人工ニューラルネットワーク原子間相互作用ポテンシャルの分子動力学法への応用と課題

島村，孝平
下條，冬樹
田中，成典

(Citation)
日本神経回路学会誌,26(4):145-155

(Issue Date)
2019-12-05

(Resource Type)
journal article

(Version)
Version of Record

日本神経回路学会誌　Vol. 26, No. 4（2019），145–155

解　　説

人工ニューラルネットワーク原子間相互作用ポテンシャルの分子動力学法への応用と課題

島　村　孝　平[*1]，下　條　冬　樹[*2]，田　中　成　典[*1]

神戸大学大学院システム情報学研究科　計算科学専攻　計算生物学講座[*1]，
熊本大学大学院先端科学研究部　基礎科学部門　物理科学分野[*2]

Recent Progress and Current Issues in Development of Artificial Neural Network Interatomic Potential for Molecular Dynamics Simulation

Kohei Shimamura,[*1] Fuyuki Shimojo[*2] and Shigenori Tanaka[*1]

Graduate School of System Informatics, Kobe University[*1]
Department of Physics, Kumamoto University[*2]

概要

分子動力学（MD）法は，原子毎に立てられたNewtonの運動方程式を逐次的に解くことで系全体の原子ダイナミクスを追跡できる計算機シミュレーション手法であり，ミクロな現象の解明に役立つことから材料分野や生物分野では標準的な手法として認識されている．近年，人工ニューラルネットワーク（ANN）の万能近似性を活用して，従来のMD法が抱えていた精度と計算コストの難点を克服するANN原子間相互作用ポテンシャル（ANN potential）の開発が活発に行われ新局面を迎えている．本稿では，ANN potentialの基本的なアルゴリズムについて，応用例を交えながら，現在直面している回帰学習のデータ不均衡問題などの解決すべき課題について述べる．

1. は　じ　め　に

人工ニューラルネットワーク（Aritificial Neural Network: ANN）が持つ万能近似性の応用の一つとして，分子動力学（Molecular Dynamics: MD）法への応用がある．MD法は，系に含まれる原子一つ一つに対して立てられたNewtonの運動方程式（下式(1)，後述する）を短いtimestep Δtで数値的に積分しこれを繰り返すことにより，原子系全体のダイナミクスを追跡する計算機シミュレーション手法である．Newton方程式を解くだけでは系の全エネルギーが一定のMDシミュレーションとなるが，温度や圧力が主たる環境設定である実際の実験条件には合わせ辛い．だが，統計力学と解析力学の理論を組み合わせることにより温度一定[1)]や温度及び圧力一定[2)]のMDシミュレーションが可能になり様々に使いやすい方向へ拡張されている．これらの高い拡張性に加えて原子レベルのミクロな現象を解析することができることから，材料学や生物学分野などでは基本的な手法であり，これまで幅広く使われてきた．

MD法では，次式(1)に含まれるN個の原子座標（位置ベクトルを$\mathbf{R}^N$で表す）間に働く相互作用potential $U(\mathbf{R}^N)$をどのように設定するかに精度の全てがかかっている．

$$m_i \frac{d^2\mathbf{R}_i}{dt^2} = \mathbf{F}_i = -\frac{dU(\mathbf{R}^N)}{d\mathbf{R}_i} \tag{1}$$

ここでm_i, $\mathbf{R}_i$, $\mathbf{F}_i$はそれぞれ原子iの質量，位置ベクトル及び原子iに働く力を意味する．$U(\mathbf{R}^N)$の与え方は大きく分けて2通りあり，1つは密度汎関数理論（Density Functional Theory: DFT）等の量子論に基づく第一原理MD（First-Principles MD: FPMD）法であり，もう1つは研究者の物理化学的な知識に基づき妥当だと思われた関数形で$U(\mathbf{R}^N)$をモデル化した古典MD（Classical MD: CMD）法である．両者の間には精度と計算コストにトレードオフの関係がある．すなわち，FPMD法は高精度の保証があるものの計算コ

[*1] 〒 657–8501　兵庫県神戸市灘区六甲台町 1–1
[*2] 〒 860–8555　熊本県熊本市中央区黒髪 2–39–1

ストが N 原子系に対して DFT の場合 $O(N^3)$ に比例するため，計算規模及び計算時間に大きな制約がある．一方 CMD 法は $O(N)$ を実現するが，精度は $U(\mathbf{R}^N)$ のモデル化がどれほど正確であるかに依存する．

これら 2 種の MD 法が存在する中で，機械学習のようなデータ駆動型手法を駆使し，FPMD 法と CMD 法の長所を融合させた手法が近年登場してきている．それらは万能近似性を活かして FPMD 法が作る複雑な potential $U^{\mathrm{FPMD}}(\mathbf{R}^N)$ を学習し精度を保ちながらも，単なる関数形へ落とし込み $O(N)$ という極めて低い計算コストの MD 計算を達成することを目的にしている．現在 ANN を利用した ANN potential とガウス過程を利用した Gaussian Approximation Potential (GAP)[3,4] の開発が主として活発に行われている．本稿では，前者の ANN potential について基本的なアルゴリズムと適用例，そして現在抱えている課題について述べたい．以下 ANN potential を用いた MD を ANN-MD と呼称する．適用例としては，主として我々が最近構築した Ag_2Se[5] という熱電材料として使われる物質の ANN potential を紹介したい[6]．以降の各手法の説明だけでは読者は捉え辛いと思われるので，この適用例を交えながら解説していく形式を採ることとする．

我々の開発動機もここで少し述べる．Ag_2Se は相転移温度 $T_c = 406\,\mathrm{K}$ で低温 β 相（β-Ag_2Se，図 1(a)）から高温 α 相（α-Ag_2Se，図 1(b)）に構造相転移することが知られている．β-Ag_2Se は直方晶の結晶である

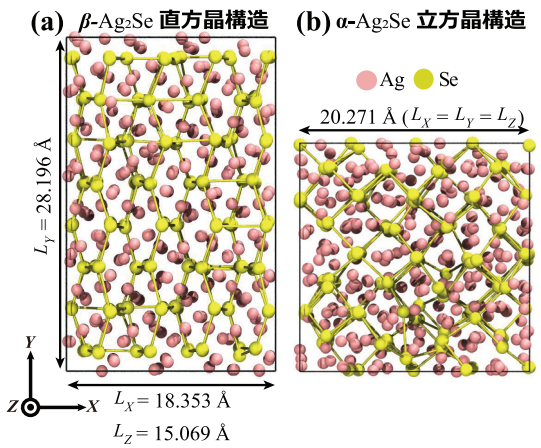

図 1　384 原子から構成される (a) 直方晶 β-及び (b) 立方晶 α-Ag_2Se の原子構造．我々の研究[6]では後者を MD 計算系として用いている．また黒枠は周期的境界条件の境界である．

