

23 合同セッションN「インフォマティクス応用」 | 一般セッション(口頭講演): 23.1 合同セッションN「インフォマティクス応用」

## [7p-S202-1~12] 23.1 合同セッションN「インフォマティクス応用」

[7p-S202-1]

所望の結晶構造を有する多結晶TiO<sub>2</sub>薄膜の自律的な合成

○長谷部 颯汰<sup>1</sup>、西尾 和記<sup>5</sup>、相場 諒<sup>2</sup>、鈴木 陽太<sup>3</sup>、中山 亮<sup>1</sup>、簾 智仁<sup>1</sup>、小林 成<sup>1</sup>、清水 亮太<sup>4</sup>、一杉 太郎<sup>1,5</sup> (1.東大院理、2.株式会社リガク、3.科学大CFC、4.分子研、5.科学大物質理工)

[7p-S202-2]

材料構造類似度マップの材料プロセス探索への応用

○橋本 佑介<sup>1</sup>、Xue Jia<sup>2</sup>、Hao Li<sup>2</sup>、筈居 高明<sup>1</sup> (1.東北大学際研、2.東北大AIMR)

[7p-S202-3]

GaN気相成長における既存条件を考慮した目的関数による信頼性の高い成膜条件最適化

○関 翔太<sup>1,2</sup>、水野 文彬<sup>1</sup>、松岡 毅<sup>1</sup>、園田 勉<sup>3</sup>、高橋 言緒<sup>3</sup>、山田 永<sup>3</sup>、阿澄 玲子<sup>3</sup>、沓掛 健太郎<sup>2</sup>、宇治原 徹<sup>2</sup> (1.アイクリスタル、2.名大院工、3.産総研)

[7p-S202-4]

強化学習を用いたデータ駆動型自動FZ結晶成長炉の開発と実証実験

○柳沢 壮吾<sup>1</sup>、松本 遼平<sup>2</sup>、炭谷 翔悟<sup>2</sup>、稲垣 拓弥<sup>3</sup>、坂内 英典<sup>4</sup>、沓掛 健太郎<sup>1,5</sup>、宇治原 徹<sup>1,5</sup>、原田 俊太<sup>1,5</sup> (1.名大院工、2.アナモルフォーシスネットワークス、3.三幸、4.物質・材料研究機構、5.名大未来研)

[7p-S202-5]

半導体製造多段工程における誤差蓄積構造に着目した機械学習モデル設計

○山田 拓実<sup>1</sup>、沓掛 健太郎<sup>1,2</sup>、原田 俊太<sup>1,2</sup>、宇治原 徹<sup>1,2</sup> (1.名大院工、2.名大未来研)

[7p-S202-6]

Jacobianに着目した正則化による物理系の状態時間発展予測に関する深層学習の口バスト化

○関 翔太<sup>1,2</sup>、中西 佑児<sup>1</sup>、高石 将輝<sup>1</sup>、沓掛 健太郎<sup>2</sup> (1.アイクリスタル、2.名大院工)

[7p-S202-7]

深層学習を活用した半導体加工形状の自動評価パイプライン

○野島 聖耀<sup>1</sup>、岩満 一功<sup>1</sup>、赤瀬 善太郎<sup>1</sup>、菊池 昭彦<sup>2</sup>、富谷 茂隆<sup>1</sup> (1.奈良先端大、2.上智大)

[7p-S202-8]

三次元アトムプローブと透過電子顕微鏡法のバイモーダル相関顕微手法の開発

○赤瀬 善太郎<sup>1</sup>、別所 泰成<sup>1</sup>、大竹 義人<sup>1</sup>、埋橋 淳<sup>2</sup>、岩満 一功<sup>1</sup>、中島 宏<sup>3</sup>、山崎 順<sup>3</sup>、大久保 忠勝<sup>2</sup>、富谷 茂隆<sup>1</sup> (1.奈良先端大、2.物質・材料研究機構、3.大阪大学)

[7p-S202-9]

深層学習エリプソメトリー：光学定数・薄膜構造・バンドギャップの超高速・高精度解析

岡 亮輔<sup>1</sup>、山本 祐輝<sup>1</sup>、Hilfiker James N.<sup>2</sup>、反保 衆志<sup>3</sup>、永井 武彦<sup>3</sup>、林 真弘<sup>1</sup>、○藤原 裕之<sup>1</sup> (1.岐阜大、2.J.A. Woollam Co.、3.産総研)

[7p-S202-10]

拡張型自由エネルギーモデルを用いたInGaAs-HEMTのデバイス動作に対するゲート長依存性の研究

○中村 晴登<sup>1</sup>、町田 陽太郎<sup>2</sup>、長岡 竜之輔<sup>2</sup>、渡邊 一世<sup>3</sup>、小嗣 真人<sup>2</sup>、吹留 博一<sup>1</sup> (1.東北大工、2.東京理科大、3.情報通信研究機構)

---

[7p-S202-11]

カルマンフィルターによるデータフュージョンの測定系最適化

○竹内 靖<sup>1</sup> (1.アイクリスタル(株))

---

[7p-S202-12]

社会キャピタル理論に基づく画期的研究の発生メカニズム解明

○品川 啓介<sup>1</sup> (1.福岡女子大)

---

## 所望の結晶構造を有する多結晶 $\text{TiO}_2$ 薄膜の自律的な合成

### Autonomous synthesis of polycrystalline $\text{TiO}_2$ thin films with desired crystal structure

東大院理<sup>1</sup>, 科学大物質理工<sup>2</sup>, 株式会社リガク<sup>3</sup>, 科学大 CFC<sup>4</sup>, 分子研<sup>5</sup>

○(M2)長谷部 颯汰<sup>1</sup>, 西尾 和記<sup>2</sup>, 相場 諒<sup>3</sup>, 鈴木 陽太<sup>4</sup>, 中山 亮<sup>1</sup>, 簾 智仁<sup>1</sup>, 小林 成<sup>1</sup>, 清水 亮太<sup>5</sup>, 一杉 太郎<sup>1,2</sup>

Univ. Tokyo<sup>1</sup>, Science Tokyo<sup>2</sup>, Rigaku Corp.<sup>3</sup>, CFC<sup>4</sup>, IMS<sup>5</sup>, ○

Sota Hasebe<sup>1</sup>, Kazunori Nishio<sup>2</sup>, Akira Aiba<sup>3</sup>, Yota Suzuki<sup>4</sup>, Ryo Nakayama<sup>1</sup>, Tomohito Sudare<sup>1</sup>, Shigeru Kobayashi<sup>1</sup>, Ryota Shimizu<sup>5</sup>, Taro Hitosugi<sup>1,2</sup>

E-mail: hasebe-s@g.ecc.u-tokyo.jp

【緒言】: 機械学習とロボットを活用し、実験を「自律的」に行う時代が到来した。そのような中、我々は薄膜新材料を探索するために、X線回折 (XRD)、ラマン分光、紫外可視分光、走査型電子顕微鏡、電気抵抗など様々な計測・分析装置を接続した自動・自律実験システム (dLab, Fig.1) を構築した<sup>1)</sup>。

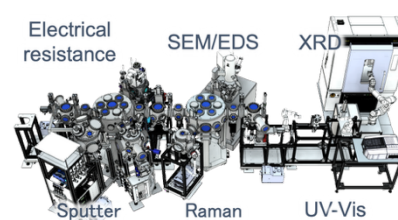


Figure 1. dLab system

中でも、XRD は得られた物質の結晶構造を精密に決定する重要な手法である。しかし、薄膜材料では回折ピークの数に限られることから、粉末材料のように自動的に構造決定することが困難である。そこで本研究では、粉末 XRD パターンと比較して相関係数を算出し、それをベイズ最適化の目的変数として自律合成を行った。本発表では多結晶  $\text{TiO}_2$  薄膜に着目し、ルチル型構造とアナターゼ型構造を自律的に作り分けることに成功した。

【実験方法】: 本研究では、 $\text{TiO}_2$  薄膜 (膜厚 ~220 nm) を RF マグネトロンスパッタ法でガラス基板上に作製した (基板温度: 200–600°C, 圧力: 1.0 Pa, Ar: 12.6 sccm,  $\text{O}_2$ : 7.4 sccm, RF 出力: 80 W)。この実験で、ルチル型構造  $\text{TiO}_2$  薄膜を得ることを目的にした。薄膜試料に対して斜入射 XRD 測定を行った。得られた XRD パターンのバックグラウンドとノイズを除去した後、ICSD に登録されているルチル型構造  $\text{TiO}_2$  粉末 XRD パターンとの相関係数( $r$ )を算出した。この  $r$  を目的変数、基板温度 ( $T_s$ ) を説明変数として、ベイズ最適化による自律実験を実施した。

【結果および考察】: Figure 2 に自律実験の結果を示す。 $T_s \geq 380^\circ\text{C}$  で  $r$  が高い値 (~0.8) を示した。このとき、 $\text{TiO}_2$  薄膜は単相のルチル型構造だった。一方で、 $220^\circ\text{C} \leq T_s \leq 360^\circ\text{C}$  の範囲で  $r < 0.6$  を示した。このとき、ルチルとアナターゼが混在した。 $T_s \leq 220^\circ\text{C}$  のときに  $r = 0$  を示し、単相のアナターゼ型構造だった。以上より、XRD パターンにおける相関係数を指標として狙いのルチル型構造を有する  $\text{TiO}_2$  薄膜の自律的な合成に成功した。当日は、合成パラメータを二つ ( $T_s$  とポストアニール温度) に増やした場合におけるアナターゼ型構造  $\text{TiO}_2$  薄膜の自律的な合成を含め、詳細を報告する。

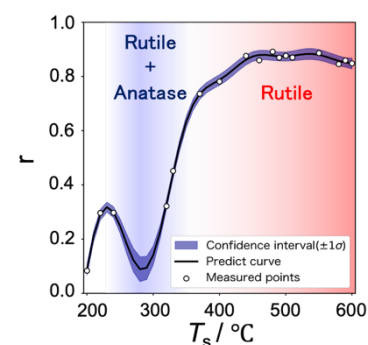


Figure 2. 自律的な条件探索

【参考文献】 K. Nishio, T. Hitosugi *et al.*, Digit. Discov. 2025, DOI: 10.1039/D4DD00326H.

## 材料構造類似度マップの材料プロセス探索への応用

### Application of a Material Structural Similarity Map to Materials Process Exploration

東北大学際研<sup>1</sup>, 東北大 AIMR<sup>2</sup>, <sup>○</sup>橋本佑介<sup>1</sup>, Xue Jia<sup>2</sup>, Hao Li<sup>2</sup>, 筈居高明<sup>1</sup>

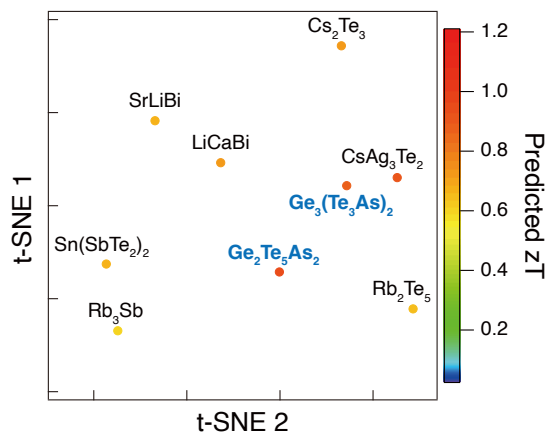
FRIS Tohoku Univ.<sup>1</sup>, AIMR Tohoku Univ.<sup>2</sup>, <sup>○</sup>Yusuke Hashimoto<sup>1</sup>, Xue Jiu<sup>2</sup>, Hao Li<sup>2</sup>, Takaaki Tomai<sup>1</sup>

E-mail: yusuke.hashimoto.b8@tohoku.ac.jp

材料科学における最大の課題の一つは、所与の特性を持つ最適材料の探索である。近年、データ科学と材料科学を融合したマテリアルズインフォマティクスが注目を集めており、蓄積されたビッグデータを活用して最適材料を提案する試みが活発化している。しかしながら、提案された材料の作製方法については、依然として研究者の経験と勘に頼る部分が多い。

我々はこれまでに、実験データと計算データの両方に基づく材料特性予測モデルの開発を進めてきた。特に、材料構造のグラフ表現と深層学習を組み合わせた材料特性機械学習フレームワーク「MatDeepLearn」を基に、材料物性と構造特性を入力として、それぞれの類似性を反映する材料マップの作成に成功している。材料科学は歴史的に、構造類似性を考慮して材料を分類し発展してきた。また、類似した構造を持つ材料は類似した手法やプロセスで作製される傾向がある。この知見に基づき本研究は、これまでに開発してきた材料マップをプロセス条件探索に活用する。

右図に、Materials Project から得られる結晶構造データと材料実験データベースである Starrydata2 から得られる材料実験データを機械学習を通して組み合わせ、Messenger Passing Neural Network と t-distributed Stochastic Neighbor Embedding (t-SNE)での次元削減に基づく MatDeepLearn フレームワークを使って作成した材料マップを示す。我々は機械学習による材料スクリーニングにより、 $\text{Ge}_3(\text{Te}_3\text{As})_2$  と  $\text{Ge}_2\text{Te}_5\text{As}_2$  を有望な新規熱電材料として提案した[1]が、その作成方法については不明だった。材料マップ中においてこれらの材料は近接しており、その近くに熱電材料としての研究が報告された  $\text{Sn}(\text{SbTe}_2)_2$  が見えている[2]。この研究を足がかりにすることで、未知の新材料でもその作製プロセス条件を効率的に探索できる。材料マップ上で近接する既知材料の合成条件を参考にすることで、効率的なプロセス探索が可能となる[3]。



[1] Xue Jia, Alex Aziz, Yusuke Hashimoto and Hao Li. Sci. China Mater. 67, 1173 (2024).

[2] Sora-at Tanusilp, *et al.*, Phys. Status Solidi RRL, 16: 2100482 (2022)

[3] Yusuke Hashimoto, Xue Jia, Li Hao, Takaaki Tomai, Submitted and arXiv:2503.07378



# GaN 気相成長における既存条件を考慮した目的関数による信頼性の高い成膜条件最適化

## Reliable optimization of MOCVD condition of GaN

### by objective function considering conventional conditions

アイクリスタル<sup>1</sup>, 名大院工<sup>2</sup>, 産総研<sup>3</sup>

○関 翔太<sup>1, 2</sup>, 水野 文彬<sup>1</sup>, 松岡 毅<sup>1</sup>, 園田 勉<sup>3</sup>,  
高橋 言緒<sup>3</sup>, 山田 永<sup>3</sup>, 阿澄 玲子<sup>3</sup>, 沓掛 健太郎<sup>2</sup>, 宇治原 徹<sup>2</sup>

Aixtal<sup>1</sup>, Nagoya Univ.<sup>2</sup>, AIST<sup>3</sup>

○Shota Seki<sup>1, 2</sup>, Fumiaki Mizuno<sup>1</sup>, Tsuyoshi Matsuoka<sup>1</sup>, Tsutomu Sonoda<sup>3</sup>, Tokio Takahashi<sup>3</sup>,  
Hisashi Yamada<sup>3</sup>, Reiko Azumi<sup>3</sup>, Kentaro Kutsukake<sup>2</sup>, Toru Ujihara<sup>2</sup>

E-mail: s.seki@aixtal.com

【緒言】機械学習モデルを用いて製造条件の最適化を行う際、既存条件からの大幅な条件変更を回避するためには既存条件との近さの指標を目的関数に加えた最適化が有効である[1]。また、一般に実験データが疎な領域では機械学習の予測精度が低下するため、既存データとの距離を考慮することは最適解の信頼性を向上させるために有効であると考えられる。本研究では窒化ガリウム (GaN) の気相成長プロセスに対して本手法を適用し、有効性を検証した。

【方法】我々はこれまでに GaN の気相成長プロセスに対し、シミュレーションと実験データを組み合わせ、各原料ガスの流量から成膜速度および表面品質を予測する機械学習モデルを開発した[2]。表面品質は A~D の 4 段階のクラスに分けられ、品質 A がデバイス作製に必要な平坦な表面である。本モデルを用いて、1) 成膜速度の最大化と 2) 原料の総流量の最小化という競合する 2 目的の最適化および、それらに加え、既知の実験条件とのユークリッド距離の総和を小さくする最適化を行なった。いずれの最適化にも、結晶全面で表面品質クラスの予測値が A となる制約を付与した。最適解にて実際に成膜を行い、成膜速度と表面品質の測定を行なった。

【結果と考察】Fig. 1 に従来条件および最適解における実験結果の成膜速度と原料の総流量を示す。従来条件は青の四角点が結晶全面で品質 A のもの、灰色点がそうでないもので、最適解は×の点が 2 目的の最適化、赤星点は既存条件との距離を考慮した最適化の結果である。提案手法では全面 A を保ちつつ、既存条件よりも原料流量を抑えながら成膜速度の向上を達成できた。既存条件との距離を考慮しない最適解では予測値は全面 A でも実際の成膜では品質 B や D が混ざり成膜速度も予測値より小さくなった。これにより、既存条件の考慮は最適解の信頼性の向上に有効であると言える。

【謝辞】この成果は、国立研究開発法人新エネルギー・産業技術総合開発機構 (NEDO) の委託業務 (JPNP20006) の結果得られたものです。

#### 【参考文献】

[1] 笠原ら、第 72 回応用物理学会春季学術講演会 17a-K505-3、[2] 関ら、第 84 回応用物理学会秋季学術講演会 20p-C601-3

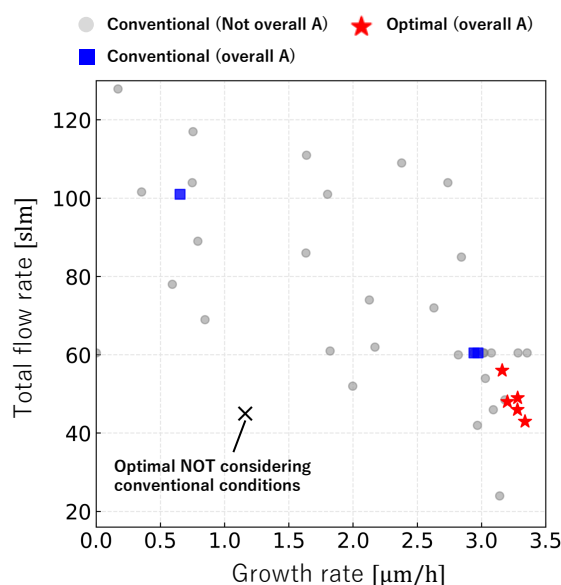


Fig.1 Growth rate and total flow rate of experiments of conventional conditions and optimal conditions. Blue square plots are the conventional conditions with overall class A, gray plots are those other than overall class A and red plots are the optimal conditions.

# 強化学習を用いたデータ駆動型自動 FZ 結晶成長炉の開発と実証実験

## Experimental demonstration of a data-driven automated floating zone crystal growth furnace using reinforcement learning

名大院工<sup>1</sup>, アナモルフォーシスネットワークス<sup>2</sup>, 三幸<sup>3</sup>, 物質・材料研究機構<sup>4</sup>, 名大未来研<sup>5</sup>

○柳沢 壮吾<sup>1</sup>, 松本 遼平<sup>2</sup>, 炭谷 翔悟<sup>2</sup>, 稲垣 拓弥<sup>3</sup>, 坂内 英典<sup>4</sup>, 沓掛 健太郎<sup>1,5</sup>,  
宇治原 徹<sup>1,5</sup>, 原田 俊太<sup>1,5</sup>

Grad. Sch. Eng., Nagoya Univ.<sup>1</sup>, Anamorphosis Networks<sup>2</sup>, Sanko<sup>3</sup>, NIMS<sup>4</sup>, IMASS Nagoya Univ.<sup>5</sup>

○Sogo Yanagisawa<sup>1</sup>, Ryohei Matsumoto<sup>2</sup>, Shogo Sumitani<sup>2</sup>, Takuya Inagaki<sup>3</sup>, Eisuke Bannai<sup>4</sup>,  
Kentaro Kutsukake<sup>1,5</sup>, Toru Ujihara<sup>1,5</sup>, Shunta Harada<sup>1,5</sup>

E-mail: yanagisawa.sogo.jn@unno.material.nagoya-u.ac.jp

最近の研究で、生産性の向上や人手不足の解消などを目的とした製造プロセスへのインフォマティクス応用が数多く検討されている。一方で floating zone (FZ) 法による単結晶成長プロセスは未だ人の手による操作が行われているプロセスの一つで、これは結晶の状態や操作のわずかな際によって全く異なる挙動を示し操業者の経験や勘による操作が必要なためである。これまでの我々の研究においてこのプロセスを対象に強化学習を用いた自動化手法を提案し、シミュレーション環境においてその有効性を実証している[1, 2]。強化学習を用いることで、自動操業モデルの試行錯誤によって操業全体を通して最適な操作を学習することができる。本研究ではこの手法を実際に適用可能な FZ 結晶成長炉を新規に開発し、Si 単結晶成長の自動操業を実験的に実証した。

本研究では、強化学習を利用した自動操業モデルで制御可能な FZ 結晶成長炉を新たに開発した (Fig. 1)。炉内の結晶および融液の状態をカメラで連続的に観察し、画像処理により結晶直径や融液のくびれ部などの特徴量を抽出する。これらを入力として、自動操業モデルが加熱出力や原料供給速度などの最適な操作条件を決定する。このモデルは、オペレーターによる Si 単結晶成長の実験データをもとに理想の形状の結晶を作製するよう強化学習で構築した。実験では、この制御サイクルを繰り返し、Si 単結晶の自動成長を実証した。

Fig. 2 は自動操業によって作製した Si 単結晶の直径の変化を表している。操業実験を複数回行った結果、安定して Si 単結晶を作製することに成功した。この結果は強化学習が FZ 法での異なる材料の結晶成長や他の製造プロセスにも応用できることを示唆している。

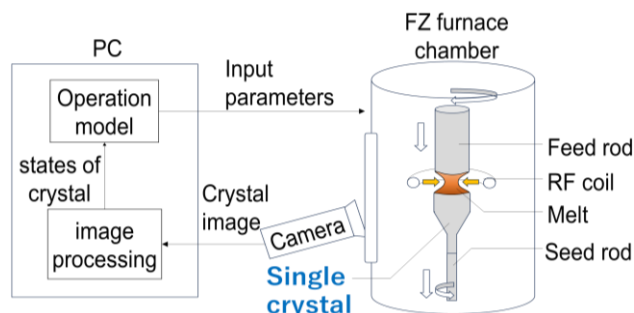


Fig. 1. Schematic illustration of a FZ furnace.

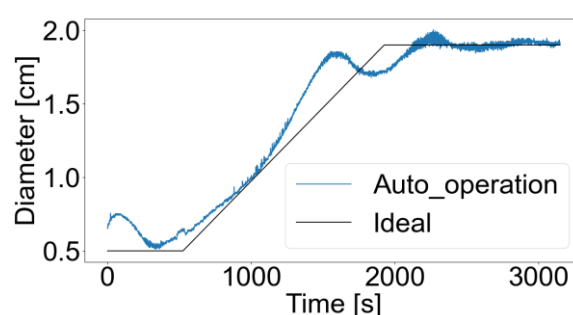


Fig. 2 Si crystal diameter in a growth experiment.

### 【参考文献】

- [1] Omae, R. et al. Sci. Technol. Adv. Mater.: Methods. 2(1), 294-301(2022).
- [2] Tosa, Y. et al. Sci. Rep. 13(1), 7517(2023).

