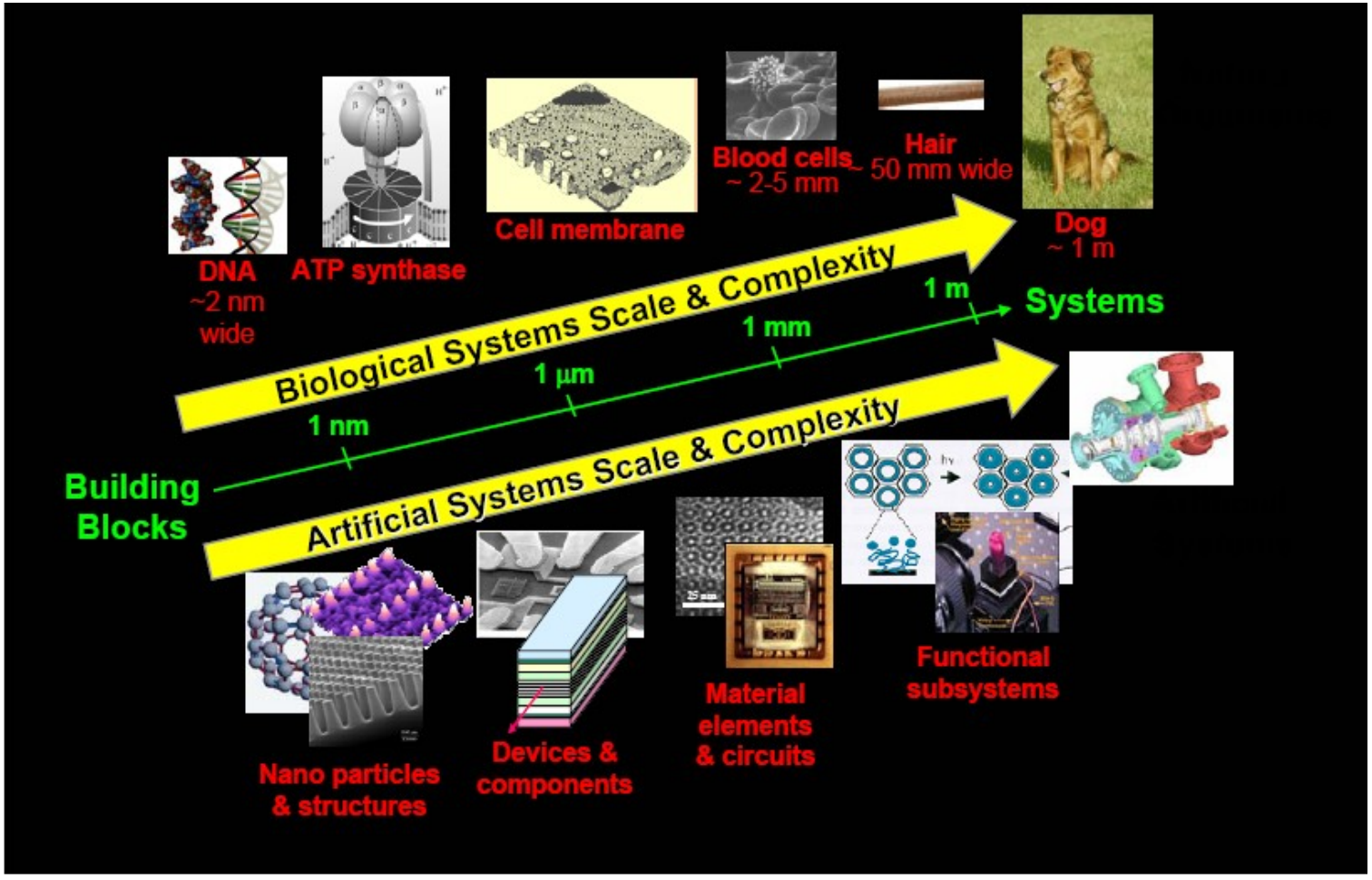
Machine Learning Interatomic Potentials: General concepts

Noel Jakse
Université Grenoble-Alpes, CNRS, Grenoble INP
SIMaP, F-38000 Grenoble, France.

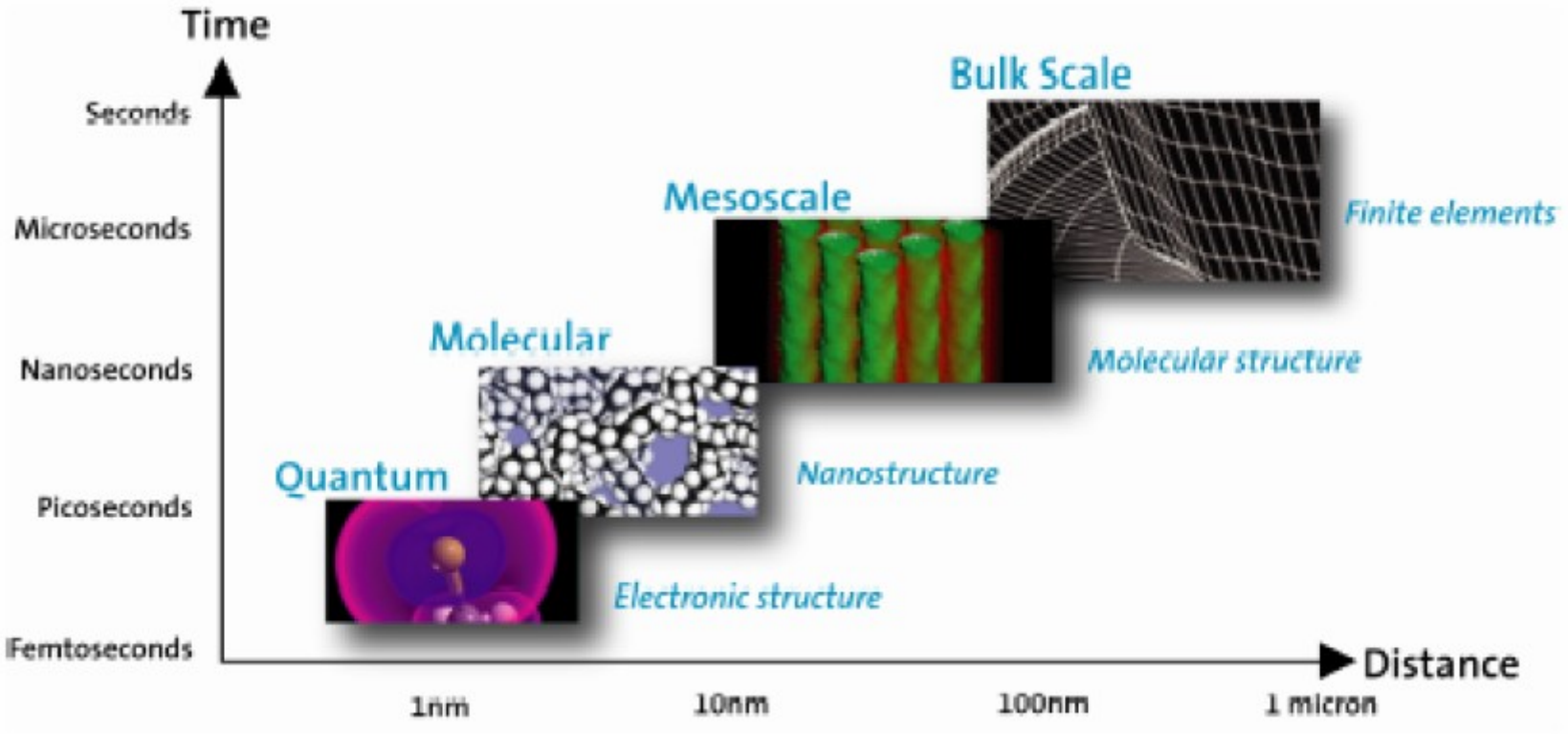
Outline

- Introduction
- Interatomic interactions and potential energy landscape
- Machine learning interatomic potentials (MLIP)
  - Supervised ML : regression tasks
  - representation of the data
  - building a dataset for training
  - Active learning
- Example: homogeneous nucleation
  - Binary alloys: Al-Ni
- Outlook