が，α-Ag_2Se は Ag 原子が液体様に振る舞う立方晶の超イオン伝導体であり，系のダイナミクスが大きく異なる．また，構造相転移付近での異常な比熱や熱伝導度の上昇が実験的に報告されており[7,8]，その原子機構の解明に興味が持たれている．だが，これらの物理量自体の計算には 10 万 timestep 以上の高精度な MD 計算が必要であり[9]，既存の FPMD 及び CMD 手法では達成できず，それ故に ANN potential に可能性を求めた．超イオン伝導体の ANN potential の作成報告がそもそも無かったこともあり，まずは α-Ag_2Se に対する ANN potential の構築を試みている．加えて，本研究では FPMD 及び ANN-MD 計算用 package として我々が開発した QXMD[10] を用いている．また ANN potential の学習用 package として Artrith らが開発した Aenet[11] を拡張して用いている．

2. ANN potential

この章では，一般的な ANN potential の構築方法[12]を述べる．FeedForward Neural Network（FFNN）を用いたものが最もシンプルな ANN potential であろう．FFNN では入力層から出力層への前方方向への情報伝搬のみを許す．一つの入力層と出力層の間に一つ以上の隠れ層が設けられ，各層のノードは隣接する層との間のみ重みパラメータ w によって接続されている．層 l のノード j の持つ値 y_j^l は次のように与えられる．

$$y_j^l = f_j^l \left(B_j^l + \sum_i w_{l,j}^{k,l} \cdot y_i^k \right) \tag{2}$$

ここで $w_{l,j}^{k,l}$ は層 k に含まれるノード i から層 l のノード j の重みパラメータであり，B_j^l はバイアスである．加えて，f は非線形な活性化関数を意味する．hyperbolic tangent（tanh）を始め多くの活性化関数が用いられている．例えば，我々の研究では tanh よりも精度が良かった modified tanh[13] を用いた．ただし，**4** 章で述べるように，深層学習等で多用される活性化関数である Rectified Linear Unit（ReLU）[14] は使えない．

ANN potential における FFNN の運用方法には 2 つの特徴があり，1 つは基本的には Multiple Neural Network（MNN）の形で使われることである．これは FFNN の入力に用いる記述子を元素毎に用意する必要性があるため，目標とする物理系を構築する元素毎に FFNN が用意されるためである．例えば Ag_2Se に対する ANN potential を作成する場合には Ag と Se 元素についてそれぞれの FFNN を構築した．記述子については **3** 章で述べる．

もう 1 つは現在の深層学習との対比であるが，入力

層と出力層間の隠れ層の数が少ない点である．1 層以上の隠れ層が用いられるけれども，ANN potential を MD 計算での運用することを踏まえると，数を増やし過ぎると計算時間が長くなってしまう．それゆえ先行研究を調べた限りでは，隠れ層を 2 層程度に設定することが最も多いようである（ノード数は 20 個程度である）．しかし例外も存在し Single NN（SNN）を用いた先行研究であるが，5 層の隠れ層を設定している例がある（ノード数は最大で 240 個使われている）[15]．

3. Symmetry Functions

ここでは代表的な記述子である Symmetry Functions（SFs）[12] を紹介する．下の式 (3) 及び (4) で定義される M 個の SFs に，各原子 i からカットオフ距離 R_c 以内の球内に存在する原子 j の座標を代入して，数値化された $\{G_{i,\alpha}\}$ $(\alpha \in 1,2,\cdots,M)$ により原子 i の周囲原子構造を特徴付けることを目的としている．（α を省略した $\{G_{i,\alpha}\}=\{\mathbf{G}_i\}$ というベクトル表記も以降で用いる．）M は後述のように，研究者によって適切な数が指定されなければならない．SFs による特徴抽出の基本的な考え方は，Convolutional NN（CNN）で行われる畳み込みによる特徴抽出や，Chemoinformatics における molecular fingerprint に通ずるものがある．具体的には，下の動径方向 G_i^{rad} と角度方向 G_i^{ang} の 2 種類の SFs が使われる．

$$G_i^{\mathrm{rad}} = \sum_j e^{-\eta\left(R_{ij}-R_c\right)^2} \cdot f_c\left(R_{ij}\right) \tag{3}$$

$$G_i^{\mathrm{ang}} = \sum_{j,k\neq i} \left(1+\lambda\cos\theta_{ijk}\right)^{\zeta} \cdot e^{-\eta\left(R_{ij}^2+R_{ik}^2+R_{jk}^2\right)} \cdot f_c(R_{ij})\cdot f_c(R_{ik})\cdot f_c(R_{jk}) \tag{4}$$

ここで $R_{ij}=|\mathbf{R}_{ij}|$ であり，原子 i と j 間の距離を意味する．θ_{ijk} は $\mathbf{R}_{ij}$ と $\mathbf{R}_{ik}$ の角度である．η，R_s，λ，ζ はハイパーパラメータであり，研究者によって対象の物理系に対して適切な値が選択される．f_c は次式で定義されるカットオフ関数であり，カットオフ距離 R_c で滑らかに 0 に $G_{i,\alpha}$ を収束させる．

$$f_c\left(R_{ij}\right) = \begin{cases} 0.5\left[\cos\left(\frac{\pi R_{ij}}{R_c}\right)\right] & \left(R_{ij}\leq R_c\right) \\ 0 & \left(R_{ij}\geq R_c\right) \end{cases} \tag{5}$$

ただし SFs は，動径方向 G_i^{rad} であれば特定の 2 元素毎に（例えば Ag-Ag や Ag-Se の組み合わせ），角度方向 G_i^{ang} であれば特定の 3 元素毎に（例えば Ag-Se-Se や Se-Ag-Ag の組み合わせ）定義される．SFs は並進及び回転対称性を有するように設計されており，これらの対称性操作によって数値が変わることが無いため，symmetry functions の名前がある．研究者は周囲原子構造の特徴を掴めるように元素の組み合わせとハイパーパラメータが設定された M 個の SFs を予め定義する．例えば我々が行った α-Ag_2Se に対する ANN potential の構築[6] に際しては，Ag と Se 元素に M=56 個の SFs をそれぞれを定義し，その内訳は G_i^{rad} を 20 個，G_i^{ang} を 36 個である．

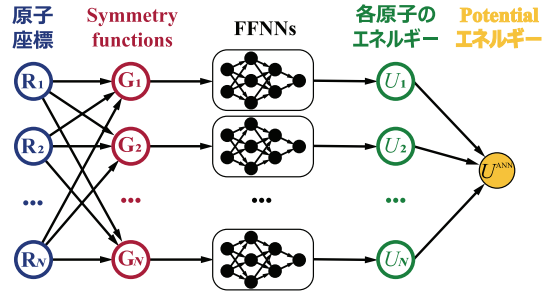

図 2　SFs を用いた ANN potential の模式図．

図 2 は SFs を用いた ANN potential の模式図を示している．系に含まれる原子の座標 $\{\mathbf{R}_i\}$ を SFs に代入して $\{\mathbf{G}_i\}$ を得て，FFNN の入力とする．ただしこのとき，学習の精度を高めるために $\{\mathbf{G}_i\}$ について α 毎に標準化の前処理が必要である．FFNN の出力を U_i として，系に含まれる全原子数 N_{atom} 分繰り返し，最終的に全系の potential エネルギーの予測値 U^{ANN} を U_i の和として得る．

$$U^{\mathrm{ANN}} = \sum_i^{N_{\mathrm{atom}}} U_i \tag{6}$$

U_i は各原子の持つ potential エネルギーとみなせる．U^{ANN} は全ての原子座標 $\{\mathbf{R}_i\}$ の関数であるので，これを式 (1) に代入して力 $\{\mathbf{F}_i\}$ を計算すれば MD 計算が可能になる．また，このように構築すると原子数 N_{atom} が異なる系に対しても同じ ANN potential を用いることができる利点がある．

