Material designing and analysing based on density functional theory calculation and machine learning

Jie Huang

January 11, 2022

Summary

To further exercise my research ability and give full play to my professional expertise, I'd like to join the research group of Prof. Yongqing Cai in the Department of Applied Physics and Materials Engineering of the University of Macau.

My research interest is material designing and analyzing based on density functional theory (DFT) calculation and machine learning. I was engaged in the intersection between machine learning and condensed matter physics as a graduate student. I have accumulated sufficient experience in applications of machine learning, and DFT calculations in the research process.

Prof. Cai is an expert in computational materials science and computational condensed matter physics. One of his research interests is material informatics where machine learning plays an important role. Therefore, I want to join Prof. Cai's research group to engage in scientific research. During my Ph.D. period, I plan to use DFT calculations to simulate materials, combined with machine learning to design materials and analyze the properties of materials.

Background

The rapid development of big data and machine learning has resulted in new data-driven materials research, which has achieved substantial progress [1]. The amount of data being generated by experiments and simulations has given rise to the fourth paradigm of science: data-driven science [2], and it unifies the first three paradigms of theory, experiment, and computation as shown in Figure 1. This massive amount of data needs to be stored and interpreted in order to advance the materials science field. Identifying correlations and patterns from large amounts of complex data is being performed by machine learning algorithms for decades [3]. It is increasingly becoming popular in the field of materials science as well and has led to the emergence of the new field of materials informatics [4, 5], which aims to discover the relations between known standard features and materials properties.

The following are two topics related to my future research.

1. High-entropy alloys. High-entropy alloys (HEAs) are alloys with five or more principal elements. Due to the distinct design concept, these alloys often exhibit unusual properties. Thus,

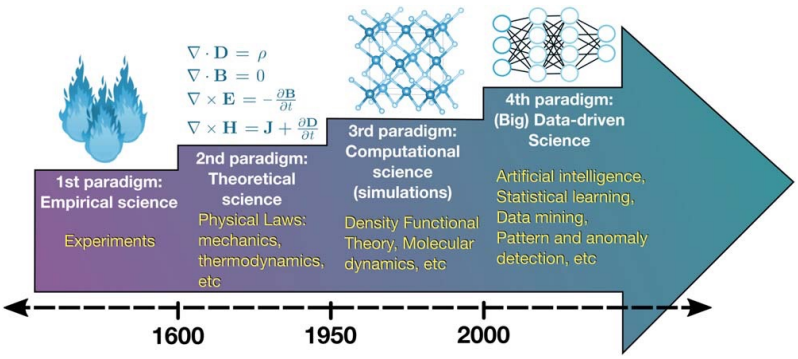

Figure 1: The four science paradigms: empirical, theoretical, computational, and data-driven[3, 5].

there has been significant interest in these materials, leading to an emerging yet exciting new field [6]. Identifying single-phase HEAs is extremely important to understanding HEA formation and their intrinsic properties, but the lack of effective guidelines has hindered their discovery. [7] Besides, the experimental process is still limited by high costs and time-consuming synthesis procedures. [8] The work of Chang et al. [9] utilizes an artificial neural network (ANN) to predict the composition of HEAs in order to achieve the highest hardness in the system. A simulated annealing algorithm is integrated with the ANN to optimize the composition. This work demonstrates that, by applying the machine learning method, new compositions of HEAs can be obtained, exhibiting hardness values higher than the best literature value for the same alloy system.

2. Nanostructures for Phonon Transport: In another work, Ju et al. [10] identified the Si/Ge composite interfacial structures that minimize or maximize the interfacial thermal conductance across Si-Si and Si-Ge interfaces by the developed framework combining the atomistic Green function and Bayesian optimization methods. The optimal structures are obtained by calculating only a few percent of the total candidate structures, considerably saving computational resources. The validity and capability of the method are demonstrated by identifying the thin interfacial structures with the optimal Si/Ge configurations among all the possible candidates. Designing materials with high thermal conductivity is the key to solving the problem of heat dissipation of electronic devices. Electronic devices need materials with good heat dissipation, which requires us to design advanced thermal management materials and devices to solve the bottleneck of heat dissipation of electronic devices in our country. This is also in line with the national strategic development requirements of electronic devices.

With regard to the design of new materials, the space for exploration is huge. Therefore, it is urgent to find a way to speed up the material discovery process. Theoretical and calculation research activities play an increasingly-important role in materials science. The combination of computer simulation and experimental data contributes to a better understanding of the physical mechanism, which therefore enables the prediction of unknown data. The applications mentioned above show that machine learning is feasible for materials science. In the future, with the increase of the amount in material data and the continuous improvement of computer computing power, machine learning will play an increasingly important role in materials science. However, unlike traditional computer fields such as computer vision, machine learning applications in materials science have their characteristics. First of all, there is a huge difference in the amount of data. The size of the data available in the field of materials science ($< 10^3$) is much smaller than that in the field of traditional machine learning ($>10^6$). The reason is that data in the field of materials science usually needs to be obtained through experiments, and the acquisition of these data is expensive. On the other hand, a large part of machine learning in the material field is looking for examples that do not exist in the data set. Therefore, a large part of machine learning in the material field studies how to use the least number of attempts to get the structure or material we desire.

