AI-accelerated discovery of high critical temperature superconductors

Xiao-Qi Han[1,2],* Zhenfeng Ouyang[1,2],* Peng-Jie Guo[1,2], Hao Sun[4], Ze-Feng Gao[1,2],† and Zhong-Yi Lu[1,2,3]‡

1. Department of Physics and Beijing Key Laboratory of Opto-electronic Functional Materials & Micro-nano Devices. Renmin University of China, Beijing 100872, China

2. Key Laboratory of Quantum State Construction and Manipulation (Ministry of Education), Renmin University of China, Beijing 100872, China

3. Hefei National Laboratory, Hefei 230088, China and

4. Gaoling School of Artificial Intelligence, Renmin University of China, Beijing, China

(Dated: September 13, 2024)

The discovery of new superconducting materials, particularly those exhibiting high critical temperature (T_c), has been a vibrant area of study within the field of condensed matter physics. Conventional approaches primarily rely on physical intuition to search for potential superconductors within the existing databases. However, the known materials only scratch the surface of the extensive array of possibilities within the realm of materials. Here, we develop an AI search engine that integrates deep model pre-training and fine-tuning techniques, diffusion models, and physics-based approaches (e.g., first-principles electronic structure calculation) for discovery of high-T_c superconductors. Utilizing this AI search engine, we have obtained 74 dynamically stable materials with critical temperatures predicted by the AI model to be $T_c \geq 15$ K based on a very small set of samples. Notably, these materials are not contained in any existing dataset. Furthermore, we analyze trends in our dataset and individual materials including B_4CN_3 and B_5CN_2 whose T_cs are 24.08 K and 15.93 K, respectively. We demonstrate that AI technique can discover a set of new high-T_c superconductors, outline its potential for accelerating discovery of the materials with targeted properties.

Introduction. Superconducting materials have numerous applications in modern society since it was discovered [1], particularly in magnetic resonance imaging [2], fueling advances in nuclear fusion technology [3]. Superconductor-based devices are demonstrating potential for achieving scalable quantum information processors, advanced sensors, and efficient communication systems [4–6]. Many of these devices use conventional Bardeen-Cooper-Schrieffer (BCS) superconductors [7], which demand costly helium-based cooling. Hence, searching superconductors with high superconducting critical temperature (T_c) is vital for propelling technological progress in these dynamic areas.

Over the past decade, substantial advancements have been achieved in searching high-T_c superconductors. For example, a superconducting transition with $T_c \sim 36$ K was experimentally observed in high-pressured Scandium, which is the highest record for elemental superconductors [8]. The discovery of superconductivity in bilayer $La_3Ni_2O_7$ under pressure raises superconducting T_c of nickelates to the liquid-nitrogen temperature zone [9]. And lots of theoretical work predicted superconductivity in hydrides [10–15], where superconductivity in H_3S under pressure was experimentally confirmed [16].

Recently, machine learning-based methods have become increasingly popular in searching potential high-T_c superconductors [17–22]. Wines et al [23]. have employed crystal diffusion variational auto-encoder (CDVAE) [24] to generate data based on the JARVIS-DFT database [25], subsequently employing the atomistic line graph neural network (ALIGNN) [26] for T_c forecasting. Using high-throughput density functional theory (DFT) calculations, 34 dynamically stable 2D superconductors with $T_c \geq 5$ K from over 1000 candidates in the JARVIS-DFT database were identified [18]. Moreover, Choudhary and Garrity [17] leveraged electron-phonon coupling (EPC) calculations, assistanted by deep-learning models for efficient prediction of superconducting properties, to identify 105 conventional superconductors with $T_c \geq 5$ K from a pre-screened set of 1736 materials. While numerous studies have highlighted the application of machine learning in this field, these approaches primarily rely on chemical formulas or searches based on the existing datasets. They often lack the intricate atomic structure details crucial for understanding superconducting behavior and are limited in exploring crystal materials beyond known databases. To truly advance the discovery of new superconductors, it is essential to incorporate detailed structural information and broaden the scope beyond existing data. So far, the conventional methods (such as elemental substitution or physical insight) have limited success in finding new high-T_c superconductors among the existing data. The rise of Al technology brings a transformative approach, potentially reshaping our path to solving this challenge.

In this work, we developed an AI search engine to explore high-T_c BCS superconductors, integrating diffusion model, formation energy prediction model, ALIGNN, pre-training and fine-tuning technique, atom docking based on pre-trained model, active learning technique, and physics-based methods (*e.g.*, first-principles electronic structure calculations), and meanwhile sufficiently incorporating detailed structural information. Leveraging a limited dataset of high-T_c BCS superconduc-

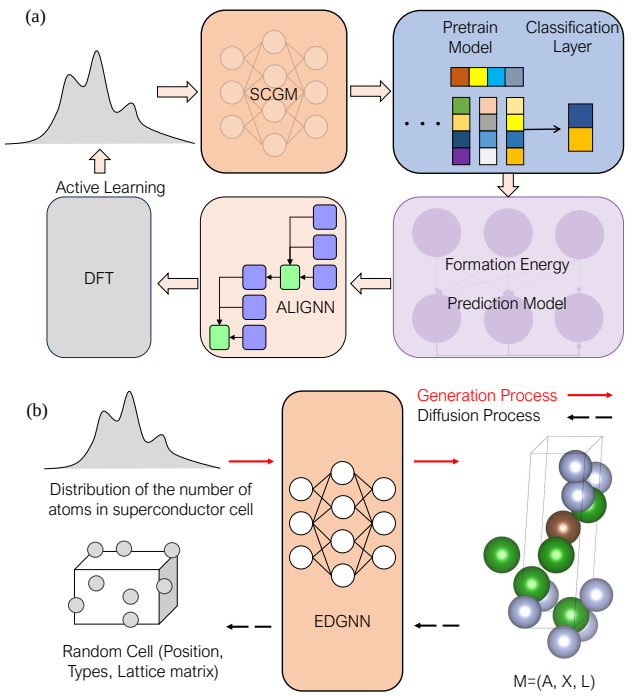

FIG. 1. (a). The AI search engine workflow overview. AI-accelerated discovery of high-T_c superconductors includes generative model for predicting crystal structures, pre-trained model for superconductivity classification, formation energy prediction model, screening model for superconducting transition temperature prediction, and validation using DFT calculation. (b). Symmetry-constrained crystal generation model. The generation of superconducting crystals defines two Markov processes: the black arrows represent the gradual addition of noise to a BCS superconducting crystal, resulting in a random unit cell, while the red arrows indicate the gradual denoising from a prior atomic distribution to generate the original superconducting crystal structure.

tors (105 superconductors with $T_c \geq 5$ K [17]), we have obtained 74 dynamically stable materials exhibiting critical temperatures predicted by the AI model to be $T_c \geq 15$ K. Furthermore, we analyze trends in our results, focusing on specific materials such as B_4CN_3 and B_5CN_2, with T_c of 24.08 K and 15.93 K, respectively. Our AI search engine stands out for its unique capability to obtain crystal structures absent from the existing material databases, effectively pioneering new avenues in the quest for high-T_c superconductors. Its adaptability allows itself to be tailored for a diverse array of functional materials, each with specific desired properties, thereby greatly expanding its utility across the field of materials science.

Overview. Our AI search engine utilizes multiple AI methods and DFT calculations (Fig 1(a)) for generating and screening high-T_c superconductors. Specifically, inspired by the DiffCSP model [27], we constructed a symmetry-constrained superconducting crystal generation model, based on diffusion generative models [28] and equivalent graph neural networks [29]. This model generates new superconducting structures. We also built a superconducting classification model using pre-training techniques [30], graph auto-encoder architectures [31], and optimal transport theory [32] to determine whether or not the generated crystals exhibit superconducting properties. To further assess the stability of the materials, we retrained a formation energy prediction model based on the MEGNET architecture [33] with improvements. Next, we used ALIGNN [26] to predict the superconducting transition temperatures of these materials and screened for high-T_c superconducting candidates. Finally, for the top-5 candidates, We validated our predictions with convergence tests and verified superconducting transition temperatures using DFT calculation. Adopting active learning, we incorporated the discovered superconductors into training set.