問題は計算対象の系をうまく表現できる適切な SFs をどのようにして選ぶかである．理論的に必要な SFs を決定することは難しく，ヒューリスティックな手法に頼っているのが現状である．ただ最近は情報理論の技術を駆使して，例えば最初に様々な元素の組み合わせと異なるハイパーパラメータを入力し多数の SFs を作っておき（この網羅的発生はヒューリスティックである），CUR 分解等の低ランク近似法を用いて削減することが行われている[4,16]．SF 数の増加は MD 計算の精度を向上させるが計算時間とトレードオフになるので，より数が少なく表現力の高い記述子が望まれて

いる．故に，SF 以外の記述子も様々提案されてきており[17]，主に元素毎に FFNN を設けた MNN 用ではあるものの，中には SNN での運用を想定した記述子も存在する[15]．

4. 学 習 方 法

ANN potential の学習データは，一般的に数百から数万 timestep 分の FPMD 計算データが用いられる．その 1 timestep 分のデータには，系の potential エネルギー U^{FPMD}，原子に働く力ベクトル $\{\mathbf{F}_i^{\mathrm{FPMD}}\}$，圧力テンソル $\{\mathrm{P}_{\mu\nu}^{\mathrm{FPMD}}\}$ $(\mu,\nu\in\{1,2,3\})$，そして原子座標ベクトル $\{\mathbf{R}_i\}$ が含まれる．学習により最小化するコスト関数 $\mathcal{L}$ は，従来は potential エネルギー U に対する損失関数（U^{FPMD} と U^{ANN} の二乗和誤差）のみで構築されていた（下式 (7) 右辺第一項）．

$$\begin{aligned}\mathcal{L} = &\frac{p_E}{2}\frac{1}{N_{\mathrm{data}}}\sum_{I}^{N_{\mathrm{data}}}\left(U_I^{\mathrm{FPMD}}-U_I^{\mathrm{ANN}}\right)^2\\ &+\frac{p_F}{2}\frac{1}{N_{\mathrm{data}}}\sum_{I}^{N_{\mathrm{data}}}\frac{1}{3N_{\mathrm{atom}}}\sum_{i}^{N_{\mathrm{atom}}}\left(\mathbf{F}_{I,i}^{\mathrm{ANN}}-\mathbf{F}_{I,i}^{\mathrm{FPMD}}\right)^2\\ &+\frac{p_P}{2}\frac{1}{N_{\mathrm{data}}}\sum_{I}^{N_{\mathrm{data}}}\frac{1}{6}\sum_{j}^{6}\left(\mathrm{P}_{I,j}^{\mathrm{ANN}}-\mathrm{P}_{I,j}^{\mathrm{FPMD}}\right)^2 \quad (7)\end{aligned}$$

ここで N_{data} は FPMD 学習データ数である．また 3×3 のテンソルである $\{\mathrm{P}_{\mu\nu}\}$ の独立成分は 6 個であり，ここでは P_j のように j を使って通し番号表記している．しかし，MD 計算では力や圧力の精度も重要であるため，最近ではこれらの損失関数（式 (7) 右辺第二及び三項）も加えて学習を行った研究が増えてくるようになった[15,18]．だが，$\mathbf{F}_i^{\mathrm{ANN}}$ と $\mathrm{P}_{\mu\nu}^{\mathrm{ANN}}$ は次式で与えられるように U^{ANN} とは微分積分の関係にあり，単純なマルチタスク学習ではない．後述のように ReLU が使えないのは，その困難さの例だとみなせる．

$$\begin{aligned}\mathbf{F}_i^{\mathrm{ANN}} &= -\frac{dU^{\mathrm{ANN}}}{d\mathbf{R}_i}\\ &= -\frac{d}{d\mathbf{R}_i}\sum_{i}^{N_{\mathrm{atom}}}U_i \quad (8)\end{aligned}$$

$$\mathrm{P}_{\mu\nu}^{\mathrm{ANN}} = -\frac{1}{V}\sum_{i}^{N_{\mathrm{atom}}}R_{i,\mu}F_{i,\nu}^{\mathrm{ANN}} \quad (9)$$

式 (9) の圧力はビリアル定理に基づき算出され，V は MD 計算系 (MD セルと以下で呼称する）の体積である（図 1 を参照のこと）．

学習は，誤差逆伝搬法に基づいて各重みパラメータ w_m に対するコスト関数の微分値 $\{d\mathcal{L}/dw_m\}$ を求めた後，研究者が選択した最適化アルゴリズムにより w_m が更新される．これについては代表的な確率的勾配降下法である Adam 法[15]，2 次収束法である準 Newton 法[11]や，誤差勾配の事後分布が小さくなるように最適化を行う Kalman Filter[18]等が使われている．

3 つの損失関数の次元が異なる点は，U^{FPMD}，$\{\mathbf{F}_i^{\mathrm{FPMD}}\}$, $\{\mathrm{P}_{\mu\nu}^{\mathrm{FPMD}}\}$ を学習前に予め標準化や正規化によって無次元化しておくことで解決できる．しかし，損失関数間に生じている学習対象数の不均衡を解消して効率的に学習を行うための手法が必要である．U^{FPMD} の数は N_{data} 個であるが，F^{FPMD} は「$3\times N_{\mathrm{atom}}\times N_{\mathrm{data}}$」個，$P^{\mathrm{FPMD}}$ は「$6\times N_{\mathrm{data}}$」個もある．故に，最適化対象の物理量総数は「$N_{\mathrm{data}}(1+3N_{\mathrm{atom}}+6)$」個になる．式 (7) において各損失関数の係数 p_E，p_F，p_P はこの不均衡解消のために用意された一つの解決策である[15]．学習中にこれら 3 つの係数の値を変化させる意図で導入されたものであるものの[15]，提案された変化方法は物理化学的な知見に基づいて定義されたものではなく経験的な方法であるため使用には注意が必要である．

