

精度保証された機械学習分子動力学法： 自己学習ハイブリッドモンテカルロ法

原子力機構システム計算科学センター
理研AIP

永井佑紀

YN, M. Okumura, K. Kobayashi, and M. Shiga,
“Self-learning Hybrid Monte Carlo: A First-principles Approach”,
arxiv:1909.02255

このセミナーについて



国内ワークショップ

2018年6月 場所：大阪大学

Y. Nagai : Self-learning Monte Carlo Method



国際ワークショップ

2019年10月31日~ 場所：京都大学



“物理学者、機械学習を使う — 機械学習・深層学習の物理学への応用 — ”

朝倉書店(2019/10)

第5章”自己学習モンテカルロ法”

ディープラーニングと物理学2020 オンライン

最新の進展の紹介

場所：インターネット

本日のお話

原子や分子の運動をシミュレーションしたい！

日常に現れるあらゆる物質（固体、液体、気体、生体）をシミュレーション

分子動力学 原子や分子に働く力とは？

原子は電子と原子核からなる→量子力学と電磁気学を使えば良い

電子や原子の数はアボガドロ数：無理！

3N次元シュレーディンガー方程式($N \sim 10^{23}$)

第一原理計算（密度汎関数理論）

3N次元波動関数と3次元電子密度の一対一対応：固体の計算が可能に

“第一原理計算”：原子位置を与えればエネルギーが計算できる

実用上は近似が入る…化学の人は“第一原理計算”と呼ばないetc

力はエネルギーの微分！

第一原理分子動力学計算：高精度なシミュレーション

めっちゃめっちゃ計算が重いのでなんとかしたい！

本研究

機械学習使ってみよう！→ 学習結果に依存しない精度が欲しい！

アウトライン

- 機械学習分子動力学法
- 自己学習ハイブリッドモンテカルロ法
- 適用例
- まとめ

シミュレーション中にニューラルネットワークを改善し、かつ、第一原理分子動力学計算精度で計算を行う

YN, M. Okumura, K. Kobayashi, and M. Shiga,
“Self-learning Hybrid Monte Carlo: A First-principles Approach”,
arxiv:1909.02255

機械学習分子動力学法

クイズ1：この動物は何？



猫です

Basile Morin / CC BY-SA (<https://creativecommons.org/licenses/by-sa/4.0>)

色々な猫を見たことがあるので、猫とわかる

クイズ2：この動物は何？



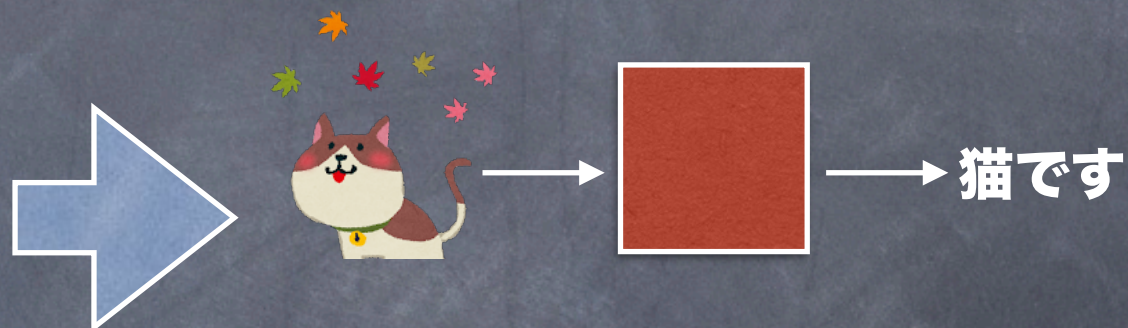
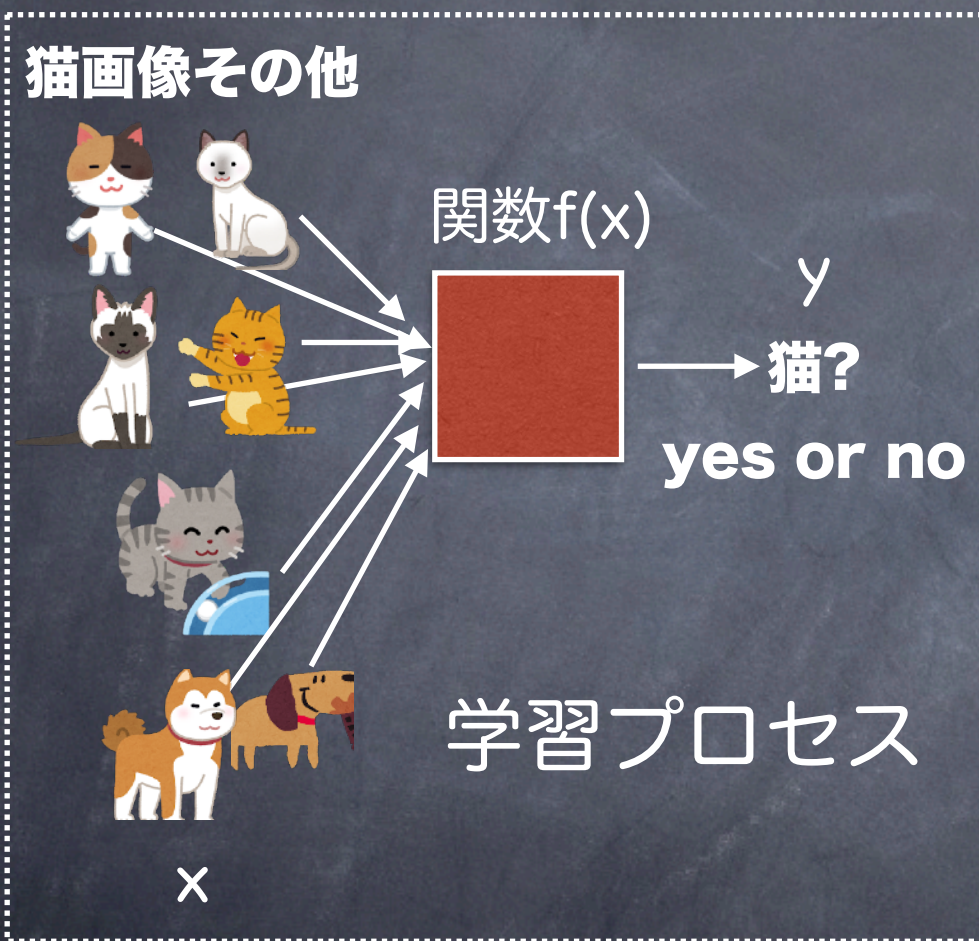
サイガです

ウシ科サイガ属に分類される偶蹄類（モンゴル）：絶滅危惧種

<https://ja-jp.facebook.com/9GAGCute/photos/saiga-antelope-a-priority-species-for-conservationthe-break-up-of-the-former-uss/816439751883761/>

見たことがなければわからない

機械学習とは



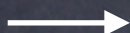
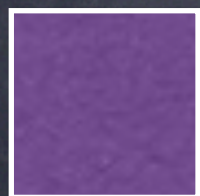
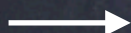
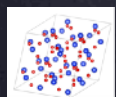
大量の入力データ x を用いて、
 $f(x)=y$ を満たす関数 $f(x)$ を決める

例：囲碁棋譜データ→勝利の方程式

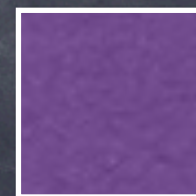
囲碁プロ棋士に勝利

機械学習分子動力学計算

原子配置



ある原子配置における
エネルギー



を使って高速に
シミュレーション

機械学習とは

教師あり学習

大量の入力データ x を用いて、 $f(x)=y$ を満たす関数 $f(x)$ を決める

一番シンプルな例

$$y = ax + b$$

直線で近似（線形回帰）

インプットが複数 $\rightarrow x$ がベクトル

$$y = Wx + b$$

もっと表現力を高めたい

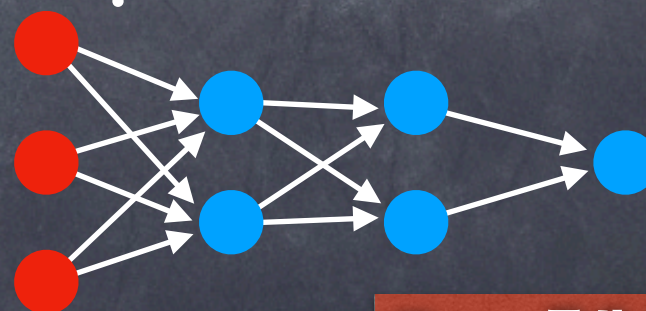
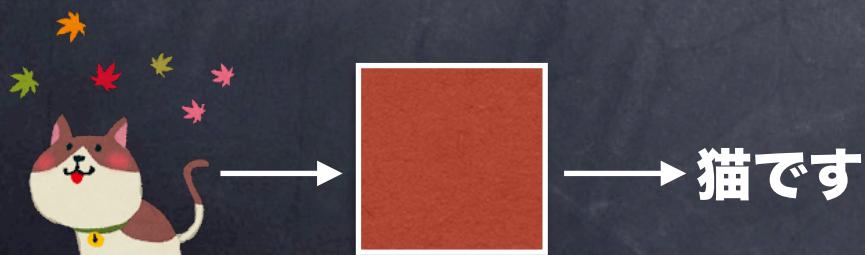
$$y = W_2 f(W_1 x + b_1) + b$$

