Supplementary Materials for Learning Superconductivity from Ordered and Disordered Material Structures

Anonymous Author(s)

Contents

The appendix is organized as follows: Section 1 details the collection method and distribution of the SuperCon3D dataset. Section 2 presents more details of disordered graph and models. Section 3 elaborates on the implementation specifics of both property prediction and generative models. The identification of potential superconductors and their corresponding DFT computational outcomes are presented in Section 4. A systematic approach for the design of practical superconductors is expounded in Section 5. We present the limitation of our data and models in Section 6. Section 7 provides the repository link for the associated coding resources.

1 SuperCon3D data details

We extracted approximately 33,000 superconductors, including their chemical formulas and critical temperatures, from the SuperCon database[1]. After eliminating duplicates and non-superconductors, we retained 11,949 superconducting materials. Over 200,000 ordered and disordered crystal structures were collected from the ICSD database. We then matched the 11,949 SuperCon entries with 208,425 ICSD entries based on chemical composition, space group and lattice parameter. Specifically, we first performed an initial matching based on chemical composition, which may result in one-to-one or one-to-many matches. We then further refined the matches using additional information provided in the literature, such as space groups and lattice constants. Additionally, T_c values and structural data for hydrogen-enriched superconductors were obtained from literature sources. Ultimately, we compiled 1,578 superconductors with both T_c and crystal structure information.
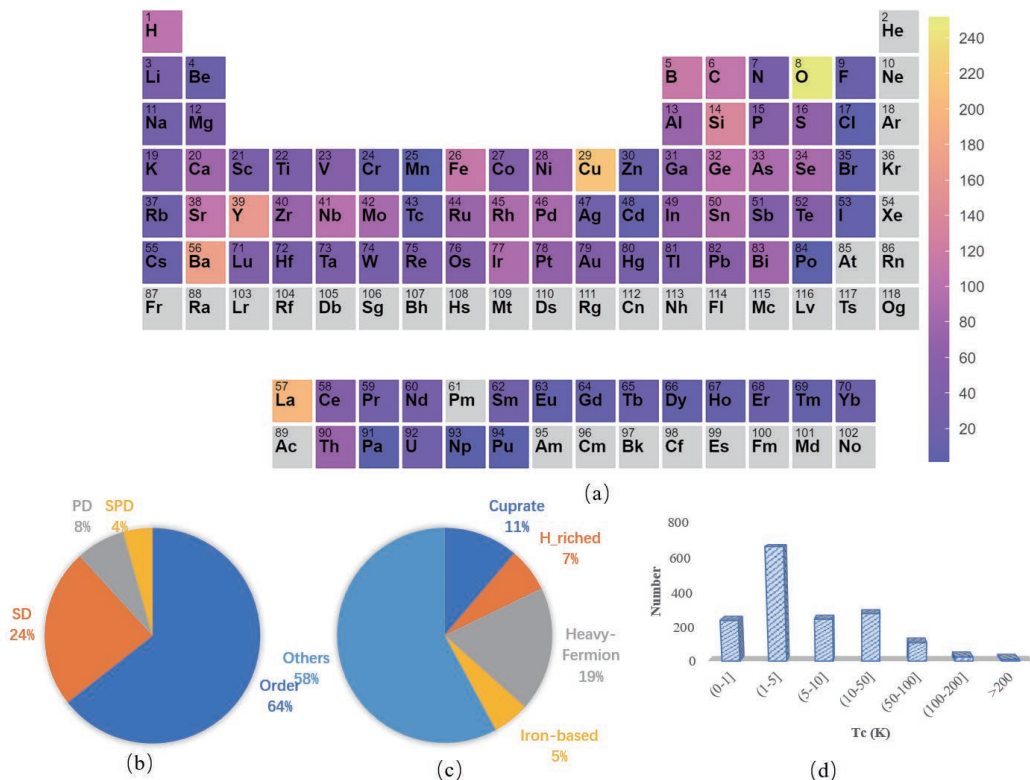
Figure 4: The data distribution of SuperCon3D dataset. (a). The probability of crystals containing a given element in the dataset. (b). The distribution of ordered and disordered superconductors. (c). The distribution of superconducting types. (d). The distribution of T_c values.

We plot the data distribution of SuperCon3D dataset in Fig. 4. In dataset, there are 83 different elements, which encompass most of elemental types found in the periodic table. The most frequent

elements are O, Cu, La, Ba, Y as shown in Fig. 4a. Fig. 4b depicts the order and disorder distribution. We classify superconducting materials according to cuprate, H riched, heavy fermion, iron based, and others, and distribute the types in Fig. 4c. The distribution of the Tc values of superconducting materials is shown in Fig. 4d. The SuperCon3D dataset can be obtained from the source code package, and the access address is proveide in Sec. 7.

2 Methods

2.1 Interstitial disorder

Interstitial disorder (ID). ID refers to the presence of atoms occupying interstitial sites within a crystal lattice, which are not part of the regular lattice positions. These interstitial atoms introduce additional disorder into the structure. The total occupancy, including both regular lattice sites and interstitial sites, can be expressed as:

$$\boldsymbol{w}_{i,1} + \boldsymbol{w}_{i,2} + \cdots + \boldsymbol{w}_{i,m_i} + \boldsymbol{w}_{i,interstitial} = 1 + \Delta \tag{1}$$

where $\boldsymbol{w}_{i,m_i}$ represents the occupancy weight of m_i at site i, $\boldsymbol{w}_{i,interstitial}$ represents the occupancy weight of interstitial atoms at site i, Δ is the excess occupancy due to interstitial atoms, with $\Delta > 0$ indicating the presence of ID. In this case, our disordered graph encoding method remains effective.

ID mixed with substitutional disorder (SD) and positional disorder (PD), would result in more new types. However, given the lack of observation of ID in the SuperCon3D dataset, we will not elaborate further on it.

