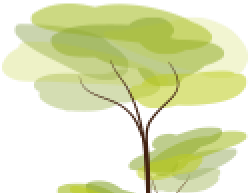
2018/4/9（Mon） 5:00 PM～

CSCI699: EXTREME-SCALE QUANTUM SIMULATIONS

Construction of Interatomic Potential based on Artificial Neural Network

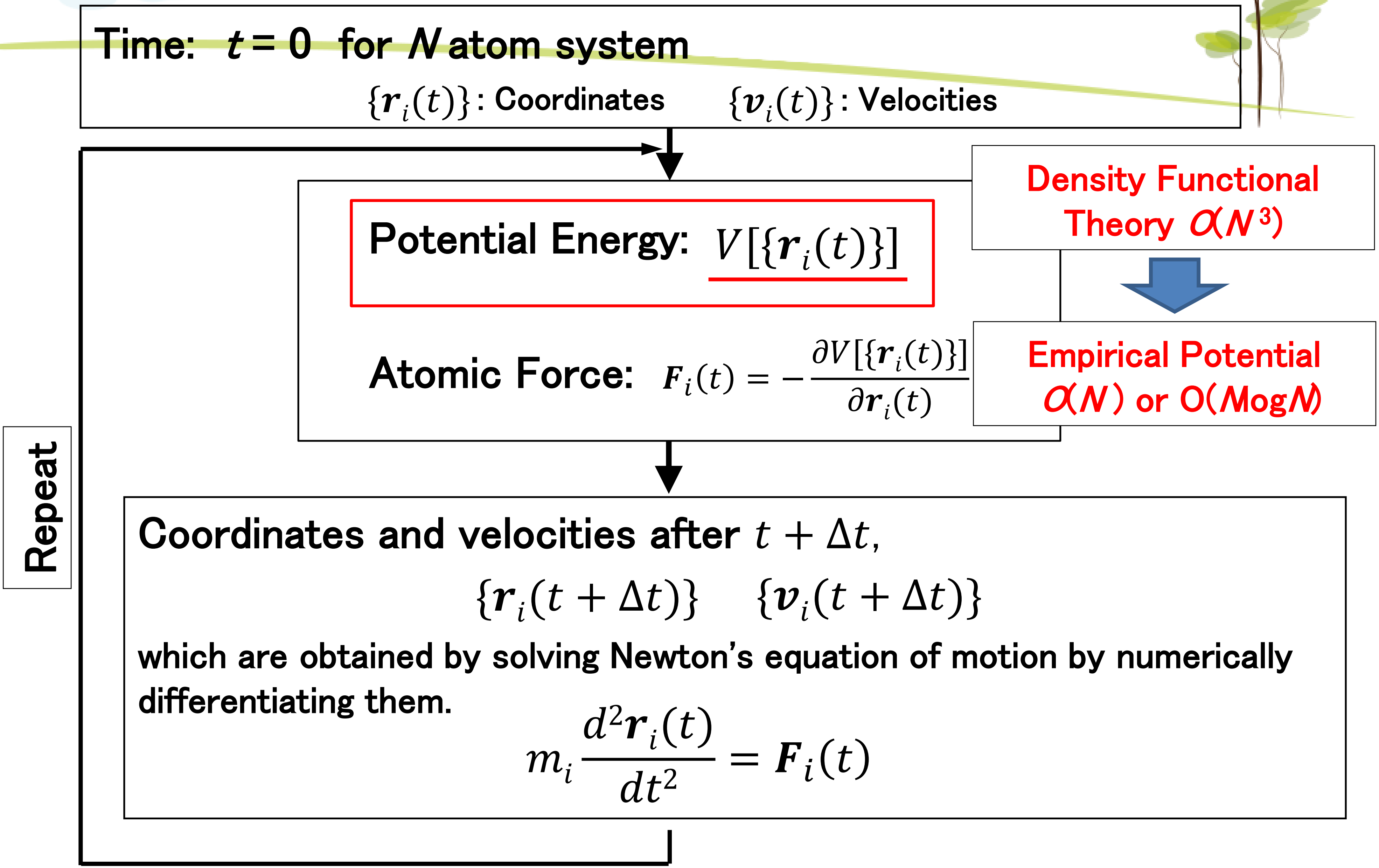
Graduate School of System Informatics
Kobe Univ. Kohei Shimamura

Outline



- **(*Ab Initio*) Molecular Dynamics**
  - **Empirical potential**
- **Artificial Neural Network potential: ANN potential**
  - **Feedforward Neural Network**
  - **Merits using ANN potential**
- **Atomic Energy NETwork (aenet) package**
  - **Atomic Fingerprint (AF: σ_{α})**
  - **How to calculate Force**
  - **How to make AF set**
- **Future works (on going)**
- **Summary**
- **How to run aenet**
  - **Aenet-KU2 on Github**

(*Ab Initio*) Molecular Dynamics



Time: $t = 0$ for N atom system

$\{\boldsymbol{r}_i(t)\}$: Coordinates $\{\boldsymbol{v}_i(t)\}$: Velocities

Potential Energy: $V[\{\boldsymbol{r}_i(t)\}]$

Atomic Force: $\boldsymbol{F}_i(t) = -\dfrac{\partial V[\{\boldsymbol{r}_i(t)\}]}{\partial \boldsymbol{r}_i(t)}$

Density Functional Theory $O(N^3)$

Empirical Potential $O(N)$ or O($M$og$N$)

Coordinates and velocities after $t + \Delta t$,

$$\{\boldsymbol{r}_i(t+\Delta t)\} \quad \{\boldsymbol{v}_i(t+\Delta t)\}$$

which are obtained by solving Newton's equation of motion by numerically differentiating them.

$$m_i \frac{d^2 \boldsymbol{r}_i(t)}{dt^2} = \boldsymbol{F}_i(t)$$

Repeat

【Empirical potential】

represents interatomic potential *V* with some kind of function.

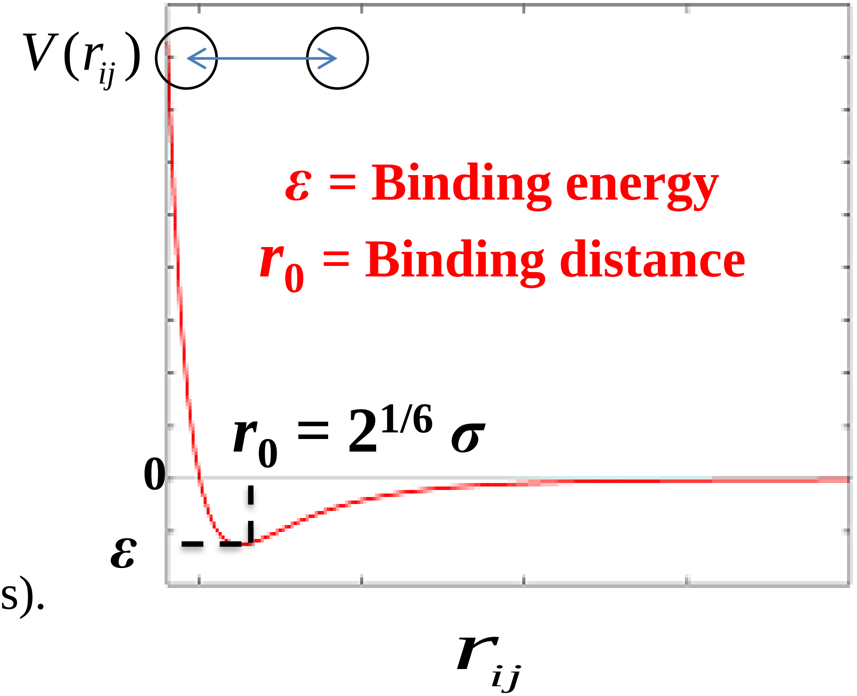
⟶ Significant reduction in calculation cost

e.g. Rare gas elements (He, Ne, Ar, Kr, Xe, Rn)

Only van der Waals interaction is working between atoms.

e.g. Lennard-Jones potential

$$V(r_{ij}) = 4\varepsilon\left[\underbrace{\left(\frac{\sigma}{r_{ij}}\right)^{12}}_{\text{Repulsive term}} - \underbrace{\left(\frac{\sigma}{r_{ij}}\right)^{6}}_{\text{Attractive term}}\right]$$



ε = **Binding energy**

r_0 = **Binding distance**

$r_0 = 2^{1/6}\,\sigma$

Fit the two parameters of σ and ε on the basis of first principles calculation（or experimental values).

