IMI Workshop of the Joint Usage Research Projects

Perspectives on Artificial Intelligence and Machine Learning in Materials Science

Editors: **Daniel PACKWOOD, Pierluigi CESANA, Shigenori FUJIKAWA, Yasuhide FUKUMOTO, Petros SOFRONIS, Alex STAYKOV**

九州大学マス・フォア・インダストリ研究所

IMI Workshop of the Joint Usage Research Projects

Perspectives on Artificial Intelligence and Machine Learning in Materials Science, February 4-6, 2022

■**Editors:**

Principal Investigator: Daniel PACKWOOD, iCeMS Kyoto University

Co-organizers (Alph.):

Pierluigi CESANA, IMI Kyushu University

Shigenori FUJIKAWA, I^{2}CNER Kyushu University

Yasuhide FUKUMOTO, IMI Kyushu University

Petros SOFRONIS, UIUC and I^{2}CNER Kyushu University

Alex STAYKOV, I^{2}CNER Kyushu University

About MI Lecture Note Series

The Math-for-Industry (MI) Lecture Note Series is the successor to the COE Lecture Notes, which were published for the 21st COE Program "Development of Dynamic Mathematics with High Functionality," sponsored by Japan's Ministry of Education, Culture, Sports, Science and Technology (MEXT) from 2003 to 2007. The MI Lecture Note Series has published the notes of lectures organized under the following two programs: "Training Program for Ph.D. and New Master's Degree in Mathematics as Required by Industry," adopted as a Support Program for Improving Graduate School Education by MEXT from 2007 to 2009; and "Education-and-Research Hub for Mathematics-for-Industry," adopted as a Global COE Program by MEXT from 2008 to 2012.

In accordance with the establishment of the Institute of Mathematics for Industry (IMI) in April 2011 and the authorization of IMI's Joint Research Center for Advanced and Fundamental Mathematics-for-Industry as a MEXT Joint Usage / Research Center in April 2013, hereafter the MI Lecture Notes Series will publish lecture notes and proceedings by worldwide researchers of MI to contribute to the development of MI.

October 2022
Kenji Kajiwara
Director, Institute of Mathematics for Industry

IMI Workshop of the Joint Usage Research Projects

Perspectives on Artificial Intelligence and Machine Learning in Materials Science, February 4-6, 2022

Perspectives on Artificial Intelligence and Machine Learning in Materials Science, February 4-6, 2022

Artificial Intelligence has led to a paradigm shift in investigation in Materials Science, with Machine Learning allowing informatics-based systematic calculations, predictions and discovery based on material databases pushing beyond the intrinsic limitations of first-principles calculations. However, the successful application requires development of novel methodologies inspired by the frontends of materials development in close synergy between physical science and Information Technology.

For example, demand is growing for functional molecules capable of performing specific work cycles such as assembly and transportation of molecules, cyclization and cycloreversion. These specific functionalities require the design and customization of molecular architectures at the atomistic level. This is a task which requires quantum chemistry modeling (so that chemical and physical properties are computed as the result of the atomic composition of the architecture) and inverse design (so that the atomic composition→ property mapping can be effectively inverted). Nowadays, complex numerical platforms blending first-principle computations, genetic algorithms for the exploration of high-dimensional chemical spaces, machine learning models for regression and classification over large pools of molecules and bioinformatics techniques for the digitalization and codification of molecular formulas have been largely exploited for the discovery of dielectric materials, perovskites, nanoparticles and more. However, outstanding questions remain unexplored, such as the coordination of nano- and meso- scale optimization programs, the identification of effective molecular descriptors capable of passing on relevant information to train machine learning models, the definition of suitable surrogate models and optimization strategies and many more.

The purpose of this conference was to gather an international group of scientists and researchers in industry and academia bringing their own distinct perspectives on problems at the intersection of Materials Physics and Information Technology, two areas where interdisciplinary collaborations both at the academic and industrial level are crucial and yet to date in their early phases.

Despite the conference was held in on-line version, the 3 days were rich in discussions and saw the participation of approximately 50 participants including from North America and China. Among topics covered were the exploitation of supervised learning techniques for high-throughput chemical calculations, regression methods to discover relations among meso-scale phenomena and nano/micro properties, genetic algorithms, materials characterization, data-driven approaches to dynamical systems, design of materials with Artificial Intelligence techniques.

The conference was organized as an IMI joint usage Research Seminar 1 and co-sponsored by the MOZES program. The conference consisted of 12 invited lectures and a students' session which saw 11 students' talks. We hope this note and the material contained herein remain as a reference and a guide for researchers operating at the overlap of these areas.

Principal Investigator: Daniel PACKWOOD (Junior Associate Professor/Institute for Integrated Cell-Material Sciences, Kyoto University)

Co-organizers (Alph.)
Pierluigi Cesana, IMI Kyushu University
Shigenori Fujikawa, I^{2}CNER Kyushu University
Yasuhide Fukumoto, IMI Kyushu University
Petros Sofronis, UIUC and I^{2}CNER Kyushu University
Alex Staykov I^{2}CNER Kyushu University

FY2021 IMI Joint Usage Research Project

Perspectives on Artificial Intelligence and Machine Learning in Materials Science

2022.2/4▸6
(Fri) (Sun)

Invited Speakers:

Sherry Chen (Hong Kong University of Science and Technology)
Data-driven approach for indexing-free X-ray crystallography

Xiangdong Ding (Xi'an Jiaotong University)
Artificial intelligence enhanced atomic simulations of complex phase transformation

Shigenori Fujikawa (I²CNER, Kyushu University)
Membrane design for negative CO2 emissions technology

Kulbir Kaur Ghuman (INRS)
Designing energy materials via genetic algorithm based machine learning approach

Gergely Juhasz (Tokyo Institute of Technology)
Electrochemistry of nanoparticles and nanotubes

Yu Kaneko (DAICEL Co.)
Introduction of machine learning approach for chemical compound search in Daicel

Yoshinobu Kawahara (IMI, Kyushu University and RIKEN)
Data-driven analysis of dynamical systems: from an operator-theoretic perspective

Ippei Obayashi (RIKEN)
3D data analysis of X-ray CT images with persistent homology and nonnegative matrix factorization

Daniel Packwood (iCeMS, Kyoto University)
Data-driven approaches for surface materials and beyond

David Rivera (Hiroshima University)
Density functional theory and machine learning description and prediction of oxygen atom chemisorption on platinum nanoparticles

Petros Sofronis (I²CNER, Kyushu University and UIUC)
Mathematical challenges in linking the micro-with macro-scale behavior of materials in adverse chemomechanical environments

Ichigaku Takigawa (RIKEN and Hokkaido University)
Machine learning for molecules: lessons and challenges of data-centric chemistry

Venue: Ito Campus, Kyushu University (Lecture room/W1-C512) and Online

Institute of Mathematics for Industry
Kyushu University

MOZES Moonshot for beyond Zero-Emission Society

Joint Research Center for Advanced and Fundamental Mathematics-for-Industry
文部科学大臣認定「産業数学の先進的・基礎的共同研究拠点」
九州大学マス・フォア・インダストリ研究所

Perspectives on Artificial Intelligence and Machine Learning in Materials Science

Date ： February 4 (Fri) – 6 (Sun), 2022
Place ： Lecture Room (W1-C-512), West Zone 1,
　　　　Ito campus, Kyushu University and Live streaming by Zoom
Principal Investigator ： Daniel PACKWOOD (Institute for Integrated Cell-Material Sciences, Kyoto University)

February 4 (Fri), 2022 (Chair: Alex STAYKOV)

08:30-08:40
Opening Remarks (Yasuhide FUKUMOTO)

08:40-09:40
Mathematical Challenges in Linking the Micro-with Macro-Scale Behavior of Materials in Adverse Chemomechanical Environments
Petros SOFRONIS (I^{2}CNER, Kyushu University and UIUC)

09:40-10:40
Designing Energy Materials via Genetic Algorithm Based Machine Learning Approach
Kulbir Kaur GHUMAN (INRS)

10:40-11:00 Break

11:00-12:00
Density Functional Theory and Machine Learning Description and Prediction of Oxygen Atom Chemisorption on Platinum Nanoparticles
David RIVERA (Hiroshima University)

12:00-13:00
Electrochemistry of Nanoparticles and Nanotube
Gergely JUHASZ (Tokyo Institute of Technology)

13:00-14:00 Lunch Break

14:00-16:00 Student Talks Session

February 5 (Sat), 2022 (Chair: Daniel PACKWOOD)

08:30-09:30
Membrane Design for Negative CO2 Emissions Technology
Shigenori FUJIKAWA (I^{2}CNER, Kyushu University)

09:30-10:30
Introduction of Machine Learning Approach for Chemical Compound Search in Daicel
Yu KANEKO (Daicel Corporation)

10:30-11:00 Break

11:00-12:00
Machine Learning for Molecules: Lessons and Challenges of Data-Centric Chemistry
Ichigaku TAKIGAWA (RIKEN and Hokkaido University)

12:00-13:00
3D Data Analysis of X-Ray CT Images with Persistent Homology and Nonnegative Matrix Factorization
Ippei OBAYASHI (RIKEN）

13:00-14:00 Lunch Break

14:00-16:00
Student Talks Session

February 6 (Sun), 2022 (Chair: Pierluigi CESANA)

08:30-09:30
Data-Driven Approaches for Surface Materials and Beyond
Daniel PACKWOOD (iCeMS, Kyoto University)

09:30-10:30
Data-Driven Analysis of Dynamical Systems: From an Operator-Theoretic Perspective
Yoshinobu KAWAHARA (IMI, Kyushu University and RIKEN)

10:30-11:00 Break

11:00-12:00
Data-Driven Approach for Indexing Free X-Ray Crystallography
Sherry CHEN (Hong Kong University of Science and Technology)

12:00-13:00
Artificial Intelligence Enhanced Atomic Simulations of Complex Phase Transformation
Xiangdong DING (Xi'an Jiaotong University)

13:00-13:10
Closing Remarks

13:10-14:00 Lunch Break

14:00-16:00
Room Breakout/Free Discussions/Student Talks Session

Table of contents

Combining Machine Learning and Density Functional Theory to Predict the Oxygen Atom Chemisorption on Platinum Nanoparticles

David Samuel RIVERA ROCABADO

Graduate School of Advanced Science and Engineering,
Hiroshima University, Japan

INTRODUCTION

Predicting the performance of a catalyst based on theoretical models is crucial for faster and less expensive design of materials with specific catalytic functions and increased lifetime. When the strong interaction between the catalyst and adsorbate scales to the catalyst reactivity, it is essential to elucidate the variables that can suitably describe the energetics of the adsorbent-adsorbate interaction.[1,2] For polymer electrolyte fuel cells (PEFCs), the energies of the O atom and OH adsorption onto various metals exhibit a linear scaling relationship with the limiting potential.[3] To such extent that experimental and density functional theory (DFT) results suggest that a surface capable of binding the O atom up to 0.4 eV less stable than Pt(111) should exhibit enhanced activity for the oxygen reduction reaction (ORR), leading to an increased performance of the PEFC.[4,5] During the ORR, O_2 is activated by proton and electron transfer to form OOH before O-O bond dissociation can occur. Following the OOH dissociation, the electrocatalyst should bind O and OH with moderate strength for ensuing the H_2O formation and desorption to be rapid; otherwise, the active sites will be covered by these species and will become inactive for O_2 dissociation, henceforth decreasing the output power density of the PEFC. First-principles computational modeling methods can help provide insight into the atomic structure, morphology, electronic properties, etc., hence would provide valuable information on crucial parameters in catalyst design. Among the theoretical approaches, the Hammer-Nørskov model,[6] linearly scales the d-band center with the adsorption energy displaying interesting linear correlations between the adsorption energies of atoms and small molecules on various metal surfaces.[7,8] Oppositely, the d-band center fails to account for the effect of low coordinated atoms, such as those located at the vertices and ridges of nanoparticles, especially in small cluster particles.[9,10] Different descriptors of the adsorption energy, such as the generalized coordination number (GCN), showed to be more appropriate for describing the O atom, O_2, OOH, H_2O, and H_2O_2 adsorption on the top site of Pt nanoparticles.[11] A similar limitation to the interaction on only the top site was reported for the energetics decomposition of the CO adsorption on platinum nanoparticles via machine learning.[12] On the other hand, more robust and complete models have been proposed using supervised learning to describe the adsorption energy between NO and 4d- and 5d-transitions metals,[13] and O atoms, and bimetallic nanoparticles with a Pt skin configuration,[14] considering all the adsorption sites.

Herein, DFT calculations were used to quantify the O atom interaction on Pt nanoparticles. Next, multiple linear regression analysis was used to describe the O atom binding energy on Pt nanoparticles based on a linear combination of five structural and electronic properties of the nanoparticles before the O atom interaction, which have been reported elsewhere.[15] The predicted O atom binding energy exhibits a significant linear correlation with the DFT-calculated one. From the model validation analysis, it was corroborated the accuracy and robustness of the model. This is the first model that describes and estimates the O binding energy on Pt nanoparticles considering for the first time the nanoparticle size effect and all the adsorption sites of the nanoparticles.

COMPUTATIONAL DETAILS

DFT calculations

All calculations performed in this study were based on the plane wave DFT method implemented in the Vienna *ab initio* Simulation Package (VASP 5.3.5).[16-18] Perdew−Burke−Ernzerhof parametrization under generalized gradient approximation was employed as the exchange-correlation functional together with the projector-augmented wave method.[19] Spin-polarized calculations were performed throughout the study with a plane wave cutoff energy of 400 eV. The convergence criteria for all calculations were set as the point at which the difference in the total energy between the two ionic steps was less than 10^{-4} eV/atom, and 10^{-5} eV/atom for self-consistent field iterations. Pt nanoparticles of varying sizes containing 13, 55, 201, and 405 atoms were modeled and optimized until reaching convergence. To avoid interactions between periodic images, the minimum distance between the cell boundaries and Pt atoms was set to 6 Å, that is, a minimum of 12 Å between neighboring image Pt nanoparticles. Nanoparticle optimization was performed at the Γ point in reciprocal space, owing to the significant spatial extent of the systems, wherein all the Pt atoms were allowed to relax.

The binding energy between O and isolated Pt nanoparticles (E_{Pt_n}), E_{bind}, was calculated as:

$$E_{bind} = E_{Pt_n+O} - E_{Pt_n} - E_O \qquad (1)$$

where E_{Pt_n+O}, E_{Pt_n}, and E_O denote the total energy of O atom interaction with the Pt nanoparticle, that of the isolated Pt nanoparticles, and of the isolated O atom, respectively. Using this definition of the binding energy, negative values denote a more stable interaction between the O atom and the Pt nanoparticles. The O atom interaction was calculated on the symmetric sites and all their possible combinations for each Pt nanoparticle. The Pt nanoparticles models and the symmetric adsorption sites of each nanoparticle are shown in Figure 1a. During the geometry optimization of the O binding to Pt nanoparticles, the same constraints were set for the Pt atoms motion as in the case of the isolated nanoparticles; all atoms were allowed to relax. Charge and bond orders analyses were performed using the 6th generation density-derived electrostatic and chemical (DDEC6) method.[20]
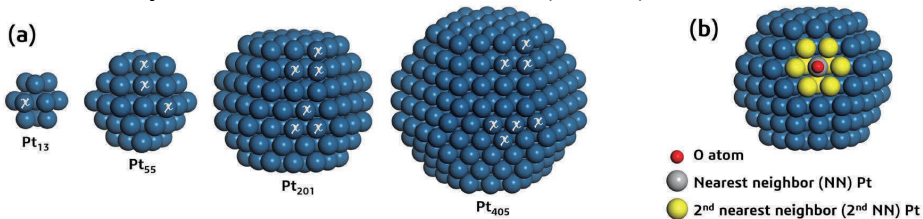
Figure 1. (a) models of the Pt nanoparticles and their symmetric adsorption sites are shown with white "x"s and (b) exemplification of the NN and 2nd NN of the adsorption site.

Multiple linear regression analysis

In this study, the model describing the O atom binding to Pt nanoparticles is obtained from linear combinations of the geometrical and electronic properties of isolated nanoparticles before their interaction with the O atom. The descriptors associated with the O atom binding energy were selected from among 20 descriptors related to the geometrical features and electronic properties of the nearest neighbors (NN) and second nearest neighbors (2nd NN) of the adsorption site. The definitions of NN and NN + 2nd NN of the adsorption site are shown in Figure 1b. Various combinations of descriptors were tested to estimate the O binding. Two criteria were carefully considered in constructing our models. The first one is the absence of multicollinearity; the highly correlated descriptors are not considered in the model. The second one is to omit the combinations of variables that do not significantly contribute to the R^2 value. Thus, descriptors with p-values above 0.05 were rejected because they were not statistically significant. Because multiple competing models can be considered, the Akaike's information criterion (AIC) is used to estimate the relative amount of information lost and simplicity of the model by including a penalty for each descriptor used to estimate a dependent variable.[21] Thus, the lower the AIC, the proposed model is simpler and closer to the optimal model, which is free of multicollinearity.[22] When the ratio between the data sample and the

number of descriptors is less than 40, the AIC may become inaccurate,[23] and the corrected AIC (AICC) should be used that is defined as follows:[24]

$$AIC_C = 2k + n\ ln\left\{\frac{\sum_{i=1}^{n}[E_{bind_i} - \widehat{E}_{bind_i}(x_i)]^2}{n}\right\} + \frac{2k(k+1)}{n-k-1} \qquad (2)$$

where k, n, E_{bind_i}, and $\widehat{E}_{bind}(x_i)$ are the number of descriptors in the model, the total number of samples, DFT-calculated, and predicted values of the O binding energy, respectively.

The model validation was performed using the holdout method, wherein the data sample was randomly divided into 3/4 of the data points for the training set and the remaining 1/4 for the test set. Multiple regression analysis was performed on the training set, and the prediction model was used to estimate and validate the test set values. Ten different training and test sets were randomly selected. To check the accuracy of the prediction, the descriptors were kept fixed during the validation process. The degree of discrepancy between the predicted and observed values was approximated by the mean absolute error (MAE).

RESULTS AND DISCUSSION

As the nanoparticle size increases, the strength of the O atom interaction decreases. The O atom chemisorption is thermodynamically more stable at the bridge sites, followed by 3-fold (3F) coordinated sites (FCC and HCP), and then the top sites of the Pt nanoparticles. For the same size Pt nanoparticles, the O atom interaction is more stable on under-coordinated atoms. For example, for the O atom binding on the top sites of Pt_{201}, the O binding energy decreases in the order: top of the {111} facet (-3.18 eV) > top of the {100} facet (-3.43 eV) > top of the ridge atoms (-3.92 eV) > top of a vertex atom (-4.13 eV). Such correlation between binding energy and adsorption site becomes difficult when considering the nanoparticle size effect because as the nanoparticle size increases, new adsorption sites comprised of atoms with different coordination numbers become available. All the configurations corresponding to the O atom chemisorption on Pt nanoparticles are summarized elsewhere.[25] When the *d*-band center or the GCN were used to describe and predict the O atom interaction to each adsorption site (top, bridge, and 3F) of the Pt nanoparticles, both the *d*-band center and GCN exhibited coefficients of determinations (R^2) larger than 0.50, and up to 0.88.[25] However, when all the adsorption sites were considered together, the *d*-band center and the GCN as sole descriptors of the O-Pt interaction, led to low coefficients of determination (R^2) of 0.33 and 0.50 and high MAEs, 0.31 and 0.29 eV.[25] Thus, the *d*-band center and the GCN cannot be used as the unique descriptor of the O atom binding energy on Pt nanoparticles.[25] To overcome this limitation, the O atom binding energy is described as a linear combination of five descriptors; the average NN Pt-Pt distance, GCN, the sum of NN bond orders and their coordinating atoms, the sum of NN + 2nd NN bond orders and their coordinating atoms, and the *d*-band centers of the atoms at the adsorption site. In general, as the nanoparticle size increases, the interatomic distance increases proportionally to the effective nanoparticle radius.[26] For the ORR, the activity improves with increasing Pt nanoparticle size because the interaction of oxygenated species becomes less stable with increasing the nanoparticle size.[9,27] Because the GCN considers information about the coordination number of the 2nd NN, the GCN can account for the changes in the adsorption sites and the nanoparticle size effect. Bond orders can provide valuable information regarding stability and activity trends. The sum of bond orders accounts for the interaction of all the atoms in the coordination sphere of the adsorption site, including those of the 3rd and 4th NNs. In the case of the *d*-band center, as the Pt nanoparticle size increases, the *d*-band center of the entire nanoparticle decreases toward the bulk Pt value.[15] Thus, the selected descriptors and their combination account for the changes in nanoparticle size and the different adsorption sites. Based on these five descriptors, the model for predicting the binding energy of O atom on Pt nanoparticles is defined as:

$$E_{bind}^{pred} = -6.681 - 0.197d_{Pt-Pt} + 0.521GCN + 0.226BO^{NN} - 0.058BO^{NN+2^{nd}\ NN} - 0.427\varepsilon_d \qquad (4)$$

where d_{Pt-Pt}, GCN, BO^{NN}, $BO^{NN+2^{nd}\ NN}$, and ε_d are the interatomic Pt-Pt distance, the GCN, the sum of NN bond orders, the sum of NN + 2nd NN bond orders, and the averaged *d*-band center of the NN of the adsorption site, respectively. Figure 2 illustrates the DFT-

calculated and predicted O atom binding energy relationship. The high R^2 (0.93) and low MAE (0.11 eV) values indicate an excellent correlation between the predicted and DFT-calculated O binding energies. From the model validation, the R^2 values for the test sets ranged from 0.847 to 0.969, and the MAEs from 0.085 to 0.167 eV, which confirm that the model is robust and appropriate for describing and predicting the O atom binding energy on Pt nanoparticles. This is the first model that can describe the O atom binding energy taking into consideration the Pt nanoparticle size effect and all the possible adsorption sites.
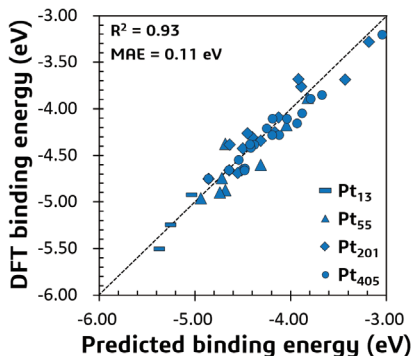

Figure 2. Relationship between the DFT-calculated and predicted O atom binding energies on Pt nanoparticles. The linear regression line, the coefficient of determination, and the mean absolute value are shown.

SUMMARY

Multiple regression analysis was performed to describe the O atom binding energy on all the possible adsorption sites of Pt nanoparticles of different sizes. The proposed model that describes and predicts the O atom binding on Pt nanoparticles employs a linear combination of the Pt nanoparticles' structural and electronic properties. The model provided an excellent correlation with the DFT-calculated adsorption energies. Model validation confirmed its accuracy and robustness in estimating the O atom binding energy.

References

[1] Han, B. C.; Miranda, C. R.; Ceder, G. Effect of Particle Size and Surface Structure on Adsorption of O and OH on Platinum Nanoparticles: A First-Principles Study. Phys. Rev. B 2008, 77 (7), 075410.

[2] İnoğlu, N.; Kitchin, J. R. New Solid-State Table: Estimating d-Band Characteristics for Transition Metal Atoms. Mol. Simul. 2010, 36 (7–8), 633.

[3] A. Kulkarni, S. Siahrostami, A. Patel, J.K. Nørskov, Understanding Catalytic Activity Trends in the Oxygen Reduction Reaction, *Chem. Rev.* 118 (2018) 2302–2312

[4] Greeley, J.; Stephens, I. E. L.; Bondarenko, A. S.; Johansson, T. P.; Hansen, H. A.; Jaramillo, T. F.; Rossmeisl, J.; Chorkendorff, I.; Nørskov, J. K. Alloys of Platinum and Early Transition Metals as Oxygen Reduction Electrocatalysts. *Nat. Chem.* **2009**, *1* (7), 552–556

[5] Krüger, S.; Vent, S.; Rösch, N. Size Dependence of Bond Length and Binding Energy in Palladium and Gold Clusters. *Ber. Bunsenges. Phys. Chem.* **1997**, *101* (11), 1640–1643.

[6] Hammer, B.; Nørskov, J. K. Theoretical Surface Science and Catalysis—calculations and Concepts. In Advances in Catalysis; Elsevier, 2000; Vol. 45, pp. 71–129

[7] Lu, C.; Lee, I. C.; Masel, R. I.; Wieckowski, A.; Rice, C. Correlations between the Heat of Adsorption and the Position of the Center of the d-Band: Differences between Computation and Experiment. *J. Phys. Chem. A* **2002**, *106* (13), 3084–3091

[8] Kim, G.; Kawazoe, Y.; Lee, K.-R. Controlled Catalytic Properties of Platinum Clusters on Strained Graphene. *J. Phys. Chem. Lett.* **2012**, *3* (15), 1989–1996

[9] Han, B. C.; Miranda, C. R.; Ceder, G. Effect of Particle Size and Surface Structure on Adsorption of O and OH on Platinum Nanoparticles: A First-Principles Study. *Phys. Rev. B* **2008**, *77* (7), 0754101

[10] Takigawa, I.; Shimizu, K.; Tsuda, K.; Takakusagi, S. Machine Learning Predictions of Factors Affecting the Activity of Heterogeneous Metal Catalysts. In *Nanoinformatics*; Tanaka, I., Ed.; Springer Singapore: Singapore, 2018; pp 45–641

[11] Calle-Vallejo, F.; Martínez, J. I.; García-Lastra, J. M.; Sautet, P.; Loffreda, D. Fast Prediction of Adsorption Properties for Platinum Nanocatalysts with Generalized Coordination Numbers. *Angew. Chem. Int. Ed.* **2014**, *53* (32), 8316–83191

[12] Gasper, R.; Shi, H.; Ramasubramaniam, A. Adsorption of CO on Low-Energy, Low-Symmetry Pt Nanoparticles: Energy Decomposition Analysis and Prediction via Machine-Learning Models. *J. Phys. Chem. C* **2017**, *121* (10), 5612–56191

[13] Nanba, Y.; Koyama, M. NO Adsorption on 4d and 5d Transition-Metal (Rh, Pd, Ag, Ir, and Pt) Nanoparticles: Density Functional Theory Study and Supervised Learning. *J. Phys. Chem. C* **2019**, *123*, 28114–28122

[14] Nanba, Y.; Koyama, M. An Element-Based Generalized Coordination Number for Predicting the Oxygen Binding Energy on Pt_3M (M = Co, Ni, or Cu) Alloy Nanoparticles. *ACS Omega* **2021**, *6* (4), 3218–3226

[15] Rivera Rocabado, D. S.; Ishimoto, T.; Koyama, M. The Effect of $SnO_2(110)$ Supports on the Geometrical and Electronic Properties of Platinum Nanoparticles. *SN Appl. Sci.* **2019**, 1 (11), 1485

[16] Kresse, G.; Hafner, J. Ab Initio Molecular Dynamics for Liquid Metals. *Phys. Rev. B: Condens. Matter. Phys.* **1993**, *47*, 558-561

[17] Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for Ab Initio Total-Energy Calculations Using a Plane-Wave Basis Set. *Phys. Rev. B: Condens. Matter. Phys.* **1996**, *54*, 11169-11186

[18] Kresse, G.; Furthmüller, J. Efficiency of Ab-Initio Total Energy Calculations for Metals and Semiconductors Using a Plane-Wave Basis Set. *Comput. Mater. Sci.* **1996**, *6*, 15-50

[19] Blöchl, P. E. Projector Augmented-Wave Method. *Phys. Rev. B: Condens. Matter. Phys.* **1994**, *50*, 17953-17979

[20] Manz, T. A. Introducing DDEC6 Atomic Population Analysis: Part 3. Comprehensive Method to Compute Bond Orders. *RSC Adv.* **2017**, *7* (72), 45552–45581

[21] Akaike, H. Information Theory and an Extension of the Maximum Likelihood Principle; In: Parzen E., Tanabe K., Kitagawa G. (eds) Selected Papers of Hirotugu Akaike. Springer Series in Statistics (Perspectives in Statistics). Springer, New York, NY; **1998**; pp 199–213

[22] Hurvich, C. M.; Tsai, C.-L. Regression and Time Series Model Selection in Small Samples. *Biometrika* **1989**, *76* (2), 297–307

[23] Xin-She, Y. *Introduction to Algorithms for Data Mining and Machine Learning*; Elsevier, **2019**

[24] Burnham, K. P.; Anderson, D. R. *Model Selection and Multimodel Inference*; Burnham, K. P., Anderson, D. R., Eds.; Springer New York: New York, NY, 2004

[25] Rivera Rocabado, D. S.; Nanba, Y.; Koyama M. Density Functional Theory and Machine Learning Description and Prediction of Oxygen Atom Chemisorption on Platinum Surfaces and Nanoparticles. *ACS Omega* **2021**, *6* (27), 17424-17432.

