Computational investigation and generation of site-disordered sodium ion cathode materials

Martin Hoffmann Petersen [1,2] **Chiku Parida** [1] **Diptendu Roy** [1]
Arghya Bhowmik [1] **Juan Maria Garcia Lastra** [1]

[1]*Technical University of Denmark, Department of Energy Conversion and Storage, Lyngby, Denmark 2800* [2]*School of Materials Science and Engineering, Nanyang Technological University, Singapore 639798*. Correspondence to: Martin Hoffmann Petersen mahpe@dtu.dk.

1. Introduction

Sodium-ion batteries (SIBs) are a promising alternative to lithium-ion technology due to the abundance and low cost of sodium; however, their performance still lags behind that of lithium-based systems[1]. Improving SIBs requires the discovery of cathode materials that combine high capacity, stability, and ionic conductivity[2]. Among these, polyanionic sodium cathodes are particularly attractive for their high structural stability and voltage[3]. Nevertheless, their limited ionic conductivity remains a major bottleneck. One strategy to overcome this involves doping transition metal (TM) sites, introducing controlled atomic site-disorder that can enhance ionic transport and structural resilience.

Atomic site-disorder is a ubiquitous feature in functional materials, often introduced intentionally through doping to enhance performance and stability[4, 5]. In SIB cathodes, TM ion substitution can improve capacity retention, ionic transport, and cycling stability; however, the underlying mechanisms remain poorly understood[6]. This knowledge gap largely arises from the difficulty of accurately modeling disordered atomic configurations using conventional computational approaches. In this study, we aim to identify optimal atomic configurations of polyanion sodium cathode materials using a range of computational methods, as outlined in Figure 1.

We aim to identify optimal configurations of polyanion sodium cathode materials, with the end goal of using generative models for inverse design of sodium-ion cathode materials. Achieving this goal requires the development of an effective database of first-principles calculations, which can be used to train machine learning interatomic potentials (MLIPs) capable of efficiently and accurately describing various site-disordered phases within the polyanion sodium cathode phase space. Once the limitations of these MLIPs have been thoroughly assessed, the final aim is to design a novel generative model for generating site-disordered cathode materials with targeted compositions and functional properties.

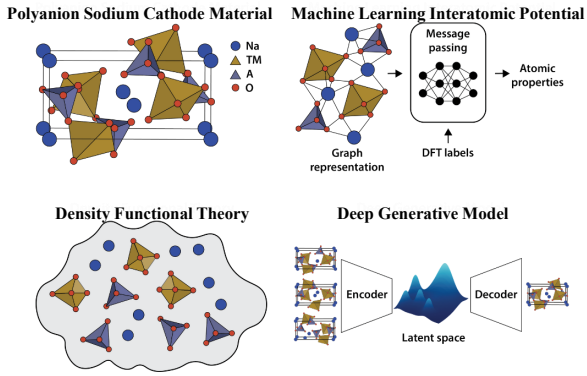

Fig. 1: Overview of all overall methods used for the research. The polyanion sodium cathodes materials, with transition metal (TM) ions and anions (A) are investigated, optimized and discovered using density functional theory calculations, machine learning interatomic potentials and deep generative models.

2. Main results

2.1 *The polyanion sodium cathode materials database*

In recent era, MLIPs are emerged as the cost-effective methods compared to the enumerative search in a huge chemical space of materials using first-principle methods. But the performance of MLIPs towards a specific domain of chemical space depend on the system-specific data. In order to search the structures from the out of distribution chemical space, active learning method is proven to be effective[7]. For polyanion cathode materials, the search explores all possible configurations arising from different levels of sodiation and various TM dopants (TM $\in$ {Fe, Mn, Co, Ni}) involving four polyanion sodium-ion cathode systems: $NaTMPO_4$ (olivine), $NaTMPO_4$ (maricite), Na_2TMSiO_4, and $Na_{2.56}TM_{1.72}(SO_4)_3$. This polyanion sodium cathode materials dataset[8] is collected using single point DFT calculations and 1000 K ab initio molecular dynamic simulations for iterative active learning search. By including both single- and multi-TM compositions, the dataset provides a diverse representation of the polyanion sodium cathode material phase space. Beyond its immediate use for direct analysis, this dataset establishes a solid foundation for training MLIPs, enabling computationally efficient modeling of both static and dynamic properties of polyanion sodium cathode materials.

