Proceedings of CONF-MCEE 2026 Symposium: Geomaterials and Environmental Engineering
DOI: 10.54254/2755-2721/2026.MH30032

From Quantum to Data-Driven Models: The Role of Computational Methods in Battery Design

Jiahua Kang

Faculty of Mathematical & Physical Sciences, University College London, London, United Kingdom
18659225738@163.com

Abstract. Battery systems are important to humanity's daily life; among them, the lithium-ion battery is widely used in electric vehicles and portable electronics. The pursuit of better battery materials must remain ongoing to meet the growing demands of future society's needs. Computational study, unlike a traditional experimental study, is based on a theoretical study that relies on the prediction of material behavior and atomic-scale structural analysis through computational simulation. Recent advances in first-principles calculations and machine learning will bring more opportunities to the battery research field. This review focuses on the density functional theory and machine learning, discussing their recent advances, limitations, and future outlook on battery science, and analyzing their position on material development. In conclusion, the first principle calculation has significantly improved the efficiency for the new material development by analyzing their behavior theoretically, saving both time and cost. By integrating machine learning with the first principle simulation, it will significantly advance the progress of battery research and pave the way for future data-driven methodologies, such as neural networks that can address complex systems more efficiently, and quantum computing.

Keywords: computational study, lithium-ion battery, machine learning, density function theory

1. Introduction

In the modern era, the growing energy demand from the increase in the global population has become a challenge to human society. To meet the increasing global demand for efficient energy, various energy storage technologies have experienced rapid advancements in recent years. Besides the energy requirement from population growth, there is another problem, which is global warming and environmental change. The development of better and efficient energy storage systems helps to create a carbon-neutral environment in the future [1].

One of the most common systems is lithium-ion batteries. The lithium-ion battery is famous for its widespread use in electric vehicles. Apart from the electric vehicle application, lithium-ion batteries have also been used in portable devices like personal phones and laptops. There are also other battery systems, such as nickel-cadmium, nickel-metal hydride, and lead-acid [1].

Ongoing research continues to focus on improving those energy storage systems, as enhancing their capacity, efficiency, and overall performance is essential for future society. In the past, the new

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.MH30032

potential electrode material and the enhanced technique were studied traditionally inside the experimental laboratory. Due to its reliance on a trial-and-error methodology that is strongly affected by the researcher's prior experience, this technique has proven to be inefficient and cost-intensive [1]. However, in the modern era, the computational capabilities of modern society have brought a revolution to material science, where the computer simulation tool plays a vital role in the material design [1]. This review has discussed and reviewed the recent development of the first-principle calculation and machine learning on battery science, and analyzed their potential impact on future research if combined. The atomic simulation relies on first-principle calculations in theoretical laws of physics that provide insight, non-related to experimental data, will improve the understanding of the material in real-world behavior and provide guidelines for experimental work [2].

2. Background of computational modeling and machine learning on battery science

In the modern era, the development of a completely new material with computer assistance has advanced rapidly. This is thanks to the automated high-throughput calculations, which allow the fast screening of the properties of thousands of target materials [3]. In material science, the most successful and commonly applied functions are density functional theory (DFT), which was developed by Kohn and coworkers [4,5]. In the development of DFT, there are many other quantum mechanical theories for computational study that have established the basis for it. Like Schrödinger's Equation (SE), Heisenberg's uncertainty principle, and Pauli's exclusion principle. Those methods are essential for the development of modern simulation methods like Hartree–Fock and DFT [6].

The application of the density function theory in the field of batteries was aimed at explaining the properties of current materials, and also to predict the behavior and electronic properties of the new material theoretically. First-principles simulation help to understand the electronic density of states distribution, the ion migration energy barriers, pathways, structural stability, vibrational heat capacity, surface point defects, magnetic properties, and elastic properties of the target materials. It's been used in all three components in a battery, which are the cathode, anode, and electrolyte [7].