2.2 Gate layer

We employ the gate activation mechanism Weiler et al. [2018] for the equivariant activation function. Standard activation functions are applied to type-0 vectors. For higher order vectors (L >0), we achieve equivariance by multiplying them with non-linearly transformed type-0 vectors. Specifically, for an input x comprising non-scalar C_L type-L vectors (where $0 < L \leq L_{\max}$) and $(C_0 + P_L \sum_{L=1}^{L_{\max}} C_L)$ type-0 vectors, we apply SiLU Elfwing et al. [2018] to the first C_0 type-0 vectors and a sigmoid function to the remaining $P_L \sum_{L=1}^{L_{\max}} C_L$ type-0 vectors. This process generates non-linear weights, which are then used to scale each type-L vector. After gate activation, the number of channels for type-0 vectors is reduced to C_0.

2.3 The denoising method of DiffCSP

We introduce the denoising model $\phi(\boldsymbol{L}, \boldsymbol{F}, \boldsymbol{A}, t)$ as part of the original DiffCSP model, which is related to the *Transformer-based Architecture* section in the main text.

Node representations in the s-th layer, $\boldsymbol{H}^{(s)} = [\boldsymbol{h}_1^{(s)}, \cdots, \boldsymbol{h}_N^{(s)}]$, are initialized as $\boldsymbol{h}_i^{(0)} = \rho(f_{\text{atom}}(\boldsymbol{a}_i), f_{\text{pos}}(t))$, combining atomic embeddings f_{atom} and sinusoidal positional encoding f_{pos} Ho et al. [2020], Vaswani et al. [2017], processed by MLP ρ.

Incorporating EGNN Satorras et al. [2021], the message-passing in layer s is:

$$\boldsymbol{m}_{ij}^{(s)} = \varphi_m(\boldsymbol{h}_i^{(s-1)}, \boldsymbol{h}_j^{(s-1)}, \boldsymbol{L}^\top \boldsymbol{L}, \psi_{\text{FT}}(\boldsymbol{f}_j - \boldsymbol{f}_i)), \tag{2}$$

$$\boldsymbol{m}_i^{(s)} = \sum_{j=1}^N \boldsymbol{m}_{ij}^{(s)}, \tag{3}$$

$$\boldsymbol{h}_i^{(s)} = \boldsymbol{h}_i^{(s-1)} + \varphi_h(\boldsymbol{h}_i^{(s-1)}, \boldsymbol{m}_i^{(s)}). \tag{4}$$

Here, φ_m and φ_h are MLPs. ψ_{FT} executes Fourier Transformation on relative fractional coordinates, ensuring periodic translation invariance.

Following S message-passing layers, lattice noise $\hat{\epsilon}_L$ is computed as follows:

$$\hat{\epsilon}L = L\varphi_L\left(\frac{1}{N}\sum i = 1^N h_i^{(S)}\right),\tag{5}$$

with φ_L shaping output as 3×3. For fractional coordinate score $\hat{\epsilon}_F$, we have:

$$\hat{\epsilon}_F[:,i] = \varphi_F(h_i^{(S)}),\tag{6}$$

where $\hat{\epsilon}_F[:,i]$ is the i-th column, and φ_F operates on the final layer's output.

The inner product $L^\top L$ in Eq.(2) ensures O(3)-invariance, as $(QL)^\top(QL) = L^\top L$ for any orthogonal $Q \in \mathbb{R}^{3\times 3}$. This guarantees the O(3)-invariance of φ_L in Eq.(6), and L left-multiplied with φ_L ensures O(3)-equivariance of $\hat{\epsilon}_L$. Thus, $\phi(L, F, A, t)$ satisfies the proposed properties. More details are described in Jiao et al. [2023].

3 Hyper-parameters and training details

In this section, we provide the training details of property predicting models and generative models.

3.1 Property predicting models

We employ the codebase from SchNet Schütt et al. [2018][2], CGCNN Xie and Grossman [2018][3], DimNet++ Gasteiger et al. [2020][4], SphereNet Liu et al. [2022][5],ALIGNN Choudhary and DeCost [2021][6], Matformer Yan et al. [2022][7] and MEGNet Chen et al. [2019][8] for baseline implementations. All models are conducted 10-fold experiments based data splited method of 8:1:1. The training details of each model are as follows:

3.1.1 SchNet.

Employing the SchNet framework, our method integrates six 64-dimensional message passing layers. SchNet was trained over 500 epochs, using a 5e-4 learning rate and 64 batch size. We optimized using Adam with 1e-5 weight decay, and a one-cycle learning rate scheduler. Atomic radii were determined by the 12th smallest distance between an atom and its neighbors.

3.1.2 CGCNN.

A batch size of 64 is employed, and the model consists of three layers of CGCNN message passing layer with 128 hidden dimensions. The training process utilizes the Adam optimizer. Initially, a learning rate of 1e-3 is set for the 200 epochs. During the training, a radius cutoff of 8.0 is applied to all crystals, and the 32 nearest neighbors are selected.

3.1.3 DimNet++.

In our approach, we apply a radius cutoff of 8.0 to all crystals and select the 12 nearest neighbors. To represent each node, we utilize Gaussian radial basis function (RBF) kernels. This results in a 64-dimensional embedding for each node. To optimize the model, we employ the Adam optimizer with a weight decay of 1e-6. The model is trained for 500 epochs using a batch size of 128.

---

[2]https://github.com/atomistic-machine-learning/SchNet
[3]https://github.com/txie-93/cgcnn
[4]https://github.com/gasteigerjo/dimenet
[5]https://github.com/divelab/DIG
[6]https://github.com/usnistgov/alignn
[7]https://github.com/YKQ98/Matformer
[8]https://github.com/materialsvirtuallab/megnet

3.1.4 SphereNet.

In our method, we utilize multi-graph representations of materials as inputs to SphereNet models. The input embedding size is set to 256, and the output embedding size is set to 64 for both the 8 LB2 and LB blocks. A cutoff distance of 6 is used. For each model, we initially perform a warm-up on the learning rate, starting at 1e-3. Subsequently, two learning rate strategies—ReduceLROnPlateau and StepLR—are employed for training. In the StepLR strategy, the learning rate is decayed by a specified ratio every fixed number of epochs, known as the step size. The batch size is set to 32, and training is conducted for 300 epochs.

