^{short}(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k) + U_{TTM} + U_{disp}$$


Number of Polynomials

- WHBB: 1380 (5th Order)+5849 (6th Order)
- MB-pol: 13(2nd Order)+202 (3rd Order)+948 (4th Order)

Possible for Bulk
Calculation

Y. Wang, X. Huang, B. C. Shepler, B. J. Braams, J. M. Bowman, *J. Chem. Phys.* **134**, 094509 (2011)

V. Babin,,C. Leforestier, F. Paesani, *J. Chem. Theory Comput.* **9**, 5395 (2013)

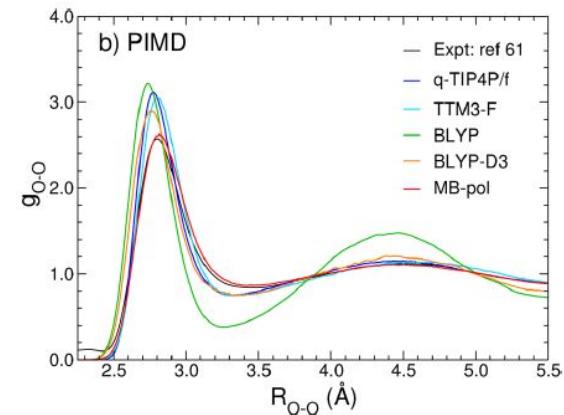
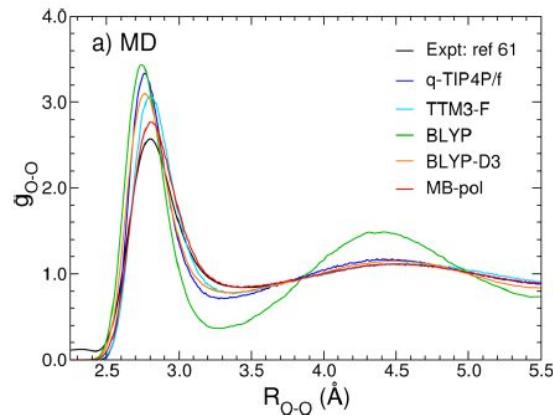
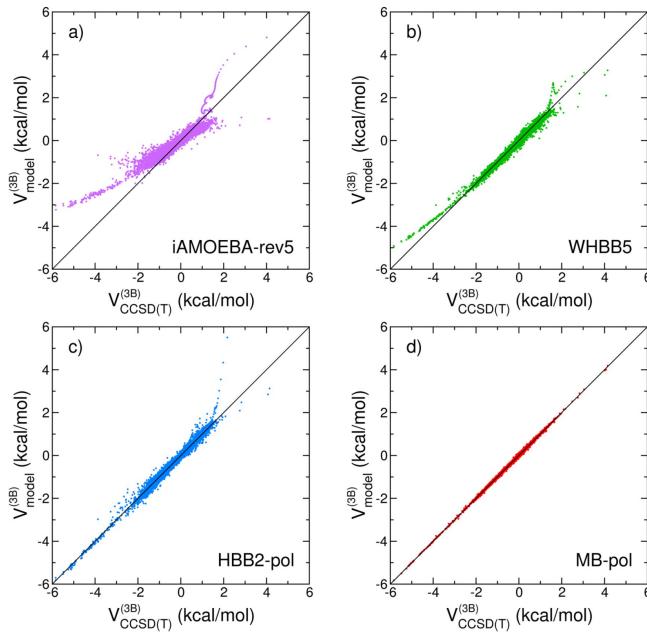
V. Babin, G. R. Medders, F. Paesani, *J. Chem. Theory Comput.* **10**, 1599 (2014)

G. R. Medders, V. Babin, F. Paesani, *J. Chem. Theory Comput.* **10**, 2906

Training MB-pol Potential

Training data:

- Configurations from simulations
- Energy calculated in CCSD(T)/CBS and CCSD(T)/aug-cc-pVTZ