2.2 *Machine learning interatomic potentials*

We then used this database to evaluate the state-of-the-art universal MLIPs such as CHGNet[9], MACE[10, 11] and M3GNET[12, 13]) and system-specific MLIPs trained on the polyanion sodium cathode materials dataset. To assess their suitability for cathode

material investigations, three key simulation tasks were considered: structural optimization, MD simulations, and nudged elastic band (NEB) calculations. The study showed that the system-specific charge-PaiNN (cPaiNN) model[14] performed equal to, or better than, the universal MLIPs across all tasks. A key advantage of cPaiNN is its explicit descriptor of atomic charges, making it well suited for cathode materials with changing oxidation states of TM ions during battery cycling. By employing cPaiNN to predict the physical and chemical properties of cathode materials, we achieve both computational efficiency and accuracy comparable to first-principles methods, while also gaining interpretability into the underlying redox process.

2.3 Generative models

To enable inverse design of promising cathode materials, we developed a generative modeling framework capable of producing site-disordered crystal structures. The Dis-GEN model[15] generates symmetry-consistent structures that incorporate both compositional disorder and vacancies while preserving crystallographic symmetry. This capability enables systematic exploration of site-disordered functional materials beyond the configurational space accessible through conventional enumeration approaches.

In the context of polyanionic sodium cathode materials, Dis-GEN can be applied to identify promising dopant configurations both at the TM and A sites, with their physical and chemical properties subsequently evaluated using MLIPs. While Dis-GEN significantly expands the accessible design space for disordered cathode materials, it currently does not explicitly account for thermodynamic stability across different charge states. Nevertheless, the model provides a powerful framework for exploring optimal doping strategies and discovering novel disordered cathode structures, thereby enabling a systematic search for high-performance polyanionic sodium cathode materials.

2.4 The search for promising polyanion sodium cathode materials

While recent advances in generative modeling have shown strong potential for materials discovery, only a limited number of experimentally realized functional materials have emerged so far[16, 17]. This is partly because most workflows rely on general post-processing criteria, commonly summarized as the stability–uniqueness–novelty (S.U.N.) framework[18, 19, 20], which focuses on thermodynamic plausibility and structural novelty but does not explicitly account for application-specific performance requirements relevant to battery materials.

To address this, we introduce an integrated battery discovery pipeline, BPline, which combines generative structure design with physics-based screening tailored for battery applications. Beyond standard S.U.N. filtering, BPline includes stability metrics such as energy above the convex hull and phonon stability, together with bond valence site-energy (BVSE)[21, 22] analysis to estimate ionic conductivity and enforce charge neutrality in systems with multiple oxidation states. The pipeline also enables classification of candidates as cathodes or electrolytes and supports evaluation of charge–discharge states for assessing redox behavior.

Overall, combining system-specific MLIPs trained with our own generated dataset, Dis-GEN as efficient generative model for disordered Na-cathode materials and BPline as a overall estimation workflow for stability and efficiency of the SIB cathode materials, offers a coherent, accelerated SIB discovery method while providing broader insights into the design of site-disordered functional materials.

Acknowledgments

The authors acknowledge support from the Det Frie Forskningsråd under Project "Data-driven quest for TWh scalable Na-ion battery (TeraBatt)" (Ref. Number 2035-00232B). All authors acknowledges support from the Novo Nordisk Foundation Data Science Research Infrastructure 2022 Grant: A high-performance computing infrastructure for data-driven research on sustainable energy materials, Grant no. NNF22OC0078009.

References

[1] Hayley S. Hirsh, Yixuan Li, Darren H. S. Tan, Minghao Zhang, Enyue Zhao, and Y. Shirley Meng. Sodium-ion batteries paving the way for grid energy storage. *Advanced Energy Materials*, 10(32), July 2020.