f : 非線形関数

$$y = W_3 f(W_2 f(W_1 x + b_1) + b_2) + b$$

$\vdots$

深層学習



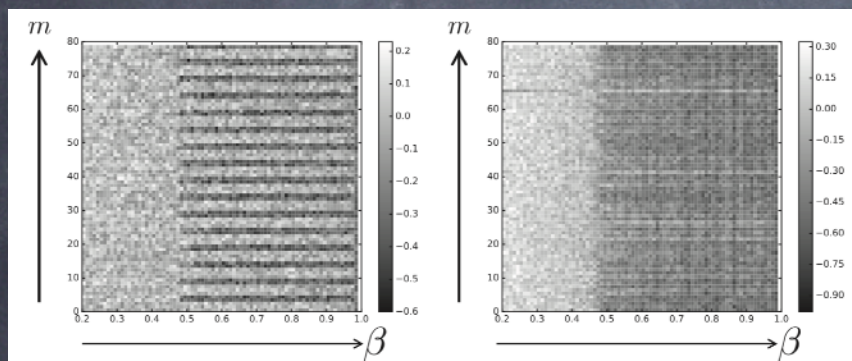
ニューラルネットワーク

機械学習の物理学への応用

例：相転移の検出

イジング模型の相転移検出

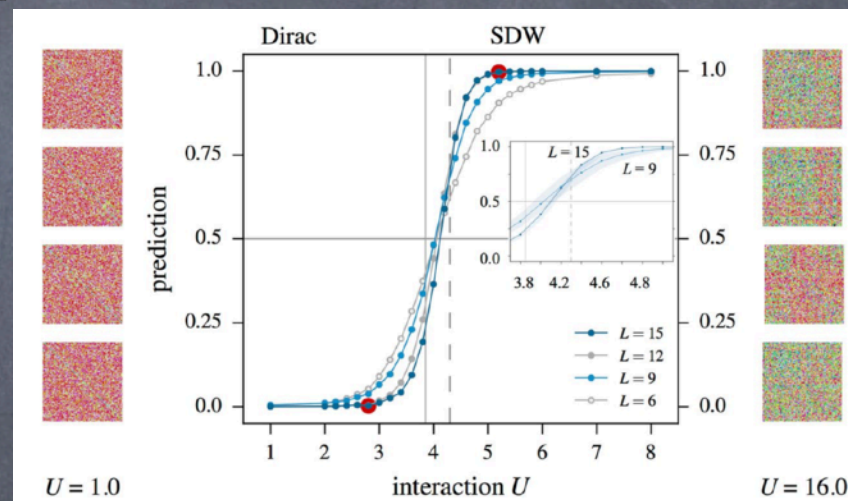
A. Tanaka and Y. Tomita, J. Phys. Soc. Jpn. 86, 063001 (2017)



スピン分布の解析

フェルミオン負符号問題のある量子系での 量子モンテカルロ法による相転移検出

P. Broecker et al. Scientific Reports 7 8823 (2017)



ハニカムハバード模型

同時刻グリーン関数の解析

多くの論文が画像認識技術をベースにしている

機械学習の物理学への応用

画像認識的な方法：相転移の検出等

別のアプローチはないか？

自然界は複雑すぎる

単純なモデルを作って解析する

例：ボールを投げたらどこに落ちる？

風は吹いていない「とする」 -> まあそんなに間違っていない結果



Done is better than perfect

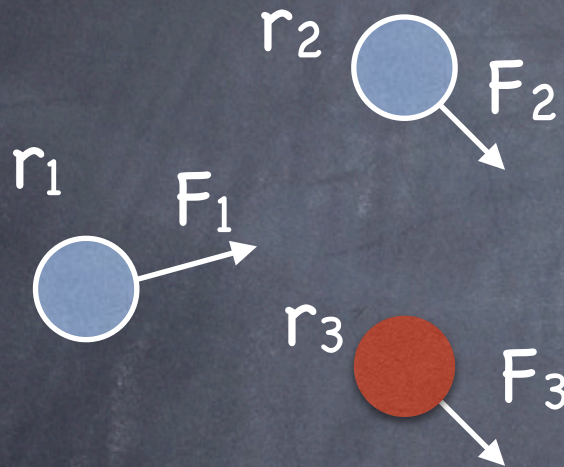
物理学者が長年やってきた方法：現象を記述する簡単なモデルを作る

機械が簡単なモデルを作ることにも可能では？

分子動力学法 (MD)

原子や分子に働く力を仮定して時間発展を追うシミュレーション

Lennard-Jones ポテンシャルetc.



力はエネルギーを偏微分すれば得られる

$$F_1 = \frac{\partial E(r_1, \dots, r_N)}{\partial r_1} \quad F_2 = \frac{\partial E(r_1, \dots, r_N)}{\partial r_2} \quad F_3 = \frac{\partial E(r_1, \dots, r_N)}{\partial r_3}$$

正確なエネルギーがあれば正確な力を計算できる

原子分布 $\{R\}$ の関数としてのエネルギー $E(\{R\})$

第一原理計算：多体系における高精度なエネルギー計算

密度汎関数理論:Hohenberg-Kohn theorems

3N次元波動関数と3次元電子密度が一対一対応

多電子系の基底状態の性質が3つの空間座標だけに依存する電子密度によって一意に決定される

固体の電子状態の計算が可能に！

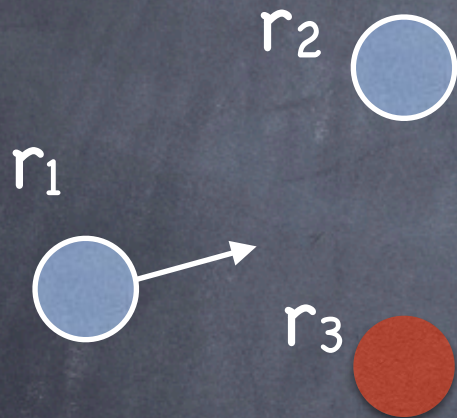
時間ステップごとに第一原理計算して力を求める
非常に計算が重い！

機械学習を使えないか？

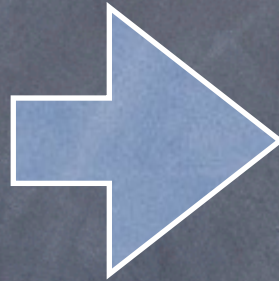
機械学習をどう使う？

インプット

原子の配置 $\{r_i\}$



第一原理計算etc

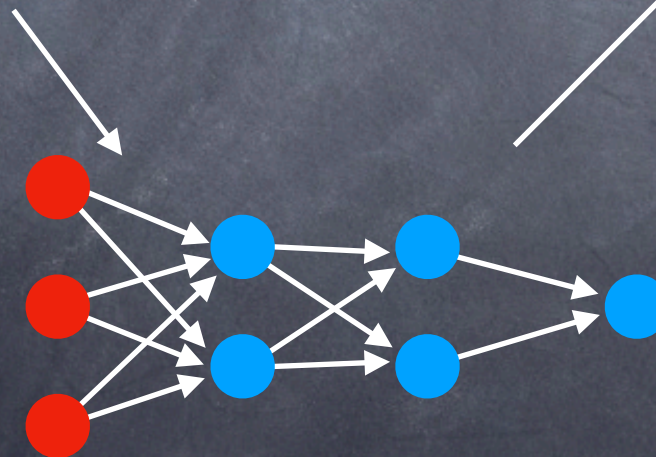


アウトプット

ポテンシャルエネルギー
 $E(\{r_i\})$

原子間に働く力

$$F_1 = \frac{\partial E(r_1, \dots, r_N)}{\partial r_1}$$



ニューラルネットワーク

インプットを入れればアウト
プットが出てくるニューラル
ネットを作ればいいのね

そう簡単ではない

機械学習をどう使う？

第一原理分子動力学で機械学習が使いたい時



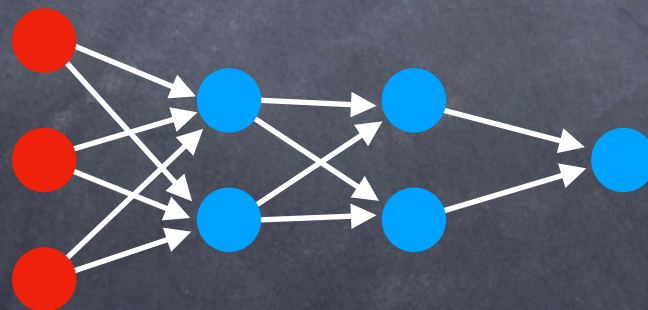
従来では重すぎて実行できないような
大きな系の計算を行いたい時

学習で作ったニューラルネット
は**原子数の異なるより大きな系**
で使いたい！

例：H₂O 32個で学習 -> H₂O 1000個で実行

インプット

アウトプット



どんなインプット、どんな形状のニューラルネットなら可能か？

Behler-Parrinello型ニューラルネットワーク

第一原理計算：高精度な力の評価→高精度シミュレーション

時間ステップ1fsのMD → 1nsのシミュレーションに 10^6 回の第一原理計算が必要

とても重たい → ニューラルネットでフィットしたい

“Neural network models of potential energy surfaces”