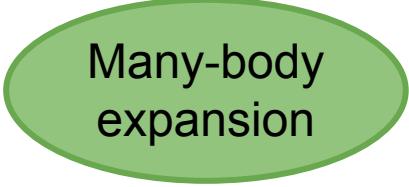


V. Babin,,C. Leforestier, F. Paesani, *J. Chem. Theory Comput.* **9**, 5395 (2013)

V. Babin, G. R. Medders, F. Paesani, *J. Chem. Theory Comput.* **10**, 1599 (2014)

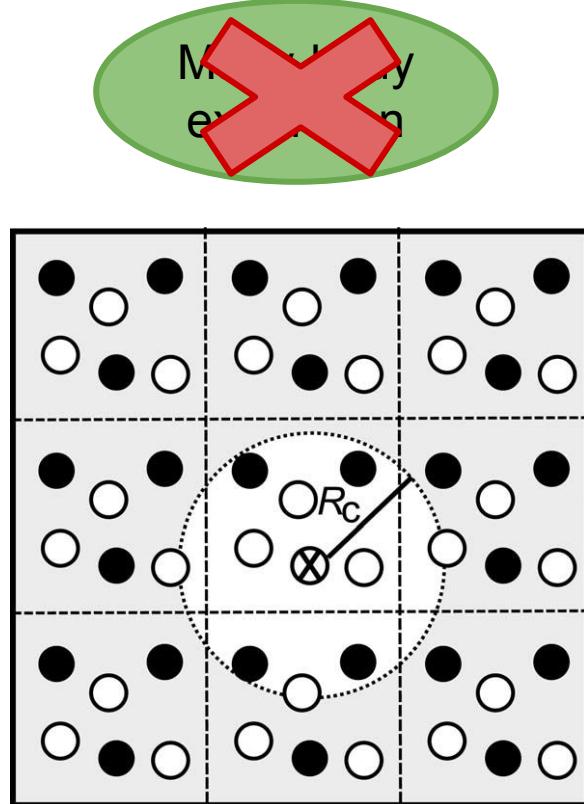
G. R. Medders, V. Babin, F. Paesani, *J. Chem. Theory Comput.* **10**, 2906

Strong Correlated System: Local Environment



Many-body
expansion

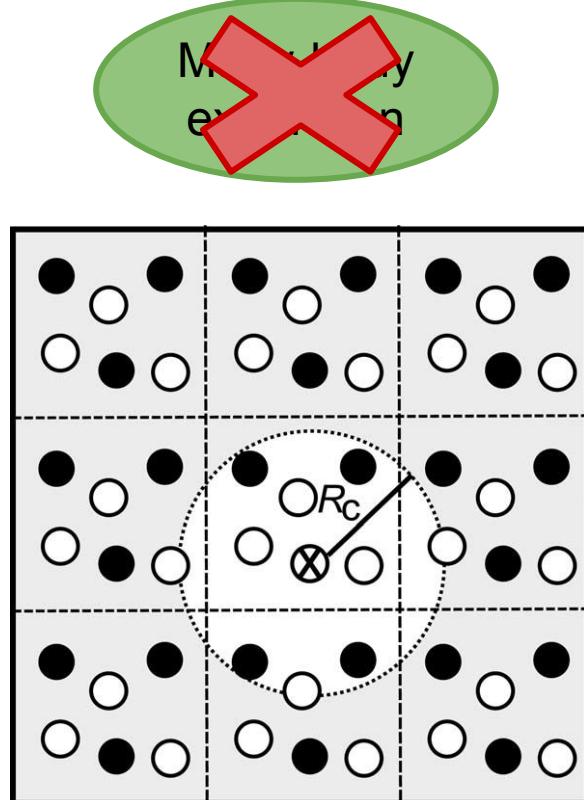
Strong Correlated System: Local Environment



$$E = \sum_{i=1}^{N_{atom}} E_i + \text{long range}$$

$$E_i = E_i(G_1, \dots, G_{N_{feature}})$$

Strong Correlated System: Local Environment



$$E = \sum_{i=1}^{N_{atom}} E_i + \text{long range}$$
$$E_i = E_i(G_1, \dots, G_{N_{feature}})$$

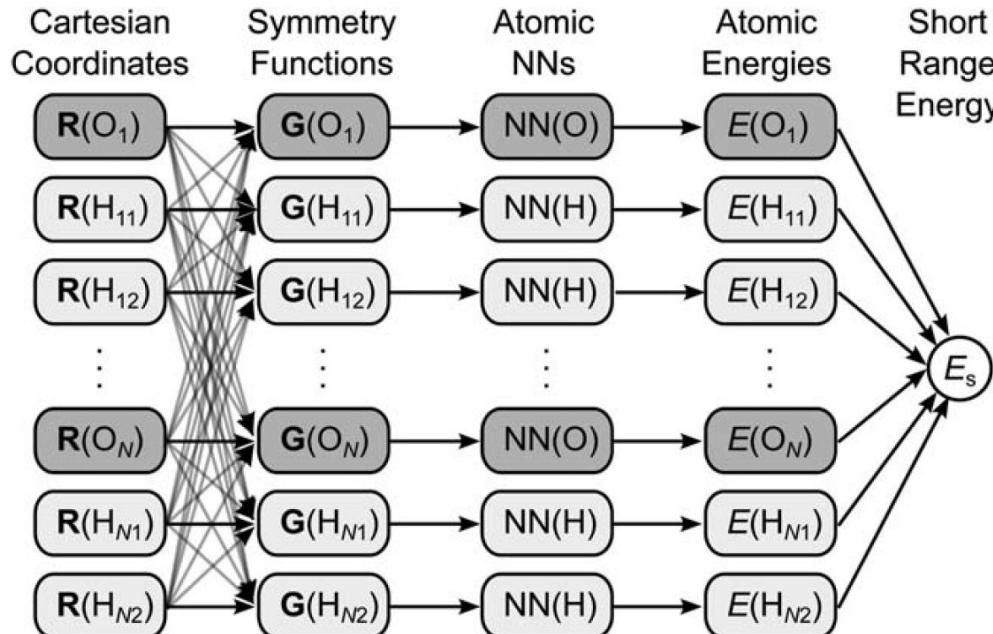
Permutational Symmetry

Rotational/
Translational
Symmetry

Arrows from the equation point to the two green ovals below it.