Symmetry-constrained crystal generation model. In crystal structures (Fig 1(b)), the atoms exhibit a periodic distribution, with the smallest repeating unit being the unit cell, denoted by $\mathcal{M}$, which can be represented as $\mathcal{M} = (\boldsymbol{A}, \boldsymbol{X}, \boldsymbol{L})$. Here, $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_N] \in \mathbb{R}^{h\times N}$ denotes the atomic types within the unit cell, $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_N] \in \mathbb{R}^{3\times N}$ represents the Cartesian coordinates of each atom, and $\boldsymbol{L} = [\boldsymbol{l}_1, \boldsymbol{l}_2, \boldsymbol{l}_3] \in \mathbb{R}^{3\times 3}$ is the lattice matrix used to describe the periodicity of the crystal. We employed an ab initio crystal generation approach to generate superconducting crystal structures. Specifically, this involves generating a superconducting crystal $\mathcal{M}$ from a given number of atoms N within the unit cell, with a sampling distribution defined as:

$$p(\mathcal{M}, N) = p(N)p(\mathcal{M}|N), \tag{1}$$

where N remains unchanged during the generation process. The distribution $p(N)$ is calculated from the training set, while $p(\mathcal{M}|N)$ is generated based on the model. Standard denoising diffusion probabilistic model (DDPM) [28] can be used to generate $\boldsymbol{L}$ and $\boldsymbol{A}$, and their loss functions take the same form as:

$$\mathcal{L}_{\boldsymbol{L}/\boldsymbol{A}} = \mathbb{E}_{\boldsymbol{\epsilon}\sim\mathcal{N}(0,\boldsymbol{I})}[\|\boldsymbol{\epsilon} - \hat{\boldsymbol{\epsilon}}_{\boldsymbol{L}/\boldsymbol{A}}(\mathcal{M}_t, t)\|_2^2]. \tag{2}$$

The denoising terms $\hat{\boldsymbol{\epsilon}}_{\boldsymbol{L}}(\mathcal{M}_t, t)$ and $\hat{\boldsymbol{\epsilon}}_{\boldsymbol{A}}(\mathcal{M}_t, t)$ are predicted by an equivalent denoising graph neural networks (short as EDGNN Fig 1(b)), and $\mathcal{N}(0, \boldsymbol{I})$ represents a standard normal distribution. Given the periodicity of $\boldsymbol{X}$, it is generated using a score-matching based framework [34]. Details are in the supplementary material (SM). Utilizing 105 BCS superconductors [17], we trained the model to generate novel crystal structures, excluding those in the training set and with overlapping compositions in the Materials Project (MP) database [35]. Since generative models often produce non-ground-state structures, we performed geometry optimization using the neural networks atomic simulation environment [36] and L-BFGS algorithm [37] to refine the generated structures.

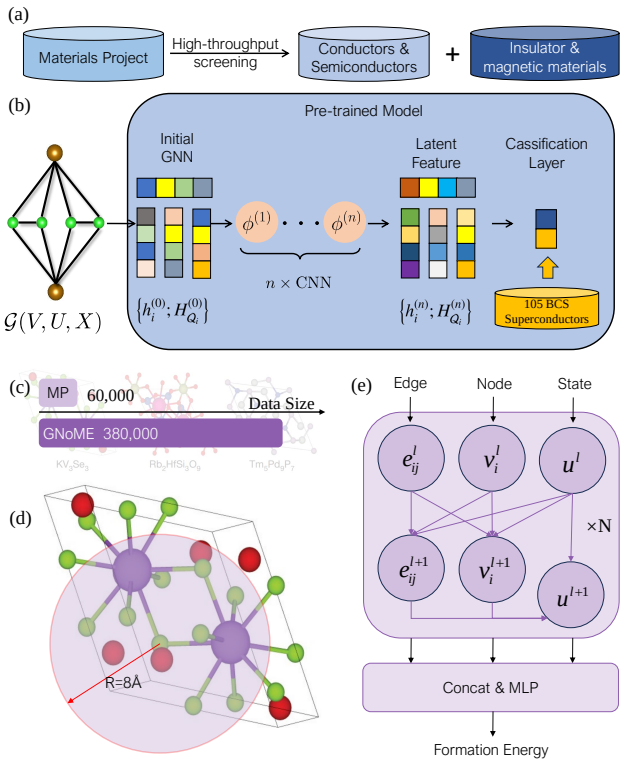

FIG. 2. Superconducting classification model: (a) High-throughput screening process for pre-training and fine-tuning data. (b) Graph auto-encoder architecture based on a graph neural network. Formation energy prediction model: (c) Training data sources. (d) Crystal data representation using atomic graphs with an 8 Å cutoff radius. (e) Interaction of node, edge, and global state representations in the model's architecture.

Superconducting classification model. Initially, we extract 144,595 crystal data entries from the MP database [35]. We first classified the materials into two groups: magnetic and non-magnetic. Subsequently, we refined the non-magnetic category into conductors, semiconductors, and insulators. Then, we designated insulators and magnetic materials as negative samples, and conductors and semiconductors as positive samples, as illustrated in Fig 2**a**. The model is based on a pre-trained graph neural network (GNN) that utilizes material crystal structure information to predict materials [31, 38], consists of a graph convolutional network encoder and a decoder that reconstructs the graph features based on optimal transport theory (see Fig 2**b** and SM). To obtain hidden layer representations related to superconductivity, we pre-trained the model using the positive samples. During the fine-tuning stage, we employed the pre-trained encoder and used up-sampling techniques to balance the number of the BCS superconductors and negative samples for binary classification model. Subsequently, we obtained the classifier model that achieved a discrimination success rate of 99.04% for the 105 BCS superconductors. Utilizing this model , we evaluated the candidate structures generated by the generative model.

Formation energy prediction model. To further assess the stability of potential superconductors, we predict the formation energy of crystals as an indicator of their stability. The AI algorithms like CGCNN [31] and SchNet [39], while fast, lack the required precision for formation energy predictions. Inspired by MEGNET [33], we trained the model using 380,000 crystal structures from GNoME [40] and 60,000 crystal structures [35]. Next, we increased the cutoff radius for constructing atomic graphs from 5Å to 8Å, enabling the model to capture more long-range atomic interactions to more accurately simulate atomic interactions. Recognizing the direct correlation between crystal formation energy and atomic bonding strength, we have incorporated eight new atomic features into our prediction model. This enhancement offers a more comprehensive representation of crystal data, as elaborated in the SM. The original MEGNET benchmark reported a mean absolute error (MAE) of 28 meV per atom, while our improved model achieved an MAE of 21 meV per atom. Since we are particularly interested in high-T_c superconducting materials, we used ALIGNN [26] to predict the superconducting transition temperatures of these materials and applied a 15 K threshold, resulting in top-5 candidate high-T_c superconductors.

Predicted high-T_c materials. By performing the DFT calculations, we studied the electronic structures, phonon properties, and EPC of B_5CN_2 and B_4CN_3 (See SM for crystal structures and additional results.). In Fig. 3, we show the band structure of B_5CN_2 and B_4CN_3 (5 GPa). The results of DFT calculations and Wannier projection show good consistence and suggest that B_5CN_2 and B_4CN_3 (5 GPa) are metallic. The atomic-orbital resolved density of states (DOS) shows that the $2p$ orbitals of B, C, and N atoms mainly contribute the Fermi surfaces.