我々が行った α-Ag_2Se の場合では，よりシンプルな変化法として p_E と p_F の 2 つのみを調整する方法を採った[6]．式 (9) を踏まえると，圧力 P^{ANN} は系全体でただ一つ決まる示強性変数でありながら局所量の力の関数でもある．したがって，系全体の量である potential エネルギー U^{ANN} と局所量である力 F^{ANN} の 2 つがバランス良く学習されるのであれば，必然的に圧力の誤差も小さくなるはずであるため，p_P の調整を省略した．一方 p_E と p_F の調整では，$(p_E, p_F) = (1.0, 1.0)$ を初期値として選び，学習中 500 epoch 毎に p_F のみを 1/10 にする方法を採った．ここで最適化アルゴリズムは準 Newton 法[11]を採用している．力の学習対象数が potential エネルギーの対象数よりも圧倒的に多いことから，$(p_E, p_F) = (1.0, 1.0)$ と設定すると学習の焦点は力に偏る．その後，徐々に p_F を小さくすることで偏りを無くしていけばバランスの良い学習を達成できる．学習が potential エネルギーと力で均等に行われているかについては圧力の誤差により判断することができる．だが圧力誤差は学習中，常にモニターする必要があり，potential エネルギーと力の誤差の両方が小さくても圧力誤差が突然大きくなるといったことが起こる[6]．また，上記の方法では p_F の減少率の設定に任意性があるため，学習の自動化を行いたい場合にはまだ課題がある．

この章の最後に ReLU について少し述べる．誤差逆伝搬法により $\{d\mathcal{L}/dw_m\}$ を算出する過程で，FFNN

中の活性化関数に対して 2 階微分を実行する必要がある．このとき 2 階微分で値が常に 0 となる ReLU のような活性化関数を使うと誤差が逆伝搬されなくなってしまい学習ができない．式 (7) のコスト関数を最小化するために最適な活性化関数を開発することは一つの課題でもある．

5. 学習データの収集方法

この章では，ANN potential の関数形を形作るための要となる学習用 FPMD 計算データをどのように準備するべきかについて述べたい．研究者の目的とする物質や状態に応じて準備方法が異なることが予想されるため，より一般的な方法を確立するために試行錯誤が続けられているところでもある．最近は能動学習（AL）に位置づけられる機械学習手法を駆使して，必要な FPMD データを効率的に識別・蓄積しようとする研究が主流になってきている[18,19]．これらの研究では例えば以下の手順を経て ANN potential を構築する．

1. 特定の熱力学的条件設定（温度，圧力，原子密度等）の下で ANN-MD 計算を実行し，timestep 毎に何らかの方法によって出力誤差を出す．例えば potential エネルギーや力の誤差である．
2. 出力誤差が大きい timestep は，そのときの原子構造が未学習の可能性があることを意味する．したがってその timestep にラベルを付ける．
3. ラベルが付いた timestep における原子構造に対して同じ熱力学的条件設定下で FPMD 計算を実行して学習データを得て，ANN potential を再学習する．
4. 熱力学的条件設定を変更して 1. に戻る．

AL を用いれば幅広い熱力学的状態に渡って量子論の精度を有する ANN potential が構築できることが示されたため大きな成果と言えるだろう[18,19]．

ただ，必要な学習データを識別することはできるが，多様性のある学習データを生成することに関してはまだ課題があるように思える．これは ANN-MD 計算を実行する際に，研究者が設定した熱力学的条件の探索領域に依存するためである．MD 法では熱力学的条件の設定を，例えば温度を 50 K ずつ大きくするなど，基本的に離散的にしか行えないので考慮されない条件区間が存在する．これまでの AL を用いた研究では探索領域に関しては，補足資料に一覧が掲載される程度に留まり，それらの設定を選択した理由については重要視されていない．しかし，将来的により複雑な系に対する ANN potential の構築を行う際には，どのような学習データが物理化学的に必要不可欠であるのかを説明できなければならないと思われる．

そこで，我々は一つの熱力学的状態（ここでは温度 500 K，圧力 1 atm，原子数 384 の α-Ag_2Se（図 1(b)））を再現するために必要なデータが何であるかについて，物理化学的な観点から考察した[6]．結論から言えば，以下が必要であった．

1. 部分動径分布関数 $g_{\alpha\beta}(R)$ が収束するまでの 500 K での FPMD 計算データ（1,000 FPMD timestep 分）．
2. 原子間の反発力を補正するための「構造-MD セル最適化」FPMD 計算データ（25 FPMD timestep 分）．

次式で示される部分動径分布関数 $g_{\alpha\beta}(R)$ は，ある元素 α の原子から距離 R 離れた元素 β の原子の存在確率を表した，ミクロな原子構造を知るための基本的な物理量である．

$$g_{\alpha\beta}(R)=\frac{1}{4\pi R^2\Delta R}\frac{V}{N_\alpha N_\beta}\langle\sum_{i=1}^{N_\alpha}\Delta N_i^{\alpha\beta}(R)\rangle \quad (10)$$

N_α と N_β はそれぞれ系を構成する元素 α と元素 β の原子数，V は MD セルの体積である．$\Delta N_i^{\alpha\beta}(R)$ は注目している元素 α の原子 i を中心とした半径 $R-1/2\Delta R$ と $R+1/2\Delta R$ の球で囲まれた球殻内に存在する元素 β の原子数である．$g_{\alpha\beta}(R)$ を得ると各原子間の局所的な構造情報を詳細に得られたことになることから，我々は $g_{\alpha\beta}(R)$ の値が変動しなくなる timestep までの FPMD 計算データを ANN potential の学習データに用いることは物理化学的観点から妥当であると考えた．図 3 に示すように，1,000 timestep 分の FPMD 計算データがあれば $g_{\alpha\beta}(R)$ が収束したとみなせた．

だが，**4** 章で述べた方法を用いて学習し ANN-MD 計算を実行したものの（温度などの設定は FPMD 計算と同条件），途中で原子間の反発力が機能しなくなり原子間距離が 0 になる現象が発生するという非現実的な結果を得た．これは図 4(a) の保存量（赤線）で確認することができ，46.0 ps（ps は 10^{-12} 秒）付近で逸脱している．保存量は MD 計算が物理化学的に妥当である限り保存されるものであり[20]，計算のチェックに利用されている．この破綻の原因は，MD 中に極めて低い確率で原子間の距離が著しく近づくことにある．α-Ag_2Se の学習データ中の Ag-Ag，Ag-Se，Se-Se 原子間の最短距離は 2.5，2.3，3.1 Å（Å は 10^{-10} m）で

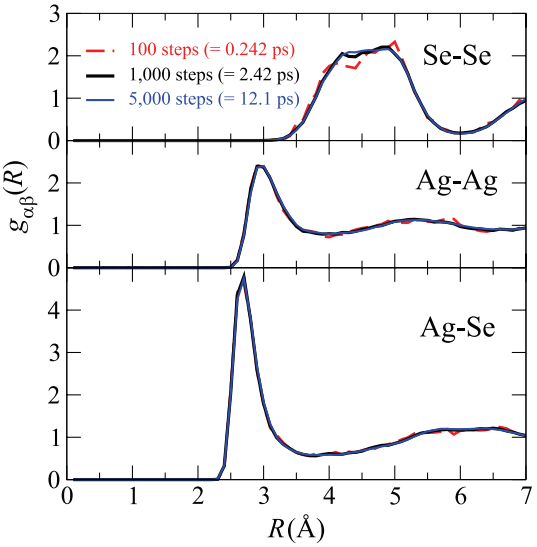

図3 100，1,000，5,000 timestep 分の FPMD 計算データにより算出した Ag-Ag，Ag-Se，Se-Se 間の $g_{\alpha\beta}(R)$.