# 半導体製造多段工程における誤差蓄積構造に着目した機械学習モデル設計

## Machine learning model design focusing on error accumulation structure in multi-stage semiconductor manufacturing process

名大院工<sup>1</sup>, 名大未来研<sup>2</sup> ○山田 拓実<sup>1</sup>, 杵掛 健太朗<sup>1,2</sup>, 原田 俊太<sup>1,2</sup>, 宇治原 徹<sup>1,2</sup>

Grad. Sch. Eng., Nagoya Univ.<sup>1</sup>, IMASS Nagoya Univ.<sup>2</sup>

○Takumi Yamada<sup>1</sup>, Kentaro Kutsukake<sup>1,2</sup>, Shunta Harada<sup>1,2</sup>, Toru Ujihara<sup>1,2</sup>

E-mail: yamada.t@unno.material.nagoya-u.ac.jp

【はじめに】半導体製造プロセスなどの多段工程に機械学習を適用する場合、各段階で生じる予測誤差が累積し、最終的な予測精度が低下することが知られているが、その効果的な抑制手法は研究途上である。本研究では、入力分布の構造に注目し、誤差蓄積の要因を解析し、その抑制手法を検討した。

【方法】イオン注入に伴う欠陥生成と熱拡散を繰り返す多段工程を対象に、シリコンウェーハ内部の点欠陥密度の一次元分布を計算する簡易シミュレーターを構築した。入力となる欠陥密度分布に対し、空孔と格子間原子を付加し拡散方程式を解いた後、更新された分布を出力する (Fig.1)。プロセスパラメータは、それぞれの欠陥導入の範囲 (始点と終点) と量、熱処理の温度・時間の計 8 つである。

プロセスパラメータをランダムに生成し、シミュレーターで 100 段工程・1000 データを作成し、それを教師データとしてニューラルネットワークにより予測モデルを構築した。誤差蓄積の要因として、入力分布に特定の「構造」が含まれると誤差が増幅することが示唆されたため、多様なノイズ構造を人工的に付与して影響を評価した。さらに、評価結果に基づき、特定の「構造」の発生を抑制する正則化項を損失関数に加え、性能比較を行った。

【結果】導入したノイズは、ランダムノイズ、定偏差ノイズ (全体のシフト)、および分布構造を顕著化・平坦化する構造ノイズの 3 種類である。解析の結果、本系では構造ノイズが誤差蓄積を増幅させることが明らかとなった。この解析結果に基づき、構造ノイズの蓄積を抑制するため、欠陥密度分布の移動平均との差異に基

づく正則化項を加えた損失関数を設計し予測モデルを改良した。Fig.2 に誤差蓄積の比較を示す。一段ごとの予測誤差 (RMSE) は従来モデルが 0.429、提案モデルが 0.433 であり、一段あたりの誤差はわずかに増加するものの、多段工程における誤差蓄積は抑制された。

【考察】本研究により、多段工程における誤差蓄積が入力分布に内在する特定構造によって顕著に増幅されることを示した。また、これを抑制する正則化項の導入により、工程全体の予測精度向上が可能であることが確認された。この結果は、多段工程予測における誤差制御手法の設計指針として有用であると考えられる。

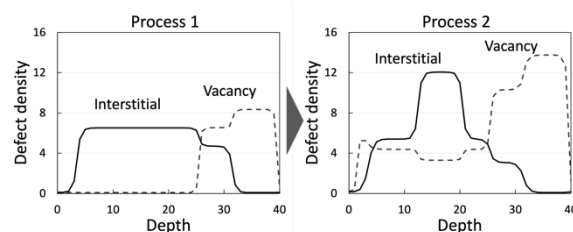


Fig.1 Example of simulator output

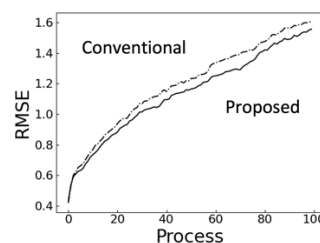


Fig.2 Error accumulation in multi-stage process prediction

謝辞：本研究の一部は、NEDO 先導研究プログラム(JPNP14004)により実施された。

# Jacobian に着目した正則化による物理系の状態時間発展予測に関する 深層学習のロバスト化

Robust deep learning prediction for physical dynamics with Jacobian oriented  
regularization

アイクリスタル<sup>1</sup>, 名大院工<sup>2</sup>

○関 翔太<sup>1,2</sup>, 中西 佑児<sup>1</sup>, 高石 将輝<sup>1</sup>, 沓掛 健太郎<sup>2</sup>

Aixtal<sup>1</sup>, Nagoya Univ.<sup>2</sup>

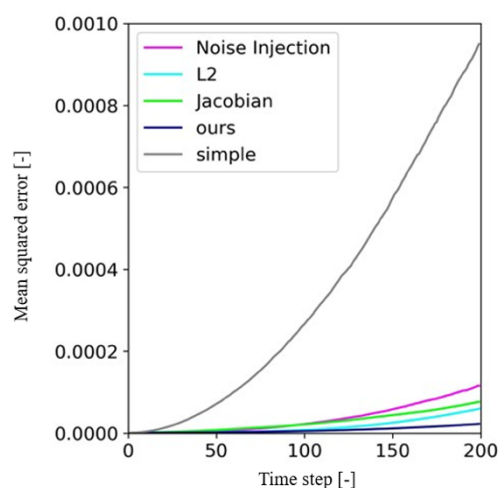
°Shota Seki<sup>1</sup>, Yuji Nakanishi<sup>1</sup>, Masaki Takaishi<sup>1</sup>, Kentaro Kutsukake<sup>2</sup>

E-mail: s.seki@aixtal.com

物理系の状態の時間発展を予測する際、時間発展に伴う誤差の蓄積が問題となる。これは数値計算力学全般に言えることだが、特に自己回帰型深層学習モデルのように、前時刻の出力が次の時刻の入力となるようなサロゲートモデルでは顕著に現れる[1]。蓄積誤差は時間ステップごとの入力される分布のわずかなブレによる誤差の蓄積であるため、蓄積誤差低減にサロゲートモデルの入出力に関する Jacobian ノルムを低減させるような正則化手法が有効であると考えられる。

本研究では、単結晶シリコン内の空孔および格子間シリコンの拡散現象を題材とし、空孔と格子間シリコン濃度の時間発展を予測する深層学習モデルを作成した。(i)正則化なし、(ii)入力にノイズを加えることによるロバスト化、(iii)L2 正則化、(iv)Jacobian ノルム正則化[2]、の4条件における学習結果を比較し、正則化の効果を検証した。また、従来の Jacobian ノルム正則化は訓練データ点での Jacobian ノルムしか計算しておらず、訓練データ点以外でも Jacobian ノルムを評価する新しい正則化手法である(v)修正 Jacobian 正則化を開発した。

Fig. 1 に各正則化手法における誤差の推移を示す。いずれの正則化手法も正則化なしの場合よりも低い誤差を示していることがわかる。したがって Jacobian に着目した正則化は誤差蓄積に有効だと言える。さらに、今回の数値実験では提案手法である修正ヤコビアン正則化が最もよい性能を示した。講演ではより詳細な結果と数学的考察を紹介する予定である。



謝辞

本研究は国立研究開発法人新エネルギー・産業技術総合開発機構(NEDO)先導研究プログラム(JPNP14004)の結果に基づいている。

[1] Z. Chen, M. Ma, T. Li, H. Wang, and C. Li, Information Fusion, **97**, (2023), 101819.

[2] J. Hoffman, D. A. Roberts, and S. Yaida, ArXiv, vol. abs/1908.02729, (2019)

## 深層学習を活用した半導体加工形状の自動評価パイプライン

### Deep Learning-Based Automated Evaluation Pipeline for Semiconductor Etched Geometry

奈良先端大<sup>1</sup>, 上智大<sup>2</sup>, <sup>○</sup>野島 聖耀<sup>1</sup>, 岩満 一功<sup>1</sup>, 赤瀬 善太郎<sup>1</sup>, 菊池 昭彦<sup>2</sup>, 富谷 茂隆<sup>1</sup>

NAIST<sup>1</sup>, Sophia Univ.<sup>2</sup>, <sup>○</sup>Seiyo Nojima<sup>1</sup>, Kazunori Iwamitsu<sup>1</sup>, Zentaro Akase<sup>2</sup>, Shigetaka Tomiya<sup>1</sup>

E-mail: [nojima.seiyo.nr0@naist.ac.jp](mailto:nojima.seiyo.nr0@naist.ac.jp)

半導体デバイスにおいて、微細構造の寸法や形状はデバイス特性に大きく影響を与えるため、正確な計測が不可欠である。特に電子顕微鏡を使った形状評価は、高分解能で詳細な構造を確認できる。しかし、従来の手動評価では作業に時間がかかるうえに、作業者によってバイアスが生じる課題が存在する。近年、深層学習による顕微鏡画像の自動解析手法が提案されているものの、その検証は主に構造が明瞭な画像に限られている[1]。また、画像解析手法の構成に関しては、先行研究において比較検討は十分とは言えず、複数の構成を評価する必要があると考えられる。本研究では、多様な画像にも対応可能な自動評価パイプラインの構築を目指した。図1にパイプラインの全体像を示す。さらに、画像解析に関して、複数の構成を実装してロバスト性を比較検討した。

異なるエッチング条件の HEATE 法[2]により加工された GaN トレンチ構造の断面 SEM 画像 92 枚を用い、2つの深層学習モデルを構築した。セグメンテーションモデルである U-Net では画像の背景と基板を識別し、物体検出モデルである Faster R-CNN では画像からトレンチ領域を検出する。2つのモデルを組み合わせることによって画像解析を行い、トレンチ形状の評価を可能にする。本研究では、2つのモデルの適用順序を検討し、それぞれのパイプラインを比較した。画像解析の定量評価には Intersection over Union (IoU) を用いた。

パイプラインの比較の結果としては、U-Net を先に適用した場合は IoU が 95%であり、Faster R-CNN を先に適用した場合は IoU が 97%であった。高精度な画像認識により、正確な寸法測定や形状異常の検出が可能になることが期待される。講演では、取り組みの詳細に加えて、画像認識後の形状評価としての異常形状の検出、FIB-SEM による 3 次元構造観察への展開、自動的に取得されるデータの体系的な蓄積と利活用についても触れる予定である。

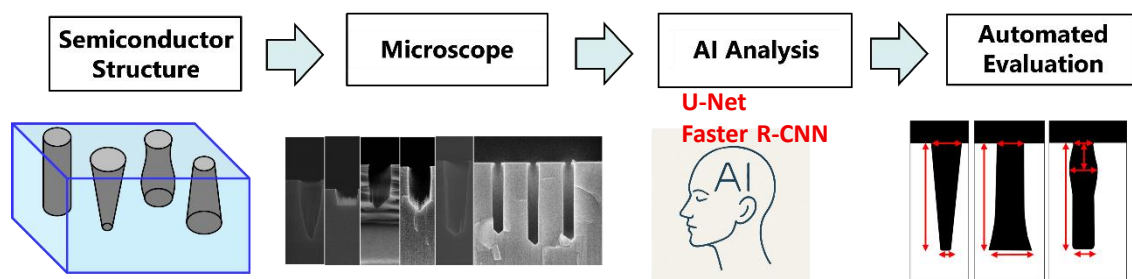


Fig. 1: Overview of the pipeline. Images of semiconductor structures are captured via microscopy, processed using AI, enabling automated evaluation of dimensions and shapes.

[1] Y. Okuyama and T. Ohmori, Jpn. J. Appl. Phys. 62 SA1016/1-9, (2023)

[2] R. Kita et al., Jpn. J. Appl. Phys. 54 046501/1-5, (2015)



## 三次元アトムプローブと透過電子顕微鏡法のバイモーダル相関顕微手法の開発

### Development of a Bimodal Correlative Microscopy Method Using Three-Dimensional Atom Probe and Transmission Electron Microscopy Analysis

奈良先端大<sup>1</sup>, 物質・材料研究機構<sup>2</sup>, 大阪大学<sup>3</sup>,<sup>○</sup>赤瀬 善太郎<sup>1</sup>, 別所泰成<sup>1</sup>, 大竹 義人<sup>1</sup>,  
埋橋 淳<sup>2</sup>, 岩満 一功<sup>1</sup>, 中島 宏<sup>3</sup>, 山崎 順<sup>3</sup>, 大久保忠勝<sup>2</sup>, 富谷 茂隆<sup>1</sup>

NAIST<sup>1</sup>, NIMS<sup>2</sup>, Osaka Univ.<sup>3</sup>, <sup>○</sup>Zentaro Akase<sup>1</sup>, Taisei Bessho<sup>1</sup>, Yoshito Otake<sup>1</sup>, Jun Uzuhashi<sup>2</sup>,  
Kazunori Iwamitsu<sup>1</sup>, Hiroshi Nakajima<sup>3</sup>, Jun Yamasaki<sup>3</sup>, Tadakatsu Ohkubo<sup>2</sup>, Shigetaka Tomiya<sup>1</sup>

E-mail: z.akase@naist.ac.jp

先端半導体デバイスは、省エネルギー化と高速化の要求に応じて、構造の微細化および三次元的複雑化が急速に進んでいる。このような背景から、性能に影響を及ぼす微量不純物の三次元分布を高精度に把握する必要性が高まっている。当研究グループでは、三次元アトムプローブ (3DAP) と電子線トモグラフィー (ET: 透過電子顕微鏡法 TEM または走査透過電子顕微鏡法 STEM の連続傾斜像による三次元再構成法) を組み合わせた相関顕微鏡法の確立に取り組んでいる。3DAP は、試料中の構成原子を高感度かつ三次元的に再構築可能な手法であるが、半導体デバイスの解析に応用する際、再構成像には不可避免的に歪みが導入される。そこで本研究では、形状の再現性に優れた ET データを基準とし、3DAP データに対して非剛体変形を伴う位置合わせ (非剛体レジストレーション) を適用することで、構造歪を補正し、微量不純物の正確な三次元分布の把握を可能とする手法の実現を目指している。

モダリティの異なるデータ間の相関を評価するため、3DAP データは非剛体変形後に一度 TEM 様データへ変換し、勾配相関により両者の一致度を評価する。前回の発表では、TEM 様データの各画素における画素値を、3DAP データ上でその位置の近くに存在する原子が与える像強度値 (元素によって異なる値を仮定) から補間して求めていた。しかしこの手法では、異種原子の混合比はコントラストに反映される一方で、同種原子の濃度分布は変換像のコントラストに反映されないという課題があった。そこで今回、同種原子間の濃度差も変換像のコントラストに反映させるため、ガウスぼかしによって強度の重ね合わせを行うモダリティ変換手法 (Fig. 1) を導入したところ、30% のイテレーション回数で、勾配相関値が約 2 倍に向上することを確認した。Fig. 1 (c)、(d) に、収束後の旧手法と新手法による変換結果を示す。新手法では AlN/GaN 界面から AlGa<sub>2</sub>N<sub>3</sub> へ向けてのなだらかなコントラスト変化が再現されたことにより、勾配相関値が向上したと考えられる。発表では、このようなアルゴリズムの改善点に加え、複雑な変形への正則化項の影響や、実験データを用いた解析事例についても詳述し、今後の同一試料同一視野での相関顕微観察への展開について議論する。

本研究の一部は、キオクシア株式会社奨励研究、NIMS 連携拠点推進制度、奈良先端科学技術大学院大学データ駆動型サイエンス創造センター新規課題発掘活動助成金および増屋記念基礎研究振興財団研究助成金の助成、ならびに NIMS 電子顕微鏡ユニットからの支援を受けて実施したものである。

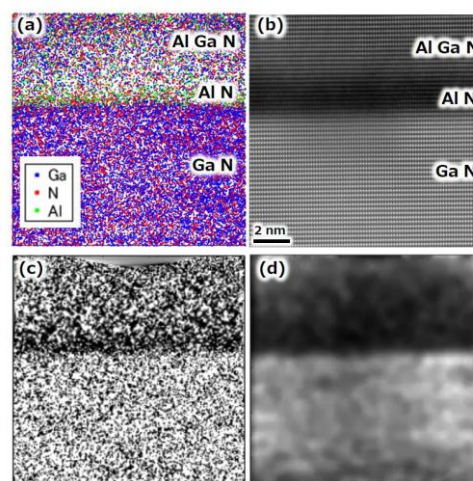


Fig. 1 AlGa<sub>2</sub>N<sub>3</sub>/AlN/GaN interface: (a) 3DAP data, (b) HAADF-STEM image (acquired from a different field of view), (c) TEM-like conversion using our previous method, and (d) conversion result using the improved method developed in this study.

# 深層学習エリプソメトリー：光学定数・薄膜構造・バンドギャップの 超高速・高精度解析

## Deep Learning Ellipsometry: Ultrafast and High-Accuracy Determination of Optical Constants, Thin-Film Structures, and Bandgaps

岐阜大学<sup>1</sup>, J.A. Woollam Co., Inc.<sup>2</sup>, 産業技術総合研究所<sup>3</sup>

岡 亮輔<sup>1</sup>, 山本 祐輝<sup>1</sup>, James N. Hilfiker<sup>2</sup>, 反保 衆志<sup>3</sup>, 永井 武彦<sup>3</sup>, 林 真弘<sup>1</sup>, 藤原 裕之<sup>1</sup>

Gifu University<sup>1</sup>, J.A. Woollam Co., Inc.<sup>2</sup>, AIST<sup>3</sup>

Ryosuke Oka<sup>1</sup>, Yuki Yamamoto<sup>1</sup>, James N. Hilfiker<sup>2</sup>, Hitoshi Tampo<sup>3</sup>, Takehiko Nagai<sup>3</sup>,

Masahiro Hayashi<sup>1</sup> and Hiroyuki Fujiwara<sup>1</sup>

E-mail: fujiwara.hiroyuki.u7@f.gifu-u.ac.jp

【はじめに】半導体材料の光学特性を迅速かつ高精度に評価することは、光電子デバイスの開発において極めて重要である。分光エリプソメトリー (SE) は高精度な光学評価手法であるが、特に薄膜に対しては、解析が複雑で時間を要する本質的な課題がある。この問題を克服するため、深層学習 (DL) 手法が提案されてきたが、従来の DL 手法は滑らかな光学定数の解析のみに対応しており[1]、高吸収領域では解析誤差が非常に大きいという問題があった[2]。本研究では、従来の DL 手法を大幅に改良し、専門知識を必要とせず、わずか 2 秒で光学定数、膜厚、バンドギャップを高精度に予測可能な革新的な DL 解析法を提案する。

【深層学習法】本研究では、表面ラフネス層/バルク層/SiO<sub>2</sub>/結晶 Si からなる多層構造を光学モデルとして仮定し、1 億と極めて大規模な仮想データを用いて学習を行った。さらに、従来の DL アルゴリズムを大きく改良し、非常に深い DL 構造を適用した。特に図 1 に示すように、ノイズを意図的に導入したエリプソメトリースペクトルを作成し、これを転移学習させることにより、ノイズを含む実験スペクトルにおいても高精度解析が可能なアルゴリズムを構築した。教師データの作成に関しては、薄膜層の光学定数は、複数の Tauc-Lorentz ピークをランダムに組み合わせることにより生成した。結果として、5 種類の入射角 (50~70°) におけるエリプソメトリースペクトルを入力とし、出力として屈折率-消衰係数スペクトル( $n, k$ )、対数での吸収係数スペクトル( $\log \alpha$ )、バルク層膜厚( $d_b$ )および表面ラフネス層膜厚( $d_s$ )を予測するモデルを構築した。

【結果】図 2(a)は、デバイス用 CuInGaSe<sub>2</sub>(CIGS)薄膜に対する解析結果を示しており、開発 DL 手法により予測された光学定数は、一般に使用されているマニュアル解析(B スプライン解析)による結果と非常に高い一致を示した。同様の結果は他の試料解析においても確認されており、膜厚の絶対誤差は a-Si:H において 1 nm 未満、CIGS において 5 nm 以内であり、十分に小さい値が確認された。図 2(b)は、32 種類の実験的な半導体薄膜材料に対して DL 解析およびマニュアル解析から算出したバンドギャップの関係を示している。この図におけるバンドギャップの平均絶対誤差 (MAE) は 31 meV であり、決定係数は 0.997 と非常に高い値を示し、開発手法の極めて高い精度が確認された。さらに DL 解析の解析時間は 1 試料あたり 2 秒以下であり、従来の手動解析(5 分)に比べて超高速解析が可能である。また、DL の解析結果を B スプライン解析の初期値として利用することで、さらに高精度解析を行うことが可能であることを確認した。

【謝辞】a-Si:H 試料を提供していただいた産業技術総合研究所の松井卓矢氏に深く感謝します。

[1] Liu et al. Light: Sci, Appl. 10 (2021) 55, [2] Li et al., Adv. Photonics Res. 2 (2021) 2100147

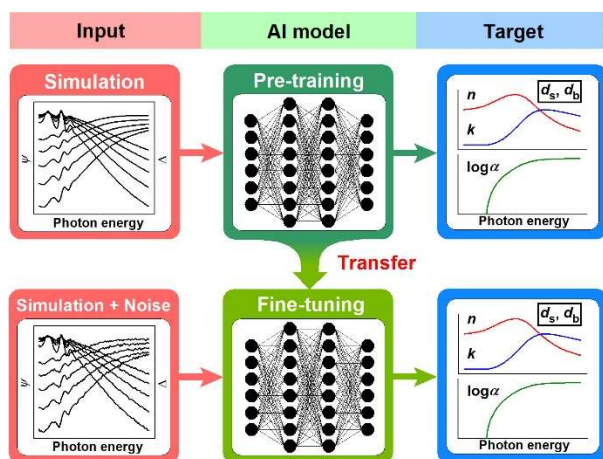


Figure 1. Procedure of developed DL analysis.

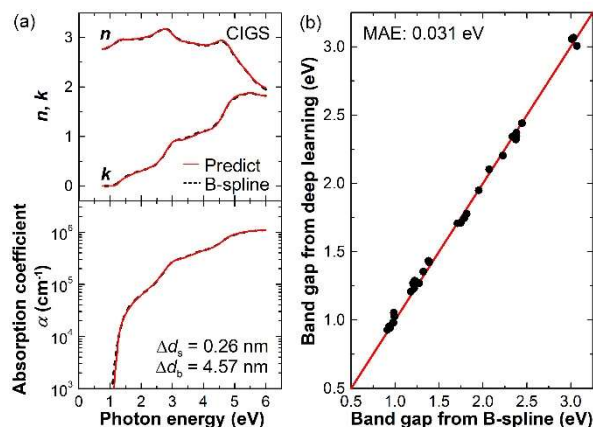


Figure 2. Analysis results of deep-learning and manual (B-Spline) analyses.



# 拡張型自由エネルギーモデルを用いた InGaAs-HEMT のデバイス動作に対するゲート長依存性の研究

Study on the Gate Length Dependence of InGaAs-HEMT Device Operation Using Extended Free Energy Model<sup>1</sup>,

中村晴登<sup>1</sup>, 町田陽太郎<sup>2</sup>, 長岡竜之輔<sup>2</sup>, 渡邊一世<sup>3</sup>, 小嗣真人<sup>2</sup>, 吹留博一<sup>1</sup>

Tohoku Univ.<sup>1</sup>, Tokyo University of Science<sup>2</sup>, NICT<sup>3</sup>

E-mail: nakamura.haruto.s7@dc.tohoku.ac.jp

ミリ波帯やテラヘルツ帯で動作する超高周波デバイスは、次世代無線通信(Beyond 5G)のキーデバイスであり、更なる高機能化のために、物理に根差したデバイス動作機構の解析手法の創発が希求されている。我々は、拡張型自由エネルギーフレームワーク[1]を用いて、機械学習に基づく超高周波デバイスの新たな解析に取り組んでいる[2]。本発表では、Persistent Homology (PH) を用いて、InGaAs をチャネルとした高電子移動度トランジスタ(InGaAs-HEMT)内部の微視的な電位分布と巨視的な電流増幅機能の解析結果を報告する。

InGaAs-HEMT 内部の物性（例：電位）の微視的分布をデバイス・シミュレータ（Silvaco 社 Atlas）を用いて取得し、同時に電流利得遮断周波数  $f_T$  を算出した。データ取得時のパラメータは、ゲート長  $L_g = 35 \sim 700$  nm、ゲート電圧  $V_{gs} = -0.6 \sim 0.6$  V、ドレイン電圧  $V_{ds} = 0.6$  V とした。

得られた電位分布に関する画像データに対して、PH を適用し、主成分分析を行って、特徴量の第1主成分 (PC1) と第2主成分 (PC2) を抽出した。PC1 と PC2 から成る情報空間上に、 $f_T$  にディラック定数  $\hbar$  を乗じた拡張型自由エネルギー曲面  $\hbar \times f_T$  を描画した (Fig. 1)。その結果、この曲面に顕著なゲート依存性が認められた。詳細な解析のために、この曲面を第2主成分に対して射影した (Fig. 2)。Fig. 2 について、立ち上がり領域を線形回帰して、その傾きをプロットしたものが Fig. 3 である。Fig. 3 において、短ゲート領域

( $L_g < 100$  nm) と長ゲート領域 ( $L_g > 100$  nm) で傾きのゲート長依存性に大きな違いがみられた。この結果は、HEMT の動作効率に大きな影響を及ぼす「短チャネル効果」を従来とは異なる新たな視点で説明できることを示唆する。

以上のようにして、我々は超高周波トランジスタにおける微視的な電位分布と巨視的な電流増幅機能を、新たな観点から双方向的に解析できる手法を開発した。

【参考】[1] S. Kunii *et al.* Sci. Rep. **12**, (2022) 29892. [2] 志賀研仁, 谷脇三千輝, 長岡龍乃輔, 渡邊一世, 小嗣真人, 吹留博一, 2024 年春季応用物理学会 24p-P07-4.