[2] Mingzhe Chen, Qiannan Liu, Shi-Wen Wang, Enhui Wang, Xiaodong Guo, and Shu-Lei Chou. High-abundance and low-cost metal-based cathode materials for sodium-ion batteries: Problems, progress, and key technologies. *Advanced Energy Materials*, 9(14), February 2019.

[3] Ting Jin, Huangxu Li, Kunjie Zhu, Peng-Fei Wang, Pei Liu, and Lifang Jiao. Polyanion-type cathode materials for sodium-ion batteries. *Chemical Society Reviews*, 49(8):2342–2377, 2020.

[4] Zhaoyang Wang, Zijuan Du, Luoqing Wang, Guanjie He, Ivan P. Parkin, Yanfei Zhang, and Yuanzheng Yue. Disordered materials for high-performance lithium-ion batteries: A review. *Nano Energy*, 121:109250, March 2024.

[5] Cai Ning, Qun Ji, Yilei Wu, Jinlan Wang, and Ming-Gang Ju. Disorder on mixed cation halide perovskite for photovoltaic applications. *The Journal of Physical Chemistry Letters*, 14(36):8034–8042, August 2023.

[6] Chunliu Xu, Junmei Zhao, Chao Yang, and Yong-Sheng Hu. Polyanionic cathode materials for

practical na-ion batteries toward high energy density and long cycle life. *ACS Central Science*, 9(9):1721–1736, September 2023.

[7] Xin Yang, Martin Hoffmann Petersen, Renata Sechi, William Sandholt Hansen, Sam Walton Norwood, Yogeshwaran Krishnan, Smobin Vincent, Jonas Busk, Francois Raymond J Cornet, Ole Winther, Juan Maria Garcia Lastra, Tejs Vegge, Heine Anton Hansen, and Arghya Bhowmik. Curator: Building robust machine learning potentials for atomistic simulations autonomously with batch active learning. *chemrxiv*, February 2024.

[8] Martin Hoffmann Petersen, Jin Hyun Chang, Arghya Bhowmik, and Juan Maria García Lastra. Dataset exploring the atomic scale structure and ionic dynamics of polyanion sodium cathode materials. *Scientific Data*, 12(1), August 2025.

[9] Bowen Deng, Peichen Zhong, KyuJung Jun, Janosh Riebesell, Kevin Han, Christopher J. Bartel, and Gerbrand Ceder. Chgnet as a pretrained universal neural network potential for charge-informed atomistic modelling. *Nature Machine Intelligence*, 5(9):1031–1041, September 2023.

[10] Ilyes Batatia, David P Kovacs, Gregor Simm, Christoph Ortner, and Gábor Csányi. MACE: Higher order equivariant message passing neural networks for fast and accurate force fields. *Advances in Neural Information Processing Systems*, 35:11423–11436, 2022.

[11] Ilyes Batatia, Philipp Benner, Yuan Chiang, Alin M. Elena, Dávid P. Kovács, Janosh Riebesell, Xavier R. Advincula, Mark Asta, Matthew Avaylon, William J. Baldwin, Fabian Berger, Noam Bernstein, Arghya Bhowmik, Samuel M. Blau, Vlad Cărare, James P. Darby, Sandip De, Flaviano Della Pia, Volker L. Deringer, Rokas Elijošius, Zakariya El-Machachi, Fabio Falcioni, Edvin Fako, Andrea C. Ferrari, Annalena Genreith-Schriever, Janine George, Rhys E. A. Goodall, Clare P. Grey, Petr Grigorev, Shuang Han, Will Handley, Hendrik H. Heenen, Kersti Hermansson, Christian Holm, Jad Jaafar, Stephan Hofmann, Konstantin S. Jakob, Hyunwook Jung, Venkat Kapil, Aaron D. Kaplan, Nima Karimitari, James R. Kermode, Namu Kroupa, Jolla Kullgren, Matthew C. Kuner, Domantas Kuryla, Guoda Liepuoniute, Johannes T. Margraf, Ioan-Bogdan Magdău, Angelos Michaelides, J. Harry Moore, Aakash A. Naik, Samuel P. Niblett, Sam Walton Norwood, Niamh O'Neill, Christoph Ortner, Kristin A. Persson, Karsten Reuter, Andrew S. Rosen, Lars L. Schaaf, Christoph Schran, Benjamin X. Shi, Eric Sivonxay, Tamás K. Stenczel, Viktor Svahn, Christopher Sutton, Thomas D. Swinburne, Jules Tilly, Cas van der Oord, Eszter Varga-Umbrich, Tejs Vegge, Martin Vondrák, Yangshuai Wang, William C. Witt, Fabian Zills, and Gábor Csányi. A foundation model for atomistic materials chemistry. *arXiv*, 2024.