Next, we investigate the dynamical stability of B_5CN_2 and B_4CN_3. At ambient pressure, we find that B_5CN_2 is dynamically stable, while B_4CN_3 shows a maximum imaginary-frequency phonon of ~ -7.7 meV along the R-Z path. By applying pressure of 5 GPa, the imaginary phonons of B_4CN_3 disappear. Hence, we show the phonon spectrum of B_5CN_2 and B_4CN_3 (5 GPa) in Fig. 4(a) and (c) and further study the EPC of these two materials. The calculated Eliashberg spectral function $\alpha^2F(\omega)$ and accumulated EPC constant $\lambda(\omega)$ are exhibited in Fig. 4(b) and (d). And the mode-resolved $\lambda_{q\nu}$ is added in the phonon spectrum. The EPC constants λ of B_5CN_2 and B_4CN_3 (5 GPa) are integrated to be 0.61 and 0.72, respectively. Using the McMillan-Allen-Dynes formula [41, 42]

$$T_c = \frac{\omega_{log}}{1.2} \exp\left[\frac{-1.04(1+\lambda)}{\lambda(1-0.62\mu^*)-\mu^*}\right], \quad (3)$$

the superconducting T_c of B_5CN_2 and B_4CN_3 (5 GPa) is

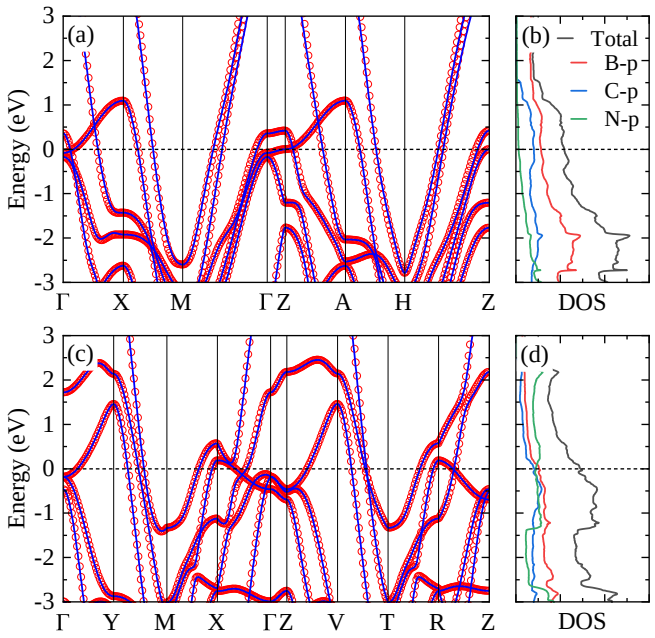

FIG. 3. (a)-(b) Electronic structure and DOS of B_5CN_2 at ambient pressure. (c)-(d) Electronic structure and DOS of B_4CN_3 under 5 GPa. The blue soild lines and red circles represent the bands obtained by DFT and Wannier projection, respectively. The Fermi level is set to be zero.

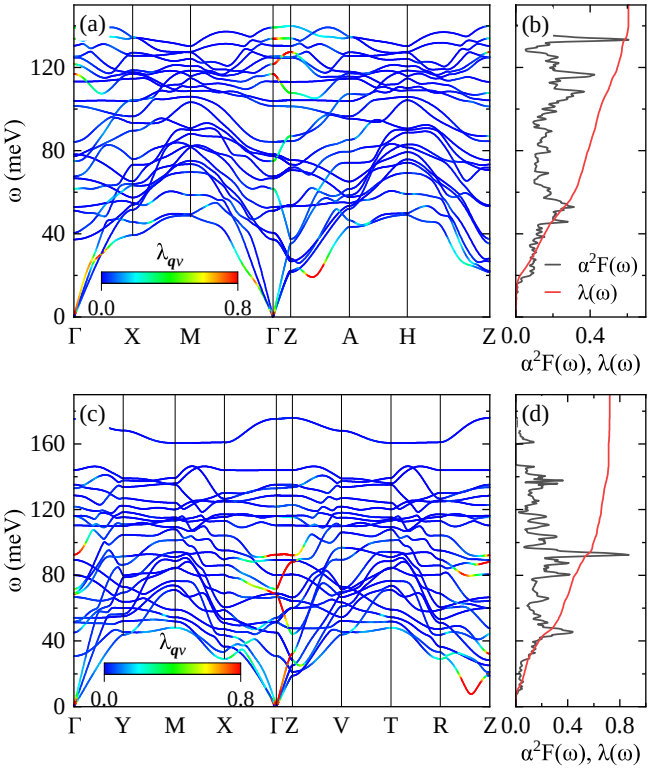

FIG. 4. (a)-(b) Phonon spectrum with a color representation of $\lambda_{q\nu}$, Eliashberg spectral function $\alpha^2F(\omega)$, and accumulated EPC constant $\lambda(\omega)$ for B_5CN_2 at ambient pressure under 5 GPa. (c)-(d) Phonon spectrum with a color representation of $\lambda_{q\nu}$, Eliashberg spectral function $\alpha^2F(\omega)$, and accumulated EPC constant $\lambda(\omega)$ for B_4CN_3 at ambient pressure under 5 GPa. The scale of $\alpha^2F(\omega)$ is omitted.

respectively estimated to be 15.93 K and 24.08 K when the Coulomb pseudopotential μ^* is set to be 0.1.

Discussion. Recently, several studies have utilized generative models to explore high-T_c superconductors [17, 18, 20, 23]. Wines *et al.* [23] employed CDVAE to generate data on the JARVES-DFT dataset [25], subsequently employing the ALIGNN [26] for T_c forecasting. Compared to existing methods, our proposed AI search engine has seen improvements in three aspects. Firstly, our method is capable of effective generation based on a few positive samples (*i.e.*, 105 samples with $T_c \geq 5$ K). Unlike CDVAE [24], which randomly generates chemical formulas before predicting structures, our approach directly generates structures. By directly generating structural configurations, our method adeptly navigates the spatial intricacies of superconductors, facilitating the genesis of plausible chemical entities. Secondly, we have integrated a sophisticated post-processing phase employing the DPA-2 model [43] for atom docking. This step meticulously circumvents atomic clashes, refines bond lengths to more rational values, and guarantees the equilibrium of forces exerted on each atomic constituent. Existing methods often predict superconducting transition temperatures without aforehand confirming the materials' superconductivity, which is problematic. We address this by introducing a superconducting classification model. We improved the formation energy prediction model under GNoME [40], increasing its precision from 28 meV to 21 meV. Lastly, active learning progressively expands the chemical space of high-T_c superconducting materials in iterative reinforcement generative learning. These refinements enhance our method's effectiveness and establish a new standard for exploring and predicting high-T_c superconductors.

In conclusion, our proposed AI search engine integrates a suite of advanced methodologies, including generative model, formation energy prediction model, pre-training and fine-tuning strategy, ALIGNN, and first-principles electronic structure calculations. This AI search engine has not only predicted 74 superconducting material candidates with T_c >15 K based on a modest set of positive samples (105 samples with T_c ≥5 K), but also identified two ideal high-T_c candidates: B_5CN_2 (T_c=15.93 K) and B_4CN_3 (T_c=24.08 K). Notably, this engine is capable of discovering crystal structures that are not yet documented in existing material dataset, thereby opening up new horizons in the search for high-T_c superconductors. Moreover, the AI search engine's versatility allows it to be adapted for exploring a wide range of functional materials with various target properties, significantly expanding its potential applications in materials science.

ACKNOWLEDGMENTS

This work was financially supported by the National Natural Science Foundation of China (Grant No.62476278, No.62206299, No.11934020, and No.12204533). Computational resources have been provided by the Physical Laboratory of High Performance Computing at Renmin University of China.

* These authors contributed equally to this work.
† zfgao@ruc.edu.cn
‡ zlu@ruc.edu.cn

[1] H. K. Onnes, The resistance of pure mercury at helium temperatures, Commun. Phys. Lab. Univ. Leiden, b **120** (1911).

[2] Y. Lvovsky, E. W. Stautner, and T. Zhang, Novel technologies and configurations of superconducting magnets for mri, Supercond. Sci. Technol. **26**, 093001 (2013).

[3] P. Bruzzone, W. H. Fietz, J. V. Minervini, M. Novikov, N. Yanagi, Y. Zhai, and J. Zheng, High temperature superconductors for fusion magnets, Nucl. Fusion **58**, 103001 (2018).

[4] M. Mirhosseini, A. Sipahigil, M. Kalaee, and O. Painter, Superconducting qubit to optical photon transduction, Nature **588**, 599 (2020).

[5] J. M. Gambetta, J. M. Chow, and M. Steffen, Building logical qubits in a superconducting quantum computing system, NPJ Quantum Inf. **3**, 2 (2017).

[6] C. L. Degen, F. Reinhard, and P. Cappellaro, Quantum sensing, Rev. Mod. Phys. **89**, 035002 (2017).

[7] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Theory of superconductivity, Phys. Rev. **108**, 1175 (1957).

[8] J. Ying, S. Liu, Q. Lu, X. Wen, Z. Gui, Y. Zhang, X. Wang, J. Sun, and X. Chen, Record high 36 k transition temperature to the superconducting state of elemental scandium at a pressure of 260 gpa, Phys. Rev. Lett. **130**, 256002 (2023).

[9] H. Sun, M. Huo, X. Hu, J. Li, Z. Liu, Y. Han, L. Tang, Z. Mao, P. Yang, B. Wang, J. Cheng, D.-X. Yao, G.-M. Zhang, and M. Wang, Signatures of superconductivity near 80 K in a nickelate under high pressure, Nature **621**, 493 (2023).