3.1.5 ALIGNN.

ALIGNN is trained for 150 epochs with a learning rate of 5e-4 and a batch size of 64. The model architecture follows the original paper, consisting of four GCN layers and four ALIGNN layers. The atom feature dimension is set to 92, and the edge feature dimension is set to 80. The training process utilize the Adam optimizer with a weight decay of 1e-5. Additionally, a one-cycle learning rate scheduler is employed. For all crystals, a radius cutoff of 8.0 is applied, and the nearest 12 neighbors are selected.

3.1.6 Matformer.

In constructing the crystal graph, we follow a specific procedure. The radius for the neighborhood of a given atom is determined by the 12-th smallest distance between that atom and its neighboring atoms. All atoms within this radius are considered part of the neighborhood for the given atom. Each node is then represented by mapping its atomic number to a 92-dimensional embedding using the CGCNN atomic embedding. This embedding is further transformed into a 128-dimensional vector through a linear transformation. Similarly, for each edge, we utilize a 128-dimensional embedding mapping of the Euclidean distance. This mapping is achieved by employing 128 radial basis function (RBF) kernels with centers ranging from 0.0 to 8.0. During the training process, we employ the Adam optimizer with a weight decay of 1e-5. Additionally, a one-cycle learning rate scheduler is utilized. A batch size of 64 is employed and trained for 150 epochs.

3.1.7 MEGNet.

To construct the crystal graph, we employ three layers of the MEGNET message passing with with 64,32,16 hidden units, and utilize the Set2Set readout function. Following the configuration described in the original paper, MEGNET is trained for 200 epochs using a batch size of 64 and a learning rate of 1e-3. The Adam optimizer with a weight decay of 1e-5 is used for optimization, and a one-cycle learning rate scheduler is implemented. A radius of 8.0 is set for all crystals.

Table 7: Hyper-parameters for SODNet.

Hyper-parameters	Value or description
Batch size	32, 64, 128
Number of epochs	150, 300
Number of attention heads	4, 8
Dropout rate	0.0, 0.1, 0.2
Cutoff radius ()	8, 12, 16
Number of radial bases	128
Number of transformer blocks	6
Weigh decay	0.5×10^{-3}, 1×10^{-3}

3.1.8 SODNet

During training, we use a batch size of 64 and trained the model for 150 epochs. A radius of 8.0 is applied to define the neighborhood of each crystal. We utilize 128 basis functions to capture the

Table 8: Recently discovered superconductors (not included in the training data).

No.	Material	Type	T_c^{exp} (K)
1	CaH_6 @172 GPa	Order	215 Ma et al. [2022]
2	Ti @248 GPa	Order	26 Zhang et al. [2022]
3	CsV_3Sb_5	Order	2.3 Ge et al. [2022]
4	$Cs(V_{0.93}Nb_{0.07})_3Sb_5$	SD	4.45 Li et al. [2022]
5	Zr_4Rh_2O	Order	3.73 Watanabe et al. [2023]
6	Zr4Pd2O	Order	2.73 Watanabe et al. [2023]
7	$LaFeSiO_{0.9}$	PD	10 Hansen et al. [2022]

features of the crystals. To control overfitting, a weight decay of 5e-3 is applied. The learning rate is set to 5e-5, with a minimum learning rate of 1e-6. We employ the AdamW optimizer for efficient optimization. The model architecture consisted of 6 Transformer blocks, each with 8 attention heads. This allowed the model to effectively capture the relationships and dependencies within the crystal structures. Irreps features consist of channels of vectors with degrees up to $L_{\max}$. We denote C_L type-L vectors as (C_L, L) and $C_{(L,p)}$ type-(L, p) vectors as $(C_{(L,p)}, L, p)$. Brackets denote concatenations of vectors. we set irreps features containing 512 type-0 vectors and 128 type-1 vectors, which can be expressed as $[(512, 0), (128, 1)]$.

Table 7 summarizes the hyper-parameters for the model.

Table 9: The predicted potential candidates of high-T_c cuprate and h-riched superconductors. Candidates of high confidence are marked in gray.

Type	ICSD code	Chemical formula	O/SD/PD	T_c (K)	Reported SC.
Cuprate	68675	$CuO_2Sr_{0.075}$	PD	93.42	CuO_2Sr 91K Takahashi et al. [1994]
	50774	$Ca_{0.779}CuO_2Y_{0.041}$	PD	65.70	
	50773	$Ca_{0.82}CuO_2$	PD	64.72	$CaCuO_2$ 89K Schön et al. [2001]
	68217	Ba_2CuO_3	O	59.89	$Ba_2CuO_{3.2}$ 70K Li et al. [2019]
	67394	Ba_2CuIO_2	O	43.80	-
H-riched	187375	ErH_3	O	193.03	-
	635802	GdH_3	O	143.19	-
	623739	$H_{2.57}Co_{0.14}U_{0.84}$	PD	136.76	-
	42009	$TbH_{2.25}$	SD	135.13	-
	424154	$H_6Mg_{1.02}Ti_{1.98}$	O	134.34	-
	230140	$Li_{0.14}Y_{0.86}H_{2.7}$	PD	125.94	-
	93250	YFe_2H_5	PD	125.00	-

3.2 Generative models

We apply the codebases from CDVAE Xie et al. [2022][9], SyMat Luo et al. [2024][10] and DiffCSP Jiao et al. [2023][11] for baseline implementations. All models are conducted experiments based data splited method of 6:2:2. For pretraining, we obtain crystal structures from the databases of Materials Project[12], Open Quantum Materials Database[13], Matgen[14], and ICSD[15]. Molecular crystals are excluded from the dataset. Subsequently, we perform deduplication on all crystal structures, resulting in approximately 1.14 million unique structures. The training specifics for each model are outlined below:

---

[9] https://github.com/txie-93/cdvae
[10] https://github.com/divelab/AIRS
[11] https://github.com/jiaor17/DiffCSP
[12] https://next-gen.materialsproject.org
[13] https://www.oqmd.org
[14] https://matgen.nscc-gz.cn
[15] https://icsd.products.fiz-karlsruhe.de/

3.2.1 CDVAE.

For CDVAE model, We replaced the original DimNet++ Gasteiger et al. [2020] with SODNet to ensure a fair comparison with other generation models. Regarding the decoder, we utilize the GemNet-T Gasteiger et al. [2021], which consists of 3 layers and 128 hidden states.