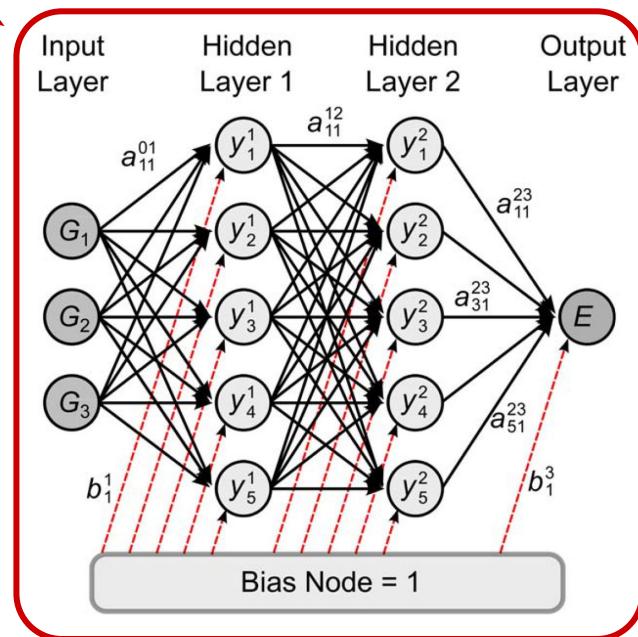
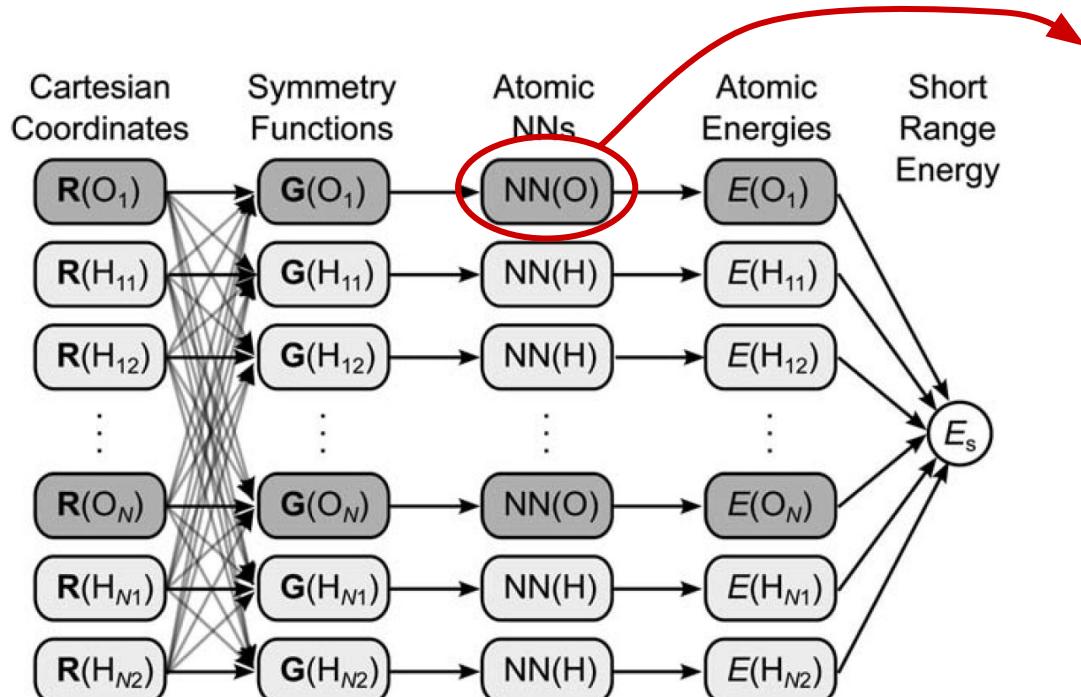
J. Behler and M. Parrinello, *Phys. Rev. Lett.* **98**, 146401 (2007)
J. Behler, *Int. J. Quantum Chem.*, **115**, 1032 (2015)

Modeling a PEF with Artificial Neural Network



$$\mathbf{G} = (G_1, \dots, G_{N_{feature}})$$

Modeling a PEF with Artificial Neural Network



$$\mathbf{G} = (G_1, \dots, G_{N_{feature}})$$

J. Behler and M. Parrinello, *Phys. Rev. Lett.* **98**, 146401 (2007)
J. Behler, *Int. J. Quantum Chem.*, **115**, 1032 (2015)