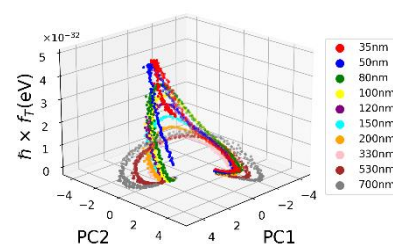


Fig. 1 Extended Free Energy analysis for the operation of InGaAs-HEMT.

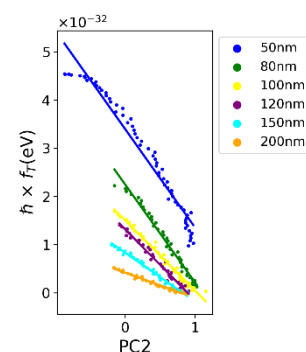


Fig. 2 Projection of onto PC2

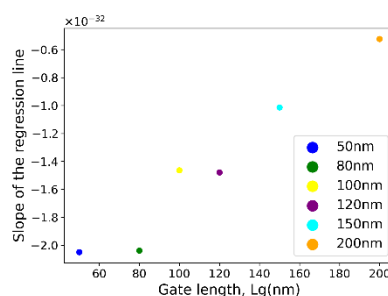


Fig. 3 Relationship between the slope of the regression line and  $L_g$

# カルマンフィルタによるデータフュージョンの測定系最適化

## Optimization of Sensor Locations by Kalman Filters in Data Fusion

アイクリスタル株式会社<sup>1</sup> ○ 竹内 靖<sup>1</sup>

Aixtal Corporation<sup>1</sup>, ○ Yasushi Takéuchi<sup>1</sup>

E-mail: y.takeuchi@aixtal.com

カルマンフィルタは、制御工学やデータ同化の研究を通じて発展し利用されてきた。データ同化は、予測モデル(シミュレーションなど)の出力情報を現実の測定データと同化させることにより、予測モデルの不確定な入力情報(初期条件や境界条件、パラメータ)を修正して予測を改善し、同時に、現実の時空間における測定値の欠損を予測値で補完する技法である。本講演の報告する理論は、特に複数のセンサー(測定点)を対象系に設置する場合に、センサーの設置位置など測定系の最適化を、カルマンフィルタを用いるデータ同化により可能にする。

複数のセンサーからの測定データの融合(データフュージョン)において、センサーへ対応して複数のサブシステムに分散したカルマンフィルタ[1]を考える(図1)。ステップ(時刻) $k$ における状態 $\mathbf{x}(k)$ に対して、カルマンゲイン $\mathbf{K}_k$ は、観測行列 $\mathbf{C}_k$ と、状態の予測値 $\mathbf{x}^f(k)$ および観測値 $\mathbf{y}(k)$ の誤差の共分散行列 $\mathbf{P}_k^f$ ,  $\mathbf{R}_k$ から与えられる;  $\mathbf{K}_k = \mathbf{P}_k^f \mathbf{C}_k^t (\mathbf{R}_k + \mathbf{C}_k \mathbf{P}_k^f \mathbf{C}_k^t)^{-1}$ . ( $t$ は転置を表す。)そして、データ同化による修正値 $\mathbf{x}^a(k)$ は次式で与えられる;

$$\mathbf{x}^a(k) = \mathbf{x}^f(k) + \mathbf{K}_k \{ \mathbf{y}(k) - \mathbf{C}_k \mathbf{x}^f(k) \}. \quad (1)$$

ここで、ステップ $k$ のセンサー位置を変更すると、対応する $\mathbf{x}(k)$ の摂動は式(1)に従い伝搬してゆく。この場合、センサー位置の最適化は、状態遷移行列 $\mathbf{A}_k$ 、観測行列 $\mathbf{C}_k$ から可観測性グラム行列 $\mathbf{G} = \sum_k \mathbf{A}_k^t \mathbf{C}_k^t \mathbf{C}_k \mathbf{A}_k$ を求め、 $\mathbf{G}$ の最小固有値を最大化する操作により達成される[2, 3]。

図1のようなカスケード系の分散型カルマンフィルタにおいて、 $\mathbf{A}_k$ ,  $\mathbf{C}_k$ が非可換であるので、原理的に、各ステップ $k = 1, 2, \dots, N$ のセンサー位置の最適化は互いに独立ではない。講演では、ローレンツモデルなど簡単な例を対象系として、測定系の最適化計算を例示する。

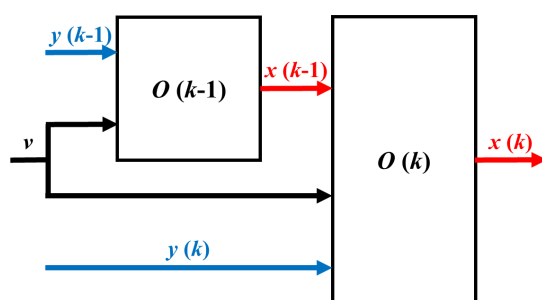


Figure 1: Cascaded observers;  $k$ th state  $\mathbf{x}(k) (= \mathbf{A}_{k-1}\mathbf{x}(k-1) + \mathbf{B}_{k-1}\mathbf{v}(k-1))$  and its observation  $\mathbf{y}(k) (= \mathbf{C}_k\mathbf{x}(k) + \mathbf{w}(k))$  with the state transition matrix  $\mathbf{A}_k$ , the forcing matrix  $\mathbf{B}_k$ , the forcing vector  $\mathbf{v}(k)$ , the observation matrix  $\mathbf{C}_k$ , and the observation error  $\mathbf{w}(k)$ .

[1] Z. Lendek, R. Babuška, B. De Schutter, Eng. Appl. Artificial Intelligence **21**, 457-469 (2008).

[2] W. Kang, L. Xu, Tellus A **64**, 17133 (2012).

[3] S. King, W. Kang, L. Xu, Int. J. Dynamics Control **3**, 416-424 (2015).

社会キャピタル理論に基づく画期的研究の発生メカニズム解明  
The mechanism of generation of ground-breaking research based on  
social capital theory

福岡女子大, 品川啓介

Fukuoka women’s Univ.

E-mail: k.shinagawa@fwu.ac.jp

自然科学の研究で生じた画期的研究および、同時期に注目集めつつも画期的研究に至らなかった研究を判別する方法が確立されれば国や企業の研究投資政策に大きく貢献する。このため前者については膨大な科学論文の書誌情報について早期発見を目指す研究が増えているが、後者についてはほとんど見られない。そこで本研究では学論文の書誌情報の分析において経営学の有力な理論の一つである社会キャピタル理論を用い、両研究発の発生過程を比較し成り立ちの違いを議論する。一般にこの理論は、人や組織が構成する社会的繋がりから情報の移動や融合のし易さを説明するが、本研究では人の繋がりではなく個々の論文が付与するキーワードの繋がり方（共起関係）を観察しそのネットワークの形成過程を比較する。なお、この比較では以下の問いを設ける。それは「画期的研究を生じさせた研究群のキーワードのネットワーク構造はブリッジング型構造ではないか」また「注目を集めつつも画期的研を生じさせなかった研究群のキーワードのネットワーク構造は網目状のボンディング型構造ではないか」というものである。この論拠は社会キャピタル理論ではブリッジング型において多様な情報が隔々まで行き渡りやすくイノベーションが生じやすいとされ、ボンディング型は相互監視や暗黙の行動規範が生じやすいとされることからそこで生まれた規範に囚われやすい、とされることによる（図 1 参照）。なお本研究では青色発光ダイオード開発において、競合した GaN（Gallium nitride）開発研究と ZnSe（Zinc selenide）開発研究を対象として選び比較する。これらの研究に用いられたキーワードのネットワークの構造と当時の研究者へのインタビュー文献の記述の相関関係を検証する。

| 型    | ブリッジング型               | ボンディング型              |
|------|-----------------------|----------------------|
| モデル図 |                       |                      |
| 特徴   | 多様な情報が効率的に隔々まで行き渡りやすい | 相互監視、暗黙の行動規範が共有されやすい |

図 1. 社会ネットワーク理論におけるブリッジング型とボンディング型ネットワークの比較図  
(Colman, J. S. (1988) American J. of Sociology, Vol.94, pp95-120 をもとに筆者が解釈した図)

23 合同セッションN「インフォマティクス応用」 | 一般セッション(ポスター講演): 23.1 合同セッションN「インフォマティクス応用」

## [8p-P11-1~13] 23.1 合同セッションN「インフォマティクス応用」

[8p-P11-1]

グラフ畳み込みネットワークを用いた多様な有機化合物の原子電荷予測

○(B)本間 悠吾<sup>1</sup>、本間 友<sup>1</sup>、松井 弘之<sup>1</sup> (1.山形大ROEL)

[8p-P11-2]

類似分子との差分に基づくHOMO,LUMOエネルギーの機械学習

○(B)男澤 泰斗<sup>1</sup>、佐々木 蓮<sup>1</sup>、岡田 智悠<sup>1</sup>、望月 祐樹<sup>1</sup>、松井 弘之<sup>1</sup> (1.山形大ROEL)

[8p-P11-3]

高速量子化学計算のための回転同変行列の線形結合に基づくハミルトニアン行列の機械学習

○(B)遠藤 仁太郎<sup>1</sup>、尾沢 昂輝<sup>1</sup>、松井 弘之<sup>1</sup> (1.山形大ROEL)

[8p-P11-4]

機械学習分子動力学シミュレーションによるチタニアシリカ膜でのガス分子の拡散挙動の解析

○山崎 周<sup>1</sup>、吉本 勇太<sup>1</sup>、五社 一心<sup>2</sup>、藤澤 拓<sup>2</sup>、松田 友也<sup>1,2</sup>、松村 直樹<sup>1</sup>、岩崎 有登<sup>1</sup>、井上 淳樹<sup>2</sup>、吉岡 朋久<sup>2</sup>、川口 博<sup>2</sup>、坂井 靖文<sup>1</sup> (1.富士通株式会社、2.神戸大学)

[8p-P11-5]

機械学習分子動力学シミュレーションによるSrTiO<sub>3</sub>の酸素欠陥の形成エネルギーの予測

○西口 和孝<sup>1,2</sup>、山本 峻太<sup>2</sup>、山崎 周<sup>1</sup>、松村 直樹<sup>1</sup>、吉本 勇太<sup>1</sup>、天能 精一郎<sup>2</sup>、坂井 靖文<sup>1</sup> (1.富士通、2.神戸大工)

[8p-P11-6]

PFP descriptorsとShapley値を組み合わせた原子レベルでの予測の解釈

○張 凡<sup>1</sup> (1.(株) Preferred Computational Chemistry)

[8p-P11-7]

機械学習を用いたGaN-HEMT用アモルファスSiN薄膜成膜プロセスの最適化

○栗林 壮太郎<sup>1</sup>、本間 健司<sup>1</sup>、丸尾 昭人<sup>1</sup>、樋口 博之<sup>1</sup>、實宝 秀幸<sup>1</sup>、山田 敦史<sup>1</sup>、多木 俊裕<sup>1</sup> (1.富士通株式会社)

[8p-P11-8]

AIモデルによるプロセス評価値を用いた半導体乾燥プロセス最適化

○久保 友輔<sup>1</sup>、中野 佑宇唯<sup>1</sup>、村島 広衛<sup>1</sup>、國枝 省吾<sup>1</sup>、佐々木 悠太<sup>1</sup>、塙 洋祐<sup>1</sup> (1.株式会社SCREENホールディングス)

[8p-P11-9]

パーシステントホモロジーで可視化するイオンビーム照射による半導体微細構造のプロセスインフォマティクス

○走出 和寛<sup>1</sup>、宮田 滯<sup>1</sup>、新田 紀子<sup>1</sup>、大石 脩人<sup>2</sup>、大林 一平<sup>3</sup> (1.高知工大、2.高知高専、3.岡山大)

[8p-P11-10]

偏光情報と温度変化を基にした深層学習による複屈折像の解析手法の開発

○真中 浩貴<sup>1</sup>、本田 宗一郎<sup>1</sup>、三浦 陽子<sup>2</sup> (1.鹿児島大学院理工、2.鈴鹿高専)

[8p-P11-11]

LLMを使った材料開発(III)：科学論文からの材料データ抽出

○岡 博之<sup>1</sup>、石井 真史<sup>1</sup> (1.NIMS)

---

[8p-P11-12]

ローカル LLMを用いた論文からの成長層構造抽出

○権 晋寛<sup>1</sup> (1.東大ナノ量子)

---

[8p-P11-13]

LLMとシミュレータを用いた材料開発支援システムの構築に向けて

Jun Jin Choong<sup>1</sup>、○小澤 圭右<sup>1</sup>、石黒 勝彦<sup>1</sup>、三好 利昇<sup>1</sup>、井尻 善久<sup>1</sup>、出村 雅彦<sup>2</sup>、源 聡<sup>2</sup>、門平 卓也<sup>2</sup> (1.SB Intuitions、2.NIMS)

---



# グラフ畳み込みネットワークを用いた多様な有機化合物の原子電荷予測

## Graph Convolutional Network for Predicting Atomic Charges of Various Organic Compounds

山形大 ROEL<sup>1</sup>, °(B)本間 悠吾<sup>1</sup>, 本間 友<sup>1</sup>, 松井弘之<sup>1</sup>

ROEL, Yamagata Univ.<sup>1</sup>, °Yugo Homma<sup>1</sup>, Yu Homma<sup>1</sup>, Hiroyuki Matsui<sup>1</sup>

E-mail: t221191@st.yamagata-u.ac.jp, h-matsui@yz.yamagata-u.ac.jp

【緒言】近年、機械学習を用いた原子電荷の高精度予測に関する研究が盛んに行われている。しかし、既存研究の多くはコンピューター上で生成された薬剤研究用の分子データセットに依存しており<sup>[1]</sup>、有機半導体研究などで重要な共役系・芳香族系化合物への適用可能性は不明である。そこで本研究では、共役系・芳香族系化合物をより多く含むケンブリッジ結晶構造データベース (CSD) を使用し、グラフ畳み込みネットワーク (GCN)<sup>[2]</sup>を用いた新たな原子電荷予測モデルの構築を試みた。

【方法】CSD から炭素・水素・窒素のみで構成された分子をスクリーニングした後、Gaussian を用いて B3LYP/6-31(d,p)で構造最適化を行い、Antechamber を用いて RESP 電荷を算出した。機械学習では、原子番号および general amber force field (GAFF)の原子タイプ (42 種) の 1-hot エンコーディングを入力特徴量とした GCN を使い、出力前に電荷の総和が分子全体で 0 となるよう補正した。また、学習時には、原子ごとのデータ数の偏りを考慮して重み付けを行った。

【結果】データセットの総分子数およびナフタレン構造が含まれる分子の割合は、先行研究の ZINC データセットでは 1056 個、3.60%であったのに対し、今回のデータセットでは 9372 個、13.26%であり、データ数の増大と縮環などの多様性の向上が見られた (表 1)。RESP 電荷の予測値と真の値を比較した結果を図 1 に示す。決定係数は 0.8578、二乗平均平方根誤差は 0.1036 であった。半経験的手法である AM1-BCC 電荷の決定係数は 0.5747、二乗平均平方根誤差は 0.1794 であったため、AM1-BCC と比較してそれぞれ 73.2%および 49.3%の精度向上が確認された。

### 【参考文献】

1. J. Wang et al., Briefings in Bioinformatics 22, bbaa183 (2021).
2. T.N. Kipf and M. Welling, arXiv 1609.02907v4 (2017)

Table 1. Percentage of molecules containing the following features in the datasets

| Molecular features         | This work | ZINC   |
|----------------------------|-----------|--------|
| at least one aromatic ring | 89.24%    | 92.71% |
| two or more aromatic rings | 79.15%    | 75.47% |
| naphthalene                | 13.26%    | 3.60%  |
| anthracene                 | 2.97%     | 0.28%  |
| quinoline                  | 5.04%     | 2.75%  |

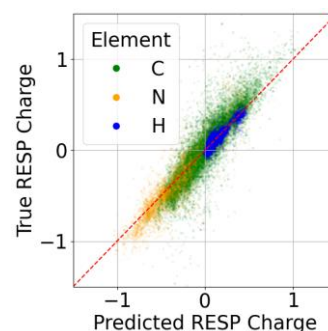


Fig. 1 Prediction results

## 類似分子との差分に基づく HOMO,LUMO エネルギーの機械学習

### Machine Learning of HOMO and LUMO Energies Based on Differences Between Similar Molecules

山形大 ROEL<sup>1</sup>, <sup>○</sup>(B)男澤 泰斗<sup>1</sup>, 佐々木 蓮<sup>1</sup>, 岡田 智悠<sup>1</sup>, 望月 祐樹<sup>1</sup>, 松井弘之<sup>1</sup>

ROEL, Yamagata Univ.<sup>1</sup>, <sup>○</sup>Hiroto Otokozaawa<sup>1</sup>, Ren Sasaki<sup>1</sup>, Yuki Mochizuki<sup>1</sup>,

Tomoharu Okada<sup>1</sup>, Hiroyuki Matsui<sup>1</sup>

E-mail: [t221198@st.yamagata-u.ac.jp](mailto:t221198@st.yamagata-u.ac.jp), [h-matsui@yz.yamagata-u.ac.jp](mailto:h-matsui@yz.yamagata-u.ac.jp)

【緒言】近年、機械学習を活用したマテリアルズ・インフォマティクスにより、材料開発の効率化が進められている。当研究室では、HOMO/LUMO エネルギーの密度汎関数理論 (DFT) 計算値を機械学習で高速予測する分子設計アプリ「YU canvas」を開発してきたが、予測精度に課題があった。そこで本研究では、DFT 計算値が既知の類似分子との差分に基づく、より高精度な予測モデルの構築を試みた。

【方法】(1) 分子全体の HOMO/LUMO を直接予測するモデル、(2) 類似分子との差分を予測し、類似分子の DFT 計算値に加算するモデルの二系統からなり、両者のアンサンブル平均を最終出力とした。それぞれにおいて、グラフ畳み込みネットワーク (GCN) と mordred 記述子を統合したモデルを構築した。学習に用いるデータセットとして、Cambridge Structural Database (CSD) から信頼性の高い有機分子を抽出し、PM7 による構造最適化の後、B3LYP/6-31G(d)で約 29 万分子の HOMO/LUMO エネルギーを計算した。類似分子として RDKit の Murcko Scaffold 関数で抽出した環状骨格を用いたが、その類似分子がデータセット内に存在しない場合には、最大環構造などで代替した。適当な類似分子が見つからなかった分子を除き、最終的に 179,885 分子を使用した。GCN のノード特徴量には元素や芳香環フラグなど 8 種、エッジ特徴量には結合タイプなど 3 種を用い、mordred 記述子は対象分子と類似分子それぞれに対して各 1020 次元で計算し、その差分も使用した。

【結果】平均絶対誤差は直接予測で 0.133 eV、差分予測で 0.135 eV、アンサンブル予測で 0.124 eV であった。平均絶対誤差、平均二乗誤差、決定係数のいずれにおいてもアンサンブル予測が最も優れ、全体として高精度かつ安定した予測性能を示した。アンサンブル平均による精度向上は、直接予測と差分予測の誤差要因が互いに異なるためと考えられる。

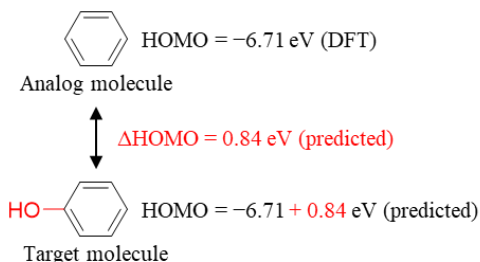


Fig1. Similarity-Based Differential Prediction

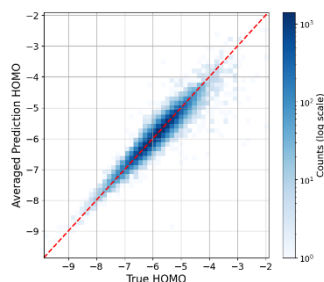


Fig2. Ensemble Prediction of HOMO

## 高速量子化学計算のための回転同変行列の線形結合に基づく ハミルトニアン行列の機械学習

### Machine Learning of Hamiltonian Matrices by Linear Combination of Rotation-Equivariant Matrices for Rapid Quantum Chemistry Calculations

山形大 ROEL<sup>1</sup>, °(B)遠藤 仁太郎<sup>1</sup>, 尾沢昂輝<sup>1</sup>, 松井弘之<sup>1</sup>

ROEL, Yamagata Univ.<sup>1</sup>, °Jintaro Endo<sup>1</sup>, Koki Ozawa<sup>1</sup>, Hiroyuki Matsui<sup>1</sup>

E-mail: [t221141@st.yamagata-u.ac.jp](mailto:t221141@st.yamagata-u.ac.jp), [h-matsui@yz.yamagata-u.ac.jp](mailto:h-matsui@yz.yamagata-u.ac.jp)

【緒言】近年、機械学習によるハミルトニアン行列の予測に関する研究が盛んに行われている。機械学習によってハミルトニアン行列を高速かつ高精度に予測できれば、量子化学計算の計算コストを大幅に削減でき、大規模分子系や動的シミュレーションへの応用にもつながる。その際、座標軸を回転させるとハミルトニアン行列は変化するが、固有エネルギー等は変化しないという回転同変性に注意する必要がある。本研究では、5つの回転同変行列(重なり行列  $S$  および  $S^2$ 、運動エネルギー行列  $K$ 、核電子相互作用行列  $V$ 、単位行列  $I$ )の線形結合としてハミルトニアン行列を表現し、その係数を機械学習モデルによって出力することで、ハミルトニアン行列の予測を試みた。

【方法】本研究では、QH9 データセット<sup>1</sup>(QM9 に基づいたハミルトニアン行列のデータセット)を用いた。Python ライブラリ GBasis を用いて  $S$ 、 $K$ 、 $V$  を計算した。主量子数と方位量子数が同じ基底群( $2p_x$ ,  $2p_y$ ,  $2p_z$  など)をまとめた小行列毎に区切り、それぞれの小行列を異なる値でスカラー倍をしても同変性は保たれるため、このスカラー値を機械学習によって最適化することで、真のハミルトニアン行列をよく再現する行列を生成した。機械学習モデルには SchNet<sup>2</sup>を用い、原子の種類と座標を入力、小行列ごとの係数を出力とした。原子ごとの特徴量を抽出し、その特徴量を基底群毎に分割することで、小行列ごとの係数を出力するための特徴量を生成した。

【結果】Fig. 1 にジオキソラン骨格を有する分子( $C_6H_9NO_2$ )の教師データのハミルトニアン行列、予測したハミルトニアン行列、およびその差分を示す。テストデータに対する平均絶対誤差は 0.00293 Hartree、平均二乗誤差は 0.00726 Hartree、決定係数は 0.997 と高精度で予測することができた。この結果は、回転同変行列の線形結合でハミルトニアン行列をよく再現できることを示している。

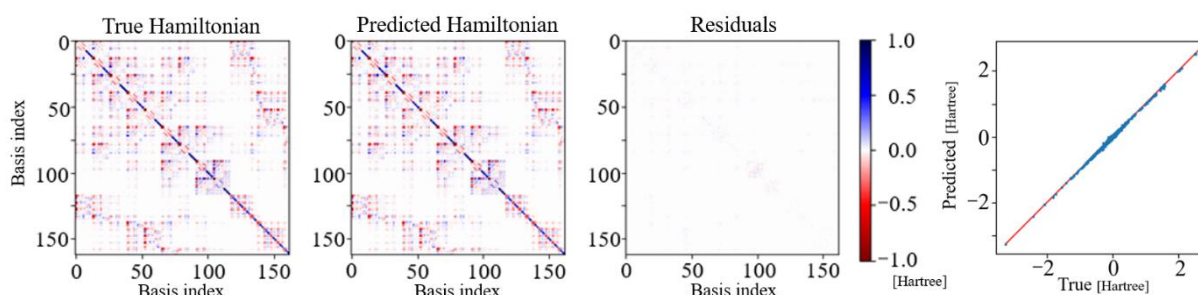


Fig 1 . Comparison between the predicted and reference Hamiltonian matrices

<sup>1</sup> H. Yu et al., arXiv:2306.09549 (2023).