T. B. Blank et al., J. Chem. Phys. **103**, 4129 (1995)

解決すべき問題

1. エネルギー $E(r_1, r_2, \dots, r_N)$ は $3N$ 次元空間の関数で複雑
2. Portability: 作ったニューラルネットを他で使えるか？
3. 原子が持つ対称性（回転対称性、並進対称性）を持つか？

2007 (13年前)

“Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces”, J. Behler and M. Parrinello, Phys. Rev. Lett. **98**, 146401 (2007)

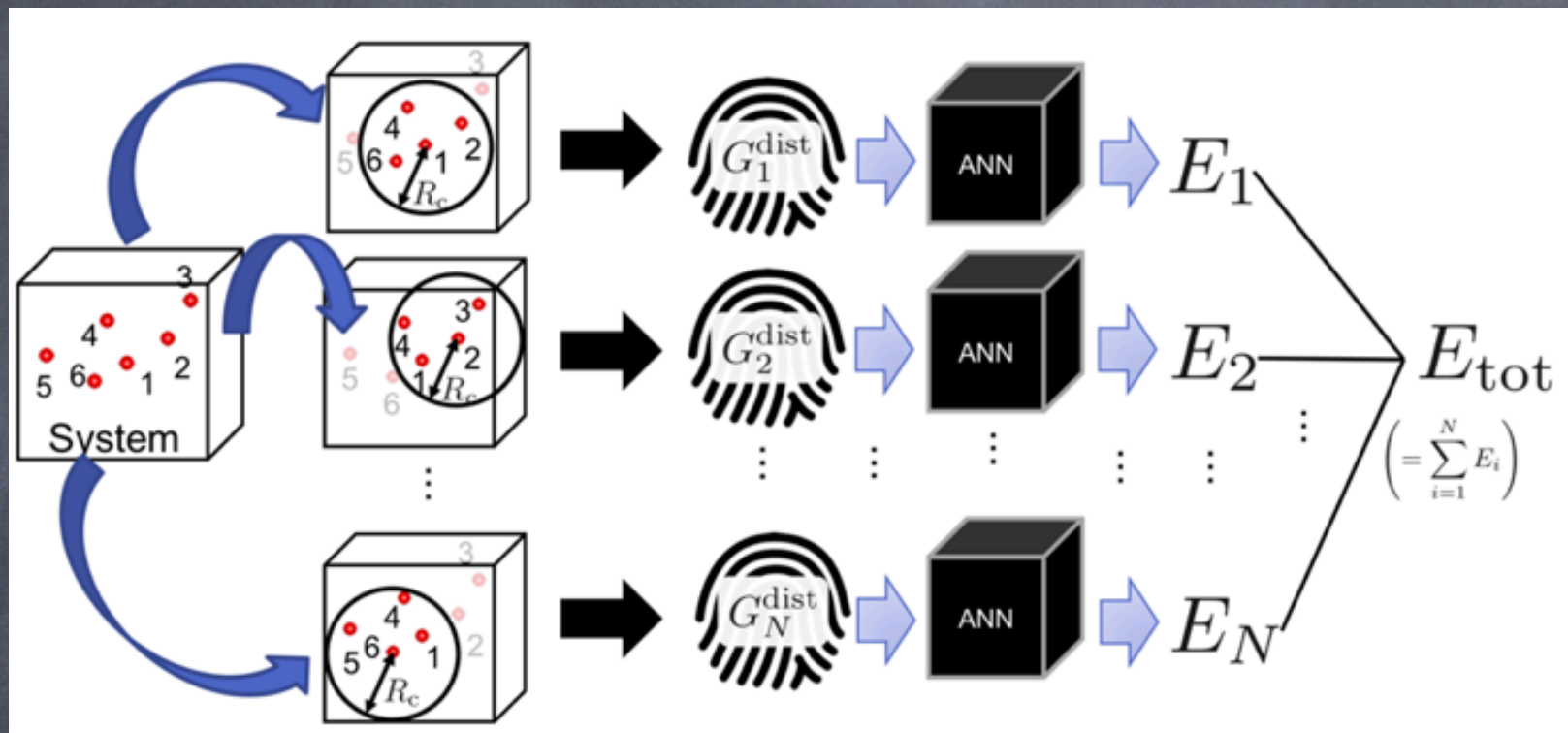
原子配置 → 第一原理計算 → エネルギー



原子配置 → ニューラルネット → エネルギー

1000倍以上の高速化

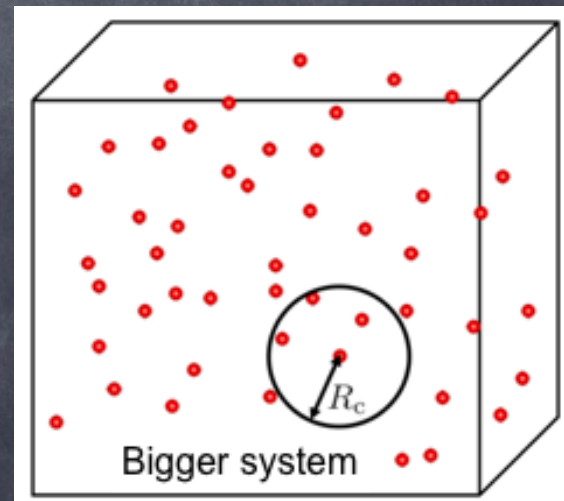
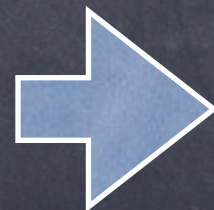
Behler-Parrinello型ニューラルネットワーク



基本アイディア

1. インプットは原子ごとに局所的な位置情報だけから作る：原子ごとの”指紋”
2. 同じ原子種は同じニューラルネットワークを使う

原子数増減
に対応可能

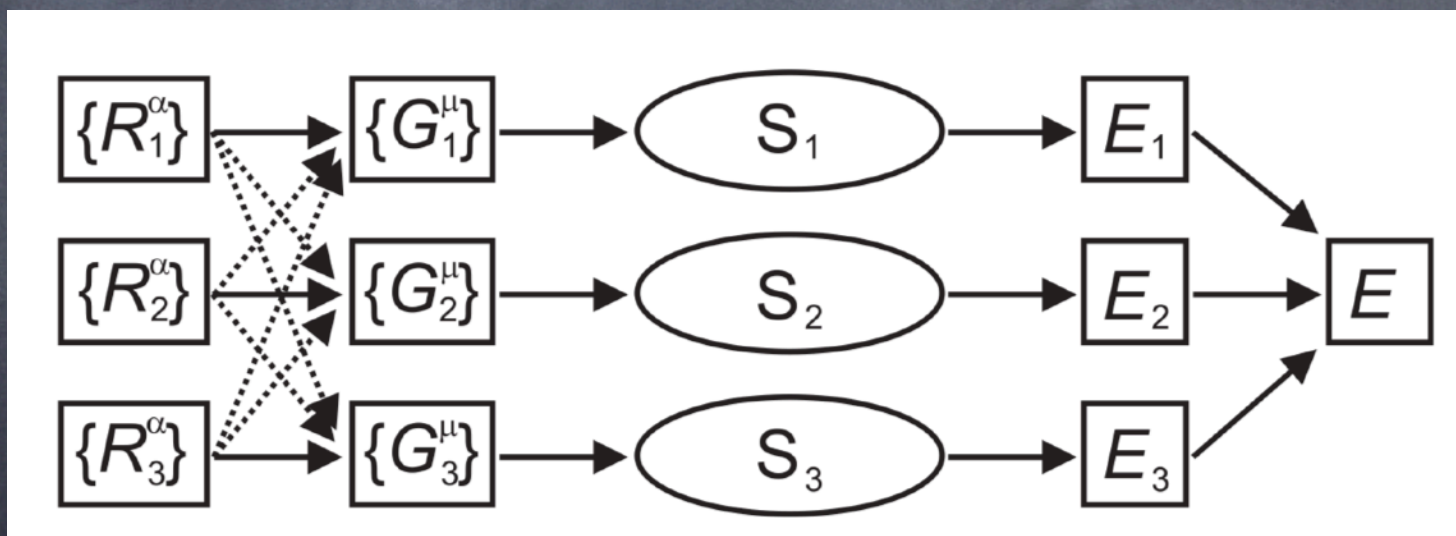


Behler-Parrinello型ニューラルネットワーク

(r_1, r_2, r_3) の代わりに $(G_1(r_1, r_2, r_3), \dots, G_3(r_1, r_2, r_3))$ をインプットに使う

$$G_i^1 = \sum_{j \neq i}^{\text{all}} e^{-\eta(R_{ij} - R_s)^2} f_c(R_{ij}), \quad (2)$$