[10] D. Duan, Y. Liu, F. Tian, D. Li, X. Huang, Z. Zhao, H. Yu, B. Liu, W. Tian, and T. Cui, Pressure-induced metallization of dense $(H_2S)_2H_2$ with high-T_c superconductivity, Sci. Rep. **4**, 6968 (2014).

[11] H. Liu, I. I. Naumov, R. Hoffmann, N. W. Ashcroft, and R. J. Hemley, Potential high-T_c superconducting lanthanum and yttrium hydrides at high pressure, Proc. Natl. Acad. Sci. U.S.A **114**, 6990 (2017).

[12] F. Peng, Y. Sun, C. J. Pickard, R. J. Needs, Q. Wu, and Y. Ma, Hydrogen clathrate structures in rare earth hydrides at high pressures: Possible route to room-temperature superconductivity, Phys. Rev. Lett. **119**, 107001 (2017).

[13] M. Gao, X.-W. Yan, Z.-Y. Lu, and T. Xiang, Phonon-mediated high-temperature superconductivity in the ternary borohydride kb_2h_8 under pressure near 12 gpa, Phys. Rev. B **104**, L100504 (2021).

[14] A. Drozdov, P. Kong, V. Minkov, S. Besedin, M. Kuzovnikov, S. Mozaffari, L. Balicas, F. Balakirev, D. Graf, V. Prakapenka, *et al.*, Superconductivity at 250 k in lanthanum hydride under high pressures, Nature **569**, 528 (2019).

[15] A. M. Shipley, M. J. Hutcheon, R. J. Needs, and C. J. Pickard, High-throughput discovery of high-temperature conventional superconductors, Phys. Rev. B **104**, 054501 (2021).

[16] A. Drozdov, M. Eremets, I. Troyan, V. Ksenofontov, and S. I. Shylin, Conventional superconductivity at 203 kelvin at high pressures in the sulfur hydride system, Nature **525**, 73 (2015).

[17] K. Choudhary and K. Garrity, Designing high-tc superconductors with bcs-inspired screening, density functional theory, and deep-learning, NPJ Comput. Mater. **8**, 244 (2022).

[18] D. Wines, K. Choudhary, A. J. Biacchi, K. F. Garrity, and F. Tavazza, High-throughput dft-based discovery of next generation two-dimensional (2d) superconductors, Nano Lett. **23**, 969 (2023).

[19] V. Stanev, C. Oses, A. G. Kusne, E. Rodriguez, J. Paglione, S. Curtarolo, and I. Takeuchi, Machine learning modeling of superconducting critical temperature, NPJ Comput. Mater. **4**, 29 (2018).

[20] J. Zhang, Z. Zhu, X.-D. Xiang, K. Zhang, S. Huang, C. Zhong, H.-J. Qiu, K. Hu, and X. Lin, Machine learning prediction of superconducting critical temperature through the structural descriptor, J. Phys. Chem. C **126**, 8922 (2022).

[21] D. Menon and R. Ranganathan, A generative approach to materials discovery, design, and optimization, ACS Omega **7**, 25958 (2022).

[22] B. Roter and S. Dordevic, Predicting new superconductors and their critical temperatures using machine learning, Physica C Supercond **575**, 1353689 (2020).

[23] D. Wines, T. Xie, and K. Choudhary, Inverse design of next-generation superconductors using data-driven deep generative models, J. Phys. Chem. Lett. **14**, 6630 (2023).

[24] T. Xie, X. Fu, O. Ganea, R. Barzilay, and T. S. Jaakkola, Crystal diffusion variational autoencoder for periodic material generation, arXiv **2110.06197** (2021).

[25] K. Choudhary, K. F. Garrity, A. C. Reid, B. DeCost, A. J. Biacchi, A. R. Hight Walker, Z. Trautt, J. Hattrick-Simpers, A. G. Kusne, A. Centrone, *et al.*, The joint automated repository for various integrated simulations (jarvis) for data-driven materials design, NPJ Comput. Mater. **6**, 173 (2020).

[26] K. Choudhary and B. DeCost, Atomistic line graph neural network for improved materials property predictions, NPJ Comput. Mater. **7**, 185 (2021).

[27] R. Jiao, W. Huang, P. Lin, J. Han, P. Chen, Y. Lu, and Y. Liu, Crystal structure prediction by joint equivariant diffusion, Advances in Neural Information Processing Systems **36** (2024).

[28] J. Ho, A. Jain, and P. Abbeel, Denoising diffusion probabilistic models, NeurIPS **33**, 6840 (2020).

[29] V. G. Satorras, E. Hoogeboom, and M. Welling, E (n) equivariant graph neural networks, ICML **139**, 9323 (2021).

[30] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding, ACL **1**, 4171 (2019).

[31] T. Xie and J. C. Grossman, Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties, Phys. Rev. Lett. **120**, 145301 (2018).

[32] L. Rüschendorf, The wasserstein distance and approximation theorems, Probab Theory Relat Fields **70**, 117 (1985).

[33] C. Chen, W. Ye, Y. Zuo, C. Zheng, and S. P. Ong, Graph networks as a universal machine learning framework for molecules and crystals, Chem. Mater **31**, 3564 (2019).

[34] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, Score-based generative modeling through stochastic differential equations, ICLR (2021).

[35] A. Jain, J. Montoya, Dwaraknath, *et al.*, The materials project: Accelerating materials design through theory-driven data and tools, Handbook of Materials Modeling , 1751 (2020).

[36] Y. Yang, O. A. Jiménez-Negrón, and J. R. Kitchin, Machine-learning accelerated geometry optimization in molecular simulation, J. Chem. Phys **154**, 234704 (2021).

[37] D. C. Liu and J. Nocedal, On the limited memory bfgs method for large scale optimization, Math Program **45**, 503 (1989).

[38] Z.-F. Gao, S. Qu, B. Zeng, Y. Liu, J.-R. Wen, H. Sun, P.-J. Guo, and Z.-Y. Lu, Ai-accelerated discovery of altermagnetic materials, arXiv **2311.04418** (2024).

[39] K. Schütt, P.-J. Kindermans, H. E. Sauceda Felix, S. Chmiela, A. Tkatchenko, and K.-R. Müller, Schnet: A continuous-filter convolutional neural network for modeling quantum interactions, NeurIPS **30**, 991 (2017).

[40] A. Merchant, S. Batzner, S. S. Schoenholz, M. Aykol, G. Cheon, and E. D. Cubuk, Scaling deep learning for materials discovery, Nature **624**, 80 (2023).

[41] P. B. Allen, Neutron spectroscopy of superconductors, Phys. Rev. B **6**, 2577 (1972).

[42] P. B. Allen and R. C. Dynes, Transition temperature of strong-coupled superconductors reanalyzed, Phys. Rev. B **12**, 905 (1975).

[43] D. Zhang, X. Liu, X. Zhang, C. Zhang, C. Cai, H. Bi, Y. Du, X. Qin, J. Huang, B. Li, *et al.*, Dpa-2: Towards a universal large atomic model for molecular and material simulation, arXiv preprint arXiv:2312.15492 (2023).

[44] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. D. Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, Quantum espresso: a modular and open-source software project for quantum simulations of materials, Journal of Physics: Condensed Matter **21**, 395502 (2009).

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

[46] V. D. Bortoli, E. Mathieu, M. J. Hutchinson, J. Thornton, Y. W. Teh, and A. Doucet, Riemannian score-based generative modelling, in *Advances in Neural Information Processing Systems*, edited by A. H. Oh, A. Agarwal, D. Belgrave, and K. Cho (2022).

[47] K. Choudhary and B. DeCost, Atomistic line graph neural network for improved materials property predictions, npj Computational Materials **7**, 10.1038/s41524-021-00650-1 (2021).

[48] T. Xie and J. C. Grossman, Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties, Phys. Rev. Lett. **120**, 145301 (2018).

[49] M. Tang, P. Li, and C. Yang, Graph auto-encoder via neighborhood wasserstein reconstruction, in *ICLR* (2022).

[50] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, Score-based generative modeling through stochastic differential equations, arXiv preprint arXiv:2011.13456 (2020).

[51] D. R. Hamann, Optimized norm-conserving vanderbilt pseudopotentials, Phys. Rev. B **88**, 085117 (2013).