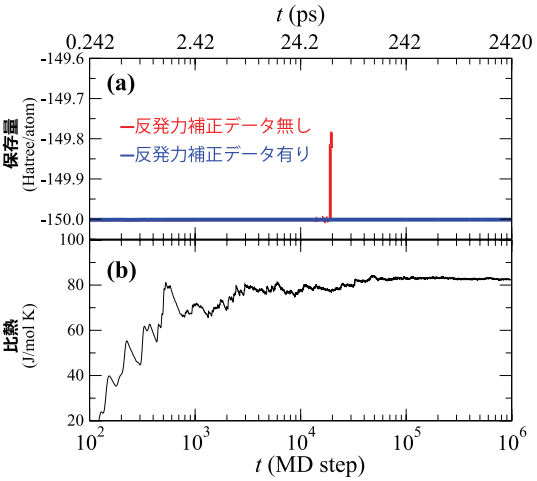

図4 (a) 反発力補正データが加えられた場合（青線）と無い場合（赤線）で学習された ANN potential による ANN-MD 計算中の保存量推移．(b) 反発力補正データ有りの ANN-MD 計算中で計算される比熱の推移．

あったが，その出現回数は 1,000 FPMD timestep 中にたったの 3 回程度であった．このようなマイノリティサンプルの学習が不十分であったために原子間の反発力が不完全に働き原子間距離が 0 になったと思われる．

上記のようなマイノリティデータが生じる要因は MD

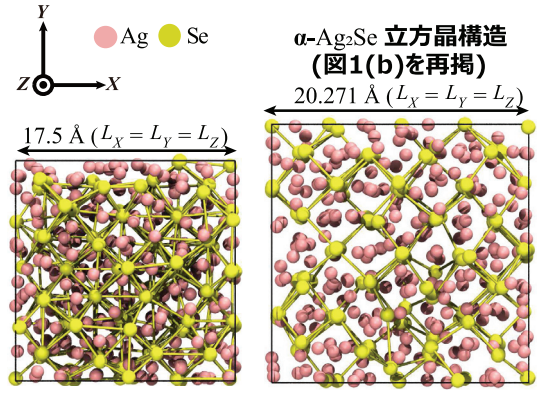

図5 図 1(b) の α-Ag_2Se 構造の MD セルサイズを等方的に一辺が 20.271 Å（右図）から 17.5 Å まで縮小した原子構造（左図）．また黒枠は周期的境界条件の境界である．

法のアルゴリズムから定性的な説明が可能である．我々が行った温度一定のアルゴリズム[1]に基づく 500 K の MD 計算では，全エネルギー E（運動エネルギーと potential エネルギーの和）を示す状態が出現する確率 $P(E)$ は以下のカノニカル分布に従う．

$$P(E) = \frac{1}{Z_{\mathrm{NVT}}} \mathrm{e}^{\frac{-E}{k_{\mathrm{B}}T}} \tag{11}$$

ここで T，k_{B}，Z_{NVT} は系の温度，ボルツマン定数，そしてカノニカル分布の分配関数である．これは統計力学からの要請であり，温度 T の熱力学的平衡状態にある系はこの分布に従う．本来原子間距離が極めて小さくなると原子核反発力が生じて E が著しく大きくなるため，$P(E)$ は指数関数的に小さくなる．それ故，この現象は（少なくとも FPMD 計算では）稀に起こるイベントであると言える．温度・圧力一定のアルゴリズム[2]に関しても同じ議論が可能である．

問題解決の糸口は，短距離の結合が豊富に存在する高密度の原子構造に対する FPMD データを oversampling することである．これは図 5 に示すように，オリジナルの α-Ag_2Se の MD セルを等方的に小さくすることで達成される．具体的にはマイノリティデータである原子間距離が 2.4，2.2，3.0 Å の Ag-Ag，Ag-Se，Se-Se 結合が 50 本以上存在するように MD セルの一辺の大きさを 17.5 Å まで縮小した．ただ，この原子構造だけでは縮小前の α-Ag_2Se よりも極めて高い potential エネルギーを有する極端な FPMD データを得るだけであるため，再学習させても逆に精度が落ちる可能性があった．そこで，「構造・MD セル最適化」と

呼ばれる系の potential エネルギーが小さくなる方向へ原子位置と MD セル自体の大きさを徐々に変化させるアルゴリズムを適用した[6]．すると，徐々に縮小前の原子構造及び MD セルサイズに戻っていく過程（図 5 の右図から左図へ戻って行く様子を想像して欲しい）の「FPMD データ」が得られるため，元の系との間のギャップを自然に埋めることができる．構造・MD セル最適化アルゴリズムは MD アルゴリズムとは厳密には異なり，Newton の運動方程式ではなく準 Newton 法を時間発展に利用している．しかし timestep 毎に得られるデータが同じであることから，これも FPMD 計算データと本稿では呼称する．構造・MD セル最適化により得られた FPMD データ（25 timestep 分）も学習データに加えると ANN-MD 計算の破綻が無くなり，少なくとも 2,420 ps 以上の ANN-MD 計算が可能になることが明らかになった[6]．図 4(a) に示すように保存量の破綻も見られない．目的であった比熱もこの時間スケールの ANN-MD 計算が可能になると図 4(b) に示すように収束し，実験値（83.6 J/molK[8]）と一致する 83.0 J/molK を示した．他にも $g_{\alpha\beta}(R)$ や拡散係数なども FPMD 計算結果や実験値と一致することを確認したため[6]，（少なくとも α-Ag_2Se の）一熱力学状態を再現する ANN potential に必要な学習データは上記の 1. 及び 2. であったと結論づけた．

また，学習データは上述した 1. 及び 2. の FPMD データだけであるが，思わぬ効果があり，相転移温度 $T_c = 406$ K 付近の比熱も再現することが分かった[6]．図 6 に示しているように 500 K に加えて 450 K, 410 K, 400 K の計算を行い，実験値と同様に比熱の上昇が見られている．400 K の ANN-MD 計算における原子構造を見ると 2,420 ps の間に，図 7 に示すように，構造変化を度々繰り返している．最終的に至った構造は直方晶の β-Ag_2Se（図 1(a)）に類似し（Ag 原子が拡散しているため β-Ag_2Se ではない），その時に示した比熱（89.0 J/molK）が実験値（89.3 J/molK[8]）と一致した．また，β-Ag_2Se，α-Ag_2Se の両方に属さない中間構造も一時的に出現しており比熱は 107.2 J/molK と最も高かった．**1** 章で述べたように構造相転移温度付近で比熱等に異常な上昇が見られるのは，このような比熱が高い構造が出現しているためではないかと考えている．