<sup>2</sup> K. T. Schütt et al., arXiv:1706.08566 (2017).

# 機械学習分子動力学シミュレーションによる チタニアシリカ膜でのガス分子の拡散挙動の解析

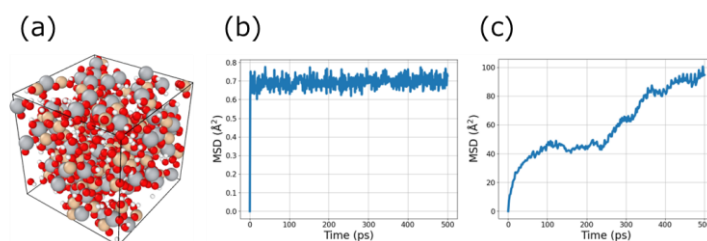
## Analysis of Gas Molecule Diffusion Behavior in Titanium-Silica Membranes Using Machine-Learning Molecular Dynamics Simulations

富士通株式会社<sup>1</sup>, 神戸大学<sup>2</sup>, <sup>○</sup>山崎 周<sup>1</sup>, 吉本 勇太<sup>1</sup>, 五社 一心<sup>2</sup>, 藤澤 拓<sup>2</sup>, 松田 友也<sup>1</sup>,  
松村 直樹<sup>1</sup>, 岩崎 有登<sup>1</sup>, 井上 淳樹<sup>2</sup>, 吉岡 朋久<sup>2</sup>, 川口 博<sup>2</sup>, 坂井 靖文<sup>1</sup>  
Fujitsu Ltd.<sup>1</sup>, Kobe University<sup>2</sup>, <sup>○</sup>Meguru Yamazaki<sup>1</sup>, Yuta Yoshimoto<sup>1</sup>, Isshin Gosha<sup>2</sup>,  
Taku Fujisawa<sup>2</sup>, Tomoya Matsuda<sup>2</sup>, Naoki Matsumura<sup>1</sup>, Yuto Iwasaki<sup>1</sup>, Atsuki Inoue<sup>2</sup>,  
Tomohisa Yoshioka<sup>2</sup>, Hiroshi Kawaguchi<sup>2</sup>, Yasufumi Sakai<sup>1</sup>

E-mail: yamazaki.meguru@fujitsu.com

ニューラルネットワークポテンシャル (NNP) を用いた分子動力学 (MD) シミュレーションは、高精度かつ大規模な系の解析を可能にする手法として期待されている。しかしながら、大規模系は原子間の相互作用が複雑化することから NNP が各原子に働くの力を高精度に予測することが難しく、安定な MD シミュレーションを実現するための、質の高い NNP を構築することは困難である。これを解決する手法として、能動学習を用いた NNP の構築フレームワークが提案されている[Zhang et al., *Comput. Phys. Commun.* 253, 107206 (2020), Matsumura et al., *J. Chem. Theory Comput.* 21, Issue 8 (2025)]。これらのフレームワークでは、訓練データの多様性を確保することで、モデルの安定性を向上させている。しかし、訓練データの多様性を確保するだけでは、データの質やモデルの表現力に課題が残る場合がある。

本研究では、能動学習を用いた NNP を構築フレームワークである GeNNIP4MD[Matsumura et al., arXiv:2411.17191 (2024)]を用いてチタニアシリカ膜とガス分子(水素、メタン)の訓練データを構築した。このデータに対して、不要な構造を取り除く原子間の距離ベースのフィルタリングを適用してデータ品質を向上させた。この訓練データを用いて、ガス分子の表現力を高める手法を適用して NNP を訓練した。訓練された NNP を用いて MD シミュレーションを実行した結果、4 ns の安定したシミュレーションを実行することができた。これにより、図 1(b), (c)に示すように従来の NNP では困難であった水素分子の拡散挙動を再現することができ、自己拡散係数を求めることができた。



**Figure 1.** (a) Titanium-Silica Membrane used for MD simulation, which has 960 atoms. (b) MSD calculated from trajectory of MD simulation using NNP model trained without data filtering. (c) MSD calculated from trajectory of MD simulation using NNP model trained with data filtering.



# 機械学習分子動力学シミュレーションによる SrTiO<sub>3</sub> の酸素欠陥の形成エネルギーの予測

## Prediction of Formation Energy of Oxygen Defects in SrTiO<sub>3</sub>

### Using Machine-Learning Molecular Dynamics Simulations

富士通<sup>1</sup>, 神戸大工<sup>2</sup> °西口 和孝<sup>1,2</sup>, 山本 峻太<sup>2</sup>, 山崎 周<sup>1</sup>,

松村 直樹<sup>1</sup>, 吉本 勇太<sup>1</sup>, 天能 精一郎<sup>2</sup>, 坂井 靖文<sup>1</sup>

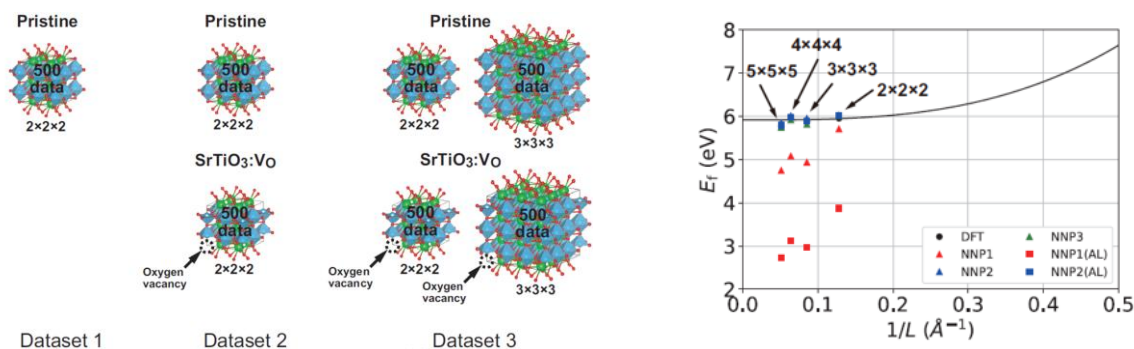
Fujitsu Ltd.<sup>1</sup>, Kobe Univ.<sup>2</sup>, °Kazutaka Nishiguchi<sup>1,2</sup>, Ryota Yamamoto<sup>2</sup>, Meguru Yamazaki<sup>1</sup>,

Naoki Matsumura<sup>1</sup>, Yuta Yoshimoto<sup>1</sup>, Seiichiro L. Ten-no<sup>2</sup>, Yasufumi Sakai<sup>2</sup>

E-mail: k.nishiguchi@fujitsu.com

ペロブスカイト酸化物 SrTiO<sub>3</sub> の酸素欠陥は n 型ドーピングとして作用することが知られており、SrTiO<sub>3</sub> の機能物性を大きく左右するとともに、半導体中の点欠陥の典型例として注目されてきた。一方で、酸素欠陥のような点欠陥を記述するためにはスーパーセルモデルを用いた大規模系の数値計算が必要となり、密度汎関数理論 (DFT) に基づく第一原理計算では多大な計算コストが要求される。また、このような系の分子動力学 (MD) を扱う場合、第一原理 MD シミュレーションは計算コストが高く、古典的 MD シミュレーションは低計算コストであるが精度が限られるため、酸素欠陥の拡散などの動力学解析は非常に困難となる。これに対して、近年ではニューラルネットワークポテンシャル (NNP) を用いた MD シミュレーション (NPP-MD) が注目されており、高精度かつ低計算コストで大規模系の長時間シミュレーションの実行が期待されている。

本研究では、能動学習(AL) を用いた NNP 構築フレームワークである GeNNIP4MD [Matsumura *et al.*, J. Chem. Theory Comput. **2025**, 21, 3832-3846.] を用いて、SrTiO<sub>3</sub> の酸素欠陥の形成エネルギーを NNP-MD シミュレーションを用いて予測した。精度検証のため、3つの異なるデータセット(Dataset 1, 2, 3) から NNP モデル(NNP 1, 2, 3) を構築した(図 1(左))。NNP-MD で予測された酸素欠陥の形成エネルギーは、酸素欠陥のデータセットが考慮されている場合、DFT 計算の結果を高精度に再現した。特に、DFT 計算が困難なスーパーセルサイズ(5×5×5)においても、NNP-MD は DFT の外挿値を良く再現する(図 1(右))。[Nishiguchi *et al.*, arXiv:2506.09372.]



**Figure 1.** (Left) Dataset 1, 2, and 3. (Right) NNP-MD prediction of the formation energy for 2×2×2, 3×3×3, 4×4×4, and 5×5×5 supercell sizes. The black line is the extrapolation of the DFT values for 2×2×2, 3×3×3, and 4×4×4 supercell sizes.



# PFP descriptorsとShapley値を組み合わせた原子レベルでの予測の解釈

## Atomic Level Explanation via PFP Descriptors and Shapley Values

(株) Preferred Computational Chemistry<sup>1</sup>, °張 凡<sup>1</sup>

Preferred Computational Chemistry, Inc.<sup>1</sup>, °Bon Cho<sup>1</sup>

E-mail: boncho@pfcc.co.jp

機械学習を用いた材料開発では、機械学習モデルのブラックボックス性が新たな科学的知見の獲得を阻害している。この課題に対し、本研究では、ユニバーサルな機械学習原子間ポテンシャルであるPFP<sup>1)</sup>から得られる原子特徴量PFP descriptorsとShapley値を組み合わせた解釈手法を提案する。PFP descriptorsは、大量のデータを学習したPFPから計算される原子特徴量であるため、学習データが限られていても高い予測精度が期待できる<sup>2)</sup>。それに加え、Shapley値を応用し、原子をマスクした際の予測値の変化から、物性に対する各原子の寄与度を計算できる。これは材料設計に有用な解釈性を与える手法であると考えられる。

本手法の有効性を、有機分子の水溶解度データセット<sup>3)</sup>を用いて検証した。まず、原子ごとに計算される256次元のPFP descriptorsを分子全体で集約（平均・合計）した場合と、半径2で2048bitにhashしたカウントベースのMorgan Fingerprint(CMFP)を特徴量に用いた場合の予測精度を比較した。評価は5fold CVを行い、全予測値と実測値のMAEを用いて比較した (Table 1)。その結果、PFP descriptorsがCMFPを上回る精度を示した。次に全データのうち一部をガウス過程回帰に学習させ、学習していない分子に対して予測値と原子をマスクすることでShapley値を計算した。これによって各原子が物性に対して、どのくらい寄与するか計算できる。その結果、水酸基が溶解性の向上に寄与し、メチル基が低下に寄与するという化学的知見と一致する結果となり、本手法の有用性を示す結果となった(Figure 1)。

Table 1. Comparison of MAE for different descriptors and models.

|            | Ridge       | PLS         | Random Forest | Gaussian Process |
|------------|-------------|-------------|---------------|------------------|
| CMFP       | 0.87        | 0.79        | 0.68          | 0.76             |
| PFP (mean) | 0.60        | 0.65        | 0.66          | 0.47             |
| PFP (sum)  | <b>0.58</b> | <b>0.62</b> | <b>0.59</b>   | <b>0.45</b>      |

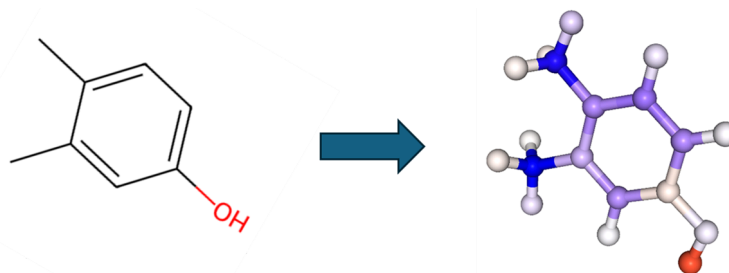


Figure 1. Atomic level explanation via PFP descriptors and Shapley values. Red and blue indicate positive and negative contributions to the property, respectively.

### References:

- 1) S. Takamoto *et al.*, *Nat. Commun.* **13**, 2991 (2022).
- 2) Z. Mao *et al.*, *npj Comput. Mater.* **10**, 265 (2024).
- 3) J. S. Delaney, *J. Chem. Inf. Comput. Sci.* **44**, 1000 (2004).

# 機械学習を用いた GaN-HEMT 用アモルファス SiN 薄膜成膜プロセスの最適化 Machine Learning for High-Performance GaN-HEMTs: Optimizing amorphous SiN Thin Film Deposition

富士通株式会社, 栗林壮太郎, 本間健司, 丸尾昭人, 樋口博之,  
實宝秀幸, 山田敦史, 多木俊裕

Fujitsu Limited, Sotaro Kuribayashi, Kenji Homma, Akito Maruo, Hiroyuki Higuchi,  
Hideyuki Jippo, Atsushi Yamada, and Toshihiro Ohki

E-mail: s\_kuribayashi@fujitsu.com

## 【緒言】

機械学習に基づく材料最適化はブラックボックスプロセスとして行われることが多く、最適化されたプロセス条件の背後にある理論的根拠の説明を困難にしている。本研究では、最適な GaN 高電子移動度トランジスタデバイスのためのアモルファス窒化シリコン(a-SiN)堆積パラメータを、相関分析と因果発見による解釈可能性も検討して決定した。

## 【実験】

成膜装置の設定値(SiH<sub>4</sub> 体積流量( $V_{SiH_4}$ ), 印加高周波パワー( $P_{RF}$ ), ポストアニール温度( $T_{anneal}$ ), アニール時間( $t_{anneal}$ ))を変えて成膜したそれぞれの a-SiN 薄膜に対し、特性値(破壊電圧( $V_{BD}$ )), ヒステリシス電位差( $\Delta V$ ), シート抵抗( $R_{sh}$ ))を測定した。次に、成膜装置の設定値を独立変数、a-SiN 薄膜の特性値を目的変数とするモデルをガウス過程(GP)によって求め、遺伝的アルゴリズム(GA)によって、最適な成膜条件を決定した。また、a-SiN 薄膜密度( $\rho$ ), 二次元電子ガス移動度( $\mu_{2DEG}$ ), 結合モードなどの測定データも中間変数として追加し、独立変数→中間変数→目的変数の関係の解釈可能性を相関分析と因果発見によって検討した。

## 【結果】

GP と GA により、最適な成膜条件を  $V_{SiH_4}$ : 1.45-1.65 SCCM,  $P_{RF}$ : 58-62 W,  $T_{anneal}$ : 700°C,  $t_{anneal}$ : 1-5 min. と決定した。

相関分析と因果発見により、例えば  $V_{SiH_4}$  の増大が  $\rho$  や  $\mu_{2DEG}$  も伴って  $V_{BD}$  を減少させる事が分かった。なお、Figure 1 は因果発見が出力した因果関係のグラフである。

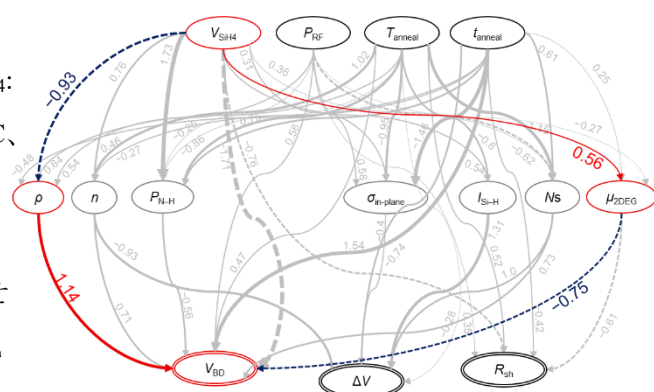


Figure 1 Graph of Causal relationships

## 【参考文献】

1: K. Homma, A. Maruo, S. Kuribayashi, H. Higuchi, H. Jippo, A. Yamada, and T. Ohki "Optimizing and Understanding amorphous SiN Thin Film Deposition for GaN-High-Electron-Mobility Transistors Using Machine Learning and Causal Discovery", *ACS Appl. Electron. Mater.* **2025**, 7, 11, 4961–4971.

# AI モデルによるプロセス評価値を用いた半導体乾燥プロセス最適化

## Optimization of Semiconductor Drying Process Using AI-Driven Process Evaluation Metrics

株式会社 SCREEN ホールディングス<sup>1</sup>,

○久保 友輔<sup>1</sup>, 中野 佑宇唯<sup>1</sup>, 村島 広衛<sup>1</sup>, 國枝 省吾<sup>1</sup>, 佐々木 悠太<sup>1</sup>, 塙 洋祐<sup>1</sup>

SCREEN Holdings Co., Ltd.<sup>1</sup>, ○Yusuke Kubo<sup>1</sup>, Yuui Nakano<sup>1</sup>,

Hiromori Murashima<sup>1</sup>, Shogo Kunieda<sup>1</sup>, Yuta Sasaki<sup>1</sup>, Yosuke Hanawa<sup>1</sup>

E-mail: yus.kubo@screen.co.jp

従来の半導体洗浄後の乾燥技術開発は、微細パターンの倒壊抑制性能を定量的に評価するために、プロセス評価のたびに、電子顕微鏡（SEM）で基板の微細パターンを撮影し、パターンの倒壊率を計測することでプロセスの定量評価を行っていた。しかしながら、SEM 像の撮影には長時間を要してしまうという課題がある。プロセスインフォマティクス技術を用いて乾燥プロセスのパラメータを最適化する場合、十分な性能のプロセス条件を発見するまでに、数十回～百回程度のプロセス実験を行う必要がある。そのような場合、SEM 撮像がボトルネックとなり、実験回数が限られてしまいプロセス最適化後の倒壊抑制性能が不十分となる課題があった。

そこで、本研究では SEM と比べて高速に撮像可能な光学顕微鏡（図 1 左図）で基板を撮影し、事前に学習した画像認識 AI によって干涉縞の強弱から倒壊率を予測することで、プロセス最適化の実験サイクルを効率化する。

図 1 右図は、テストデータに対する AI による倒壊率の予測結果である。1 点が 1 つのプロセス条件に対応し、複数枚の基板画像から予測した倒壊率の平均値をそのプロセスの評価値として予測性能を評価した。RMSE 0.145, 決定係数 0.831 と高精度で倒壊率を予測できることを確認した。

また、この AI の評価値をベイズ最適化の目的関数として使い、実際の乾燥プロセス最適化の実験も行った。全 75 回のプロセス実験を行い、そのうち AI の予測値が 30%を下回る 10 条件をサンプリングし、SEM による精密な倒壊率評価を行った。その結果、従来の最良条件である倒壊率 1.39%を下回る 0.58%のプロセス条件を発見することができた。本手法により、従来の評価方法と比べて、約 68%の作業時間を短縮し、プロセス品質を向上させることができた。

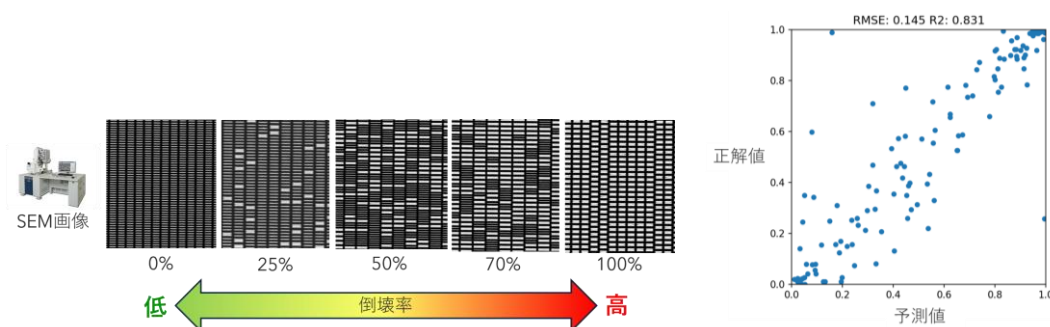


Fig. 1. Left: Relation of collapse rate and SEM image appearance.

Right: Predicted collapse rate results for test data. This sample is provided by Imec.