[26] Roduner, E. Size Matters: Why Nanomaterials Are Different. *Chem. Soc. Rev.* **2006**, *35* (7), 583-592.

[27] Mayrhofer, K. J. J.; Blizanac, B. B.; Arenz, M.; Stamenkovic, V. R.; Ross, P. N.; Markovic, N. M. The Impact of Geometric and Surface Electronic Properties of Pt-Catalysts on the Particle Size Effect in Electrocatalysis. *J. Phys. Chem. B* **2005**, *109* (30), 14433–14440

Introduction to the machine learning approach to chemical compound search at Daicel

Yu Kaneko

Simulation department, Process Technology Division, Production
Management Headquarters Daicel Corporation

Mitsuru Ohno

Business Development Center, Innovation and Business
Development Headquarters, Daicel Corporation, Japan

The elementary machine learning activities of Daicel was presented from the perspective of chemical industrial R&D stage. The four examples related to new molecular / process design problem upon novel material development were shown in this presentation.

 Daicel corporation established at September 8-th of 1919. We celebrated its 100-th anniversary at 2019. Starting from celluloid, "Dainippon Celluloid Co. Ltd." was founded through the merger of eight celluloid manufacturing companies. The term of celluloid is same as cellulose nitrate. Although celluloid was an easy-to-process plastic with excellent physical properties, the flammability of this nitrified cotton product presented a hazard. In order to overcome this disadvantage with a noncombustible formula, the Company developed an acetate plastic made from acetate cotton. In 1933, we decided to enter the cellulose acetate business. We adopted the major policy of becoming self-sufficient in production of the raw materials of acetic acid and other products through integrated production from carbide and to develop related chemicals. This represents the origins of our current organic chemicals business. Daicel also planned to enter the photographic film business as a new venture intended to complement the existing celluloid business. In order to ensure that photographic film would eventually become a major product, Fuji Photo Film Co., Ltd. (currently, FUJIFILM Corporation) was established in 1934 as a comprehensive photography industry company. In 1960s, we started engineering plastics business and established Polyplastics Co. Ltd, which has been manufacturing polyacetal resin. In 2000s, the integrated production center was established in Abosh plant and began the integrated operation method known as "Production Innovations the Daicel way".

 The new R&D base called "Innovation Park" was made in 2017, integrating with central research center and Himeji technology head office. Its function is for creation of new products and development of innovative process technologies and promotes technological innovation throughout Daicel group. Many kinds of technological field people are gathered in one site. It is not so easy to corporate one issue with different kinds of people, since those points of view are too different. Specially, the machine learning technology is quite new. In order to show its effectiveness and extend the application range, we needed to have a clear objective of machine learning methods with their applications. The machine learning toward material development can be used for optimization of experimental parameters by tuning of controllable parameters or used for chemical material search, as reported everywhere. The first one can be achieved by material blending ratio variation or temperature alternation. The other one can be a difficult part, since it includes chemical material search via new chemical

synthesis. The shortest way to find new materials is an elucidation of what kinds of chemical functional group are needed to express a certain function. It requires many chemical analysis, hypothesis creation and computer chemistry simulation. If only data analysis of chemical compounds by machine learning can solve those difficulties in a very short time, it is expected to reduce many efforts. Therefore, the objective has been set as material search via machine learning methods.

There was no example at the starting period of this activity. A proof of its effectiveness was a problem to convince experimental researcher of using machine learning. The nature of a chemical compound must be originated from chemical structure itself. It consists of atoms and chemical bindings with specific meanings. For example, cellulose has a glucose repeating units. There are hydrophilic functional group which enables to form hydrogen bond. This hydroxyl group can be used as alternation of cellulose as seen in our company history, celluloid to cellulose acetate. As another example, propylene glychol diacetate (PGDA) is a neutral and small compound which behaves hydrophobic nature. It is used for ink solution or cleaner. But it is weakly polarized due to ester groups. Then, it can solve in water or ethanol solution too. Therefore, functional groups have meaning and determine the nature of chemical compound. From cheminformatics side, there are many studies to investigate chemical nature by functional group. The famous library known as RDkit [1] can decompose the chemical structure into numbers of arrays known as fingerprint, chemical descriptor. Each element of array indicates the existence of certain functional group. This classical representation method is used as a descriptor to perform machine learning studies in this presentation. During the talk, the four examples were explained.

The first one was a solvent material search example. Some specific solid catalyst requires a suitable solvent. Depending on solvent choices, the reaction rate and its magnitudes of biproduct suppression was changed. Identification of an effective solvent was difficult because a common functional group character could not be guessed in the human point of view. The catalyst reaction takes place on the solid catalyst surface. At the reaction event, there would be solvent-agent interaction on the catalyst surface. Solvent effect of reactivity and suppression of biproducts are hard to be determined from simple quantum chemical calculations due to lack of the knowledge of surface information. However, there would be effective substructures of solvent. The effective substructures can be decomposed in parts by the usage of fingerprint and can be extracted through machine learning. The prediction model was made from a ridge regressor using Morgan fingerprints. Then we obtained the contributing factors of a solvent's chemical structure to reactivity and biproduct suppression. The contribution of each functional group was visualized in a picture. It was found that the ester group would enhance reactivity. This could not be found from even exert engineers because the compound including ester group showed poor behavior in some cases. The candidate solvent was tested experimentally and was found to be superior to others. Within 2 weeks, the one of candidate solvents was shown as PGDA, which is easy to be obtained in Daicel, since it is our product, CELTOL® (Fig.1).

The next example was also related to the molecular design. The concept of thermosetting oligomer is consisted of cross-linking functional group and engineering plastic region. Due to the cross linker, it is easy to mold the resin in any shape with heating. The machine learning method was applied to the engineering plastic region's molecular design for having lower dielectric dispersion and higher heat resistance. All the data was used for training until experimental analysis report could meet the prediction result of candidates. The evaluation cycle consisted of data gathering, machine learning, prediction, chemical synthesis, and evaluation. It took three cycles to be completed. At the final stage of cycle, the dielectric loss tangent prediction model by support vector regressor and Layered fingerprint can predict the experimental data of new compounds. Also, other physical value prediction models from the oligomer series data were created to draw a physical value map to find its application range. Even a material of new concept such as our oligomer, physical value blending with internal values, less than 50 points and outside values, about hundreds can be used for the specific prediction.

As shown above two examples, a simple model can tell undiscovered functional group

identification from small amount of data. Therefore, we imagine that such model making could lead innovative knowledge within short period. Most of chemists tends to do many experiments and gathering evidences and making their own hypothesis other than mathematic modeling. If machine learning can be used casually without programing, R&D cycle can be shortened. PyCaret [2] is known as a low-code machine learning library. The library defines test, training and validation datasets automatically. It performs model comparison, tuning, and evaluation. The autoML function returns the best model of all models created in the active environment base on the metric defined as optimize parameter.

Next two example was performed by the usage of PyCaret. The organic functional products are used in a wide range of industrial fields such as coating, plastic modification, and electronic materials. One of those is a monomer like alicyclic epoxy compounds, CEL2021P®. It produces highly heat resistant cured polymer. Heat and UV cation can produce rapid polymerization. Almost no chlorine derived from the manufacturing method gives wide range usage. But sometimes, the epoxy group inclusion can be a cause of mutagenicity. The mutagenicity is defined as capacity of a chemical or physical agent to cause permanent genetic alterations. Every chemical compound should be tested to estimate mutagenicity. The AMES test report must be attached as a result of hazard investigation. The AMES assay is defined as a test for identifying potential carcinogens by studying their mutagenic effect on bacteria. Our past results indicate that all the epoxy compound will not cause such symptoms. Its prediction in R&D stage could be used for molecular design, avoiding unworthy chemical synthesis. The mutagenicity database was obtained from Berlin's machine learning group's website [3]. Usage of PyCaret's auto ML, the classification model was constructed from published data and our inhouse data using Morgan fingerprint. The model performance metric of 10-fold cross validation result was chosen as Area Under the Curve, in short, AUC. When the value of AUC is close to one, it could be a good model. The AUC is about 0.9, so it shows moderate predictor performance (Fig. 2) The confusion matrix of test data in Fig.2 shows reasonable behavior with internal data. It can be used for raking of candidate compound synthesis. Having inhouse data is an advantage of this prediction construction, since the predictor behaves better performance than only public data.

Then, material of suspected toxicity can be avoided from molecular design stage. However, all the developed material will be tested in the point of function and toxicity in experiment. The prediction performance is almost same as published result [4], although our code is quite simple without using deep learning technique. When database would be updated by users, users can make new prediction model easily, changing database file name.

The final example is cellulosic derivatives. Daicel has developed cellulosic derivative since 1919. It is focused and considered as sustainable material. The substituent functional group and degree of substitution per glucose (DS) are important factors of functional alternation. The glass transition temperature, Tg can be changed by such alternation. We prepared published polymer data and gathering Daicel's Cellulosic derivative data about 30 points. Then, we constructed a predictor by the usage of PyCaret. We changed the functional group largely and set DS range from two to three. We draw the population of Tg along temperature and DS in Fig.3. As DS reaches to 3.0, Tg peak locates around 100 Celsius degree. On the other hand, as DS reaches to 2.0, its peak located around ninety to 120-130 Celsius degree. Therefore, As DS increases, the average Tg gets lower. The reason is not so obvious, but in the point of hydrogen bond formation, lower contents of hydroxyl group might give weaker interaction between cellulose chains. Also, Tg can be changes largely as R functional group changes. Such understanding of physical function is very important knowledge to determine the substitution methods.

The simple machine learning method is quite useful in the case that one could make a reasonable model in a short time period. It should be noted that there is a limitation to use simple RDkit fingerprint only. Usage of Pycaret library enables of non-experts of machine learning to use predictors. Thanks to this simplification, rapid trial can be done inside the R&D sections. In the next step, we will expand its application area toward process engineering for suitable and rapid business plan construction.

Instant predictor performance
Morgan fingerprint/Ridge regression

y = 0.9753x
R² = 0.9265

Reaction rate (predicted) [mmol/h]

Reaction rate (measured) [mmol/h]

Contribution factor's map

Green: Enhance reactivity
Red: Suppress reactivity

Candidate structure

CELTOL®
Propylene glychol diacetate （PGDA）

Two ester group

Fig.1. Result of example 1

Alicyclic epoxy compound example (CEL2021P)

Specific features of an alicyclic epoxy compound
1) Generation of highly heat-resistant cured product
2) High polymerization performance due to heat and UV cation
3) Almost no chlorine derived from the manufacturing method

10-fold cross validation result
(published and inhouse data, ~8k)
AUC=0.88
LGBM Classifier / Morgan fingerprint

LGBMClassifier Confusion Matrix

True Class

0	662	182
1	154	752

Predicted Class

Fig.2. AMES prediction result including our epoxy compounds

□ 0-2 □ 2-4 ■ 4-6 ■ 6-8

POPULATION

PREDICTED TG [°C]

DS
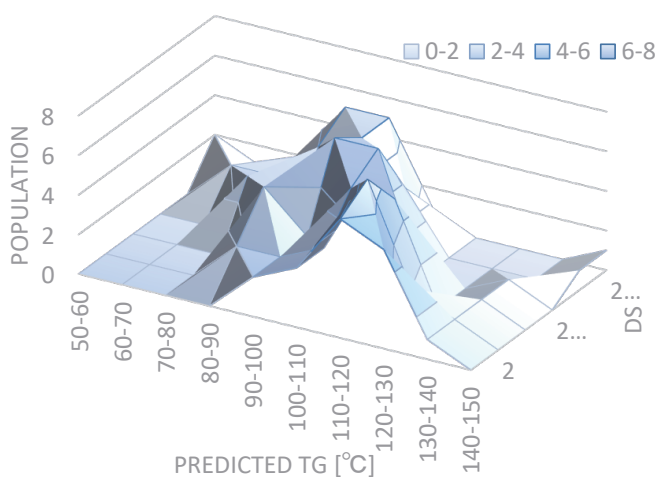
Fig.3. Population analysis of predicted Tg and DS for the candidate cellulose substitution series.

References

[1] RDKit: Open-source cheminformatics ; http://www.rdkit.org

[2] PyCaret: Open source low-code machine learning library: https://pycaret.org/

[3] Technische Universität Berlin's machine learning group website: http://doc.ml.tu-berlin.de/toxbenchmark/index.html

[4] 半田 千彰，小沢 知永，福澤 薫，米持 悦生，化合物のAmes予測におけるGraph Convolutional Networkの特徴評価, J. Comput. Che. Jpn., 20, 1-9 (2021)

Machine Learning for Molecules: Lessons and Challenges of Data-Centric Chemistry

Ichigaku TAKIGAWA

RIKEN Center for Advanced Intelligence Project, Japan
Institute for Chemical Reaction Design and Discovery, Hokkaido University

1. INTRODUCTION: MACHINE LEARNING WITH DISCRETE STRUCTURES

The current machine learning (ML) is grounded on a statistical principle, where we basically assume that data take a spreadsheet-like tabular form and each datapoint is regarded as a multidimensional random vector $x \in \mathbb{R}^d$ drawn from an unknown probability distribution. However, the practical data in chemistry often come with non-tabular and non-numeric information such as point sets of different sizes, chemical structures, reaction formulas, monomer sequences, and reaction networks. In particular, discrete structures, i.e., combinatorial or algebraic structures such as sets, groups, permutations, combinations, sequences, trees, and graphs are the most frequent and versatile data forms. However, it is technically challenging to leverage this discrete combinatorial information in statistical ML because we need to effectively fuse two different aspects of the *continuous* nature of statistical ML and the *discrete* nature of structural and combinatorial chemical information.

Here we briefly give a quick overview of how ML with discrete structures is important and promising for handling molecular datasets in chemistry-related fields. Molecules clearly have a combinatorial aspect. First, a molecule is a set of atoms and bonds and made up of combining small building blocks such as functional groups or any substituents like LEGO bricks. In this sense, it has clear modularity, compositionality, and hierarchy as we might see in natural language processing: We combine words in our vocabulary to make any complicated sentences, while chemists combine basic building blocks (atoms, bonds, scaffolds, functional groups, and substituents) to make any complex molecules. Chemical reactions between compounds are the main player in chemistry converting compounds to another and can be seen as a recombination pattern of bonds for the set of atoms. Second, a very interesting point for computer scientists is that the underlying rules that actually determine the combinations are (largely) governed by many-body quantum chemistry of electrons, which has *deductive* first principles requiring *approximation* in any case. In contrast, ML is learning from examples, which is totally an *inductive* principle.

2. THE DARK SIDE AND LIGHT SIDE OF MACHINE LEARNING

For a start, we first summarize practical insights and implications from machine learning research over past years. Leo Breiman's influential paper [5] published 20 years ago [1] introduces the "algorithmic modeling" to the statistics community, and now we see it is more dominating as ML than traditional statistical methods. The paper was very influential, but also somewhat controversial at that moment since "algorithmic modelling" encourages the use of black-box algorithms such as neural networks and

tree ensembles, which are basically task-agnostic, and contrastively different from the traditional "data modelling" approach in statistics. Interestingly, after 20 years we still see this conflict between two cultures of "algorithmic modelling" and "data modelling" in modern ML. Interestingly, "the three lessons" by iconic names discussed in Breiman's paper would be still valid when we think of applying ML in natural sciences:

- *Rashomon*: the multiplicity of good models;
- *Occam*: the conflict between simplicity and accuracy;
- *Bellman*: dimensionality—curse or blessing.

2.1. **Rashomon.** As we see in many papers or ML competitions, we typically can have many good but different ML/AI models having the same level of prediction capability for a given dataset. Breiman called this situation as the *Rashomon effect* named for a classic Japanese film *Rashomon* where we see contradictory interpretations or descriptions by the individuals involved but still all seem equally possible.

One of the reasons behind this would be *underspecification*. The current ML models are overrepresented and have a huge representation capability with an incredibly large number (several hundred million or several hundred billion) of parameters. Hence both in theory and practice, they are very data-hungry and require large data for them to work. But in reality, we can have finite, limited, noisy, and biased data at hand, and it's way smaller than these numbers in many cases. This would also mean that most data in daily cases are insufficient to fully specify the optimal model representing the true underlying mechanism even if any. Even most cases including so-called "big data" scenarios at commercial levels [7] might simply lack enough data, and often underspecify the model, which further exacerbates the Rashomon effect. ML is grounded on interpolatively capturing multidimensional correlations in the given data, and this fact makes any causal inference or learning challenging [23].

2.2. **Occam.** Overrepresentation and underspecification also cause *interpretability* problems in practice. Many ML models are too complicated for us to interpret what's going on inside. If we expect interpretability, we need to compromise on model complexity within the scope of our (very limited) recognition capability. Meanwhile, we use ML because we do not fully understand the target phenomena, which implies that the mechanism of the target phenomena is not simple, and is expected to be very complicated. That's why we need large-capacity ML models and large-scale datasets to realize successful applications. For example, the current ML models have too high expressive power and can have zero training error on random labels [13, 24].

2.3. **Bellman.** We tend to use many input variables because ML is completely unaware of any information not in the input variables. Missing relevant factors merely results in capturing spurious correlation. In the situations where deep learning works, we often use direct raw observations for input as it is. Typical cases are pixel values of images. An 100×100 RGB image corresponds to 30 thousand variables, and 1000×1000 to 3 million variables. This means that we're fitting a function model with hundreds million parameters in a several million dimensional space. Curiously, this shouldn't work in theory, but it sometimes works in practice. If we assume a mild continuity just imposing similar outputs for similar inputs, the required number of samples for sufficient fitting is of exponential order with respect to the number of input variables [4]. On any high-dimensional (> 100) dataset, the probability that a sample falls inside the given dataset's convex hull is zero [12], which challenges our concept of interpolation.

3. Molecular and Geometric Representation Learning

All three aspects of Rashomon, Occam, and Bellman suggest that ML is not trying to describe an underlying reality, and "essentially, all models are wrong, but some are useful." It is not guaranteed that ML captures any meaningful and insightful patterns in data, and we always have a risk of being trapped by spurious correlation [25]. We need to further constrain the model complexity not to unintentionally represent any chemically invalid functions, and that's exactly why additional considerations of discrete structures such as graphs and combinations [15] as well as physics-derived heuristics, assumptions, or first-principle theories [19, 22, 9, 20, 10, 17] are needed.

3.1. **Molecular and Geometric Graphs.** One of the most difficult problems is how to bettter represent a molecule for ML. We need to represent a molecule in a computer-readable form and input it to ML. For this purpose, the long history in chemoinformatics and computational chemistry has been devoted to develop several thousands of hand-crafted molecular descriptors [26]. However, we need to carefully choose descriptors for individual problems and targets, and here is why *representation learning* by deep architectures [6] is deemed as promising. For molecules, a general form representing all related information is *graphs* that, in molecular tasks, are roughly divided into two types: *molecular graphs* and *geometric graphs*. Molecular graphs are a graph, nodes for atoms and edges for bonds, with node and edge feature vectors encoding relevant information. Geometric graphs are molecular graphs embedded in Euclidean space, and each atom comes with the corresponding xyz coordinates in 3D. For example, standard 2D molecular graphs would be useful for drug screening for complex properties such as toxicity, efficacy, and potency, whereas 3D geometric graphs would be a good fit for DFT approximation, ML potentials, and ML force fields.

3.2. **Representation Learning with Graph Neural Networks.** In this direction, *geometric deep learning (GDL)* [14] is one of the recent hottest topics in ML. In particular, any geometric objects should be invariant under rotations, translations, and some local symmetries. In geometry, symmetry is all about the invariance of objects under transformations or group actions. Considering this general symmetry in representation learning is one of the big open issues as "the Erlangen Programme of ML"[14].

For DGL or molecular representation learning, one of the current fundamental methods are *graph neural networks (GNNs)* [27, 8, 14]. GNNs are one of the general modules in ML that directly takes graphs as inputs and outputs graphs or vectors. They are already used in various applications from travel time estimation in Google Maps or Baidu Maps to Siri triggering in Apple's iPhone to knowledge collection at Amazon, and also they are the technical core to consider ML for combinatorial optimization and reasoning [15]. Also, GNNs are tightly related to Transformers that is a core module that revolutionized natural language processing and computer visions, and Transformers can be seen as a special case of GNNs. Much research is now going on to fuse or redesign GNNs and Transformers more effectively and efficiently.

GNNs are both applicable to molecular graphs and geometric graphs [11] as well as other discrete structures such as sets, groups, trees, and graphs. An interesting usage is to generate molecular structures [16] that are similar to a given dataset and navigate the generation towards a desireble direction, for example, molecules having a specific property. Also, it can be useful to design ML potentials or ML force fields [2, 3] , or even ML density functionals [18, 21].

References

[1] Special Issue: Commentaries on Breimen's Two Cultures paper. *Observational Studies*, 7(1), 2021. https://muse.jhu.edu/issue/45147

[2] Special Issue: Machine Learning at the Atomic Scale, *Chem. Rev.* https://pubs.acs.org/toc/chreay/121/16

[3] Special Issue: Data Science Meets Chemistry, *Acc. Chem. Res.* https://pubs.acs.org/page/achre4/data-science-meets-chemistry

[4] David L Donoho and Others. High-dimensional data analysis: The curses and blessings of dimensionality. *AMS math challenges lecture*, 1(2000):32, 2000.

[5] Leo Breiman. Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Stat. Sci.*, 16(3):199–231, August 2001.

[6] Yoshua Bengio, Aaron Courville, and Pascal Vincent. Representation learning: a review and new perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.*, 35(8):1798–1828, August 2013.

[7] Alexander D'Amour, Katherine Heller, Dan Moldovan, Ben Adlam, Babak Alipanahi, Alex Beutel, Christina Chen, Jonathan Deaton, Jacob Eisenstein, Matthew D Hoffman, Farhad Hormozdiari, Neil Houlsby, Shaobo Hou, Ghassen Jerfel, Alan Karthikesalingam, Mario Lucic, Yian Ma, Cory McLean, Diana Mincu, Akinori Mitani, Andrea Montanari, Zachary Nado, Vivek Natarajan, Christopher Nielson, Thomas F Osborne, Rajiv Raman, Kim Ramasamy, Rory Sayres, Jessica Schrouff, Martin Seneviratne, Shannon Sequeira, Harini Suresh, Victor Veitch, Max Vladymyrov, Xuezhi Wang, Kellie Webster, Steve Yadlowsky, Taedong Yun, Xiaohua Zhai, and D Sculley. Underspecification presents challenges for credibility in modern machine learning. November 2020, arXiv:2011.03395 [cs.LG]

[8] William L Hamilton. Graph representation learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 14(3):1–159, September 2020.

[9] Frank Noé, Alexandre Tkatchenko, Klaus-Robert Müller, and Cecilia Clementi. Machine learning for molecular simulation. *Annu. Rev. Phys. Chem.*, 71:361–390, April 2020.

[10] David Pfau, James S Spencer, Alexander G D G Matthews, and W M C Foulkes. Ab initio solution of the many-electron schrödinger equation with deep neural networks. *Phys. Rev. Research*, 2(3), September 2020.

[11] Kenneth Atz, Francesca Grisoni, and Gisbert Schneider. Geometric deep learning on molecular representations. *Nat. Mach. Intell.*, 3(12):1023–1032, December 2021.

[12] Randall Balestriero, Jerome Pesenti, and Yann LeCun. Learning in high dimension always amounts to extrapolation. October 2021, arXiv:2110.09485 [cs.LG]

[13] Julius Berner, Philipp Grohs, Gitta Kutyniok, and Philipp Petersen. The modern mathematics of deep learning. May 2021, arXiv:2105.04026 [cs.LG]

[14] Michael M Bronstein, Joan Bruna, Taco Cohen, and Petar Veličković. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges, April 2021, arXiv:2104.13478 [cs.LG]

[15] Quentin Cappart, Didier Chételat, Elias Khalil, Andrea Lodi, Christopher Morris, and Petar Veličković. Combinatorial optimization and reasoning with graph neural networks. February 2021, arXiv:2102.09544 [cs.LG]

[16] Faezeh Faez, Yassaman Ommi, Mahdieh Soleymani Baghshah, and Hamid R Rabiee. Deep graph generators: A survey. *IEEE Access*, 9:106675–106702, 2021.

[17] Bing Huang and O Anatole von Lilienfeld. Ab initio machine learning in chemical compound space. *Chem. Rev.*, 121(16):10001–10036, August 2021.

[18] Bhupalee Kalita, Li Li, Ryan J McCarty, and Kieron Burke. Learning to approximate density functionals. *Acc. Chem. Res.*, 54(4):818–826, February 2021.

[19] George Em Karniadakis, Ioannis G Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. Physics-informed machine learning. *Nat. Rev. Phys.*, 3(6):422–440, May 2021.

[20] John A Keith, Valentin Vassilev-Galindo, Bingqing Cheng, Stefan Chmiela, Michael Gastegger, Klaus-Robert Müller, and Alexandre Tkatchenko. Combining machine learning and computational chemistry for predictive insights into chemical systems. *Chem. Rev.*, 121(16):9816–9872, August 2021.

[21] James Kirkpatrick, Brendan McMorrow, David H P Turban, Alexander L Gaunt, James S Spencer, Alexander G D G Matthews, Annette Obika, Louis Thiry, Meire Fortunato, David Pfau, Lara Román Castellanos, Stig Petersen, Alexander W R Nelson, Pushmeet Kohli, Paula

Mori-Sánchez, Demis Hassabis, and Aron J Cohen. Pushing the frontiers of density functionals by solving the fractional electron problem. *Science*, 374(6573):1385–1389, December 2021.

[22] Felix Musil, Andrea Grisafi, Albert P Bartók, Christoph Ortner, Gábor Csányi, and Michele Ceriotti. Physics-Inspired structural representations for molecules and materials. *Chem. Rev.*, 121(16):9759–9815, August 2021.