$g_{\alpha\beta}(R)$ の収束という基準のみで集めた FPMD 計算データに必要なデータは確かに含まれていたが数の不均衡が発生しており，物理化学的に不可欠な原子間反発力の再現に必要なデータ（原子間距離が極めて近いイベント）が極めて少なかった．構造・MD セル最適化

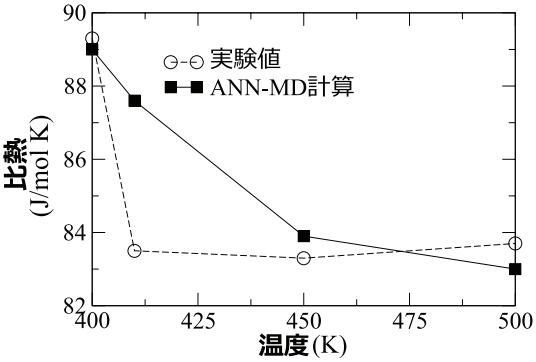

図 6　構造相転移温度 406 K 付近までの ANN-MD 計算と実験[8]の比熱の変化．

アルゴリズムにより高密度の原子構造に対する FPMD 計算データを生成すれば，マイノリティデータを oversampling することに対応するため，その結果に一つの熱力学的状態を再現する ANN potential の構築につながったと思われる．ただ，目的とした状態以外にも相転移温度付近での比熱の実験値との一致や構造変化が見られることを踏まえると，単に oversampling だけでなく他の物理化学的に重要なデータも一緒に得られた可能性がある．例えば，構造・MD セル最適化アルゴリズムの実行により得られた FPMD 計算データには，極めて小さくなった原子間距離を元の自然長に戻していく過程のデータが含まれているはずである．つまりこのデータを学習させるとその ANN potential は原子間距離を自然長に戻す方法を獲得すると考えられる．したがってこれから検証を続けていくことが必要であるが，$g_{\alpha\beta}(R)$ の収束を基準に「大雑把に」FPMD データを蓄積しつつその一方で，高密度原子構造からの緩和過程を発生させ，そこで AL を用いれば，さらに効率的に必須な FPMD 計算データを得られるのではないかと考えている．

ただし，**6** 章で少し述べるように，系が分子で構成される場合には高密度領域だけでなく低密度領域も生成する必要があると思われる．また，現在の AL における誤差評価法にも少し問題がある．現在主流であるのは，重みパラメータの初期値は異なるが同じデータを学習させた複数の ANN potential によって出力の分散を算出する方法である[18,19]．もう一つの万能近似性を使った potential 作成法である GAP[3,4]ではベイズ理論に基づいて誤差算出を実現していることを踏まえると，やはり ANN potential でも理論に基づいて誤差評価を考えたいところである．例えば Monte Carlo

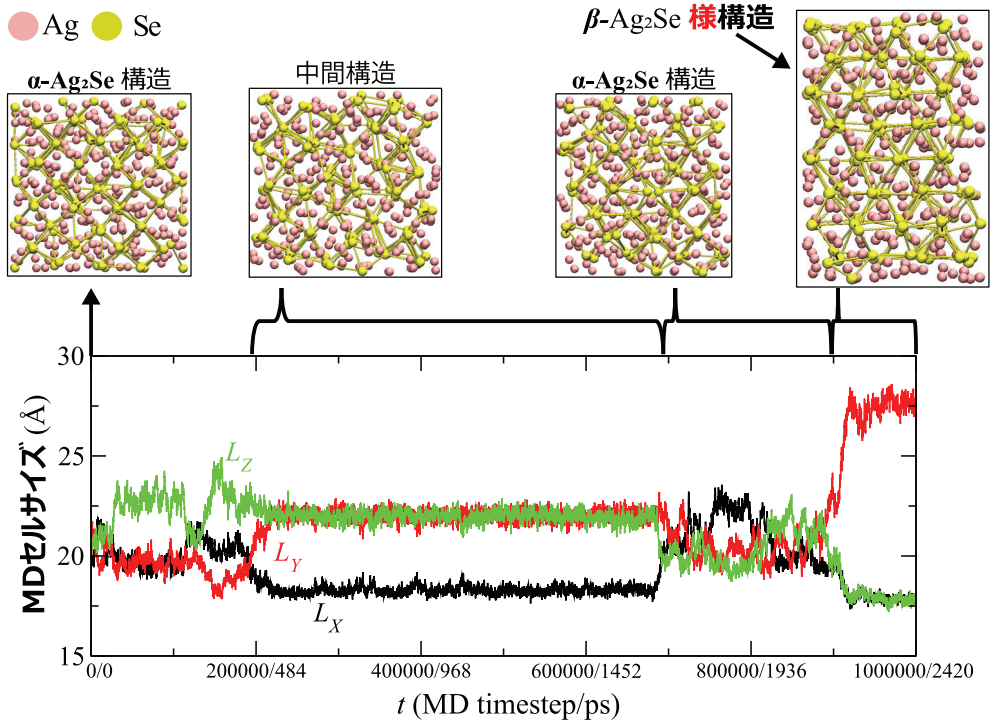

図7　400 K の ANN-MD 計算における MD セルの各辺の長さ L_x，L_y，L_z の変化と特徴的な原子構造のスナップショット．

dropout 法[21]は実装の容易さや計算コストが低い点などを考慮すると現実的なベイズ推定法であるかもしれない．

上記の AL を用いた方法以外にも，熱平衡状態では全エネルギー E を持つ系の出現確率が式 (11) のように分かっていることから，モンテカルロ法の利用も大変有効であると思われる．最近ではマルコフ連鎖モンテカルロ（MCMC）法の流れを汲んだ自己学習モンテカルロ（SLMC）法と呼ばれる方法により，高速に重要な原子構造をサンプリングする方法が提案されており，ANN potential の構築に使われている[22]．通常 MCMC 法では研究者が予め設定した提案確率分布に従って原子構造を生成し，詳細つり合い条件が満たされた場合のみ，目的とする確率分布に属する原子構造と判断してサンプルを獲得する．提案確率分布に依存して採択率は非常に小さくなる場合がある．一方，SLMC 法は採択率が高くなるような提案確率分布を学習により獲得することを可能とし，少ないサンプリングでも目的確率分布に属する十分な数の原子構造を得られると見込まれる．式 (11) のような確率分布に従わない非平衡状態（化学反応や構造相転移過程等）に対しても ANN potential を構築したい場合には SLMC 法だけでは困難であるが，AL と組み合わせることで構築の可能性を検討して行けるのではないかと考えている．