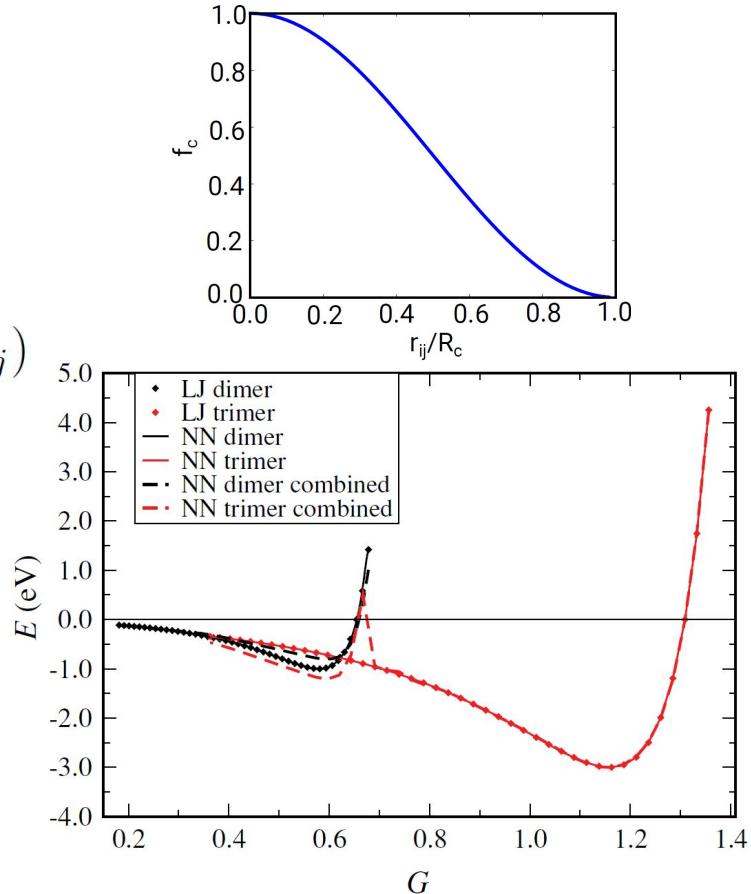
$$y_i^j = f_i^j \left(b_i^j + \sum_{k=1}^{N_{j-1}} a_{k,i}^{j,j-1} y_k^{j-1} \right)$$

Symmetry Functions as Features

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

$$G_i^{(1)} = \sum_{j=1}^{N_{atom}} f_c(r_{ij}), G_i^{(2)} = \sum_{j=1}^{N_{atom}} e^{-\eta(r_{ij}-R_s)^2} f_c(r_{ij})$$

$$\begin{aligned} G_i^{(3)} = & 2^{1-\zeta} \sum_{j \neq i} \sum_{k \neq i,j} (1 + \lambda \cos \theta_{ijk})^\zeta e^{-\eta(r_{ij}^2 + r_{ik}^2 + r_{jk}^2)} \\ & \times f_c(r_{ij}) f_c(r_{ik}) f_c(r_{jk}) \end{aligned}$$

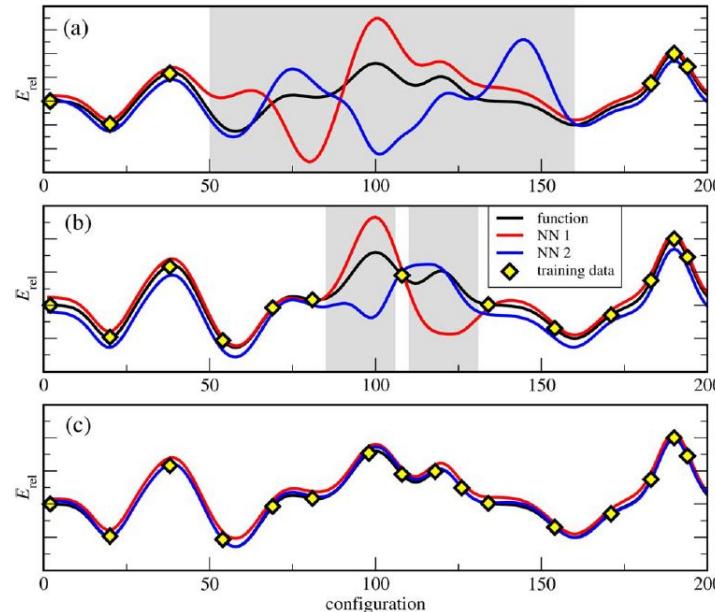
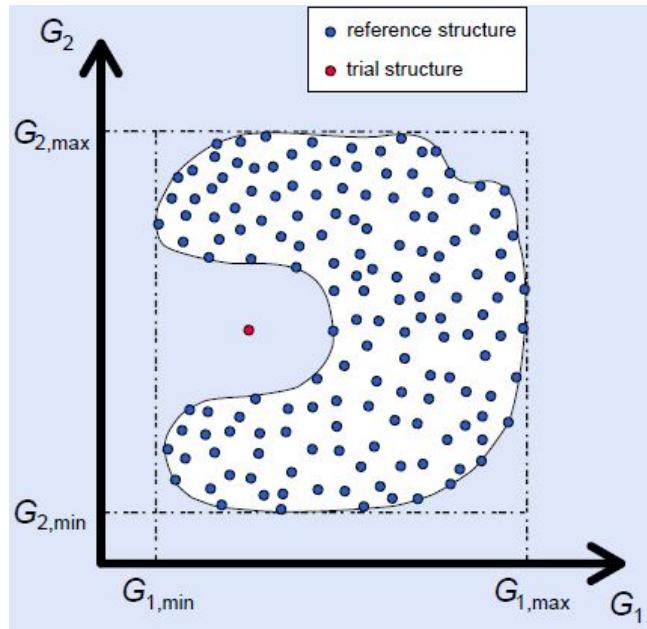