[52] M. Methfessel and A. T. Paxton, High-precision sampling for brillouin-zone integration in metals, Phys. Rev. B **40**, 3616 (1989).

[53] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, Phonons and related crystal properties from density-functional perturbation theory, Rev. Mod. Phys. **73**, 515 (2001).

[54] G. Pizzi, V. Vitale, R. Arita, S. Blügel, F. Freimuth, G. Géranton, M. Gibertini, D. Gresch, C. Johnson, T. Koretsune, J. Ibañez-Azpiroz, H. Lee, J.-M. Lihm, D. Marchand, A. Marrazzo, Y. Mokrousov, J. I. Mustafa, Y. Nohara, Y. Nomura, L. Paulatto, S. Poncé, T. Ponweiser, J. Qiao, F. Thöle, S. S. Tsirkin, M. Wierzbowska, N. Marzari, D. Vanderbilt, I. Souza, A. A. Mostofi, and J. R. Yates, Wannier90 as a community code: new features and applications, Journal of Physics: Condensed Matter **32**, 165902 (2020).

[55] S. Poncé, E. Margine, C. Verdi, and F. Giustino, Epw: Electron–phonon coupling, transport and superconducting properties using maximally localized wannier functions, Computer Physics Communications **209**, 116 (2016).

CALCULATION METHOD

The first-principle calculations were performed by QUANTUM-ESPRESSO package [44]. The generalized gradient approximation of Perdew-Burke-Ernzerhof formula [45] and the optimized norm-conserving Vanderbilt pseudopotentials [51] were used. We set the kinetic energy and the charge density cutoff to be 80 Ry and 320 Ry, respectively. An unshifted 12×12×6 **k**-points grid was used in the self-consistent calculations. The Methfessel-Paxton smearing method [52] with a width of 0.02 Ry was set for the Fermi surface broadening. The dynamical matrices and the perturbation potentials were calculated on a Γ-centered 4×4×2 **q**-points grid based on the density functional perturbation theory [53].

The maximally localized Wannier functions (MLWFs) [54] were constructed on a 4×4×2 **k**-points grid of the Brillouin zone. The MLWFs of B_5CN_2 and B_4CN_3 include the $2p$ orbitals of B, C, and N atoms. We used EPW package [55] to perform electron-phonon coupling (EPC) calculations. The EPC constant λ was determined through fine electron (72×72×36) and phonon (24×24×12) grids. And the smearing widths of electron and phonon Dirac δ functions are 0.1 eV and 0.5 eV, respectively.

The mode and wavevecter-dependent coupling $\lambda_{\mathbf{q}\nu}$ reads:

$$\lambda_{\mathbf{q}\nu} = \frac{2}{\hbar N(0) N_{\mathbf{k}}} \sum_{nm\mathbf{k}} \frac{1}{\omega_{\mathbf{q}\nu}} |g^{nm}_{\mathbf{k},\mathbf{q}\nu}|^2 \delta(\epsilon^n_{\mathbf{k}}) \delta(\epsilon^m_{\mathbf{k}+\mathbf{q}}). \quad (4)$$

$N(0)$ is the density of states (DOS) of electrons at the Fermi level. $N_{\mathbf{k}}$ is the total number of **k** points in the fine **k**-mesh. $\omega_{\mathbf{q}\nu}$ is the phonon frequency and $g^{nm}_{\mathbf{k},\mathbf{q}\nu}$ is the EPC matrix element. $(n,\ m)$ and ν denote the indices of energy bands and phonon mode, respectively. $\epsilon^n_{\mathbf{k}}$ and $\epsilon^m_{\mathbf{k}+\mathbf{q}}$ are the eigenvalues of the Kohn-Sham orbitals with respect to the Fermi level.

The EPC constant λ was determined by the summation of $\lambda_{\mathbf{q}\nu}$ over the first Brillouin zone, or the integration of the Eliahberg spectral function $\alpha^2F(\omega)$ [41, 42],

$$\lambda = \frac{1}{N_{\mathbf{q}}} \sum_{\mathbf{q}\nu} \lambda_{\mathbf{q}\nu} = 2 \int \frac{\alpha^2F(\omega)}{\omega} d\omega, \quad (5)$$

where $N_{\mathbf{q}}$ represents the total number of **q** points in the fine **q** mesh. The Eliashberg spectral function $\alpha^2F(\omega)$ was calculated with:

$$\alpha^2F(\omega) = \frac{1}{2N_{\mathbf{q}}} \sum_{\mathbf{q}\nu} \lambda_{\mathbf{q}\nu} \omega_{\mathbf{q}\nu} \delta(\omega - \omega_{\mathbf{q}\nu}). \quad (6)$$

And $\omega_{\log}$ is the logarithmic average frequency that is defined as

$$\omega_{\log} = \exp\left[\frac{2}{\lambda} \int \frac{d\omega}{\omega} \alpha^2F(\omega) ln(\omega)\right]. \quad (7)$$

CRYSTAL STRUCTURE

The optimized crystal structures are shown in Fig. 5. The crystal parameters of B_5CN_2 at ambient pressure are $a = 2.49497$ Å, $b = 2.52950$ Å, $c = 8.35250$ Å, $\alpha = 90°$, $\beta = 90°$, and $\gamma = 90°$, respectively. As for B_4CN_3 under 5 GPa, crystal parameters are $a = 2.52034$ Å, $b = 2.59156$ Å, $c = 8.13945$ Å, $\alpha = 90°$, $\beta = 98.9°$, and $\gamma = 90°$, respectively.

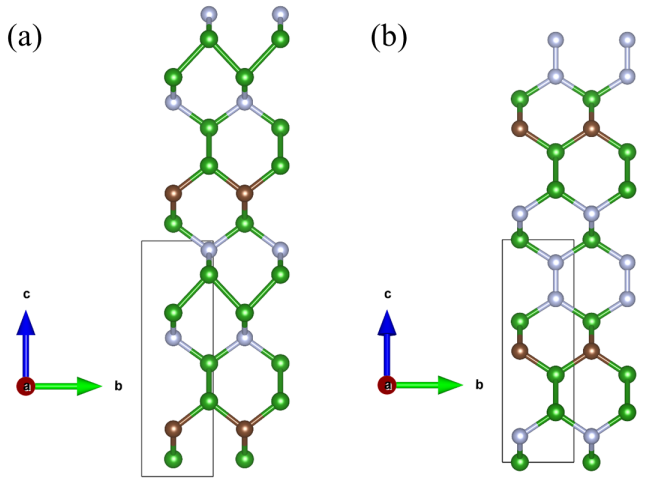

FIG. 5. Crystal structures. (a) B_5CN_2 at ambient pressure. (b) B_4CN_3 under 5 GPa. The black lines represent the unit cell. And the green, brown, and gray balls denote B, C, and N atoms, respectively.

FERMI SURFACES

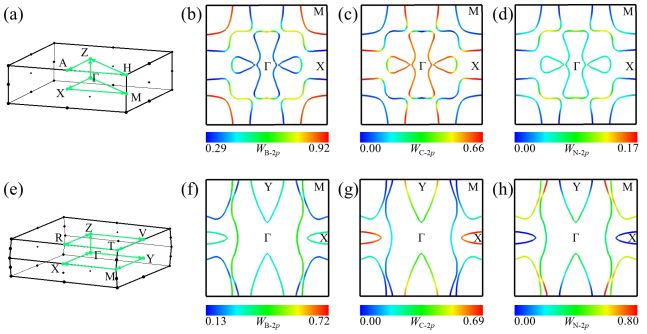

FIG. 6. (a) Brillouin zone and high-symmetry path of B_5CN_2 at ambient pressure. (c)-(d) Distributions of different orbitals on the $k_z = 0$ Fermi surfaces. (e)-(h) Brillouin zone and Fermi surfaces of B_4CN_3 under 5 GPa.

CANDIDATE SUPERCONDUCTORS.

The table I- II present the candidate superconducting formulas and their corresponding predicted superconducting transition stability values, obtained from our AI search engine.

TABLE I. Candidate superconductors.