6. その他の課題

詳細は文献に譲ることにし，簡潔に 3 つほど述べたい．まずは，分子系に対する ANN potential 構築についてである．前章とは逆に，原子間の距離が広がるときにマイノリティデータを生むことがあるのは分子系を取り扱う場合である．例えば想像し易いように簡単なケースとして 2 つの球をバネでつないだ図 8(a) のような物体を考える．これは H_2 のような 2 原子分子の運動を表す良いモデルである．バネ運動が行われているとき，2 球間の距離は自然長から離れ，最短になる最小短縮点と最長になる最大伸長点が存在する．この 2 点において potential エネルギーが最大になり，加えて 2 点は同じエネルギーを持つ．前章で oversampling の対象になったのが前者の最小短縮点に対応するが，分子にはもう一つ最大伸長点という oversampling が必要と思われる状態が存在することになる．この oversampling には低密度構造を生成して構造・MD セル

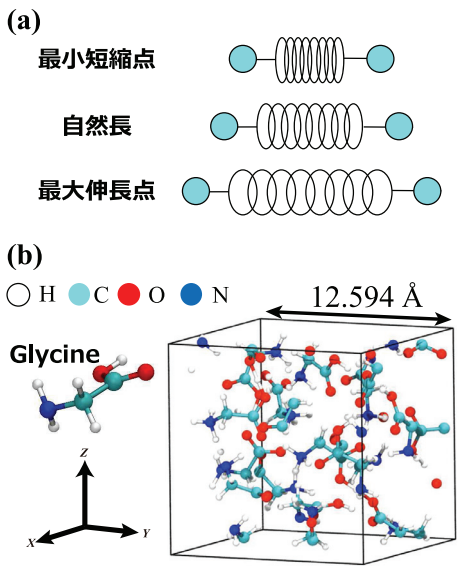

図 8　(a) バネ運動における最小短縮点，自然長，最大伸長点．(b) 立方体 MD セル中の 20 分子から構成された glycine アモルファス系．

最適化アルゴリズムなどにより緩和過程の FPMD 計算データを得る方法が考えられる．実際我々の試みでは最小短縮点を oversampling する FPMD 計算データ（準備方法は前章と同様）に加えて，最大伸長点を oversampling する低密度の FPMD 計算データを学習データに追加することで，常温常圧（300 K，1 atm）下で 100 万 ANN-MD timestep（= 242 ps）規模の計算が可能な glycine アモルファス系（図 8(b)）に対する ANN potential の構築が可能であることが分かっている．

次に，非一様系への ANN potential 構築方法の確立についてである．例えば Ag_2Se に少数の硫黄原子を不純物として添加し，熱電材料としての機能を改善するといったことがより実用段階では行われている．この時，硫黄原子のデータは Ag や Se 原子に比べて少ないため，データ数の不均衡が生じる．硫黄原子のダイナミクスに関するデータは SLMC 法により効率的にサンプリングが可能かもしれないが，一方で，既存の枠組みに囚われない工夫を行っている研究もある[23]．そこでは不純物の有無で 2 通りの FPMD 計算データを用意し，それらの「potential エネルギー差」を学習することで少数派の原子ダイナミクスを効率的に学習することを達成している（学習手法自体は図 2 と同じである）．構築技術を複雑化せずに発想次第で ANN potential の精度向上を達成した例として参考にするべき研究である．

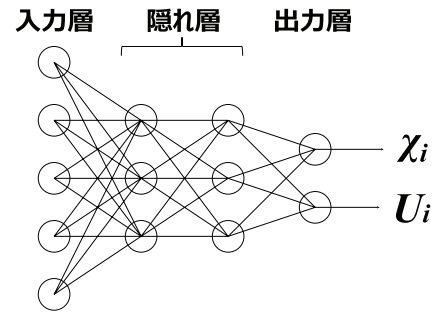

図 9　短距離相互作用 potential エネルギー U_i と電気陰性度 χ_i を出力する FFNN の例．

最後に，静電相互作用の取り扱いについてである．**3** 章で述べたように ANN potential で使われる記述子は基本的にはカットオフ距離 R_c 内に存在する原子構造を特徴化しているため，物理化学的には $U^{\mathrm{FPMD}}(\mathbf{R}^N)$ の短距離原子間相互作用のみを模倣できることになる．電荷間に働く静電相互作用のような長距離相互作用が考慮されていない理由は，電荷遮蔽効果が十分働く系を取り扱っているか，もしくはカットオフ距離 R_c を十分長く取ることによって対処できるためである．しかしカットオフ距離の延長は計算コストの増大を招くため限界がある．静電相互作用の影響が大きいイオン性物質や化学反応により電荷移動が頻繁に起こる非平衡過程を取り扱いたい場合には問題になるだろう．解決の糸口として考えているのが，イオン性物質に対する ANN potential 構築方法として CENT potential と名付けられた手法である[24]．本手法では，FFNN の個々の出力値（式 (6) で言えば U_i に対応するもの）を各原子の potential の持つエネルギーではなく，各原子の持つ電気陰性度 χ_i とみなす（記述子は SFs が用いられている）．電気陰性度は原子がマイナス電荷を引き付ける強度を意味するため原子電荷の変化を支配するものであるが，この物理量自体は周囲の近接原子構造にのみ影響を受けると考えられるため SFs でも記述できる．次に電荷平衡法と呼ばれる電気陰性度から電荷及び系の potential エネルギーを求める反復法を用いることで，イオン性物質を記述することに成功している．だが，この CENT potential の枠組みを用いると逆に静電相互作用以外を記述できなくなる．そこで例えば，記述子は SFs を採用したまま図 9 のように FFNN の出力層のノードを 2 値にして片方を電気陰性度 χ_i の出力，もう片方を短距離相互作用 potential エネルギー U_i とすることが考えられる．このような形で CENT potential との融合した枠組みを考えていく

ことで，化学反応過程をも記述できる ANN potential の構築が実現できるかもしれない．

7. おわりに

本稿では ANN が持つ万能近似性の応用の一つとして，MD 法に用いるための ANN potential を構築する研究を挙げ，その構築方法について我々の最近の研究[6]を引き合いに出しながら，現状抱える課題も併せて述べた．

ANN potential の精度向上につながる方法が徐々に明らかになってきており，有効な記述子（**3** 章），コスト関数（**4** 章），学習データを生成・識別する方法（**5** 章）が提案されている．その一方で，これらの方法を最適化していくことも必要である．特に下に挙げるようなデータ数の不均衡問題が多く見られ，現在は物理化学的な観点から効果的な方法を検討しているところである．

1. コスト関数を構成する 3 つの損失関数は学習対象数に大きな差がある（**4** 章）．
2. バネ運動に例えると最小短縮点と最大伸長点に対応する物理化学的に不可欠な状態に対する学習データが極めて少ない場合がある（**6** 章）．
3. 非一様系の場合，系に含まれる原子数の違いにより少数派の原子ダイナミクスの学習が困難になると予想される（**6** 章）．

回帰学習におけるデータ不均衡問題は少し前から機械学習分野でも取り上げられるようになってきており解決手法の提案もある[25]．このようなデータ不均衡問題も含めて情報科学的観点と物理化学的な観点の両側面から有効な手法を探り出し，一つ一つの課題解決を試みていくべきだと考える．