3.2.2 SyMat.

For the SyMat model, the property predictor employs SphereNet, which consists of four message-passing layers with a hidden size of 128. The VAE decoder utilizes MLP models composed of two linear layers with a ReLU activation function between them and a hidden size of 256. During training, we use a learning rate of 0.001, a batch size of 128, and run for 1,000 epochs. We assign different weights to various loss terms: 1.0 for atom type set size, 30.0 for atom types, 1.0 for the number of each atom type, and 10.0 for lattice items. Additionally, we apply a weight of 0.01 for the KL-divergence loss and 10.0 for the denoising score matching loss.

3.2.3 DiffCSP.

We employ a configuration of 6 layers with 512 hidden states for datasets other than specified ones. The dimension of the Fourier embedding is set to 256. To control the variance of the DDPM (Diffusion-Driven Probabilistic Modeling) process on L_t, we utilize the cosine scheduler with 0.008. Additionally, we use an exponential scheduler with $\sigma_1 = 0.005, \sigma_T = 0.5$ to control the noise scale of the score matching process on F_t. The diffusion step is set to 1000. Our model is trained for 1000 epochs, employing the same optimizer and learning rate scheduler as CDVAE.

3.2.4 DiffCSP-SC.

We utilize SODNet as the property predictor, and the parameter configuration aligns with Table 7. The parameters for the diffusion process also follow the original DiffCSP setup. The difference lies in the message passing layer, where we employ a transformer. Specifically, we use a 512-dimensional hidden state encoding and set the number of heads to 8.

3.3 Pre-training dataset

we pre-trained our model on approximately 1.14 million unique 3D crystals sourced from existing databases, including Project[16], OQMD[17], Matgen[18], and ICSD[19].

4 Potential superconductors

In this section, we initially validate our model using the T_c values of superconducting materials reported in recent literature, noting that these data points are not included in the SuperCon3D dataset. Subsequently, we present the potential superconducting materials using property prediction model based on SODNet and generative model based on DiffCSP-SC, respectively.

4.1 Real-world superconductors validation

As shown in Table 8, we have collected the structures of superconducting materials along with their corresponding Tc values, as reported in the latest literature over the past three years. This includes a total of seven superconducting materials with both ordered and disordered structures.

---

[16] https://next-gen.materialsproject.org
[17] https://www.oqmd.org
[18] https://matgen.nscc-gz.cn
[19] https://icsd.products.fiz-karlsruhe.de

Table 10: The predicted potential candidates of high-T_c heavy-fermion, iron-based and others superconductors. Candidates of high confidence are marked in gray.

Type	ICSD code	Chemical formula	O/SD/PD	T_c (K)	Reported SC.
Heavy-Fermion	168466	$LaMg_{12}$	O	23.83	-
	161141	$LaMg_{11.196}Al_{0.804}$	SD	21.13	-
	69897	$C_2Ce_{0.75}U_{0.25}$	PD	11.88	-
	647197	$Np_{1.1}Pu_{0.9}$	SD	11.75	-
	614236	$TmFe_4B$	O	10.81	-
Iron-based	427163	$Ba_{0.83}Fe_2Rb_{0.17}As_2$	SD	23.21	$Ba_{0.6}Fe_2Rb_{0.4}As_2$ 37.5K Peschke et al. [2014]
	188347	$BaFe_2As_2$	O	23.27	-
	39530	$FeCl_7Te$	O	19.57	-
	633401	$FeSb_{0.4}Te_{1.6}$	SD	16.83	-
	165523	$As_2Ba_{0.777}Fe_2-K_{0.126}Sn_{0.096}$	PD	15.55	-
Others	96031	$Ba_{1.1432}Co_{0.1429}-O_{3.0009}Rh_{0.8574}$	PD	202.12	-
	58639	$Ba_{0.515}Ca_{0.485}$	SD	160.95	-
	616160	$BaSr$	SD	123.51	-
	106111	$SrTl_2$	O	63.52	-
	428028	$Ge_{0.6}Sb_{0.27}Te$	SD	47.48	-

4.2 Screening based method

We apply our superconductivity predicting model for screening the entire ICSD database. Potential superconductors are show in Table 9 and 10. To elaborate on the candidates with high confidence, we provide the subsequent details:

1. $CuO_2Sr_{0.075}$ and $Ca_{0.82}CuO_2$ exhibit disordered structures, and their respective parent compounds demonstrate superconductivity Takahashi et al. [1994], Schön et al. [2001]. Consequently, these disordered structures are more likely to be superconducting materials as well.

2. $Ba_2CuO_{3.2}$ exhibits superconductivity with a Tc of 70K Li et al. [2019]. Its corresponding parent structure Ba_2CuO_3 may also be a superconductor, albeit with a comparatively lower probability.

3. $Ba_{0.83}Fe_2Rb_{0.17}As_2$ and $Ba_{0.6}Fe_2Rb_{0.4}As_2$ share the same parent structure and have closely related compositions. Given that $Ba_{0.6}Fe_2Rb_{0.4}As_2$ exhibits superconductivity with a Tc of 37.5K Peschke et al. [2014], it is highly likely that $Ba_{0.83}Fe_2Rb_{0.17}As_2$ is also a superconducting material.

4.3 Interpretability on SODNet

We attempt to interpret our SODNet predictor by determining which feature(s) a given model weighs most heavily when making the prediction. As shown in Fig. 5, we extract the node embedding of the whole graph in the last layer of SODNet, and present the contributions of each atom to T_c values. We can observe that the B sites contributes more significantly to the property of Tc compared to the Mg site in Fig. 5 (a-d). Moreover, conducting atomic doping and atomic translation on the cation Mg results in a decrease in T_c with 39.0 K $\rightarrow$ 38.4 K $\rightarrow$ 34.3 K. This phenomenon demonstrates that attempting to enhance the T_c value by disrupting the symmetry of Mg site within the lattice may be not workable. Another case of cuprate superconductor has shown in Fig. 5 (e-f), there are three types of oxygen sites that contribute significantly to the T_c value: Hg-O-Hg (PD disorder), Cu-O-Cu (order), and Hg-O-Ba (order). Among them, the contribution of disordered Hg-O-Hg is the greatest, indicating that disrupting the symmetry of oxygen atoms within the lattice might potentially further enhance the property of T_c.

