

# 適応的実験計画と高分子物性自動計算の融合

南條 舜

総合研究大学院大学 統計科学専攻 博士課程 (3年次編入学) 4年

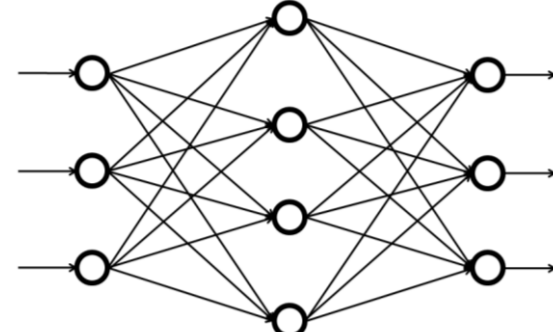
## 背景

### ■ 高分子を対象としたデータ駆動型材料研究における問題点

- データ数が少ない
  - 特定の物性に着目した場合、わずか100データにも満たない場合が存在<sup>1</sup>
- 予測したい領域のデータが存在しない
  - 高熱伝導化 ( $> 10$  [W/m·K]) が求められる一方、既存の高分子の熱伝導率は低い ( $< 1$  [W/m·K])

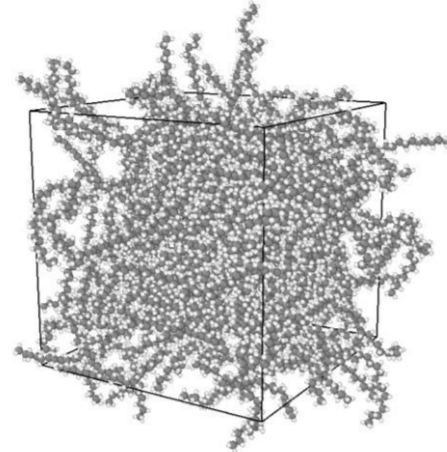
### ■ データの不足を補うための方法：機械学習と分子シミュレーションの融合

機械学習



- ✓ 大量の予測が可能
- × 入力が高ければ出力も近いという原理に基づくため 外挿領域の予測が困難

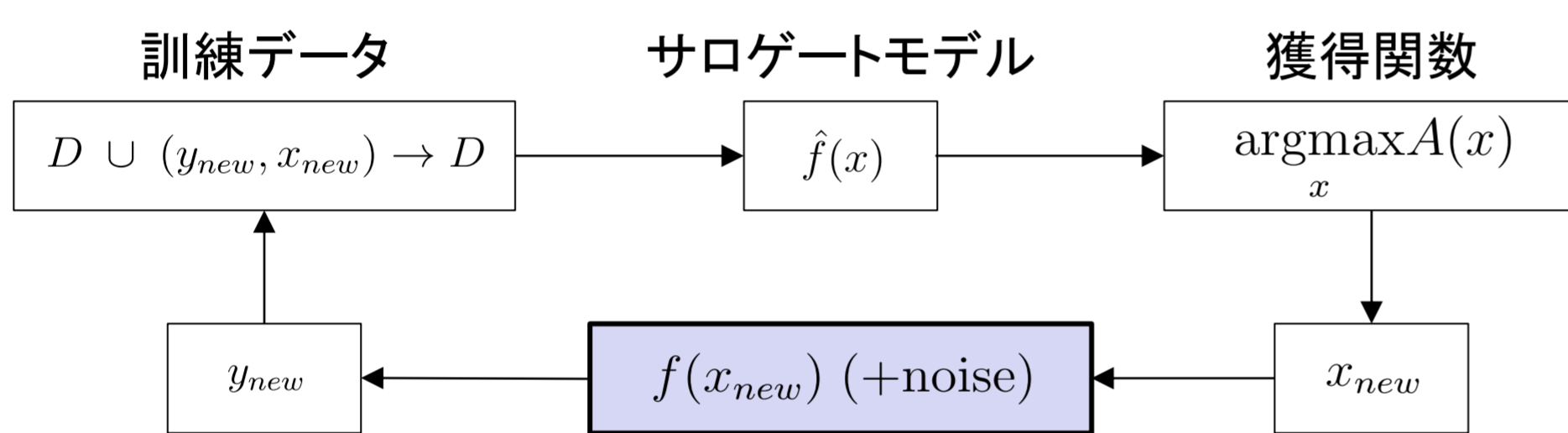
分子シミュレーション



- ✓ 物理法則に基づくため、外挿領域の予測がある程度可能
- × 大量の予測が困難

## 手法

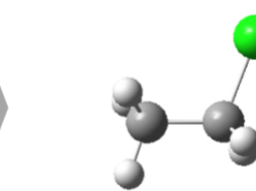
### ■ 適応的実験計画と高分子物性自動計算の循環



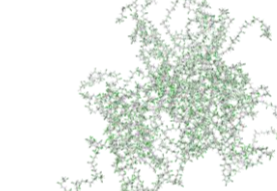
RadonPyでは高分子物性の全自動ハイスループット計算が可能

Input

- Repeat unit
- Degree of polymerization etc.



DFT calculation



MD simulation

Output

- Thermal conductivity
- Density etc.