[12] Chi Chen and Shyue Ping Ong. A universal graph deep learning interatomic potential for the periodic table. *Nature Computational Science*, 2(11):718–728, November 2022.

[13] Ji Qi, Tsz Wai Ko, Brandon C. Wood, Tuan Anh Pham, and Shyue Ping Ong. Robust training of machine learning interatomic potentials with dimensionality reduction and stratified sampling. *npj Computational Materials*, 10(1), February 2024.

[14] Martin Hoffmann Petersen, Chiku Parida, Jonas Busk, Jin Hyun Chang, Arghya Bhowmik, and Juan Maria Garcia Lastra. Limitations of foundation models in energy materials simulations: A case study in polyanion sodium cathode materials. *Advanced Intelligent Discovery*, February 2026.

[15] Martin Hoffmann Petersen, Ruiming Zhu, Haiwen Dai, Savyasanchi Aggarwal, Nong Wei, Andy Paul Chen, Arghya Bhowmik, Juan Maria Garcia Lastra, and Kedar Hippalgaonkar. Dis-gen: Disordered crystal structure generation, 2025.

[16] Mikkel Juelsholt. Continued challenges in high-throughput materials predictions: Mattergen predicts compounds from the training dataset. June 2025.

[17] Anthony K. Cheetham and Ram Seshadri. Artificial intelligence driving materials discovery? perspective on the article: Scaling deep learning for materials discovery. *Chemistry of Materials*, 36(8):3490–3495, April 2024.

[18] Tian Xie, Xiang Fu, Octavian-Eugen Ganea, Regina Barzilay, and Tommi Jaakkola. Crystal diffusion variational autoencoder for periodic material generation. *ICLR*, 2022.

[19] Nikita Kazeev, Wei Nong, Ignat Romanov, Ruiming Zhu, Andrey Ustyuzhanin, Shuya Yamazaki, and Kedar Hippalgaonkar. Wyckoff transformer: Generation of symmetric crystals. In *Proceedings of the 46nd International Conference on Machine Learning (ICML)*. arXiv, 2025.

[20] Claudio Zeni, Robert Pinsler, Daniel Zügner, Andrew Fowler, Matthew Horton, Xiang Fu, Zilong Wang, Aliaksandra Shysheya, Jonathan Crabbé, Shoko Ueda, Roberto Sordillo, Lixin Sun, Jake Smith, Bichlien Nguyen, Hannes Schulz, Sarah Lewis, Chin-Wei Huang, Ziheng Lu, Yichi Zhou, Han Yang, Hongxia Hao, Jielan Li, Chunlei Yang, Wenjie Li, Ryota Tomioka, and Tian Xie. A generative model for inorganic materials design. *Nature*, 639(8055):624–632, January 2025.

[21] Artem D. Dembitskiy, Dmitry A. Aksyonov, Artem M. Abakumov, and Stanislav S. Fedotov. Nh+-based frameworks as a platform for designing electrodes and solid electrolytes for na-ion batteries: A screening approach. *Solid State Ionics*, 374:115810, January 2022.

[22] Artem D. Dembitskiy, Innokentiy S. Humonen, Roman A. Eremin, Dmitry A. Aksyonov, Stanislav S. Fedotov, and Semen A. Budennyy. Benchmarking machine learning models for predicting lithium ion migration. *npj Computational Materials*, 11(1), May 2025.