[23] Bernhard Schölkopf, Francesco Locatello, Stefan Bauer, Nan Rosemary Ke, Nal Kalchbrenner, Anirudh Goyal, and Yoshua Bengio. Toward causal representation learning. *Proc. IEEE*, 109(5):612–634, May 2021.

[24] Chiyuan Zhang, Samy Bengio, Moritz Hardt, Benjamin Recht, and Oriol Vinyals. Understanding deep learning (still) requires rethinking generalization. *Commun. ACM*, 64(3):107–115, February 2021.

[25] Wiktor Beker, Rafał Roszak, Agnieszka Wołos, Nicholas H Angello, Vandana Rathore, Martin D Burke, and Bartosz A Grzybowski. Machine learning may sometimes simply capture literature popularity trends: A case study of heterocyclic Suzuki–Miyaura coupling. *J. Am. Chem. Soc.*, 144(11):4819–4827, March 2022.

[26] Viviana Consonni and Roberto Todeschini. *Molecular Descriptors for Chemoinformatics: Volume I: Alphabetical Listing / Volume II: Appendices, References.* John Wiley & Sons, October 2009.

[27] Yao Ma and Jiliang Tang. *Deep learning on graphs.* Cambridge University Press, 2021.

Data-Driven Approaches for Surface Materials and Beyond

Daniel Packwood

Institute for Integrated Cell-Material Sciences (iCeMS), Kyoto University, Kyoto 606-8501, Japan

Electrical circuits and other device components have undergone remarkable diminution over the last several decades. It is astonishing to think that a computer from 20 years ago could be embedded into the circuitry of a modern smartphone several times over. Yet this diminution cannot continue indefinitely, as there are physical limits to how small transistors, capacitors, conducting channels and other device components can be.

Pushing this diminution to its extreme requires that device components be built from individual molecules. In materials chemistry, the idea of constructing devices by sticking molecules together like Lego blocks is referred to as bottom-up device fabrication. Bottom-up device fabrication cannot be carried out with existing industrial or laboratory processes, because modern equipment cannot manipulate individual molecules with precision. However, nature provides us with a potential indirect route which does not require any special equipment at all: on-surface molecular self-assembly. On-surface molecular self-assembly is a natural phenomenon in which molecules adsorbed on a surface spontaneously assemble into an ordered structure. On-surface molecular self-assembly has been reported a vast number of times and has been the subject of intense research (see [1] and [2] for recent reviews). In terms of bottom-up device fabrication, molecular self-assembly sounds absurdly simple. Providing that the molecules and surface have been selected properly, all one needs to do is deposit the molecules onto the surface and leave it to nature to build the desired structures. However, if molecular self-assembly is indeed so simple, where then are our self-assembled electronics? Behind this question lies two serious problems.

The first problem arises from the spontaneous nature of molecular self-assembly. The only control that a scientist has over the molecular self-assembly process is in the choice of molecules and certain parameters such as surface type or temperature. As soon as these choices are made and the molecules have been deposited onto the surface, the self-assembly process begins and cannot be affected thereafter. As a methodology for creating device components, molecular self-assembly therefore has a trial-and-error quality, in which different candidate molecules must be tested in the hopes of finding ones which assemble as desired. This problem is exacerbated by our inability to predict the outcome of the self-assembly process for a given molecule, which prevents us from narrowing the list of candidates in order to expedite the search.

The second problem arises from the physical requirements for a self-assembled structure to serve as a component for a device. As a component of an electrical device, the self-assembled structure should possess physical properties such as electrical conductivity or magnetic order. These physical properties are collective in the sense that they are the result of the collective interactions between molecules within the structure, and are not present in any of the molecules when they are isolated. Yet, the vast majority of self-assembled structures reported so far are nothing more than aggregates of molecules and lack any collective properties at all.

Unfortunately, guidelines for choosing molecules which assemble into structures showing novel collective properties are lacking.

This talk summarises recent efforts by collaborators and myself to address these problems. For the first problem, we have been working on computational methods for predicting the outcome of the molecular self-assembly process for a given type of molecule. These methods use simple models with atomistic detail embedded into their parameters *via* machine learning. Novel methods combining annealing, genetic algorithms, and Markov chain Monte Carlo are then use to obtain predictions from these models. For the second problem, we have been exploring the use of metal
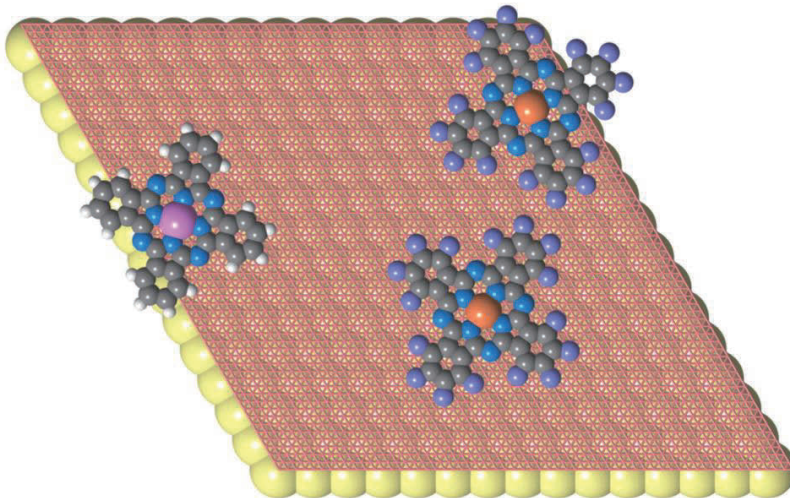
Figure 1. Illustation of the model. One MnPc and two FeFPc molecules are adsorbed on a gold surface. A segment of the grid is shown by the red-orange mesh. Yellow spheres = gold atoms, grey spheres = carbon atoms, blue spheres = nitrogen atoms, white spheres = hydrogen atoms, teal spheres = fluorine atoms, orange spheres = iron atoms, purple sphere = manganese atom.

ion-containing organic molecules for building self-assembled structures with magnetic ordering. The following is a summary of our results so far, which will appear later as a full publication.

1. Model construction with machine learning

Our model for a system of surface-adsorbed molecules is summarized in Figure 1. In brief, it considers an infinite grid on top of which a finite number of molecules reside. n_1 of these molecules are of one type, and the remaining n_2 are of another type. At each grid point the molecules may take on one of a finite number of orientations. The molecules are treated as rigid objects. The energy of the system is defined as

$$\epsilon = \sum_{k=1}^{n_1+n_2} u_k + \sum_{i \neq j} v_{ij}$$

(1)

the second sum runs over all unique pairs of molecules. In equation (1), u_k is interpreted as the energy arising from the interaction between molecule k and the surface, and v_{ij} is interpreted as the energy arising from the interaction between molecules i and j. The formation probability is defined as

$$p = C \exp\left(-\epsilon/k_B T\right),$$ (2)

where $C > 0$ is a constant, k_B is the Boltzmann constant, and $T > 0$ is the temperature. Predictions are made from the model by sampling configurations (that is, assignments of grid points and orientations to each molecule) according to the probability distribution p.

In the present study, the interaction parameters u_k and v_{ij} are set using a procedure which combines density functional theory (DFT) with machine learning. DFT is a technique from computational physics which computes energies and other physical properties for molecules according to the equations of quantum mechanics. Using DFT, the parameters u_k can be computed from an atomistic model of the surface and an appropriately placed molecule. Due to the translation symmetry of the surface, the range of potential values for u_k is finite. DFT calculations can therefore be performed for every potential value of u_k and the results tabulated for future reference. The parameters v_{ij} can also be computed according to DFT using atomistic models of two interacting molecules. However, compared to u_k, the range of potential values for v_{ij} is immensely large, and it is difficult to compute every possibility and tabulate the results. We therefore adopt a machine learning procedure, in which a sample of around 9000 potential values for v_{ij} are computed from DFT and the results used to train a Gaussian process regression (GPR) model. Using this GPR model, the other potential values for v_{ij} can be computed at negligible computation cost. The combination of pre-tabulated values for u_k and the GPR model for assigning values of v_{ij} therefore completes the specification of the model in (1) for the specific molecules and surface under study. Similar techniques were used in reference [3], as well as in work by other research groups [4, 5].

## 2. Model predictions using genetic algorithms and Markov chain Monte Carlo

Markov chain Monte Carlo (MCMC) is a standard method for sampling from probability distributions such as the one in equation (2). In essence, MCMC involves simulating a random walk (a Markov chain) on the configuration space of the model in such a way that, over a sufficiently long computation time, the fraction of visits to a specific configuration converges to the probability p in equation (2). While the convergence properties of MCMC are well established, in practice this convergence may require impractical computational times to achieve, especially when the configuration space is infinitely large and high dimensional. Efficient computational performance therefore requires that an initial configuration which is close to the mode of the distribution.

We developed a genetic algorithm which can generate a suitable initial configuration for MCMC for the model described in the previous section. In this genetic algorithm, the molecules in the model are partitioned into subsets. Within each these subset, each molecule is undergoing a non-negligible interaction with at least one other molecule from the subset, and undergoes negligible interactions with all molecules belonging to other subsets. The genetic algorithm starts by generating a population of so-called chromosomes, which are sets of these subsets, and proceeds by mutating the subsets and exchanging them between chromosomes in a stochastic manner. Coupled with a bias towards energy minimization, our genetic algorithm arrives at a configuration which is close to the mode of the distribution in (2).

3. Model predictions

Figure 2 shows the result of our MCMC sampling procedure, as carried out using four iron fluoro-phthalocyanine (FeFPc) molecules, four manganese phthalocyane (MnPc) molecules, a gold(111) surface, and a temperature T of 300 K are shown in Figure 2. Here, the t-distributed stochastic neighborhood embedding (t-SNE) method was used to visualise the MCMC sample. Each point corresponds to one configuration, and points which are close together in the plot are similar according to a similarity metric. It can be seen that the MCMC

sample concentrates mainly into two clusters, which contain configurations in which the molecules are compactly and regularly assembled. These assemblies strongly resemble the ones observed by high-resolution scanning tunneling microscopy studies reported by other authors [6]. In one of these clusters the molecules are condensed into square-shaped assemblies, and in the other cluster the molecules are condensed into rectangular assemblies. On the right-hand side of the t-SNE plot a small group of points can be seen, which correspond loosely packed configurations which appear to be intermediates to the ones described above. These results suggest that at 300 K, there exists two compact assemblies in equilibrium, with transitions between assemblies involving the intermediate configurations (refer to the green arrows in Figure 2).
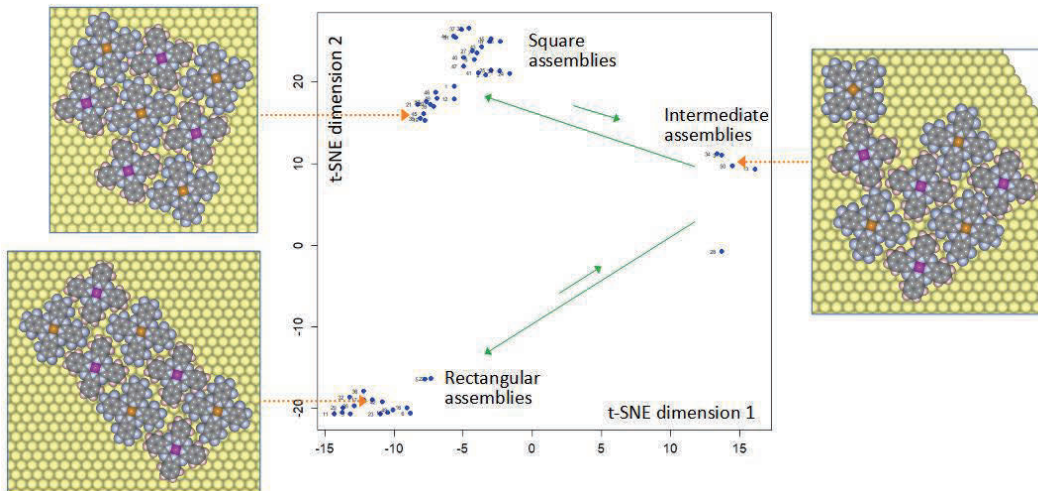
Figure 2. Summary of the results of MCMC sampling from the probability distribution in equation (2) for a gold surface, four MnPc and FePc molecules, and temperature 300 K. Each point corresponds to one configuration from the sample. Three representative configurations are shown in the surrounding inserts.

The metal ions (Fe and Mn) of the FeFPc and MnPc molecules possess unpaired electrons, which can interact *via* the surface-mediated Ruderman-Kittel-Kasuya-Yoshida (RKKY) interaction. In agreement with the experimental results of [6], our predicted assemblies show an antiferromagnetic ordering, in which the electrons of neighboring molecules point in opposite directions. This antiferromagnetic ordering is 'collective' in the sense that it arises from the interaction between molecules, and means that these assemblies could be used as components for ultra-high density memory devices or spintronics.

With the computational method presented above, we have made significant headway towards solving problem (1) described in the introduction. However, at present this method demands huge overheads which must be overcome in order to assist the development of self-assembled materials. These overheads arise from the need to perform many DFT calculations before setting the model parameters in equation (1). An important next step is therefore to develop highly general machine-learned models on the basis of rich datasets which can make predictions for a wide range of molecules and surfaces. These models would not need to be re-trained whenever a new molecule or surface is under consideration, and would permit us to immediately proceed to the prediction phase of the calculation using MCMC sampling. In addition to problem (1), we have made headway towards solving problem (2) by demonstrating how molecular assemblies with RKKY-interaction-induced magnetic ordering can be predicted. As a next step, we need to use this computational method to explore how magnetic order depends upon molecule structure, and hence deduce strategies for realizing molecular assemblies with desired magnetic orderings for device fabrication.

References

[1] Wakayama, Y. *Jpn. J. Appl. Phys.* **55**, 2016, 1102AA

[2] Ariga, K. *et al*. *Sci. Technol. Adv. Mater.* **20**, 2019, 52.

[3] Packwood, D. M., Han, P., Hitosugi, T. *Nat. Commun.* **8**, 2017, 14463.

[4] Hormann, L. *et al*. *Nano Lett.* **17**, 2017, 4453.

[6] Girovsky, J. *et al. Nat. Commun.* **8**, 2017, 15388.

Data-Driven Analysis of Dynamical Systems: From an Operator-Theoretic Perspective

Yoshinobu KAWAHARA

Institute of Mathematics for Industry, Kyushu University, and
Center for Advanced Intelligence Project, RIKEN

Analysis and prediction of dynamic processes using data are fundamental in a variety of scientific and industrial problems, and have been discussed actively in several fields including machine learning. Recently, operator-theoretic analysis of dynamical systems, in particular analysis with Koopman operator [16, 2, 19], has attracted much attention due to the generality, the connection to physical concepts [24, 18, 25], and advances of estimation methods such as dynamic mode decomposition (DMD), which was first proposed in the field of fluids [21, 22, 3].

In this talk, I first overview the basics of this research topics, focusing on spectral analysis of dynamical systems with Koopman operator. Then, I describe several recently-proposed related algorithms based on machine learning [15, 29, 30, 28, 27, 17, 5, 26, 10, 33, 14, 34], the ones for prediction based on those [4, 12, 32, 1, 6, 31], and the application to control and reinforcement learning [13, 20, 35]. In the talk, I occasionally show several applications of these method to several real-world data, with a focus on the ones we have worked on [7, 9, 8, 23, 11].

REFERENCES

[1] T. Bito, M. Hiraoka, and Y. Kawahara. Learning with coherence patterns in multivariate time-series data via dynamic mode decomposition. In *Proc. of the 2019 Int'l Joint Conf. on Neural Networks (IJCNN'19)*, number 19278, 2019.

[2] M. Budišić, R. Mohr, and I. Mezić. Applied koopmanism. *Chaos*, 22:047510, 2012.

[3] K.K. Chen, J.H. Tu, and C.W. Rowley. Variants of dynamic mode decomposition: Boundary condition, koopman, and fourier analyses. *Journal of Nonlinear Science*, 22:887–915, 2012.

[4] K. Fujii, Y. Inaba, and Y. Kawahara. Koopman spectral kernels for comparing complex dynamics: Application to multiagent sport plays. In *Proc. of the 2017 European Conf. on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD'17)*, pages 127–139, 2017.

[5] K. Fujii and Y. Kawahara. Dynamic mode decomposition with vector-valued reproducing kernels for extracting dynamical structures among observables. *Neural Networks*, 117:94–103, 2019.

[6] K. Fujii and Y. Kawahara. Supervised dynamic mode decomposition via multitask learning. *Pattern Recognition Letters*, 122:7–13, 2019.

[7] K. Fujii, T. Kawasaki, Y. Inaba, and Y. Kawahara. Prediction and classification in equation-free collective motion dynamics. *PLOS Computational Biology*, 14(11):e1006545, 2018.

[8] K. Fujii, N. Takeishi, M. Hojo, Y. Inaba, and Y. Kawahara. Physically-interpretable classification of biological network dynamics for complex collective motions. *Scientific Reports*, 10:3005, 2020.

[9] K. Fujii, N. Takeishi, B. Kibushi, M. Kouzaki, and Y. Kawahara. Data-driven spectral analysis for coordinative structures in periodic human locomotion. *Scientific Reports*, 9:16755, 2019.

[10] Y. Hashimoto, I. Ishikawa, M. Ikeda, Y. Matsuo, and Y. Kawahara. Krylov subspace method for nonlinear dynamical systems with random noise. *Journal of Machine Learning Research*, 21(172):1–29, 2020.

[11] S. Ikeda, K. Kawano, S. Watanabe, O. Yamashita, and Y. Kawahara. Predicting behavior through dynamic modes in resting-state fMRI data. *NeuroImage*, 247:118801, 2022.

[12] I. Ishikawa, K. Fujii, M. Ikeda, Y. Hashimoto, and Y. Kawahara. Metric on nonlinear dynamical systems with Perron-Frobenius operators. In *Advances in Neural Information Processing Systems 31*, pages 2856–2866. 2018.

[13] T. Iwata and Y. Kawahara. Controlling nonlinear dynamical systems with linear quadratic regulator-based policy networks in Koopman space. In *Proc. of the 60th IEEE Conference on Decision and Control (CDC'21)*, pages 5086–5091.

[14] T. Iwata and Y. Kawahara. Neural dynamic mode decomposition for end-to-end modeling of nonlinear dynamics. *arXiv:2012.06191*, 2020.

[15] Y. Kawahara. Dynamic mode decomposition with reproducing kernels for Koopman spectral analysis. In *Advances in Neural Information Processing Systems 29*, pages 911–919. 2016.

[16] B. Koopman. Hamiltonian systems and transformation in Hilbert space. *Proceedings of the National Academy of Sciences of the United States of America*, 17(5):315–318, 1931.

[17] B. Lusch, J.N. Kutz, and S.L. Brunton. Deep learning for universal linear embeddings of nonlinear dynamics. *Nature Communications*, 9:4950, 2018.

[18] A. Mauroy and I. Mezić. Global computation of phase-amplitude reduction for limit-cycle dynamics. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 28:073108, 2018.

[19] I. Mezić. Analysis of fluid flows via spectral properties of the koopman operator. *Annual Review of Fluid Mechanics*, 45:357–378, 2013.

[20] M. Ohnishi, I. Ishikawa, K. Lowrey, M. Ikeda, S. Kakade, and Y. Kawahara. Koopman spectrum nonlinear regulator and provably efficient online learning. *arXiv:2106.15775*, 2021.

[21] C.W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D.S. Henningson. Spectral analysis of nonlinear flows. *J. of Fluid Mechanics*, 641:115–127, 2009.

[22] P.J. Schmid. Dynamic mode decomposition of numerical and experimental data. *J. Fluid Mechanics*, 656:5–28, 2010.

[23] H. Shiraishi, Y. Kawahara, R. Fukuma, O. Yamashita, S. Yamamoto, Y. Saitoh, H. Kishima, and T. Yanagisawa. Neural decoding of ECoG signals using dynamic mode decomposition. *Journal of Neural Engineering*, 17(3):036009, 2020.

[24] S. Shirasaka, W. Kurebayashi, and H. Nakao. Phase-amplitude reduction of transient dynamics far from attractors for limit-cycling systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 27:023119, 2017.

[25] K. Taga, Y. Kato, Y. Kawahara, Y. Yamazaki, and H. Nakao. Koopman spectral analysis of elementary cellular automata. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 31:103121, 2021.

[26] N. Takeishi and Y. Kawahara. Learning multiple nonlinear dynamical systems with side information. In *Proc. of the 59th IEEE Conf. on Decision and Control (CDC'20)*, pages 3206–3211, 2020.

[27] N. Takeishi, Y. Kawahara, Y. Tabei, and T. Yairi. Bayesian dynamic mode decomposition. In *Proc. of the 26th Int'l Joint Conf. on Artificial Intelligence (IJCAI'17)*, pages 2814–2821, 2017.

[28] N. Takeishi, Y. Kawahara, and T. Yairi. Sparse nonnegative dynamic mode decomposition. In *Proc. of the 2017 IEEE International Conference on Image Processing (ICIP'17)*, pages 2381–8549.

[29] N. Takeishi, Y. Kawahara, and T. Yairi. Learning koopman invariant subspaces for dynamic mode decomposition. In *Advances in Neural Information Processing Systems 30*, pages 1130–1140. 2017.

[30] N. Takeishi, Y. Kawahara, and T. Yairi. Subspace dynamic mode decomposition for stochastic Koopman analysis. *Physical Review E*, 96:033310, 2017.

[31] N. Takeishi, K. Takeuchi, K. Fujii, and Y. Kawahara. Discriminant dynamic mode decomposition for labeled spatio-temporal data collections. *SIAM Journal on Applied Dynamical Systems*, in press.

[32] N. Takeishi, T. Yairi, and Y. Kawahara. Factorially-switching dynamic mode decomposition for koopman analysis of time-variant systems. In *Proc. of the 57th IEEE Conf. on Decision and Control (CDC'18)*, pages 6402–5408.

[33] I. Ul Haq, K. Fujii, and Y. Kawahara. Dynamic mode decomposition via dictionary learning for foreground modeling in videos. *Computer Vision and Image Understanding*, 199:103022, 2020.

[34] I. Ul Haq, T. Iwata, and Y. Kawahara. Dynamic mode decomposition via convolutional autoencoders for dynamics modeling in videos. *Computer Vision and Image Understanding*, 216:103355, 2022.

[35] M. Weissenbacher, S. Sinha, A. Garg, and Y. Kawahara. Koopman Q-learning: Offline reinforcement learning via symmetries of dynamics. *arXiv:2111.01365*, 2021.

Data-driven approach for indexing-free X-ray crystallography

Xian CHEN

Mechanical and Aerospace Engineering, Hong Kong University of Science and Technology

We propose a novel data-driven approach for analyzing the synchrotron X-ray microdiffraction scans based on machine learning algorithms. The basic architecture and major components of the method are formulated mathematically. We demonstrate it through the typical examples including polycrystalline solids, multiphase phase transforming alloys and finely twinned martensite. The computational pipeline is implemented for the Beamline 12.3.2 at the Advanced Light Source, Lawrence Berkeley National Lab. The conventional analysis of X-ray diffraction is based on the crystallographic study and the pattern-by-pattern indexing processing. This work underlies a new way of X-ray diffraction analysis independent of the indexing results. It motivates further studies of X-ray diffraction patterns from the prospective of machine learning for suitable feature extraction, clustering and labeling algorithms.

Reconstruction of log-aesthetic curve parameters

Sebastián Elías GRAIFF ZURITA

Graduate School of Mathematics, Kyushu University, Japan

(joint work with Kenji Kajiwara)

Log-aesthetic curves (LAC) constitute a family of planar spirals including the logarithmic spiral, Nielsen's spiral, Cornu spiral, and the circle involute, among others. This family is defined to best represent the properties observed by T. Harada et al. in a quantitative study of aesthetically pleasing curves used in industrial design. Many works have been written since then, mainly focused on the construction of the LAC for given constraints. In this work, as a first step to approximate a planar curve segment by Log-aesthetic curves, we show how to identify the parameters that uniquely define a LAC segment.

1. Introduction

In 1994, T. Harada et al. set up an experiment to quantitatively analyze a curve's character from the viewpoint of the observer (see [3], in Japanese). Their main result may be described as follows: the curves that car designers regard as aesthetic have the common property that the frequency histogram of the radius of curvature follows a piecewise linear relation in a log-log scale. In 2006, K. T. Miura et al. in [5], and N. Yoshida and T. Saito in [6] provided an analytic formulation of the work by T. Harada, defining what will later be coined as *log-aesthetic curves (LAC)*. Several other works have been written regarding the implementation and construction of LAC for fixed boundary conditions, motivated for their application in design with computer-aided design (CAD) softwares, see for example [1, 2, 7]. In this work, we emphasize the reverse engineering side of the problem, by providing tools that help characterize a general curve segment by LAC.

2. Log-aesthetic curves

Let $\gamma(s) \in \mathbb{R}^2$ (s: arc length) be an arc length parameterized plane curve, the tangent and normal vectors are $T(s) = \gamma'(s)$ and $N(s) = \mathrm{R}_{\pi/2}\,T(s)$, where $\mathrm{R}_{\pi/2}$ is a $\pi/2$ rotation matrix and $' = \mathrm{d}/\mathrm{d}s$. Denote as $\theta(s)$ and $\kappa(s)$ to the angle and curvature function, respectively.

Given a curve $\gamma(s)$ parameterized by arc length $s \in [0,1]$, define $R(s) := \log \rho(s)$, where $\rho(s)$ is the radius of curvature. By taking the continuous limit of the logarithmic curvature histogram, it can be seen[5] that

$$(1) \qquad \log\left(\frac{\mathrm{d}s}{\mathrm{d}R}\right) = \alpha R - \log A,$$

holds for some $\alpha \in \mathbb{R}$ and $A > 0$. Assuming that functions are well behaved, we have $\frac{ds}{dR} = \rho/\rho'$, and the previous equation can be rewritten as

$$(2) \qquad \rho'\rho^{(\alpha-1)} = A,$$

or equivalently, by talking the derivative of (2), as

$$(3) \qquad \rho\rho'' + (\alpha - 1)(\rho')^2 = 0.$$

Equation (3), in terms of the curvature, is what we consider as the defining equation for log-aesthetic curves.

Definition 2.1 (Log-aesthetic curve). An arc length parameterized curve $\gamma(s)$ with strictly monotonic radius of curvature is called a *log-aesthetic curve (LAC)* if its curvature satisfies

$$(4) \qquad \kappa\kappa'' - (\alpha + 1)\left(\kappa'\right)^2 = 0$$

for some constant $\alpha \in \mathbb{R}$.

Note that, for a given LAC, its slope α is invariant under similarity transformations and reflections. Indeed, knowing that the curvature of a planar curve is invariant under rotations and translations, we only check invariance under scales and reflections. For the scale transformations, consider the arc length parameterized LAC $\gamma(s)$ satisfying (4), for some $\alpha \in \mathbb{R}$, and define $\tilde{\gamma}(\tilde{s}) := S\gamma(\tilde{s}/S)$, where $S > 0$. Then, we have $\tilde{\kappa}(\tilde{s}) = S^{-1}\kappa(\tilde{s}/S)$ which satisfies the equivalent differential equation (4) for $\tilde{\kappa}(\tilde{s})$. For the reflections, note that interchanging x and y in the curve is equivalent to changing the sign of the curvature, and (4) is invariant under that change.