Matter and complexity

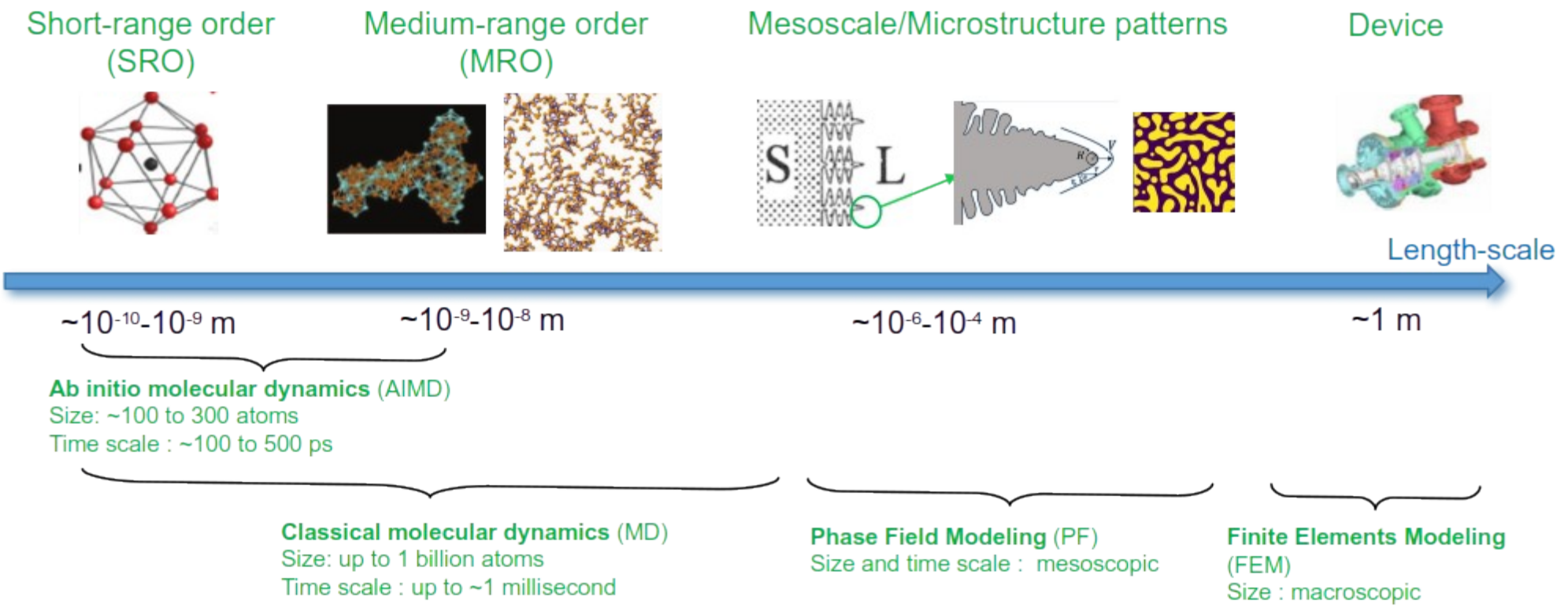


Scales and methods



Refs

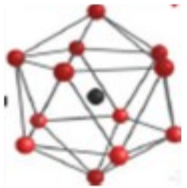
Scales and methods



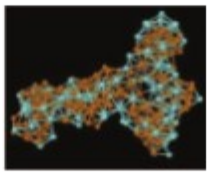
Refs

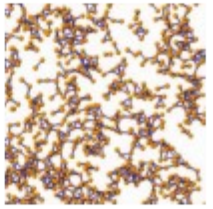
Scales and methods

Short-range order (SRO)

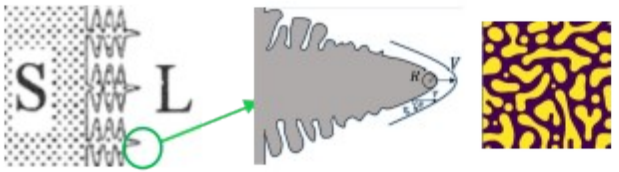


Medium-range order (MRO)





Mesoscale/Microstructure patterns



Device

Length-scale

~10^{-10}-10^{-9} m

~10^{-9}-10^{-8} m

~10^{-6}-10^{-4} m

~1 m

Ab initio molecul dynamics (AIMD)
Size: ~100 to 30
Time scale : ~100 t s

M L I P

Cla molecular dynamics (MD)
Size: up to 1 billion atoms
Time scale : up to ~1 millisecond

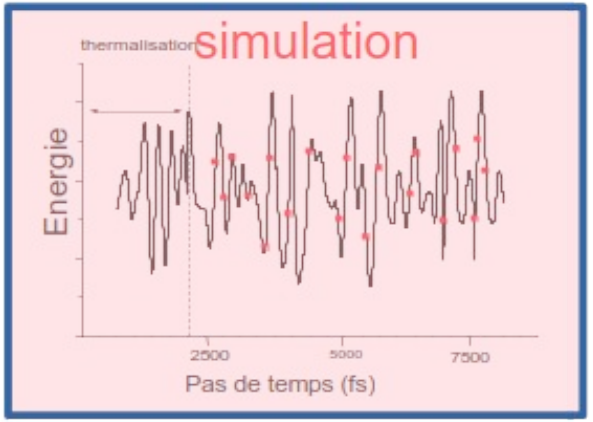
Phase Field Modeling (PF)
Size and time scale : mesoscopic

Finite Elements Modeling (FEM)
Size : macroscopic

Refs

Atomistic simulations

Quantum Physics
Electronic
Structure

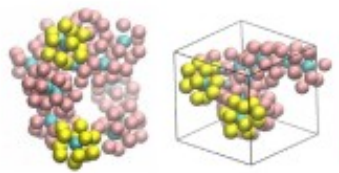


Properties of
Materials
and
Processing

Simplified Hamiltonian
Structure Matching
Force Matching
Machine Learning

Energetics :
E, H
Fluctuations :
C_P, C_V

Dynamics :
Diffusion
Vibrational spectrum
Etc..

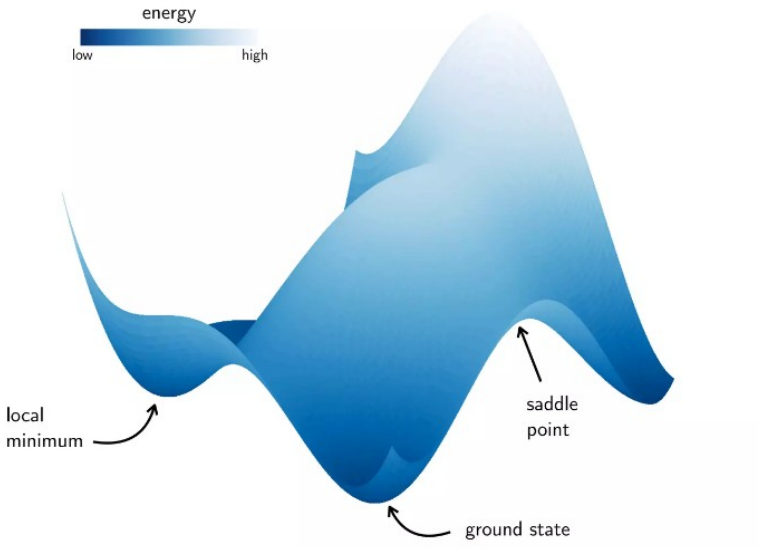


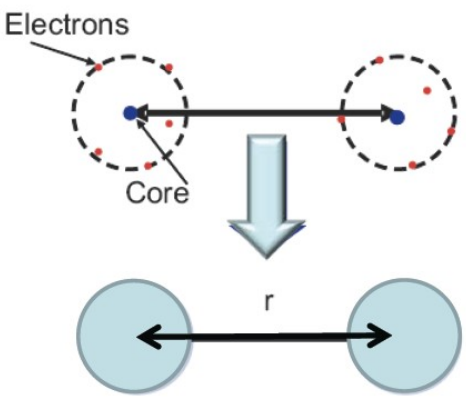
Atomic structure

Interatomic interactions

$$U_N(\mathbf{r}^N) = u_0(\rho) + \sum_{i_1} u_1(\mathbf{r}_{i_1}) + \frac{1}{2!}\sum_{i_1 \neq i_2} u_2(\mathbf{r}_{i_1}, \mathbf{r}_{i_2}) + \frac{1}{3!}\sum_{i_1 \neq i_2 \neq i_3} u_3(\mathbf{r}_{i_1 i_2}, \mathbf{r}_{i_1 i_3}, \mathbf{r}_{i_2 i_3}) + \cdots$$

Potential Energy Landscape (PEL)





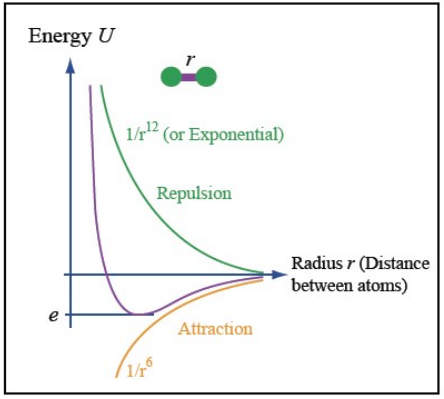


Image by MIT OpenCourseWare.