【Empirical potential】

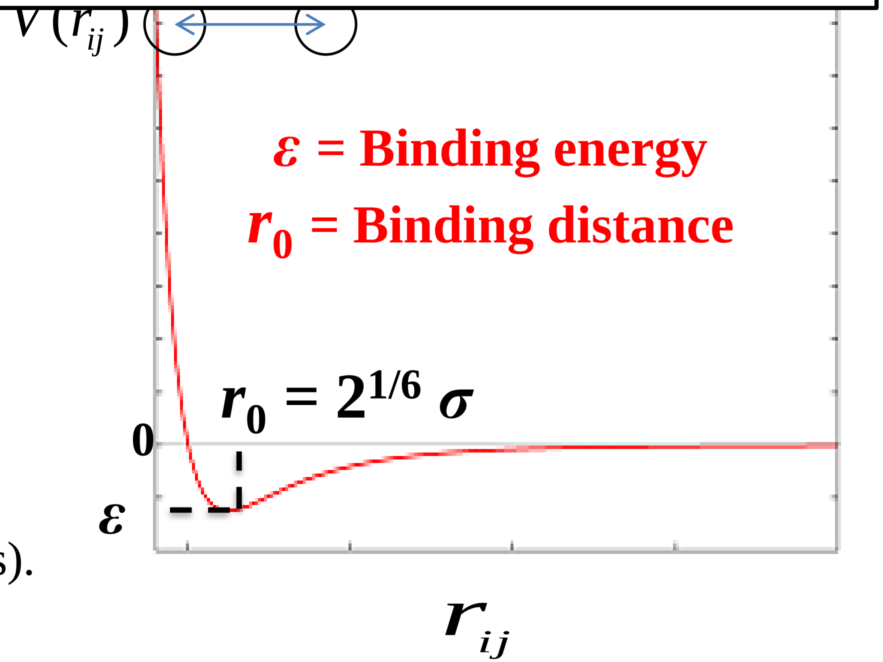
For repulsive term, Lennard-Jones potential: $\left(\frac{\sigma}{r_{ij}}\right)^{12}$

Buckingham potential: $A\exp(-Br_{ij})$

Although the functional form was fixed, it must be changed according to the situation.

e.g.

$$\underline{V(r_{ij})} = 4\varepsilon\left[\underbrace{\left(\frac{\sigma}{r_{ij}}\right)^{12}}_{\text{Repulsive term}} - \underbrace{\left(\frac{\sigma}{r_{ij}}\right)^{6}}_{\text{Attractive term}}\right]$$



Fit the two parameters of σ and ε on the basis of first principles calculation（or experimental values).

Artificial Neural Network potential: ANN potential

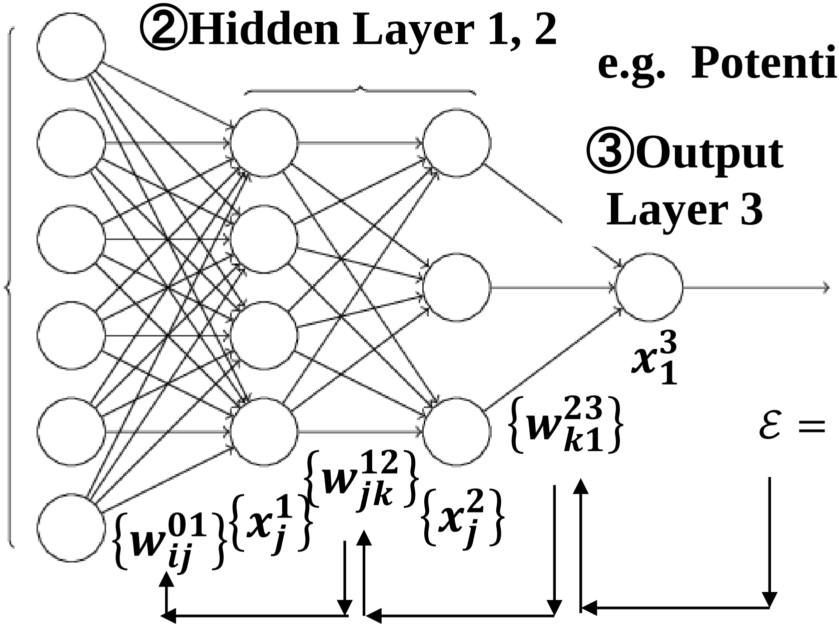
It is "artificial" because it modeled the neural tissue (neural network) of living organisms.

Feedforward Neural Network

Information transmission direction →

○**Node (or Perceptron)**

$\{w_{ij}^{lm}\}$: weights



②Hidden Layer 1, 2

e.g. Potential energy $V^{\text{Predict}}(\{x_i^0\})$

①Input Layer 0

$\{x_i^0\}$

Atomic coordinates

③Output Layer 3

x_1^3

$\{w_{ij}^{01}\}$ $\{x_j^1\}$ $\{w_{jk}^{12}\}$ $\{x_j^2\}$ $\{w_{k1}^{23}\}$

④Cost function $\mathcal{E}$

e.g. Sum of square errors

$$\mathcal{E} = \frac{1}{2}\sum_{I}^{\text{Training Atomic Structures}} \left(V_I^{\text{Predict}} - \underline{V_I^{\text{Reference}}}\right)^2$$

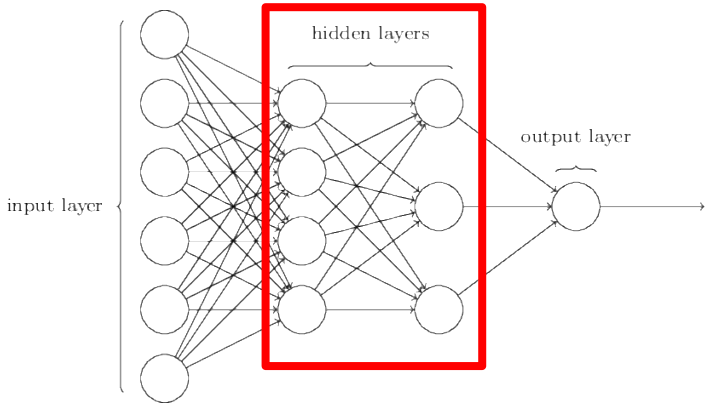
Calculated by *Ab initio* MD

⑤Using back-propagation method, we can calculate weight gradients $\left\{\frac{\partial\mathcal{E}}{\partial w_\alpha}\right\}$.

⑥Update $\{w_{ij}^{lm}\}$ using $\left\{\frac{\partial\mathcal{E}}{\partial w_\alpha}\right\}$.

⑦ Repeat ①-⑥ until the cost function $\mathcal{E}$ is minimized.

Merits using ANN potential



① **Universal approximation theorem**

ANN can approximate arbitrary functions if it has at least one hidden layer.

It is not necessary to limit the functional form like the LJ potential.

⇒ High accuracy

② **Dimensional compression of information**

By increasing the number of hidden layers, it is possible to remarkably reduce the information having the exponential dimension to the polynomial dimension like the tensor network.

⇒ Low calculation cost

Acceleration

Ab Initio → ANN

For MD simulation, it is possible to calculate the physical quantity that requires a long time to converge (e.g. Free energy) with the precision of *Ab initio* calculation.

Combined with LDC-DFT, MD simulation for protein (consists of over 10,000 atoms) with *ab initio* accuracy is possible.