J. Behler, *J. Chem. Phys.*, **134**, 074106 (2011)

J. Behler, *Int. J. Quantum Chem.*, **115**, 1032 (2015)

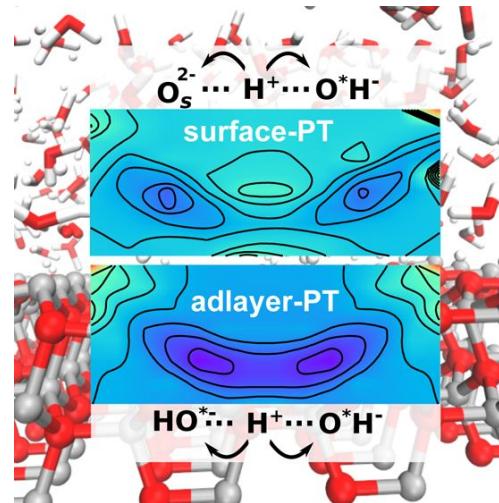
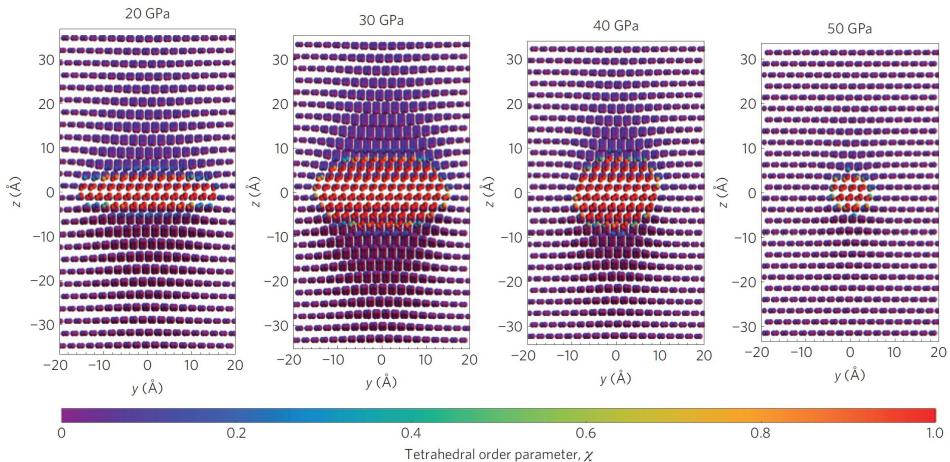
Training the Neural Network Potential

$$\Gamma = \frac{1}{N_{\text{struct}}} \sum_{i=1}^{N_{\text{struct}}} \left[(E_{\text{NN}}^i - E_{\text{Ref}}^i)^2 + \frac{\beta}{3N_{\text{atom}}^i} \sum_{j=1}^{3N_{\text{atom}}^i} (F_{j\text{NN}}^i - F_{j\text{Ref}}^i)^2 \right]$$

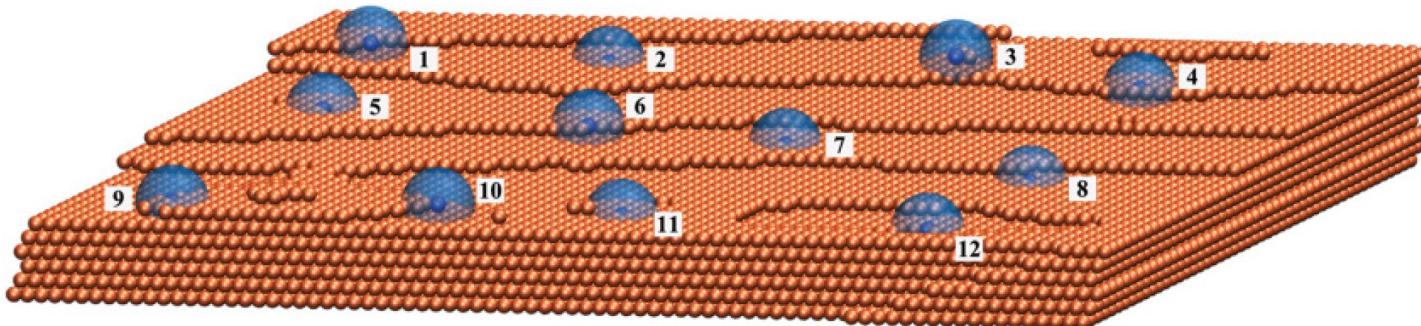


J. Behler, *Angew. Chem. Int. Ed.* 10.1002/anie.201703114 (2017)
J. Behler, *Int. J. Quantum Chem.*, **115**, 1032 (2015)