Objectives and Methods

The goal of this research is to use DFT calculations and machine learning to

- aid the design of materials, and discover new materials;
- analyze the simulated data and explore the properties of the materials.

The methods used in this research mainly include the following two aspects.

Method 1: DFT is a mature theory that is currently the undisputed choice of method for electronic structure calculations. A number of papers and reviews are presented in the literature [11, 12, 13], facilitating the widespread of the theory and, thus, the entry of researchers into the field of computational solid-state physics, materials science, and quantum chemistry. On the one hand, DFT calculation can solve the problem of insufficient data in materials science to a certain extent. On the other hand, I am familiar with the DFT simulations. Therefore, in my research, I will use the DFT method to simulate materials, such as two-dimensional materials [14, 15, 16], which is one topic of Prof. Cai's projects, study its phonon characteristics and thermal conductivity.

Method 2: As for machine learning, combined with previous experience, the tools I may use include ANN, Monte Carlo tree search, Bayesian optimization. Reinforcement learning, as a method of providing optimal strategies, is promising in the material designing. So I plan to explore the

methods of reinforcement learning in material designing. Thanks to many data analysis software nowadays, I will choose the Python that I am good at combined with the widely used machine learning software library, like Tensorflow, etc. to build machine learning models and analyze data.

References

[1] Jianxin Xie, Yanjing Su, Dezhen Xue, Xue Jiang, Huadong Fu, and Haiyou Huang. Machine learning for materials research and development. *Acta Metall Sin*, 57(11):1343, 2021.

[2] Tony Hey, Stewart Tansley, and Kristin Tolle. *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Microsoft Research, October 2009.

[3] Gabriel R Schleder, Antonio C M Padilha, Carlos Mera Acosta, Marcio Costa, and Adalberto Fazzio. From DFT to machine learning: recent approaches to materials science–a review. *Journal of Physics: Materials*, 2(3):032001, May 2019.

[4] Krishna Rajan. Materials informatics. *Materials Today*, 8(10):38–45, October 2005.

[5] Ankit Agrawal and Alok Choudhary. Perspective: Materials informatics and big data: Realization of the "fourth paradigm" of science in materials science. *APL Materials*, 4(5):053208, April 2016.

[6] Ming-Hung Tsai and Jien-Wei Yeh. High-entropy alloys: A critical review. *Materials Research Letters*, 2(3):107–123, April 2014.

[7] Michael C. Gao, Peter K. Liaw, Jien Wei Yeh, and Yong Zhang. *High-entropy alloys: Fundamentals and applications*. 2016.

[8] Ruixuan Li, Lu Xie, William Yi Wang, Peter K. Liaw, and Yong Zhang. High-throughput calculations for high-entropy alloys: A brief review. *Frontiers in Materials*, 7, September 2020.

[9] Yao-Jen Chang, Chia-Yung Jui, Wen-Jay Lee, and An-Chou Yeh. Prediction of the composition and hardness of high-entropy alloys by machine learning. *JOM*, 71(10):3433–3442, July 2019.

[10] Shenghong Ju, Takuma Shiga, Lei Feng, Zhufeng Hou, Koji Tsuda, and Junichiro Shiomi. Designing nanostructures for phonon transport via bayesian optimization. *Physical Review X*, 7(2), May 2017.

[11] R. O. Jones. Density functional theory: Its origins, rise to prominence, and future. *Reviews of Modern Physics*, 87(3):897–923, August 2015.

[12] Kieron Burke. Perspective on density functional theory. *The Journal of Chemical Physics*, 136(15):150901, April 2012.

[13] John P. Perdew and Adrienn Ruzsinszky. Fourteen easy lessons in density functional theory. *International Journal of Quantum Chemistry*, 110(15):2801–2807, August 2010.

[14] Yongqing Cai, Gang Zhang, and Yong-Wei Zhang. Polarity-reversed robust carrier mobility in monolayer MoS2 nanoribbons. *Journal of the American Chemical Society*, 136(17):6269–6275, April 2014.

[15] Yongqing Cai, Gang Zhang, and Yong-Wei Zhang. Electronic properties of phosphorene/graphene and phosphorene/hexagonal boron nitride heterostructures. *The Journal of Physical Chemistry C*, 119(24):13929–13936, June 2015.

[16] Jiaren Yuan, Yuanping Chen, Yuee Xie, Xiaoyu Zhang, Dewei Rao, Yandong Guo, Xiaohong Yan, Yuan Ping Feng, and Yongqing Cai. Squeezed metallic droplet with tunable kubo gap and charge injection in transition metal dichalcogenides. *Proceedings of the National Academy of Sciences*, 117(12):6362–6369, March 2020.