N. Artrith, *et al.*, *Comp. Mat. Sci.* **114**, 135 (2016).

Atomic Energy NETwork (aenet) package

Interatomic potential generation software using ANN provided by Nongnuch Artrith et al. (UC Berkley) (GNU GPL, fortran95/2003 Flat MPI)

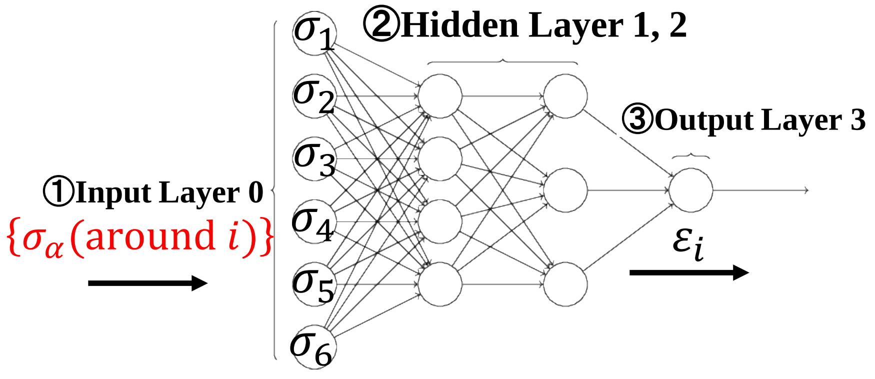
N. Artrith, *et al.*, *Comp. Mat. Sci.* **114**, 135 (2016).

(i) Here, the interatomic potential energy V of the N atom system is defined as the sum of the energy contribution $\{\varepsilon_i\}$ from each atom.

$$V = \sum_i^N \varepsilon_i(\{\sigma_\alpha\})$$

(ii) Use Atomic Finger Prints (AF: $\{\sigma_\alpha\}$) to represent $\{\varepsilon_i\}$.

→ Numerical representations of the characteristic local structure around each atom



***N* atoms**

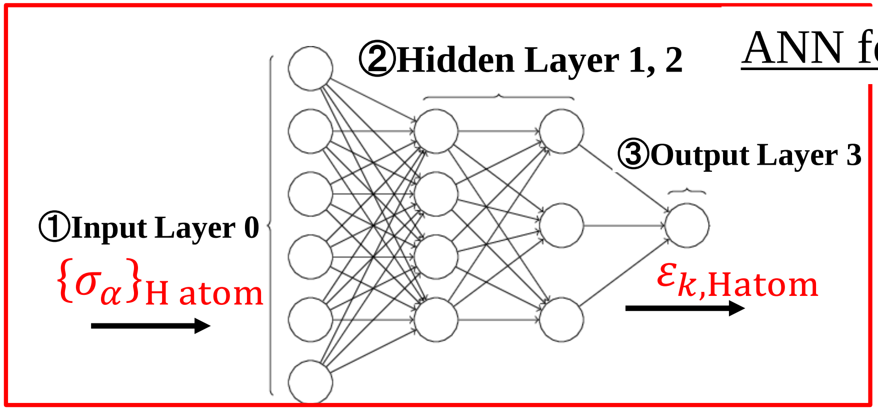
$$V = \sum_i^N \varepsilon_i(\{\sigma_\alpha\})$$

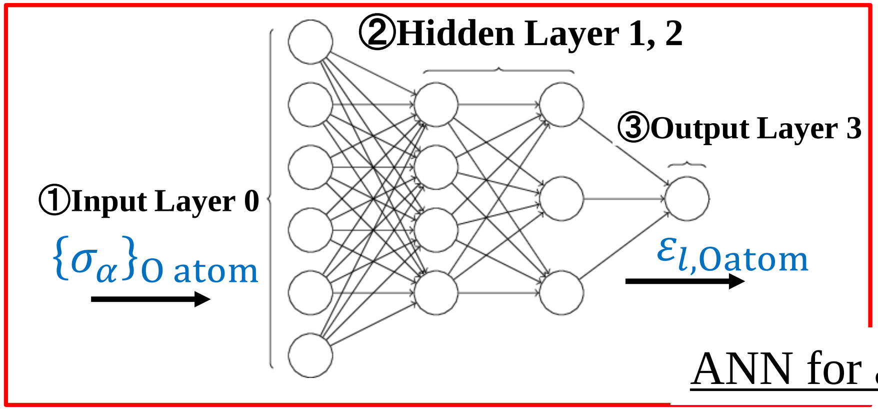
Example: H_2O

N. Artrith, *et al.*, *Comp. Mat. Sci.* **114**, 135 (2016).

Define AF set and make ANN <span>for each element</span>.

ANN for a H atom


②Hidden Layer 1, 2
③Output Layer 3
①Input Layer 0
$\{\sigma_\alpha\}_{\text{H atom}}$
$\varepsilon_{k,\text{Hatom}}$


②Hidden Layer 1, 2
③Output Layer 3
①Input Layer 0
$\{\sigma_\alpha\}_{\text{O atom}}$
$\varepsilon_{l,\text{Oatom}}$

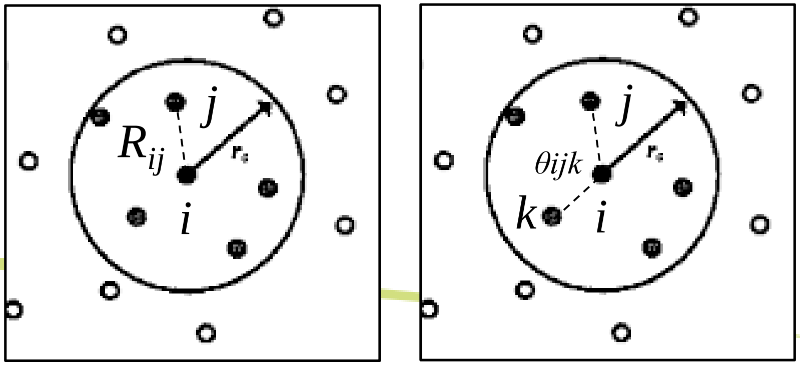
ANN for a O atom

N atoms

$\boldsymbol{N = 3n}$

$$V = \sum_{k}^{2n} \varepsilon_{k,\text{Hatom}}(\{\sigma_\alpha\}) + \sum_{l}^{n} \varepsilon_{l,\text{Oatom}}(\{\sigma_\alpha\})$$

Atomic Fingerprint (AF: σ_i)



▪Behler–Parrinello (BP) method

With the ith atom as the center, define functions using distances $\{R_{ij}\}$ and angles $\{\theta_{ijk}\}$ for atoms within the cutoff distance R_c.

【①Radial AF $\{\sigma_\alpha\} = \{G^r_{\alpha=\{i,k\}}\}$】 Adjustable parameters: $\{\eta_k\}, \{R^s_k\}, R_c$

$$G^r_{i,k} = \sum_{j\neq i}^{\text{neighbors}} \underbrace{\exp\left\{-\eta_k\left(R_{ij}-R^s_k\right)^2\right\}}_{\text{Gaussian function}} \cdot \underbrace{f_c\left(R_{ij}, R_c\right)}_{\text{Cutoff function}}$$

【② Angular AF $\{\sigma_\alpha\} = \{G^a_{\alpha=\{i,k\}}\}$ 】Adjustable parameters: $\{\eta_k\}, \{\zeta_k\}, \{\lambda_k\}, R_c$

$$G^a_{i,k} = 2^{1-\zeta_k}\sum_{k\neq j\neq i}^{\text{neighbors}}\left(1+\lambda_k\cos\theta_{ijk}\right)^{\zeta_k}\exp\{-\eta_k\left(R_{ij}{}^2+R_{ik}{}^2+R_{jk}{}^2\right)\}\cdot f_c\left(R_{ij}, R_c\right)\cdot f_c\left(R_{ik}, R_c\right)\cdot f_c\left(R_{jk}, R_c\right)$$

By setting different parameters, it is possible to express various local structures.
⇒ $\{\sigma_j\}$ is a basis set for expressing the local structure.