Number	Materials	ALIGNN Prediction T_c/K
1	B_4C_3N	18.3
2	B_4CN_3	18.6
3	B_5CN_2	17.9
4	ReN_9	18.9
5	$Ti(ReN_2)_3$	15.7
6	TiN_9	18.6
7	$TiReN_8$	18.3
8	$ZrBN_2$	15.6
9	$B_3C_2N_3$	19.5
10	$B_3C_2N_4$	19.6
11	$B_3C_3N_3$	19.2
12	B_3CN_4	18.4
13	B_3CN_5	18.2
14	B_3N_5	17.8
15	B_3N_6	17.9
16	$B_4C_2N_3$	18.7
17	B_4CN_2	17.9
18	B_4CN_4	18.2
19	B_4N_5	17.7
20	$B_5C_2N_2$	18.2
21	B_5N_3	16.9
22	B_5N_4	17.5
23	B_6CN_2	16.3
24	B_6N_2	16.4
25	B_6N_3	15.9
26	$Eu_2C_3N_8$	19.2
27	Eu_2CN_{10}	20.1
28	Eu_2N_{11}	19.5
29	$Eu_2Re_2C_2N_7$	17.4
30	$Eu_2Re_2CN_8$	17.6
31	$Eu_2Re_2N_9$	18.4
32	$Eu_2Re_3C_2N_6$	16.0
33	$Eu_2Re_3CN_7$	16.1
34	$Eu_2ReC_2N_8$	19.1

TABLE II. Candidate superconductors.

Number	Materials	ALIGNN Prediction T_c/K
35	$Eu_2ReC_3N_7$	17.9
36	Eu_2ReCN_9	19.2
37	Eu_2ReN_{10}	19.5
38	Eu_2TiCN_9	19.1
39	$Eu_3C_2N_8$	19.5
40	Eu_3CN_9	19.9
41	Eu_3N_{10}	19.3
42	$Eu_3Re_2CN_7$	17.5
43	$Eu_3Re_2N_8$	18.7
44	$Eu_3Re_3N_7$	15.0
45	Eu_3ReCN_8	19.0
46	Eu_3ReN_9	18.7
47	Eu_4CN_8	19.2
48	Eu_4N_9	19.1
49	Eu_4ReCN_7	18.9
50	$EuTiCN_{10}$	17.4
51	$EuTiRe_2CN_8$	18.1
52	Li_2CO	15.4
53	$Np_2Ti_2C_4N_5$	15.3
54	$Np_2TiC_4N_6$	16.9
55	$Np_2TiC_6N_4$	16.3
56	$Np_3C_4N_6$	17.7
57	$Np_3C_6N_4$	17.3
58	$Np_3TiC_2N_7$	17.6
59	$Np_3TiC_3N_6$	16.6
60	$Np_3TiC_4N_5$	16.0
61	$Np_3TiC_5N_4$	15.9
62	$Np_4C_4N_5$	17.3
63	$Np_4C_5N_4$	17.1
64	Ti_2CN_8	16.1
65	Ti_2N_7	16.6
66	Ti_2N_8	16.3
67	$Ti_2Re_2CN_7$	16.7
68	$Ti_2ReC_2N_6$	15.1
69	Ti_2ReCN_7	15.5
70	Ti_2ReCN_8	16.6
71	Ti_2ReN_6	16.4
72	$TiRe_2N_6$	16.2
73	$TiReBN_6$	15.4
74	V_2Ru_4	15.1

SYMMETRY-CONSTRAINED CRYSTAL GENERATION MODEL DETAIL

Equivariant denoising model

Any atom in the crystal can be expressed by its Cartesian coordinates and type as $\{(\boldsymbol{a}_i', \boldsymbol{x}_i') | \boldsymbol{a}_i' = \boldsymbol{a}_i, \boldsymbol{x}_i' = \boldsymbol{x}_i + \boldsymbol{L}\boldsymbol{k}, \forall \boldsymbol{k} \in \mathbb{Z}^{3\times 1}\}$. There is a relationship between Cartesian coordinates and fractional coordinates given by $\boldsymbol{x} = \sum_{i=1}^{3} f_i \boldsymbol{l}_i$. The following generation process employs the fractional coordinate system $\mathcal{M} = (\boldsymbol{A}, \boldsymbol{F}, \boldsymbol{L})$. Given the periodicity of $\boldsymbol{F}$, it is generated using a score-matching (SM) based framework [34]. and its loss function is:

$$\mathcal{L}_{\boldsymbol{F}} = \mathbb{E}_{\boldsymbol{F}_t \sim q_t} \left[\lambda_t \| \nabla_{\boldsymbol{F}_t} \log q(\boldsymbol{F}_t | \boldsymbol{F}_0) - \hat{\boldsymbol{\epsilon}}_{\boldsymbol{F}}(\mathcal{M}_t, t) \|_2^2 \right]. \quad (8)$$

Similar with DiffCSP [27], the distribution q_t here uses the wrapped normal(WN) [46] distribution. where λ_t is approximated via Monte-Carlo sampling. In this work, we employed EGNN [29] as an equivariant denoising model. The superconducting unit cell structure is represented as an atomic graph, where the encoding of atomic types and the diffusion time step t is processed through an MLP, serving as the model's input: $\boldsymbol{h}_i^{(0)} = \mathbf{MLP}(f_{\text{atom}}(\boldsymbol{a}_i), f_{\text{pos}}(t))$. The representation of the i-th node at the s-th layer is denoted as $\boldsymbol{h}_i^{(s)}$.

$$\boldsymbol{h}_i^{(s)} = \boldsymbol{h}_i^{(s-1)} + \varphi_h(\boldsymbol{h}_i^{(s-1)}, \boldsymbol{m}_i^{(s)}). \quad (9)$$

φ_h denotes MLP, and $\boldsymbol{m}_i^{(s)}$ represents the aggregated representation from all nodes to the i-th node.

$$\boldsymbol{m}_i^{(s)} = \sum_{j=1}^{N} \boldsymbol{m}_{ij}^{(s)}, \quad (10)$$

the representation of the edge ij at layer s is given by

$$\boldsymbol{m}_{ij}^{(s)} = \varphi_m(\boldsymbol{h}_i^{(s-1)}, \boldsymbol{h}_j^{(s-1)}, \boldsymbol{L}^\top \boldsymbol{L}, \psi_{\mathrm{FT}}(\boldsymbol{f}_j - \boldsymbol{f}_i)). \quad (11)$$

The term $\boldsymbol{L}^\top \boldsymbol{L}$ ensures that the input is invariant under O(3) transformations. The edge vector $(\boldsymbol{f}_j - \boldsymbol{f}_i)$ is transformed using the Fourier transform $\psi_{\mathrm{FT}} : (-1, 1)^3 \to [-1, 1]^{3\times K}$, which guarantees the translational invariance of the edge vector. The representation of the final layer nodes predicts the atomic type noise $\hat{\boldsymbol{\epsilon}}_{\boldsymbol{A}}$, the lattice matrix noise $\hat{\boldsymbol{\epsilon}}_{\boldsymbol{L}}$, and the fractional coordinate score $\hat{\boldsymbol{\epsilon}}_{\boldsymbol{F}}$ using different MLPs.

Algorithms for Training and Sampling

Algorithm 1 provides an overview of the forward diffusion process and the training procedure for the denoising model ϕ. Meanwhile, Algorithm 2 details the backward sampling process. These algorithms are designed to preserve symmetries, provided that ϕ is carefully constructed. Specifically, the predictor-corrector sampler [50] is employed to sample $\boldsymbol{F}_0$. In Algorithm 2, Line 8 represents the predictor step, while Lines 10-11 correspond to the corrector steps.

Algorithm 1 Training procedure of symmetry-constrained crystal generation model.