Attraction: Formation of chemical bond by sharing of electrons
Repulsion: Pauli exclusion (too many electrons in small volume)

D. Wales et al., Journal of Chemical Theory and Computation (2024)

Interatomic interactions

Put as much as you need for describing the physics of the system

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

Harmonic potential (Hook) :

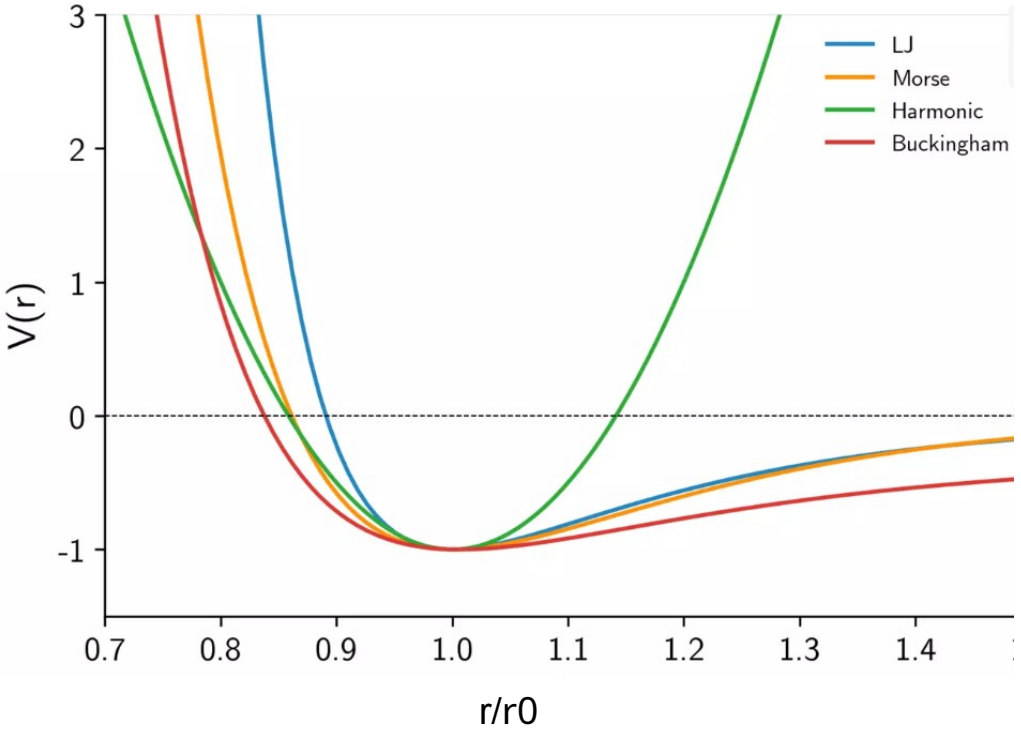
$$V(r) = A(r - r_{eq})^2$$

Morse :

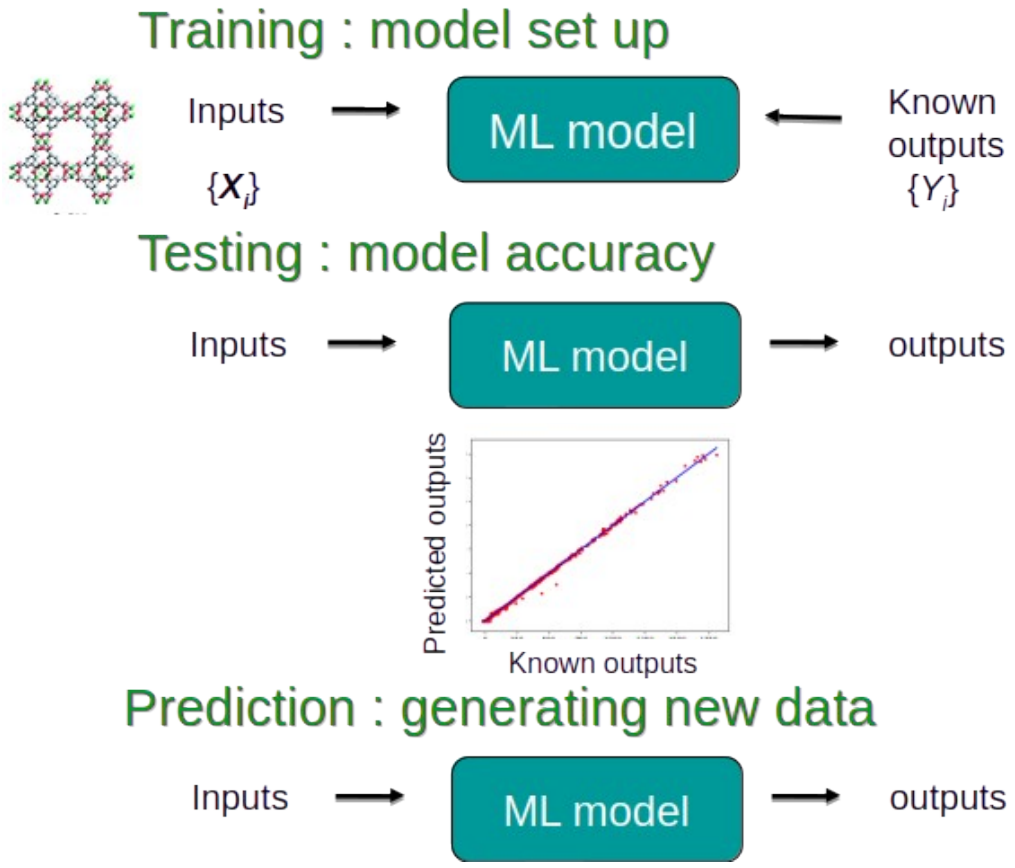
$$V(r) = D\exp\left[-2\alpha(r - r_o\right] - 2D\exp\left[-\alpha(r - r_o\right]$$

Buckingham :

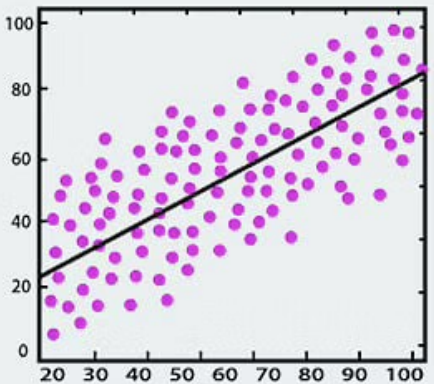
$$V(r) = A\exp\left[-\frac{r}{\rho}\right] - \frac{C}{r^6} + \frac{q_1 q_2}{r}$$