Neighbor atoms: $R_{ij}, R_{ik} < R_c$ (= 6 ~ 8 Å)

N. Artrith, *et al.*, *Comp. Mat. Sci.* **114**, 135 (2016).

Example: Diatomic molecule

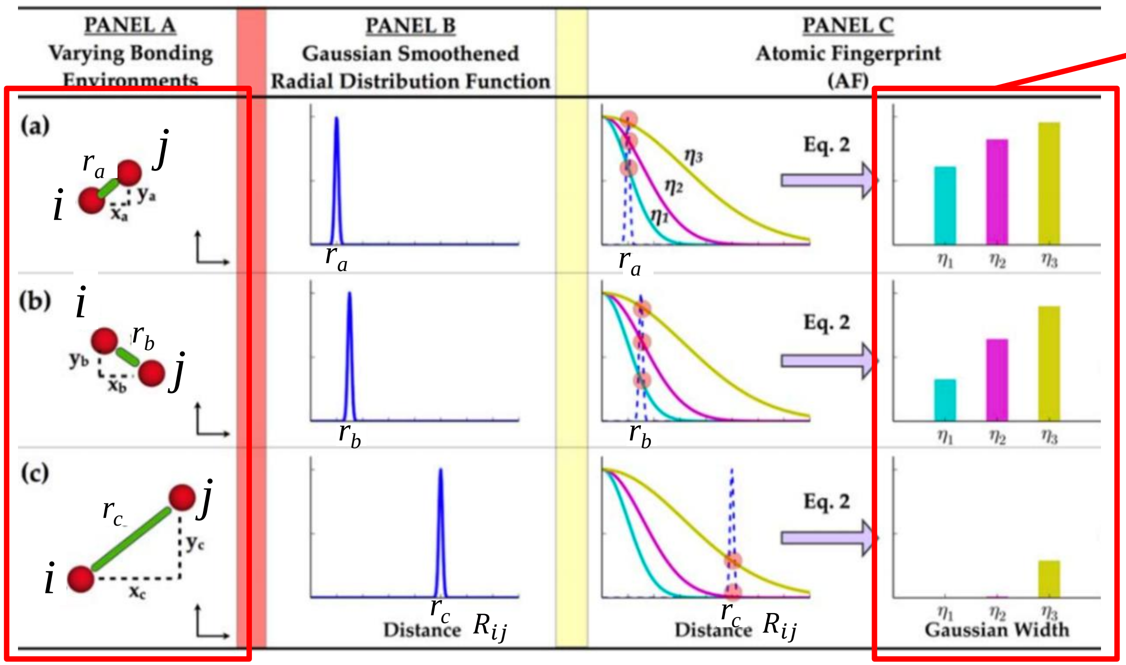
【①Radial AF $\{\sigma_\alpha\} = \{G_\alpha^r\}$】 Adjustable parameters: $\{\eta_k\}, \{R_k^s\}, R_c$

$$G_{i,k}^r = \sum_{j \neq i}^{\text{neighbors}} \underbrace{\exp\left\{-\eta_k \left(R_{ij} - R_k^s\right)^2\right\}}_{\text{Gaussian function}} \cdot f_c\left(R_{ij}, R_c\right)$$

Feature space of local structure

Here, define three Radial AFs with different parameters η.
$G_{i,1}^r(\eta_1), G_{i,2}^r(\eta_2), G_{i,3}^r(\eta_3)$

The shape of the Gaussian function can be changed.



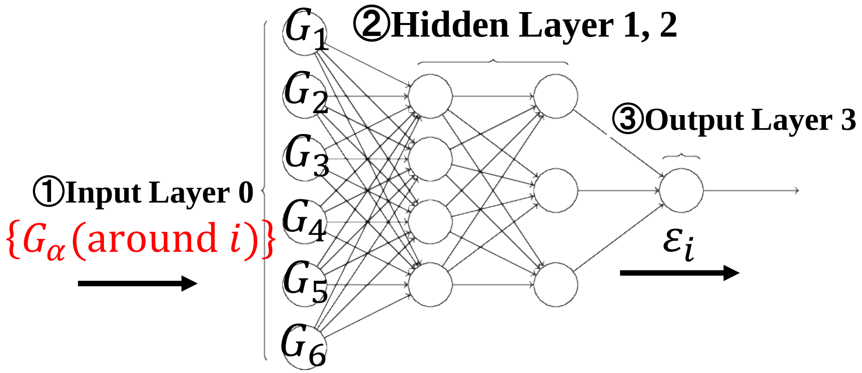
V. Botu, *et al.*, *Int. J. Quan. Chem.* **115**, 1074 (2015).

Structures are very similar!

It differs greatly from (a) and (b).

Difference in local structure is discriminated based on AF values with different parameters.

How to calculate Force



***N* atoms**

$$V = \sum_{i}^{N} \varepsilon_i(\{G_\alpha\})$$

$$\underline{F_{k,\beta}} = -\frac{\partial V}{\partial R_{k,\beta}} = -\sum_{i}^{N} \frac{\partial \varepsilon_i\,(\{G_\alpha\})}{\partial R_{k,\beta}}$$

$$= -\sum_{i}^{N} \underline{\sum_{\alpha}^{M_i} \frac{\partial \varepsilon_i\,(\{G_\alpha\})}{\partial G_\alpha} \frac{\partial G_\alpha}{\partial R_{k,\beta}}}$$

$F_{k,\beta}$ (kth atom$\beta \in (x, y, z)$)

M_i is the number of AFs defined for *i* th atom.

If the calculation cost becomes enlarged for the large number of AFs,
it is necessary to construct a minimum AF set.

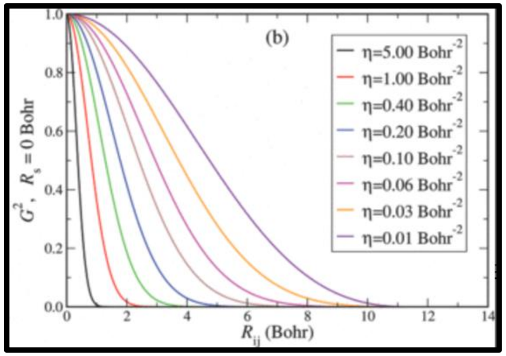
J. Behler, J. Phys.: Condens. Matter, **26**, 183001 (2014).

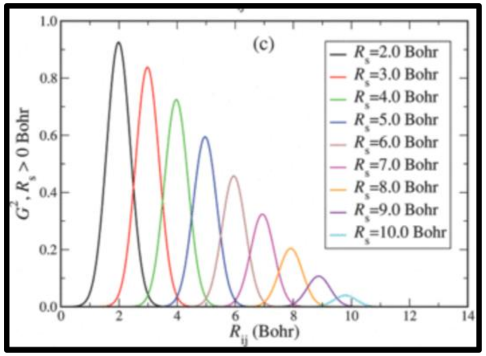
How to make AF set

①Prepare AFs at equal intervals and highly dense

【Radial AF $\{\sigma_\alpha\} = \{G^r_{\alpha=\{i,k\}}\}$】 Adjustable parameters: $\{\eta_k\}, \{R^s_k\}, R_c$

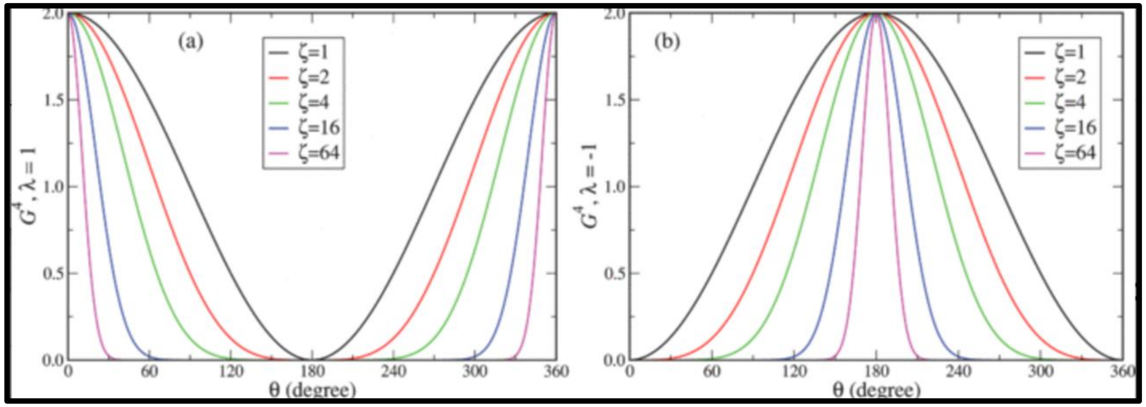
$$G^r_{i,k} = \sum_{j\neq i}^{\text{neighbors}} \exp\left\{-\eta_k\left(R_{ij} - R^s_k\right)^2\right\} \cdot f_c\left(R_{ij}, R_c\right)$$





【Angular AF $\{\sigma_\alpha\} = \{G^a_{\alpha=\{i,k\}}\}$】 Adjustable parameters: $\{\eta_k\}, \{\zeta_k\}, \{\lambda_k\}, R_c$

$$G^a_{i,k} = 2^{1-\zeta_k} \sum_{k\neq j\neq i}^{\text{neighbors}} \left(1 + \lambda_k \cos\theta_{ijk}\right)^{\zeta_k} \exp\left\{-\eta_k\left(R_{ij}{}^2 + R_{ik}{}^2 + R_{jk}{}^2\right)\right\} \cdot f_c\left(R_{ij}, R_c\right) \cdot f_c\left(R_{ik}, R_c\right) \cdot f_c\left(R_{jk}, R_c\right)$$



② Reduce the number of AFs by using correlation coefficients between AFs.