最近は ANN potential の学習用 package が充実してきており，コードも公開されているものも多い．Aenet[11]，AMP[26]，DeepMD[15]，n2p2[18]，TensorMol[27] などがある．これらの様々なプログラミング言語で書かれた package は，ANN potential や MD 法に触れたことが無くても自身の手法を試しに導入してみる契機を与えるかもしれない．それ故，名前だけであるが最後にご紹介しておきたい．

謝辞

本稿執筆の機会をいただいた松原崇博士に深く感謝する．また本稿で紹介した研究は，MEXT/JSPS 科研費（16K05478，17H06353，18K03825，19K14676）及び JST CREST（JPMJCR18I2）の助成を受けて行われている．加えて計算機資源として，東京大学物性研究所及び九州大学情報基盤研究開発センター所有のスーパーコンピューターを使わせていただいたので，ここで感謝申し上げる．

参考文献

1) Nosé, S. (1984): A molecular dynamics method for simulations in the canonical ensemble, Molecular Physics, Vol.52, pp.255–268.
2) Martyna, G.J., Tobias, D.J., Klein, M.L. (1994): Constant pressure molecular dynamics algorithms, The Journal of Chemical Physics, Vol.101, pp.4177–4189.
3) Bartók, A.P., Payne, M.C., Kondor, R., Csányi, G. (2010): Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons, Physical Review Letters, Vol.104, Article Number 136403.
4) Jinnouchi, R., Karsai, F., Kresse, G. (2019): On-the-fly machine learning force field generation: Application to melting points, Physical Review B, Vol.100, Article Number 014105.
5) Ding, Y., Qiu, Y., Cai, K., Yao, Q., Chen, S., Chen, L., He, J. (2019): High performance n-type Ag_2Se film on nylon membrane for flexible thermoelectric power generator, Nature Communications, Vol.10, Article Number 841.
6) Shimamura, K., Fukushima, S., Koura, A., Shimojo, F., Misawa, M., Kalia, R.K., Nakano, A., Vashishta, P., Matsubara, T., Tanaka, S. (2019): Guidelines for creating artificial neural network empirical interatomic potential from first-principles molecular dynamics data under specific conditions and its application to α-Ag_2Se, The Journal of Chemical Physics, Vol.151, Article Number 124303.
7) Aliev, S.A., Aliev, F.F. (2008): Effect of fluctuations on electron and phonon processes and thermodynamic parameters of Ag_2Te and Ag_2Se in the region of phase transition, Semiconductors, Vol.42, pp.394–400.
8) Grønvold, F., Stølen, S., Semenov, Y. (2003): Heat capacity and thermodynamic properties of silver(I) selenide, oP-Ag_2Se from 300 to 406 K and of cI-Ag_2Se from 406 to 900 K: Transitional behavior and formation properties, Thermochimica Acta, Vol.399, pp.213–224.
9) Chen, J., Zhang, G., Li, B. (2010): How to improve the accuracy of equilibrium molecular

dynamics for computation of thermal conductivity?, Physics Letters A, Vol.374, pp.2392–2396.

10) Shimojo, F., Fukushima, S., Kumazoe, H., Misawa, M., Ohmura, S., Rajak, P., Shimamura, K., Bassman, L., Tiwari, S., Kalia, R.K., Nakano, A., Vashishta, P. (2019): QXMD: An open-source program for nonadiabatic quantum molecular dynamics, SoftwareX, Vol.10, Article Number 100307.
11) Artrith, N., Urban, A. (2016): An implementation of artificial neural-network potentials for atomistic materials simulations: Performance for TiO_2, Computational Materials Science, Vol.114, pp.135–150.
12) Behler, J. (2011): Atom-centered symmetry functions for constructing high-dimensional neural network potentials, The Journal of Chemical Physics, Vol.134, Article Number 074106.
13) LeCun, Y.A., Bottou, L., Orr, G.B., Müller, K. (2012): Efficient backprop, Neural Networks: Tricks of the Trade: Second Edition, pp.9–48 (Springer Berlin Heidelberg).
14) Glorot, X., Bordes, A., Bengio, Y. (2011): Deep sparse rectifier neural networks, proceedings of the fourteenth international conference on artificial intelligence and statistics, PMLR, Vol.15, pp.315–323.
15) Zhang, L., Han, J., Wang, H., Car, R., E, W. (2018): Deep potential molecular dynamics: A scalable model with the accuracy of quantum mechanics, Physical Review Letters, Vol.120, Article Number 143001.
16) Imbalzano, G., Anelli, A., Giofré, D., Klees, S., Behler, J., Ceriotti, M. (2018): Automatic selection of atomic fingerprints and reference configurations for machine-learning potentials, The Journal of Chemical Physics, Vol.148, Article Number 241730.
17) Reveil, M., Clancy, P. (2018): Classification of spatially resolved molecular fingerprints for machine learning applications and development of a codebase for their implementation, Molecular Systems Design & Engineering, Vol.3, pp.431–441.
18) Singraber, A., Morawietz, T., Behler, J., Dellago, C. (2019): Parallel multistream training of high-dimensional neural network potentials, The Journal of Chemical Theory and Computation, Vol.15, pp.3075–3092.
19) Zhang, L., Lin, D., Wang, H., Car, R., E, W. (2019): Active learning of uniformly accurate interatomic potentials for materials simulation, Physical Review Materials, Vol.3, Article Number 023804.
20) Martyna, G.J., Tuckerman, M.E., Tobias, D.J., Klein, M.L. (1996): Explicit reversible integrators for extended systems dynamics, Molecular Physics, Vol.87, pp.1117–1157.
21) Gal, Y., Ghahramani, Z. (2016): Dropout as a Bayesian approximation: Representing model uncertainty in deep learning, proceedings of the 33rd international conference on machine learning, PMLR, Vol.48, pp.1050–1059.
22) Nagai, Y., Okumura, M., Kobayashi. K., Shiga, M. (2019): Self-learning hybrid Monte Carlo: A first-principles approach, arXiv preprint, arXiv:1909.02255.
23) Li, W., Ando, Y., Watanabe, S. (2017): Cu diffusion in amorphous Ta_2O_5 studied with a simplified neural network potential, Journal of the Physical Society of Japan, Vol.86, Article Number 104004.
24) Faraji, S., Ghasemi, S.A., Rostami, S., Rasoulkhani, R., Schaefer, B., Goedecker, S., Amsler, M. (2017): High accuracy and transferability of a neural network potential through charge equilibration for calcium fluoride, Journal of the Physical Society of Japan, Vol.95, Article Number 104105.
25) Krawczyk, B. (2016): Learning from imbalanced data: Open challenges and future directions, Progress in Artificial Intelligence, Vol.5, pp.221–232.
26) Khorshidi, A., Peterson, A.A. (2016): Amp: A modular approach to machine learning in atomistic simulations, Computer Physics Communications, Vol.207, pp.310–324.
27) Yao, K., Herr, J.E., Toth, D.W., Mckintyre, R., Parkhill, J. (2018): The TensorMol–0.1 model chemistry: A neural network augmented with long-range physics, Chemical Science, Vol.9, pp. 2261–2269.