Definition 2.2. Let $\xi^\alpha(s)$ be an LAC defined over an open interval $I \subset \mathbb{R}$, such that $\{0\} \in I$, and satisfying

$$(5) \qquad \begin{cases} \kappa'(s) = -(\kappa(s))^{(\alpha+1)} < 0, & \forall s \in I, \\ \kappa(0) = 1, \\ \theta(0) = 0, \\ \xi^\alpha(0) = 0. \end{cases}$$

We call $\xi^\alpha(s)$ a *basic LAC* with slope α.

Let us see a more explicit expression for the basic LAC and its related quantities. In what follows, we use the sub-index ξ^α, as for example κ_{ξ^α}, to denote those quantities related to their respective basic LAC. Taking the initial condition into consideration, the explicit equation for the curvature is

$$(6) \qquad \kappa_{\xi^\alpha}(s) = \begin{cases} \exp(-s), & \alpha = 0, \\ (1 + \alpha s)^{-1/\alpha}, & \alpha \neq 0. \end{cases}$$

Then, the turning angle is obtained from the curvature by $(\theta_{\xi^\alpha})' = \kappa_{\xi^\alpha}$,

(7)
$$\theta_{\xi^\alpha}(s) = \begin{cases} 1 - \exp(-s), & \alpha = 0, \\ \log(s+1), & \alpha = 1, \\ \frac{(1+\alpha s)^{\frac{\alpha-1}{\alpha}} - 1}{\alpha - 1}, & \alpha \neq 0, 1. \end{cases}$$

Finally, the position vector can be obtained from the tangent vector as

(8)
$$\xi^\alpha(s) = \int_0^s \begin{pmatrix} \cos\theta_{\xi^\alpha}(\bar{s}) \\ \sin\theta_{\xi^\alpha}(\bar{s}) \end{pmatrix} d\bar{s}.$$

Although it is not used in this work, we note that the position vector can be expressed in terms of the incomplete gamma function, see for example [7]. For simplicity we consider the case in which $1 + \alpha s > 0$. In this case, the maximal interval $I_\alpha \subset \mathbb{R}$ on which the basic LAC can be defined is

$$I_\alpha = \begin{cases} (-\infty, -1/\alpha), & \alpha < 0, \\ (-\infty, \infty), & \alpha = 0, \\ (-1/\alpha, \infty), & \alpha > 0. \end{cases}$$

In all cases, the image of κ_{ξ^α} is $\kappa_{\xi^\alpha}[I_\alpha] = (0, \infty)$.

Proposition 2.3. *Any LAC with $\alpha \neq 1$ and positive and decreasing curvature can be expressed as a basic LAC after applying similarity transformations. In particular, if $\gamma(s)$, $s \in [0, L]$, is an LAC of length L, there exist unique $\gamma_0 \in \mathbb{R}^2$, $\phi \in [0, 2\pi)$, $S \in \mathbb{R}\backslash\{0\}$, and $s_0 \in \mathbb{R}$, such that*

(9)
$$\gamma(s) = \gamma_0 + S\,\mathrm{R}_\phi\,\xi^\alpha(s/S + s_0), \qquad s \in [0, L],$$

where $\xi^\alpha(s)$ is a basic LAC of length $l := L/S$.

In consideration of the later proposition, we denote as ξ^p the LAC segment uniquely defined by the parameters

$$p = (\alpha, S, s_0, l, \phi, x_0, y_0),$$

where $\gamma_0 = (x_0, y_0)$. With this notation, note that $\xi^p = \xi^\alpha$ in the case that $p = (\alpha, 1, s_0, l, 0, 0, 0)$ for some s_0 and l satisfying that $(s_0, s_0 + l) \subset I_\alpha$.

Proof of Proposition 2.3. For a given LAC $\gamma(s)$, $s \in [0, L]$ with $\alpha \neq 1$ and positive and decreasing curvature, from Definition 2.1 its curvature satisfies (4), which can be integrated once to obtain

$$\kappa'(s) = -A(\kappa(s))^{(\alpha+1)},$$

for some $A > 0$. Next, we consider the curve $\bar{\gamma}(\bar{s}) := S^{-1}\gamma(\bar{s}S)$, $\bar{s} \in [0, L/S]$, and we set $S = A^{1/(\alpha-1)}$. We see that its curvature satisfies

$$\text{(10)} \qquad \begin{cases} \bar{\kappa}'(\bar{s}) = -(\bar{\kappa}(\bar{s}))^{(\alpha+1)}, \\ \bar{\kappa}(0) = A^{1/(\alpha-1)}\kappa(0), \end{cases}$$

which can be integrated to obtain

$$\text{(11)} \qquad \bar{\kappa}(\bar{s}) = \begin{cases} \exp(-(\bar{s} - \log\bar{\kappa}(0))), & \alpha = 0, \\ \left(1 + \alpha\left(\bar{s} + \frac{(\bar{\kappa}(0))^{-\alpha}-1}{\alpha}\right)\right)^{-1/\alpha}, & \alpha \neq 0. \end{cases}$$

By comparing $\bar{\kappa}$ with κ_{ξ^α} (6) there exists a unique $s_0 \in \mathbb{R}$ such that $\bar{\kappa}(\bar{s}) = \kappa_{\xi^\alpha}(\bar{s}+s_0)$, hence the curves $\bar{\gamma}$ and ξ^α are congruent up to rigid transformations, i.e.

$$\bar{\gamma}(\bar{s}) = \bar{\gamma}_0 + \mathrm{R}_\phi\,\xi^\alpha(\bar{s}+s_0), \qquad \bar{s} \in [0, L/S],$$

for some $\bar{\gamma}_0 \in \mathbb{R}^2$ and $\phi \in [0, 2\pi)$. Finally, we use that $\gamma(s) = S\bar{\gamma}(s/S)$ to obtain (9) with $\gamma_0 := S\bar{\gamma}_0$. $\qquad \square$

Remark 2.4. Let us denote by X to the reflection $X : \mathbb{R}^2 \to \mathbb{R}^2 / (x, y) \mapsto (y, x)$. Then, if $\gamma(s)$, $s \in [0, L]$, is an LAC then also are

$$\text{(12)} \qquad \begin{cases} \gamma^{(1)}(s) := \gamma(L - s), \\ \gamma^{(2)}(s) := X\gamma(s), \\ \gamma^{(3)}(s) := X\gamma(L - s), \end{cases}$$

and their respective curvature satisfies that

$$\begin{cases} \kappa^{(1)}(s) = -\kappa(L - s), \\ \kappa^{(2)}(s) = -\kappa(s), \\ \kappa^{(3)}(s) = \kappa(L - s). \end{cases}$$

This allows as to use the previous proposition for those cases in which the curvature is not positive and decreasing, by applying one of the previous transformations.

2.1. **Recovering the parameters of an LAC segment.** We focus our attention to the problem of finding the parameters that uniquely identify a given LAC segment. We proceed in three steps, where we solve several linear equations in the least squares sense, with the objective of constructing an algorithm that can be applied to general curves.

Given an LAC segment $\gamma(s)$, $s \in [0, L]$, by possibly applying one of the transformations (12), we assume that its curvature is positive and strictly monotonic decreasing. From Proposition 2.3, there exists $p = (\alpha, S, s_0, l, \phi, x_0, y_0)$ such that $\gamma(s) = \xi^p(s)$, $s \in [0, L]$. We further define $s_1 := L/S + s_0$ and assume that $\alpha \neq 0, 1$. From (9),

$$\gamma(s) = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + S\,\mathrm{R}_\phi\,\xi^\alpha(s/S + s_0), \qquad s \in [0, L],$$

implies that

$$(13) \qquad \kappa(s) = S^{-1}\kappa_{\xi\alpha}(s/S + s_0).$$

Recall that $R = -\log\kappa$, hence $R(s) = \log S + R_{\xi\alpha}(s/S + s_0)$ and using (5) we have that $\log R'_{\xi\alpha} + \alpha R_{\xi\alpha} = 0$, thus we obtain

$$(14) \qquad \log R' + \alpha R = (\alpha - 1)\log S.$$

Step 1 — Let $c_1 := \alpha$ and $c_0 := (\alpha - 1)\log S$. Then, from (14), in the least squares sense we have

$$(c_0, c_1) = \arg\min_{(\bar{c}_0, \bar{c}_1)} \left\{ \frac{1}{2} \int_0^L (\log R' + \bar{c}_1 R - \bar{c}_0)^2 \, \mathrm{d}s \right\},$$

which leads to

$$(15) \qquad c_1 = \frac{L \int_0^L R\log R' \, \mathrm{d}s - \int_0^L R \, \mathrm{d}s \int_0^L \log R' \, \mathrm{d}s}{(\int_0^L R \, \mathrm{d}s)^2 - L \int_0^L R^2 \, \mathrm{d}s},$$

and

$$(16) \qquad c_0 = \frac{1}{L} \int_0^L (\log R' - c_1 R) \, \mathrm{d}s.$$

Then, $\alpha = c_1$ and $S = \exp(c_0/(c_1 - 1))$.

Step 2 — From (6) and (13), we have that $\kappa(s) = S^{-1}(1 + \alpha(s/S + s_0))^{-1/\alpha}$, which allows us to isolate the parameter s_0 as

$$s_0 = \frac{(S\kappa(s))^{-\alpha}}{\alpha} - \frac{1}{\alpha} - \frac{s}{S}.$$

Then,

$$s_0 := \arg\min_{\bar{s}_0} \left\{ \frac{1}{2} \int_0^L \left(\frac{1}{\alpha S^\alpha \kappa(s)^\alpha} - \frac{1}{\alpha} - \frac{s}{S} - \bar{s}_0 \right)^2 \mathrm{d}s \right\}$$

gives

$$(17) \qquad s_0 = \frac{1}{\alpha L S^\alpha} \int_0^L (\kappa(s))^{-\alpha} \, \mathrm{d}s - \frac{1}{\alpha} - \frac{1}{LS} \int_0^L s \, \mathrm{d}s.$$

Similarly, we compute $s_1 := L/S + s_0$ from $\kappa^{(3)}(s) = \kappa(L-s) = S^{-1}(1 + \alpha(s_1 - s/S))^{-1/\alpha}$. In the least squares sense, we obtain

$$(18) \qquad s_1 = \frac{1}{\alpha L S^\alpha} \int_0^L (\kappa(L - s))^{-\alpha} \, \mathrm{d}s - \frac{1}{\alpha} + \frac{1}{LS} \int_0^L s \, \mathrm{d}s,$$

equivalently,

$$(19) \qquad s_1 = s_0 + \frac{2}{LS} \int_0^L s \, \mathrm{d}s.$$

Then, $l = s_1 - s_0$.

Step 3 — At this point, we are left with finding the rotation and translation parameters. For the former, note that the angle function of γ and ξ^α differ only by a constant ϕ, as

$$\theta(s) = \phi + \theta_{\xi^\alpha}(s/S + s_0).$$

Thus, in the least squares sense we obtain

$$(20) \qquad \phi = \frac{1}{L} \int_0^L (\theta(s) - \theta_{\xi^\alpha}(s/S + s_0)) \, \mathrm{d}s.$$

Finally, for the translation (x_0, y_0) we solve (9)c in the least squares sense,

$$(21) \qquad \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \frac{1}{L} \int_0^L (\gamma(s) - S \, \mathrm{R}_\phi \, \xi^\alpha(s/S + s_0)) \, \mathrm{d}s.$$

3. Follow-up work: Approximation by LAC

If time permits, we will show how to incorporate this algorithm as the initial guess of an approximation process, in which a planar curve is given as an input and we seek to find a Log-aesthetic curve that is the closest in a L^2-distance sense.

References

[1] R. U. Gobithaasan and K. T. Miura, Aesthetic spiral for design, Sains Malaysiana **40** (2011) 1301–1305.

[2] R. U. Gobithaasan, Y. Siew Wei, K. T. Miura, Log-aesthetic curves for shape completion problem, J. Appl. Math. **2014** (2014) 960302.

[3] T. Harada, N. Mori, K. Sugiyama, Study of quantitative analysis of curve's character, Bull. JSSD **40** (1994) 9–16. (in Japanese)

[4] J. Inoguchi, Y. Jikumaru, K. Kajiwara, K. T. Miura and W. K. Schief, Log-aesthetic curves: similarity geometry, integrable discretization and variational principles, preprint, arXiv:1808.03104v2 (2021)

[5] K. T. Miura, J. Sone, A. Yamashita, T. Kaneko, Derivation of a general formula of aesthetic curves, Proceedings of the Eighth International Conference on Humans and Computers (HC2005), 166–171, 2005

[6] N. Yoshida and T. Saito, Interactive aesthetic curve segments, Visual Comput **22** (2006) 896—905.

[7] R. Ziatdinov, N. Yoshida, T. Kim, Analytic parametric equations of log-aesthetic curves in terms of incomplete gamma functions. Comput. Aided Geom. Des. **29** 2 (2012) 129–140.

Solutions to a new strongly coupled phase-field model for nematic liquid crystals with variable degree of orientation

Junzhi ZHENG

Department of Mathematics,
Shanghai University, China

(joint work with Pierluigi Cesana, IMI, Kyushu University)

1. Introduction

The term nematic liquid crystal (LC) indicates an intermediate state of matter which exhibits liquid-like and crystalline properties at the same time [12], [28], [10], [20], [13], [3], [14], [1]. Indeed, similarly to fluids, nematic LC mesophases have no positional orders and are not capable of supporting shear. These mesophases have however orientational order as well as some crystalline properties such as anysotropic response to electrical, optical and magnetic stimuli. In a typical nematic mesophase molecules are, locally, aligned along a common direction identified by a unit vector named *director* and denoted with $n \in \mathbb{S}^2$. In most situations, nematic liquid crystal molecules exhibit order states [12] which can be described in terms of an order tensor, which is a 3×3, symmetric matrix.

1.1. **LC Models.** We recall the three main theories on liquid crystals available in the literature which are characterized through a symmetric 3×3 matrix in the sets $\mathcal{Q}_{Fr}$ (Frank tensor model [18]), $\mathcal{Q}_U$ (Ericksen uniaxial tensor model [15]), $\mathcal{Q}_B$ (de Gennes biaxial tensor model [12]), respectively. We have

$$\mathcal{Q}_{Fr} \subset \mathcal{Q}_U \subset \mathcal{Q}_B$$

with $\mathcal{Q}_B$ being a compact and convex set, while $\mathcal{Q}_U$ and $\mathcal{Q}_{Fr}$ are non-convex and compact sets. Observe $0 \in \mathcal{Q}_U$ (and consequently, also in $\mathcal{Q}_U$) and the zero matrix denotes the state of perfect optical isotropy. Frank set $\mathcal{Q}_{Fr}$ is that of tensors with eigenvalues identically equal to $\{2/3, -1/3, -1/3\}$. Frank model describes the local directions of the molecules through the eigenvectors of the tensor Q. This is suitable to describe perfectly ordered molecules, that is, systems where molecules are perfectly aligned along the director $n \in \mathbb{S}^2$. On the other hand, Ericksen and de Gennes models can be interpreted as generalization of Frank theory, in the sense that they allow the description of disordered systems, that is when the eigenvalue of Q are allowed to deviate from $\{2/3, -1/3, -1/3\}$ (perfect order) and possibly approach $\{0, 0, 0\}$ (isotropy). The connection of Frank model to the one of de Gennes is that the director n represents one distinguished eigenvector of the order tensor Q.

In what follows, we adopt Ericksen's parameterization of Q [15], that is,

$$(1) \qquad Q = s\left(n \otimes n - \frac{1}{3}I\right)$$

where $s \in [-1/2, 1]$ is the degree of order. Observe Frank's tensor can be recovered by assuming $s \equiv 1$ in (1).

1.2. **The model.** In this proceeding we present a novel strongly coupled model for the evolution of the order-director physics in a system of of uniaxially nematic liquid crystals with variable degree of orientation. We consider a 2-dimensional setting and parametrize the director $n = (\cos\phi, \sin\phi)$ with $\phi \in [0, \pi]$, leaving the full 3D version of the model as well as full analysis for a forthcoming paper [31].

The system of partial differential equations for the Ericksen model read [15], [9], [29]

$$(2) \qquad \begin{aligned} s_t &= \left(k_1\Delta s - k_2 s|\nabla\phi|^2 - W'(s)\right), \\ s^2\phi_t &= k_3\text{div}(s^2\nabla\phi), \end{aligned}$$

In a nutshell, model (2) establishes a complicate interplay between order and director. Indeed, n can even be discontinuous with little energy involved provided s vanishes when ∇n jumps and provided $s^2|\nabla\phi|^2$ is controlled. When s is bounded away from zero and ϕ is smooth, essentially (2) shows reaction-diffusion type behavior, with the non-convex function W acting as Landau de Gennes potential condensation potential [12].

Based on the recent development of Alber and Zhu, [2] [27], we introduce a modified system of equations characterized by a strongly no-linear driving force for the variable s. To keep our notation concise, we write the system of time-space evolution equations

$$(3) \qquad s_t = \left(k_1\Delta s - k_2 s|\nabla\phi|^2 - W'(s)\right)\alpha(x,t),$$
$$(4) \qquad s^2\phi_t = k_3\text{div}(s^2\nabla\phi),$$

which must be satisfied in $Q_{t_e} = \Omega \times (0, t_e)$, with the initial value condition

$$(5) \qquad s(x,0) = s_0(x), \qquad (x,t) \in \Omega,$$
$$(6) \qquad \phi(x,0) = \phi_0(x), \qquad (x,t) \in \Omega,$$

and periodic boundary conditions on s, ϕ on $\partial\Omega$. We implement our new version of the model by setting $\alpha(x,t) = |\nabla s|$ while we recover Ericksen model by setting $\alpha(x,t) \equiv 1$ in (3).

1.2.1. *Phase-field model.* To overcome singularities and degeneracy for vanishing s in (3) and (4), we introduce the modified system of equations

$$(7) \qquad s_t = \left(k_1\Delta s - k_2 s|\nabla\phi|^2 - W'(s)\right)\alpha(x,t),$$
$$(8) \qquad (s^2 + \varepsilon^2)\phi_t = k_3\text{div}((s^2 + \varepsilon^2)\nabla\phi)$$

$$(9) \qquad s(x,0) = s_0(x), \qquad (x,t) \in \Omega,$$
$$(10) \qquad \phi(x,0) = \phi_0(x), \qquad (x,t) \in \Omega,$$

for $\varepsilon > 0$ and periodic boundary conditions for both variables on $\partial \Omega$.
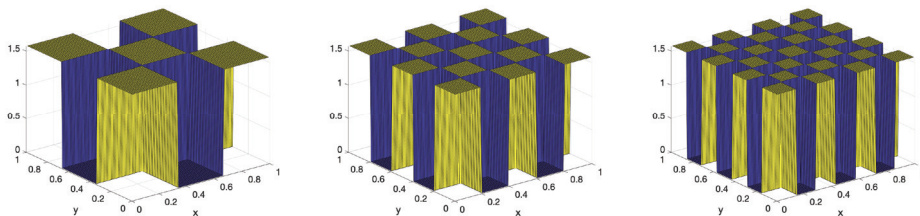
2. Numerical computations

FIGURE 1. Profiles of the initial condition $\phi_o(x)$ for three sets $(1), (2)$ and (3) of numerical computations.
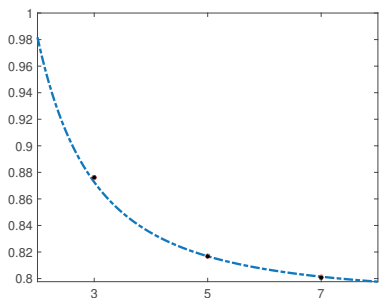

FIGURE 2. dashed line: mean value of the angle $\phi_o(x)$ over Ω for our three different initial simulations (label: 3 stands for the 9-square chessboard, label: 5 stands for the 25-square chessboard,) label: 7 stands for the 49-square chessboard). Circles (stars) correspond to the average of $\phi(x)$ over Ω at the final time step of each of the three simulations for the Ericksen (novel) model, respectively.
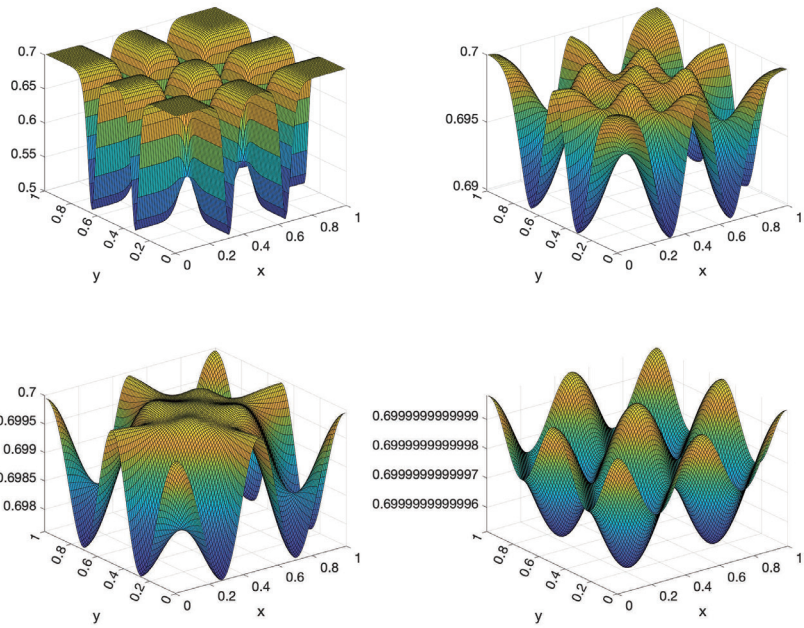
FIGURE 3. Evolution of profiles for the oder parameter s for the Ericksen model, Simulation (1).
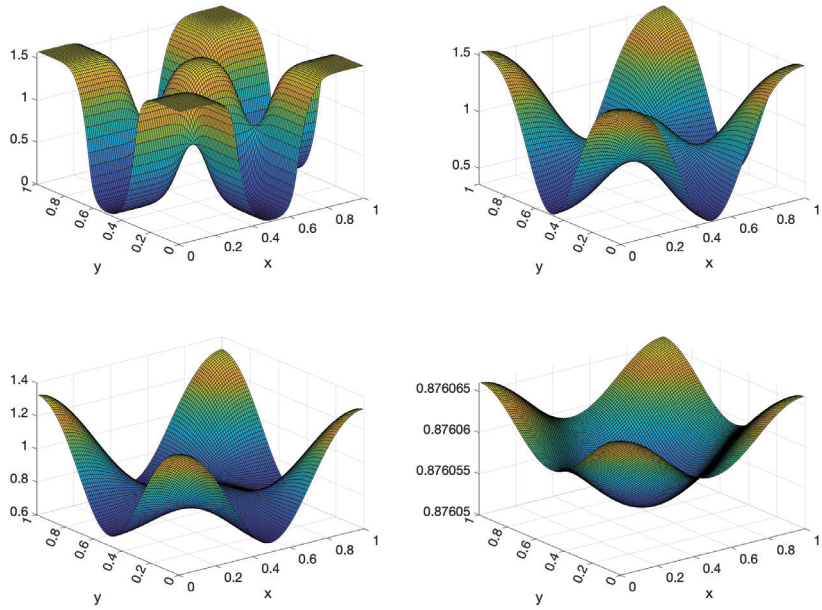
FIGURE 4. Evolution of profiles for the angle ϕ, Ericksen model, Simulation (1).
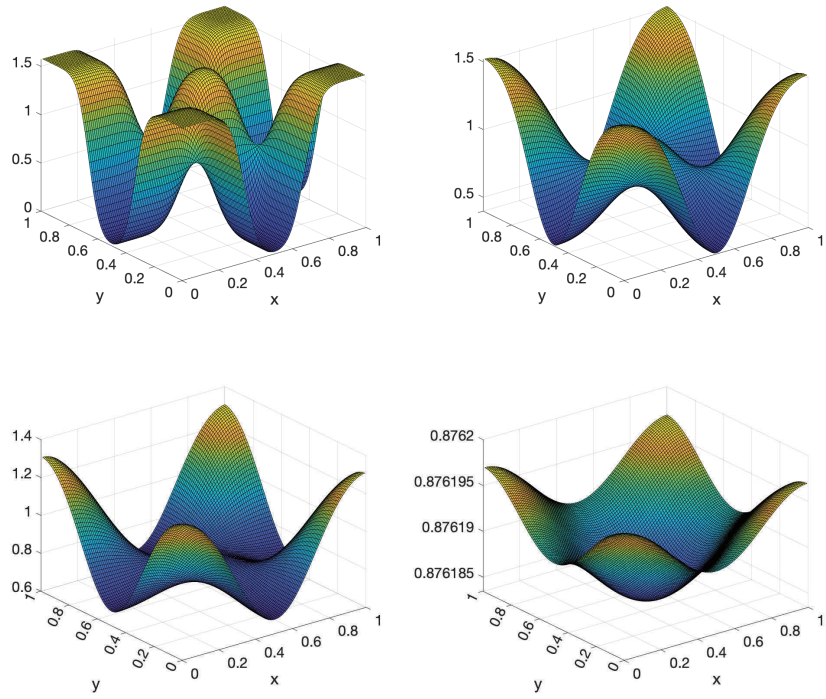
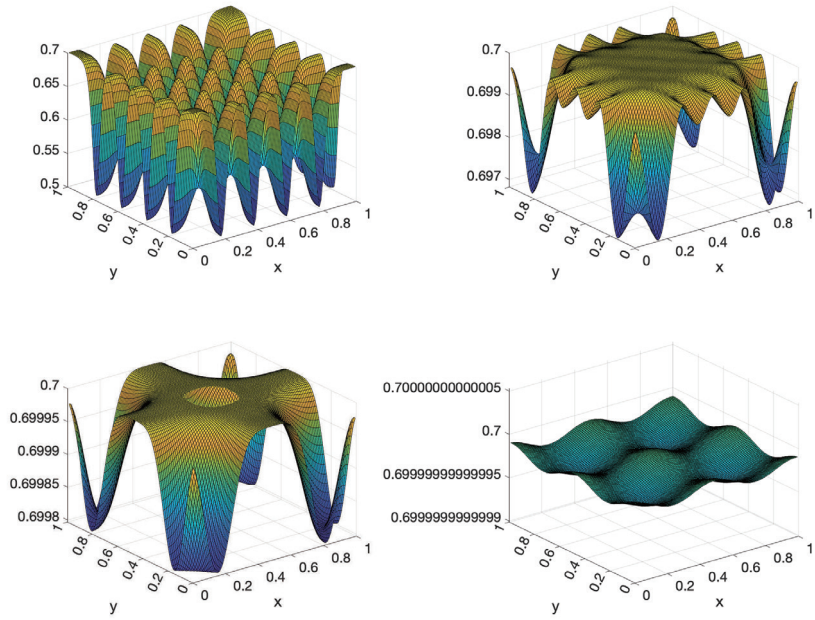FIGURE 5. Evolution of profiles for the angle ϕ, novel model, Simulation (1).