Building a ML potential : Supervised learning

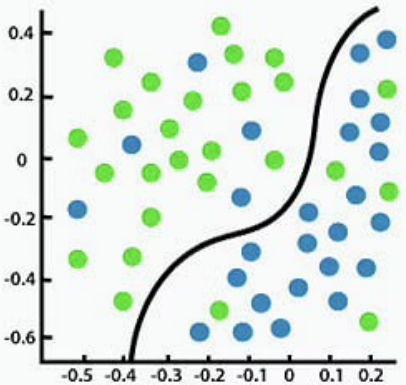


Building a ML potential : regression task



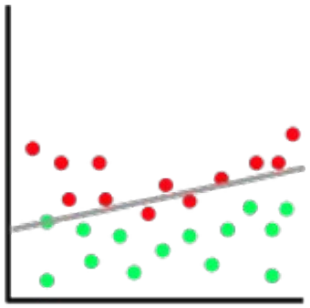
Regression

versus

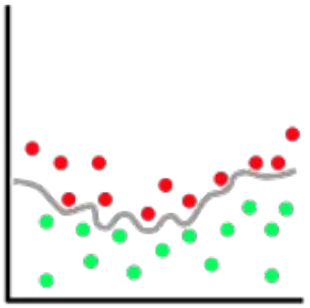


Classification

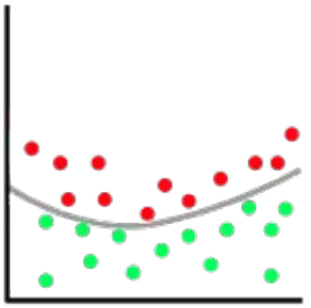
Building a ML potential : regression task



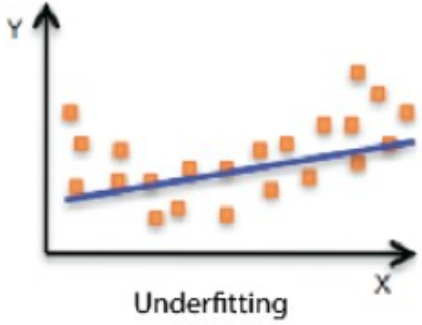
Underfitting



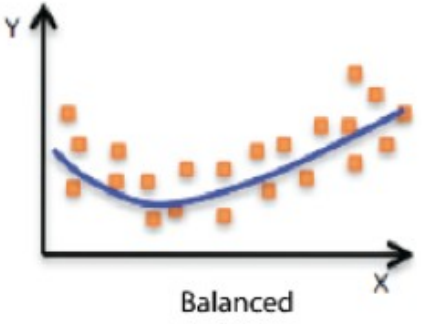
Overfitting



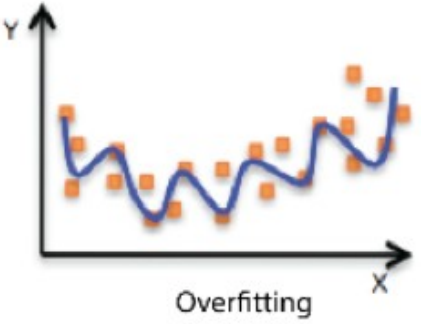
Balanced



Underfitting

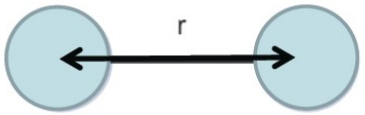


Balanced



Overfitting

Building a ML potential : regression task



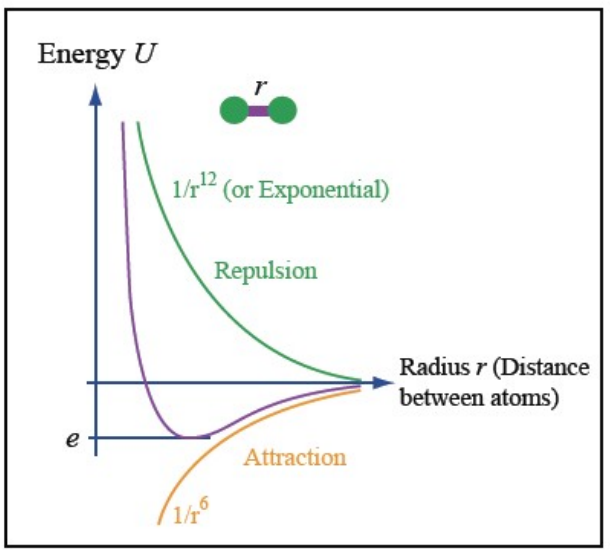
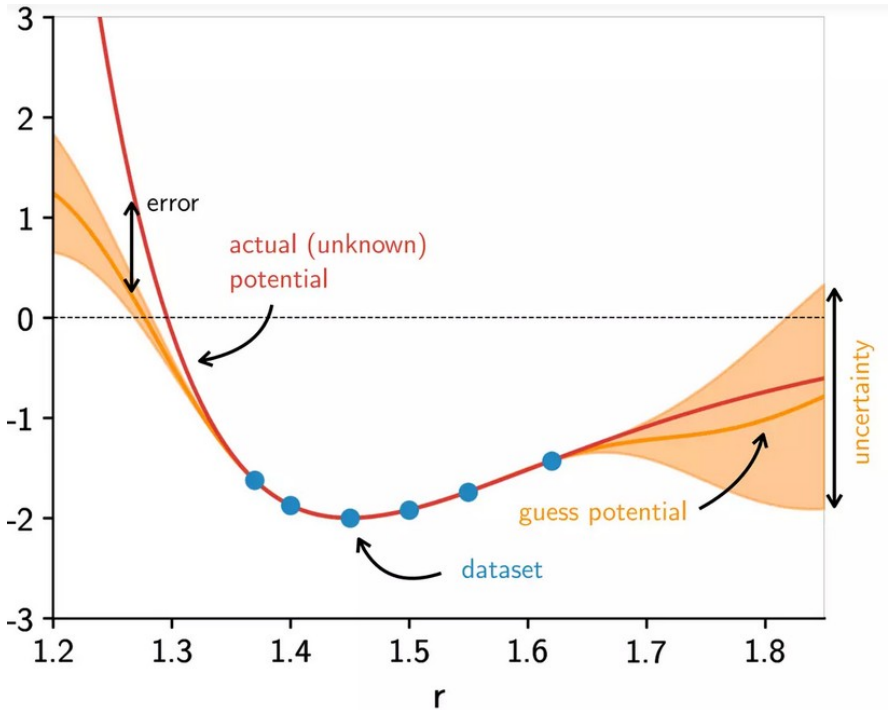
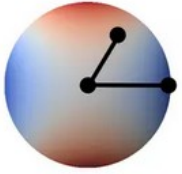


Image by MIT OpenCourseWare.



Building a ML potential : regression task

Linear methods

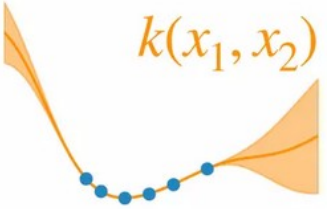


Polynomial on many-body terms

Simple and fast

Relies crafting a representation for the inputs

Kernel / Gaussian process regression

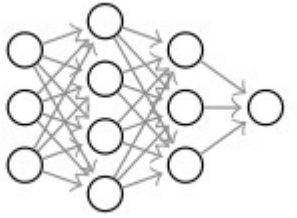


Computes an explicit similarity between points

Fewer data points (+uncertainty for GPR)

$O(N^3)$ complexity for training for GPR

Neural Networks



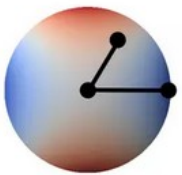
"Universal approximator" with non-linear mappings

High accuracy

Large number of trainable parameters

Building a ML potential : regression task

Linear methods

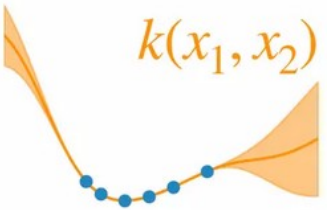


SNAP (Thompson et al.)

MTP (Shapeev)

ACE (Drautz, Kovács et al.)

Kernel / Gaussian process regression



GPR:

GAP (Bartok et al.)

MLOTF (Li et al.)

FLARE (Vandermause et al.)

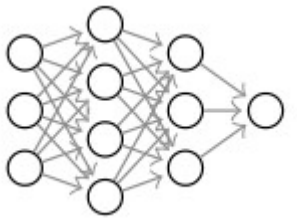
Other kernels with:

sGDML (Chmiela et al.)

FCHL repres. (Faber et al.)

Coulomb matrices (Rupp et al.)