However, the first principal calculation based on density function theory is not the only computer-related technique that has been used in battery material investigations. Machine learning has also been widely applied in battery research. It uses algorithms and statistical models that let computers learn from data, detect patterns, and make predictions or decisions automatically [1]. In 2011, the Materials Genome Initiative (MGI) was developed and introduced to the public. It aimed to accelerate the discovery of new materials. MGI integrates experimental methods with computational techniques and theoretical insights, facilitating a more efficient development of new materials. This method has brought a significant impact on energy storage materials and the catalyst field. Allowing for a more fundamental breakthrough for the materials and creating a material database. The development of large-scale material databases has significantly advanced battery research, particularly by allowing data-driven insights. This is accelerating material discovery significantly through machine learning [8].

3. Computational study approaches

3.1. Density functional theory

Kohn-Sham density functional theory (DFT) provides a practical estimate for the electron density and ground-state energy for a many-electron material system, such as an atom, molecule, or solid. In

DOI: 10.54254/2755-2721/2026.MH30032

a simulation, it can predict the molecular structure of the target substance, like shape and size [9].

In all ab initio method simulations, it assumes that electrons and nuclei are basic particles that are not externally disturbed, and the physical interaction between different particles is studied using quantum mechanics, because the formulation of the event requires the use of complex partial differential equations [6]. But instead of using many body-wave functions, DFT considers electron density as the main variable, which allows it to simulate electronic and magnetic properties [6]. DFT relies on a series of approximations to achieve computationally practical calculations. Those approximations speed up the calculations, dramatically improve efficiency, but they also introduce limitations on accuracy [6].

The basis approximation that DFT relies on is the exchange–correlation functional, which is responsible for most of the approximations of DFT simulations. Exchange–correlation functionals are typically organized into Perdew's "Jacob's Ladder," a conceptual framework in which each rung corresponds to a higher level of approximation, covering a range of functionals, including the Local Density Approximation (LDA), Generalized Gradient Approximation (GGA), meta-GGA, and hybrid forms [9]. It is based on simplified energy expressions, which will cause a mismatch to the real material system. However, its efficiency and capability are still beyond the capability of non-interacting computational theories, thus it remains an effective approach for estimating the total energy of a material's system [6]. The mean-field approximation and the One-electron model have been used with the exchange–correlation functional to estimate the electronic potential of the target material. The one-electron Schro¨ dinger equation model visualizes each orbital of the material electron as a wave that has its own different amplitude and phase in which match with its space and position [6]. One-electron model helps to improve the accuracy of the electron potential simulation based on the mean-field approximation, especially for complex systems.

Despite DFT being a practical tool to enhance the efficiency of the development of battery material. But it still has limitations. To save time and cost, DFT's approximation simplifies the calculation, but this simplified results in inaccuracy in the realistic situation. DFT's limitation will arise with the complex material system, where there is strong electronic interaction. The additional models, like the Car–Perriello method and the 1-RDM model, need to be applied to overcome those limitations [6].

3.2. Machine learning approach

Machine learning has been used to improve the quality of material microscope data and extract information from modeling or experimental data [8]. The process of using machine learning in material development involves several major steps. Machine learning begins by gathering and preparing the given data for analysis. The data is obtained and cleared by the terminal, and then the data will be ready to be used in further model training. After data preparation, a suitable machine learning algorithm is selected depending on the target characteristic, and the model is subsequently trained using the prepared dataset [10]. During training, the algorithm repeatedly updates its parameters through iterative optimization to reduce the error between predicted and actual values. During the process, several safeguard methods, including regularization and cross-validation, were implemented during training to prevent overfitting and improve the model's performance on unseen data [1]. There are many model training methods, like unsupervised learning, supervised learning, semi-supervised learning, and reinforcement learning [1]. Within the field of materials research, the selection of an appropriate training method is governed by the nature of the problem, the specific research goals and the availability of labeled data [1]. After training is complete, the model's

DOI: 10.54254/2755-2721/2026.MH30032

performance will be assessed using appropriate evaluation metrics. Once its performance criteria are satisfied, the model is deployed for practical implementation [1].