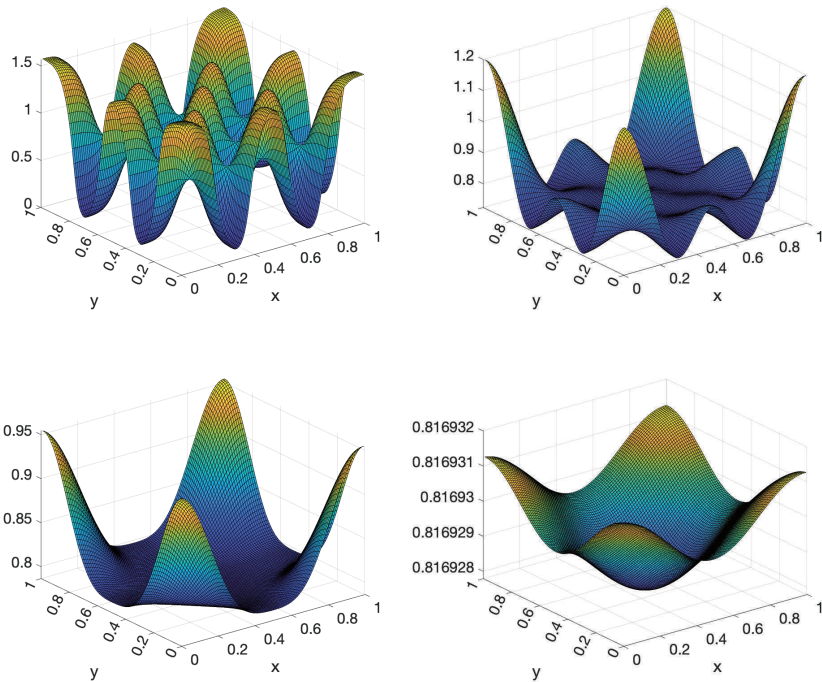FIGURE 6. Evolution of profiles for the oder parameter s for the Ericksen model, Simulation (2).

FIGURE 7. Evolution of profiles for the angle ϕ, Ericksen model, Simulation (2).

3. NUMERICAL SIMULATIONS

We solve ε-regularized versions of equations (3)-(4) both for the classical model of Ericksen as well as for our ε-regularized novel version of the nematic liquid crystal model, for three different initial conditions. The Ericksen model is obtained by setting $\alpha \equiv 1$ in (3) while our novel version corresponds to taking $\alpha \equiv |\nabla s|$.

To emphasize the differences and peculiarities of the classical Ericksen model and our novel version, we impose constant initial condition for the order parameter, that is, $\lim_{t \to 0^+} s(x, t) =: s(x, 0) \equiv s_o$, $\forall x \in \Omega$. Here s_o is the (unique) minimizer of the Landau de Gennes potential (see caption to Fig. 3.1) Consequently, we implement three realizations of numerical experiments for different shapes of the initial condition ϕ_0 (see, illustrated in Fig. 1). To drive the dynamics for the pair $(s(x, t), \phi(x, t))$ we prescribe discontinuous initial conditions $\phi_o(x) := \lim_{t \to 0^+} \phi(x, t)$. In this way, on one hand, we test stability of our implementation of the differential equations of the system in the presence of jumps of the solution resulting in singular gradients. On the
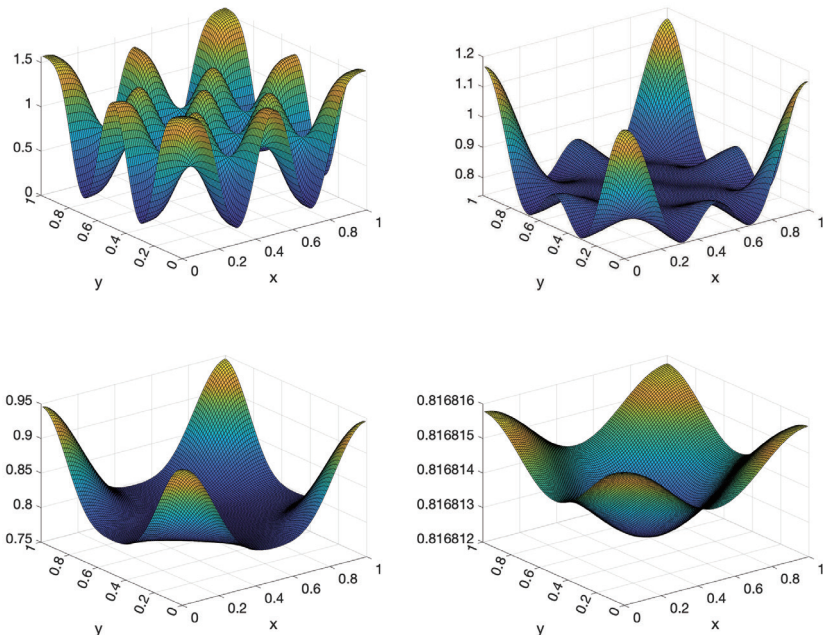
FIGURE 8. Evolution of profiles for the angle ϕ, novel model, bottom panel, Simulation (2).

other hand, presence of with steep gradients of ϕ is useful in testing the order-director coupling via strong, non-liner, coupling with the evolution equation for the variable $s(x,t)$.

3.1. **Numerical parameters.** We report all numerical parameters of the simulation in Table 1. In our simulations, implemented in Matlab environment, the domain of the problem is a unit square. The parameters at the bottom of Table 3.1 indicate the discretization of the domain and the time step.

3.2. **Discussion.** Figs. 3-11 display solutions for variables $s(x,t)$ and $\phi(x,t)$ for various time steps, for both LC models under consideration and for the three initial conditions.

3.2.1. *Order parameter.* As the initial condition for $s(x,t)$ is constant over Ω and equal to $s_o = 0.7$, we observe $s(x,t)$ remains constant, both in space and time, for the novel model (7), (8) obtained for $\alpha \equiv |\nabla s|$. Indeed, observe in Eq. (7) for $\alpha \equiv |s_x|$ one has, for $t = 0$, the value of s_t is equal to zero and consequently the dynamics for s is frozen. Consequently, the solutions obtained for this novel model coincide with the
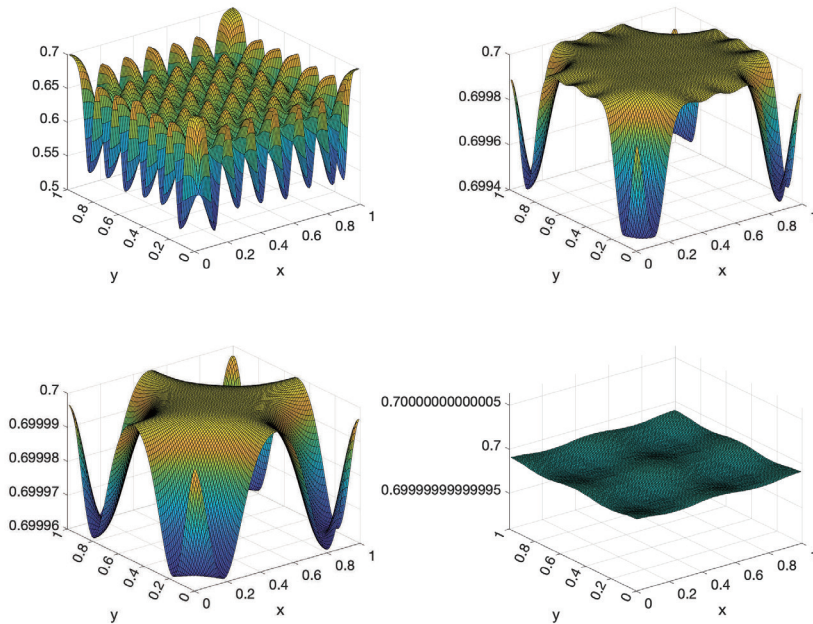
FIGURE 9. Evolution of profiles for the oder parameter s for the Ericksen model, Simulation (3).

solutions for classical Frank model obtained by setting $s(x,t) \equiv 1$ for every $t \geq 0$ and $x \in \overline{\Omega}$. Instead, in the classical Ericksen model, the evolution for $s(x,t)$ is driven by the non-negative right hand side of Eq. (7). Fig. 3 shows time-evolution of the order parameter $s(x,t)$ at four different time steps for the classical Ericksen model. We can see $s(x,t)$ tends to a constant value close to s_o for large time. Similar behavior is observed for the remaining two simulations, see Figs. 6, 9.

3.2.2. *Director.* Plots of the solution for the angle $\phi(x,t)$ show no significant qualitative difference for the two available models, see Figs. 4, 5, 7, 8, 10 and 11. Observe for large values of t, the solution $\phi(x,t)$ tends to the average value of $\phi_o(x,t)$ over Ω). We show this behavior in the plot of Fig. 2. The average value of ϕ_o in our three experiments depends on the number of squares of the chessboards.

REFERENCES

[1] D. Andrienko, *Introduction to liquid crystals*, Journal of Molecular Liquids 267 (2018) 520–541 (2018)
[2] Alber H D, Zhu P. Solutions to a model with nonuniformly parabolic terms for phase evolution driven by configurational forces. SIAM Journal on Applied Mathematics, 2005, 66(2): 680-699.
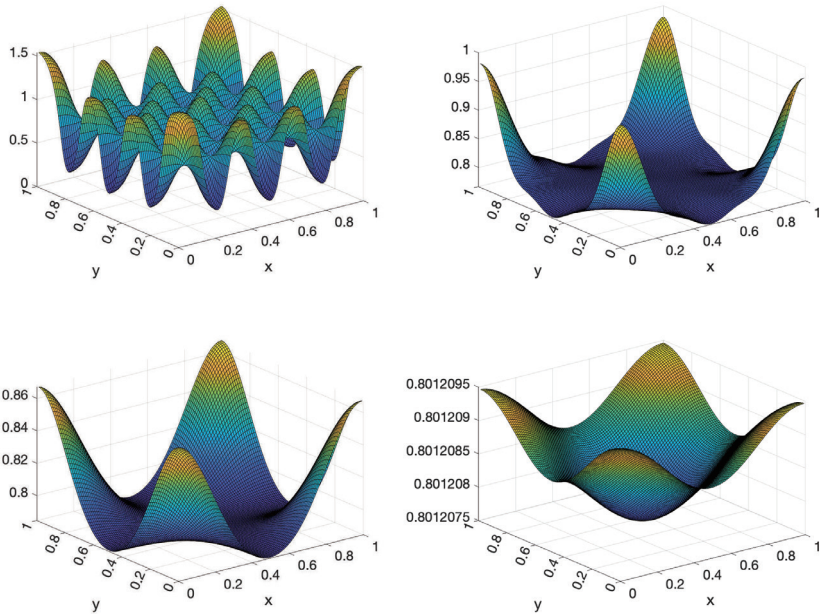
FIGURE 10. Evolution of profiles for the angle ϕ, Ericksen model, bottom panel, Simulation (3).

Parameter	Values
A	$0.21/2$
B	$-1/3$
C	$1/4$
ε	10^{-3}
k_1	10^{-4}
k_2	10^{-4}
k_3	10^{-4}
s_o	0.7
dx	0.01
dy	0.01
dt	0.02

TABLE 1. Numerical values of the simulations.

[3] M.P. Allen, D.J. Tildesley, Computer Simulation of Liquids, 2nd edition ed., Oxford University Press, Oxford, New York, 2017.
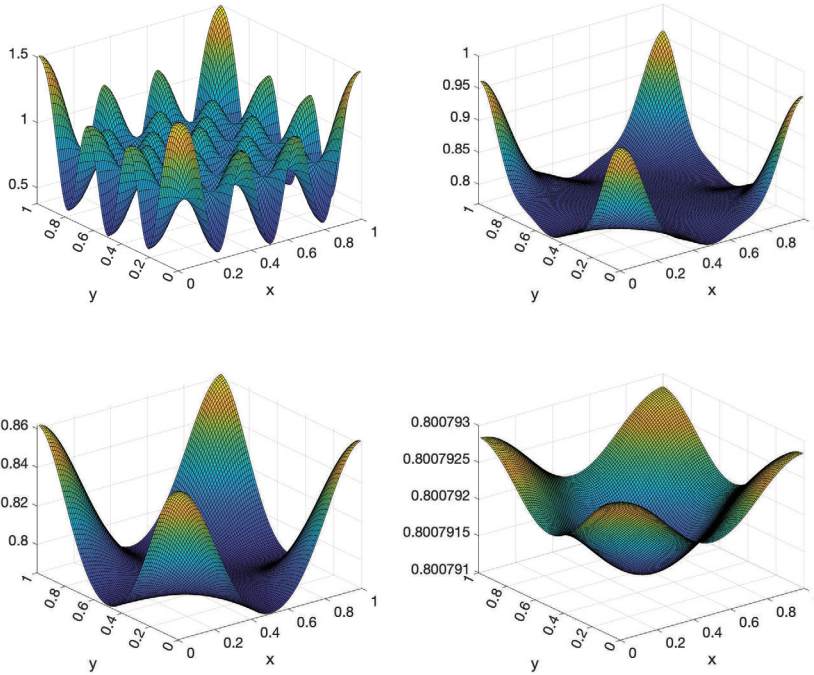
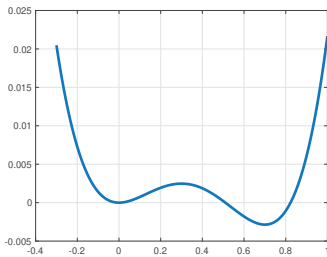FIGURE 11. Evolution of profiles for the angle ϕ, novel model, bottom panel, Simulation (3).

FIGURE 12. Profile of the Landau-de Gennes potential, $W(s) = As^2 + Bs^3 + Cs^4$. Numerical values of the parameters of the potential (displayed in Table 1) are designed so that so that W is bounded below and has the minimum at $s = s_o = 0.7$ and a local minimum at $s = 0$.

[4] J.M. Ball, A. Zarnescu, *Orientable and non-orientable director fields for liquid crystals*, Proc. in Appl. Math. and Mech. (2007)

[5] L. Ambrosio, *Existence of minimal energy configurations of nematic liquid crystals with variable degree of orientation*, Manuscr. Math., **68** 215-228 (1990).

[6] G. Barles, B. Perthame, Discontinuous solutions of deterministic optimal stopping time problems, ESAIM: Math. Model. Numer. Anal. 21 (4) (1987) 557C579.

[7] Barrett J W , Feng X , Prohl A . Convergence of a fully discrete finite element method for a degenerate parabolic system modelling nematic liquid crystals with variable degree of orientation. Esaim Mathematical Modelling and Numerical Analysis, 2006, 40(1): 175-199.

[8] P. Biscari, P. Cesana, *Ordering effects in electric splay Freedericksz transitions*, Cont. Mech. and Thermod., **19** 285 298 (2007)

[9] Calderer M.C., Golovaty D. , Lin F.-H. and Liu C. , Time evolution of nematic liquid crystals with variable degree of orientation. SIAM J. Math. Anal. 2002, 33, 1033-1047.

[10] P.M. Chaikin, T.C. Lubensky, Principles of Condensed Matter Physics, Cambridge University Press. 2000.

[11] De Gennes, R G . : Phenomenology of short-range-order effects in the isotropic phase of nematic materials. Phys. Lett. 1969, 30 A 454-455.

[12] P.G. de Gennes, J. Prost, The Physics of Liquid Crystals, 2nd edition ed., Clarendon Press. 1995.

[13] I. Dierking, Textures of Liquid Crystals, 1st edition ed., Wiley-VCH, Weinheim, 2003.

[14] D. Dunmur, T. Sluckin, Soap, Science, and Flat-screen TVs: A History of Liquid Crystals, OUP Oxford. 2010.

[15] Ericksen J L . Liquid crystals with variable degree of orientation. Archive for Rational Mechanics and Analysis, 1991, 113(2): 97-120.

[16] Fan, C.: Disclination lines in liquid crystals, Phys. Lett. 1972, 34A 335-336.

[17] Fan, C. and Stephen, M. J. : Isotropic-nematic phase transitions in liquid crystals, Phys. Rev. Lett. 1970, 25 500-503.

[18] F.C. Frank, P.J. Wojtowicz, P. Sheng, *On the theory of liquid crystals*, Discuss. Faraday Soc., **25** 19-28 (1958)

[19] R. Hardt, D. Kinderlehrer, F.H. Lin, *Existence and partial regularity of static liquid crystal Configurations*, Commun. Math. Phys., **105** 547-570 (1986)

[20] I.-C. Khoo, S.-T. Wu, Optics and Nonlinear Optics of Liquid Crystals, World Scientific. 1993.

[21] O.D. Lavrentovich, P. Pasini, C. Zannoni, S. Zumer (Eds.), Defects in Liquid Crystals: Computer Simulations, Theory and Experiments, Springer, Dordrecht; Boston., 2001,

[22] F.H. Lin, *On nematic liquid crystals with variable degree of orientation*, Comm. on Pure and App. Math., **XLIV** 453-468 (1991)

[23] L. Longa, D. Monselesan, H.R. Trebin, *An extension of the Landau-Ginzburg-de Gennes theory for liquid crystals*, Liquid Crystals, **2** N. 6 769-796 (1987)

[24] E.B. Priestly, P.J. Wojtowicz, P. Sheng, *Introduction to liquid crystals*, Plenum Press New York, (1975)

[25] Maugin G. A., Configurational forces: Modern Mechanics and Mathematics, Chapman and Hall/CRC, 2010.

[26] Ladyzenskaya O, Solonnikov V, Uralceva N. Linear and Quasilinear Equations of Parabolic Type. Translations of Mathematical Monographs. Providence: AMS; 1968.

[27] Sheng W, Zhu P. Viscosity solutions to a model for solid-to-solid phase transitions driven by material forces. Nonlinear Analysis: Real World Applications, 2018, 39: 14-32.

[28] M.J. Stephen, J.P. Straley, Physics of liquid crystals, Rev. Mod. Phys. 46 (4) (1974) 617–704.

[29] Virga E G . Variational Theories for Liquid Crystals. Chapman and Hall, 1994.

[30] X. Xu, Existence for a model arising from the in situ vitrification process. J. Math. Anal. Appl. 271. 2002: 333–342.

[31] Junzhi Zheng and Pierluigi Cesana, Solutions to a new strongly coupled phase-field model for nematic liquid crystals with variable degree of orientation, in preparation.

Solutions to a new strongly coupled phase-field model for nematic liquid crystals with variable degree of orientation

Zheng Junzhi

Department of Mathematics,
Shanghai University, China

2022.1.5

Catalogue

Background

Method and Model

Difficulties and Results

Reference

Background

- Since the second half of the 20th century, the world has entered the era of new technological revolution, new materials have become a breakthrough in the development of various high-tech fields.
- Liquid crystal is an intermediate substance with some properties of both liquid and solid crystal. Its unique properties can be expressed by coupling relationship between deformation and degree of order.

Phase-Filed Method

Although phase-field method is a "young" method, studied from 1980s, they have a wide range of applications.

Phase-field method combines phase-field variables with other field variables (such as stress field, temperature field, etc.) to describe the formation and evolution of microstructure. Phase-field variables can describe the system interface implicitly at various scales, and are coupled with other variables to achieve a unified description of the system, which makes up for the defects of the sharp interface model.

Phase-field model has become a powerful tool for meso-scopic simulation of microstructure evolution such as solidification, precipitation growth and coarsening, solid-solid phase transformation, and crack propagation of various materials.

The quasi-static equations

$$\dot{s} = \left(\text{div} \left\{ \frac{\partial W_1}{\partial \nabla s} \right\} - \frac{\partial W_1}{\partial s} - W'(s) \right) \cdot \alpha(s) \quad (1)$$

$$\beta(s) \cdot \dot{\mathbf{n}} \times \mathbf{n} = \left(\text{div} \left\{ \frac{\partial W}{\partial \nabla \mathbf{n}} \right\} - \frac{\partial W}{\nabla \mathbf{n}} \right) \times \mathbf{n} \quad (2)$$

$$|\mathbf{n}|^2 = 1 \quad (3)$$

▸ $s(x,t) \in \mathbb{R}$ (the order parameter) characterizes the averaged orientation of molecules compared to the unit vector $\mathbf{n}(x,t)$ (the director).

▸ $W_1(s, \nabla s, \nabla \mathbf{n}) = \frac{k_1}{2} |\nabla s|^2 + \frac{k_2}{2} s^2 |\nabla \mathbf{n}|^2$.

The phase-field model for nematic liquid crystal

$$
\begin{aligned}
s_t &= -(W'(s) - k_1 \Delta_x s - k_2 \phi)|\nabla_x s|, & (x,t) \in Q_T \\
\phi_t &= k_3 \, \text{div}(s^2 \nabla_x \phi), & (x,t) \in Q_T \\
s(x,0) &= s_0(x), \quad \phi(x,0) = \phi_0(x), & x \in \Omega, \\
s(x,t) &= 0, \qquad \phi(x,t) = 0. & (x,t) \in \partial\Omega \times (0,T),
\end{aligned}
$$

▸ $\Omega \subset R^d, d = 1, 2, 3.$ $Q_T := \Omega \times (0,T).$

▸ $k_1, k_2, k_3 = const > 0.$

▸ $W(s) \in C^2(\mathbb{R}, [0, \infty))$ is a double-well potential which has at least two minimum points, and one maximum point in-between.

Difficulties

- The liquid crystal model is strongly coupled and degenerate nonlinear parabolic equations;
- $W'(s)$ is the partial derivative of $W(s)$ respect to s and without exact function expression;
- The degradation and high-nonlinearity of the equations make numerical calculations difficult.

Results (1-D)

Model I

$$s_t = \left(k_1 s_{xx} - k_2 s|\phi_x|^2 - W'(s)\right)|s_x|,$$
$$(s^2 + \varepsilon^2)\phi_t = k_3((s^2 + \varepsilon^2)\phi_x)_x.$$

Model II

$$s_t = k_1 s_{xx} - k_2 s|\phi_x|^2 - W'(s),$$
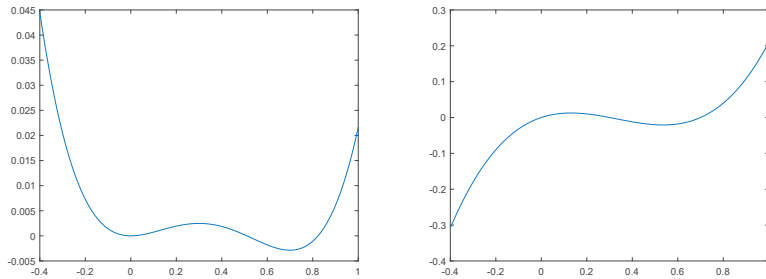$$(s^2 + \varepsilon^2)\phi_t = k_3((s^2 + \varepsilon^2)\phi_x)_x.$$

Fig1. Relation between the chemical free energy density function and order parameter (left); the derivative of the chemical free energy density function and order parameter (right).

Conclusion

▶ The evolutionary process of two models is similar;

▶ The directivity of nematic liquid crystal will be different due to the odd and even difference of the number of discontinuities contained in the initial value;

▶ The more discontinuities the nematic liquid crystal contains, the greater the change of the total energy during evolution.

Reference

- Kelker H. History of liquid crystals[J]. Molecular Crystals and Liquid Crystals, 1973, 21(1-2): 1-48.

- Ericksen J L . Liquid crystals with variable degree of orientation[J]. Archive for Rational Mechanics and Analysis, 1991, 113(2): 97-120.

- Calderer M.C. ,Golovaty D. , Lin F.-H. and Liu C. , Time evolution of nematic liquid crystals with variable degree of orientation[J]. SIAM J. Math. Anal. 2002, 33, 1033-1047.

Partial covalent interactions effect on oxygen diffusion

Nguyen Xuan Thi

Institute of International Carbon-Neutral Energy
Research, Kyushu University, Japan

1, Introduction:

Reducing the dependence on fossil fuels requires more efficient Solid Oxide Fuel Cells (SOFCs). One of the determining factors of SOFCs' efficiency is how fast can electrolyte conduct oxygen. Each electrolyte material has different oxygen conductivity because of varying interactions between oxygen and surroundings ions. Most oxygen conducting materials need dopants, but those dopants hinder oxygen conduction. Understanding how oxygen interacts with these ions can guide the choice of electrolyte material and so we need accurate models of oxygen diffusion.

There are currently two models of oxygen diffusion – the classical ionic model and the partial covalent conduction model. The conventional model assumes that all ions in the lattice are fully ionized, and oxygen conduction is hindered by steric hindrance from cations and electrostatic interactions. According to this model, the most favorable scenario is oxygen travelling through cations with small radius and small positive charge. The classical theory completely ignores oxygen-cation electronic interactions but not all electrolyte materials are fully ionized. The partial covalent model, however, involves partially ionized atom with partial covalent bonds. This partial covalent theory considers electronic interactions along with steric hindrance and electrostatic force. Including these partial covalent bonds is likely to produce very different conclusions to those of the conventional model. Recent investigation on cathode materials is supporting this partial covalent theory.[1,2]

We examined the effect of partial covalent interactions on oxygen migration, activation energy and electron density of all possible migration pathways were calculated using plane wave-DFT in this research.

The most common electrolyte materials are oxides with fluorite structure. This research focuses on investigating three materials: Y doped ZrO_2, Y doped CeO_2, δ-Bi_2O_3. ZrO_2 and CeO_2 are conventionally used for electrolytes in SOFCs, and the four valent ions (Zr, Ce) need to be substituted with two trivalent ions like Y to generate one oxygen vacancy and allow oxygen diffusion. δ-Bi_2O_3 is a less commonly used material because of its instability under 730°C, but it has intrinsic oxygen vacancies and is the best oxygen conducting oxide reported.

This research will explore how electronic interactions affect the oxygen diffusion in these three common fluorite lattices using plane-wave density functional method. Then, this research will suggest how machine learning can further explore this issue.

The diffusion of O inside a lattice can be considered as random jumps between adjacent sites. The diffusion coefficient indicates how easy oxygen diffuse in a material, and can be calculated using Einstein relation in the form of equation (1):[3]

$$D = \frac{n_P}{2d} l^2 \Gamma \quad (1)$$

Where D (cm^2s^{-1}) is the diffusion coefficient, n_P is the number of jump sites available, d is the dimension of lattice, l (cm) is the jumped distance, and Γ (s^{-1}) is the frequency of jumps.

The frequency of jumps can be calculated using Arrhenius-like equation:

$$\Gamma = v^* e^{\frac{-E_a}{k_B T}} \quad (2)$$

Here E_a (eV) is the activation energy of oxygen diffusion, T (K) is temperature, k_B is the Boltzmann constant, in this case $k_B = 8.617 \times 10^{-5}$ (eV K^{-1}). v^* (Hz) is the attempt frequency. In solid state system, Vineyard's harmonic transition state theory is suitable for calculating v^*:[4]

$$v^* = \frac{\prod_{i=1}^{N} v_i^o}{\prod_{j=1}^{N-1} v_j^s} \quad (3)$$

The numerator is the product of all vibrational frequencies of the normal modes of the ground state. The denominator is the product of all vibrational frequencies of the normal modes of the transition state, excluding one imaginary frequency.

2. Methods:

Linear Combination of Atomic Orbitals Density Functional Theory (LCAO-DFT) calculations were performed using Quantum ATK software. The SCAN-metaGGA functional and Pseudodojo pseudopotential with Medium basis set was applied for all atoms. The calculations were performed with $4 \times 4 \times 3$ k points for Y doped ZrO_2 and δ-Bi_2O_3, $5 \times 5 \times 3$ k point for Y doped CeO_2. The valence electron configuration is $4s^2 4p^6 4d^2 5s^2$ for Zr, $4s^2 4p^6 4d^1 5s^2$ for Y, $4f^1 5s^2 5p^6 5d^1 6s^2$ for Ce, $5d^{10} 6s^2 6p^3$ for Bi, $2s^2 2p^4$ for O. Lattice were optimized with 0.05 eV/Å force tolerance and LBFGS optimizer method. Density mesh cut off is set at 125 Hartree. The electronic energies were converged to 10^{-6} eV with the Gaussian smearing method.