Neural Networks



Behler-Parrinello

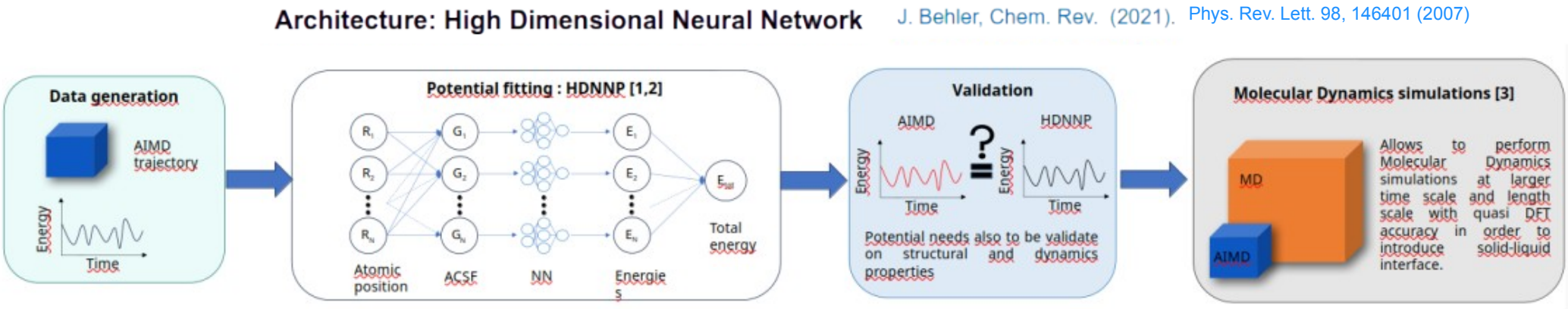
Representation + NN (DeepMD, ANI etc.)

Deep learning-based NNFF (SchNet etc.)

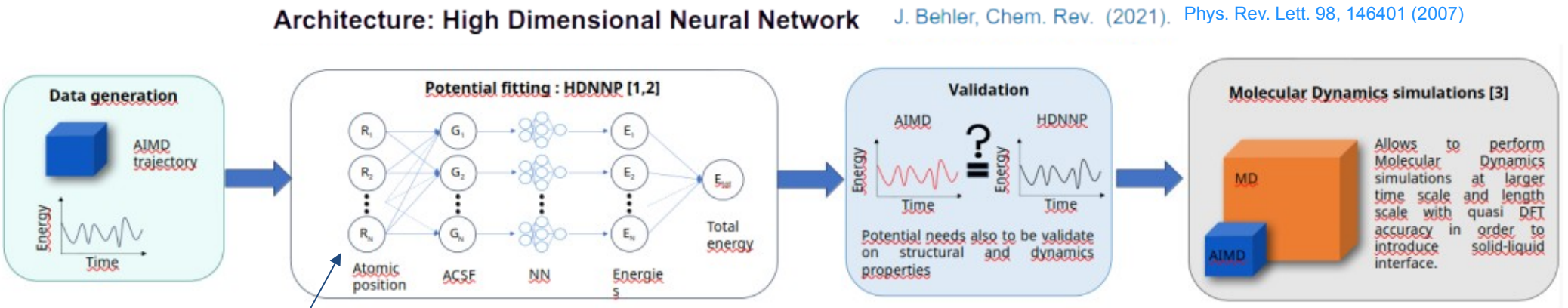
Deep learning + equivariance (NequIP, PaiNN etc.)

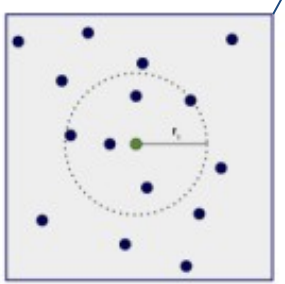
Deep learning + many-body expansion (MACE etc.)

Building a ML potential : regression task



Building a ML potential : regression task



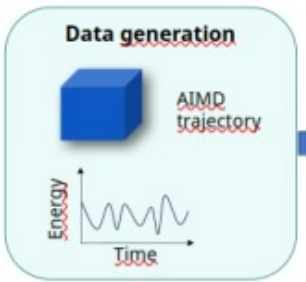


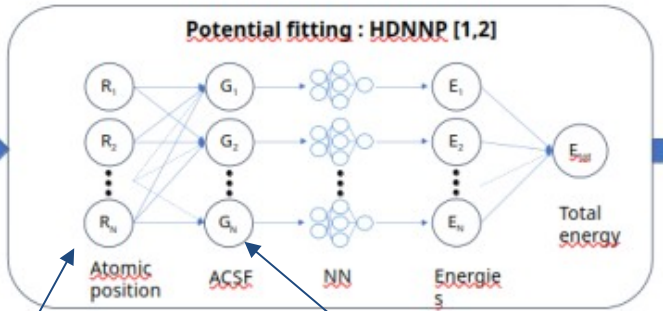
$$E = \sum_i E_i$$

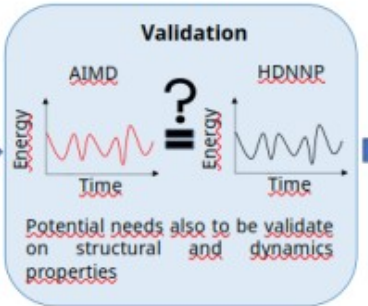
Building a ML potential : regression task

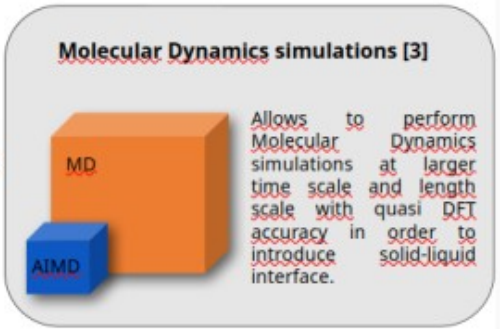
Architecture: High Dimensional Neural Network

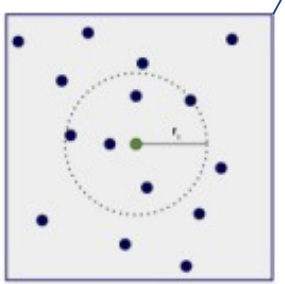
J. Behler, Chem. Rev. (2021). Phys. Rev. Lett. 98, 146401 (2007)



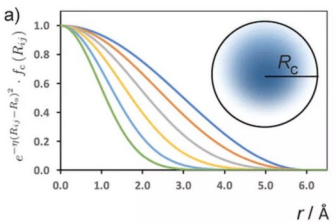






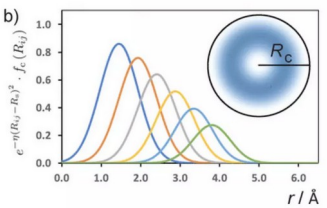


cutoff $$f_c(r_{ij}) = \frac{1}{2}\cos\left(\frac{\pi r_{ij}}{r_c}\right) + \frac{1}{2},\ r_{ij} < r_c$$



radial $$G_i^1 = \sum_{i \neq j} e^{-\eta(r_{ij}-r_s)^2} f_c(r_{ij})$$

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

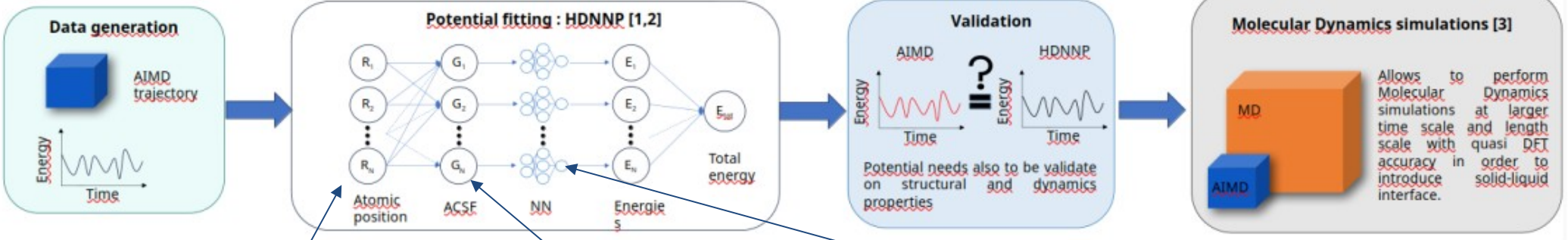


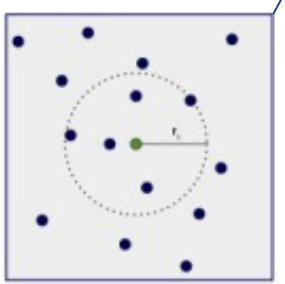
$$E = \sum_i E_i$$

Building a ML potential : regression task

Architecture: High Dimensional Neural Network

J. Behler, Chem. Rev. (2021). Phys. Rev. Lett. 98, 146401 (2007)



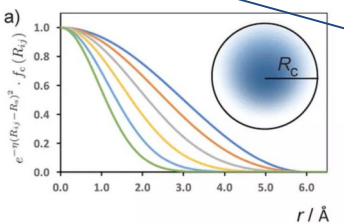


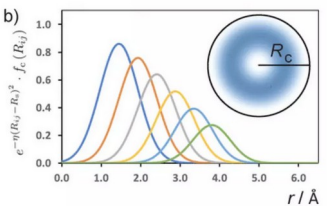
$$E = \sum_i E_i$$

cutoff $$f_c(r_{ij}) = \frac{1}{2}\cos\left(\frac{\pi r_{ij}}{r_c}\right) + \frac{1}{2},\ r_{ij} < r_c$$

