Diffusion-Driven Generation of Novel Crystalline Materials with Target Optical Properties

Ivan Kruglov[1], Liudmila Klimova[1], Alexey Arsenin[1], Valentyn Volkov[1], Konstantin Novoselov[2]

[1] *Emerging Technologies Research Center, XPANCEO, Dubai Investment Park First, Dubai, United Arab Emirates*

[2] *Department of Materials Science and Engineering, National University of Singapore, Singapore, 03-09 EA, Singapore*

Correspondence to: Ivan Kruglov ikruglov@xpanceo.com

1. Generative models in materials science

Recent breakthroughs in generative artificial intelligence have fundamentally reshaped the landscape of natural sciences and opened new paradigms for discovery. Materials science stands to benefit particularly strongly from these advances, as progress in photonics, computing, and electronics critically depends on the availability of materials with tailored properties, yet the space of potentially stable crystalline compounds is astronomically large. Over the past decade, graph neural networks have become a cornerstone of materials informatics, enabling accurate prediction of diverse physical properties directly from crystal structure and supporting large-scale screening across existing materials databases. While such approaches have revealed many compounds with exceptional optical characteristics, these materials are rarely “new” in a structural sense - they already exist in repositories, but their functional potential remained hidden. This gap highlights a fundamental limitation of purely predictive pipelines and motivates the development of generative frameworks capable of proposing genuinely novel materials optimized for target optical functionalities.

In this work we present a diffusion-based generative framework that enables both unconstrained exploration of crystal structure space and property-conditioned generation of materials with target optical responses. By combining modular diffusion sampling with state-of-the-art graph neural networks, the framework generates physically consistent and structurally novel candidates optimized for transparency and extreme birefringence. Operating in a closed-loop generation–screening–DFT validation cycle, the developed framework continuously enriches its training data, enabling efficient navigation of data-sparse regions of materials space and accelerating the discovery of optical materials beyond existing databases.

2. Diffusion-driven generation of novel optical materials

To enable property-conditioned generation, we first constructed a high-fidelity database of van der Waals materials with accurately computed optical properties. Van der Waals crystals are of particular interest for advanced photonics due to their pronounced in-plane and out-of-plane optical anisotropy, which underpins a wide range of novel photonic functionalities. However, identifying compounds with extreme birefringence by manual exploration is prohibitively labor-intensive and incompatible with large-scale discovery. We therefore assembled a curated dataset of vdW materials and evaluated their optical responses using first-principles DFT calculations, achieving close agreement with available experimental measurements, thereby providing a reliable foundation for data-driven modeling and conditional generation [1-3].

For the rapid evaluation of newly generated structures, we developed a suite of physics-informed machine learning models for key optical properties, including birefringence [4], refractive index [5], and second-harmonic generation (SHG) [6]. These models are based on state-of-the-art graph neural network architectures and are trained on high-quality DFT datasets with experimental validation, achieving near *ab initio* accuracy across diverse material classes. In particular, we employ geometry-aware networks for predicting optical anisotropy in van der Waals materials, Cauchy-model–based ML for near-IR refractive indices of 2D systems, and symmetry-aware equivariant networks (OptiXNet) for accurate SHG tensor prediction with built-in uncertainty quantification.

Building on these datasets and models, we developed a diffusion-based framework for *de novo* generation of crystalline materials with targeted optical responses (Figure 1). Starting from a large dataset of near-stable structures, we first train an unconditional generator to capture stability and crystal-chemical regularities, and then fine-tune it for property-conditioned sampling toward the transparent, strongly birefringent regime. Generated candidates are rapidly filtered using our GNN-based optical models and subsequently validated with first-principles DFT calculations to confirm structural stability and optical tensors. Using this pipeline, we identify a set of previously unreported high-birefringence candidates, including $CaMoN_2$, and verify their optical anisotropy at the DFT level, demonstrating that diffusion-based inverse design can move beyond database screening toward genuinely novel optical materials.

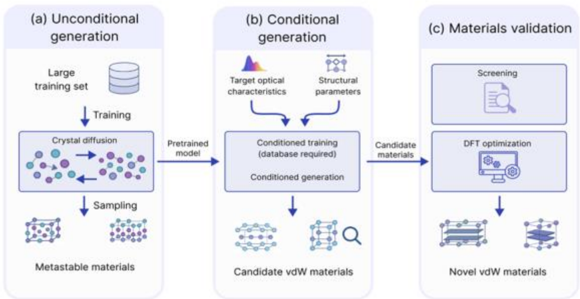

Fig. 1: Three-staged workflow of the developed framework.

References

[1] Ermolaev, G. A., Grudinin, D. V., Stebunov, Y. V., Voronin, K. V., Kravets, V. G., Duan, J., Kruglov I., ... & Volkov, V. S. (2021). Giant optical anisotropy in transition metal dichalcogenides for next-generation photonics. Nature communications, 12(1), 854.

[2] Ermolaev, G. A., Voronin, K. V., Toksumakov, A. N., Grudinin, D. V., Fradkin, I. M., Mazitov, A., Kruglov I., ... & Novoselov, K. S. (2024). Wandering principal optical axes in van der Waals triclinic materials. Nature communications, 15(1), 1552.

[3] Slavich, A. S., Ermolaev, G. A., Tatmyshevskiy, M. K., Toksumakov, A. N., Matveeva, O. G., Grudinin, D. V., Kruglov I., ... & Novoselov, K. S. (2024). Exploring van der Waals materials with high anisotropy: geometrical and optical approaches. Light: Science & Applications, 13(1), 68.

[4] Bereznikova, L. A., Kruglov, I. A., Ermolaev, G. A., Trofimov, I., Xie, C., Mazitov, A., ... & Novoselov, K. S. (2025). Artificial intelligence guided search for van der Waals materials with high optical anisotropy. Materials Horizons.

[5] Klimova, L. A., Kruglov, I. A., Ermolaev, G. A., Grudinin, D., Arsenin, A., Volkov, V. S., & Novoselov, K. S. (2026). Materials Informatics Framework for Accelerated Discovery of High-Refractive-Index 2D Materials. ACS Nano.

[6] Klimova, L. A., Trofimov, I. S., Jin, W., Song, Q., Arsenin, A., Xie, C., Kruglov I., & Volkov, V. (2025). Symmetry-Aware Equivariant Network for Discovering SHG-Active Materials. Advanced Functional Materials, e23683.