Examples of Neural Network Potential



V. Quaranta, M. Hellström
and J. Behler, *J. Phys. Chem. Lett.*, **8**, 1476 (2017)



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

R. Z. Khaliullin, H. Eshet, T. D. Kühne, J. Behler, and M. Parrinello, *Nat. Mater.*, **10**, 693 (2011)

Gaussian Approximation Potential

Gaussian Process Regression/Kernel Ridge Regression

$$E = \sum_{i=1}^{N_{atom}} E_i + \text{long range}$$

$$E_i = \sum_k^{N_{sample}} C_k K(\mathbf{G}, \mathbf{G}_k)$$

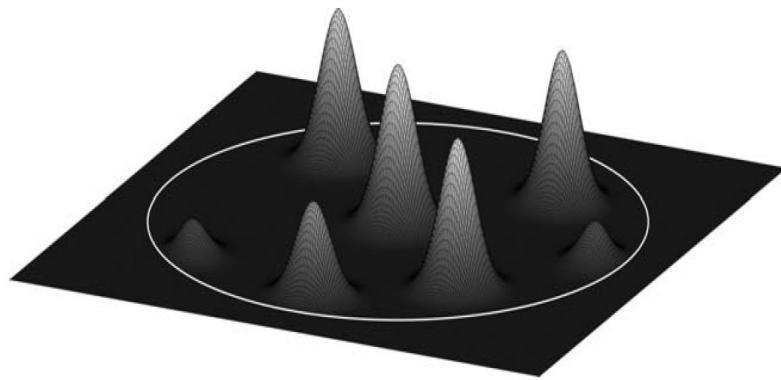
A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, *Phys. Rev. Lett.* **104**, 136403 (2010)
A. P. Bartók and G. Csányi, *Int. J. Quantum Chem.*, **115**, 1051 (2015)

- There are no direct atomic energies so the training are based on total energy and forces.
- Local environments are highly correlated. A sparse subset is selected as representative atomic neighbourhood environments

Descriptor: Smooth Overlap of Atomic Positions

$$\rho_i(\mathbf{r}) \equiv \sum_j^{\text{neigh.}} \exp\left(-\frac{|\mathbf{r} - \mathbf{r}_{ij}|^2}{2\sigma_{\text{atom}}^2}\right)$$

- Permutational symmetry ✓
- Translational symmetry ✓
- Rotational symmetry ✗



A. P. Bartók and G. Csányi, *Int. J. Quantum Chem.*, **115**, 1051 (2015)

A. P. Bartók, R. Kondor, and G. Csányi, *Phys. Rev. B*, **87**, 184115 (2013)

Descriptor: Smooth Overlap of Atomic Positions

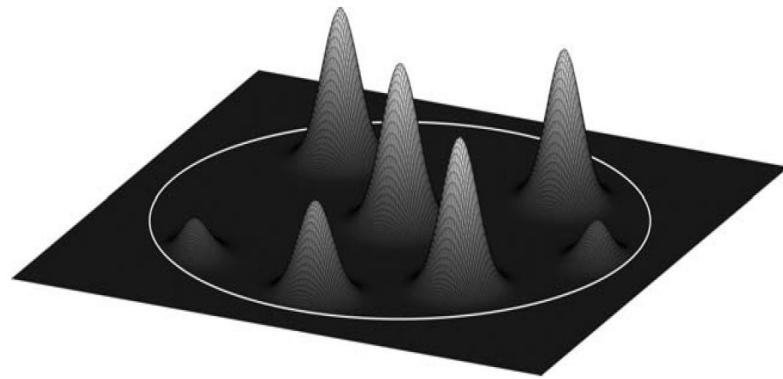
$$\rho_i(\mathbf{r}) \equiv \sum_j^{\text{neigh.}} \exp\left(-\frac{|\mathbf{r} - \mathbf{r}_{ij}|^2}{2\sigma_{\text{atom}}^2}\right)$$


$$\rho(\mathbf{r}) = \sum_i \exp(-\alpha|\mathbf{r} - \mathbf{r}_i|^2) = \sum_i \sum_{lm} c_{lm}^i(r) Y_{lm}(\hat{\mathbf{r}})$$

$$c_{lm}^i(r) \equiv 4\pi \exp[-\alpha(r^2 + r_i^2)] \iota_l(2\alpha r r_i) Y_{lm}^*(\hat{\mathbf{r}}_i)$$