J. Behler, J. Phys.: Condens. Matter, **26**, 183001 (2014).

(i) Check the correlation coefficients r_{xy} between AFs for **Training Set**.

(ii) Delete one if r_{xy} value between two arbitrary AFs is too high large (> 0.9).

Correlation coefficients r_{xy}

$-1 < r_{xy} < 1$

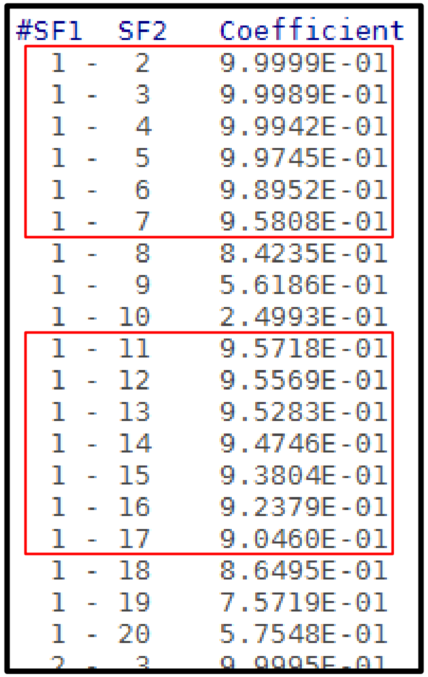
$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\left(\left(\sum_{i=1}^{n}(x_i - \bar{x})^2\right)\left(\sum_{i=1}^{n}(y_i - \bar{y})^2\right)\right)^{1/2}}$$

$0 < r_{xv} < 1$: Positive correlation

$r_{xv} = 0$: No correlation

$-1 < r_{xy} < 0$: Negative correlation

Example



```
#SF1  SF2    Coefficient
  1 -  2     9.9999E-01
  1 -  3     9.9989E-01
  1 -  4     9.9942E-01
  1 -  5     9.9745E-01
  1 -  6     9.8952E-01
  1 -  7     9.5808E-01
  1 -  8     8.4235E-01
  1 -  9     5.6186E-01
  1 - 10     2.4993E-01
  1 - 11     9.5718E-01
  1 - 12     9.5569E-01
  1 - 13     9.5283E-01
  1 - 14     9.4746E-01
  1 - 15     9.3804E-01
  1 - 16     9.2379E-01
  1 - 17     9.0460E-01
  1 - 18     8.6495E-01
  1 - 19     7.5719E-01
  1 - 20     5.7548E-01
  2 -  3     9.9995E-01
```

$G_2 \sim G_7,\ G_{11} \sim G_{17}$

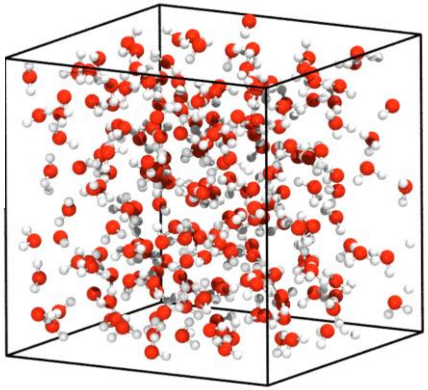
Delete

It is possible to reduce redundant AFs at once.

H_2O AF set

$\{\sigma_\alpha\}_{\mathrm{H}\ atom}$: **20 AFs ($R_c$ = 3.00 Å)**
$\{\sigma_\alpha\}_{\mathrm{O}\ atom}$: **92 AFs ($R_c$ = 6.00 Å)**

216 H_2O system
(648 atoms)
T = 300 K, P = 1 atm



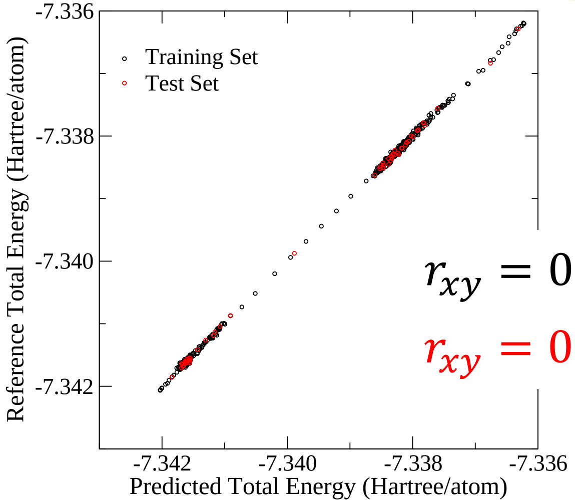
【10^4 *acceleration*】
Tokyo Univ. Sekirei 64 cores (3 ps)
ANN-MD: 233 seconds
Ab initio-MD: 2,397,600 seconds (28 days)

【Physical values】

Accuracy of potential energy is very good for any materials.
(**Na, GaAs, CH_4, H_2O**).
⇒○ Specific heats, ○Radial distribution functions

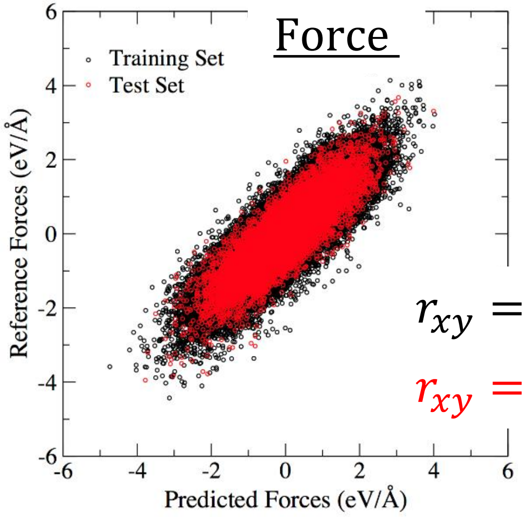
On the other hand, the accuracy of the force depends on the materials.
(**Na, GaAs / CH_4, H_2O**)
⇒×Thermal conductivity, ×Viscosity coefficient

Potential energy *V*



r_{xy} = 0.999 (train)
r_{xy} = 0.999 (test)

Force



r_{xy} = 0.711 (train)
r_{xy} = 0.711 (test)

Include force in the cost function (on going)

Arxiv:1707.09571

The cost function used in *Aenet* consists only of potential energy *V*.