Input: lattice matrix $\boldsymbol{L}_0$, atom types $\boldsymbol{A}_0$, atom number N in cell,fractional coordinates $\boldsymbol{F}_0$, denoising model ϕ, and the number of sampling steps T.
Sample $\boldsymbol{\epsilon}_{\boldsymbol{L}} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$,$\boldsymbol{\epsilon}_{\boldsymbol{A}} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$,$\boldsymbol{\epsilon}_{\boldsymbol{F}} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$ and $t \sim \mathcal{U}(1, T)$.
$\boldsymbol{L}_t \leftarrow \sqrt{\bar{\alpha}_t}\boldsymbol{L}_0 + \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon}_{\boldsymbol{L}}$
$\boldsymbol{A}_t \leftarrow \sqrt{\bar{\alpha}_t}\boldsymbol{A}_0 + \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon}_{\boldsymbol{A}}$
$\boldsymbol{F}_t \leftarrow w(\boldsymbol{F}_0 + \sigma_t \boldsymbol{\epsilon}_{\boldsymbol{F}})$
$\hat{\boldsymbol{\epsilon}}_{\boldsymbol{L}}, \hat{\boldsymbol{\epsilon}}_{\boldsymbol{A}}, \hat{\boldsymbol{\epsilon}}_{\boldsymbol{F}} \leftarrow \phi(\boldsymbol{L}_t, \boldsymbol{F}_t, \boldsymbol{A}_t, N, t)$
$\mathcal{L}_{\boldsymbol{L}} \leftarrow \|\boldsymbol{\epsilon}_{\boldsymbol{L}} - \hat{\boldsymbol{\epsilon}}_{\boldsymbol{L}}\|_2^2$
$\mathcal{L}_{\boldsymbol{A}} \leftarrow \|\boldsymbol{\epsilon}_{\boldsymbol{A}} - \hat{\boldsymbol{\epsilon}}_{\boldsymbol{A}}\|_2^2$
$\mathcal{L}_{\boldsymbol{F}} \leftarrow \lambda_t \|\nabla_{\boldsymbol{F}_t} \log q(\boldsymbol{F}_t|\boldsymbol{F}_0) - \hat{\boldsymbol{\epsilon}}_{\boldsymbol{F}}\|_2^2$
Minimize $\mathcal{L}_{\boldsymbol{L}} + \mathcal{L}_{\boldsymbol{A}} + \mathcal{L}_{\boldsymbol{F}}$

Algorithm 2 Sampling procedure of symmetry-constrained crystal generation model.

Input: atom number N in cell, denoising model ϕ, number of sampling steps T, step size of Langevin dynamics γ.
Sample $\boldsymbol{L}_T \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$,$\boldsymbol{A}_T \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$,$\boldsymbol{F}_T \sim \mathcal{U}(0, 1)$.
for $t \leftarrow T, \cdots, 1$ **do**
  Sample $\boldsymbol{\epsilon}_{\boldsymbol{L}}, \boldsymbol{\epsilon}_{\boldsymbol{A}}, \boldsymbol{\epsilon}_{\boldsymbol{F}}, \boldsymbol{\epsilon}'_{\boldsymbol{F}} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$
  $\hat{\boldsymbol{\epsilon}}_{\boldsymbol{L}}, \hat{\boldsymbol{\epsilon}}_{\boldsymbol{A}}, \hat{\boldsymbol{\epsilon}}_{\boldsymbol{F}} \leftarrow \phi(\boldsymbol{L}_t, \boldsymbol{F}_t, \boldsymbol{A}_t, N, t)$.
  $\boldsymbol{L}_{t-1} \leftarrow \frac{1}{\sqrt{\alpha_t}}(\boldsymbol{L}_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}}\hat{\boldsymbol{\epsilon}}_{\boldsymbol{L}}) + \sqrt{\beta_t \cdot \frac{1-\bar{\alpha}_{t-1}}{1-\bar{\alpha}_t}}\boldsymbol{\epsilon}_{\boldsymbol{L}}$.
  $\boldsymbol{A}_{t-1} \leftarrow \frac{1}{\sqrt{\alpha_t}}(\boldsymbol{A}_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}}\hat{\boldsymbol{\epsilon}}_{\boldsymbol{A}}) + \sqrt{\beta_t \cdot \frac{1-\bar{\alpha}_{t-1}}{1-\bar{\alpha}_t}}\boldsymbol{\epsilon}_{\boldsymbol{A}}$.
  $\boldsymbol{F}_{t-\frac{1}{2}} \leftarrow w(\boldsymbol{F}_t + (\sigma_t^2 - \sigma_{t-1}^2)\hat{\boldsymbol{\epsilon}}_{\boldsymbol{F}} + \frac{\sigma_{t-1}\sqrt{\sigma_t^2 - \sigma_{t-1}^2}}{\sigma_t}\boldsymbol{\epsilon}_{\boldsymbol{F}})$
  $\_, \hat{\boldsymbol{\epsilon}}_{\boldsymbol{F}} \leftarrow \phi(\boldsymbol{L}_{t-1}, \boldsymbol{F}_{t-\frac{1}{2}}, \boldsymbol{A}, t-1)$.
  $d_t \leftarrow \gamma\sigma_{t-1}^2/\sigma_1^2$
  $\boldsymbol{F}_{t-1} \leftarrow w(\boldsymbol{F}_{t-\frac{1}{2}} + d_t\hat{\boldsymbol{\epsilon}}_{\boldsymbol{F}} + \sqrt{2d_t}\boldsymbol{\epsilon}'_{\boldsymbol{F}})$.
end for
Return $\boldsymbol{L}_0, \boldsymbol{A}_0, \boldsymbol{F}_0$.

Hyper-parameters and training details

In this section, we detail the training process of the superconducting generative model. First, we perform an up-sampling operation on 105 BCS superconducting crystals, replicating the training dataset 100 times to create the superconductors dataset. We utilize a model architecture with 6 layers and 512 hidden units. The dimension of the Fourier embedding is set to $k = 256$. A cosine scheduler with $s = 0.008$ is applied to control the variance in the DDPM process on $\boldsymbol{L}_t$, and an exponential scheduler with $\sigma_1 = 0.005$ and $\sigma_T = 0.5$ is used to manage the noise scale in the score-matching process on $\boldsymbol{F}_t$. The diffusion step is set to $T = 1000$. Our model is trained for 1000 epochs. For ab initio generation, we apply a scaling factor of $\gamma = 5 \times 10^{-6}$. The model training was conducted on a GeForce RTX 3090 GPU.

SUPERCONDUCTING CLASSIFICATION MODEL

Model architecture

We developed a superconductor classification model inspired by MatAltMag [38]. We define the graph representation $\mathcal{G}(V, U, X)$ to encapsulate crystal structure information, where V represents the set of nodes, U the set of edges, and X the set of features. In this representation, atoms in the crystal structure are depicted as nodes v_i, where $i = 1, \ldots, |V|$. Due to the consideration of periodic boundary conditions, equivalent nodes are merged, resulting in a set of irreducible nodes. For each node v_i, we identify its neighboring nodes v_j, where $j = 1, \ldots, |\mathcal{N}_i|$, and $\mathcal{N}_i$ denotes the set of neighbors of node v_i. The connections between nodes v_i and v_j are represented by the

edges $u_{(i,j)_k}$, where k denotes the number of bonds between the nodes. The initial features of each node v_i are denoted by $\{h_i^{(0)}\}_{i=1}^{|V|}$ and are derived through one-hot encoding based on the atomic sequence in the crystal structure. The neighbor node features for node v_i are represented as $H_{\mathcal{N}_i}^{(0)}$. Each edge $u_{(i,j)_k} \in U$ is associated with a feature vector $u_{(i,j)_k}$, corresponding to the k-th bond between nodes v_i and v_j. Finally, each node $v_i \in V$ is represented by a feature vector $h_i \in X$, which encodes the attributes of the atom corresponding to that node.

The encoder maps the input crystal structure information to a high-dimensional matrix, using n convolutional layers. In each layer t, the node feature vector $h_i^{(t)}$ is updated through the convolution function $h_i^{(t+1)} = \mathrm{Conv}\big(h_i^{(t)}, h_j^{(t)}, u_{(i,j)_k}\big)$. The graph convolution function g iteratively refines the feature vector h_i, feeding the output from one step as input to the next. The structure and length of h_i remain consistent across steps. At each step t, the concatenated neighbor vector is defined as $z_{(i,j)_k}^{(t)} = h_i^{(t)} \oplus h_j^{(t)} \oplus u_{(i,j)_k}$. The feature update is then performed via the convolution operation:

$$h_i^{(t+1)} = h_i^{(t)} + \sum_{v_j \in \mathcal{N}_i, v_m \in \mathcal{M}_i, k} \sigma\left(z_{(i,j)_k}^{(t)} W_f^{(t)} + b_f^{(t)}\right) \odot g\left(z_{(i,j)_k}^{(t)} W_s^{(t)} + h_{i,m}^{(t)} W_m^{(t)} + b_s^{(t)}\right), \quad (12)$$

where $\odot$ represents element-wise multiplication, and σ is the sigmoid activation function. The term $W_m^{(t)}$ accounts for the importance of magnetic atoms, crucial for superconducting properties. Residual connections via $h_i^{(t)}$ are included to facilitate neural network training.