Two densities: ρ and ρ'

$$S(\rho, \rho') = \int \rho(\mathbf{r}) \rho'(\mathbf{r}) d\mathbf{r}$$



Still **NOT** rotational invariant

A. P. Bartók and G. Csányi, *Int. J. Quantum Chem.*, **115**, 1051 (2015)
A. P. Bartók, R. Kondor, and G. Csányi, *Phys. Rev. B*, **87**, 184115 (2013)

Descriptor: Smooth Overlap of Atomic Positions

$$k(\rho, \rho') = \int |S(\rho, \hat{R}\rho')|^n d\hat{R}$$

Integrate over SO(3) group, with respect to Haar measure. Take n=2 as example

$$= \int d\hat{R} \left| \int \rho(\mathbf{r}) \rho'(\hat{R}\mathbf{r}) d\mathbf{r} \right|^n$$

$$S(\hat{R}) \equiv S(\rho, \hat{R}\rho') = \int d\mathbf{r} \rho(\mathbf{r}) \rho'(\hat{R}\mathbf{r})$$

$$= \sum_{i,i'} \sum_{\substack{l,m \\ l',m',m''}} D_{m'm''}^{l'}(\hat{R}) \int dr c_{lm}^{i*}(r) c_{l'm'}^{i'}(r) \int d\hat{\mathbf{r}} Y_{lm}^*(\hat{\mathbf{r}}) Y_{l'm''}(\hat{\mathbf{r}})$$

Wigner matrices, irreducible representations of SO(3)

$$= \sum_{i,i'} \sum_{l,m,m'} \tilde{I}_{mm'}^l(\alpha, r_i, r_{i'}) D_{mm'}^l(\hat{R}) = \sum_{l,m,m'} I_{mm'}^l D_{mm'}^l(\hat{R}),$$

$$\tilde{I}_{mm'}^l(\alpha, r_i, r_{i'})$$

$$= 4\pi \exp[-\alpha(r_i^2 + r_{i'}^2)/2] \iota_l(\alpha r_i r_{i'}) Y_{lm}(\hat{\mathbf{r}}_i) Y_{lm}^*(\hat{\mathbf{r}}_{i'})$$

A. P. Bartók and G. Csányi, *Int. J. Quantum Chem.*, **115**, 1051 (2015)
A. P. Bartók, R. Kondor, and G. Csányi, *Phys. Rev. B*, **87**, 184115 (2013)

Descriptor: Smooth Overlap of Atomic Positions

$$\begin{aligned} k(\rho, \rho') &= \int d\hat{R} S^*(\hat{R})S(\hat{R}) \\ &= \sum_{l,m,m'\lambda,\mu,\mu'} (I_{mm'}^l)^* I_{\mu\mu'}^\lambda \int d\hat{R} D^*(\hat{R})_{mm'}^l D(\hat{R})_{\mu\mu'}^\lambda \\ &= \sum_{l,m,m'} (I_{mm'}^l)^* I_{mm'}^l, \end{aligned}$$

↗

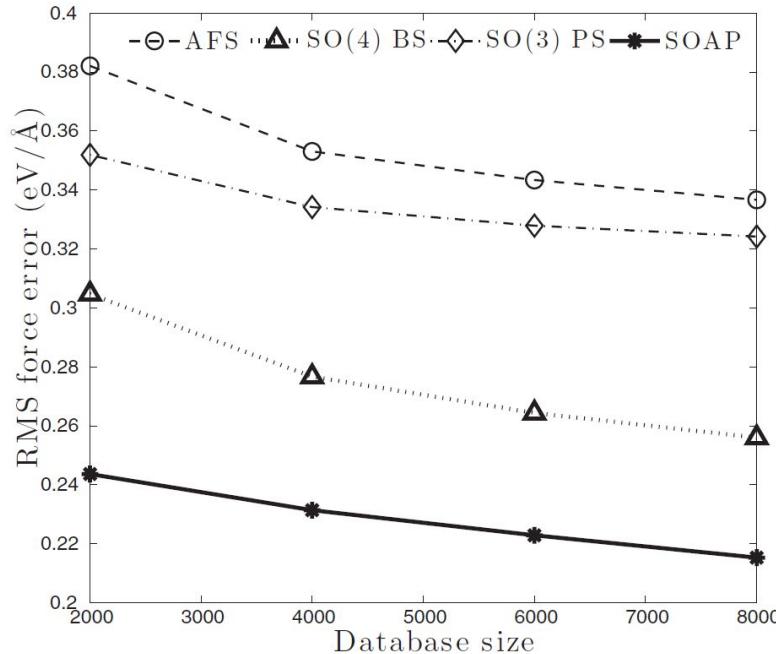
By Peter-Weyl theorem, coefficients of Wigner matrices form orthonormal basis in $L^2(SO(3))$