Since the force can be obtained from $\varepsilon_i(\{G_\alpha\})$, we redefine the cost function as follows.

$$\boxed{F_{k,\beta} = -\frac{\partial V}{\partial R_{k,\beta}} = -\sum_i^N \frac{\partial \varepsilon_i(\{G_\alpha\})}{\partial R_{k,\beta}}}$$

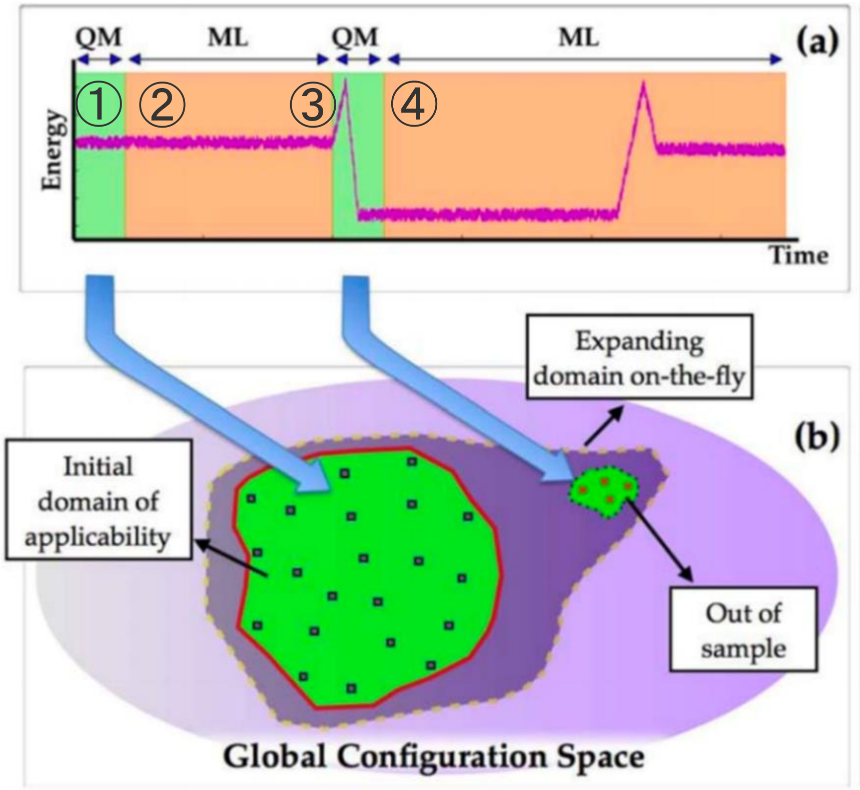
$$\mathcal{E} = \underbrace{\frac{1}{2}\sum_I^{M\text{:Training Structures}} \left(\sum_i^{N\text{:atoms}} \varepsilon_i(\{G_\alpha\}) - V_I^{\text{Reference}} \right)^2}_{\text{Potential Energy term}} + \underbrace{\frac{1}{2}\sum_I^{M\text{:Training Structures}} \sum_i^{N\text{:atoms}} \left(\vec{F}_{I,i}^{\text{Predict}} - \vec{F}_{I,i}^{\text{Reference}} \right)^2}_{\text{Force term}}$$

・We implemented this cost function on aenet, but accuracy of force improved by only a few percent (in the case of H_2O).
・Since the above two terms have different dimensions and different magnitudes of values, it is not easy to minimize both.

⇒ We are seeking better optimization methods to minimize both terms efficiently (discuss later).

Combine *Ab initio* with ANN potentials to speed up MD simulation (on going)

For example, many proteins cause chemical reactions over a long time of nanoseconds or more.
⇒ Such a chemical reaction can not be handled by the *ab initio* MD.


QM
ML
QM
ML
(a)
①
②
③
④
Energy
Time
Expanding domain on-the-fly
(b)
Initial domain of applicability
Out of sample
Global Configuration Space

V. Botu, *et al.*, *Int. J. Quan. Chem.* **115**, 1074 (2015).

① Proceed with learning ANN
while performing *Ab initio* MD simulation.

② Switch from *ab initio* to ANN
when energy and force accuracies of ANN are assured.

③ Switch from ANN to *ab initio*
when reaching a new atomic configuration which has not been learned.

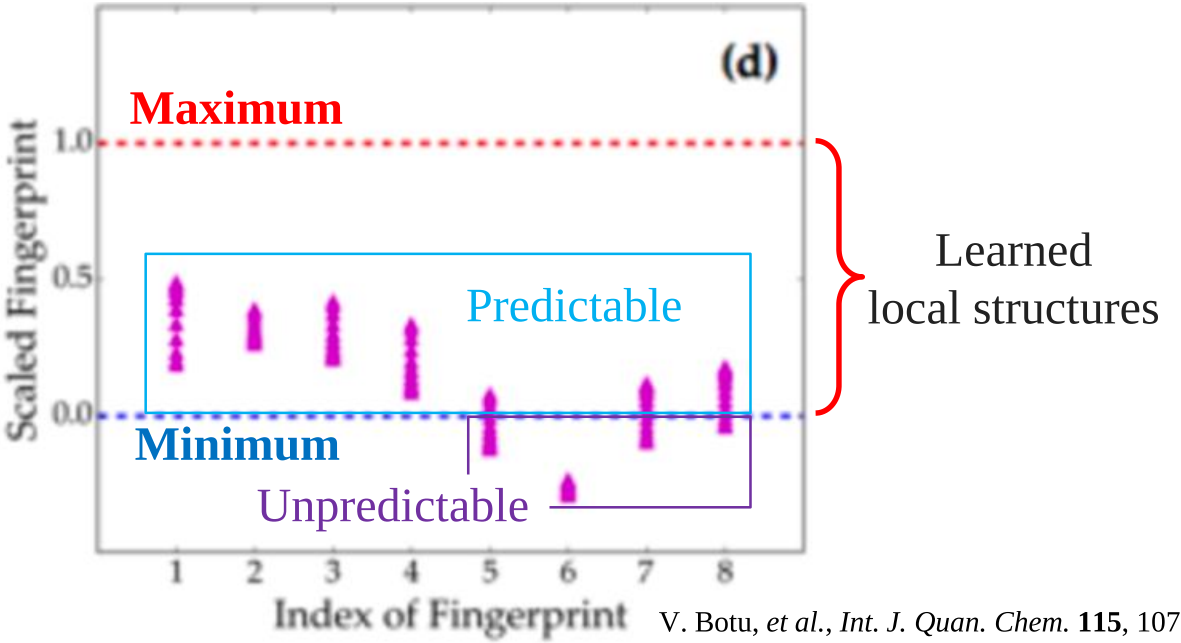
④ Switch from *ab initio* to ANN again
when energy and force accuracies of ANN are assured.

How to determine new atomic configuration using AFs

AFs represents the features of the local structure.

Therefore, the interval between the minimum and maximum values of AF with respect to the training data corresponds to the learned local structures.

In the case of 8 AFs



Maximum

Predictable

Minimum

Unpredictable

Learned local structures

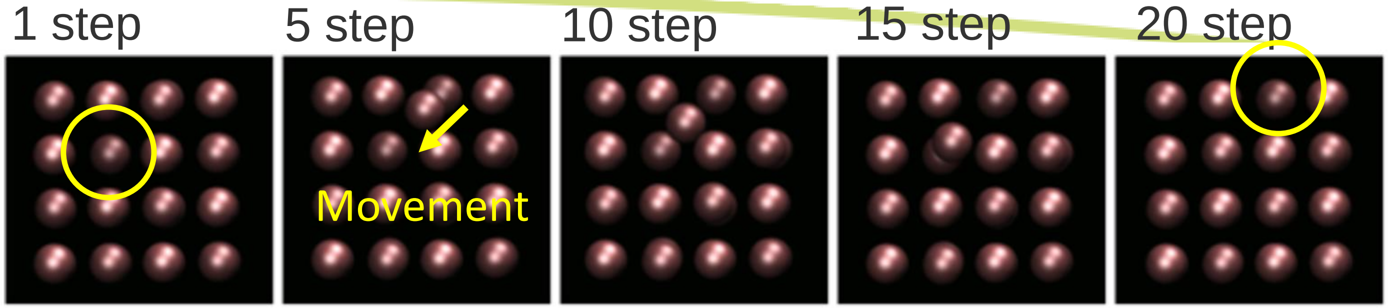
V. Botu, *et al.*, *Int. J. Quan. Chem.* **115**, 1074 (2015).

When a value of AF deviating from the interval appears, it is judged that a new atomic structure is appeared.