**パーシステントホモロジーで可視化する  
イオンビーム照射による半導体微細構造のプロセスインフォマティクス  
Process Informatics of Ion Beam Irradiated Semiconductor Nanostructures  
Visualized by Persistent Homology**

高知工科大学<sup>1</sup>, 高知工業高等専門学校<sup>2</sup>, 岡山大学<sup>3</sup>

<sup>○</sup>(M1) 走出 和寛<sup>1</sup>, 宮田 滯<sup>1</sup>, 新田 紀子<sup>1</sup>, 大石 脩人<sup>2</sup>, 大林 一平<sup>3</sup>

Kochi University of Technology<sup>1</sup>,

National Institute of Technology, Kochi College<sup>2</sup>, Okayama University<sup>3</sup>

<sup>○</sup>Kazuhiro Hashiride<sup>1</sup>, Rei Miyata<sup>1</sup>, Noriko Nitta<sup>1</sup>, Naoto Oishi<sup>2</sup>, Ippei Obayashi<sup>3</sup>

E-mail: hashiride02@gmail.com

近年、材料開発における製造プロセス最適化を行うために、データ解析と情報科学を駆使するプロセスインフォマティクスが急速に発展している。本研究では、半導体材料にイオンビームを照射した際に形成される微細構造をパーシステントホモロジーを用いて解析を行い、照射条件との関連を明らかにした。パーシステントホモロジーとは、トポロジーを考慮した形態を捉えることの出来る数学的手法である。原子配列の違いを捉えた構造解析や材料物性予測などへの応用がなされている。同手法を用いて、イオンビーム照射誘起微細構造の特徴量埋め込み画像解析を行った。

試料は単結晶 Ge を用いた。基板表面に対してイオンビームによる重畳照射を行うことでポーラス構造を作製した。照射には FIB (Focused Ion Beam, FEI, QUANTA3D200i) を用いた。この際、イオン種は Ga<sup>+</sup>、加速電圧は 30 kV、照射電流値は 526–580 pA である。照射量は 2 つのグループに分けて実験した。A グループ 1 回目: 照射角度 60°,  $1 \times 10^{20}$  ions/m<sup>2</sup>、2 回目: 照射角度 0°,  $2.5 \times 10^{19}$ – $1 \times 10^{21}$  ions/m<sup>2</sup>、B グループ 1 回目: 照射角度 60°,  $2.5 \times 10^{19}$ – $1 \times 10^{21}$  ions/m<sup>2</sup>、二回目: 照射角度 0°,  $1 \times 10^{20}$  ions/m<sup>2</sup> である。形成された構造評価には SEM (Scanning Electron Microscope, HITACHI, SU8020) を用いて表面観察を行った。解析手法は、まず得られた SEM 画像に対して前処理として OpenCV 4.11.0 によって閾値 100 で二値化を行った。HomCloud 4.7.0 によってパーシステントホモロジー解析を行った。解析には連結成分の 0 次と穴の 1 次を用いた。その後、scikit-learn 1.6.1 を使って、得られた埋め込みベクトルで主成分分析を行った。

Ge 表面にイオンビームを照射すると、ポーラス構造は照射角度 0° では穴、60° では線状の形態を示す。パーシステントホモロジーの 0 次における主成分分析の第一主成分を解析したところ、照射条件によって穴を示すか線状を示すかを表すことができた。この時の第一主成分は、A グループで分散が小さく、B グループでは分散が大きかった。これは、照射量の違いに起因する Ge 表面のポーラス構造の変化を捉えることが出来たと考えている。

# 偏光情報と温度変化を基にした深層学習による 複屈折像の解析手法の開発

鹿児島大院理工<sup>A</sup>, 鈴鹿高専<sup>B</sup>  
真中 浩貴<sup>A</sup>, 本田 宗一郎<sup>A</sup>, 三浦 陽子<sup>B</sup>

Image Analysis of Polarized Light Microscopy Using Convolutional Neural Networks  
Kagoshima Univ.<sup>A</sup> and NIT, Suzuka Col.<sup>B</sup>  
Hirotaka Manaka<sup>A</sup>, Soichiro Honda<sup>A</sup>, Yoko Miura<sup>B</sup>

これまで本研究グループでは、複屈折イメージング装置を活用して、様々な物質における相転移現象を観察してきた。本装置の原理は偏光顕微鏡と同様であり、偏光させた光を試料に入射し、複屈折によって変化した偏光状態の空間分布を観察する。図 1 に示すように、通常の顕微鏡では二つの点光源からの波が干渉し合いながら拡がる場合、両者を識別できる最小間隔を光学的分解能と定義する。しかし偏光顕微鏡では、空間的に隣接した領域の偏光状態が重ね合わさることで、観測される偏光状態自体が変質してしまう問題がある。このような偏光の空間的混合を前提とした画像解析手法は、現在のところ一般的ではなく、従来の空間分解能ベースの画像解析のやり方とは異なるアプローチが求められる。

本発表では、温度変化に伴って連続的に取得された複屈折画像に対し、偏光状態の温度依存性の情報も保持したまま解析する新たな手法を提案する。図 1 に示すような偏光状態の重なり合いに対しては、超解像などの空間補完的なアプローチも考えられるが、それらは主に画素間の連続性を補う手法であり、全体的な偏光構造の抽出には適さない。そこで本研究では、偏光を直接分離するのではなく、畳み込みニューラルネットワーク(CNN)を用いて隣接画素の情報を積極的に統合し、画像中に内在する偏光状態の分布パターンを抽出することを試みた。具体的には、空間受容野(SRF)のサイズを段階的に変化させ、その出力に対してクラスタリングを適用することで、SRF サイズに依存しない安定した偏光状態をグループ化できる事を目指した。これにより物理的に意味のある偏光状態の分布像が可視化できる。

実際の解析では、応力誘起強誘電体  $\text{SrTiO}_3$  に対して 130.0 K~14.1 K の範囲で取得した 233 枚の複屈折画像に対して、 $(x, y)$  の画素位置の情報に加えて温度情報も加えることで、3D-CNN によるオートエンコードを構築した。さらにこの出力に対して、温度変化の傾向が類似する画素同士をグループ化するために「温度系列フォレスト」を用いてクラスタリングを行った。図 2 に示す最終的なクラスタリング結果からは、強誘電分域に起因するストライプ構造が抽出できた[1]。特に赤色の領域では、応力が集中しているため局所的に強誘電転移温度が上昇している領域とも一致していることから、本手法が単なる画像処理にとどまらず、物理的整合性を伴った偏光状態の抽出手法として有効であることが確認できた。

[1] K. Toyoda 他: Sci. Technol. Adv. Mater Meth. 3, 2278322 (2023), H. Manaka 他: Sci. Technol. Adv. Mater Meth. 4, 2342234 (2024).

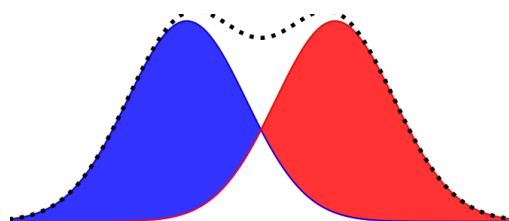


図1. 顕微鏡における分解能の概念

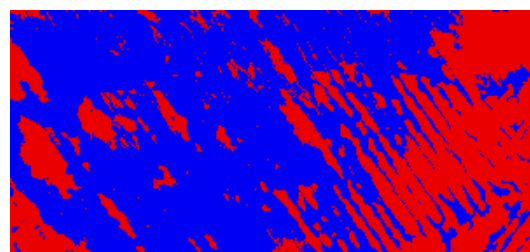


図2. 温度系列フォレストを用いたクラスタリングの結果



## LLM を使った材料開発(III)：科学論文からの材料データ抽出

## Materials Development with LLM (III): Extraction of Materials Data from Scientific Articles

NIMS ○岡 博之, 石井 真史

NIMS ○Hiroyuki Oka, Masashi Ishii

E-mail: OKA.Hiroyuki@nims.go.jp

## 1. 序

AI の普及に伴って、既存データを活用した材料開発、マテリアルズ・インフォマティクス(MI)が盛んになっている。特に、最近では生成 AI と呼ばれる大規模言語モデル(LLM)の利用が盛んで、文書からのデータ抽出にも用いられている。我々は以前にオープンソース LLM の 1 つである Gemma 2 (Google) を用いて、米国特許の本文から磁石製造条件の抽出を検討した<sup>[1]</sup>。検討では、米国特許はいくつかの製造方法がセクション毎に分けて書かれているため、各セクションに対して抽出対象の製造条件を指定したプロンプトを与えることで、高精度で抽出できることが分かった。このため、次はこれを材料論文からのデータ抽出に使うことを考えている。ただ、論文は特許と違って文書構成が多様なため、LLM に投入する本文はセクション単位だけでは不十分で、また、具体的な名前の無い試料の認識や図からのデータ抽出が必要なこともある。そのため、現在、論文用に LLM を利用する方法を確立中である。特に図は画像であることから、テキストしか扱えない Gemma 2 では処理できない。しかし、最近リリースされた Gemma 3 では画像も扱えるため、これを利用して検討している。本発表では、これらの結果について報告する。

## 2. 検討抽出方法および結果

LLM には Gemma 3 の 12B のモデルを用いている。LLM に投入する論文内容は本文と図とし、それぞれプレーンテキストと画像形式で PDF から抽出している。尚、図は散布図のみを対象とし、表は現時点では扱っていない。抽出した本文と図を用いて、次の 3 つのステップでデータ抽出を検討している：①本文からの試料名の抽出、②得られた各試料の作成条件と物性値の抽出、③散布図からのデータ抽出。①では、試料名の記載がない試料についても適当な名前を内容から推測して付与するようにプロンプトで指示している。例えば、図 1(a)に示す結果が得られた論文では、FePt 薄膜試料を室温と 300°C で作成しており、さらにその後、いくつかの温度でアニーリングした試料も作成しているが、後者の試料名は具体的には記載されていない。しかし、LLM は図 1(a)のよう

に適当に試料名を付与して抽出できるため、①では LLM のテキスト生成能力を利用している。次に②では、①で得られた各試料の作成条件と物性値の抽出を米国特許と同様の方法で検討している。ここでの詳細な記述は省略する。最後に③では、Gemma 3 で散布図からのデータ抽出を検討している。図 1(b)に示すように凡例ごとにプロットの x, y 値を出力するようにプロンプトで指示している。現時点では正しい x, y 値が認識できていないため、画像上でのプロットの位置情報などの補助データの入力が必要と考えている。

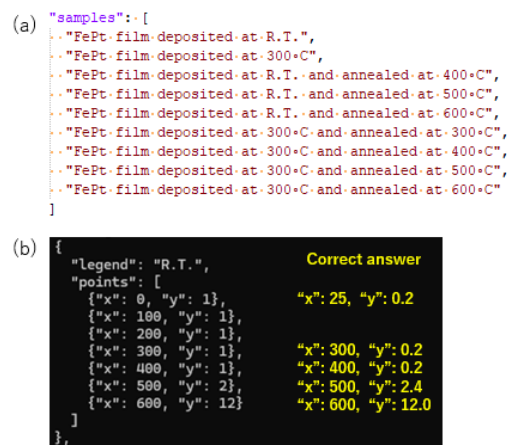


図 1 抽出された(a)試料名と(b)散布図中のデータ。使用論文の DOI: 10.1143/JJAP.40.L1367.

## 3. まとめ

上記のようにして有用な抽出方法を検討中であり、発表当日は最新の結果を報告する。

## 謝辞

本研究は、データ創出・活用型マテリアル研究開発プロジェクト事業 (Grant Number: JPMXP1122715503) の助成で行っています。ここに感謝の意を表します。

- [1] 岡博之他, LLM を使った材料開発(II): 米国特許からの NdFeB 磁石データの抽出, 応物 72 春, 理科大野田キャンパス(野田市), 2025.3.14-17.

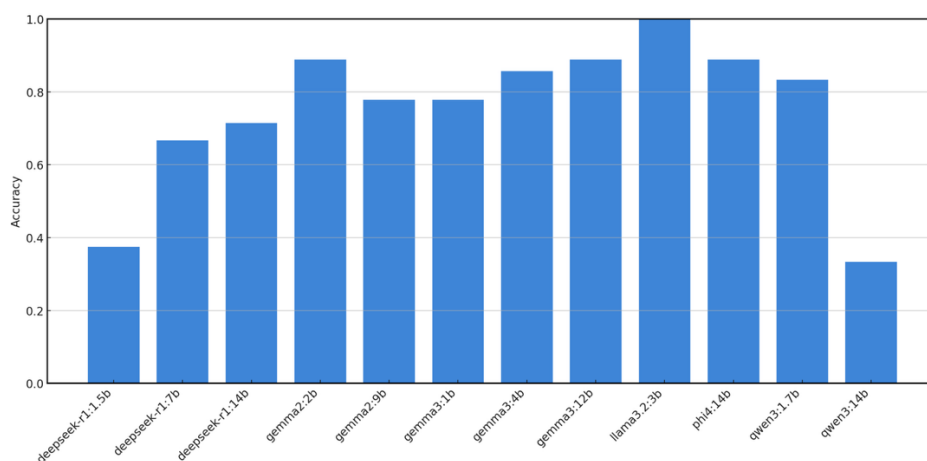
## ローカル LLM を用いた論文からの成長層構造抽出

### Extraction of Growth Layer Structures from Research Papers Using a Local LLM

東大ナノ量子, °權 晋寛 NanoQuine, Univ., °Jinkwan Kwoen

E-mail: jkkwoen@iis.u-tokyo.ac.jp

LLM (Large Language Model、大規模言語モデル) は、近年自然言語処理の分野で注目されている技術であり、文書生成、文書分類、質問応答、要約、翻訳など、さまざまなタスクにおいて優れた性能を示している。また、LLM は科学論文からの情報抽出にも応用されており、特に材料科学分野では、合成条件や物性データなどの構造化情報を自動的に抽出する用途で活用されている [1, 2]。LLM は大規模なデータセットで学習されたニューラルネットワークであるが、一般的な LLM は特定のドメインに特化した知識抽出にはプロンプトエンジニアリングが必要である。本研究では、半導体および材料科学といった専門分野において、論文に記載された成長層構造に関する情報を抽出することを目的とし、ローカル LLM を活用する手法を提案する。具体的には、論文中に記載された基板の材料や面方位、層の材料、厚さ、積層順序など、成長層構造に関する要素をローカル LLM を用いて自動的に抽出する手法を構築する。本研究ではノートパソコンのような一般的な計算資源でも実行可能な軽量 LLM を活用し遂行した。Figure は、さまざまな軽量言語モデルに対して基板材料の識別精度を比較した結果の一例である。特に gemma2:2b および llama3:8b は、より大規模なモデルを上回る精度を達成しており、軽量モデルでありながら高い性能を示した。



**Figure** Substrate material identification performance of various lightweight language models.

[1] Dunn, A. et al. Structured information extraction from complex scientific text with fine-tuned large language models. Preprint at <https://doi.org/10.48550/arXiv.2212.05238> (2022).

[2] Verma, A. K. et al. Structured Extraction of Process Structure Properties Relationships in Materials Science. Preprint at <https://doi.org/10.48550/arXiv.2504.03979> (2025).

## LLM とシミュレータを用いた材料開発支援システムの構築に向けて

Developing a Support System for Materials Scientists through LLM–Simulator Synergy

SB Intuitions<sup>1</sup>, NIMS<sup>2</sup>

○Jun Jin Choong<sup>1</sup>, ○小澤 圭右<sup>1</sup>, 石黒 勝彦<sup>1</sup>, 三好 利昇<sup>1</sup>, 井尻 善久<sup>1</sup>, 出村 雅彦<sup>2</sup>, 源 聡<sup>2</sup>, 門平 卓也<sup>2</sup>

SB Intuitions<sup>1</sup>, NIMS<sup>2</sup>, ○Jun Jin Choong<sup>1</sup>, Keisuke Ozawa<sup>1</sup>, Katsuhiko Ishiguro<sup>1</sup>,

Toshinori Miyoshi<sup>1</sup>, Yoshihisa Ijiri<sup>1</sup>, Masahiko Demura<sup>2</sup>, Satoshi Minamoto<sup>2</sup>, Takuya Kadohira<sup>2</sup>

E-mail: [jun-jin.choong@sbintuitions.co.jp](mailto:jun-jin.choong@sbintuitions.co.jp), [keisuke.ozawa@sbintuitions.co.jp](mailto:keisuke.ozawa@sbintuitions.co.jp)

材料開発において専門家の知識と経験が中核的な役割を果たしている。知識や経験に基づく判断は初期案の創出とその改善における要だが、同時に、物理現象を一定の近似のもと再現可能とするシミュレータを適切に活用し、その結果を知識と比較・統合または再検証することも求められる。しかし、シミュレータの活用にあたっては、現象に対する知識に加え、ユーザの目的に適したシミュレータの選定と設定の技量も求められ、実務活用への障壁となっている。

近年の大規模言語モデル (LLM) の進歩により、自然言語で与えられた目的を機械が自動的に解釈できるようになりつつある。また、知識共有が容易化されることで属人性の解消も期待される。LLM の活用により材料・実験設計を効率化、加速できる可能性がある一方、ハルシネーションや物理的整合性の欠如といった、LLM に内在する問題が依然として残されている。

ユーザ目的を自然言語で柔軟に受け取りつつも物理的または数値的再現性を担保できるよう、我々は、LLM の高度な文脈理解能力と、シミュレータが有する精緻な現象再現能力とを、相互補完的に活かす材料開発支援システムの構築を開始した。その初期検討にあつて、材料設計に関する簡単な事例を紹介し、今後の展望に触れる。

Expert knowledge and experience play a central role in decision making processes for materials development. Decision making based on knowledge and experience is essential in both the creation and refinement of initial ideas. Yet, simultaneously, it is also important to leverage simulators that allow reproducing physical phenomena under certain approximations. It is often necessary to compare the results with the knowledge, integrate them, or subject them to revalidation. However, using simulators requires skills in addition to domain knowledge, specifically in selecting and configuring the appropriate simulators for the purpose, which makes it harder for practitioners to master the simulators.

Recent advancements in LLMs have enabled systems to automatically interpret purposes expressed in natural language. This progress has the potential to mitigate difficulty in knowledge sharing and reduce reliance on individual expertise. While LLMs provide a possibility to accelerate materials development, there remain fundamental issues unresolved, including hallucinations and the lack of physical consistency generated by LLMs.

To address these challenges and provide a system that flexibly responds to user instructions while ensuring reproducibility based on physical principles embedded within simulations, we have started to implement a support system for materials scientists that combines the contextual understanding ability of LLMs with the precise reproduction ability of simulators. As the first step, we present a simple use case from materials design and discuss our prospects.

23 合同セッションN「インフォマティクス応用」 | 一般セッション(口頭講演): 23.1 合同セッションN「インフォマティクス応用」

## [9a-S301-1~7] 23.1 合同セッションN「インフォマティクス応用」

[9a-S301-1]

データベース中で外挿領域周辺の物性値を有する材料の設計: 勾配ベース手法の強みと拡散生成モデルの限界

○藤井 亮宏<sup>1</sup>、牛久 祥孝<sup>2</sup>、清水 康司<sup>3</sup>、Lu Anh Khoa Augustin<sup>1,4,3</sup>、渡邊 聡<sup>1</sup> (1.東京大工、2.OMRON SINIC X Corp.、3.産総研、4.物材機構)

[9a-S301-2]

化学情報を組み込んだドメイン変換によるSim2Real転移学習

○矢作 裕太<sup>1,2</sup>、小淵 喜一<sup>1,2</sup>、高坂 文彦<sup>2</sup>、松井 孝太<sup>3</sup> (1.NEC、2.産総研、3.京大医)

[9a-S301-3]

Structure Prediction from Chemical Formula Using Periodic Descriptors

○Yen-Ju Wu<sup>1</sup>, Yibin Xu<sup>1</sup> (1.NIMS)

[9a-S301-4]

解釈性に考慮した機械学習によるCNT繊維の構造—物性関係の可視化

○木村 大輔<sup>1</sup>、田島 奈緒子<sup>1</sup>、岡崎 俊也<sup>1</sup>、室賀 駿<sup>1</sup> (1.産総研)

[9a-S301-5]

BaTiO<sub>3</sub>の構造相転移・誘電特性に対する機械学習ポテンシャルの応用

○平井 大介<sup>1</sup>、金川 朋賢<sup>1</sup>、廣瀬 左京<sup>1</sup> (1.株式会社村田製作所)

[9a-S301-6]

Extreme learning machineによるGeSn混晶半導体の全エネルギー予測

○野田 祐輔<sup>1</sup>、別宮 響<sup>2</sup>、末岡 浩治<sup>3</sup> (1.九工大院情報工、2.岡山県立大院情報系工、3.岡山県立大情報工)

[9a-S301-7]

Stable and Functional Quaternary Heusler Alloys: A Machine Learning-Assisted High-Throughput First-Principles Exploration

○(P)Enda Xiao<sup>1</sup>, Terumasa Tadano<sup>1</sup> (1.NIMS)

データベース中で外挿領域周辺の物性値を有する材料の設計:  
勾配ベース手法の強みと拡散モデルの限界

Designing Materials with Property Values Near the Extrapolation Regions in Databases:  
Strengths of Gradient-Based Optimization and Limitations of Diffusion Models

東京大学<sup>1</sup>, オムロンサイニックス<sup>2</sup>, 産業技術総合研究所<sup>3</sup>, 物質・材料研究機構<sup>4</sup>

○ 藤井亮宏<sup>1</sup>, 牛久祥孝<sup>2</sup>, 清水康司<sup>3</sup>, Anh Khoa Augustin Lu<sup>1,4</sup>, 渡邊聡<sup>1</sup>

The University of Tokyo<sup>1</sup>, OMRON SINIC X Corporation<sup>2</sup>, National Institute of Advanced Industrial Science and Technology (AIST)<sup>3</sup>, National Institute for Materials Science (NIMS)<sup>4</sup>

○ Akihiro Fujii<sup>1</sup>, Yoshitaka Ushiku<sup>2</sup>, Koji Shimizu<sup>3</sup>,

Anh Khoa Augustin Lu<sup>1,4</sup>, Satoshi Watanabe<sup>1</sup>

E-mail: akihiro.fujii@cello.t.u-tokyo.ac.jp

有用な物性を発現する新規材料の設計は極めて重要である。一方で、例えば低損失送電などへの応用が期待される高い超伝導転移温度 ( $T_c$ ) をもつ材料は稀もしくは開発途上であり、材料データベース上でもその存在は限られている。近年急速に発展した機械学習を使った材料設計手法を用いれば、従来よりも広い範囲を探索でき、そのような卓越した物性値を発現する材料を効率的に探索できる。

拡散モデルに代表される深層生成モデルは、今までにない新たな材料を設計する用途で広く使われている。特に CDVAE [1] は物性予測モデルとその勾配を使って、生成を特定の物性値へと誘導できる。しかし、生成モデルは外挿領域周辺の特性よりもデータセット上で頻出する特性へと生成を誘導する傾向がある [2]。このため、データセット上で外挿領域周辺に存在する物性（高い超伝導転移温度をもつ材料など）の生成には適していない。しかし、これらの外挿領域周辺の物性値をもつ材料は重要であるため、生成モデルの代替となる手法の開発が望まれる。

本研究では、我々が以前提案した、組成式や結晶を生成モデルを介さずに直接最適化する勾配ベースの手法 [3,4] を用いて、データセット上で外挿領域に周辺の物性値を発現する材料を提案可能かどうかを検証する。これらの手法は、深層学習モデルとその勾配を用いて組成式または結晶を目標の物性へ向けて最適化する手法である (Figure 1)。この方式では、学習データ分布による拘束が生成モデルよりも小さいため、データベース中で外挿領域周辺の物性値へと組成式や結晶を誘導しやすい。

我々は、データベース上で外挿領域に近い物性値をもつ材料を提案する実験において、開発した勾配ベースの手法と生成モデルを比較した。概念実証実験として、高い超伝導転移温度  $T_c$  を発現する組成式の提案と、高い超伝導転移温度  $T_c$  もしくは高いバンドギャップ値を発現する結晶の提案において、両者を成功率で比較した。生成モデル側には、物性値を外挿領域の方向へ誘導する損失関数 [2] を導入し、生成性能を補強した。それにもかかわらず、我々の提案手法は、生成モデルと誘導を組み合わせた手法よりも顕著に高い成功率を示した (Table 1)。この結果は、外挿領域に近い物性値をもつ材料の提案において、勾配ベース手法が生成モデルに代わる有力な手法となりうることを示唆している。

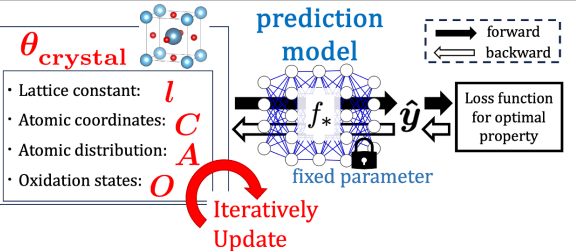


Figure 1: Overview of the proposed method to optimize crystal structures towards targeted properties.

| $T_c$<br>(Percentile) | >5.0 K<br>(>73.3%) | >10.0 K<br>(>90.1%) | >15.0 K<br>(>96.3%) |
|-----------------------|--------------------|---------------------|---------------------|
| CDVAE                 | 0.22 ± 0.07        | 0.03 ± 0.03         | 0.02 ± 0.02         |
| Proposed              | <b>0.92 ± 0.03</b> | <b>0.54 ± 0.02</b>  | <b>0.25 ± 0.02</b>  |

Table 1: Success rates for achieving predicted  $T_c$  above various thresholds.

[1] Xian, T., et al. (2022). *The International Conference on Learning Representations (ICLR) 2022*.  
[2] Um, S., et al. (2024). *The International Conference on Learning Representations (ICLR) 2024*.  
[3] Fujii, A., et al. (2023). *arXiv preprint arXiv:2304.13860*.  
[4] Fujii, A., et al. (2024). *arXiv preprint arXiv:2410.08562*.