The activation energy and transition state geometry of vacancy-mediated oxygen transport data were optimized using climbing-image nudge elastic band (CI-NEB) method with 0.03 eV/Å force tolerance. A varying number of intermediate images were used, depending on different pathway. The intermediate images were created by interpolating the position of initial and final states that were optimized.

The vibrational frequencies of normal modes of initial and transition state were calculated by combining dynamical matrix and vibrational mode calculation of Quantum ATK. Initial and transition geometry were converged to 10^{-8} eV to ensure that the vibrational frequency calculation yield proper results. Each ion in the cell is then displaced by ± 0.015 Å in all direction of cartesian vector, then construct a Hessian matrix from the force. Force derivative tolerance is set to 10^{-3} eV/Å. Once the normal mode vibrational frequencies of system were obtained, multiplying all data (excluding the imaginary one) and divide the data of initial to transition frequency to get attempt frequency.

3. Results and discussion:

Activation energy, attempt frequency, and diffusion coefficient were calculated and shown in Table 1. The smaller E_a and larger D value indicates better oxygen conductivity. Results show δ-Bi_2O_3 is the best oxygen conducting material, follows by oxygen migrating through Zr rich edge in ZrO_2, then oxygen migrating through CeO_2, and oxygen migrating through Y rich edge of ZrO_2. The trend of variation in activation energy is the opposite of that in D, which is desirable, while attempt frequency has minor effect on value of D. Thus, this research will focus on investigating only activation energy value.

In the lattice $Zr_{14}O_{31}Y_2$ and $Ce_{14}O_{31}Y_2$ (one oxygen vacancy), there are five possible oxygen migration pathways in total: two yttrium ions in the adjacent tetrahedral (model 1), one yttrium ion in the adjacent tetrahedral (model 2), no yttrium ion nearby (model 3), two yttrium ions in the common edge (model 4), and one yttrium ion in the common edge (model 5). All pathways were constructed from the same $Zr_{14}O_{31}Y_2$ and $Ce_{14}O_{31}Y_2$ models

E_a (eV)	v^* (Hz)	D (cm²s⁻¹)
δ Bi₂O₃ 0.398	ZrO₂ - Zr 0.016 x 10⁶	ZrO₂ - Y 5.069 x 10⁻²²
ZrO₂ - Zr 0.790	CeO₂ - Ce 0.029 x 10⁶	CeO₂ – Ce 2.96 x 10⁻¹⁸
CeO₂ – Ce 1.18	CeO₂ - Y 0.043 x 10⁶	CeO₂ – Y 4.91 x 10⁻¹⁸
CeO₂ – Y 1.2	ZrO₂ - Y 1.04 x 10⁶	ZrO₂ - Zr 2.467 x 10⁻¹⁶
ZrO₂ - Y 2.31	δ Bi₂O₃ 30.3 x 10⁶	δ Bi₂O₃ 1.399 x 10⁻⁹

Table 1. Summarized data of activation energy, attempt frequency and diffusion

with varying oxygen vacancy position. This paper considered oxygen migration pathways in <100>, <010>, <001> direction to cover all possible oxygen migration configuration. The oxygen migration pathways with the corresponding activation energy and electron density data will be schematically shown in Fig. 1. (in $Zr_{14}O_{31}Y_2$ lattice) and Fig. 2. (in $Ce_{14}O_{31}Y_2$ lattice).

Activation energy and electron density maps of all possible oxygen migration pathways were explored. Electron density map offers an intuitive understanding of electronic interaction. The degree of electronic interaction from electron density map is shown as blue numbers in the below figures. Larger number means more partial covalent interaction between atoms and vice versa.

3.1 Oxygen migration in Y doped ZrO₂:

There were five oxygen migration pathways in Y doped ZrO₂. Pathways 1, 2,

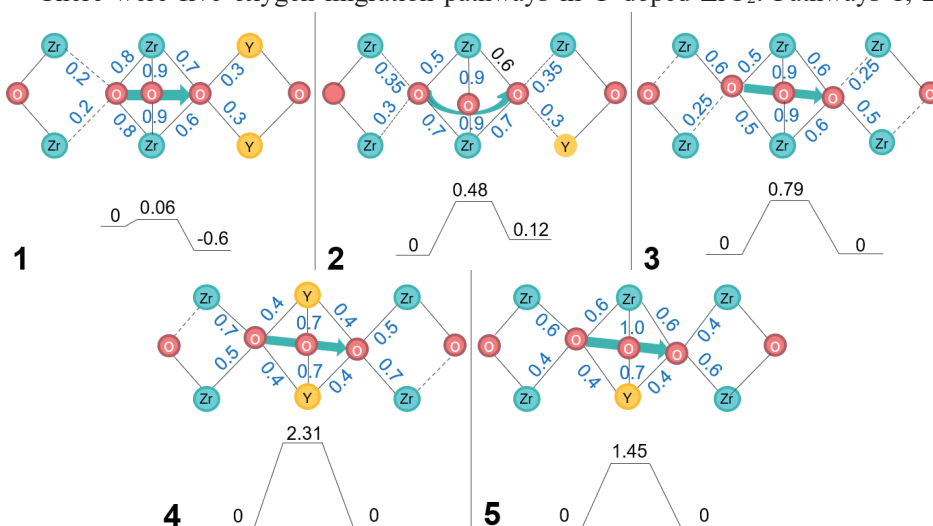
Figure 1. Possible pathways of oxygen diffusion in ZrO₂ dope Y with corresponding electron density results and activation energy results

and 3 have the same transition state configuration, and the increment of E_a from pathway 1 to 3 is attributed to different degrees of electron density overlap at the initial state. The interaction of oxygen with Zr strength increase from path 1 to 2 to 3, corresponding with lower initial energy, higher bond-breaking penalty, and higher E_a. Paths 3, 4, and 5 have the same initial configuration and different transition state configuration, and the diversity of E_a is attributed to different degrees of electron density overlap at the transition state. Zr has more overlapping with oxygen than Y, leading to a more stabilized transition state and lowered E_a in paths 3 and 5 compared to path 4.

3.2 Oxygen migration in Y doped CeO₂:

Generally, activation energy of oxygen moving in Y doped CeO_2 lattice is higher than that of Y doped ZrO_2. The notable point is when oxygen is moving through the same yttrium ions transition state in ceria and zirconia, the activation energy is different (Fig. 2.).
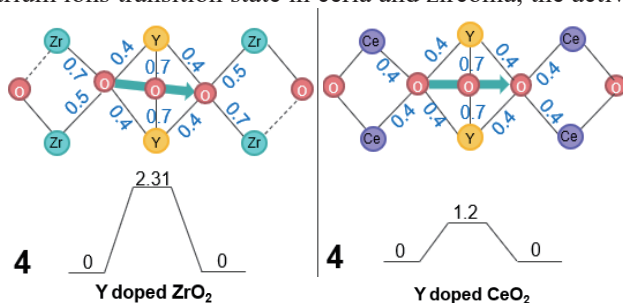

Figure 2. Oxygen diffusion in Y doped ZrO_2 and Y doped CeO_2 with corresponding electron density results, activation energy results.

2.). The activation of that in ceria is lower than zirconia because the Ce-O interactions at the initial state is weaker than the Zr-O interactions at the initial state.

3.3 Oxygen migration in δ-Bi_2O_3:

Oxygen migration in δ-Bi_2O_3 completed in two steps and has the activation energy of only 0.4 eV. The energy diagram in this case is shown in Fig. 3. It is worth to mention that the oxygen in δ-Bi_2O_3 has curved migration trajectory. Oxygen also migrated in this curved trajectory in Y doped ZrO_2 in both step 1 (Fig. 4.) and step 2 (Fig. 5.). From the electron density data, it can be concluded that oxygen moved in curve pathways because they prefer to maximize interaction with bismuth cations, rather than avoiding the cations as expected.
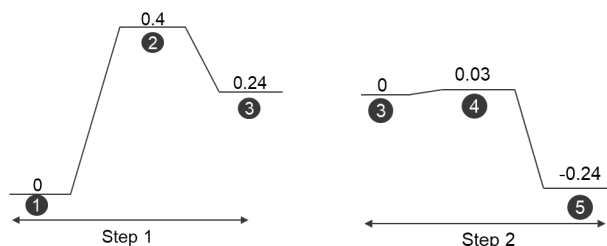

Figure 3. Energy diagram for two steps oxygen migration in δ-Bi_2O_3

3.4 Application of machine learning:

As the experimental results and computational results have shown, adding dopants are necessary, but the presence of dopant can hinder oxygen transportation. Mixing suitable dopants might mitigate the trapping effect of dopant. However, it is time consuming trying to find the
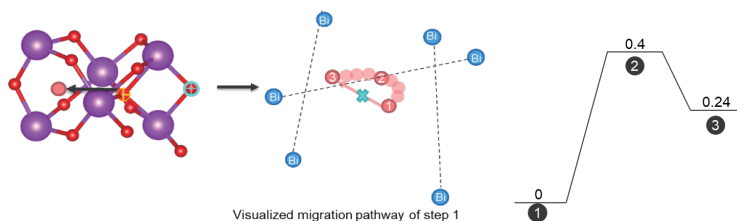

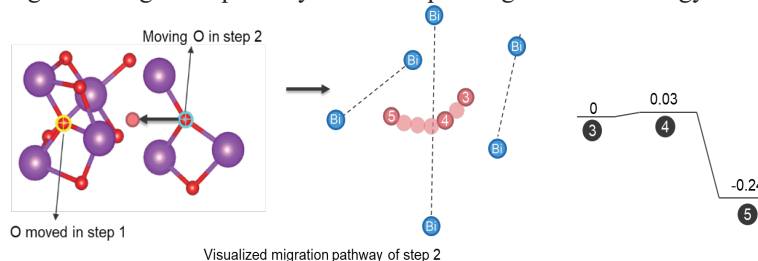Figure 4. Migration pathway and corresponding activation energy of step 1

Figure 5. Migration pathway and corresponding activation energy of step 2

suitable mixing ratio of dopants with experiments or with density functional theory. Therefore,

machine learning can be used to find such desirable combinations. We have seen a relationship between having strong partial covalent interactions at the transition state and relatively weak partial covalent interactions at the initial state, and low oxygen migration activation barrier. We can use this relationship and machine learning to predict possible good oxygen conducting dopants ratio. First, we need to generate activation energy calculations and electron density of as many as possible dopant combination that were reported before in experiments. This can be used for input data and use image analyzing algorithm to identify the strong partial covalent bond at the transition state and relatively weak partial covalent bond at initial state from the electron density images. Then we can use machine learning to predict possible combination and test the results using DFT and experiments. However, generating DFT results for the initial input will take time, and might not give a large enough input set.

4. Conclusion:

In conclusion, this research has investigated oxygen migration in three fluorite materials: Y doped ZrO_2, Y doped CeO_2, and δ-Bi_2O_3. This research has suggested a connection between partial covalent interactions and oxygen diffusion activation energy. Partial covalent interaction stabilizes transition state energy and lowers activation energy. Strong partial covalent interaction at initial state increases activation energy due to increased bond breaking energy penalty. Maximizing partial covalent interaction with cations and be in their vicinity is preferable for oxygen migration. Machine learning can offer great help to save time and energy for finding better dopants combination in oxygen conducting materials.

References

[1]   Wu, J.; Fujii, K.; Yashima, M.; Staykov, A.; Akbay, T.; Ishihara, T.; Kilner, J. A. A Systematic Evaluation of the Role of Lanthanide Elements in Functional Complex Oxides; Implications for Energy Conversion Devices. *Journal of Materials Chemistry A* **2018**, *6* (25), 11819–11829. https://doi.org/10.1039/C8TA01191E.

[2]   Zhao, Y.; Kilner, J.; Ishihara, T.; Yoshizawa, K.; Staykov, A. Effect of Electronic Interactions and Coordination Spheres on Ionic Diffusion in LaxSr1-XGayMg1-YO3-δ. *Journal of Physics and Chemistry of Solids* **2022**, *161*, 110393. https://doi.org/10.1016/J.JPCS.2021.110393.

[3]   Kong, L. T.; Lewis, L. J. Transition State Theory of the Preexponential Factors for Self-Diffusion on Cu, Ag, and Ni Surfaces. *Physical Review B - Condensed Matter and Materials Physics* **2006**, *74* (7). https://doi.org/10.1103/PhysRevB.74.073412.

[4]   Koettgen, J.; Zacherle, T.; Grieshammer, S.; Martin, M. Ab Initio Calculation of the Attempt Frequency of Oxygen Diffusion in Pure and Samarium Doped Ceria. *Physical Chemistry Chemical Physics* **2017**, *19* (15), 9957–9973. https://doi.org/10.1039/c6cp04802a.

Machine learning approach to ophthalmological circadian therapeutics from retina transcriptome dataset

Debashis Sahu[†][a], Soundhar Ramasamy[†][b], Ganesh Ram Sahu[c], Surbhi Sharma[b], Arkasubhra Ghosh[c], Swaminathan Sethu[c], Daniel Packwood[b] and Ganesh N. Pandian*[b]

[a]Mitsui Chemicals, Inc.-Carbon Neutral Research Center (MCI-CNRC), I2CNER, Kyushu University, Fukuoka 819-0395, Japan

[b]Institute for Integrated Cell-Material Sciences (iCeMS), Kyoto University, Kyoto 606-8501, Japan

[c]GROW Research Laboratory, Narayana Nethralaya Foundation, Bangalore 560-099, India

The circadian clock is a cell autonomous oscillator that coordinates daily rhythms of many living organisms[1,2]. Daily rhythms in mammalian behavior and physiology are a significant biological clock of life on Earth[3,4]. This characteristic has developed throughout the series of evolution process on this planet in accordance with periodic variation of environmental factors like the 24-hour day and night cycle. This 24-hour circadian cycles are controlled by internal autonomous oscillator, circadian clocks. In mammals, the principle pacemaker which generates a regular rhythm with an approximate 24-h period is situated in the suprachiasmatic nucleus (SCN)[5].

In this context, the rhythmic functions in human's circadian pacemaker are one of the important areas of research[6]. In the 24-hour cycle of a day, specific biological functions in human body are critically controlled where the different gene activities in human tissues varies with a daily rhythm. Some of them are active in day light time while some others are active in dark time (Figure 1). Human organ-specific understanding of biological pathways and its temporal organization is a topic of great importance for circadian based optimal drug administration time from a clinical perspective[7].
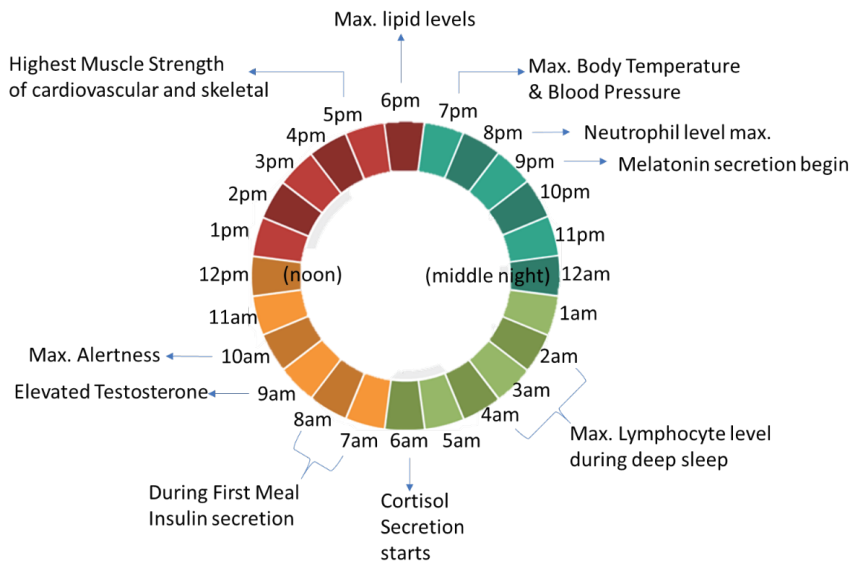
Figure 1: Outline of rhythmic functions in human circadian pacemaker over the 24-h cycle

Vision is an automatic rhythmic function which is critically influenced by the light intensity occurring over the 24-hour day[8]. Human retina plays major role to both the vision and non-vision specific functions with a comparatively less complex cellular population than the brain[9] (Figure 2). The interplay of cellular and molecular processes within the retina are harmonized to the Earth's day and night cycle[8]. A network of circadian systems in the eye mainly involves a light input pathway, circadian oscillator, and multiple output pathways for the fine-tune detection and processing of light information over the 24-hour period[10]. The recent high throughput studies with systematic analyses showed that several clock genes expression in the retina indicating at specific functions or pathways under circadian control[11]. The retinal endogenous circadian clock controls many physiological processes within the neural retina to regulate the visual sensitivity to ambient light levels[10,11]. However, the involvement of specific clock genes is poorly understood.
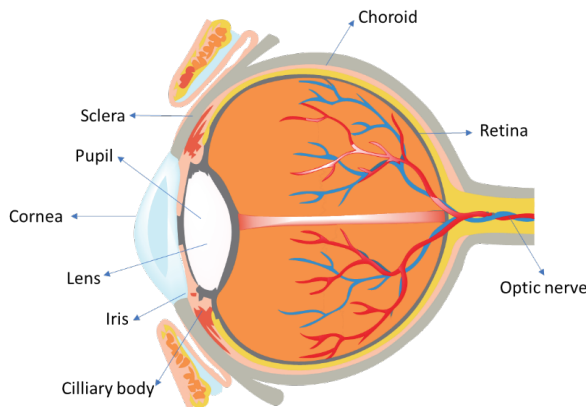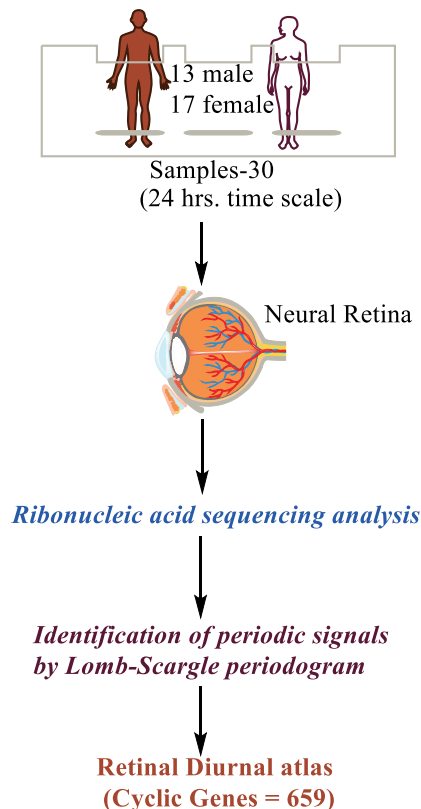
Figure 2: Human eye with a schematic enlargement of retina

Our main interests here to deal with macular degeneration from both basic scientific and the clinical therapeutic point of view[12]. Age related macular degeneration (AMD) is a common threat to vision which leads to the progressive and irreversible loss of central vision in eyes[13]. It happens when aging causes destruction of the macula in the retina which controls sharp, straight-ahead vision. The macula is the dense collection of light-sensitive tissue at the back of the eye contributes to the central vision[10].

To address the AMD under the human circadian clock[14], we need to recognize the optical retina genomes whose activities varies with a daily rhythm[15]. Because the proper detection of the circadian clock genomes could guide us to improve the dosing time of day for many existing drugs of AMD[16]. An important advantage of this approach is that some selective tissue can be readily targetable while avoiding injury to normal tissue in practice. This promising clinical principle commonly termed as circadian therapeutics is emerged into wound healing drug delivery formulation in recent days[12].

The Machine Learning (ML) approach to ophthalmological circadian therapeutics from retina transcriptome datasets is quite interesting. To understand this, we have applied the machine learning approach for the cyclic genes extraction and understanding of this rhythmicity. This on-going project was conducted with the collaboration with Prof. Namasivayam Ganesh Pandian and Prof. Daniel Packwood Labs in ICEMS, Kyoto University. First, we have collected the 659 cyclic genes from 30 samples (17 female & 13 male) with known daily oscillatory features time (Figure 3). We know that it's just impractical and dangerous to take samples from an individual around the clock to monitor how gene activity in a specific cell type varies. In this regard, machine learning approaches have shown promising results in determining the clock genes from retinal transcriptomes datasets[17].

13 male
17 female

Samples-30
(24 hrs. time scale)

Neural Retina

Ribonucleic acid sequencing analysis

*Identification of periodic signals
by Lomb-Scargle periodogram*

**Retinal Diurnal atlas
(Cyclic Genes = 659)**

Figure 3: Stepwise process for the Samples collection and extraction of the retinal cyclic

genes

From the bulk transcriptomes datasets for the human retina[18–21], we have taken four publicly available datasets which are related to human retinal transcriptome. Here, n is the no of samples. Then we have used the ML algorithm called cyclops[22] to estimate the rhythmic phases of gene by using experimental seed genes. By using this ML algorithm, we able to extract the cyclic genes from the datasets and make a comparison of this rhythmicity (Figure 4).

Publicly Available Datasets

human retinal transcriptome

SS-1 (n=160)
SS-2 (n=172)
SS-3(n=112)
SS-4 (n=61)

ML-Algorithm "CYCLOPS"
(using 659 seed genes)

Phase estimation of genes

1. Clock Genes Extraction
2.Rhythimicity Comparison
over macular samples
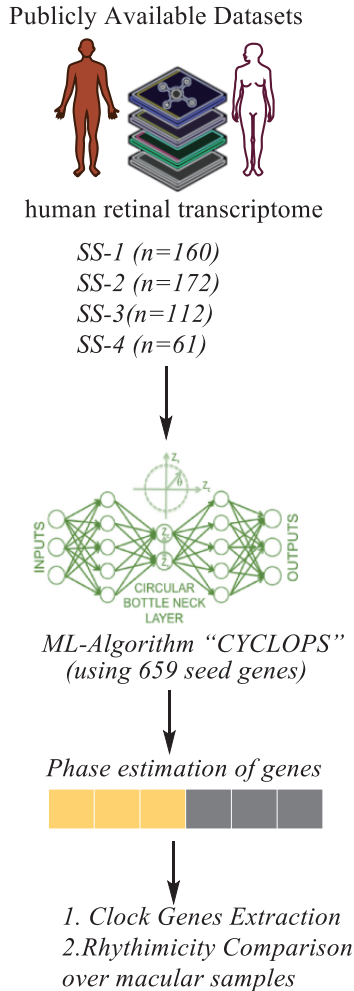
Figure 4: The modelled machine learning workflow from four publicly available datasets to identify the clock genes and their rhythmicity.

We have used the four-different sets of macular samples which is defined here as SS-1, SS-2, SS-3 and SS-4 in our studies. SS-1 is control state of sample while the SS-2, SS-3 and SS-4 are corresponding to the 2nd , 3rd and 4th are the different states of macular samples. In cyclops, the 24 hour of a day is presented into 0 to 2π time scale phase[4,22]. From the cyclops predicted results, we have noticed that the oscillatory nature of the optical retina genes are gradually decreased from the central macular sample SS-1 to SS-4. This also indicates that circadian behavior of the retinal transcriptome is significant in control state of dataset rather than the other more disease datasets. We have extracted around 34 clock genes by applying the cyclops algorithm[22] which follow the circadian rhythmicity. Based on the results, we aim to create a reliable map of the human retinal diurnal transcriptome map by light and dark cycle of earth's day explaining the temporal

organization of biological pathways in our ongoing projects. We believe that such a diurnal atlas on the human retina will be an essential resource for the researchers bent on exploring the principles of circadian therapeutics in ophthalmology. Our present research work is going in this direction.

Machine Learning Methods

The CYCLOPS (Unsupervised Learning) autoencoder and downstream analysis[22] were implemented in Julia 0.3.10. The associated files are available for download on GitHub.

Author Information

[†]Both authors contributed equally to the work.

Corresponding Author: *namasivayam.ganeshpandian.5z@kyoto-u.ac.jp

Acknowledgements

References

1. Golombek, D. A. & Rosenstein, R. E. Physiology of Circadian Entrainment. *Physiol. Rev.* **90**, 1063–1102 (2010).
2. Ramasamy, S. *et al.* Identification of novel circadian transcripts in the zebrafish retina. *J. Exp. Biol.* (2018) doi:10.1242/jeb.192195.
3. Zheng, X., Zhang, K., Zhao, Y. & Fent, K. Environmental chemicals affect circadian rhythms: An underexplored effect influencing health and fitness in animals and humans. *Environ. Int.* **149**, 106159 (2021).
4. Wu, G. *et al.* Population-level rhythms in human skin with implications for circadian medicine. *Proc. Natl. Acad. Sci.* **115**, 12313–12318 (2018).
5. Hughey, J. J. Machine learning identifies a compact gene set for monitoring the circadian clock in human blood. *Genome Med.* **9**, 19 (2017).
6. Masri, S. & Sassone-Corsi, P. The emerging link between cancer, metabolism, and circadian rhythms. *Nat. Med.* **24**, 1795–1803 (2018).
7. Ruben, M. D. *et al.* A database of tissue-specific rhythmically expressed human genes has potential applications in circadian medicine. *Sci. Transl. Med.* **10**, (2018).
8. Zang, J. *et al.* Circadian regulation of vertebrate cone photoreceptor function. *Elife* **10**, (2021).
9. Cherry, T. J. *et al.* Mapping the cis -regulatory architecture of the human retina reveals noncoding genetic variation in disease. *Proc. Natl. Acad. Sci.* **117**, 9001–9012 (2020).
10. London, A., Benhar, I. & Schwartz, M. The retina as a window to the brain—from eye research to CNS disorders. *Nat. Rev. Neurol.* **9**, 44–53 (2013).
11. McMahon, D. G., Iuvone, P. M. & Tosini, G. Circadian organization of the mammalian retina: From gene regulation to physiology and diseases. *Prog. Retin. Eye Res.* **39**, 58–76 (2014).
12. Ruan, W., Yuan, X. & Eltzschig, H. K. Circadian rhythm as a therapeutic target. *Nat. Rev. Drug Discov.* **20**, 287–307 (2021).
13. Kaarniranta, K. *et al.* Autophagy and heterophagy dysregulation leads

to retinal pigment epithelium dysfunction and development of age-related macular degeneration. *Autophagy* **9**, 973–984 (2013).

14. Felder-Schmittbuhl, M.-P. *et al.* Ocular Clocks: Adapting Mechanisms for Eye Functions and Health. *Investig. Opthalmology Vis. Sci.* **59**, 4856 (2018).

15. Schumacker, S. T., Coppage, K. R. & Enke, R. A. RNA sequencing analysis of the human retina and associated ocular tissues. *Sci. Data* **7**, 199 (2020).

16. Ratnapriya, R. *et al.* Retinal transcriptome and eQTL analyses identify genes associated with age-related macular degeneration. *Nat. Genet.* **51**, 606–610 (2019).

17. Whalen, S., Schreiber, J., Noble, W. S. & Pollard, K. S. Navigating the pitfalls of applying machine learning in genomics. *Nat. Rev. Genet.* **0123456789**, (2021).