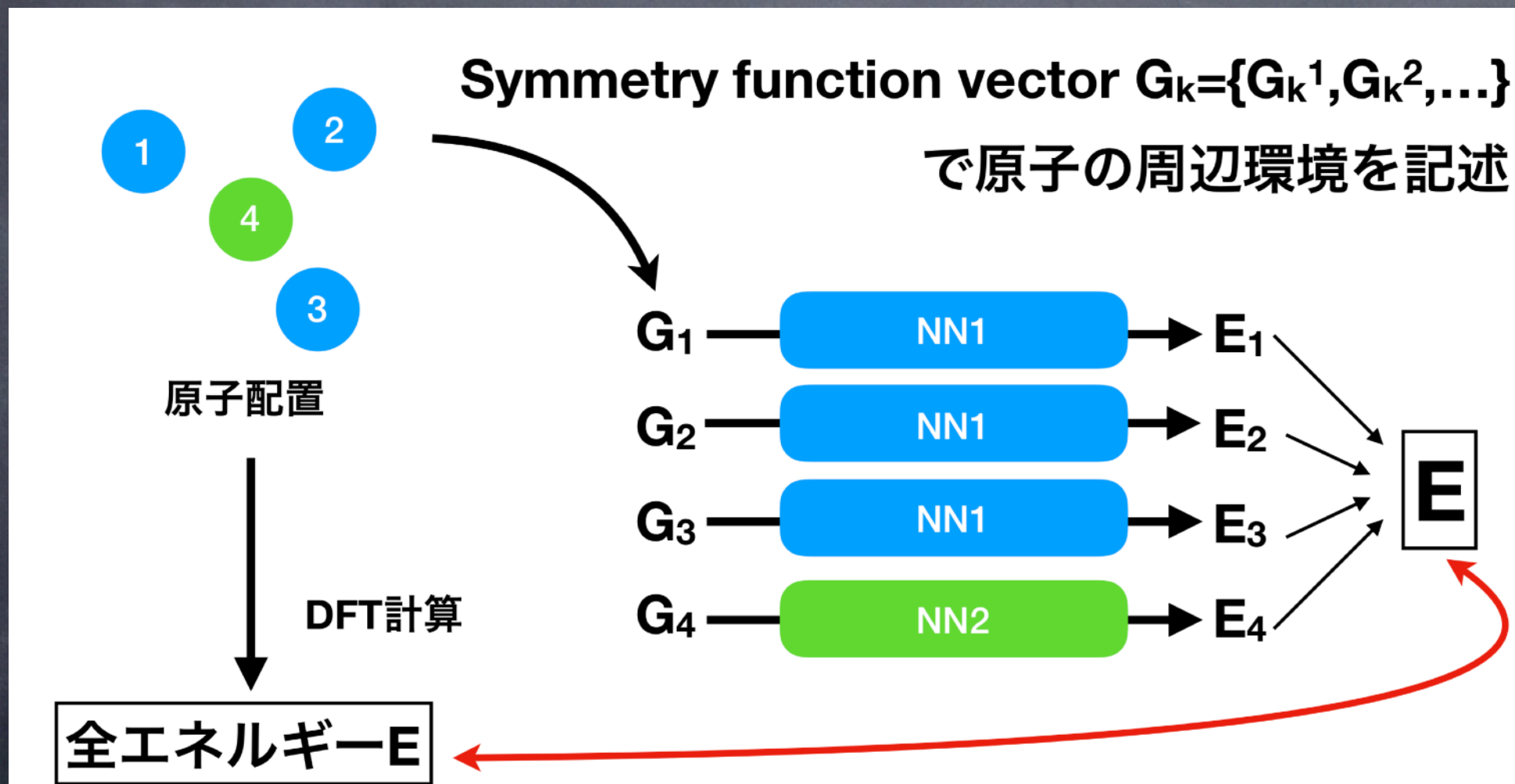
$$G_i^2 = 2^{1-\zeta} \sum_{j, k \neq i}^{\text{all}} (1 + \lambda \cos \theta_{ijk})^\zeta e^{-\eta(R_{ij}^2 + R_{ik}^2 + R_{jk}^2)} f_c(R_{ij}) f_c(R_{ik}) f_c(R_{jk}), \quad (3)$$



S_1 : 原子周りでのニューラルネットワーク
回転対称性や並進対称性が自動的に考慮されている

Behler-Parrinello型ニューラルネットワーク

教師あり学習：インプットとアウトプット



Schematic figure from Y. Ando, Deep-learning and physics 2018 @Osaka Univ.

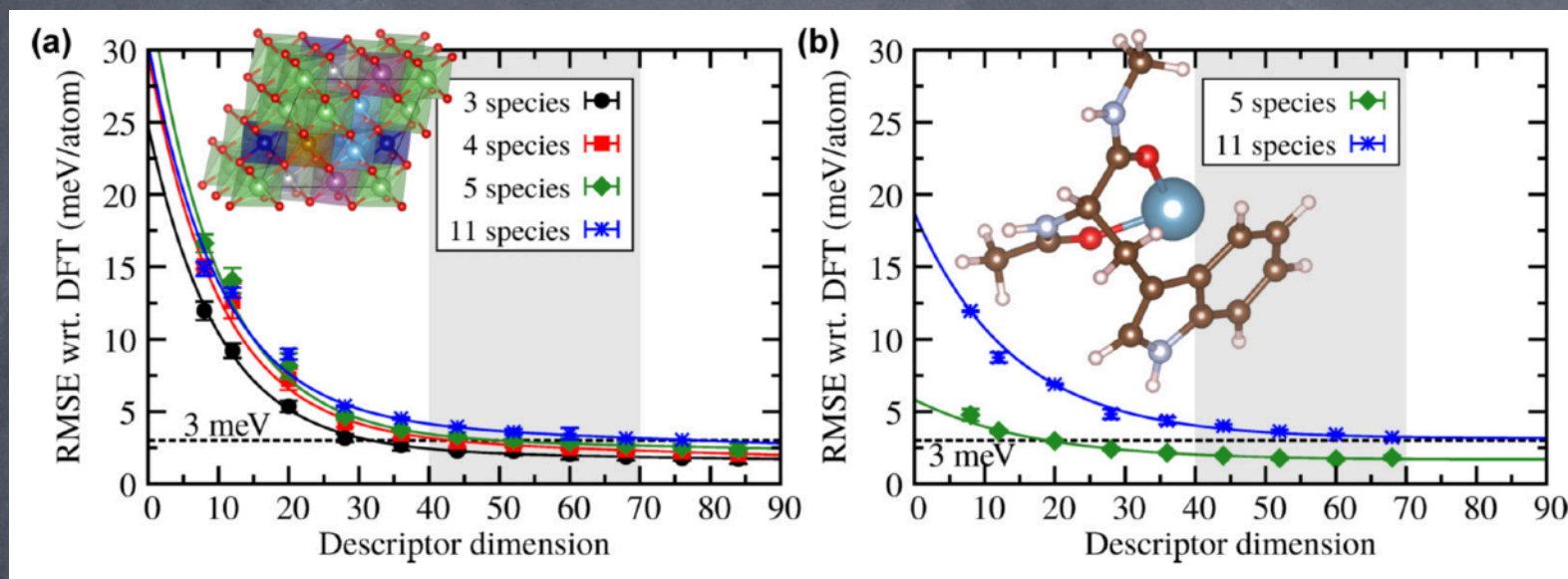
同じ原子種には同じニューラルネットワークを使用

Gの数が固定→ 原子数が増えても同じニューラルネットに対応が可能

オープンソースソフトウェアaenet

N. Artrith, A. Urban, and G. Ceder, Phys. Rev. B 96, 014112 (2017). LDA+U 16,047 LiMO_2

インプットベクトルの長さ（周辺環境を記述するGの数）とニューラルネットワークの精度



隠れ層2層(各15ユニット)のニューラルネットワーク

エネルギーの精度: 3mev/atom



The Atomic Energy Network (aenet)

<http://ann.atomistic.net>

近年：オープンソースソフトウェアが充実

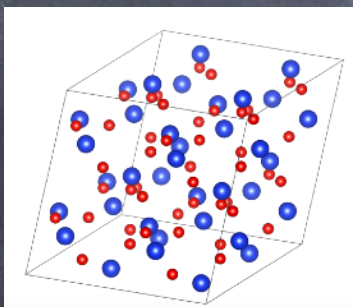
2007年から10年近く、弟子入りしないと使えないような状態だった...