## 化学情報を組み込んだドメイン変換による Sim2Real 転移学習

### Sim2Real transfer learning with chemistry-informed domain transformation

NEC<sup>1</sup>、産総研<sup>2</sup>、京大医<sup>3</sup>、

○矢作 裕太<sup>1,2</sup>、小淵喜一<sup>1,2</sup>、高坂文彦<sup>2</sup>、松井孝太

NEC<sup>1</sup>, AIST<sup>2</sup>, Kyoto Univ.,

○Yuta Yahagi<sup>1,2</sup>, Kiichi Obuchi<sup>1,2</sup>, Fumihiko Kosaka<sup>2</sup>, Kota Matsui<sup>3</sup>

E-mail: yuta-yahagi@nec.com

機械学習を用いた材料性能予測はデータ駆動型材料開発において基本的なタスクである。しかしながら、一般に実験データは希少性が高く、学習のためのデータが限られることが大きな課題であった。そこで我々は、大規模第一原理計算データセットを活用することで、少量の実験データでも効率的に予測精度を向上させる転移学習手法を提案した[1]。

本手法は計算から実験へのドメイン変換と、同質ドメイン間の転移学習の2段階からなる(図1)。ドメイン変換においては第一原理計算における微視的な特徴量を実験における巨視的な特徴量へと変換する。変換時に事前知識として化学情報(統計分布、経験則等)を用いることで、異質ドメイン間のデータ変換を可能とする。

本手法の効果を実証するために、触媒実験データおよび計算データを用いて触媒活性予測モデルを作成した。予測性能を評価したところ、転移学習を用いなかった場合と比較して予測精度が1桁程度改善され、正の転移が確認できた。

[1] YY, K. Obuchi, F. Kosaka, and K. Matsui, *Mach. Learn.: Sci. Technol.* **6** 025026 (2025).

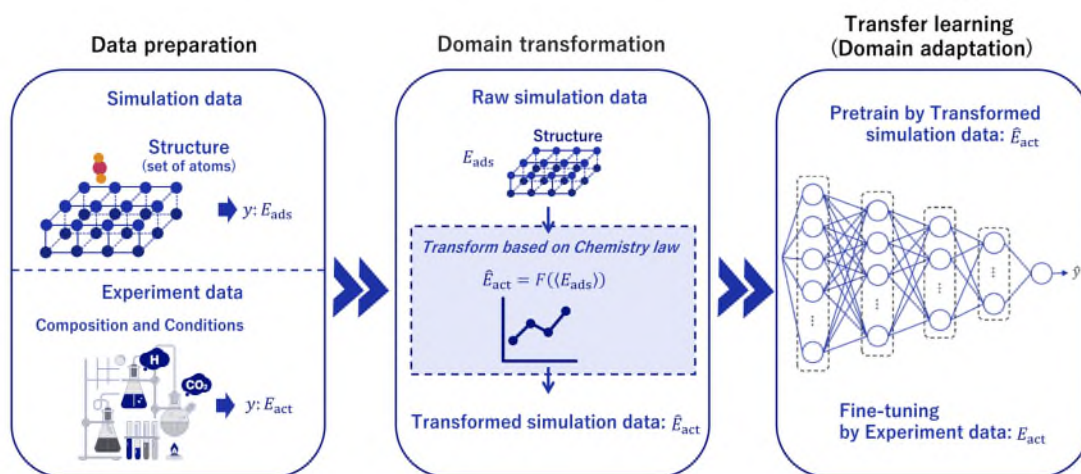


図1. 化学情報を組み込んだドメイン変換による転移学習法の概略図

## Structure Prediction from Chemical Formula Using Periodic Descriptors

◦ 吳 彦儒<sup>1</sup>, 徐 一斌<sup>1</sup>

◦ Yen-Ju Wu<sup>1</sup>, Yibin Xu<sup>1</sup>

Center for Basic Research on Materials, National Institute for Materials Science (NIMS)<sup>1</sup>

E-mail: Wu.YenJu@nims.go.jp

Recent advances in machine learning and generative models have opened new opportunities for exploring novel functional materials. However, many of the proposed candidates, especially those generated through formula-based searches, lack structural information, posing a bottleneck for simulation and experimental design. In this study, we explore the potential of our previously proposed periodic descriptors [1], which are low-dimensional and chemically structured features derived from the periodic table, for predicting crystal structure directly from chemical formulas.

We construct classification models trained on approximately 150,000 experimentally reported inorganic compounds from the AtomWork-Adv. (AWA) database [2] to predict structures, including space groups and Pearson symbols. Our models achieve top-3 accuracies of ~0.70 for space group and ~0.76 for Pearson symbol classification. These results demonstrate that compositional patterns captured through periodic trends can be effectively linked to structural tendencies, enabling prototype structure inference.

We further compare our models with benchmark classification tools and analyze prediction consistency across structure classes. This approach addresses the challenge of missing structural data in early-stage materials screening and provides a simple, composition-only method to guide structural assumptions for novel candidates. Future applications will apply this method to newly generated formulas from chemical space exploration algorithms. Structure prediction prior to simulation or experimental evaluation will support a more integrated inverse design framework in inorganic materials discovery.

### References:

- 1) Y.-J. Wu, Y. Xu, L. Fang, W. Peng, K. Sakaushi, M. Zhang, M. Arai, Y. Koyama, STAM Methods, 2513218 (2025), DOI: 10.1080/27660400.2025.2513218
- 2) AtomWork-Adv., NIMS, 2018. Available from: <https://atomwork-adv.nims.go.jp/>

# 解釈性に考慮した機械学習による CNT 繊維の構造—物性関係の可視化 Explainable Machine Learning for Elucidating Structure–Property Relationships in Carbon Nanotube Fibers

産総研<sup>1</sup> ○木村大輔<sup>1</sup>, 田島奈緒子<sup>1</sup>, 岡崎俊也<sup>1</sup>, 室賀駿<sup>1</sup>

National Institute of Advanced Industrial Science and Technology (AIST)<sup>1</sup>, ○Daisuke Kimura<sup>1</sup>,  
Naoko Tajima<sup>1</sup>, Toshiya Okazaki<sup>1</sup>, Shun Muroga<sup>1</sup>

E-mail: kimura.daisuke@aist.go.jp

近年、材料開発において人工知能（AI）の活用が進んでおり、目的に応じて多様なモデル設計が行われている<sup>[1, 2]</sup>。特に、未知領域での材料探索やメカニズム解明に資するためには、AI がどのような根拠に基づいて予測を行っているのかという「説明可能性」や「解釈性」が重要視される。本研究では、こうした説明可能な AI 技術を用いて、外挿領域も含めた構造—物性の関係性の解明を目指した。対象としたのは、軽量性と高い電気伝導性・力学物性を兼ね備え、次世代モビリティや長距離電線への応用が期待されているカーボンナノチューブ（CNT）繊維である。特に、安全性や環境負荷に配慮した水分散液を用いて作製された CNT 繊維を対象とし、その構造と物性の関係性をデータ駆動的に解析した<sup>[3]</sup>。

本研究では、多段階な製造プロセスとマルチスケールな構造情報を統合し、因子分析による特徴量抽出、回帰分析、説明可能な AI を組み合わせた解析を行った。これにより、CNT 自体の欠損、CNT 分散液の分散状態や集合構造、最終的な繊維の密度や配向など、各スケールの構造が電気伝導性や破断強度などの物性に与える影響を定量的に評価した。特に、CNT 繊維の前駆体である分散液の局所的な凝集構造が力学物性において支配的である一方、電気伝導性は CNT 自体の有効長や結晶性といったナノスケールな構造との関係が明確に示された。本手法は、複雑な材料構造に対しても解釈可能な分析を可能とし、将来的には物理モデルによる検証を行うことによって、より深い材料メカニズムの理解に貢献すると期待される。

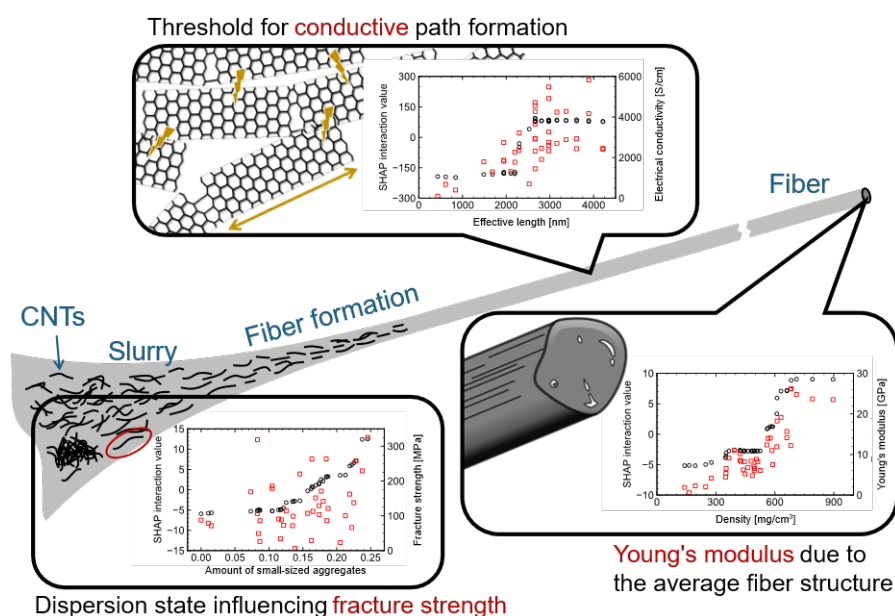


Fig. Relationship between the structure and properties of CNT fibers.

## References

- [1] S. Muroga et al., *Advanced Science*, 2302508 (2023).
- [2] H. Jintoku et al., *ACS Applied Materials & Interfaces*, 16 (9), 11800-11808 (2024).
- [3] D. Kimura et al., *Carbon*, 241, 120390 (2025).

# BaTiO<sub>3</sub>の構造相転移・誘電特性に対する機械学習ポテンシャルの応用

## Application of Machine Learning Potentials to Structural Phase Transitions and Dielectric Properties of BaTiO<sub>3</sub>

株式会社村田製作所<sup>1</sup> °平井 大介<sup>1</sup>, 金川 朋賢<sup>1</sup>, 廣瀬 左京<sup>1</sup>

Murata Manufacturing Co., Ltd.<sup>1</sup>, °Daisuke Hirai<sup>1</sup>, Tomosato Kanagawa<sup>1</sup>, Sakyo Hirose<sup>1</sup>

E-mail: daisuke.hirai@murata.com

近年、機械学習ポテンシャル（MLP）に関する研究が著しく進展しており、第一原理分子動力学（MD）計算では困難であった大規模空間スケールや長時間スケールのシミュレーション、ならびに古典 MD 計算で課題とされてきた計算精度の向上に一定の成果が得られている。また、事前学習済みの汎用機械学習ポテンシャル（uMLP）の開発も活発化しており、その計算精度の向上は顕著である [1]。uMLP は材料開発の加速に大きな可能性を秘めている一方で、原子間力や応力の精度については注意が必要である。実際、文献 [1]におけるベンチマーク計算の多くはエネルギー精度に着目したものであり、フォノンに起因する物性計算に関するベンチマーク研究では、その計算精度に大きなばらつきが存在することが報告されている [2]。この課題に対し、uMLP のファインチューニングが有効と考えられるが、適切な訓練データの選定は容易ではない。そこで本研究では、第一原理 MD 計算と並行して機械学習ポテンシャルを構築する on-the-fly 学習型 MLP に着目し、特定材料に対する高精度 MLP の構築および検証を行った。

本研究では、BaTiO<sub>3</sub>の構造相転移および誘電特性を解析可能な MLP の構築を実施した。構造最適化計算および MLP 構築には VASP コード [3]を用い、交換相関ポテンシャルには GGA-PBE を採用した。MLP の訓練には、2×2×2 の菱面体相スーパーセルを用い、第一原理 MD 計算を並行させながら on-the-fly 学習を行った。様々な温度で学習した MLP を用いて構造の安定性をシミュレーションした結果、実験で報告されている BaTiO<sub>3</sub>の菱面体晶、直方晶、正方晶、立方晶の 4 相を観測するには、8×8×8 以上のセルサイズが必要であることが明らかとなった。また、このスーパーセルを用いて誘電率の温度依存性を計算し、各相転移温度付近で誘電率ピークが顕著に現れることを確認した。さらに、各温度における局所分極の挙動を解析した結果、立方晶相では分極がランダムに配向していることが示され、BaTiO<sub>3</sub>の正方晶－立方晶間の強誘電相転移が order・disorder 型であることを示唆する結果を得た。

[1] <https://matbench-discovery.materialsproject.org/>.

[2] Antoine Loew et. al., npj Comput. Mater. **11**, 178 (2025).

[3] George Kresse and Jürgen Furthmüller, Phys. Rev. B, **54**, 11169 (1996).; George Kresse and Daniel Joubert, Phys. Rev. B, **59**, 1758 (1999).; Ryosuke Jinnouchi et. al., Phys. Rev. Lett. **122**, 225701 (2019).; Ryosuke Jinnouchi et. al., Phys. Rev. B **100**, 014105 (2019).

# Extreme learning machine による GeSn 混晶半導体の全エネルギー予測

## Total energy prediction of GeSn alloy semiconductors using extreme learning machine

九工大院情報工<sup>1</sup>, 岡山県立大院情報系工<sup>2</sup>, 岡山県立大情報工<sup>3</sup>

○野田 祐輔<sup>1</sup>, 別宮 響<sup>2</sup>, 末岡 浩治<sup>3</sup>

Kyushu Inst. Technol.<sup>1</sup>, Grad. Sch. Comput. Sci. Syst. Eng., Okayama Pref. Univ.<sup>2</sup>,

Fac. Comput. Sci. Syst. Eng., Okayama Pref. Univ.<sup>3</sup>

○Yusuke Noda<sup>1</sup>, Hibiki Bekku<sup>2</sup>, Koji Sueoka<sup>3</sup>

E-mail: noda@phys.kyutech.ac.jp

近年, 密度汎関数理論 (DFT) に基づく第一原理計算を機械学習モデルに置き換える機械学習原子間ポテンシャル計算が様々な分野で活用されている[1-3]. 特に人工ニューラルネットワーク (ANN) 原子間ポテンシャル[4]は機械学習原子間ポテンシャルの代表的な手法として注目されている. しかし, ANN 原子間ポテンシャルではモデル最適化のために NN モデルの隠れ層やノードの数を適切に決めることが困難であり, 機械学習のコストが高くなる傾向がある[5].

本研究では, ニューラルネットワーク系機械学習の一種であり, 回帰モデルを高速に構築できる Extreme Learning Machine (ELM) [6]に注目する. GeSn 混晶半導体の結晶構造の全エネルギー予測を例に, DFT 計算から得られる全エネルギーと原子周りの環境を表す記述子の関係を学習した ELM 回帰モデルの予測精度を検証する. 図 1 は ELM 回帰モデルによる  $\text{Ge}_{1-x}\text{Sn}_x$  混晶 ( $0.0 < x < 1.0$ ) の結晶構造の全エネルギー予測結果であり, 訓練データ・検証データ共に, 高い予測精度が得られた. 研究成果の詳細は, 当日の発表にて紹介する.

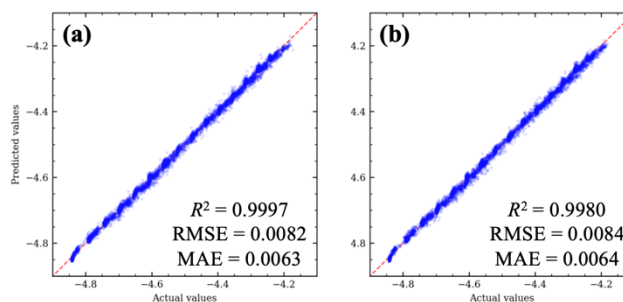


Fig. 1 Predicted values for total energies of GeSn structures. (a) Training data and (b) validation data.

### 参考文献

- [1] G. R. Schlerder et al., *J. Phys. Mater.* **2**, 032001 (2019).
- [2] T. Mueller, A. Hernandez, and C. Wang, *J. Chem. Phys.* **152**, 050902 (2020).
- [3] R. Jacobs et al., *Curr. Opin. Solid State Mater. Sci.* **35**, 101214 (2025).
- [4] J. Behler, *J. Chem. Phys.* **134**, 074106 (2011).
- [5] A. M. Tokita and J. Behler, *J. Chem. Phys.* **159**, 121501 (2023).
- [6] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, *Neurocomputing* **70**, 489 (2006).

### 謝辞

本研究の一部は JST-CREST (課題番号: JPMJCR21C2) の助成を受けた. 本研究における計算は名古屋大学スーパーコンピュータ「不老」を利用して実施した.



# Stable and Functional Quaternary Heusler Alloys: A Machine Learning-Assisted High-Throughput First-Principles Exploration

Research Center for Magnetic and Spintronic Materials, National Institute for Materials Science<sup>1</sup>

Enda Xiao<sup>1</sup>, Terumasa Tadano<sup>1</sup>

XIAO.Enda@nims.go.jp TADANO.Terumasa@nims.go.jp

Heusler alloys are promising materials for applications in spintronic devices, shape-memory devices, and thermoelectrics. Building upon our DXMag comprehensive density functional theory (DFT) database of ternary Heusler compounds [1], we extend our exploration to quaternary systems with the formula XX'YZ. Given the vast chemical space of over 131,545 potential compositions, an efficient pre-screening method is essential to mitigate computational costs.

We employ a machine learning (ML)-assisted strategy to overcome this problem. Pretrained ML force fields (MLFFs) are utilized for structural optimization and stability estimation. The models trained on our database are used to do further estimation of stability and property. Successive filters based on formation energy and convex hull distance reduce the candidate pool to about 10%. Further screening, incorporating ML-predicted phonon stability and Curie temperatures, narrows the selection to a manageable subset for first-principles validation.

In this talk, we assess the reliability of different pretrained MLFFs in structural optimization and energetic stability evaluation. We also demonstrate the predictive accuracy of our custom ML models in estimating various properties. Our findings underscore the potential of this ML-assisted data-driven approach in accelerating the discovery of stable, multifunctional magnetic materials.

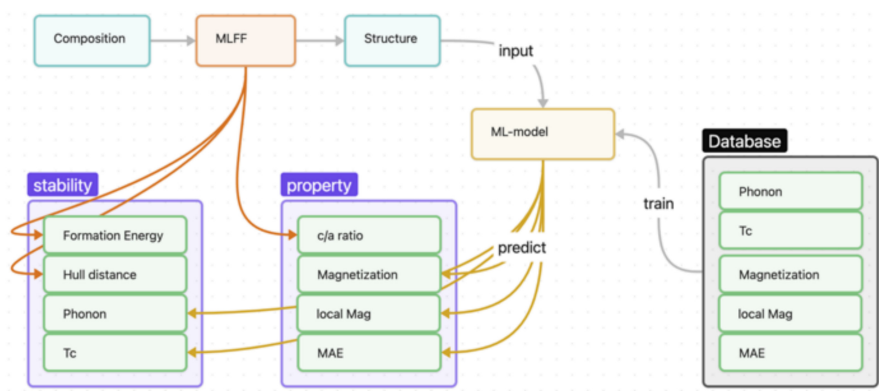


Fig. 1. Workflow of the Quaternary Heusler Space Exploration.

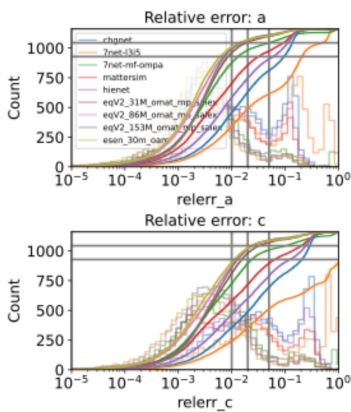


Fig. 2. Optimization performance

## References

[1] E. Xiao and T. Tadano. arXiv preprint arXiv:2502.17946 (2025)

23 合同セッションN「インフォマティクス応用」 | 一般セッション(口頭講演): 23.1 合同セッションN「インフォマティクス応用」

## [9p-S301-1~12] 23.1 合同セッションN「インフォマティクス応用」

[9p-S301-1]

GNNモデルを用いたCaTiO<sub>3</sub>の構造相転移と誘電特性の評価

○(M1)吉持 幸毅<sup>1</sup>、Alex Kutana<sup>1</sup>、旭 良司<sup>1</sup> (1.名大工)

[9p-S301-2]

機械学習ハイブリッドポテンシャル「Tersoff-NN」による第一原理計算データの学習

○(D)西村 祐亮<sup>1</sup>、久保田 翔<sup>1</sup>、竹松 孝太郎<sup>1</sup>、渡邊 孝信<sup>1</sup> (1.早大理工)

[9p-S301-3]

機械学習ハイブリッドポテンシャル「Tersoff-NN」のSi-O 2元素系への適用

○(M2)竹松 孝太郎<sup>1</sup>、西村 祐亮<sup>1</sup>、久保田 翔<sup>1</sup>、渡邊 孝信<sup>1</sup> (1.早大理工)

[9p-S301-4]

量子アニーリングによる固溶体の安定構造探索

○酒井 拓<sup>1</sup>、市場 友宏<sup>1</sup>、本郷 研太<sup>2</sup>、前園 涼<sup>3</sup> (1.北陸先端大、2.北陸先端大情報社会基盤、3.東京科学大)

[9p-S301-5]

シンボリック回帰によるバンドギャップ補正に基づく太陽電池材料の探索

○河野 貴大<sup>1</sup>、宮本 奏汰<sup>1</sup>、増田 太一<sup>1</sup>、田辺 克明<sup>1</sup> (1.京大工)

[9p-S301-6]

三元水素化物に対する機械学習を用いた高圧超伝導材料の探索

○宮本 奏汰<sup>1</sup>、徳山 和映<sup>1</sup>、増田 太一<sup>1</sup>、田辺 克明<sup>1</sup> (1.京大工)

[9p-S301-7]

汎用機械学習ポテンシャルPFPを用いたMagnéliベースチタン酸窒化物の安定構造探索

○青木 祐太<sup>1</sup> (1.株式会社Preferred Computational Chemistry)

[9p-S301-8]

第一原理計算とLightGBMによる硫化物系熱電材料のインシリコ探索

○野田 祐輔<sup>1</sup>、島津 圭音<sup>1</sup> (1.九工大院情報工)

[9p-S301-9]

金属結晶に基づくイオン結晶の構造予測手法

○横山 智康<sup>1</sup>、大内 暁<sup>1</sup>、市川 和秀<sup>1</sup> (1.パナソニック ホールディングス (株) )

[9p-S301-10]

LLMを使った材料開発(IV): 新規高密度磁気記録媒体の設計と試作

○石井 真史<sup>1</sup>、渡邊 泰史<sup>1</sup>、鈴木 一平<sup>1</sup>、高橋 有紀子<sup>1</sup> (1.NIMS)

[9p-S301-11]

ChatGPTを用いた、材料物性間の関係性知識グラフ及び関係性探索ツールの作成

○吉武 道子<sup>1,2</sup>、長田 貴弘<sup>1</sup> (1.NIMS、2.MatQ-lab)

[9p-S301-12]

元素置換系のパーシステントホモロジー検討

○吉武 道子<sup>1</sup>、長田 貴弘<sup>1</sup> (1.NIMS)

---

# GNN モデルを用いた $\text{CaTiO}_3$ の構造相転移と誘電特性の評価

## Evaluation of phase transition and dielectric properties using GNN model

名大工<sup>1</sup>, <sup>○</sup>(M1)吉持 幸毅<sup>1</sup>, Alex Kutana<sup>1</sup>, 旭 良司<sup>1</sup>

Nagoya Univ.<sup>1</sup>, <sup>○</sup>Koki Yoshimochi<sup>1</sup>, Alex Kutana<sup>1</sup>, Ryoji Asahi<sup>1</sup>

E-mail: yoshimochi.koki.u7@s.mail.nagoya-u.ac.jp

誘電率は、密度汎関数摂動法 (DFPT) によって計算可能であるが、計算負荷が大きく大規模な材料系や有限温度の誘電特性には適用が困難である。我々は最近、ボルン有効電荷テンソルおよび誘電率テンソルの予測が可能な回転等価型グラフニューラルネットワークモデル (EquivarGNN) [1,2]を開発した。本研究では、誘電体材料である  $\text{CaTiO}_3$ に着目し、その相転移と誘電率特性について調べた。

80 原子の  $\text{CaTiO}_3$  構造モデルに対して、力場は GNN モデル MACE によって、ボルン有効電荷は EquivarGNN によってそれぞれ精密に予測できるよう、第一原理 MD 計算のデータセットを用いて訓練した。電場印可のもとで各原子のボルン有効電荷と変位から分極率を計算し、クラウジウス・モソッティの関係式から誘電率を計算した。Fig. 1 に計算された  $\text{CaTiO}_3$ の誘電率の温度依存性を示す。T=0 では DFPT の結果と良く一致し、温度依存性は実験結果を再現した。

$\text{CaTiO}_3$  は 1400-1600K の温度領域で、低温の ortho 相から tetra 相、cubic 相と構造相転移することが知られている。しかしながら格子定数の変化からこれらの相転移を同定するのは困難であった。そこで、ペロブスカイト構造の隣接する octahedral との傾きの関係を PDynA モジュール[3]から得て、局所構造の相分類を行った。計算は、1280 原子のモデルを用い、GNN を用いた NPT の MD シミュレーションによって、時系列の構造データを獲得し、評価を行った。

局所構造の相分類の結果、一部の領域において異なる局所相が発現、拡大、消失する様子が観察された (Fig. 2)。これらの構造的なゆらぎを平均化することで、実験で観察された相転移を再現することができた。本結果によって、GNN モデルが局所構造やドメイン構造を精密に予測することが示された。

[1] Kutana, Shimizu, Watanabe, Asahi, Sci. Rep. 15, 16719 (2025).