18. Pinelli, M. *et al.* An atlas of gene expression and gene co-regulation in the human retina. *Nucleic Acids Res.* **44**, 5773–5784 (2016).

19. Farkas, M. H. *et al.* Transcriptome analyses of the human retina identify unprecedented transcript diversity and 3.5 Mb of novel transcribed sequence via significant alternative splicing and novel genes. *BMC Genomics* **14**, 486 (2013).

20. Orozco, L. D. *et al.* Integration of eQTL and a Single-Cell Atlas in the Human Eye Identifies Causal Genes for Age-Related Macular Degeneration. *Cell Rep.* **30**, 1246-1259.e6 (2020).

21. Macosko, E. Z. *et al.* Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **161**, 1202–1214 (2015).

22. Anafi, R. C., Francey, L. J., Hogenesch, J. B. & Kim, J. CYCLOPS reveals human transcriptional rhythms in health and disease. *Proc. Natl. Acad. Sci.* **114**, 5312–5317 (2017).

Density functional theory analysis for adsorption of hydrogen sulfide on graphene-based structures

**Takaya Fujisaki[a], Kei Ikeda[b],
Aleksandar Tsekov Staykov[c],
Hendrik Srtiwan[d], Yusuke Shiratori[e]**

[a] Institute of Multidisciplinary Research for Advanced Materials, Tohoku University
[b] Institute for Materials Chemistry and Engineering and Integrated Research Consortium on Chemical Science (IRCCS), Kyushu University
[c] International Institute for Carbon-neutral Energy Research (WPI-I2CNER), Kyushu University
[d] Hydrogen Energy Systems, Graduate School of Engineering, Kyushu University
[e] Department of Mechanical Engineering, Faculty of Engineering, Kyushu University

Ideally, hydrogen should be produced by electrolysis of water using renewable energy, and some Organisation for Economic Cooperation and Development (OECD) countries, such as Japan and Germany, have already begun demonstration experiments[1] [2] . On the other hand, there are many technical challenges to make it economically self-sustaining, and only the above countries have introduced hydrogen-based economies in the OECD, and none in the non-OECD [3] [4]. The introduction of a hydrogen-based economy to the non-OECD is a major challenge for us to build a sustainable society on a global scale, given that energy demand growth is expected to be much higher in the future than in the OECD   [5, 6]. As a means to successfully meet this challenge, a system combining biogas and fuel cells, shown in Fig. 1, has been attracting much attention in recent years [7, 8]. This system uses organic waste from agriculture and aquaculture as biomass. This biomass is converted into biogas through anaerobic wastewater treatment (methane fermentation), and the methane in it is used as an energy carrier for hydrogen, which can be used to extract electrical energy from methane using solid oxide fuel cells. As a reaction process to extract hydrogen from methane, hydrogen is produced by methane and carbon dioxide according to the reforming reaction shown in reaction equations (1) and (2) [7, 8]. The hydrogen reacts with oxide ions from the cathode side according to the reaction equation (3) on the anode side of the solid oxide fuel cell, and electrical energy can be extracted to the outside [9]. In this case, the oxide ions are supplied by oxygen on the cathode side.

$$CH_4 + H_2O \leftrightarrows CO + 3H_2 \ (1)$$
$$CH_4 + CO_2 \leftrightarrows 2CO + 2H_2 \ (2)$$
$$H_2 + O^{2-} \leftrightarrows H_2O \ (3)$$
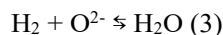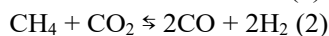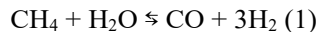
The carbon monoxide produced by the reaction equation (1) reacts with oxygen in the air and is released into the atmosphere as carbon dioxide. Since biogas is derived from the waste, biogas-based power generation is a clean way to generate electricity, considering that the carbon dioxide emitted by this reaction is released back into the atmosphere[10-12]. A system combining biomass-derived biogas and fuel cells has already been demonstrated. Shiratori *et al.* fermented biomass derived from shrimp farms in Vietnam to generate biogas, which was then used in solid oxide fuel cells to generate electricity. This effort is one of the attempts to introduce a hydrogen-based

economy to developing countries with thriving primary industries by utilizing methane as an energy carrier for hydrogen [10-12].

To efficiently run a system combining biogas and solid oxide fuel cells, it is essential to make the fuel cells more efficient in generating electricity. Ideally, only methane, carbon dioxide, and water, as shown in reaction equations (1) and (2), should be used as feedstock for this purpose, but actual biogas contains trace amounts of hydrogen sulfide [13] [14] [15]. Hydrogen sulfide is known to enter the biogas during the process of anaerobic digestion of biodegradable materials[16]. Although this amount is reported to be as low as 1 ppm, hydrogen sulfide is known to degrade the metal materials in fuel cells[17] as well as reduce the efficiency of power generation. Even if the amount is reduced to 0.001 ppm, the power generation efficiency of solid oxide fuel cells is reported to decrease, and hydrogen sulfide is recognized as a substance that should be removed as much as possible[18-22].

Nanocarbons, with their high specific surface area, pore structure, and surface functionality, have recently attracted attention as a material capable of removing hydrogen sulfide in a wide temperature range from room temperature to 900°C[23-28]. Since nanocarbons can be utilized as catalysts without metal-based active sites, they are expected to avoid degradation of catalytic performance due to hydrogen sulfide poisoning. The reaction to remove hydrogen sulfide on nanocarbons is the dissociation reaction shown in equation (5), followed by the reaction of HS$^-$ with oxygen radicals as shown in reaction equation (6), where sulfur is recovered.

$$H_2S \rightarrow H^+ + HS^- \quad (5)$$
$$HS^- + O^* \rightarrow S + OH^- \quad (6)$$
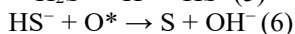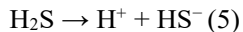
O* in the above reaction equation is the oxygen radical [29].

Recently, Shiyan *et al.* conducted comparative experiments on the effects of nitrogen and oxygen doped in nanocarbons on the oxidation reaction of H_2S. The results showed that the structure called pyridinic N, in which nitrogen is doped in nanocarbons and hydrogen is attached to the nitrogen as a terminal atom, promotes the oxidation reaction of H_2S the most. The structure called oxidized N, in which an oxygen atom is substituted for the hydrogen bonded next to the nitrogen atom of pyridinic N, was shown to promote the oxidation reaction of H_2S the least in the comparison experiments. Furthermore, Li *et al.* attempted to explain the above experimental results using density functional theory [30]. However, the crystal model targeted by this density functional theory was structurally optimized by placing only HS$^-$ on the carbon atoms of pyridinic N and oxidized N, which is insufficient to explain the actual oxidation reaction of hydrogen sulfide.

Therefore, we believe that it is not appropriate to focus on HS$^-$ to explain the adsorption properties of hydrogen sulfide on pyridinic N and oxidized N, and that we should focus on H_2S, which is the preliminary step of the reaction equation (6), and use density functional theory to clarify the difference in adsorption properties of H_2S on pyridinic N and oxidized N. This is the purpose of this study.
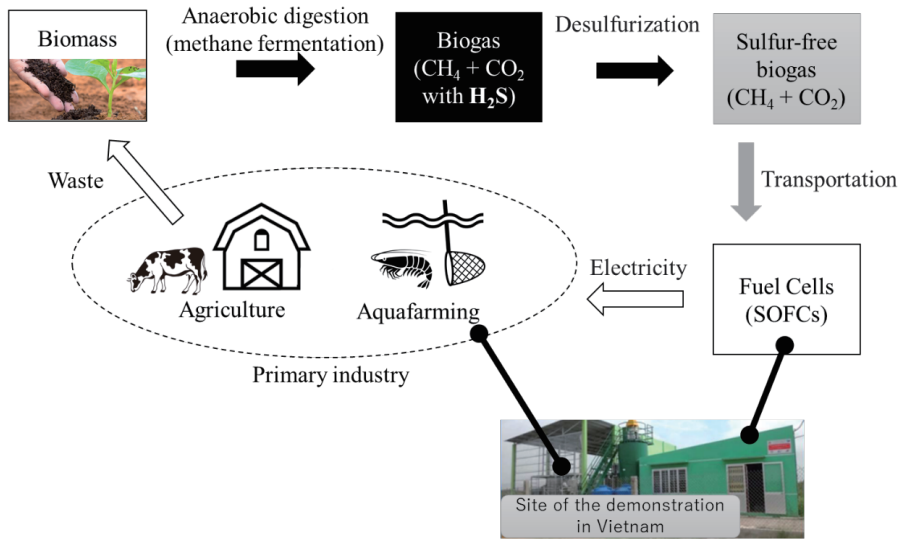
Fig. 1 Conceptual diagram of the demonstration experiment of energy circulation using biomass being conducted in Vietnam

All DFT calculations were performed using the Vienna Ab initio simulation package (VASP) [31-33]. The Perdew-Burke-Ernzerhof (PBE) [34] exchange-correlation functional was employed based on the projector-augmented wave (PAW) method [33]. The cutoff energy for the plane wave basis was set to 400 eV, referring to previous studies[35]. Firstly, 4×4 graphene sheet was constructed and optimized using 5×5×1 Monkhorst-Pack k-points meshes, respectively. Also, the spin polarization was considered. For Projector Augmented Wave (PAW) method, pseudopotentials with valence states of C ($2s^2$, $2p^2$), N ($2s^2$, $2p^3$), S ($3s^2$, $3p^4$), O ($2s^2$, $2p^4$), and H ($1s$) were used for all calculations. All ionic positions were optimized by a conjugate gradient method until the forces on each ion were below 10^{-2} eV/Å^2. Electronic energy was converged to 10^{-4} eV and ion positions were converged to forces of 10^{-2} eV/Å^2. Cell parameters were converged to stress of 10^{-2} eV/Å.

Next, to introduce nitrogen and oxygen atoms into the graphene structure, we prepared two graphene optimized by the above method and made a vacuum slab by making the c-axis direction of the unit cell 15 Å in size. The six-membered ring structure consisting of six carbons was then removed from graphene, and hydrogen was introduced at the ends of the remaining carbon atoms. It has been reported that the terminations of graphene are bonded with hydrogen atoms. Subsequently, some carbon sites were replaced with N only in Fig. 2(a) and with nitrogen and oxygen in Fig. 2(b) to create pyridinic N and oxidized N. Hydrogen sulfide molecules were placed on these pyridinic N and oxidized N. The absorptivity of hydrogen sulfide was evaluated by the magnitude of the interaction energy (E_{int}), which is shown in Equation 7 below [36].

$$E_{int} = E_{H_2S+ \text{ graphene}} - (E_{\text{graphene}} + E_{H_2S}) \qquad (7)$$

$E_{H_2S+ \text{ graphene}}$，$E_{\text{graphene}}$，and E_{H_2S} are the electrostatic potential of hydrogen sulfide molecules adsorbed with pyridinic N /oxidized N, the electrostatic potential of pyridinic N /oxidized N, and the electrostatic potential of a single hydrogen sulfide molecule, respectively. The interaction energies shown above are -0.13 eV for hydrogen sulfide adsorbed on pyridinic and -0.11 eV for oxidized N, as shown in Fig 3. This indicates that pyridinic N is a more favorable material for adsorbing hydrogen sulfide than oxidized N, a trend that is consistent with the experimental results shown by Shiyan et al. in their previous study[37].
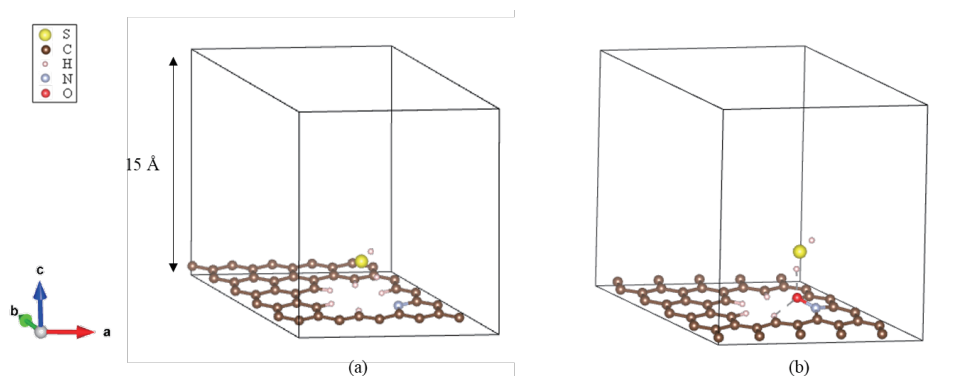
Fig. 2 Geometry optimized structure of H_2S on pyridinic N (a) and oxidized N (b). Red, brown, yellow, gray, white balls indicate oxygen, carbon, sulfur, nitrogen, and hydrogen.

	E_{int} (eV)
■ pyridinic N	-0.13
■ oxidized N	-0.11
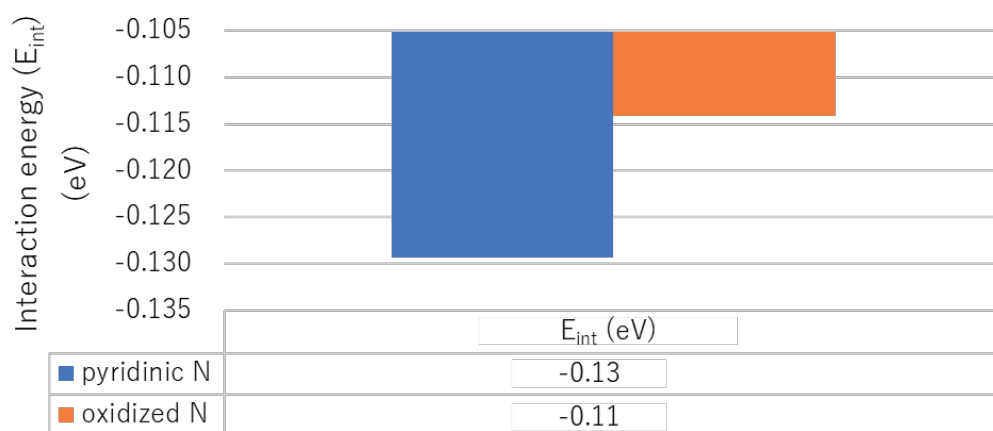
Fig. 3 The result of interaction energy between pyridinic N / oxidized N and H_2S

References

[1] R. Samsun, L. Antoni, M. Rex, Mobile fuel cell application: tracking market trends, IEA Technology Collaboration Program (2020).

[2] E. Weidner, R. Ortiz Cebolla, J. Davies, Global deployment of large capacity stationary fuel cells, Tech. Rep, 2019.

[3] Y. Shiratori, T. Quang-Tuyen, Y. Umemura, T. Kitaoka, K. Sasaki, Paper-structured catalyst for the steam reforming of biodiesel fuel, Int. J. Hydrogen Energy 38(26) (2013) 11278-11287.

[4] Ö. Aydın, A. Kubota, D.L. Tran, M. Sakamoto, Y. Shiratori, Designing graded catalytic domain to homogenize temperature distribution while dry reforming of CH4, Int. J. Hydrogen Energy 43(36) (2018) 17431-17443.

[5] U.S.D.o. Energy, April 2021 Monthly Energy Review, U.S. Energy Information Administration.

[6] J.F. Mercure, H. Pollitt, J.E. Viñuales, N.R. Edwards, P.B. Holden, U. Chewpreecha, P. Salas, I. Sognnaes, A. Lam, F. Knobloch, Macroeconomic impact of stranded fossil fuel assets, Nature Climate Change 8(7) (2018) 588-593.

[7] Y. Shiratori, M. Sakamoto, T.G.H. Nguyen, T. Yamakawa, T. Kitaoka, H. Orishima, H. Matsubara, Y. Watanabe, S. Nakatsuka, T.C.D. Doan, C.M. Dang, Biogas Power Generation with SOFC to Demonstrate Energy Circulation Suitable for Mekong Delta, Vietnam, Fuel Cells 19(4) (2019) 346-353.

[8] S.A. Saadabadi, A. Thallam Thattai, L. Fan, R.E.F. Lindeboom, H. Spanjers, P.V. Aravind, Solid

Oxide Fuel Cells fuelled with biogas: Potential and constraints, Renewable Energy 134 (2019) 194-214.

[9] Y. Hu, L. Peng, X. Li, X. Yao, H. Lin, T. Chi, A novel evolution tree for analyzing the global energy consumption structure, Energy 147 (2018) 1177-1187.

[10] D. Divya, L.R. Gopinath, P. Merlin Christy, A review on current aspects and diverse prospects for enhancing biogas production in sustainable means, Renewable and Sustainable Energy Reviews 42 (2015) 690-699.

[11] E. Bocci, A. Di Carlo, S.J. McPhail, K. Gallucci, P.U. Foscolo, M. Moneti, M. Villarini, M. Carlini, Biomass to fuel cells state of the art: A review of the most innovative technology solutions, Int. J. Hydrogen Energy 39(36) (2014) 21876-21895.

[12] H.J. Alves, C. Bley Junior, R.R. Niklevicz, E.P. Frigo, M.S. Frigo, C.H. Coimbra-Araújo, Overview of hydrogen production technologies from biogas and the applications in fuel cells, Int. J. Hydrogen Energy 38(13) (2013) 5215-5225.

[13] M. Syed, G. Soreanu, P. Falletta, M. Béland, Removal of hydrogen sulfide from gas streams using biological processes• a review, Canadian Biosystems Engineering 48 (2006) 2.

[14] C.H. Pham, S. Saggar, P. Berben, T. Palmada, C. Ross, Removing Hydrogen Sulfide Contamination in Biogas Produced from Animal Wastes, J. Environ. Qual. 48(1) (2019) 32-38.

[15] F.B. Asghari, J. Jaafari, M. Yousefi, A.A. Mohammadi, R. Dehghanzadeh, Evaluation of water corrosion, scaling extent and heterotrophic plate count bacteria in asbestos and polyethylene pipes in drinking water distribution system, Human and Ecological Risk Assessment: An International Journal 24(4) (2018) 1138-1149.

[16] Z. Cheng, J.-H. Wang, Y. Choi, L. Yang, M.C. Lin, M. Liu, From Ni-YSZ to sulfur-tolerant anode materials for SOFCs: electrochemical behavior, in situ characterization, modeling, and future perspectives, Energy & Environmental Science 4(11) (2011).

[17] J. Kupecki, D. Papurello, A. Lanzini, Y. Naumovich, K. Motylinski, M. Blesznowski, M. Santarelli, Numerical model of planar anode supported solid oxide fuel cell fed with fuel containing H2S operated in direct internal reforming mode (DIR-SOFC), Applied Energy 230 (2018) 1573-1584.

[18] R. Dehghan, M. Anbia, Zeolites for adsorptive desulfurization from fuels: A review, Fuel Process. Technol. 167 (2017) 99-116.

[19] A. Peluso, N. Gargiulo, P. Aprea, F. Pepe, D. Caputo, Nanoporous materials as H2S adsorbents for biogas purification: a review, Separation & Purification Reviews 48(1) (2019) 78-89.

[20] M. Ozekmekci, G. Salkic, M.F. Fellah, Use of zeolites for the removal of H2S: a mini-review, Fuel Process. Technol. 139 (2015) 49-60.

[21] Z. Jiang, X. Xu, Y. Ma, H.S. Cho, D. Ding, C. Wang, J. Wu, P. Oleynikov, M. Jia, J. Cheng, Y. Zhou, O. Terasaki, T. Peng, L. Zan, H. Deng, Filling metal-organic framework mesopores with TiO2 for CO2 photoreduction, Nature 586(7830) (2020) 549-554.

[22] M. Khabazipour, M. Anbia, Removal of Hydrogen Sulfide from Gas Streams Using Porous Materials: A Review, Industrial & Engineering Chemistry Research 58(49) (2019) 22133-22164.

[23] Z. Yu, X. Wang, X. Song, Y. Liu, J. Qiu, Molten salt synthesis of nitrogen-doped porous carbons for hydrogen sulfide adsorptive removal, Carbon 95 (2015) 852-860.

[24] F. Sun, J. Liu, H. Chen, Z. Zhang, W. Qiao, D. Long, L. Ling, Nitrogen-rich mesoporous carbons: highly efficient, regenerable metal-free catalysts for low-temperature oxidation of H2S, Acs Catalysis 3(5) (2013) 862-870.

[25] M. Daraee, M. Baniadam, A. Rashidi, M. Maghrebi, Synthesis of TiO2-CNT hybrid nanocatalyst and its application in direct oxidation of H2S to S, Chemical Physics 511 (2018) 7-19.

[26] M. Sun, X. Wang, X. Pan, L. Liu, Y. Li, Z. Zhao, J. Qiu, Nitrogen-rich hierarchical porous carbon nanofibers for selective oxidation of hydrogen sulfide, Fuel Processing Technology 191 (2019) 121-128.

[27] Y. Pan, M. Chen, M. Hu, M. Tian, Y. Zhang, D. Long, Probing the room-temperature oxidative desulfurization activity of three-dimensional alkaline graphene aerogel, Applied Catalysis B: Environmental 262 (2020) 118266.

[28] G. Lei, Y. Cao, W. Zhao, Z. Dai, L. Shen, Y. Xiao, L. Jiang, Exfoliation of graphitic carbon nitride for enhanced oxidative desulfurization: A facile and general strategy, ACS Sustainable Chemistry & Engineering 7(5) (2019) 4941-4950.

[29] D. Li, W. Chen, J. Wu, C.Q. Jia, X. Jiang, The preparation of waste biomass-derived N-doped carbons and their application in acid gas removal: focus on N functional groups, Journal of Materials Chemistry A 8(47) (2020) 24977-24995.

[30] S. Li, Q. Gu, N. Cao, Q. Jiang, C. Xu, C. Jiang, C. Chen, C. Pham-Huu, Y. Liu, Defect enriched N-doped carbon nanoflakes as robust carbocatalysts for H2S selective oxidation, Journal of Materials Chemistry A (2020).

[31] G. Kresse, J. Furthmüller, Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set, Computational materials science 6(1) (1996) 15-50.

[32] G. Kresse, D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, Physical review b 59(3) (1999) 1758.

[33] G. Kresse, J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, Physical review B 54(16) (1996) 11169.

[34] J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. 77(18) (1996) 3865.

[35] Q. Ren, N. Qin, B. Liu, Y. Yao, X. Zhao, Z. Deng, Y. Li, Y. Dong, D. Qian, B.-L. Su, W. Zhang, H.-E. Wang, An oxygen-deficient vanadium oxide@N-doped carbon heterostructure for sodium-ion batteries: insights into the charge storage mechanism and enhanced reaction kinetics, Journal of Materials Chemistry A 8(6) (2020) 3450-3458.

[36] M. Shahabi, H. Raissi, Investigation of the molecular structure, electronic properties, AIM, NBO, NMR and NQR parameters for the interaction of Sc, Ga and Mg-doped (6,0) aluminum nitride nanotubes with COCl2 gas by DFT study, Journal of Inclusion Phenomena and Macrocyclic Chemistry 84(1-2) (2015) 99-114.

[37] S. Li, Q. Gu, N. Cao, Q. Jiang, C. Xu, C. Jiang, C. Chen, C. Pham-Huu, Y. Liu, Defect enriched N-doped carbon nanoflakes as robust carbocatalysts for H2S selective oxidation, Journal of Materials Chemistry A 8(18) (2020) 8892-8902.

Mineralization Process in Porous Media using Random Walk with Absorption

Linh Thi Hoai NGUYEN

Institute of Mathematics for Industry,
Kyushu University

Random walk with absorbtion
Particle flow through pore rock
Mineralization process

1. Motivation & Purpose

Motivation: Active researches on permeability caused by carbonate precipitation in pororous media and its important applications, e.g., geological CO_2 storage, geothermal power.$\implies$ It is of vital important to quantify the hydraulic and elastic property of natural rocks.

Objectives: Get understanding about the dynamics of mineralization process in pore rocks caused by carbonate precipitation in porous media through the so-called random walk with absorption method.

Permeability: is a measure of the ability of a porous material (often, a rock or an unconsolidated material) to allow fluids to pass through it.

Hydraulics is a topic in engineering dealing with the mechanical properties of liquids.

Random walk with absorbtion
Particle flow through pore rock
Mineralization process

Table: Statistics on void structure

	Bentheimer	Doddington	Estaillades	Ketton
Porosity	0.217	0.196	0.127	0.133
No. connected components	33,920	280	14,8329	39,190
No. passable components	1	1	1	1
Account for	99.764%	99.484%	85.668%	99.002%

Porosity is an important measure:

$$\text{porosity} = \frac{\text{pore volume}}{\text{total volume}}$$

If pore spaces have the same porosity, the hydraulic and elastic properties can be different.

68

Random walk with absorbtion
Particle flow through pore rock
Mineralization process

Random walk method

Suppose that at time t a particle (red) are stay at position x_t. Let $\mathcal{N}(x_t)$ denote the set of neighbors (black star) of x_t. At the step $t+1$, the particle can move to one of its current position's neighbor, say y with a probability

$p(x_t \to y)$ with $\sum_{y \in \mathcal{N}(x_t)} = 1$.
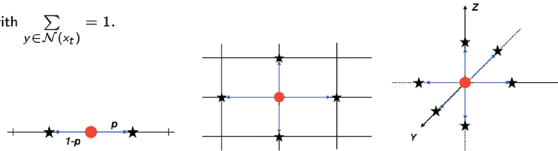
RW with absorption in porous media
Neighbor set in porous media: $\mathcal{N}_v(x_t) = \{y \in \mathcal{N}(x_t) | y \text{ is a void point}\}$.

$$p(x_t \to y) = \frac{g(x_t, y)}{\sum_{z \in \mathcal{N}_v(x_t)} g(x_t, z)},$$

where $g(x_t, z) = v + d$ if particle approches z toward absorption direction, and v otherwise. Here v : velocity and d : dispersion.

Random walk with absorbtion
Particle flow through pore rock
Mineralization process

Random walk with absorption algorithm

K: max of iteration step,
N: no. of injected particles,
v: velocity, d: dispersion, $k = 0$

$k = k + 1$

RW_k yes $k < K$?

no

Elapse time,
No. of particle pass
through each void
position in the rock

Absorpting /input surface

Absorpting direction

Escaping /output surface

69

Random walk with absorbtion
Particle flow through pore rock
Mineralization process

Travel time

$N = 100.000$ particles are injected into the rock, only a small part of which can escape from the upper surface within the given period of time ($K = 5,000.000$).
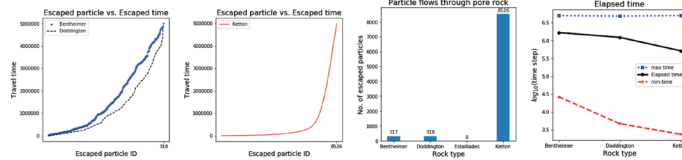
Porosity seems not to be the factor that affects the number of escaped particles (First subfigure). However, the porosity does affect the elapsed time (see last subfigure).

Random walk with absorbtion
Particle flow through pore rock
Mineralization process

Particle flow with respect to v

Definition

Elapsed time is the average of travel time of all escaped particles within an allotted time.

Question: Will the elapsed time will improve if we increase the proportion v/d?

Parameter settings: $v = 1$ and $v = 10$, $N = 100,000$ (Number of particles), $d = 0.01$ (Dispersion), $K = 5,000,000$.