Table 11: The novel high-T_c cuprate and h-riched superconducting candidates. Candidates of high confidence are marked in gray.

| Type	Index	Chemical formula	T_c (K)	Reported SC.
Cuprate	1	$Ba_2CuCl_2O_2$	33.56	-
	2	$Tl_2Ca_2Ba_2Cu_3O_{10}$	14.09	-
	3	$Ba_3CaLa_2GdCu_7O_{17}$	10.12	-
	4	YCu_3O_7	9.73	-
	5	$BaCaCu_3O_7$	9.65	-
	6	Cu_7BO_{16}	7.87	-
	7	$CsMgCu_3BiAuO_8$	7.82	-
H-riched	8	TbH_3	164.33	TbH_3 20K Hai et al. [2021] Calculated by DFT
	9	SeH_3	139.89	SeH_3 113K Novakovic et al. [2023] Predicted by ML
	10	$CaGe_2H_9$	103.55	-
	11	Ca_2MnCrH_6	58.07	-
	12	SbH_3	46.42	SbH_3 20K Fu et al. [2016] Calculated by DFT
	13	$MgCoCuH_{42}CS_2N_{16}$	44.27	-
	14	$Rb_2Ca_2H_4$	13.05	-
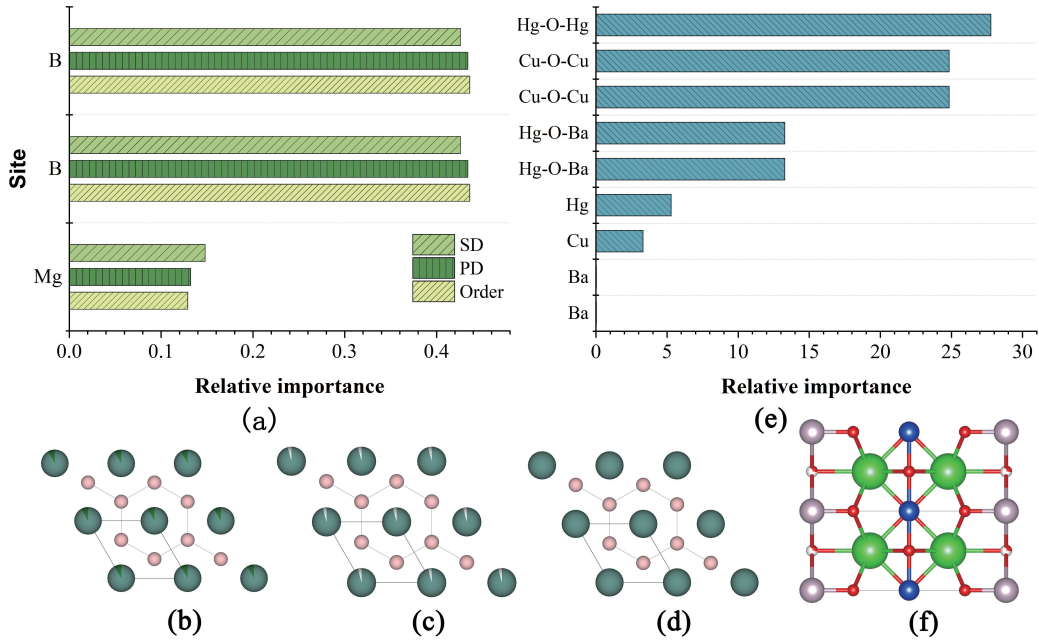
Figure 5: Contribution of each atom to T_c value. (a). Feature relative importance of each site in three type MgB_2 superconductors. Snapshots of (b) SD $Mg_{0.9}Al_{0.1}B_2$, (c) PD $Mg_{0.98}B_2$, and (d) ordered MgB_2 crystals. Here, B Mg and Al sites are colored by light pink, atrovirens and dark atrovirens. (e). Feature relative importance of $Ba_2CuHgO_{4.27}$ superconductor. (f). Snapshot of $Ba_2CuHgO_{4.27}$ superconductor (Ba: green, Cu: blue, Hg: pink, O: red).

4.4 Generative superconducting candidates

We apply our generative model for generating new superconducting candidates. We present the crystal structures of the 20 superconducting candidate materials from Table 6 in Fig. 6. Additionally, we display the 32 superconducting candidate materials in Table 11 and 12, arranged in descending order of predicted T_c values. The structures of all superconducting candidate materials can be obtained in the source code package. We collected superconducting materials that have been reported and observed that five candidates are more likely to be superconducting materials. Among them, four candidates obtained Tc through theoretical calculations, and another material displayed superconducting properties through doping. Specific descriptions are as follows:

1. SeH_3 exhibited a Tc of 113K as predicted by machine learning Novakovic et al. [2023], corroborated by DFT calculations indicating 110K Zhang et al. [2015].

2. DFT methods calculated the T_c of TbH_3, SbH_3, and KFe_2As_2 as 20K Hai et al. [2021], Fu et al. [2016], Ptok et al. [2020]. Since H-riched materials belong to conventional superconductors and show high Tc under high pressure, but the conditions for wet experimental synthesis are very stringent. Therefore, it can further verify whether superconducting materials are superconducting materials by combining DFT methods, and reduce the research and development cycle of superconducting materials.

3. The parent compound SmFeAsO underwent a superconducting Tc around 54 K Azam et al. [2023], following fluorine (F) doping at the O-site in the SmO layer. This case can provide us with a method that we can use DiffCSP-SC's generative model to generate superconducting parent structures, and then improve the Tc of materials by doping, or transform materials without superconducting properties into superconducting materials.

Table 12: The novel high-T_c heavy-fermion, iron-based and others superconducting candidates. Candidates of high confidence are marked in gray.