自己学習ハイブリッド モンテカルロ法

機械学習MDの改善点

訓練データをどう集めるか？

2通り 1. つぶしたりのぼしたりして構造を系を沢山作る



色々な原子位置に対してエネルギーを第一原理計算

->教師データを集める

問題点：どういう系を集めてくれば良い？（配置は無限にある）

問題点：何個集めれば十分に精度の良いニューラルネットができる？

1000個？10000個？100000個？

2. 第一原理分子動力学計算を実行して、構造を集める

1ステップごとに訓練データが増える

問題点：何個集めれば十分に精度の良いニューラルネットができる？

1000ステップ？10000ステップ？100000ステップ？

問題点：これを高速化したいのに…。時間がかかりすぎる

機械学習MDの改善点

作ったニューラルネットワークの精度は？

悪いニューラルネットワークは悪い力を返す

より大きな系に適用したい！

その系でも精度がいい保証は？

多成分系(4元素以上)に弱い

入力データが複雑になり特徴が捉えづらくなる→悪いポテンシャルに

ニューラルネットワークの質に踊らされたくない！

→自己学習モンテカルロ法を使おう

自己学習モンテカルロ法

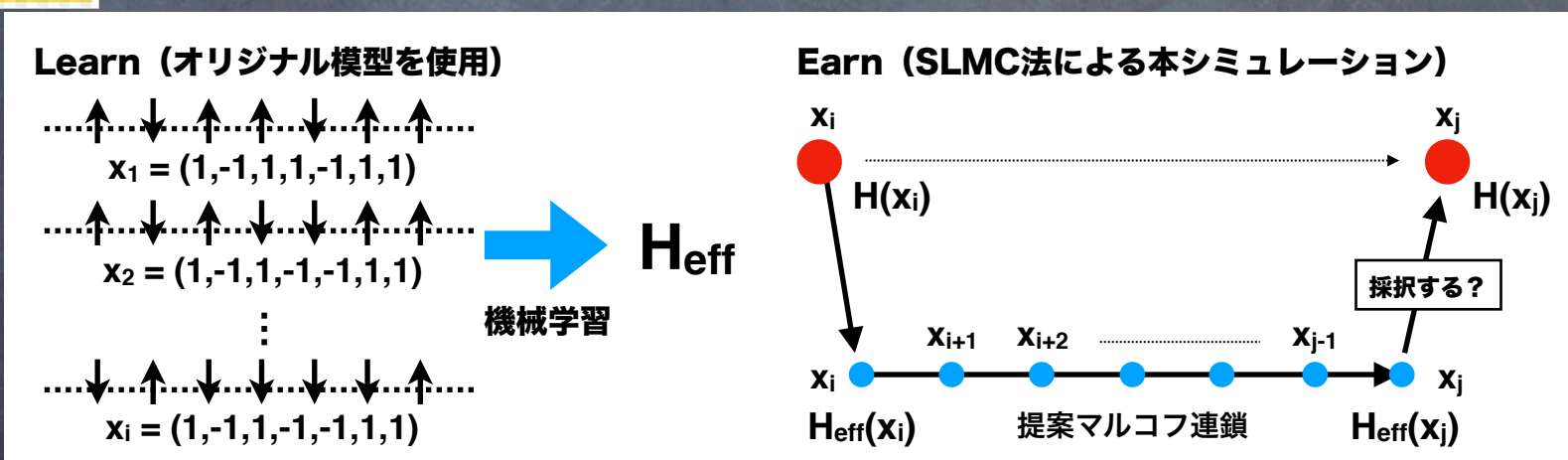
モンテカルロ法を機械学習を使って高速化する手法



“物理学者、機械学習を使う — 機械学習・深層学習の物理学への応用 — ”

朝倉書店(2019/10)

第5章”自己学習モンテカルロ法”



機械学習で有効模型を作り、計算量を削減
 固体物理、特にスピン系や電子系の模型に適用されてきた
 結果は統計的に厳密に元のモンテカルロ法に等しい

機械学習MDにも使えるのでは？

SLMCと機械学習MDの比較

自己学習モンテカルロ法(SLMC)

モンテカルロ配位C

スピン配位etc

モンテカルロ重み $W(C)$

有効ハミルトニアン
でフィッティング

有効モンテカルロ重み $W_{\text{eff}}(C)$

有効モデルは配位提案にのみ使う

厳密！

機械学習分子動力学法

原子位置 $R=(r_1, r_2, r_3, \dots)$

第一原理計算による
エネルギー $E(R)$

ニューラルネット
でフィッティング

有効エネルギー $E_{\text{eff}}(R)$

有効モデルのMDで時間発展：近似

MCを使おう！

厳密！

ハイブリッドモンテカルロ法

分子動力学法もモンテカルロ法も統計平均をとるという意味では同じ
分配関数と物理量の期待値を計算する手法

ハイブリッドモンテカルロ法

格子QCDで使われている

S. Gottlieb, W. Liu, D. Toussaint, R. L. Renken, and R. L. Sugar, Hybrid-molecular-dynamics algorithms for the numerical simulation of quantum chromodynamics, Phys. Rev. D 35, 2531 (1987).

固体物理でも考案されている

B. Mehlig, D. W. Heermann, and B. M. Forrest, Hybrid Monte Carlo method for condensed-matter systems, Phys. Rev. B 45, 679 (1992).

MDで十分だったからあまり使われてこなかった

初期原子配置 $\{R_i\}$

提案原子配置 $\{R_i'\}$ $\{R_i'\}$ を受け入れる？



メトロポリス法に従って受け入れ確率を計算

MD+MC = ハイブリッドMC

自己学習ハイブリッドモンテカルロ法 (SLHMC)

機械学習ポテンシャルを利用して、第一原理MDと統計的に**厳密**に等しい結果を得ることができる

機械学習=近似計算 **ではない**

ハイブリッドモンテカルロ法の一つ

原子の運動量{p}と位置{r}をモンテカルロ法で変更

1. Trial configurationを機械学習MDで生成

$$\dot{\mathbf{p}}_i = -\frac{\partial V_{\text{ML}}(\{\mathbf{r}\}, t)}{\partial \mathbf{r}_i}, \quad \dot{\mathbf{r}}_i = \frac{\mathbf{p}_i}{m_i},$$

2. 提案した{p'}, {r'}を以下の確率で受け入れる

$$P_{\text{acc}}(\{\mathbf{p}, \mathbf{r}\} \rightarrow \{\mathbf{p}', \mathbf{r}'\}) = \min \left(1, e^{-\beta(H_{\text{DFT}}(\{\mathbf{p}', \mathbf{r}'\}) - H_{\text{DFT}}(\{\mathbf{p}, \mathbf{r}\}))} \right)$$



DFTのエネルギーをボルツマン重みとした
メトロポリスヘイスティングス法

機械学習ポテンシャルは配位提案に使う
->第一原理HMCと結果が厳密に等しい

自己学習ハイブリッドモンテカルロ法 (SLHMC)

1. Trial configurationを機械学習MDで生成

2. 提案した $\{p'\}, \{r'\}$ を以下の確率で受け入れる

$$\dot{\mathbf{p}}_i = -\frac{\partial V_{\text{ML}}(\{\mathbf{r}\}, t)}{\partial \mathbf{r}_i}, \quad \dot{\mathbf{r}}_i = \frac{\mathbf{p}_i}{m_i},$$

$$P_{\text{acc}}(\{\mathbf{p}, \mathbf{r}\} \rightarrow \{\mathbf{p}', \mathbf{r}'\}) = \min \left(1, e^{-\beta(H_{\text{DFT}}(\{\mathbf{p}', \mathbf{r}'\}) - H_{\text{DFT}}(\{\mathbf{p}, \mathbf{r}\}))} \right)$$



MDの時間ステップ Δt が微小ならば

$$P_{\text{acc}}(\mathbf{p}, \mathbf{r}) \rightarrow \{\mathbf{p}', \mathbf{r}'\} \sim \min \left(1, \exp \left[\underbrace{-\beta(V_{\text{DFT}}(\{\mathbf{r}'\}) - V_{\text{ML}}(\{\mathbf{r}'\}))}_{\text{ポテンシャルの差}} - \underbrace{(V_{\text{DFT}}(\{\mathbf{r}\}) - V_{\text{ML}}(\{\mathbf{r}\}))}_{\text{ポテンシャルの差}} \right] \right)$$

機械学習ポテンシャルが完全に第一原理計算のポテンシャルを模倣すると、アクセプト率は1となる

自己学習ハイブリッドモンテカルロ法 (SLHMC)

機械学習ポテンシャルを利用して、第一原理MDと統計的に**厳密**に等しい結果を得ることができる

機械学習=近似計算 **ではない**

1. 学習に使うデータを自動的に効率的に収集

第一原理MDと同じアンサンブルを出力 -> 効率的なデータ集め

機械学習MDを行う目的の一つは第一原理MDの代わりに計算すること

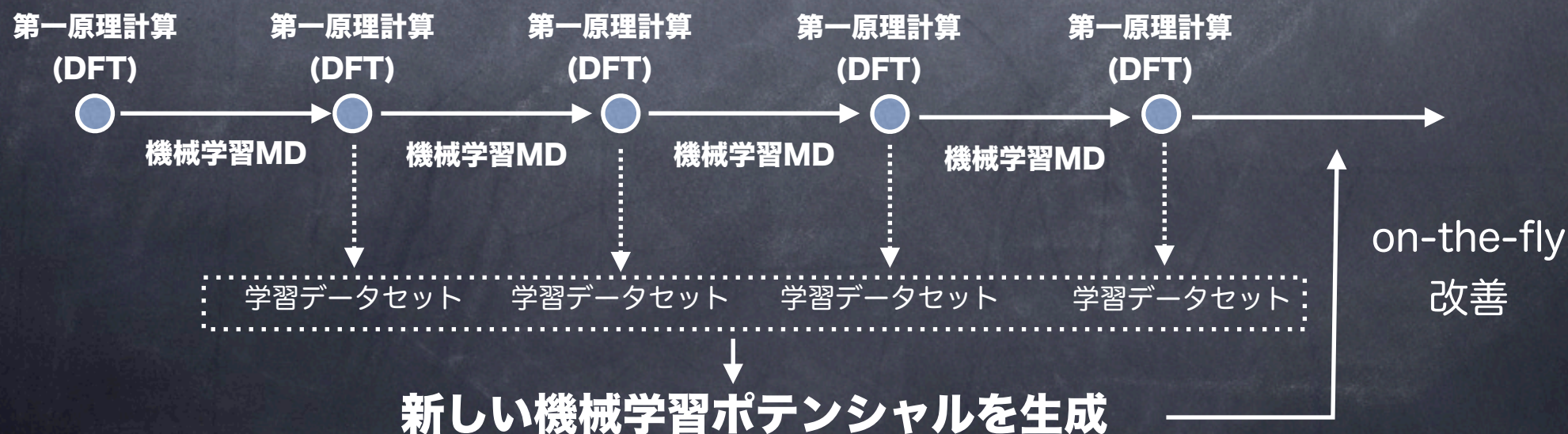
ターゲットの第一原理MDと同じアンサンブル -> 欲しいデータそのもの

大量の学習データをあらかじめ集める必要がない

ポテンシャルはon-the-flyで逐次的に改善されていく

機械学習ポテンシャルを利用して、第一原理MDと統計的に**厳密**に等しい結果を得ることができる

2. 作ったポテンシャルは近似でも結果は厳密+トレーニングしながら物理量計算



分子シミュレーション手法の比較

	計算精度	計算コスト
古典分子動力学	低い	非常に低い
第一原理分子動力学	非常に高い	非常に高い
機械学習分子動力学	高い（学習結果に依存）	低い
自己学習ハイブリッドモンテカルロ法	非常に高い（学習結果に依存しない）	中間