Random walk with absorbtion
Particle flow through pore rock
Mineralization process
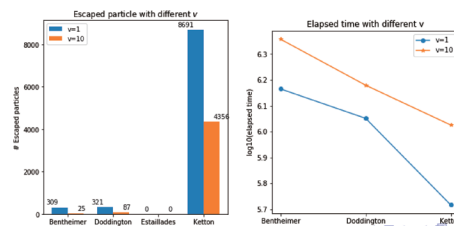
Observations from numerical result

- Number of escaped particles do not increase as the velocity *v* increases.

- The number of escaped particles from Ketton is much larger than those from Bentheimer and Doddington. Also, the elapsed time of Ketton rock is remarkably lower than those of Bentheimer and Doddington. Desspite the fact that the porosity of Ketton is lower than the other two. The reason for this can be explain through the heterogeneity of the rock. Jiang et al. [3] reported that Ketton is the most heterogonous among the three. We also consider the heterogeneity problem in another work through the so-called Gaussian Mixture method [4].

[3] Fei Jiang, Takeshi Tsuji, and Tomoyuki Shirai, Pore geometry charactererization by persistent homology theory, *Water Resource Research*, **(54)**, 4150-4163 (2018).

[4] Linh Thi Hoai Nguyen, Tomoyuki Shirai, and Takeshi Tsuji, Pore rock model and classification, (preprint).

Random walk with absorbtion
Particle flow through pore rock
Mineralization process

Mineralization algorithm

Update rule: Void point becomes rock point with probability specified by:

$$\sum_{i=1}^{k}(1-q)^{i-1}q,$$

$q \in (0, 1)$: probability of precipitation,

k : number of time that particles pass throught the corresponding void position.

[Flowchart:]
q : probability of precipitation
T: number of trials, $t = 0$,
Inject N particles to the input surface

$t = t + 1$

Add rock points (update rule) Inject more particles (equals number of escaped ones)

$t < T$?

no → Get new rock image

yes

Perform RW with absorption

Random walk with absorbtion
Particle flow through pore rock
Mineralization process
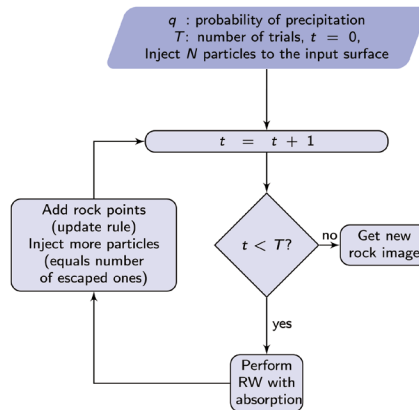
Mineralization process through pore rock

For visualizing the mineralization process through pore rock, we will show the porosity changes on a 2D slice as the simulation proceed. We will show the result of the middle slice ($z = 500$) of the Bentheimer rock with $q = 0.0001$. In these figure, rock point are in gray, void points are in white and mineralized point are in red.
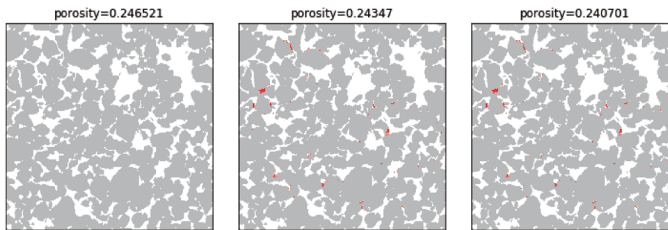
Original   At 25-th trial   At 50-th trial

Random walk with absorbtion
Particle flow through pore rock
Mineralization process

Mineralization process with respect to probability of precipitation

Simulation results show that porosities can be viewed as some smooth function of the number of trials (or according to time). This "smooth relationship" is revealed by using the curve fitting method.

Mineralization process with respect to probabiity of precipitation.

Random walk with absorbtion
Particle flow through pore rock
Mineralization process
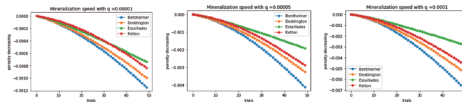
Porosity as function of trials using curved fitting tool

Porosity of Bentheimer as function of trials with $q = 0.00001$:

$$f(x) = a \exp(bx) + c \exp(dx).$$

Porosity with respect to trial of Bentheimer rock
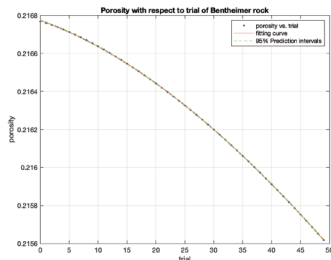
Coefficients (with 95% confidence bounds):

$a = 0.2209$		$(0.219, 0.2227)$
$b = -0.004714$		$(-0.005343, -0.004085)$
$c = -0.004514$		$(-0.00639, -0.002639)$
$d = -0.1528$		$(-0.1834, -0.1222)$

We find that the fitting curve describes the porosity changes process quite well, precisely, SSE: $8.402e - 11$, R-square: 1, RMSE: $1.352e - 06$. For more details about these goodness criteria.

We see that this sum of exponential function is also a good fit for all other type of rock and different values of probability of precipitation (q).

Random walk with absorbtion
Particle flow through pore rock
Mineralization process

Conclusions

- We have introduced the RW with absorption algorithm as a tool for studying the quantity of pore space. The algorithm is easy to implement and parallelizable.

- The algorithm are used to study the particle flow distribution and the mechanism of mineralization process on porous media.

- This algorithm can also be used to generated artificial porous media samples.

Thank you very much for your kind attention!

73

Ensemble Effect of Pd@Au Electrocatalysts on CO_2 Reduction Reaction: A DFT-based Machine-learning Study

Fuzhu Liu[1], Chao Wu[2*], Shengchun Yang[1], Xiangdong Ding[1*]

[1]State Key Laboratory for Mechanical Behavior of Materials, MOE Key Laboratory for Non-Equilibrium Synthesis and Modulation of Condensed Matter, Xi'an Jiaotong University, Xi'an 710049, People's Republic of China

[2]Frontier Institute of Science and Technology, Xi'an Jiaotong University, Xi'an 710054, People's Republic of China

Abstract

The ensemble effect due to variation of Pd content in Pd-Au alloys have been widely investigated for several important reactions, including CO_2 reduction reaction (CO_2RR), however, identifying the stable Pd arrangements on the alloyed surface and picking out the active sites are still challenging. Here we use a machine-learning (ML)-based approach coupled with density functional theory (DFT) calculations to efficiently find the low-energy configurations of Pd-Au(111) surface alloys and the potentially active sites for CO_2RR, fully covering the Pd content from 0 to 100%. The ML model is actively learning process to improve the predicting accuracy for the configuration formation energy and to find the stable Pd-Au(111) alloyed surfaces, respectively. The local surface properties of adsorption sites are classified into two classes by the K-means clustering approach, which are closely related to the Pd content on Au surface. The classification is reflected in the variation of adsorption energy of CO and H: In the low Pd content range (0 - 60 %) the adsorption energies over the surface alloys can be tuned significantly, and in the medium Pd content (37 - 68%), the catalytic activity of surface alloys for CO_2RR can be increased by increase the Pd content and attributed to the meta-stable active site over the surface. The results demonstrated that the combination of DFT and machine-learning (supervised and unsupervised) approaches are feasible to understand the surface-related properties.

Keywords: Ensemble effect, Machine-learning, DFT calculations, CO_2RR, K-means clustering

MI レクチャーノートシリーズ刊行にあたり

　本レクチャーノートシリーズは、文部科学省 21 世紀 COE プログラム「機能数理学の構築と展開」（H15-19 年度）において作成した COE Lecture Notes の続刊であり、文部科学省大学院教育改革支援プログラム「産業界が求める数学博士と新修士養成」（H19-21 年度）および、同グローバル COE プログラム「マス・フォア・インダストリ教育研究拠点」（H20-24 年度）において行われた講義の講義録として出版されてきた。平成 23 年 4 月のマス・フォア・インダストリ研究所（IMI）設立と平成 25 年 4 月の IMI の文部科学省共同利用・共同研究拠点として「産業数学の先進的・基礎的共同研究拠点」の認定を受け、今後、レクチャーノートは、マス・フォア・インダストリに関わる国内外の研究者による講義の講義録、会議録等として出版し、マス・フォア・インダストリの本格的な展開に資するものとする。

<div align="right">

2022 年 10 月
マス・フォア・インダストリ研究所
所長　梶原 健司

</div>

IMI Workshop of the Joint Usage Research Projects

Perspectives on Artificial Intelligence and Machine Learning in Materials Science, February 4-6, 2022

シリーズ既刊

Issue	Author／Editor	Title	Published
COE Lecture Note	Mitsuhiro T. NAKAO Kazuhiro YOKOYAMA	Computer Assisted Proofs - Numeric and Symbolic Approaches - 199pages	August 22, 2006
COE Lecture Note	M.J.Shai HARAN	Arithmetical Investigations - Representation theory, Orthogonal polynomials and Quantum interpolations- 174pages	August 22, 2006
COE Lecture Note Vol.3	Michal BENES Masato KIMURA Tatsuyuki NAKAKI	Proceedings of Czech-Japanese Seminar in Applied Mathematics 2005 155pages	October 13, 2006
COE Lecture Note Vol.4	宮田　健治	辺要素有限要素法による磁界解析 - 機能数理学特別講義 21pages	May 15, 2007
COE Lecture Note Vol.5	Francois APERY	Univariate Elimination Subresultants - Bezout formula, Laurent series and vanishing conditions - 89pages	September 25, 2007
COE Lecture Note Vol.6	Michal BENES Masato KIMURA Tatsuyuki NAKAKI	Proceedings of Czech-Japanese Seminar in Applied Mathematics 2006 209pages	October 12, 2007
COE Lecture Note Vol.7	若山　正人 中尾　充宏	九州大学産業技術数理研究センター キックオフミーティング 138pages	October 15, 2007
COE Lecture Note Vol.8	Alberto PARMEGGIANI	Introduction to the Spectral Theory of Non-Commutative Harmonic Oscillators 233pages	January 31, 2008
COE Lecture Note Vol.9	Michael I.TRIBELSKY	Introduction to Mathematical modeling 23pages	February 15, 2008
COE Lecture Note Vol.10	Jacques FARAUT	Infinite Dimensional Spherical Analysis 74pages	March 14, 2008
COE Lecture Note Vol.11	Gerrit van DIJK	Gelfand Pairs And Beyond 60pages	August 25, 2008
COE Lecture Note Vol.12	Faculty of Mathematics, Kyushu University	Consortium "MATH for INDUSTRY" First Forum 87pages	September 16, 2008
COE Lecture Note Vol.13	九州大学大学院 数理学研究院	プロシーディング「損保数理に現れる確率モデル」 — 日新火災・九州大学 共同研究2008年11月 研究会 — 82pages	February 6, 2009

シリーズ既刊

Issue	Author／Editor	Title	Published
COE Lecture Note Vol.14	Michal Beneš, Tohru Tsujikawa Shigetoshi Yazaki	Proceedings of Czech-Japanese Seminar in Applied Mathematics 2008 77pages	February 12, 2009
COE Lecture Note Vol.15	Faculty of Mathematics, Kyushu University	International Workshop on Verified Computations and Related Topics 129pages	February 23, 2009
COE Lecture Note Vol.16	Alexander Samokhin	Volume Integral Equation Method in Problems of Mathematical Physics 50pages	February 24, 2009
COE Lecture Note Vol.17	矢嶋　　徹 及川　正行 梶原　健司 辻　　英一 福本　康秀	非線形波動の数理と物理　66pages	February 27, 2009
COE Lecture Note Vol.18	Tim Hoffmann	Discrete Differential Geometry of Curves and Surfaces　75pages	April 21, 2009
COE Lecture Note Vol.19	Ichiro Suzuki	The Pattern Formation Problem for Autonomous Mobile Robots ―Special Lecture in Functional Mathematics―　23pages	April 30, 2009
COE Lecture Note Vol.20	Yasuhide Fukumoto Yasunori Maekawa	Math-for-Industry Tutorial: Spectral theories of non-Hermitian operators and their application　184pages	June 19, 2009
COE Lecture Note Vol.21	Faculty of Mathematics, Kyushu University	Forum "Math-for-Industry" Casimir Force, Casimir Operators and the Riemann Hypothesis 95pages	November 9, 2009
COE Lecture Note Vol.22	Masakazu Suzuki Hoon Hong Hirokazu Anai Chee Yap Yousuke Sato Hiroshi Yoshida	The Joint Conference of ASCM 2009 and MACIS 2009: Asian Symposium on Computer Mathematics Mathematical Aspects of Computer and Information Sciences　436pages	December 14, 2009
COE Lecture Note Vol.23	荒川　恒男 金子　昌信	多重ゼータ値入門　111pages	February 15, 2010
COE Lecture Note Vol.24	Fulton B.Gonzalez	Notes on Integral Geometry and Harmonic Analysis　125pages	March 12, 2010
COE Lecture Note Vol.25	Wayne Rossman	Discrete Constant Mean Curvature Surfaces via Conserved Quantities 130pages	May 31, 2010
COE Lecture Note Vol.26	Mihai Ciucu	Perfect Matchings and Applications　66pages	July 2, 2010

シリーズ既刊

Issue	Author／Editor	Title	Published
COE Lecture Note Vol.27	九州大学大学院 数理学研究院	Forum "Math-for-Industry" and Study Group Workshop Information security, visualization, and inverse problems, on the basis of optimization techniques　100pages	October 21, 2010
COE Lecture Note Vol.28	ANDREAS LANGER	MODULAR FORMS, ELLIPTIC AND MODULAR CURVES LECTURES AT KYUSHU UNIVERSITY 2010　62pages	November 26, 2010
COE Lecture Note Vol.29	木田　雅成 原田　昌晃 横山　俊一	Magma で広がる数学の世界　157pages	December 27, 2010
COE Lecture Note Vol.30	原　　　隆 松井　　卓 廣島　文生	Mathematical Quantum Field Theory and Renormalization Theory 201pages	January 31, 2011
COE Lecture Note Vol.31	若山　正人 福本　康秀 高木　　剛 山本　昌宏	Study Group Workshop 2010 Lecture & Report　128pages	February 8, 2011
COE Lecture Note Vol.32	Institute of Mathematics for Industry, Kyushu University	Forum "Math-for-Industry" 2011 "TSUNAMI-Mathematical Modelling" Using Mathematics for Natural Disaster Prediction, Recovery and Provision for the Future　90pages	September 30, 2011
COE Lecture Note Vol.33	若山　正人 福本　康秀 高木　　剛 山本　昌宏	Study Group Workshop 2011 Lecture & Report　140pages	October 27, 2011
COE Lecture Note Vol.34	Adrian Muntean Vladimír Chalupecký	Homogenization Method and Multiscale Modeling　72pages	October 28, 2011
COE Lecture Note Vol.35	横山　俊一 夫　　紀恵 林　　卓也	計算機代数システムの進展　210pages	November 30, 2011
COE Lecture Note Vol.36	Michal Beneš Masato Kimura Shigetoshi Yazaki	Proceedings of Czech-Japanese Seminar in Applied Mathematics 2010 107pages	January 27, 2012
COE Lecture Note Vol.37	若山　正人 高木　　剛 Kirill Morozov 平岡　裕章 木村　正人 白井　朋之 西井　龍映 栄　伸一郎 穴井　宏和 福本　康秀	平成23年度 数学・数理科学と諸科学・産業との連携研究ワークショップ　拡がっていく数学　〜期待される "見えない力" 〜 154pages	February 20, 2012

シリーズ既刊

Issue	Author／Editor	Title	Published
COE Lecture Note Vol.38	Fumio Hiroshima Itaru Sasaki Herbert Spohn Akito Suzuki	Enhanced Binding in Quantum Field Theory　204pages	March 12, 2012
COE Lecture Note Vol.39	Institute of Mathematics for Industry, Kyushu University	Multiscale Mathematics; Hierarchy of collective phenomena and interrelations between hierarchical structures　180pages	March 13, 2012
COE Lecture Note Vol.40	井ノ口順一 太田　泰広 筧　　三郎 梶原　健司 松浦　　望	離散可積分系・離散微分幾何チュートリアル2012　152pages	March 15, 2012
COE Lecture Note Vol.41	Institute of Mathematics for Industry, Kyushu University	Forum "Math-for-Industry" 2012 "Information Recovery and Discovery"　91pages	October 22, 2012
COE Lecture Note Vol.42	佐伯　　修 若山　正人 山本　昌宏	Study Group Workshop 2012 Abstract, Lecture & Report　178pages	November 19, 2012
COE Lecture Note Vol.43	Institute of Mathematics for Industry, Kyushu University	Combinatorics and Numerical Analysis Joint Workshop　103pages	December 27, 2012
COE Lecture Note Vol.44	萩原　　学	モダン符号理論からポストモダン符号理論への展望　107pages	January 30, 2013
COE Lecture Note Vol.45	金山　　寛	Joint Research Workshop of Institute of Mathematics for Industry (IMI), Kyushu University "Propagation of Ultra-large-scale Computation by the Domain-decomposition-method for Industrial Problems (PUCDIP 2012)" 121pages	February 19, 2013
COE Lecture Note Vol.46	西井　龍映 栄　伸一郎 岡田　勘三 落合　啓之 小磯　深幸 斎藤　新悟 白井　朋之	科学・技術の研究課題への数学アプローチ ―数学モデリングの基礎と展開―　325pages	February 28, 2013
COE Lecture Note Vol.47	SOO TECK LEE	BRANCHING RULES AND BRANCHING ALGEBRAS FOR THE COMPLEX CLASSICAL GROUPS　40pages	March 8, 2013
COE Lecture Note Vol.48	溝口　佳寛 脇　　隼人 平坂　　貢 谷口　哲至 島袋　　修	博多ワークショップ「組み合わせとその応用」　124pages	March 28, 2013

シリーズ既刊

Issue	Author／Editor	Title	Published
COE Lecture Note Vol.49	照井　　章 小原　功任 濱田　龍義 横山　俊一 穴井　宏和 横田　博史	マス・フォア・インダストリ研究所　共同利用研究集会 II 数式処理研究と産学連携の新たな発展　137pages	August 9, 2013
MI Lecture Note Vol.50	Ken Anjyo Hiroyuki Ochiai Yoshinori Dobashi Yoshihiro Mizoguchi Shizuo Kaji	Symposium MEIS2013: Mathematical Progress in Expressive Image Synthesis　154pages	October 21, 2013
MI Lecture Note Vol.51	Institute of Mathematics for Industry, Kyushu University	Forum "Math-for-Industry" 2013 "The Impact of Applications on Mathematics"　97pages	October 30, 2013
MI Lecture Note Vol.52	佐伯　　修 岡田　勘三 髙木　　剛 若山　正人 山本　昌宏	Study Group Workshop 2013 Abstract, Lecture & Report 142pages	November 15, 2013
MI Lecture Note Vol.53	四方　義啓 櫻井　幸一 安田　貴徳 Xavier Dahan	平成25年度　九州大学マス・フォア・インダストリ研究所 共同利用研究集会　安全・安心社会基盤構築のための代数構造 〜サイバー社会の信頼性確保のための数理学〜　158pages	December 26, 2013
MI Lecture Note Vol.54	Takashi Takiguchi Hiroshi Fujiwara	Inverse problems for practice, the present and the future　93pages	January 30, 2014
MI Lecture Note Vol.55	栄　伸一郎 溝口　佳寛 脇　　隼人 渋田　敬史	Study Group Workshop 2013 数学協働プログラム Lecture & Report 98pages	February 10, 2014
MI Lecture Note Vol.56	Yoshihiro Mizoguchi Hayato Waki Takafumi Shibuta Tetsuji Taniguchi Osamu Shimabukuro Makoto Tagami Hirotake Kurihara Shuya Chiba	Hakata Workshop 2014 〜 Discrete Mathematics and its Applications 〜　141pages	March 28, 2014
MI Lecture Note Vol.57	Institute of Mathematics for Industry, Kyushu University	Forum "Math-for-Industry" 2014: "Applications + Practical Conceptualization + Mathematics = fruitful Innovation"　93pages	October 23, 2014
MI Lecture Note Vol.58	安生健一 落合啓之	Symposium MEIS2014: Mathematical Progress in Expressive Image Synthesis　135pages	November 12, 2014

シリーズ既刊

Issue	Author／Editor	Title	Published
MI Lecture Note Vol.59	西井　龍映 岡田　勘三 梶原　健司 髙木　　剛 若山　正人 脇　　隼人 山本　昌宏	Study Group Workshop 2014 数学協働プログラム Abstract, Lecture & Report　196pages	November 14, 2014
MI Lecture Note Vol.60	西浦　　博	平成26年度九州大学 IMI 共同利用研究・研究集会（I） 感染症数理モデルの実用化と産業及び政策での活用のための新たな展開　120pages	November 28, 2014
MI Lecture Note Vol.61	溝口　佳寛 Jacques Garrigue 萩原　　学 Reynald Affeldt	研究集会 高信頼な理論と実装のための定理証明および定理証明器 Theorem proving and provers for reliable theory and implementations (TPP2014)　138pages	February 26, 2015
MI Lecture Note Vol.62	白井　朋之	Workshop on "β-transformation and related topics"　59pages	March 10, 2015
MI Lecture Note Vol.63	白井　朋之	Workshop on "Probabilistic models with determinantal structure" 107pages	August 20, 2015
MI Lecture Note Vol.64	落合　啓之 土橋　宜典	Symposium MEIS2015: Mathematical Progress in Expressive Image Synthesis　124pages	September 18, 2015
MI Lecture Note Vol.65	Institute of Mathematics for Industry, Kyushu University	Forum "Math-for-Industry" 2015 "The Role and Importance of Mathematics in Innovation"　74pages	October 23, 2015
MI Lecture Note Vol.66	岡田　勘三 藤澤　克己 白井　朋之 若山　正人 脇　　隼人 Philip Broadbridge 山本　昌宏	Study Group Workshop 2015 Abstract, Lecture & Report 156pages	November 5, 2015
MI Lecture Note Vol.67	Institute of Mathematics for Industry, Kyushu University	IMI-La Trobe Joint Conference "Mathematics for Materials Science and Processing" 66pages	February 5, 2016
MI Lecture Note Vol.68	古庄　英和 小谷　久寿 新甫　洋史	結び目と Grothendieck-Teichmüller 群 116pages	February 22, 2016
MI Lecture Note Vol.69	土橋　宜典 鍛治　静雄	Symposium MEIS2016: Mathematical Progress in Expressive Image Synthesis　82pages	October 24, 2016
MI Lecture Note Vol.70	Institute of Mathematics for Industry, Kyushu University	Forum "Math-for-Industry" 2016 "Agriculture as a metaphor for creativity in all human endeavors" 98pages	November 2, 2016
MI Lecture Note Vol.71	小磯　深幸 二宮　嘉行 山本　昌宏	Study Group Workshop 2016 Abstract, Lecture & Report　143pages	November 21, 2016

シリーズ既刊

Issue	Author／Editor	Title	Published
MI Lecture Note Vol.72	新井　朝雄 小嶋　　泉 廣島　文生	Mathematical quantum field theory and related topics　133pages	January 27, 2017
MI Lecture Note Vol.73	穴田　啓晃 Kirill Morozov 須賀　祐治 奥村　伸也 櫻井　幸一	Secret Sharing for Dependability, Usability and Security of Network Storage and Its Mathematical Modeling　211pages	March 15, 2017
MI Lecture Note Vol.74	QUISPEL, G. Reinout W. BADER, Philipp MCLAREN, David I. TAGAMI, Daisuke	IMI-La Trobe Joint Conference Geometric Numerical Integration and its Applications　71pages	March 31, 2017
MI Lecture Note Vol.75	手塚　　集 田上　大助 山本　昌宏	Study Group Workshop 2017 Abstract, Lecture & Report　118pages	October 20, 2017
MI Lecture Note Vol.76	宇田川誠一	Tzitzéica 方程式の有限間隙解に付随した極小曲面の構成理論 ―Tzitzéica 方程式の楕円関数解を出発点として―　68pages	August 4, 2017
MI Lecture Note Vol.77	松谷　茂樹 佐伯　　修 中川　淳一 田上　大助 上坂　正晃 Pierluigi Cesana 濵田　裕康	平成29年度　九州大学マス・フォア・インダストリ研究所 共同利用研究集会（I） 結晶の界面，転位，構造の数理　148pages	December 20, 2017
MI Lecture Note Vol.78	瀧澤　重志 小林　和博 佐藤憲一郎 斎藤　　努 清水　正明 間瀬　正啓 藤澤　克樹 神山　直之	平成29年度　九州大学マス・フォア・インダストリ研究所 プロジェクト研究　研究集会（I） 防災・避難計画の数理モデルの高度化と社会実装へ向けて 136pages	February 26, 2018
MI Lecture Note Vol.79	神山　直之 畔上　秀幸	平成29年度　AIMaP チュートリアル 最適化理論の基礎と応用　96pages	February 28, 2018
MI Lecture Note Vol.80	Kirill Morozov Hiroaki Anada Yuji Suga	IMI Workshop of the Joint Research Projects Cryptographic Technologies for Securing Network Storage and Their Mathematical Modeling　116pages	March 30, 2018
MI Lecture Note Vol.81	Tsuyoshi Takagi Masato Wakayama Keisuke Tanaka Noboru Kunihiro Kazufumi Kimoto Yasuhiko Ikematsu	IMI Workshop of the Joint Research Projects International Symposium on Mathematics, Quantum Theory, and Cryptography　246pages	September 25, 2019
MI Lecture Note Vol.82	池森　俊文	令和2年度　AIMaP チュートリアル 新型コロナウイルス感染症にかかわる諸問題の数理 145pages	March 22, 2021

シリーズ既刊

Issue	Author／Editor	Title	Published
MI Lecture Note Vol.83	早川健太郎 軸丸　芳揮 横須賀洋平 可香谷　隆 林　　和希 堺　　雄亮	シェル理論・膜理論への微分幾何学からのアプローチと その建築曲面設計への応用　49pages	July 28, 2021
MI Lecture Note Vol.84	Taketoshi Kawabe Yoshihiro Mizoguchi Junichi Kako Masakazu Mukai Yuji Yasui	SICE-JSAE-AIMaP Tutorial Advanced Automotive Control and Mathematics　110pages	December 27, 2021
MI Lecture Note Vol.85	Hiroaki Anada Yasuhiko Ikematsu Koji Nuida Satsuya Ohata Yuntao Wang	IMI Workshop of the Joint Usage Research Projects Exploring Mathematical and Practical Principles of Secure Computation and Secret Sharing　114pages	February 9, 2022
MI Lecture Note Vol.86	濱田　　直希 穴井　宏和 梅田　裕平 千葉　一永 佐藤　寛之 能島　裕介 加葉田雄太朗 一木　俊助 早野　健太 佐伯　　修	2020年度採択分　九州大学マス・フォア・インダストリ研究所 共同利用研究集会 進化計算の数理　135pages	February 22, 2022
MI Lecture Note Vol.87	Osamu Saeki, Ho Tu Bao, Shizuo Kaji, Kenji Kajiwara, Nguyen Ha Nam, Ta Hai Tung, Melanie Roberts, Masato Wakayama, Le Minh Ha, Philip Broadbridge	Proceedings of Forum "Math-for-Industry" 2021 -Mathematics for Digital Economy-　122pages	March 28, 2022

Institute of Mathematics for Industry
Kyushu University

九州大学マス・フォア・インダストリ研究所
九州大学大学院 数理学府