The decoder, denoted by ψ, reconstructs the input graph representation of a crystal from its embeddings. It consists of two main components: node feature reconstruction (ψ_s) and adjacent node feature reconstruction (ψ_p), with ψ defined as $\psi = (\psi_p + \psi_s)$. Node features are reconstructed using $\psi_s = \mathrm{MLP}_s\big(h_i^{(t)}\big)$, where MLP stands for a multilayer perceptron. The decoder block architecture follows the design in [49]. We utilize an n-hop neighboring Wasserstein decoder for graph feature reconstruction. First, we obtain the initial node features $\{h_i^{(0)}, H_{\mathcal{N}_i}^{(0)}\}$. For each node $v_i \in V$, the GNN layer in the encoder updates the node representation $h_i^{(t+1)}$ by aggregating information from $h_i^{(t)}$ and its neighbors $H_{\mathcal{N}_i}^{(t)}$, following the rule $h_i^{(t+1)} = \phi^{(t)}\big(h_i^{(t)}, H_{\mathcal{N}_i}^{(t)}\big)$. Here, $H_{\mathcal{N}_i}^{(t)}$ is assembled based on node adjacency. The network is trained by:

$$\arg\min_{\phi,\psi} \sum_{v_i \in V} \mathcal{L}\left(h_i^{(t)}, H_{\mathcal{N}_i}^{(t)}, \psi\left(h_i^{(t+1)}, H_{\mathcal{N}_i}^{(t+1)}\right)\right), \quad (13)$$

where $\mathcal{L}(\cdot,\cdot)$ denotes the reconstruction loss over $0 \le t < n$. More detailed information can be found in MatAltMag [38]. The classifier model is built by adding a pooling layer and a softmax module to the encoder of the pre-trained model. The pooling layer aggregates the embeddings from the encoder into an overall feature vector h_g, expressed as $h_g = \mathrm{Pool}(h_0^{(0)}, h_1^{(0)}, \ldots, h_N^{(0)}, \ldots, h_N^{(n)})$, where n is the number of convolutional layers and N is the number of nodes in the graph. The softmax module then outputs a probability in the range $[0, 1]$, indicating the likelihood that a given material is superconducting.

Visualization of superconducting classification model

Initially, the crystal structures undergo encoding through the Encode 12, yielding the feature representation of the final layer. Subsequently, the t-distributed stochastic neighbor embedding (t-SNE) method is employed for dimensionality reduction to render the visualization. Fig. 7 displays the visualization of the superconductivity classification model.

Hyper-parameters and Training Details

Below are the hyperparameters used for the pre-training and fine-tuning of the superconductor classification model.

TABLE III. Hyperparameter value

Hyperparameter	Auto-encoder	Classifier
Epochs	10	500
Learning rate	1.0e-3	1.0e-3
convolution layer	6	6
Batch size	128	128
Hidden dimension	512	512
Sample size	10	-
Radius	20	20
Drop rate	0.0	0.25

FORMATION ENERGY PREDICTION MODEL

Model architecture

Let $\boldsymbol{V}$, $\boldsymbol{E}$, and $\boldsymbol{u}$ represent the node, edge, and global state attributes of atoms in a crystal, respectively. $\boldsymbol{v}_i$ denotes the attribute vector of the i-th atom , and $\boldsymbol{V}$ is the set of all $\boldsymbol{v}_i$. The edges between atoms in the crystal are defined based on the geometric distance between atoms being less than a specified threshold (8Å), with $\boldsymbol{e}_{ij}$ representing the edge attribute vector between atom i and atom j, and $\boldsymbol{E}$ being the set of all edges $\boldsymbol{e}_{ij}$. The process of updating the graph from the input graph $\boldsymbol{G}^l = (\boldsymbol{E}^l, \boldsymbol{V}^l, \boldsymbol{u}^l)$ to the output graph $\boldsymbol{G}^{l+1} =$

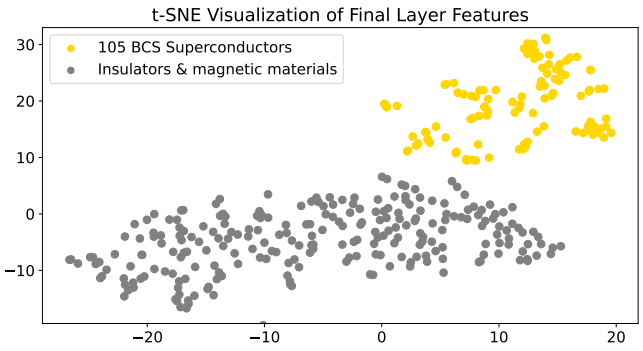

FIG. 7. Visualization of superconducting classification model. The process of visualizing the superconducting classification model initiates with the encoding of crystal structures through the Encoder 12. Following this, the features from the last layer are mapped using the t-SNE technique for dimensional reduction. Within this visual representation, superconductors are indicated by yellow points, whereas gray points correspond to insulators and other magnetic materials.

$(\boldsymbol{E}^{l+1}, \boldsymbol{V}^{l+1}, \boldsymbol{u}^{l+1})$ involves a series of state update operations. The update for the edge state is given by:

$$\boldsymbol{e}_{ij}^{l+1} = \varphi_e(\boldsymbol{v}_i^l \oplus \boldsymbol{v}_j^l \oplus \boldsymbol{e}_{ij}^l \oplus \boldsymbol{u}). \tag{14}$$

Here, φ_e represents a multi-layer perception (MLP), and $\oplus$ denotes the concatenation operation. The update for the node state is given by:

$$\boldsymbol{v}_i^{le} = \frac{1}{N_i^e} \sum_j \boldsymbol{e}_{ij}^{l+1}. \tag{15}$$

$$\boldsymbol{v}_i^{l+1} = \varphi_v(\boldsymbol{v}_i^{le} \oplus \boldsymbol{v}_i^l \oplus \boldsymbol{u}). \tag{16}$$

N_i^e represents the total number of atoms bonded to atom i, and φ_v denotes a MLP. The update for the global state is given by:

$$\boldsymbol{u}^{le} = \frac{1}{N^e} \sum_{ij} \boldsymbol{e}_{ij}^{l+1}. \tag{17}$$

$$\boldsymbol{u}^{lv} = \frac{1}{N^v} \sum_i \boldsymbol{v}_i^{l+1}. \tag{18}$$

$$\boldsymbol{u}^{l+1} = \varphi_u(\boldsymbol{u}^{le} \oplus \boldsymbol{u}^{lv} \oplus \boldsymbol{u}). \tag{19}$$

N^e represents the total number of edges, N^v represents the total number of atoms in the crystal, and φ_u denotes a MLP. Notably, we incorporated eight new atomic features into the model: electronegativity, group number, covalent radius, valence electrons, first ionization energy, electron affinity, atomic orbital, and atomic volume. These features are considered as follows: Electronegativity and electron affinity impact atomic interactions and bonding strength, thereby affecting formation energy, with higher electronegativity typically leading to stronger bonds. Valence electrons and first ionization energy influence how easily an atom loses or gains electrons, directly impacting bond strength and formation energy. Covalent radius and atomic volume affect bond lengths between atoms, influencing stability and formation energy. Group number indicates an element's reactivity and bonding modes, indirectly affecting compound formation energy. Atomic orbitals relate to formation energy indirectly by affecting chemical bond properties. All these atomic features play a role in directly or indirectly influencing crystal formation energy.

Absolute error

Here are the formation energy prediction errors of some common methods, our model demonstrates strong competitiveness.

TABLE IV. Formation energy prediction absolute error.

Method	Absolute Error/meV
MAD [47]	930
CFID [47]	104
SchNet [39]	35
CGCNN [48]	39
MEGNET [33]	28
Ours	21

Hyper-parameters and training details

We detail the hyperparameter settings employed for model training. The training dataset size was set to 440,000 records, while the test dataset size was set to 4,500 records. The batch size was configured to 256. The number of training epochs was 1,000. The bond feature dimension was established at 100. The cutoff radius was defined as 8Å. The Gaussian centers were evenly distributed from 0 to 8(Å)+1, totaling 100 centers. The Gaussian width was set to 0.5.

ACTIVE LEARNING

Using our proposed AI search engine, 74 stable superconducting candidates with $T_c \geq 15$ K were generated through three iterations of active learning. This approach progressively expanded the chemical space of high-T_c superconducting materials. Additionally, we observed that active learning improved the success rate of the DPA-2 model [43] from 15% to 53%.