Type	Index	Chemical formula	T_c (K)	Reported SC.
Heavy-Fermion	15	Th	43.61	-
	16	Ba_3Pu	44.81	-
	17	ThC_3	17.96	-
	18	Lu	4.86	-
	19	Yb3In	1.04	-
Iron-based	20	$BaFe_2Se_2$	11.99	-
	21	SmFeAsO	4.42	$SmFeAsO_{0.8}F0.2$ 54K Azam et al. [2023]
	22	KFe_2As_2	4.23	KFe_2As_2@30GPa 20K Ptok et al. [2020] Calculated by DFT
	23	NdFeAsF	4.13	-
	24	FeSe	3.36	-
Others	25	Ba_3Ca	80.04	-
	26	Ba_2Se	60.70	-
	27	Ba	52.26	-
	28	Mg_3B	43.96	-
	29	$BaCl_2O$	35.72	-
	30	Ba_2CaB	32.77	-
	31	Sb_2Ba_4	22.70	-
	32	V_3Si_{11}	16.28	-

4.5 DFT calculations

We conduct DFT calculation using the Vienna ab initio package (VASP)Wang and Pickett [1983], Chan and Ceder [2010]. The structures are fully relaxed using the generalized gradient approximation (GGA)Perdew et al. [1996] of the SCAN meta-GGA functional, employing the pseudopotentials of the projector augmented wave (PAW) methodBlöchl [1994]. A plane wave cutoff of 500 eV is employed for all simulations. Brillouin-zone integrations are performed using the τ-centered Monkhorst-Pack (MP) schemeMonkhorst and Pack [1976] . We initiate the calculations with a
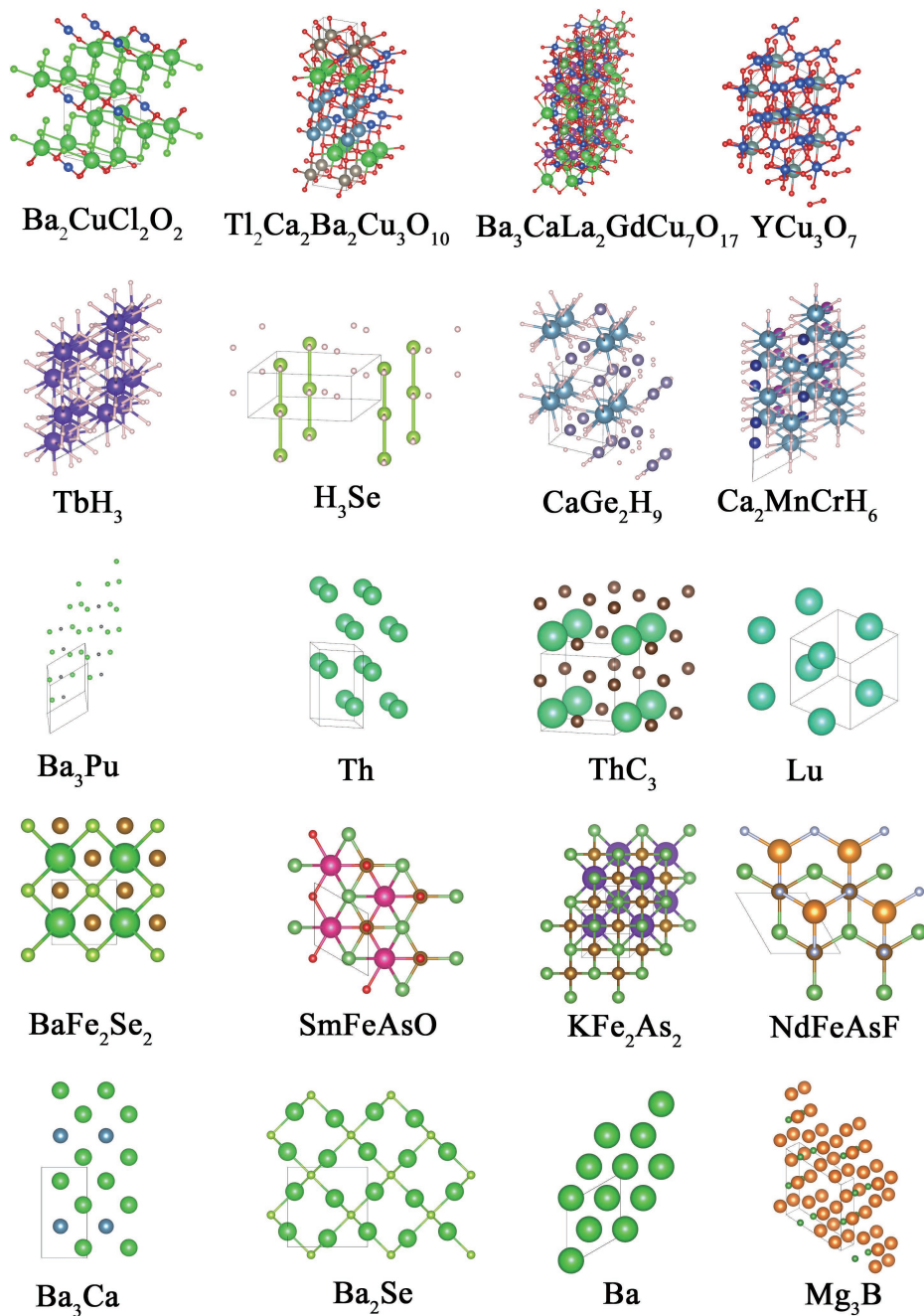
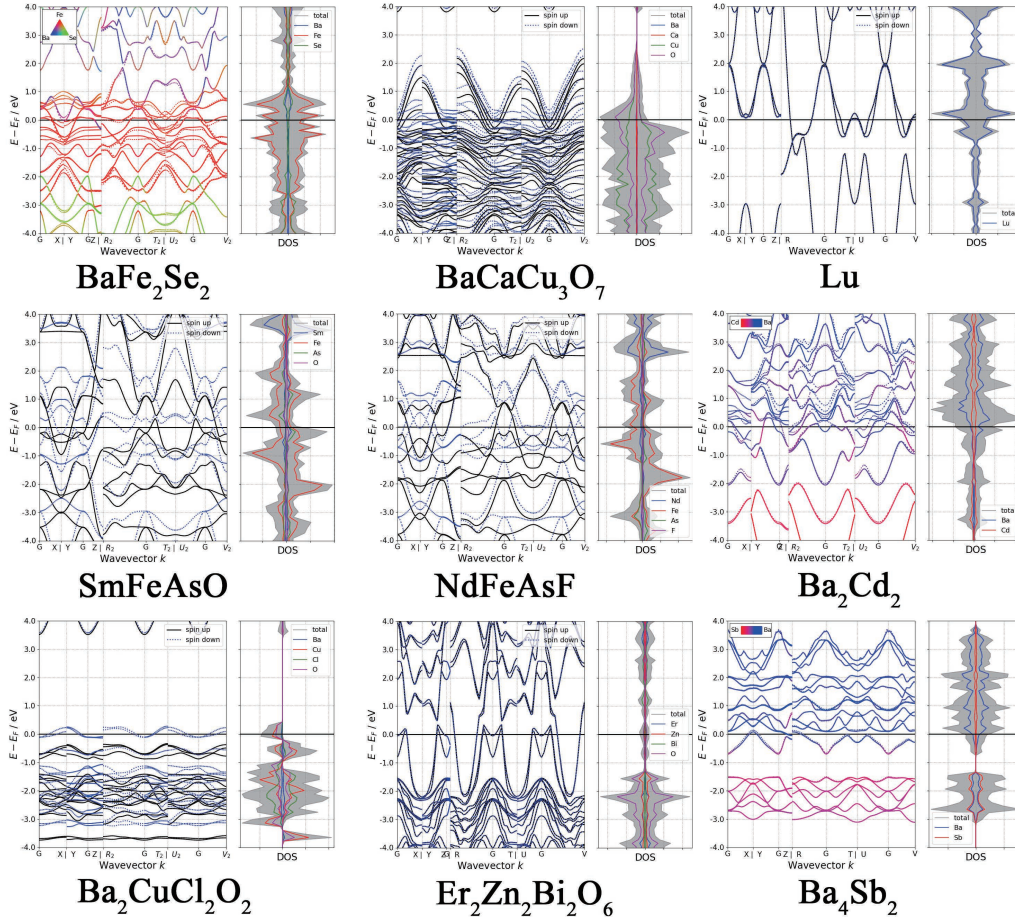Figure 6: The geometric structures of novel superconducting candidates in Table 6.