radial $$G_i^1 = \sum_{i \neq j} e^{-\eta(r_{ij}-r_s)^2} f_c(r_{ij})$$

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





$$\hat{y} = f(\mathbf{X}) = \sigma(W\mathbf{X} + \mathbf{b})$$

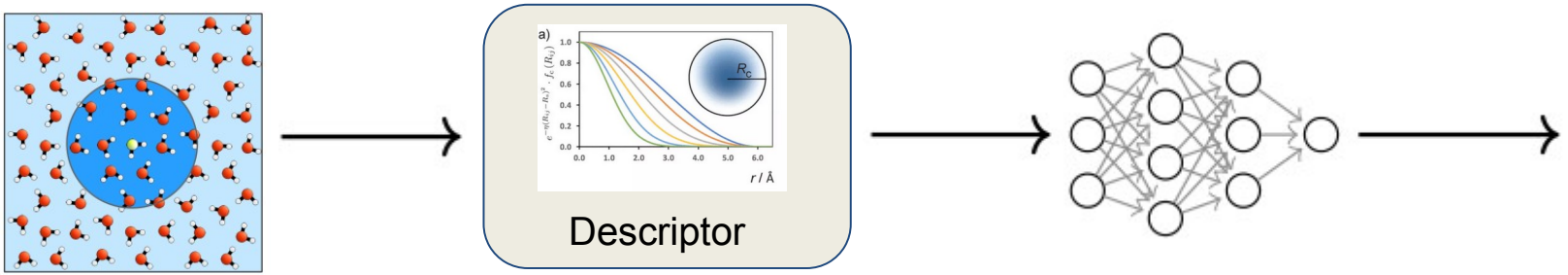
Loss function

$$\mathscr{L} = \mathbb{E}_{\mathbf{X} \sim P(X)} \left[\|\hat{y} - y\|^2 \right]$$

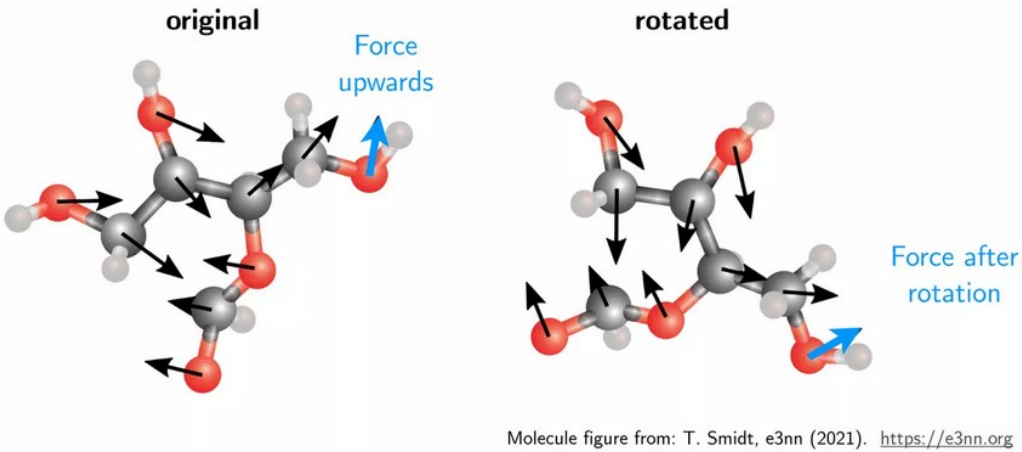
Back propagation and gradient opt.

$$w_{ij}^{(n+1)} = w_{ij}^{(n)} - \alpha \frac{\partial \mathscr{L}}{\partial w_{ij}}$$

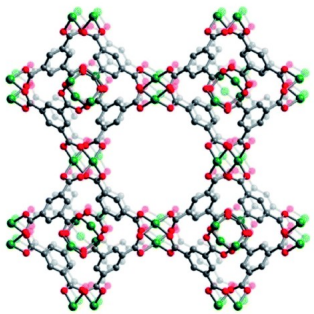
Representations of local atomic environment



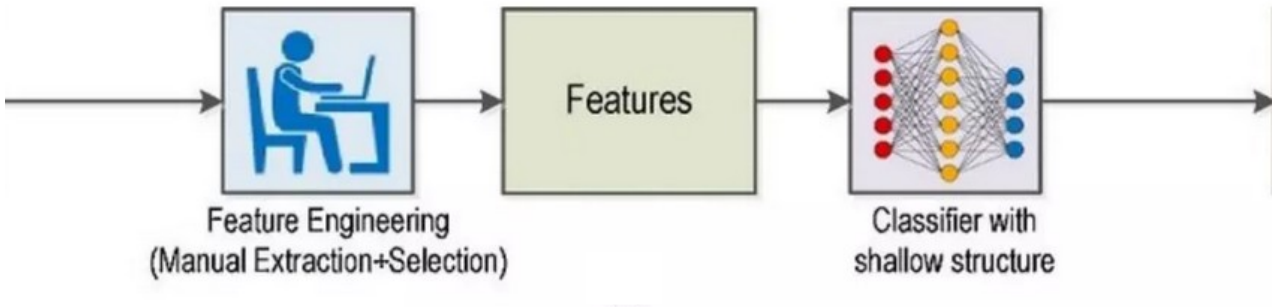
- **Direct coordinates:** are not suitable
- **Invariance:** translation, rotation, permutation
- **Equivariance :** property change, same effect



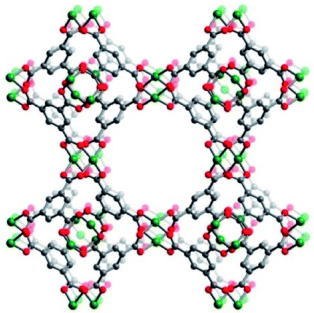
Machine Learning vs Deep Learning workflows



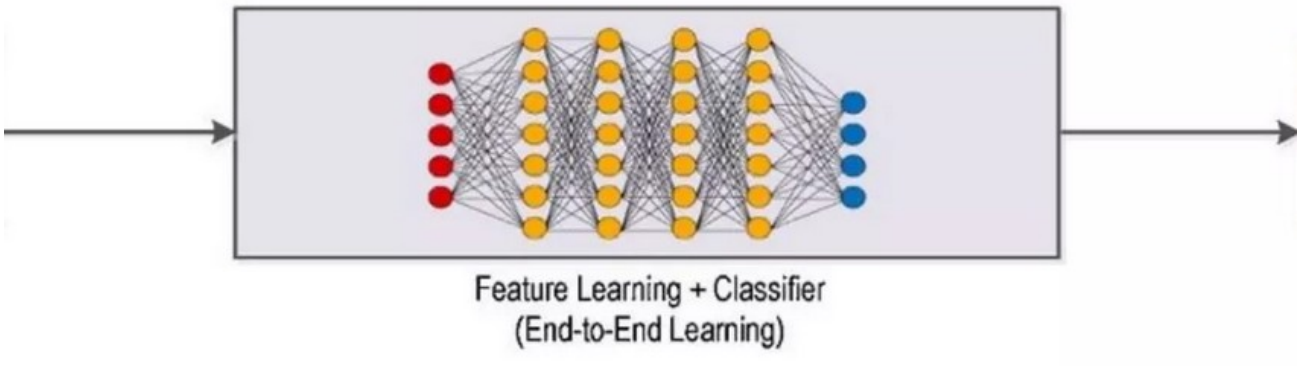
Predefined descriptors



Energy
Forces



Learning the descriptors



Energy
Forces

How to building a dataset for the training ?