V. Botu, *et al.*, *Int. J. Quan. Chem.* **115**, 1074 (2015).

How to determine new configuration using AFs

A defect in Al bulk



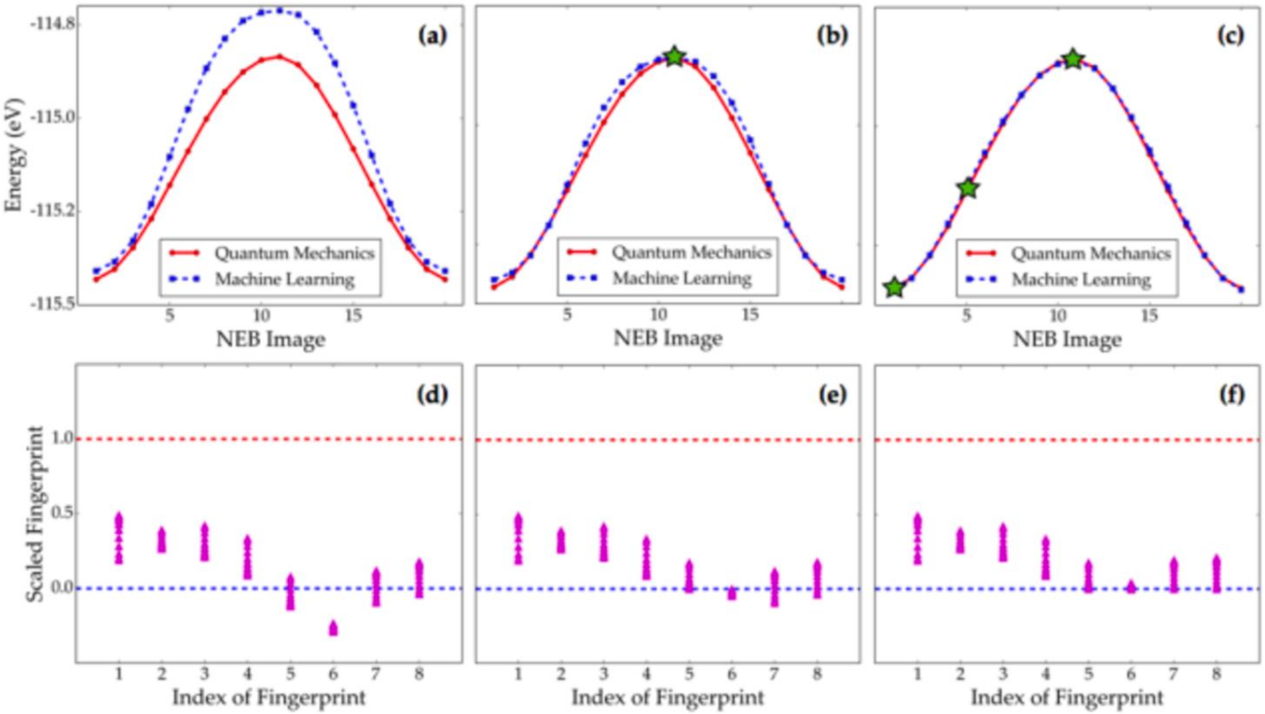
A defect in Al bulk

①Save the maximum and minimum values of each AFs in Training Set.

②If a value of AF exceeds or falls, switch from ANN-MD to AIMD.

③ Learn the unknown structure with AIMD and update ANN potential.

④Restart ANN-MD.



To python with tensorflow (on going)

It is difficult for the present aenet package to implement a framework that switches between AIMD and ANN - MD.

【Problems】

▪Online learning is necessary.
Whenever a new atomic structure appears, the ANN potential needs to be updated.
However, the online learning method (steepest descent method) implemented in present aenet has poor convergence.

▪Transfer learning is required.
If you need to increase AF, the number of weights of ANN will be increased. By freezing the original weights and learning only new weights, we can minimize the training time.

We are rewriting aenet (fortran) to python code using tensorflow library which can use the latest online learning methods, transfer learning methods, cost functions, etc.

Summary

【Merits using ANN potential】

Based on the universal approximation theorem, Artificial Neural Network (ANN) can approximate arbitrary functions. Therefore, ANN can imitate the complicated *ab initio* interatomic potential *V*.

With the ANN potential, we can considerably accelerate MD simulation with *ab initio* accuracy (at least for the states which are included in the training set).

【Atomic Energy NETwork (aenet) package】

This is an interatomic potential generation software using ANN provided by Nongnuch Artrith *et al.* at UC Berkley. (GNU GPL, fortran95/2003 Flat MPI).

Atomic Fingerprints (AFs) proposed by J. Behler indicates numerical representations of characteristic local atomic structures. Using the AFs as basis set, we can construct ANN potential.

Using the AFs, we can also determine whether a focused atomic structure data has been learned or not. Thus, for example, it is possible to speed up MD simulation by switching of ANN and *Ab initio* potentials.

Since a (rough) AF create method has been provided by J. Behler, basically, you can construct ANN potential for any materials you want.

How to run aenet

N. Artrith, *et al.*, *Comp. Mat. Sci.* **114**, 135 (2016).

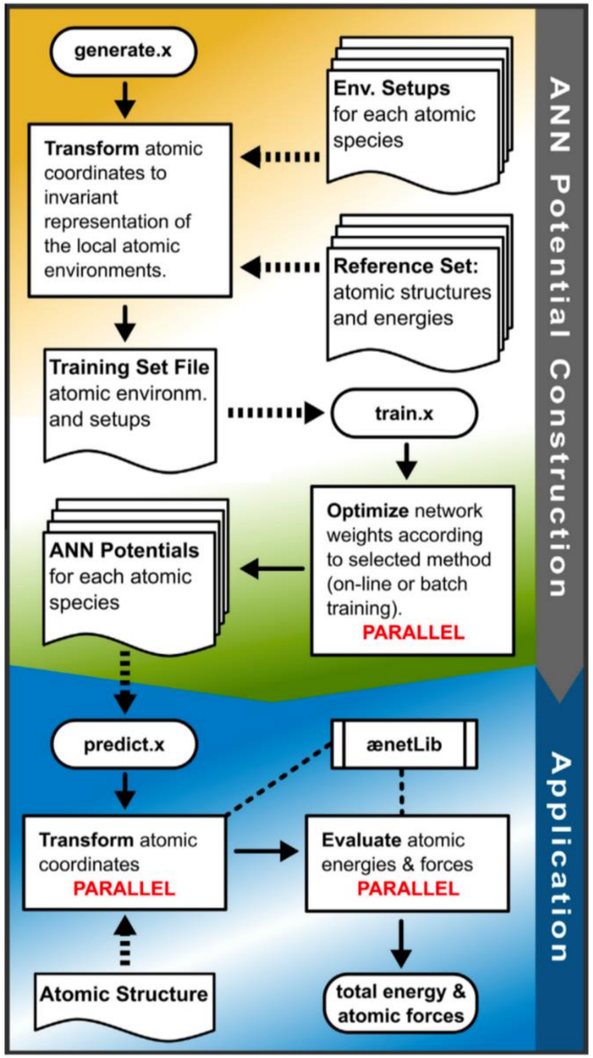
generate.x Create training data that train.x can read.

However, prepare before executing generate.x.
- Atomic Fingerprints set (Fingerprint.f90)
- Structure Files (results of QXMD and make_xsf.f90)

train.x Learn ANN Potentials (weights $\{w_{ij}^{lm}\}$)
- Setting of ANN
  - Num. of Hidden Layers
  - Num. of Nodes
- Select optimization method
  - Online learning method: Steepest descent method
  - Batch learning method: L- BFGS method
  - Levenberg-Marquardt method

predicts.x

Check the prediction accuracy of energy and force of ANN potential



generate.x

Env. Setups for each atomic species

Transform atomic coordinates to invariant representation of the local atomic environments.

Reference Set: atomic structures and energies

Training Set File atomic environm. and setups

train.x

Optimize network weights according to selected method (on-line or batch training).
PARALLEL

ANN Potentials for each atomic species

predict.x

ænetLib

Transform atomic coordinates
PARALLEL

Evaluate atomic energies & forces
PARALLEL

Atomic Structure

total energy & atomic forces

Aenet-KU2 on Github

①. Before you create ANN using aenet, you need to prepare the following.

(i) QXMD data as a reference data (specifically, qm_ion.d, qm_frc.d, qm_cel.d, md_eng.d)

(ii) Potential Energies of Isolated Atoms calculated by QXMD.
(If you want to create ANN for H_2O system,
you need to prepare Potential Energies of H and O atoms, respectively.)

②. Using make_xsf.f90 with QXMD data and the energies of isolated atoms,
create xsf files for each atomic configuration that generate.x can read.