[2] Kutana, Yoshimochi, Asahi, STAM: Methods (2025).

[3]Liang et al., J. Phys. Chem. C 127, 19141 (2023).

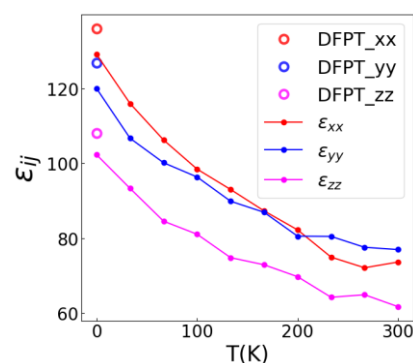


Fig. 1 Dielectric constants of  $\text{CaTiO}_3$

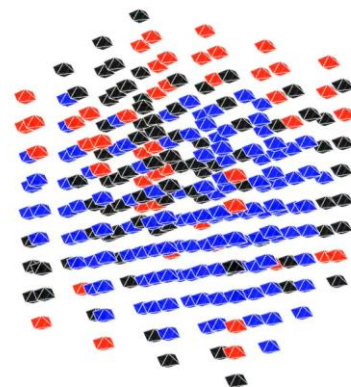


Fig. 2 Local phase mapping of  $\text{CaTiO}_3$  at 1450K

# 機械学習ハイブリッドポテンシャル「Tersoff-NN」による 第一原理計算データの学習

## Training on First-Principles Calculation Data

### Using the Machine Learning Hybrid Potential 'Tersoff-NN'

早大理工 ○西村 祐亮, 久保田 翔, 竹松 孝太朗, 渡邊 孝信

○Y. Nishimura, S. Kubota, K. Takematsu, T. Watanabe (Waseda Univ.)

E-mail: yusukeskelton@toki.waseda.jp

【はじめに】ニューラルネットワーク (NN) やガウス過程を用いた機械学習原子間ポテンシャルは大規模系の分子動力学 (Molecular Dynamics; MD) 計算を実現する一方、従来の解析型ポテンシャルより計算が100倍以上遅く、とくに多層パーセプトロン (MLP) による推論では未知の構造への外挿性能が低いという課題が残る<sup>[1,2]</sup>。そこで古典的なポテンシャル関数の一部に機械学習モデルを組み込み、高速性と高精度を両立させるハイブリッド型ポテンシャルが注目されている<sup>[3]</sup>。例えば ReaxFF のボンドオーダー計算部を機械学習モデルに置き換えた ReaxFF-NN が提案されており、従来の ReaxFF と同等の計算速度を維持しつつ、第一原理計算に迫る予測精度を実現している<sup>[4,5]</sup>。しかし ReaxFF 自体が複雑かつアドホックな関数形であるため、外挿性能が保証されているとは言えない。

前回<sup>[5,6]</sup>、共有結合のユニバーサルリティ<sup>[7,8]</sup>を考慮して設計された Tersoff ポテンシャル<sup>[9]</sup>をベースに、ボンドオーダー部を多層パーセプトロン (MLP) で置き換えたハイブリッド型ポテンシャル「Tersoff-NN ポテンシャル」を提案した。ただし前回の報告では、学習ラベルとして Tersoff ポテンシャルのエネルギー値を用いたため、Tersoff ポテンシャルの再現能力の検証にとどまった。

今回、第一原理計算の結果を用いて Tersoff-NN の学習を行い、第一原理計算に近い水準の再現能力があるかどうか検証した。

【ポテンシャルの概要】Tersoff ポテンシャルではボンドオーダー $b_{ij}$ を次式で定義する。

$$b_{ij} = (1 + \beta^n \zeta_{ij}^n)^{-0.5/n} \quad \zeta_{ij} = \sum_k \xi_{jik}$$

$$\xi_{jik} = f_c(r_{ik}) g(\theta_{jik}) \exp\left\{\lambda_3^3 (r_{ij} - r_{ik})^3\right\}$$

$$g(\theta_{jik}) = 1 + \frac{c^2}{d^2} - \frac{c^2}{d^2 + (h - \cos \theta_{jik})^2}$$

$k$ は原子 $i$ の $n_i$ 個の近接原子(結合相手である $j$ を除く)のラベル、 $f_c(r_{ik})$ はカットオフ関数、 $g(\theta_{jik})$ は角度依存項、 $n$ 、 $\beta$ 、 $c$ 、 $d$ 、 $h$ はパラメータである。前回は、原子 $i$ 、 $j$ と、それ以外の周辺原子 $k$ からなる $j$ - $i$ - $k$ トリプレットの寄与 $\xi_{jik}$ を MLP で表現したが、今回はさらに $\xi_{jik}$ の総和 $\zeta_{ij}$ を引数とする関数 $b_{ij}(\zeta_{ij})$ を MLP に置き換え、ポテンシャルの表現能力を高めた。ボンドオーダー $b_{ij}$ 決定部のネットワーク構成を Fig.1 に示す。

今回、第一原理計算プログラム VASP<sup>[10-12]</sup>を用いて

第一原理分子動力学計算を実施し、構造とエネルギーの時系列データを訓練データとして学習を行った。

【結果】全学習データに対するパリティプロットを Fig. 2 に示す。Fig.3 に学習済みの Tersoff-NN モデルによる Si 結晶 (256 原子) の融解 MD 計算の結果を示す。4Å で観測されていた動径分布関数の過剰な振動がほぼ消滅し、第一原理計算の結果に近づいた。

【謝辞】本研究は科学研究費 基盤研究 (B) (25K01295)の支援を受けて行われた。

【参考文献】[1] Y. Mishin, Acta Mater., 214, 116980 (2021). [2] J. Behler, Chem. Rev., 121, 10037-10072 (2021). [3] L.-Y. Xue et al., Phys. Chem. Chem. Phys., 23, 19457-19464 (2021). [4] Feng Guo et al., Phys. Chem. Chem. Phys., 27, 10571-10579 (2025). [5] Y. Nishimura et al., 16a-K505-1 JSAP EXPO spring 2025 (2025). [6] K. Hirai et al., 16a-K505-2 JSAP EXPO spring 2025 (2025). [7] J. H. Rose et al., PRB 29, 2963 (1984). [8] G. C. Abel, PRB 31, 6184 (1985). [9] J. Tersoff, PRB 38, 9902-9905 (1988). [10] G. Kresse et al., Phys. Rev. B., 47, 558 (1993); ibid. 49, 251 (1994). [11] G. Kresse et al., Comput. Mat. Sci., 6, 15 (1996). [12] G. Kresse et al., Phys. Rev. B., 54, 169 (1996).

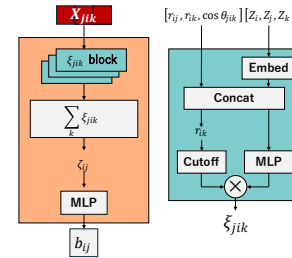


Fig. 1 Architecture of bond order prediction models

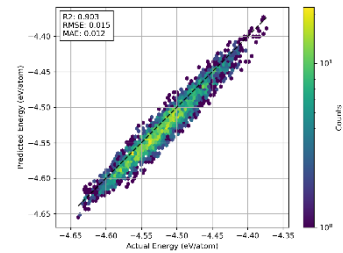


Fig. 2 Prediction accuracy of energy

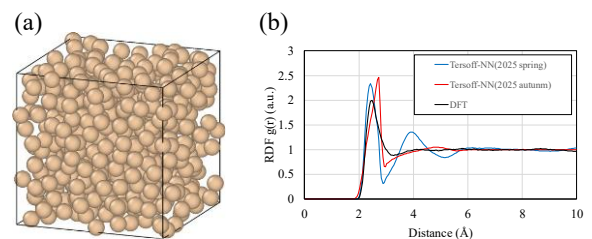


Fig. 3 (a) final snapshot of Si melting simulation

(b) the corresponding RDF



# 機械学習ハイブリッドポテンシャル「Terstoff-NN」の Si-O 2 元素系への適用

## Application of the Machine Learning Hybrid Potential 'Terstoff-NN' to Si-O Binary Systems

早大理工 ○竹松 孝太郎, 西村 祐亮, 久保田 翔, 渡邊 孝信

○K. Takematsu, Y. Nishimura, S. Kubota, T. Watanabe (Waseda Univ.)

E-mail: kotakematsu@toki.waseda.jp

【はじめに】機械学習ポテンシャルの拡張性と旧来の解析的ポテンシャル関数の計算コストの軽さを兼ね備えたハイブリッド型のポテンシャル「Terstoff-NN」の開発に取り組んでいる<sup>[1, 2]</sup>。Terstoff-NN では、オリジナルの Terstoff ポテンシャルにおけるボンドオーダーのみを Multi-layer Perceptron (MLP) を用いて予測する。前回、我々は Terstoff-NN を Si に適用できることを確認した<sup>[1]</sup>。今回は、多元素を扱えるように拡張した Terstoff-NN で Si-O 系への適用を試みた。

【ポテンシャルの概要】オリジナルの Terstoff ポテンシャル<sup>[3]</sup>は以下のように定義される。

$$E_{\text{total}} = \frac{1}{2} \sum_i \sum_{j \neq i} f_c(r_{ij}) [f_R(r_{ij}) + b_{ij} f_A(r_{ij})]$$

$$b_{ij} = (1 + \beta^n \zeta_{ij}^n)^{-0.5/n}, \quad \zeta_{ij} = \sum_k^{n_i} \xi_{jik}$$

$$\xi_{jik} = f_c(r_{ik}) g(\theta_{jik}) \exp \left\{ \lambda_3^3 (r_{ij} - r_{ik})^3 \right\}$$

$$g(\theta_{jik}) = 1 + \frac{c^2}{d^2} - \frac{c^2}{d^2 + (h - \cos \theta_{jik})^2}$$

$f_c$  はカットオフ関数、 $f_R$  と  $f_A$  はそれぞれ反発エネルギーと引力エネルギー項、 $b_{ij}$  が原子  $i$ - $j$  間のボンドオーダーを表す。 $g(\theta_{jik})$  は角度依存項、 $n$ 、 $\beta$ 、 $c$ 、 $d$ 、 $h$ 、 $\lambda_3$  はパラメータである。Terstoff-NN ではボンドオーダー  $b_{ij}$  の計算部を MLP に置き換えている。本研究では、Fig.1(a), (b) に示すように、ボンドオーダー計算部の構成が異なる 2 つのモデル(a), (b)を用意して比較した。

モデル(a)は、 $\xi_{jik}$  の計算のみを MLP に置き換える。Fig.1(c)に  $\xi_{jik}$  の計算部の構成を示す。各  $j$ - $i$ - $k$  トリプレットについて、 $i$ - $j$  間距離  $r_{ij}$ 、 $i$ - $k$  間距離  $r_{ik}$ 、原子  $i$  を中心とする  $j$ - $i$ - $k$  角の余弦  $\cos \theta_{jik}$ 、原子種  $Z_i$ 、 $Z_j$ 、 $Z_k$  を入力として、 $\xi_{jik}$  を出力する。あとは Terstoff ポテンシャルの定義通り  $b_{ij}$  を計算する。

モデル(b)では、 $\xi_{jik}$  の計算だけでなく、 $b_{ij}$  の計算も MLP が担う。 $\zeta_{ij}$ 、原子種  $Z_i$ 、 $Z_j$  を入力する MLP が  $b_{ij}$  を出力する。

64 原子のダイヤモンド型 Si 結晶、72 原子の  $\alpha$ -quartz 型  $\text{SiO}_2$  結晶を 6000K で溶融する定積 MD 計算をそれぞれ行い、MD 計算の軌跡から得られた各 2000 個の構造を用意し、オリジナルの Terstoff ポテンシャルで計算したエネルギーを訓練データとして用いた。計 4000 個の構造のうち、80%を学習データ、20%をテストデータとした。

【結果】Fig. 2(a), (b)にダイヤモンド型 Si 結晶と  $\alpha$ -quartz 型  $\text{SiO}_2$  結晶の、格子定数を変化させた際のエネルギー変化を示す。モデル(b)を用いるとダイヤモンド型 Si 結晶、 $\alpha$ -quartz 型  $\text{SiO}_2$  結晶共に、Terstoff ポテンシャルに近いポテンシャルカーブが再現された。Fig. 3(a), (b)に、64 原子のダイヤモンド型 Si 結晶と 72 原子の  $\alpha$ -quartz 型  $\text{SiO}_2$  結晶を用いて定積 MD 計算を 300K で行った結果を示す。ダイヤモンド型 Si 結晶については Terstoff ポテンシャルの RDF をほぼ再現した。 $\alpha$ -quartz 型  $\text{SiO}_2$  結晶の RDF についても、ややピーク幅が増したものの、Terstoff ポテンシャルの RDF を概ね再現できた。

【謝辞】本研究は科学研究費 基盤研究(B) (25K01295) の支援を受けて行われた。

【参考文献】[1] 西村他, 春季応物, 16a-K505-1(2025). [2] 平井他, 春季応物, 16a-K505-2, JSAP (2025). [3] J. Tersoff, PRB 37, 6991 (1988).

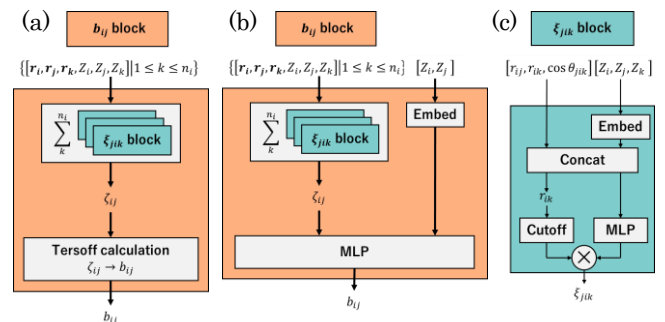


Fig. 1 Architectures of bond order prediction models

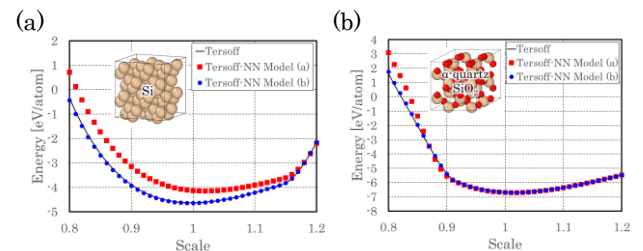


Fig. 2 Scale-Energy curve of (a) Si and (b)  $\alpha$ -quartz  $\text{SiO}_2$

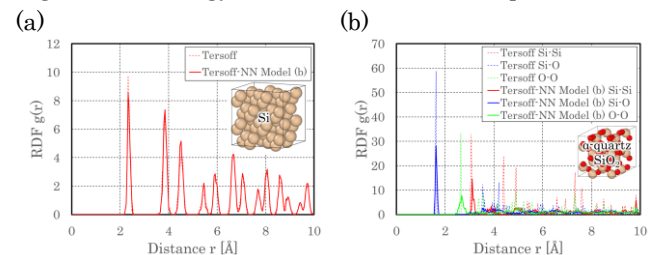


Fig. 3 RDF of (a) Si and (b)  $\alpha$ -quartz  $\text{SiO}_2$

## 量子アニーリングによる固溶体の安定構造探索

### Stable Structure Search in Solid Solutions via Quantum Annealing

北陸先端大<sup>1</sup>, 北陸先端大情報社会基盤<sup>2</sup>, 東京科学大<sup>3</sup>

○(M2) 酒井 拓<sup>1</sup>, 市場友宏<sup>1</sup>, 本郷 研太<sup>2</sup>, 前園涼<sup>3</sup>

JAIST<sup>1</sup>, JAIST RCACI<sup>2</sup>, Science Tokyo<sup>3</sup>

°Tack Saquai<sup>1</sup>, Tom Ichibha<sup>1</sup>, Kenta Hongo<sup>2</sup>, Ryo Maezono<sup>3</sup>

E-mail: mwkskt2308@icloud.com

元素やイオン、空孔のドーピングは材料特性を調整する上で有効な手法であり、このように形成された固溶体では添加元素が母体結晶を部分的に占有した構造が現れる[1]。当該系の第一原理計算においては、仮想結晶近似やスーパーセル法が用いられている[1]。特に無機化合物においては仮想結晶近似の精度が低下することが知られているため[2]、スーパーセル法が一般的である。これは、部分的に占有した元素による可能な全ての原子配置に対して第一原理計算を行う手法であり、物性評価には全ての配置の中からエネルギー的に安定な構造を採用する[1]。

一方で、固溶体や無秩序系における原子配置の組合せは容易に増加するため、候補構造全てに対する第一原理計算は困難である。そこで用いられるのが静電エネルギーによるプレスクリーニングである。静電エネルギーは各元素に割り当てられた形式電荷を用いて算定される古典的な静電ポテンシャルの総和であり、第一原理計算と比較して高速に算定可能である。また、異なる原子配置に対して安定性を評価する指標としても有用であるため、静電エネルギーの小さい構造として候補構造を削減する方策が取られる。

しかし、候補構造が増加するとその全てに対して静電エネルギー算定するだけでも膨大な時間を要し、プレスクリーニング自体も難渋してしまう。そこで我々は、プレスクリーニングにおいては安定な原子配置のみを見つければいいことに着目し、組合せ最適化問題として定式化した。この高速解法である量子アニーリングの適用に向けて、静電エネルギーをモデル化するハミルトニアンを設計してその有用性を検証している。具体的には、イオン伝導体として注目される $\beta$ -KSbF<sub>4</sub>や蛍光体のホスト材料であるBa添加SiAlONおよびCaYAlO<sub>4</sub>などの化合物群を対象に安定な原子配置のサンプリングを行っている[3-5]。当日の発表では、ハミルトニアンの定式化について紹介し、シミュレーテッドアニーリングおよび量子アニーリングによる組合せ最適化の有効性について、最適化性能および計算時間に基づいて議論する。

[1] G. Prayogo *et al.*, J. Chem. Inf. Model. **62**, 2909, (2022).

[2] J. Leeman *et al.*, PRX Energy **3**, 011002, (2024).

[3] K. Yamada *et al.*, Chem. Lett. **28**, 627, (1999).

[4] S. Esmacilzadeh *et al.*, Chem. Mater. **16**, 2113, (2004).

[5] Y. Zhang *et al.*, ACS Appl. Mater. Interfaces **7**, 2715, (2015).

# シンボリック回帰によるバンドギャップ補正に基づく太陽電池材料の探索

## Exploration of Solar Cell Materials

### Based on Bandgap Correction via Symbolic Regression

京大工, °(M2) 河野 貴大, 宮本 奏汰, 増田 太一, 田辺 克明

Kyoto Univ., °Takahiro Kono, Souta Miyamoto, Taichi Masuda, Katsuaki Tanabe

E-mail: tanabe@cheme.kyoto-u.ac.jp

バンドギャップは、太陽電池材料の発電効率を左右する最も重要な物性値の一つである [1]。しかし、実験的に測定するには高度な設備と多くの時間を要する。また、密度汎関数理論 (DFT) を用いて計算することも可能だが、実験値に対して過小評価される傾向があることが知られている [2]。本研究では、データベースに多数存在する DFT 計算値を活用しつつ、実験値に整合するバンドギャップを予測する機械学習モデルを構築し、太陽電池材料の探索を行った。

バンドギャップの DFT 計算値として Automatic Flow for Materials Discovery (AFLOW) に保存された 30,592 件及び Materials Project (MP) の 13,771 件を用い、実験値として文献 [3] で使用された 2,483 件を用いた。DFT 計算値と実験値の両方が揃っている 763 件の化合物のデータから、両者の関係を構成元素の物性値で表す式をシンボリック回帰で導出した。この際、Lasso 回帰による特徴量選択を経た後、PySR ライブラリ [4] を用いた。次に、導出した関係式を用いて、DFT 計算値を補正し、構成元素の物性値から導出した 129 種類の特徴量を説明変数として、XGBoost 回帰によるバンドギャップ予測モデルを構築した。可採埋蔵量が比較的豊富な元素である Cu および Zn を含む三元系化合物に対して本モデルを適用し、高効率が期待される太陽電池材料候補の探索を行った。加えて、有望と予測された化合物については、その熱力学的安定性を評価するために、Roost (Representation Learning from Stoichiometry) [5] を用いて生成エンタルピーを予測した。Roost の学習には、Open Quantum Materials Database (OQMD) から抽出した、DFT 計算による生成エンタルピー 644,008 件を用いた。

シンボリック回帰の結果、DFT 計算値は、構成元素の標準原子化熱や第一イオン化エネルギーの平均などを用いることで、実験値へと補正可能であることが示唆された。また、この式を用いて生成した補正值は従来よりも広範な範囲を網羅しており、学習データに多様性を与えた。スクリーニングの結果、 $\text{CuCa}_{0.8}\text{Nb}_{0.8}$  や  $\text{ZnSb}_{0.9}\text{Ca}_{0.5}$  などの化合物が、バンドギャップが 1.40 eV かつ生成エンタルピーが  $-0.40 \text{ eV atom}^{-1}$  未満と予測され、太陽電池材料として有望であると示唆された。

[1] W. Shockley and H. J. Queisser, *J. Appl. Phys.*, **32**, 510 (1961).

[2] M. K. Y. Chan and G. Ceder, *Phys. Rev. Lett.*, **105**, 196403 (2010).

[3] S. K. Kauwe *et al.*, *Integr. Mater. Manuf. Innov.*, **9**, 213 (2020).

[4] M. Cranmer, arXiv 2305.01582 (2023).

[5] R. E. A. Goodall and A. A. Lee, *Nat. Commun.*, **11**, 6280 (2020).