To train a MLIP an appropriate dataset needs to be built for the application

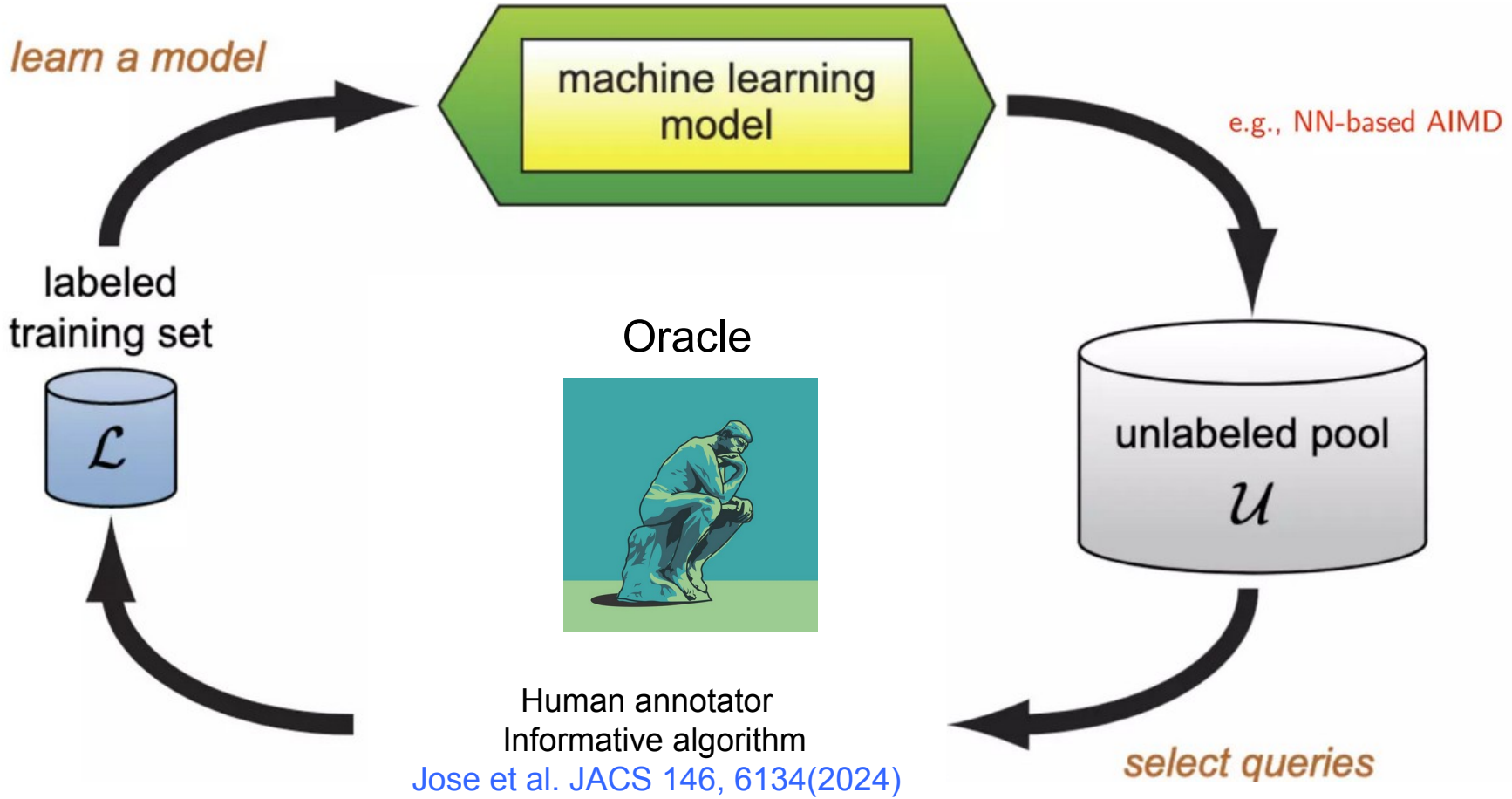
- **AIMD trajectories sampling:**
  - easy to perform but computationally costly and risk of correlated samples
- **Normal modes sampling:**
  - generate configuration along the Hessian matrix from given configurations
  - only small displacements allowed
- **PEL sampling:**
  - better sampling with enhanced sampling (for ex. Metadynamics)
  - tricky to implement and depend on the ab initio technique used
- **Active learning:**
  - Improve the dataset over the time
  - optimized and computationally cheaper to produce
  - but requires a metric for the uncertainty quantification to identify new configurations

How to building a dataset for the training ?

To train a MLIP an appropriate dataset needs to be built for the application

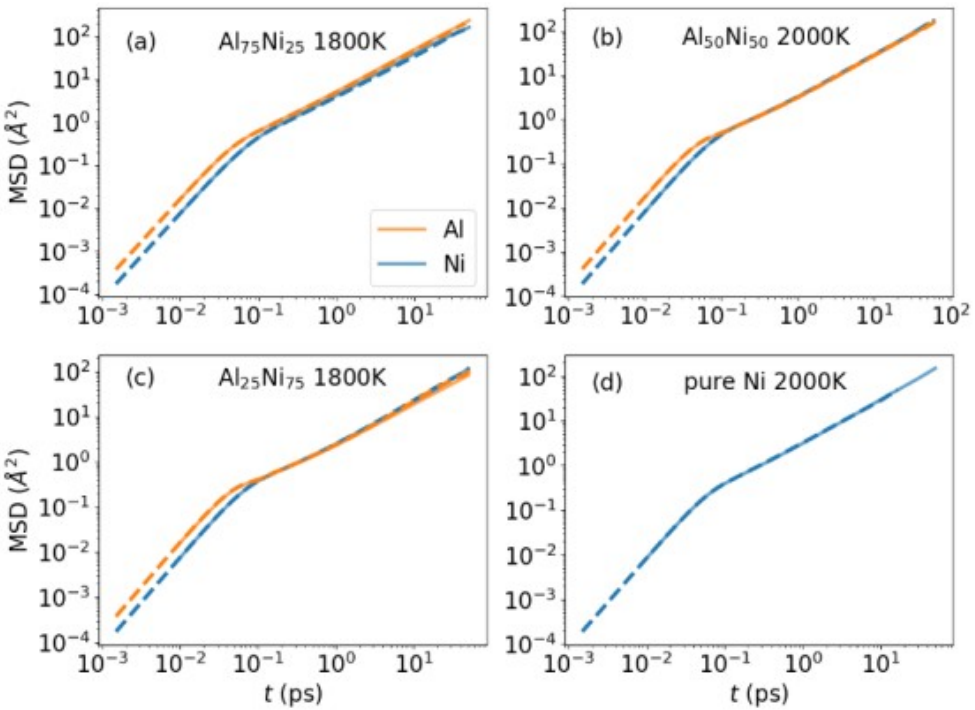
- **AIMD trajectories sampling:**
  - easy to perform but computationally costly and risk of correlated samples
- **Normal modes sampling:**
  - generate configuration along the Hessian matrix from given configurations
  - only small displacements allowed
- **PEL sampling:**
  - better sampling with enhanced sampling (for ex. Metadynamics)
  - tricky to implement and depend on the ab initio technique used
- **Active learning:**
  - Improve the dataset over the time
  - optimized and computationally cheaper to produce
  - but requires a metric for the uncertainty quantification to identify new configurations

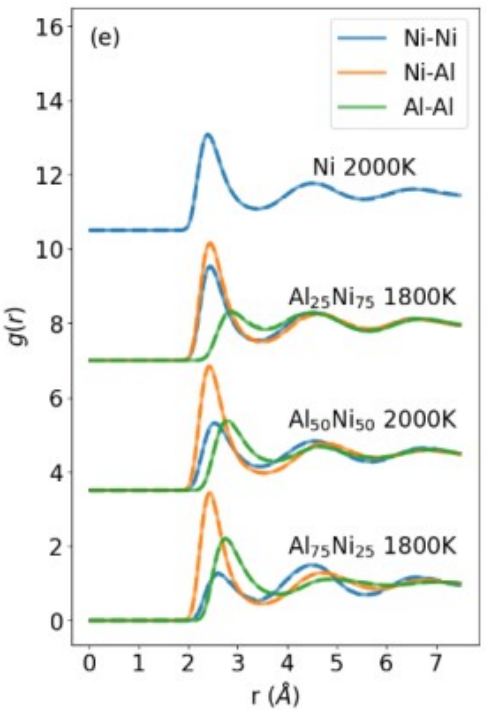
How to building a dataset for the training ?