③. Using Fingerprint.f90, create files which AF sets for each element are written.
(e.g. H.fingerprint.stp, O.fingerprint.stp)

④. Using generate.x with xsf files, fingerprint files, and generate.in,
create training data that train.x can read (e.g. H2O.train).
and files which correlation coefficients between AFs are written (correlation files).

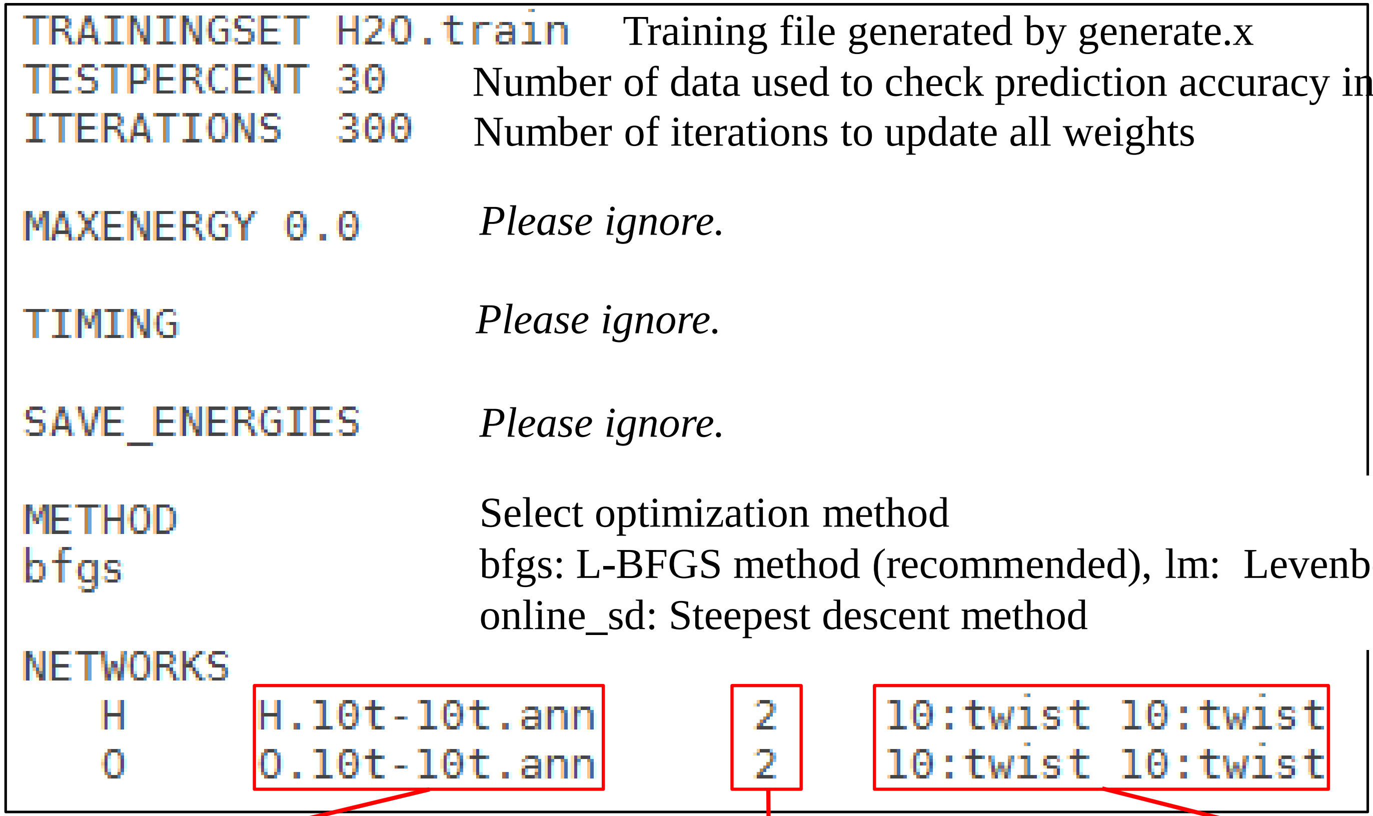
It is generated automatically by executing make_xsf.f90.

⑤. Using the correlation files, delete the redundant AFs from fingerprint files.

⑥. Using train.x with training data (e.g. H2O.train) and train.in,
train and create ANN potentials. (e.g. H.10t-10t.ann O.10t-10t.ann)

Explain later.

train.in (For example, H_2O)



```
TRAININGSET H2O.train
TESTPERCENT 30
ITERATIONS  300

MAXENERGY 0.0

TIMING

SAVE_ENERGIES

METHOD
bfgs

NETWORKS
    H     H.10t-10t.ann     2    10:twist 10:twist
    O     O.10t-10t.ann     2    10:twist 10:twist
```

Training file generated by generate.x

Number of data used to check prediction accuracy in H2O.train (%)

Number of iterations to update all weights

Please ignore.

Please ignore.

Please ignore.

Select optimization method
bfgs: L-BFGS method (recommended), lm: Levenberg-Marquardt method
online_sd: Steepest descent method

Output ANN potential file names

Num. of Hidden layers

Num. of Nodes and
types of activation functions

Predict.x

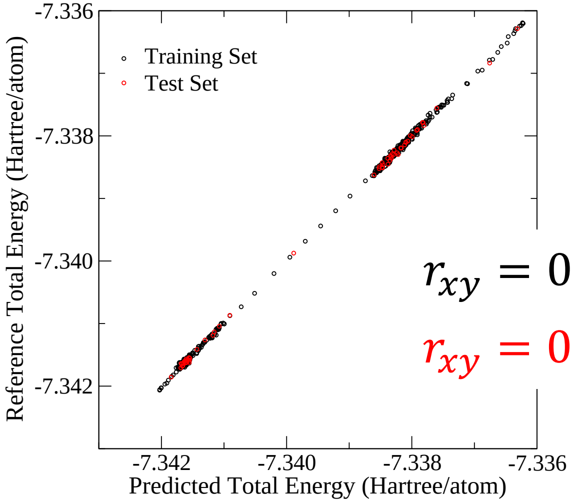
①. Before you perform predict.x, you need to prepare input file (predict.in).

Using predict.f90, you divide xsf files into those used for learning (train) and other (test), and create two predict.in that respective them are written (predict_train.in and predict_test.in)

②. Using predict.x with predict_train.in or predict_test.in, you can find out prediction accuracies for train and test data, respectively.

If you perform predict.x, energy.dat and force.dat are output.
In these files, predicted energies predicted forces as well as referenced energies with those correlation coefficients and forces are written.

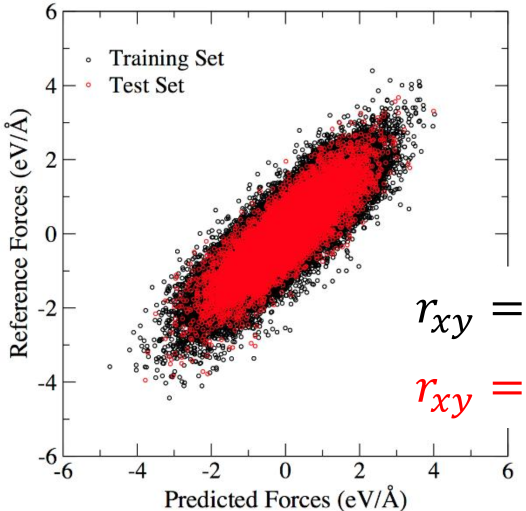
Potential energy *V*



r_{xy} = 0.999 (train)

r_{xy} = 0.999 (test)

Force



r_{xy} = 0.711 (train)

r_{xy} = 0.711 (test)

Aenet-KU2 on Github

Uploaded to Github,

(1) the ANN potential for liquid Na as an example.

(2) Readme: explaining procedure of creating ANN potential for liquid Na.

How to run MD using ANN potential

With ANN potential files, we can perform ANN based MD simulation using QXMD.

Uploaded to Github,

(1) the input file to perform QXMD with ANN potentials.

(2) Readme: how to perform using above input file.

If you have questions…

Please contact Shimamura (shimamura@port.kobe-u.ac.jp).