- 未踏領域の計算データを作りだし、そのデータを含めてモデルを訓練 → 外挿的予測性能を徐々に獲得
- 高分子材料の場合、物性計算の自動化・高速化が技術的障壁 → RadonPy<sup>2</sup>を組み合わせることにより解決

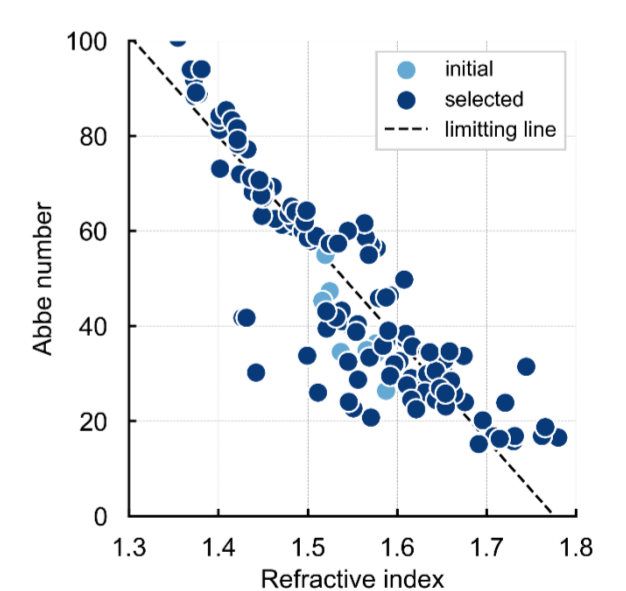
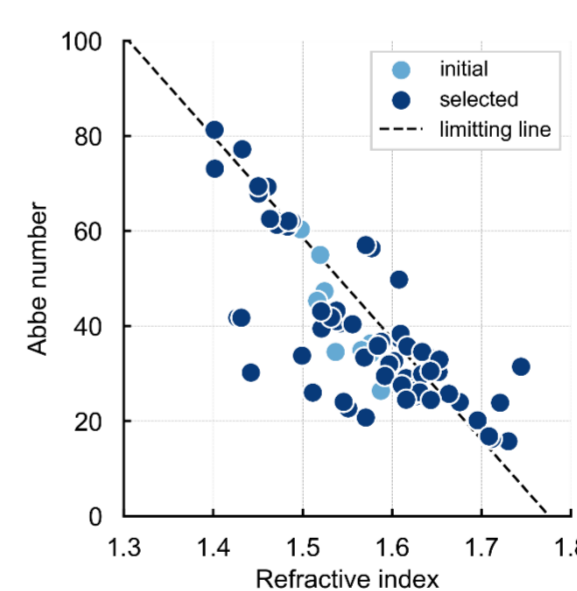
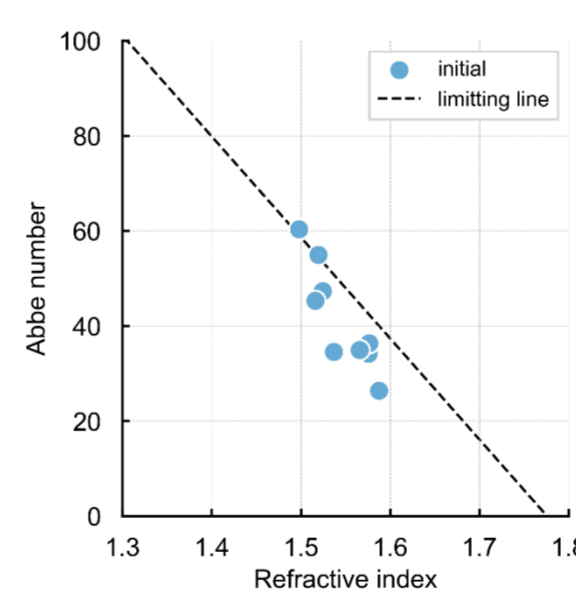
## 実験

### ■ 光学用高分子の探索

両物性の間には経験的な限界線が存在<sup>3</sup>

文献3の図1

上記手法により収集された物性計算データ



- 探索が進むにつれて、経験的な限界線を越える高分子が徐々に出現 → 現実世界での合成および物性評価を実施中

実験条件

- サロゲートモデルに通常のガウス過程帰帰モデル、モデルの訓練の入力に物性計算のパラメータをカーネル平均埋め込みにより固定長化したベクトル、獲得関数に期待超体積改善量を使用。
- プールデータに化学反応のルールに基づく生成器 SMiPoly<sup>1</sup>により生成された仮想高分子約10万構造を使用。

## まとめ

- 自動分子設計（機械学習）と高分子物性自動計算（RadonPy）を融合した初めての例となる
- RadonPyで計算可能な17種類の物性、もしくは計算結果をソースデータとして転移可能な物性であれば今回のワークフローが適用可能であるため汎用性が高い

[1] Wu, S., Kondo, Y., Kakimoto, M. A., Yang, B., Yamada, H., Kuwajima, I., ... & Yoshida, R. (2019). Machine-learning-assisted discovery of polymers with high thermal conductivity using a molecular design algorithm. *npj Computational Materials*, 5(1), 66.

[2] Hayashi, Y., Shiomi, J., Morikawa, J., & Yoshida, R. (2022). RadonPy: automated physical property calculation using all-atom classical molecular dynamics simulations for polymer informatics. *npj Computational Materials*, 8(1), 222.

[3] Ando S. (2015). Control of Refractive Indices of Optical Polymers: Theoretical Prediction and Molecular Design Method. *Japanese Journal of Optics*, 44(8), 298-303.