## 三元水素化物に対する機械学習を用いた高圧超伝導材料の探索

### Exploration of High-Pressure Superconducting Materials

#### in Ternary Hydrides Using Machine Learning

京大工 °宮本 奏汰, 徳山 和映, 増田 太一, 田辺 克明

Kyoto Univ., °Souta Miyamoto, Kazuaki Tokuyama, Taichi Masuda, Katsuaki Tanabe

E-mail: s.miyamoto@cheme.kyoto-u.ac.jp

水素化物は高圧下で高温超伝導を示す有望な材料系であり、特に、比較的低压で超伝導を示す多元水素化物が注目されている[1,2]。しかし、多元系の組成空間は広大であり、実験や理論計算のみで網羅的に探索することは難しい。そこで、機械学習によるスクリーニング手法が検討されており[3,4]、二元水素化物を対象とした研究例がある[4]。本研究では三元水素化物に着目し、機械学習モデルの構築と、候補組成の予測評価を行った。

2007年から2024年に報告された二元および三元水素化物2,059件からなるデータセットを収集した。構成元素毎の特徴量の組み合わせによる220変数と、これに圧力を加えた計221変数を入力とし、臨界温度を目的変数とした。特徴量選択には Ridge 回帰と赤池・ベイズ情報量基準を用い、回帰モデルには Nested Cross Validation で最も高い精度を示した XGBoost 回帰モデルを用いた。二乗平均平方根誤差は 38.5 K、平均絶対誤差は 24.7 K、決定係数は 0.80 であった。モデル初期化・最適化を異なる乱数シードで 30 回独立に行い、30 個の XGBoost モデルを構築した。原子番号 2-90 の元素対 A-B 全 3,916 通りに 1% 組成刻みの組成グリッドから生成した A-B-H 系候補に対し、100、200、300 GPa の各圧力で臨界温度を予測した。30 モデルの予測から 95 %信頼区間下限を計算し、候補を順位付けした。

予測結果から、訓練データに含まれない未探索の有望な候補として、 $\text{Ca}_5\text{Ti}_5\text{H}_{90}$  は 100 GPa で 233.1 K、200 GPa で 298.4 K、および 300 GPa で 295.3 K を示した。その他、Na-Mg-H、Li-K-H 系などを抽出した。XGBoost モデルにおける特徴量重要度を解析したところ、寄与率の約 50 % に上位 16 項目の変数が含まれており、支配的な要因は確認できなかった。しかし、第一イオン化エネルギー、van der Waals 半径、価電子数などが一貫して高い寄与を示した。

以上、高圧水素化物超伝導体探索に対して、機械学習手法による膨大な組成空間の効率的なスクリーニングを行った。得られた結果に基づき、密度汎函数理論計算により、候補水素化物の構造探索と、予測された臨界温度の検証を進めている。

[1] A. P. Drozdov, M. I. Erements, I. A. Troyan, V. Ksenofontov, and S. I. Shylin: *Nature* **525**, 73–76 (2015).

[2] N. W. Ashcroft: *Phys. Rev. Lett.* **92**, 187002 (2004).

[3] M. J. Hutcheon, A. M. Shipley, and R. J. Needs: *Phys. Rev. B* **101**, 144505 (2020).

[4] T. Konno, H. Kurokawa, F. Nabeshima, Y. Sakishita, R. Ogawa, I. Hosako, and A. Maeda: *Phys. Rev. B* **103**, 014509 (2021).

## 汎用機械学習ポテンシャル PFP を用いた Magnéli ベースチタン酸窒化物の 安定構造探索

株式会社 Preferred Computational Chemistry<sup>1</sup>, <sup>○</sup>青木 祐太<sup>1</sup>

Preferred Computational Chemistry, Inc.<sup>1</sup>, <sup>○</sup>Yuta Aoki

E-mail: yutaaoki@pfcc.co.jp

可視光応答型光触媒材料の候補として理論的に提案されたチタン酸窒化物は,  $\text{Ti}_n\text{N}_2\text{O}_{2n-3}$  ( $n \geq 2$ ) の一般式で表される一連の化合物である. これは,  $\text{Ti}_n\text{O}_{2n-1}$  ( $n \geq 2$ ) の一般式で表される一連のチタン亜酸化物系 (Magnéli 相) の結晶構造において,  $2n-1$  個の O 原子のうち 2 つを N 原子に置換することで得られるもので,  $\text{TiO}_2$  に対する N ドーピングと比べて高い N 濃度が実現でき, それによって可視光応答性を高めることが期待される.<sup>[1,2]</sup> 先行研究においては最も単純かつ最も N 濃度の高いケース ( $n = 2$ ) となる  $\text{Ti}_2\text{N}_2\text{O}$  については詳細に調べられており, 可視光を利用した光触媒水分解に非常に適したバンド構造を持つことが示されているが,  $n \geq 3$  の系についてはまだ詳細に調べられていない. N 濃度のチューニングを通したバンドギャップエンジニアリングの可能性を評価するには  $n \geq 3$  のケースについても系統的に調べる必要がある.  $n \geq 3$  のケースを調べるうえで大きな障壁となるのが,  $\text{Ti}_n\text{O}_{2n-1}$  から  $\text{Ti}_n\text{N}_2\text{O}_{2n-3}$  への原子置換において可能な構造パターンの膨大さである. 一般に  $n$  が大きくなるほど置換前の母構造の結晶構造も複雑になると考えられ, そこから可能な置換後の派生構造の数は母構造によっては 10000 を優に超えるケースもある. この膨大な数の候補構造の中から安定な構造を第一原理計算によってスクリーニングするのは容易ではなく, 現実的な計算コストで精度の高いスクリーニングを実行するには高精度な機械学習ポテンシャルを使った計算が必要になる.

本研究では, 汎用機械学習ポテンシャル PFP<sup>[3]</sup>を用い, 先行研究ではまだ調べられていない  $n \geq 3$  のケースにおけるチタン酸窒化物  $\text{Ti}_n\text{N}_2\text{O}_{2n-3}$  の安定構造を網羅的にスクリーニングした結果を報告する.  $\text{Ti}_n\text{N}_2\text{O}_{2n-3}$  の候補構造は Enumlib の Python インターフェースを用い, 対称性的に非等価な構造のみ網羅的に生成させている. スクリーニングされた構造については, 動的な安定性を調べるためにフォノンの計算も行った. これらの中にはソフトモードを示さない構造も見つかっており, その構造の動的な安定性を示している. さらに, どのような構造的・原子環境的な特徴があるときに構造が安定化するかについても解析を行ったので, その結果も合わせて報告する.

### References:

- [1] Yuta Aoki *et al.*, arXiv:1701.06251 (2017).
- [2] Yuta Aoki *et al.*, Phys. Rev. **B** 99, 075203 (2019).
- [3] S. Takamoto *et al.*, Nat. Commun. **13**, 2991 (2022).



# 第一原理計算と LightGBM による硫化物系熱電材料のインシリコ探索 In silico exploration of sulfide-based thermoelectric materials using first-principles calculations and LightGBM

九工大院情報工<sup>1</sup>

○野田 祐輔<sup>1</sup>, 島津 圭音<sup>1</sup>

Department of Physics and Information Technology, Kyushu Institute of Technology<sup>1</sup>

○Yusuke Noda<sup>1</sup>, Kein Shimazu<sup>1</sup>

E-mail: noda@phys.kyutech.ac.jp

電気エネルギーと熱エネルギーを相互に変換できる熱電材料は、発電用途や冷却・加熱用途の熱電変換デバイスに用いられる。特にエネルギー問題の観点から、工業廃熱を熱電変換へ利用する技術が発展することで、化石燃料の使用削減や CO<sub>2</sub> 排出量の抑制など脱炭素社会の実現に大きく貢献できることが期待されている。Cu<sub>1.97</sub>S や SnS を代表とする硫化物は、無毒性、地球上の豊富な資源、質量の軽い原子による格子熱伝導率の低減化という特徴を活かして、次世代熱電材料の有力候補の一つとなっている[1]。

本研究では、新たな硫化物系熱電材料を探索するために、理論計算と機械学習を活用したインシリコ探索のアプローチを試みた。結晶構造データベース OQMD[2]から硫化物系の結晶構造データを抽出し、第一原理計算および Boltzmann 輸送理論を用いて 2,610 件の硫化物系の性能指数  $ZT$  を算出した。図 1 は価電子帯の領域で高  $ZT$  を示す有力物質の一つである BiKS<sub>2</sub> の結晶構造および  $x, y, z$  各成分へ寄与する  $ZT$  のグラフであり、各成分の平均結果（図 1(b)黒線）の最大値として  $ZT_{\text{cal}} = 1.20$  という計算結果が得られた。様々な硫化物系の  $ZT_{\text{cal}}$  を学習した LightGBM モデル[3]を構築し、このモデルを用いて別の 3,554 件の硫化物系の  $ZT$  を予測した。研究成果の詳細は、当日の発表にて紹介する。

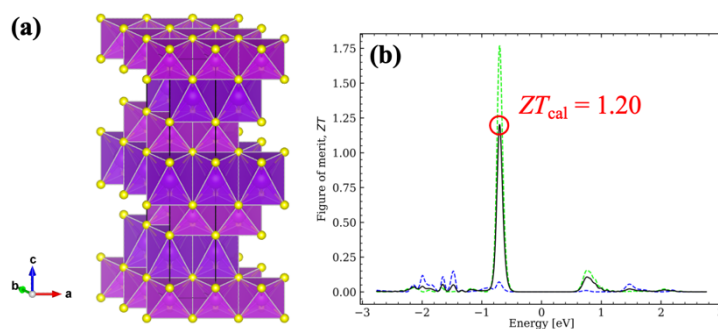


Fig. 1 (a) Crystal structure and (b) energy- and direction-dependent  $ZT$  of BiKS<sub>2</sub> exhibiting high p-type  $ZT$ .

## 参考文献

- [1] F.-H. Sun et al., *J. Materiomics*, **10**, 218 (2024).
- [2] S. Kirklin et al., *npj Comput. Mater.* **1**, 15010 (2015).
- [3] G. Ke et al., *Proc. 31st NeurIPS*, 3149 (2017).

## 謝辞

本研究における計算は名古屋大学スーパーコンピュータ「不老」を利用して実施した。

## 金属結晶に基づくイオン結晶の構造予測手法

### Crystal Structure Prediction Method Using Metal Crystals for Ionic Crystals

パナソニック ホールディングス株式会社<sup>1</sup> °横山 智康<sup>1</sup>, 大内 暁<sup>1</sup>, 市川 和秀<sup>1</sup>

Panasonic Holdings Corp.<sup>1</sup>, °Tomoyasu Yokoyama<sup>1</sup>, Satoru Ohuchi<sup>1</sup>, Kazuhide Ichikawa<sup>1</sup>

E-mail: yokoyama.tomoyasu@jp.panasonic.com

**背景：**結晶構造予測は、組成式に基づいて原子が空間内でどのように配置されるかを予測する手法であり、新材料の発見に向けた研究として近年注目を集めている。これまでの多くの研究は、既に合成が報告されている結晶構造に対して元素置換を行い得られた仮想構造を基にスクリーニングを実施してきた。しかし、このアプローチでは既知の構造と同じトポロジーを持つ結晶構造しか得られず、未知の結晶構造の予測には限界がある。そこで我々は、金属結晶とイオン結晶の構造的共通点に着目し[1]、金属結晶をテンプレート構造としてイオン結晶の構造を生成する新たな手法「Met2Ion 法」を提案した。

**計算方法：**Met2Ion 法は、金属結晶とイオン結晶の構造的共通点に基づいた結晶構造予測手法である。図にはその関係性の一例を示す。具体的には、金属結晶である  $\text{MgCu}_2$  の  $\text{Mg}$  原子を  $\text{Mg}^{2+}$  イオンに、 $\text{Cu}$  原子を  $\text{Al}^{3+}$  イオンに置換し、これらのイオンからなる多面体中心サイトに電気的中性を保つための  $\text{O}^{2-}$  イオンを配置することで、イオン結晶であるスピネル型  $\text{MgAl}_2\text{O}_4$  が得られる。Met2Ion 法では、この関係性に従い金属結晶の元素をカチオンおよびアニオンに置換し、多面体中心サイトにアニオンおよびカチオンを配置することでイオン結晶を生成する。イオンの配置には複数の組み合わせが考えられるため、クラスター展開法を用いてエネルギーが最も低くなる配置を抽出した。

**計算結果：**出発構造として  $\text{MgCu}_2$  を用いて Met2Ion 法を  $\text{MgAl}_2\text{O}_4$  の組成に適用したところ、生成された構造は実験的に観測されるスピネル型構造と一致した。また、他の金属結晶を出発構造とした場合にも、イオン結晶の構造を正しく予測することができ、本手法の妥当性を示すことができた。本手法は、合成報告のある金属結晶に限定されるものの、構造的妥当性の高い未知のイオン結晶を予測することが可能であり、新たな高機能エネルギー材料の創出が期待される。

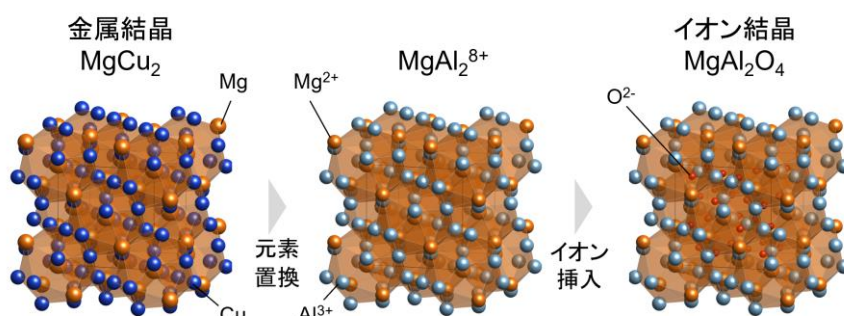


図 1. Met2Ion 法による金属結晶 ( $\text{MgCu}_2$ ) からイオン結晶 ( $\text{MgAl}_2\text{O}_4$ ) の生成過程

[1] T. Yokoyama *et al.*, *arXiv* 2407.02838 (2024).

LLM を使った材料開発(IV)：新規高密度磁気記録媒体の設計と試作

Materials Development with LLM (IV): Design and prototyping of new high-density magnetic recording media

NIMS, 渡邊泰史、鈴木一平、<sup>○</sup>石井真史、高橋有紀子

NIMS, Yasufumi Watanabe, Ippei Suzuki, <sup>○</sup>Masashi Ishii, and Yukiko Takahashi

E-mail: ISHII.Masashi@nims.go.jp

【序】 大規模言語モデル（Large Language Model, LLM）の影響が材料科学にも波及してきている。信頼性や、ハルシネーション、コストの問題など多くの課題が残っている中で、LLM を新規材料予測に使う意味や役割を、客観的に考える必要がある。

既報では、トークン数が限られている LLM であっても、RAG (Retrieval -Augmented Generation)を使うことで、専門知識の統合が可能であるものの、いわゆる外挿的な新規発想に至らなかった[1]。新規発想は、むしろ膨大なデータ空間からの内挿、すなわち疑似的な外挿（疑似外挿）から得られると考えた。ここでは、LLM の疑似外挿による FePt 系の新規高密度磁気記録媒体の開発を試行した。

【方法】本研究の新規材料開発では、思考するモデルと呼ばれ、推論に優れていると考えられる Google の Gemini 2.5 Flash 及び Gemini 2.5 Pro を使用した。有効な疑似外挿を行うために、微細な FePt 構造を作る際の課題の一つである Segregant の選択にフォーカスした。すなわち関連の 6 論文を事前に読み込ませ、課題と解決の方向性を共有したうえで予測を実施した。未知の Segregant を予測するために、ソフトマックス関数の Temperature をパラメータとした。

【予測結果】 Gemini 2.5 Flash は、新規物質予測の可能性を高めた Temperature=1.7 の時、有用な Segregant としてフッ化物（CaF<sub>2</sub>, SrF<sub>2</sub>, LaF<sub>3</sub>

など）を予測した。例えば Google scholar のタイトル検索結果には、FePt 系の Segreant にフッ化物はなく（図 1）、予測の新規性がわかる。

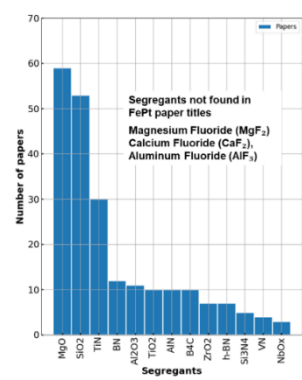


図 1 FePt 系の Segregant と論文数

【実験】LaF<sub>3</sub>を Segregant として MgO 上に FePt 粒を 400℃でスパッタリング堆積したところ、粒塊密度 2.11×10<sup>12</sup>/cm<sup>2</sup> の良好な FePt ナノグラニューラー薄膜が得られ（図 2）、LLM の予測結果を裏付けた。

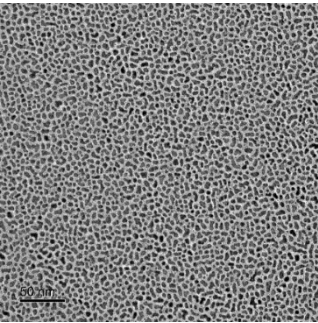


図 2 FePt-LaF<sub>3</sub> グラニューラー膜の TEM 像

【謝辞】 本研究は文部科学省 DxMT プロジェクト JPMXP1122715503 の助成を受けました。

【参考文献】 [1] 石井真史他：LLM を使った材料開発(I),第 85 回応用物理学会秋季学術講演会

# ChatGPT を用いた, 材料物性間の関係性知識グラフ及び関係性探索ツールの作成

## Constructing Materials Property Relationship Knowledge Graph and Tool for Relationship Search by ChatGPT

NIMS<sup>1</sup>, MatQ-lab<sup>2</sup> °吉武 道子<sup>1,2</sup>, 長田 貴弘<sup>1</sup>

NIMS.<sup>1</sup>, MatQ-lab.<sup>2</sup>, °Michiko Yoshitake<sup>1,2</sup>, Takahiro Nagata<sup>1</sup>

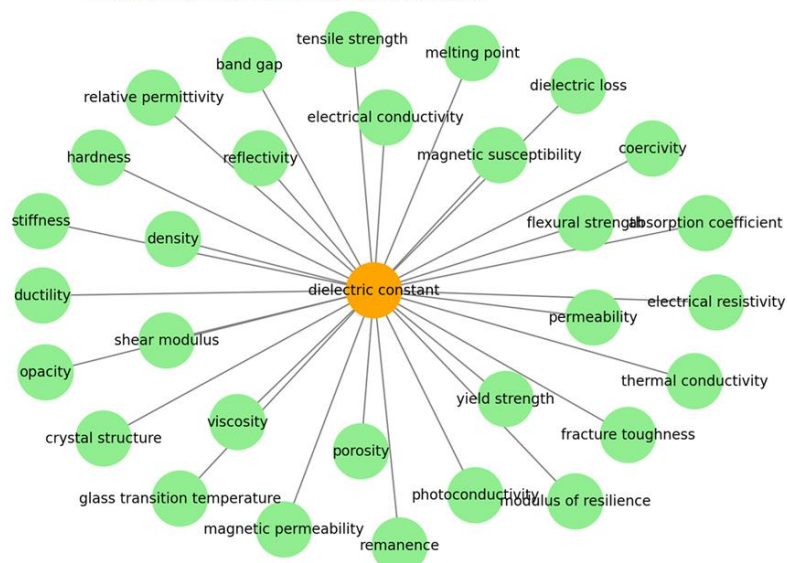
E-mail: yoshitake.michiko@nims.go.jp

著者の一人は, 物性間の関係性を知識グラフにして利用する「マテリアルキュレーション」を提唱し, 多数の NIMS 単独特許[1]の内の一部を実装したプロトタイプを富士通株式会社と共同で開発した[2]。当時は LLM の出現前で関係性の抽出は構文解析を基盤としていたが, 現在では LLM を用いて物性間の関係性の知識グラフが作成できることを前回の応用物理学会で発表した[3]。今回, ChatGPT-Plus 会員が利用可能な MyGPT の機能[4]を利用して, 前回作成した物性間の関係性の知識グラフから関係性を探索するツールを作成したので報告する。

グラフの探索は, 基本的に 2 つの物性間のつながりを探索するモード (2 つの物性が, 他どの物性との関係を介して関係を持っているか) と, 1 つの物性が別のどの物性と関係しているかを探索するモードがある。作成した MyGPT では, 最初にどちらの探索をしたいかを尋ね, 番号で選択を入力すると, 物性名のリストが出て, 物性名の入力を促される。物性名を入力すると望みの探索結果が表示される。図に示したのは, 探索の種類を聞いてくる画面と, dielectric constant と繋がっている物性を探索した結果の画面の例である。この MyGPT をブラッシュアップ後に公開して, 以前共同研究により作成したプロトタイプと似た探索が多くの方に体験いただけるようにしたいと考えている。

### どの探索を行いますか？

1. 物性間の経路探索を行う  
→ 2つの物性名の間を最短経路でつなぐ経路を探索・表示します。
2. ある物性の周りの関係を探索する  
→ 指定した1つの物性名に直接つながる物性だけを表示します。



[1] 特許第 6719748 号など 18 個

[2] Yoshitake, M.; Sato, F.; Kawano, H. J. Surf. Anal. 2020, 27, 22–33.

[3] 吉武 道子, 長田 貴弘, 2025 年第 72 回応用物理学会春季学術講演会 16p-K505-14.

[4] <https://help.openai.com/en/articles/8798878-building-and-publishing-a-gpt>



## 元素置換系のパーシステントホモロジー検討

### Persistent Diagram for Crystals with Partially Substituted Atoms

NIMS °吉武 道子, 長田 貴弘

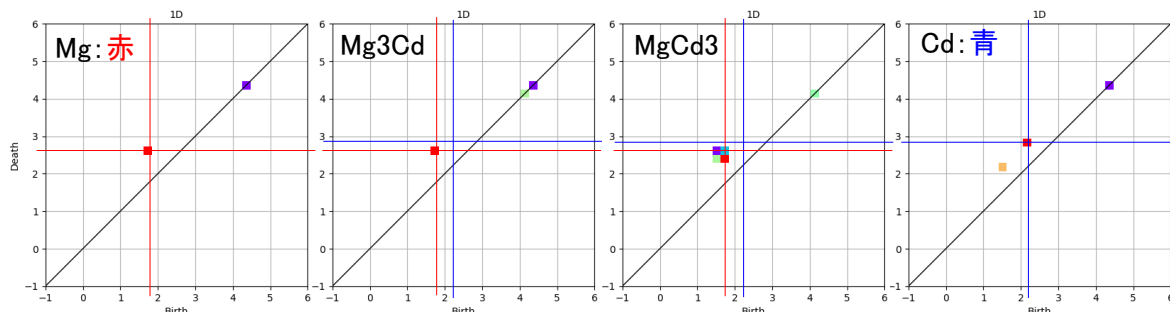
NIMS °Michiko Yoshitake, Takahiro Nagata

E-mail: yoshitake.michiko@nims.go.jp

固体イオン導電体や高エントロピー合金、誘電体など重要な材料の多くは、基本となる組成・構造の材料の一部の元素を置換して性能を飛躍的に高めることが多い。結晶情報（対称性、格子定数や原子座標）は CIF ファイルという統一した形式で表記されており、元素を一部置換した物質は、CIF ファイルの原子占有率という欄が1ではなく部分置換の割合が入っている。通常のパーシステントホモロジー（PH）では、原子占有率の考慮はできず、CIF ファイルの原子占有率の値が1でなくても1と解釈されてパーシステント図（PD）が作成される。合金のPDの作成を試みる最も手っ取り早い方法は、合金構成純金属のPDの組成に応じた加重平均を取ることだと考えられる。そこで本研究では、純金属（原子占有率が1）のPDを組成に応じて加重平均することで合金のPDが得られるかどうかを、規則合金（原子占有率が1のCIFファイルが得られ、かつ原子占有率が1でない合金とみなすこともできる）を用いて検証した。

2つの構成元素とその元素からできる規則合金の結晶構造（空間群）が同じ2元系をAtomWork[1]を用いて探したところ、Mg-Cd系、Li-Ag系、Cu-Pt系で見つけることができた。この中で空間群が同じ規則合金の組成がより1:1に近いMg-Cd系について、純元素結晶、規則合金結晶に対して、マテリアルプロジェクト[2]のCIFファイルを入力としてHomCloud[3]を用いてPD作成を行った。初期原子半径として結晶半径[Mg: 0.86, Cd: 0.92]を入力した場合について規則合金のPDが純元素結晶のPDの加重平均で表すことができるかを検討した。

図に示したのは、純元素結晶と規則合金結晶のCIFファイルを用いて作成したPDである。規則合金のPDは赤い四角がMgとCdのPDの加重平均位置には無く、純元素結晶の加重平均で表すことができないことが判明した。当日はより詳細で多彩な検討を行った結果も発表予定である。



[1] AtomWork, <http://crystdb.nims.go.jp/> (accessed on 2025.04.17)

[2] Materials Project, <https://legacy.materialsproject.org/> (accessed on 2025.04.17)

[3] Ippei Obayashi, Takenobu Nakamura, Yasuaki Hiraoka, J. Phys. Soc. Jpn. 91, 091013 (2022)