要求される精度と計算コスト（サイズ、原子数）で手法を選ぶ

SLHMCは二つの利用法がある

1. シミュレーション：第一原理分子動力学法ができない領域で精度保証で計算したい
 2. ニューラルネット作成：機械学習分子動力学用のニューラルネットを高精度に構築したい
- どちらにせよ、ニューラルネットは自動的に構築される->手間も大幅減

実際にシミュレーションしたい方へ

第一原理分子動力学パッケージ PIMD

<https://ccse.jaea.go.jp/software/PIMD/index.en.html>

経路積分第一原理分子動力学（水素の量子効果考慮等）

も可能

VASP, QuantumEspresso等様々な第一原理計算パッケージに対応

機械学習MDが可能

The Atomic Energy Network (aenet)



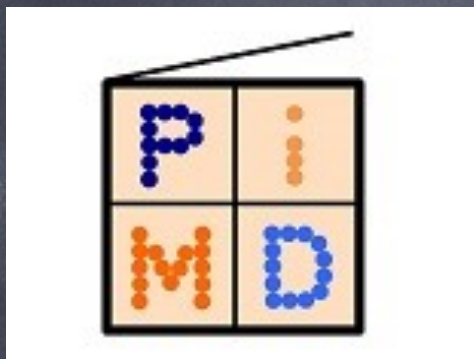
自己学習ハイブリッドモンテカルロ法が実装済み！

実習開催の実績あり（東京大学2019年）

関連記事

PIMDとAENETを使って機械学習分子シミュレーションをやってみよう

https://qiita.com/cometscome_phys/items/660a84119b5d4acce903



適用例

α -quartz (SiO_2)

α -quartz (SiO_2) 2x2x2スーパーセル 300K NVT

第一原理計算：VASP

機械学習ポテンシャル：artificial neural networks (ANN)

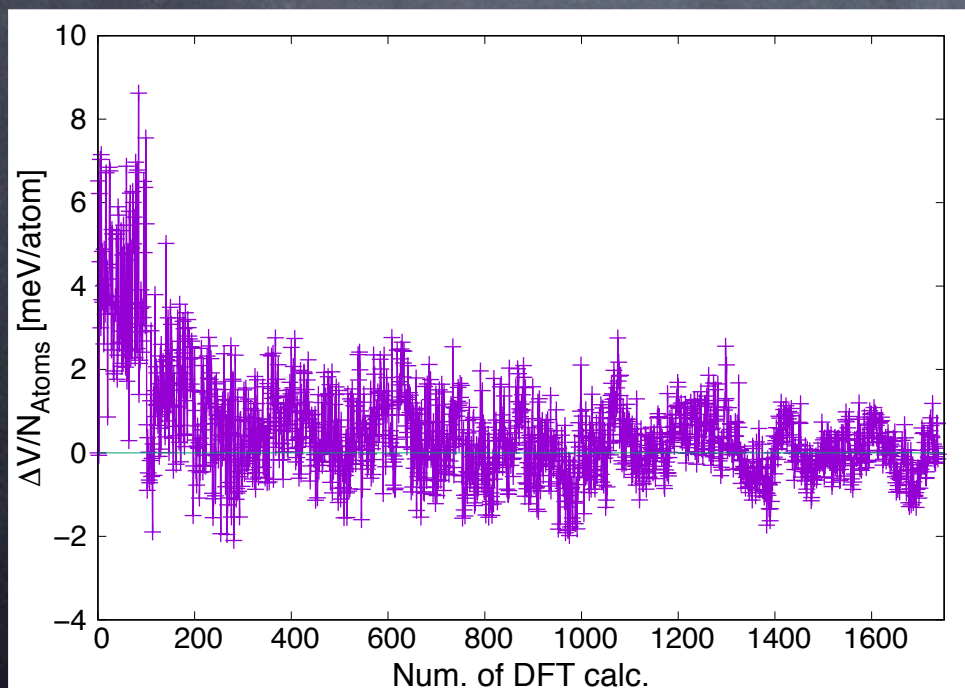
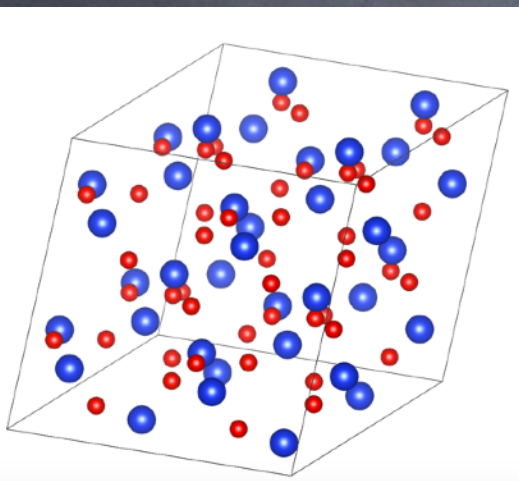
The Atomic Energy Network (***ænet***)

ニューラルネットワークは標準的な構造(15ユニットの隠れ層2層)

ハイブリッドモンテカルロ法：

PIMD: An Open Source Software for Parallel Molecular Simulation

<https://ccse.jaea.go.jp/ja/download/pimd/index.en.html>



1. 第一原理MDで300構造集めて初期ANNを作成
2. 0.25fsステップで機械学習MD
3. 5fsごとに第一原理計算してアクセプト率計算し、学習データ(構造とエネルギー)も収集
4. 同時に物理量の計算も行う
5. 500fsごとにANNをアップデート

**ANNポテンシャルの精度が改善されていく
→1 meV/atom以下に**

α -quartz (SiO_2)

α -quartz (SiO_2) 2x2x2スーパーセル 300K NVT

第一原理計算：VASP

機械学習ポテンシャル：artificial neural networks (ANN)