Figure 7: The electronic structures of novel superconducting candidates.

k-point meth featuring a dense sampling density of $2\pi \times 0.04$. The convergence criteria for energy and force is set to 0.1 meV and 0.001 eV/Å, respectively.

The van Hove singularity (VHS) is a notable occurrence in condensed matter physics, specifically in the density of states (DOS) of a material. It manifests as a distinct peak or divergence in the DOS at a particular energy level. We select materials from Table 11 and 12 for DFT calculations and display their band structures and density of states (DOS) in Fig. 7. From the density of states (DOS) plot, we can observe the van Hove singularity (VHS) phenomenon. Additionally, we can also observe the presence of flat bands in the band structures of materials such as $Ba_2CuCl_2O_2$, Lu, Ba_4Sb_2, and others. The integration of flat bands in the electronic architecture, along with the Van Hove Singularities (VHS) in the Density of States (DOS), markedly amplifies the likelihood of these candidates being superconducting materials.

5 Pipeline for designing real-world superconductors.

Fig. 8 presents a pipeline for designing SC., validating our dataset and models for real-world scenarios. We initially generate potential, ordered superconducting structures using the DiffCSP-SC model trained on the SuperCon3D database. Candidate materials are selected based on Tc values predicted by SODNet, followed by DFT verification to confirm the presence of superconducting electronic structures, such as VHS. Subsequently, selected candidates undergo wet lab synthesis, with T_c values characterized and recorded in the SuperCon3D database. Further, if a superconductor is

13

discovered, methods such as doping, which may transform ordered structures into disordered ones, are explored to enhance the T_c value. SODNet is employed to investigate the relationship between disordered structures and doping ratios, aiming to design optimal doping proportions for experimental verification. These experimental outcomes are also recorded in the database. Continuous expansion of the database will incrementally improve the accuracy of the DiffCSP-SC and SODNet models trained on this dataset, creating a reinforcing cycle of enhancement.
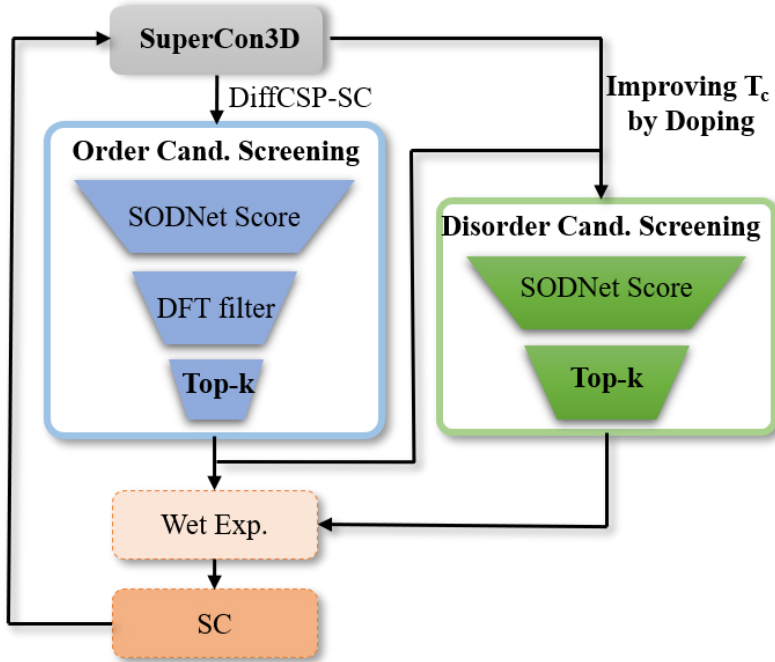
Figure 8: Flowchart for designing novel SC materials.