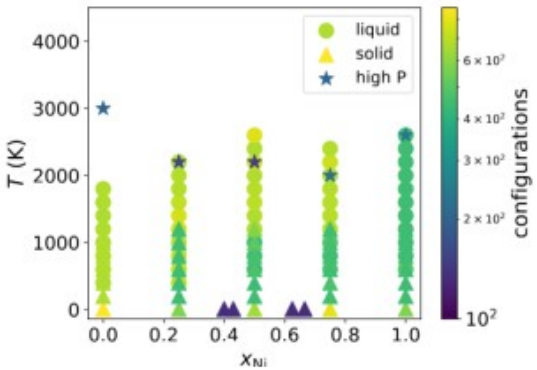


An example : nucleation pathways in Al-Ni Alloys

Regressor : HDNNP (N2P2 package),
dataset 24000 configurations (AIMD traj. Sampling),
Feature space : 64 dimensions
Performance : RMSE ~2.5 meV/atom





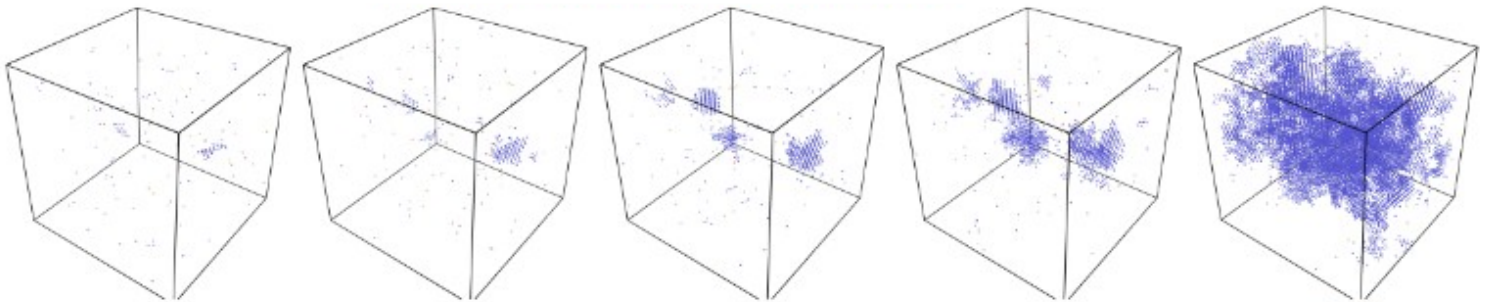


Sandberg et al arXiv/:2410.07886 (2024)

An example : nucleation pathways in Al-Ni Alloys

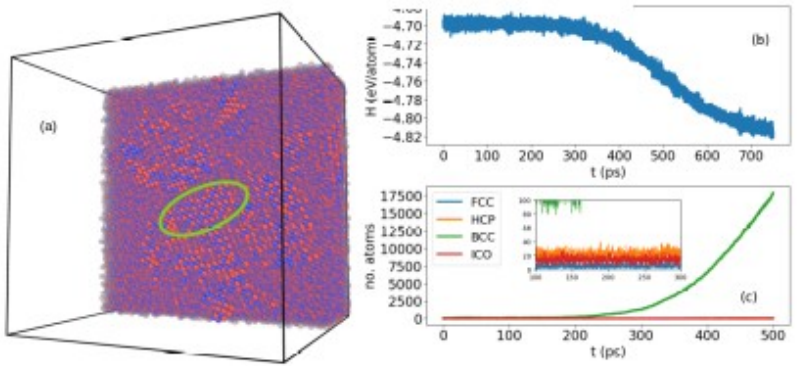
ML-MD simulation with 125000 atoms

Nucleation pathway at T = 1300 K



- One step nucleation process in the B2 structure

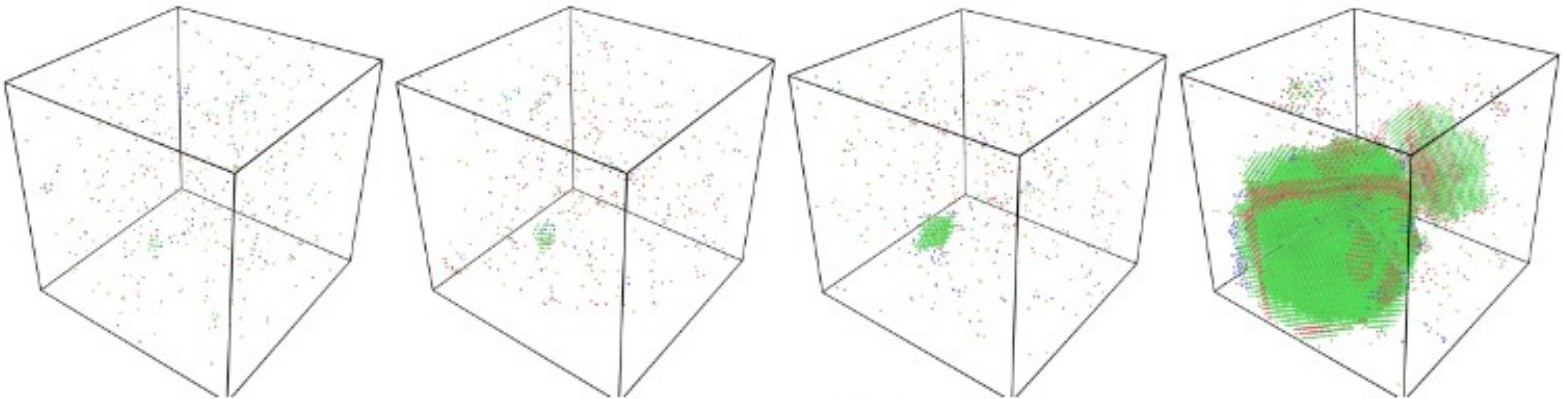
- The B2-like chemical order pre-exist in the liquid



Sandberg et al. arXiv:2410.07886 (2024)

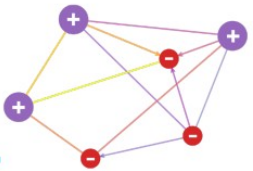
An example : nucleation pathways in Al-Ni Alloys

Pure Nickel



Outlook

CHGNet (Crystal Hamiltonian Graph neural Network)



nature machine intelligence

Article

https://doi.org/10.1038/s42256-023-00716-3

CHGNet as a pretrained universal neural network potential for charge-informed atomistic modelling

Received: 2 March 2023

Accepted: 4 August 2023

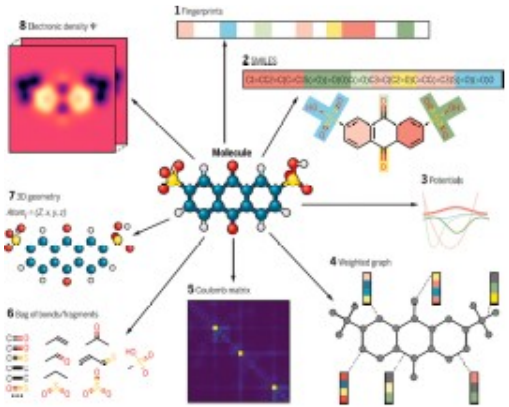
Bowen Deng[1,2], Peichen Zhong[1,2], KyuJung Jun[1,2], Janosh Riebesell[2,3], Kevin Han[2], Christopher J. Bartel[1,4] & Gerbrand Ceder[1,2]

https://github.com/CederGroupHub/chgnet
https://chgnet.lbl.gov/

Many developments towards universality : MACE-MP-0, ALIGNN, M3GNet, TeaNet, etc

Thank you for your attention

General concepts of ML for materials



Experimental + Ab initio databases

Datasets :
A resource of materials data

Human brain:
Descriptors :
A representation to describe each material quantitatively

Numerical brain
Intelligent search
(***Machine Learning tools***)

Results (Predictions):
Guide towards new materials
or at least
Define candidates
to be tested in real experiments