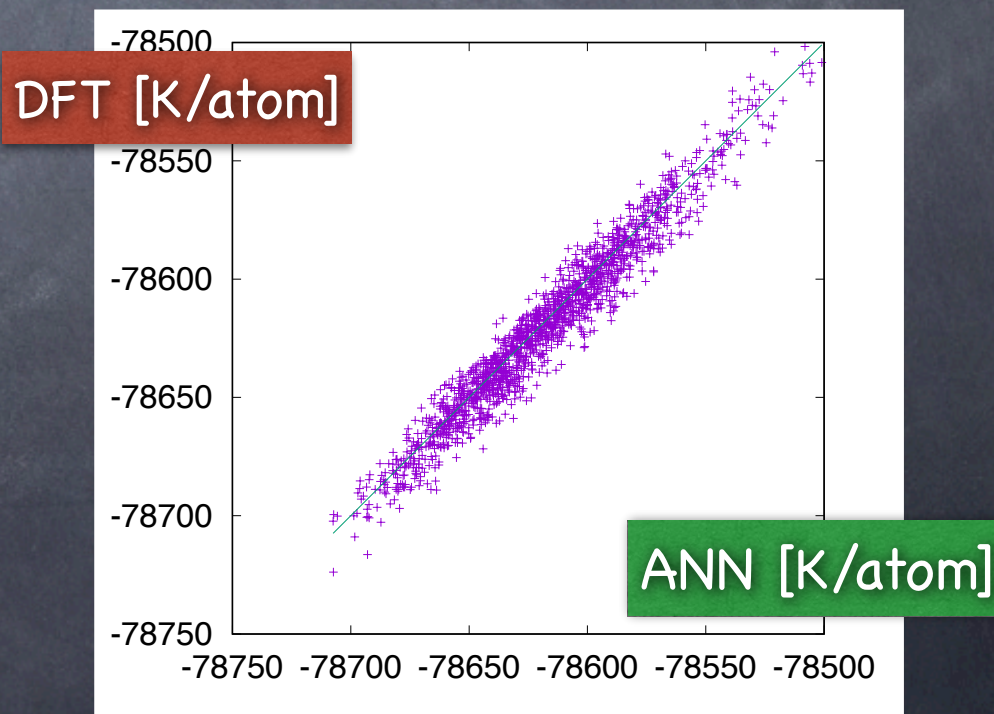
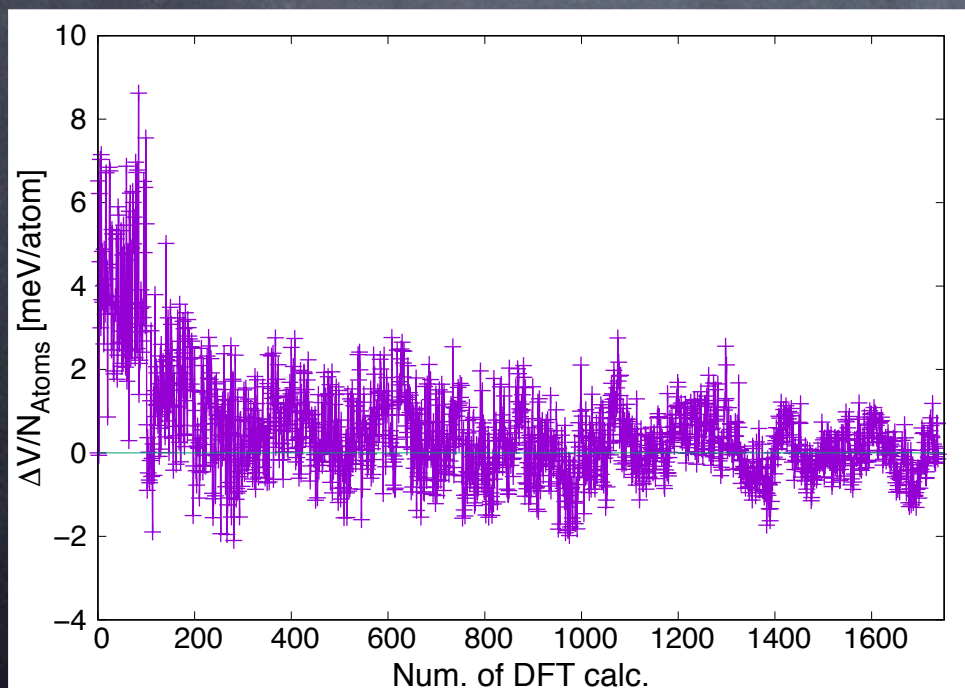
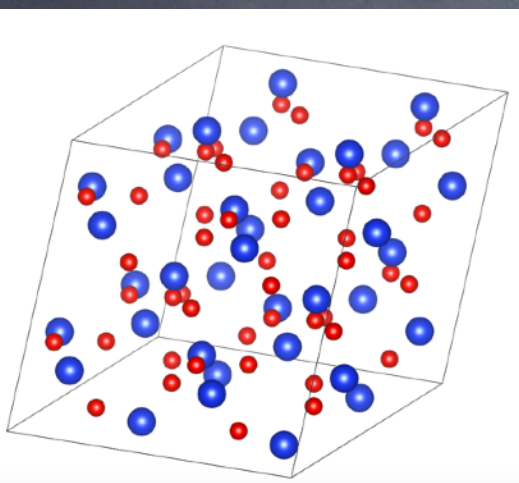
The Atomic Energy Network (***æ*net**)

ニューラルネットワークは標準的な構造(15ユニットの隠れ層2層)

ハイブリッドモンテカルロ法：

PIMD: An Open Source Software for Parallel Molecular Simulation

<https://ccse.jaea.go.jp/ja/download/pimd/index.en.html>



α -quartz (SiO_2)

α -quartz (SiO_2) 2x2x2スーパーセル 300K NVT

第一原理計算：VASP

機械学習ポテンシャル：artificial neural networks (ANN)

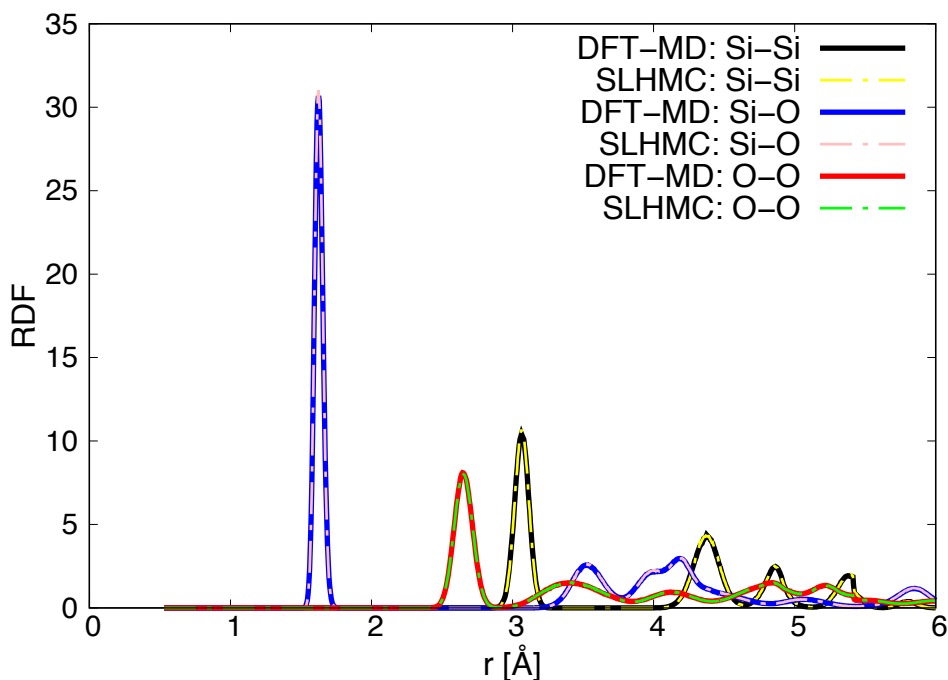
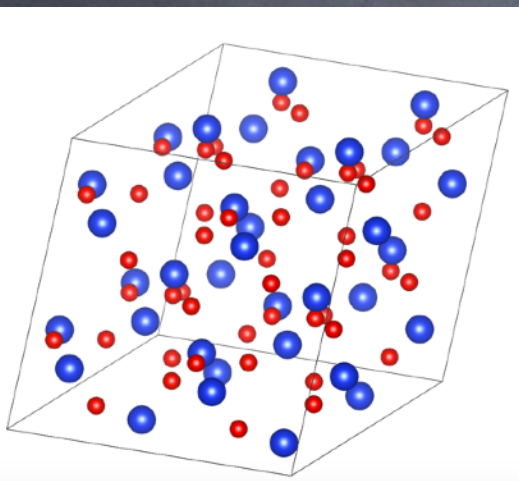
The Atomic Energy Network (***ænet***)

ニューラルネットワークは標準的な構造(15ユニットの隠れ層2層)

ハイブリッドモンテカルロ法：

PIMD: An Open Source Software for Parallel Molecular Simulation

<https://ccse.jaea.go.jp/ja/download/pimd/index.en.html>



動径分布関数（原子の位置の分布）

第一原理MDの結果を完全に再現している

DFTのエネルギーをボルツマン重みとした
メトロポリスヘイスティングス法

機械学習ポテンシャルは配位提案に使う

->第一原理HMCと結果が厳密に等しい

注：トレーニングしながら計算が可能

α -quartz (SiO_2)

α -quartz (SiO_2) 2x2x2スーパーセル 300K NVT

第一原理計算: VASP

機械学習ポテンシャル: artificial neural networks (ANN)

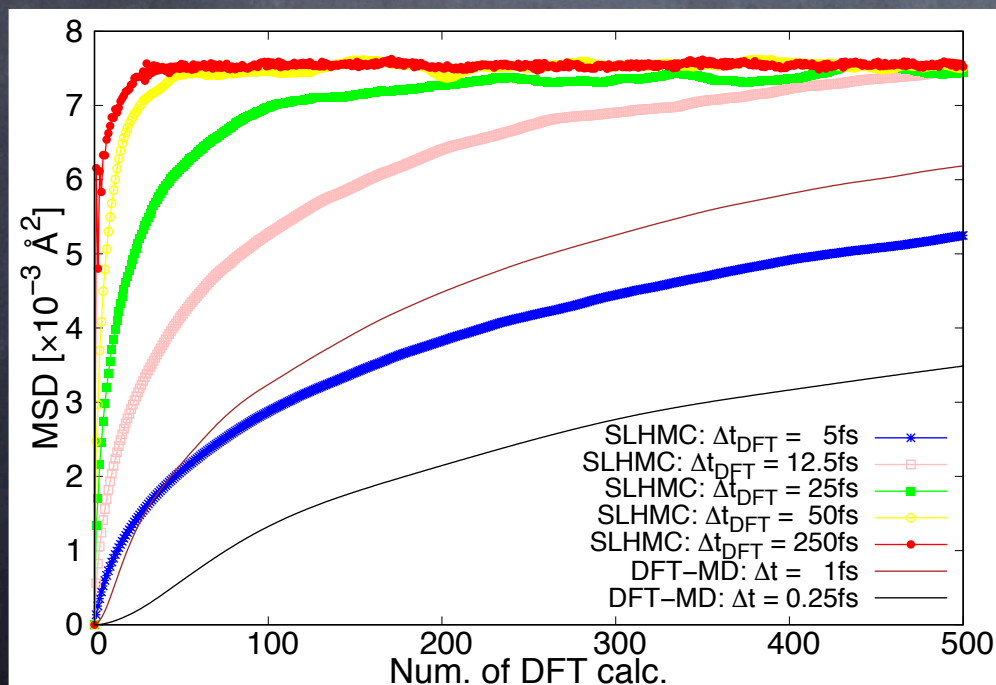
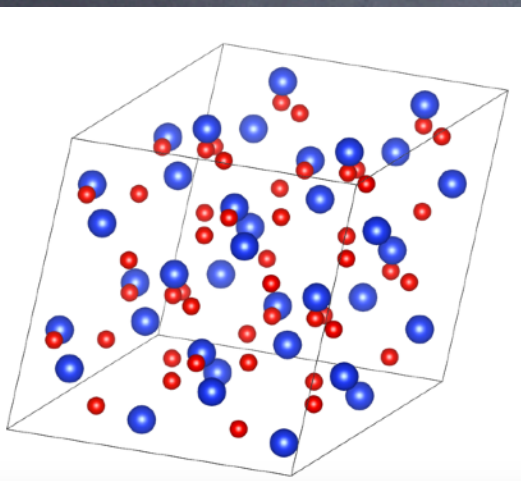
The Atomic Energy Network (***ænet***)