For the battery material development, machine learning has brought significant advancement, especially in the field of Lithium-ion batteries. It's been applied in the development of high-performance electrolytes and electrodes. Machine learning has been able to predict the properties of target materials with a quicker speed and accuracy [1]. In recent years, the development of machine learning in material electrodes has greatly advanced. In 2023, Yu et al introduced the multi-output ensemble deep learning MOEDL framework. This technique integrates deep neural networks (DNNs) for multi-output learning, incorporating Bayesian optimization (BO) for efficient hyperparameter tuning and an attention mechanism (AM) to identify and prioritize significant input features. Additionally, a deep belief network (DBN) has been developed to aid training efficiency of the model and mitigate the risk of suboptimal solutions. The primary contribution of this study lies in the integration of four enhanced models—Deep Neural Networks (DNN), Bayesian Optimization (BO), Attention Mechanism (AM), and Deep Belief Networks (DBN)—within the MOEDL framework, resulting in an R^2 value of 0.93. The capabilities of the MOEDL framework extend beyond merely predicting material properties. The framework shows versatility throughout the entire battery lifecycle, with applications spanning battery management, control, and optimization fields [11]. Other notable contributions to the field come from Duquesnoy et al, which introduce a new novel machine learning–assisted framework designed for the multi-objective optimization of lithium-ion battery (LIB) electrode properties, and for the inverse engineering of their manufacturing processes. The author of this work prepared a synthetic dataset from simulation work and used it to train reproducible machine learning models. Those models allow a rapid and more cost-effective optimization loop for the machine learning process, and concurrently minimize the electrode tortuosity and enhance properties critical to Li-ion (de)intercalation kinetics, which is important for high-performance batteries [12].

4. Current challenge and future perspective of computational learning

Despite their capabilities, there are still challenges and limitations associated with both methods. For DFT, the simulation and determination of materials behavior only consider three dimensions, because the fundamental principle of DFT is based on the fact that, in a fixed arrangement of nuclei, the total energy density is determined as a functional of the three-dimensional electron density. This means that DFT can solve a large material system pretty well compared to wave function theory (WFT). However, it is impossible to solve the Schrödinger equation for a many-body system accurately, even for DFT; the inaccuracy grows with the system size. In reality, much research that focused on DFT involves adding additional ingredients to the spin densities to correct the error of the simulations [13]. Also, the DFT simulation for a large and complex system is very time-consuming, which is also a limitation for its use on complex material structures [6].

From the machine learning perspective, computers equipped with machine learning tools are better on average than human researchers at many tasks, such as image recognition. However, there are still errors that occur in some cases, which require human investigators to check and fix them. Human researchers are required to minimize the errors of the machine learning and also perform a sanity check for the machine learning/AI tool. The limitations of machine learning at the current stage are that they cannot be efficient without human supervision, and the human instructor requires certain levels of knowledge to identify the errors of the prediction [14].

Despite the limitations of both methods, the computational study, combined with a traditional lab-based experiment, shows great potential for revolutionizing material development efficiency. The

DOI: 10.54254/2755-2721/2026.MH30032

DFT limitation to more complex or larger material systems can be overcome by developing more advanced principles. The future focus of research on improving the DFT should focus on creating a more efficient DFT function to enable the prediction of the material's thermodynamic properties [7]. The integration of machine learning into DFT simulations allows the development of a unified framework that includes data generation, acquisition, and analysis [7]. Such integration will substantially improve the efficiency of new material development for batteries, and its applicability may extend well beyond this domain. From the Machine learning perspective, current machine learning applications mainly rely on the human-machine combination [14]. But in the future, this might change to humans and machines working separately. From the prediction made by D. Morgan, the machine learning technique is still in their beginning stage; the development of machine learning will be quick, and their influence will grow rapidly through the building of the machine learning ecosystem in the field.

5. Conclusion

Overall, recent advances in computational studies have opened a new path for transformative breakthroughs in the discovery, design, and deployment of new materials that impact a wide range of modern technologies, like batteries. The continuous advancement of computational techniques and tools will benefit the progress of designing more efficient and safer energy storage systems, like lithium-ion batteries. The computational insight into material first-principles behavior will guide the following experimental work, and computational science can investigate the problem that is important to lithium-ion battery performance, such as the ions diffusion kinetics within electrode materials, the structural evolution of electrode materials during cycling, the relationship between potential and structure, and the mechanisms behind the formation of the solid electrolyte interphase (SEI). More importantly, the combination of machine learning and DFT simulations is expected to significantly accelerate progress in battery research, paving the way for future data-driven methodologies such as neural networks to address complex systems and quantum computing. Still, this review does not analyze the challenges associated with integrating first-principles calculations and machine learning; future research should prioritize developing feasible approaches to effectively combine these two methodologies. Meanwhile, continuous improvement of current methodologies and the development of more effective models remain crucial for the future progress of battery science to achieve a more sustainable and green future for human society.