SOAP: $K(\rho, \rho') = \left(\frac{k(\rho, \rho')}{\sqrt{k(\rho, \rho)k(\rho', \rho')}} \right)^\zeta$

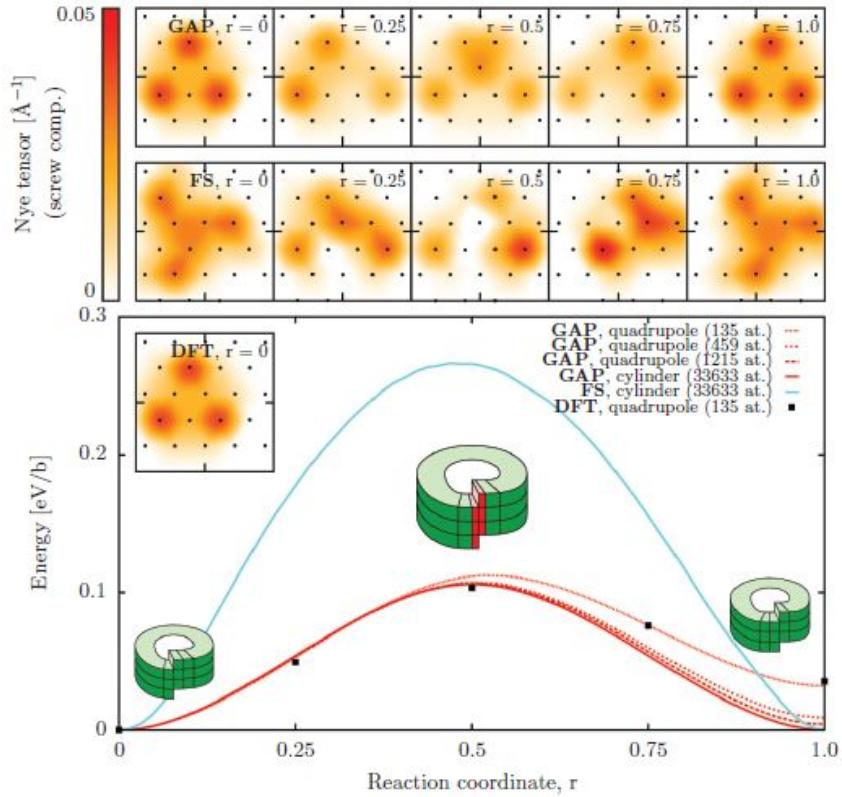
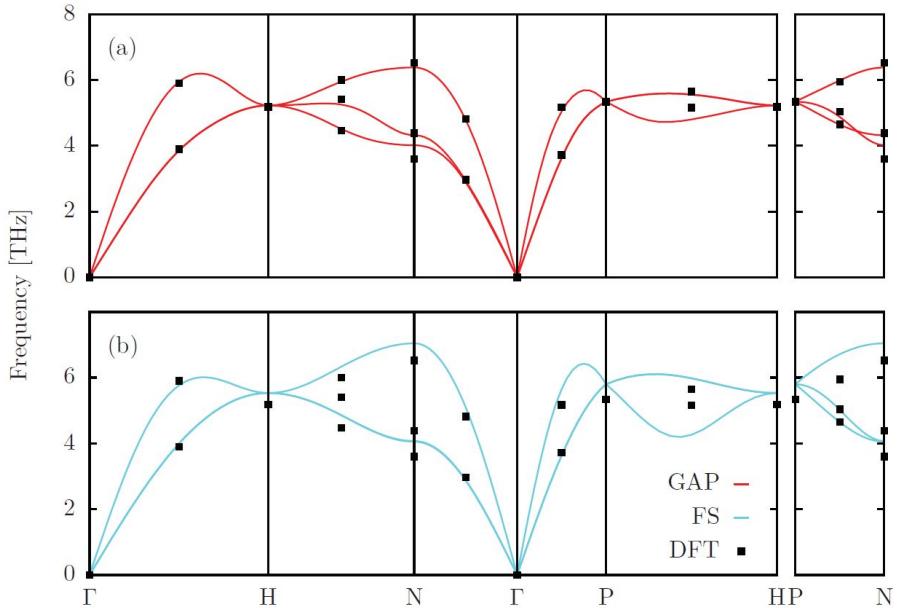
A. P. Bartók and G. Csányi, *Int. J. Quantum Chem.*, **115**, 1051 (2015)

A. P. Bartók, R. Kondor, and G. Csányi, *Phys. Rev. B*, **87**, 184115 (2013)

$$I_{mm'}^l \equiv \sum_{i,i'} \tilde{I}_{mm'}^l(\alpha, r_i, r_{i'})$$



Example of GAP+SOAP



Conclusion

- Machine learning based potential energy functions are able to model condensed matter systems accurately
- Basic symmetries (permutational/translational/rotational) need to be satisfied.
- “Delta-ML” may perform better if “relatively good” reference potentials are available.
- Keeping one-to-one mapping from Cartesian coordinates to features is important
- Extrapolating is one big issue.