ニューラルネットワークは標準的な構造(15ユニットの隠れ層2層)

ハイブリッドモンテカルロ法:

PIMD: An Open Source Software for Parallel Molecular Simulation

<https://ccse.jaea.go.jp/ja/download/pimd/index.en.html>



平均自乗変位: 位置相関

(固体では一定値に収束する)

機械学習MDの長さを変更

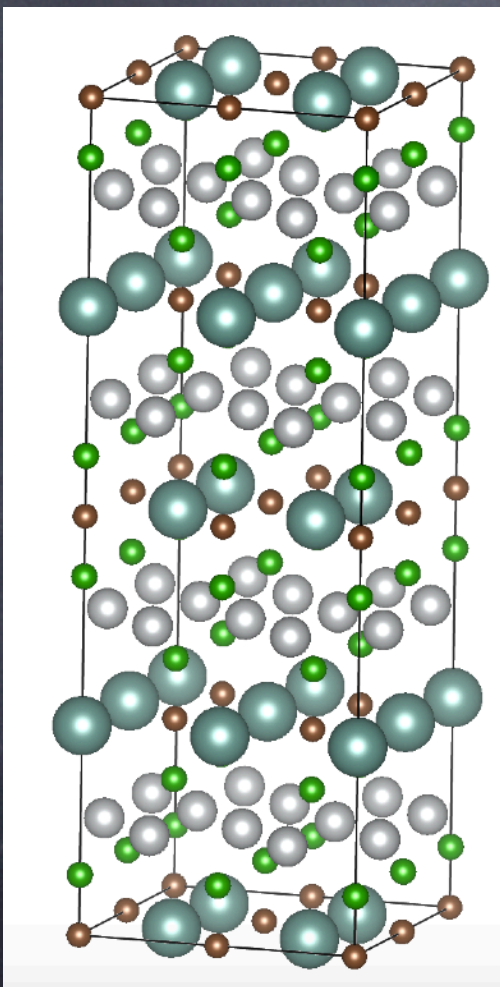


DFT計算の回数を減らすことができる

高速化

YNi₂B₂C

超伝導体($T_c \sim 15K$) フォノン起源の非従来型超伝導体(強い異方性をもつ)



2x2x2

中性子散乱実験：フォノンDOSに強い温度依存性

->非調和振動のフォノンが重要な系

第一原理計算でフォノンの温度依存性を出すには？

->第一原理分子動力学法

フォノンを計算するには長時間シミュレーションが必要

->第一原理MDでは厳しい

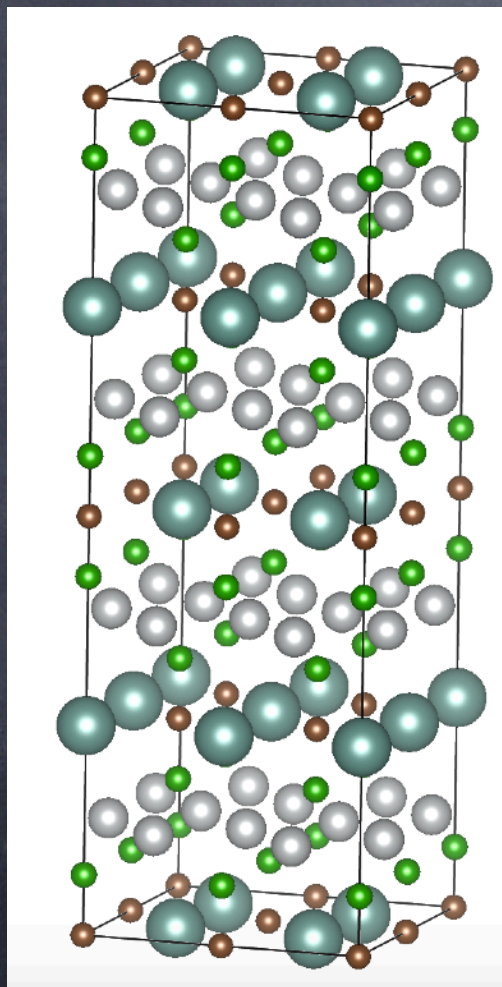
機械学習MDを使えばフォノンDOSを計算できる？

->4元素系->安定な機械学習MDは難しい

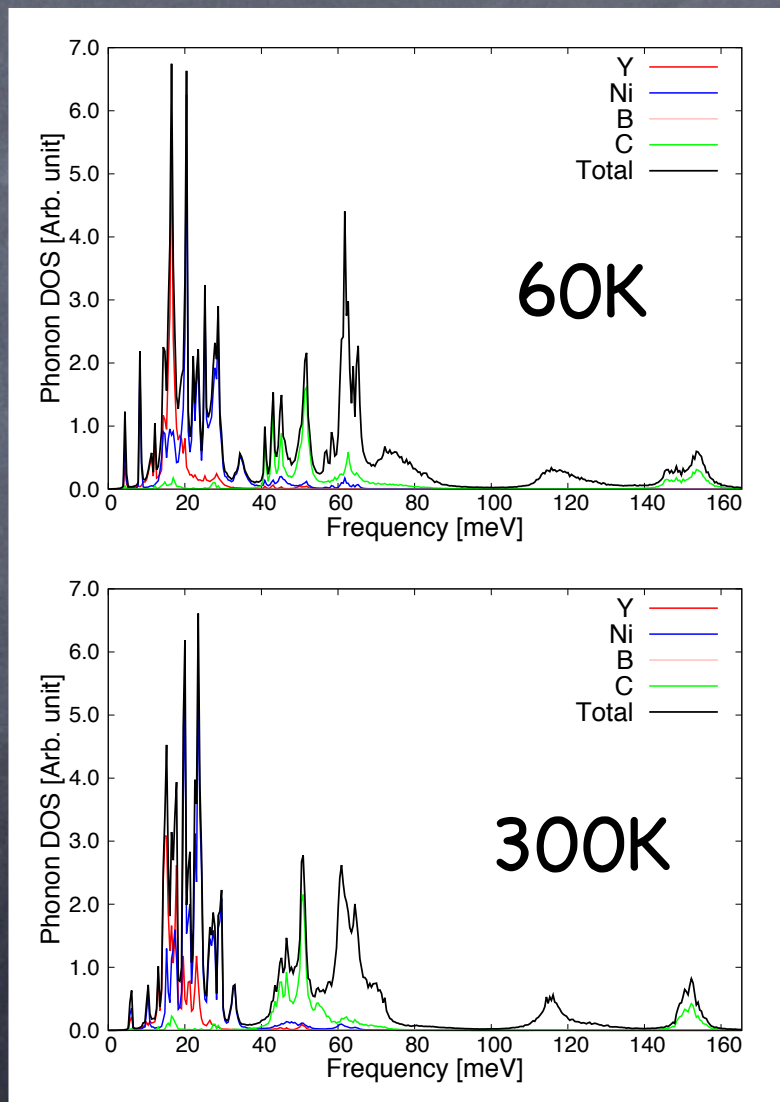
**SLHMCで「良いポテンシャル」を作成し、
機械学習MDをやってみる**

YNi₂B₂C

超伝導体($T_c \sim 15K$) フォノン起源の非従来型超伝導体(強い異方性をもつ)



2x2x2



1000KでSLHMC
well-trained ANNを作成

1step 1fsでANN-MD
10万ステップ→100ps
10回繰り返し、計算

フォノンDOSが強い
温度依存性を持つ

通常のDFT(Phononのバンド分散
を計算してDOS)では計算できない
効果

まとめ

まとめ

自己学習ハイブリッドモンテカルロ法

1. 機械学習ポテンシャルを利用して、第一原理MDと統計的に**厳密**に等しい結果を得ることができる



ポテンシャルはon-the-flyで逐次的に改善されていく

トレーニングしながら計算が可能
DFT計算の回数を減らすことができる

機械学習ポテンシャルは配位提案に使う
->第一原理HMCと結果が厳密に等しい

高速化

2. 学習で得られた機械学習ポテンシャルは、ターゲットMDからサンプルされたもの->well-trained なポテンシャル

得られたポテンシャルで長時間安定にML-MDができる