References

[1] P. C.V., A. K. Lakshminarayanan, B. Ramasubramanian, and S. Ramakrishna, "Progress of machine learning in materials design for Li-Ion battery, " Next Mater., vol. 2, no. February, p. 100145, 2024, doi: 10.1016/j.nxmate.2024.100145.

[2] A. Urban, D. H. Seo, and G. Ceder, "Computational understanding of Li-ion batteries, " npj Comput. Mater., vol. 2, no. January, 2016, doi: 10.1038/npjcompumats.2016.2.

[3] S. Curtarolo, G. L. W. Hart, M. B. Nardelli, N. Mingo, S. Sanvito, and O. Levy, "The high-throughput highway to computational materials design, " Nat. Mater., vol. 12, no. 3, pp. 191–201, 2013, doi: 10.1038/nmat3568.

[4] C. H. Chu and C. W. Leung, "The convolution equation of Choquet and Deny on [IN]-groups, " Integr. Equations Oper. Theory, vol. 40, no. 4, pp. 391–402, 2001, doi: 10.1007/BF01198136.

[5] M. C. Koch, W. and Holthausen, A Chemist's Guide to Density Functional Theory. Weinheim, Germany: Wiley-VCH Verlag GmbH, 2001.

[6] H. Heryanto, A. Ardiansyah, R. Rahmat, and D. Tahir, "Science Mapping Analysis of Density Functional Theory (DFT) for Material Design: A Review, " Jom, vol. 76, no. 8, pp. 4629–4642, 2024, doi: 10.1007/s11837-024-06644-w.

DOI: 10.54254/2755-2721/2026.MH30032

[7] H. Shen, Z. Li, Y. Shang, L. Wang, H. Jiang, and S. Lei, “Density functional theory in Li-ion battery electrodes: Bridging atomic structure and performance optimization, ” Mater. Today Commun., vol. 48, no. May, p. 113326, 2025, doi: 10.1016/j.mtcomm.2025.113326.
[8] Y. Liu et al., “Machine learning in materials genome initiative: A review, ” J. Mater. Sci. Technol., vol. 57, pp. 113–122, 2020, doi: 10.1016/j.jmst.2020.01.067.
[9] J. P. P. K. Schmidt, “Jacob’s Ladder of Density Functional for the Exchange-Correlation Energy, ” WIREs Comput. Mol. Sci., vol. 577, no. 1, pp. 1–20, 2001.
[10] S. Badillo et al., “An Introduction to Machine Learning, ” Clin. Pharmacol. Ther., vol. 107, no. 4, pp. 871–885, 2020, doi: 10.1002/cpt.1796.
[11] H. Yu et al., “Multi-output ensemble deep learning: A framework for simultaneous prediction of multiple electrode material properties, ” Chem. Eng. J., vol. 475, no. July, p. 146280, 2023, doi: 10.1016/j.cej.2023.146280.
[12] M. Duquesnoy, C. Liu, D. Z. Dominguez, V. Kumar, E. Ayerbe, and A. A. Franco, “Machine learning-assisted multi-objective optimization of battery manufacturing from synthetic data generated by physics-based simulations, ” Energy Storage Mater., vol. 56, no. September 2022, pp. 50–61, 2023, doi: 10.1016/j.ensm.2022.12.040.
[13] P. Verma and D. G. Truhlar, “Status and Challenges of Density Functional Theory, ” Trends Chem., vol. 2, no. 4, pp. 302–318, 2020, doi: 10.1016/j.trechm.2020.02.005.
[14] D. Morgan and R. Jacobs, “Opportunities and Challenges for Machine Learning in Materials Science, ” Annu. Rev. Mater. Res., vol. 50, pp. 71–103, 2020, doi: 10.1146/annurev-matsci-070218-010015.