6 Limitations

Our SuperCon3D dataset, featuring experimental structures and T_c values, paves the way for real-world superconductor applications. Combined with SODNet, which addresses disordered graph issues previously overlooked by the AI community, and DiffCSP-SC for novel designs. However, the accuracy of data-driven models remains constrained by the collected superconducting dataset. As Figure 4 in the Appendix shows, data unevenness and elemental skewness (especially in Cu and O) may bias the model. Additionally, as Table 8 indicates, atomic distributions under extreme pressures contribute to predictive errors. Addressing these, Figure 8 presents our pipeline, combining DiffCSP-SC and SODNet, to design and validate novel superconductors through wet experiments, iteratively enriching the dataset for improved model training and accuracy.

7 Code

We have provided the source code of SODNet and DiffCSP-SC, which are mentioned in this article, on an anonymous GitHub repository. The access address is as follows: https://anonymous.4open.science/r/SODNet-F569, https://anonymous.4open.science/r/DiffCSP-SC-8F3F.

References

Maurice Weiler, Mario Geiger, Max Welling, Wouter Boomsma, and Taco Cohen. 3d steerable cnns: Learning rotationally equivariant features in volumetric data. In Samy Bengio, Hanna M. Wallach, Hugo Larochelle, Kristen Grauman, Nicolò Cesa-Bianchi, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, December 3-8, 2018, Montréal, Canada*, pages 10402–10413, 2018. URL `https://proceedings.neurips.cc/paper/2018/hash/488e4104520c6aab692863cc1dba45af-Abstract.html`.

Stefan Elfwing, Eiji Uchibe, and Kenji Doya. Sigmoid-weighted linear units for neural network function approximation in reinforcement learning. *Neural Networks*, 107:3–11, 2018. ISSN 0893-6080. doi: https://doi.org/10.1016/j.neunet.2017.12.012. URL `https://www.sciencedirect.com/science/article/pii/S0893608017302976`. Special issue on deep reinforcement learning.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33:6840–6851, 2020.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

Víctor Garcia Satorras, Emiel Hoogeboom, and Max Welling. E (n) equivariant graph neural networks. In *International Conference on Machine Learning*, pages 9323–9332. PMLR, 2021.

Rui Jiao, Wenbing Huang, Peijia Lin, Jiaqi Han, Pin Chen, Yutong Lu, and Yang Liu. Crystal structure prediction by joint equivariant diffusion on lattices and fractional coordinates. In *Workshop on "Machine Learning for Materials" ICLR 2023*, 2023. URL `https://openreview.net/forum?id=VPByphdu24j`.

Kristof T Schütt, Huziel E Sauceda, P-J Kindermans, Alexandre Tkatchenko, and K-R Müller. Schnet–a deep learning architecture for molecules and materials. *The Journal of Chemical Physics*, 148(24):241722, 2018.

Tian Xie and Jeffrey C. Grossman. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Physical Review Letters*, 120(14):145301, 2018. doi: 10.1103/PhysRevLett.120.145301. URL `https://link.aps.org/doi/10.1103/PhysRevLett.120.145301`.

Johannes Gasteiger, Shankari Giri, Johannes T. Margraf, and Stephan Günnemann. Fast and uncertainty-aware directional message passing for non-equilibrium molecules. In *Machine Learning for Molecules Workshop, NeurIPS*, 2020.

Yi Liu, Limei Wang, Meng Liu, Yuchao Lin, Xuan Zhang, Bora Oztekin, and Shuiwang Ji. Spherical message passing for 3d molecular graphs. In *International Conference on Learning Representations (ICLR)*, 2022.

Kamal Choudhary and Brian DeCost. Atomistic line graph neural network for improved materials property predictions. *npj Computational Materials*, 7(1):1–8, 2021.

Keqiang Yan, Yi Liu, Yu-Ching Lin, and Shuiwang Ji. Periodic graph transformers for crystal material property prediction. *ArXiv*, abs/2209.11807, 2022.

Chi Chen, Weike Ye, Yunxing Zuo, Chen Zheng, and Shyue Ping Ong. Graph networks as a universal machine learning framework for molecules and crystals. *Chemistry of Materials*, 31 (9):3564–3572, 2019. ISSN 0897-4756. doi: 10.1021/acs.chemmater.9b01294. URL `https://doi.org/10.1021/acs.chemmater.9b01294`.

Liang Ma, Kui Wang, Yu Xie, Xin Yang, Yingying Wang, Mi Zhou, Hanyu Liu, Xiaohui Yu, Yongsheng Zhao, Hongbo Wang, et al. High-temperature superconducting phase in clathrate calcium hydride cah 6 up to 215 k at a pressure of 172 gpa. *Physical Review Letters*, 128(16): 167001, 2022.

Changling Zhang, Xin He, Chang Liu, Zhiwen Li, Ke Lu, Sijia Zhang, Shaomin Feng, Xiancheng Wang, Yi Peng, Youwen Long, et al. Record high t c element superconductivity achieved in titanium. *Nature Communications*, 13(1):5411, 2022.

Jun Ge, Pinyuan Wang, Ying Xing, Qiangwei Yin, Hechang Lei, Ziqiang Wang, and Jian Wang. Discovery of charge-4e and charge-6e superconductivity in kagome superconductor csv3sb5. *arXiv preprint arXiv:2201.10352*, 2022.

Yongkai Li, Qing Li, Xinwei Fan, Jinjin Liu, Qi Feng, Min Liu, Chunlei Wang, Jia-Xin Yin, Junxi Duan, Xiang Li, et al. Tuning the competition between superconductivity and charge order in the kagome superconductor cs (v 1- x nb x) 3 sb 5. *Physical Review B*, 105(18):L180507, 2022.

Yuto Watanabe, Akira Miura, Chikako Moriyoshi, Aichi Yamashita, and Yoshikazu Mizuguchi. Observation of superconductivity and enhanced upper critical field of η-carbide-type oxide zr4pd2o. *Scientific Reports*, 13(1):22458, 2023.

Mad F Hansen, J-B Vaney, Christophe Lepoittevin, F Bernardini, Etienne Gaudin, V Nassif, M-A Méasson, A Sulpice, H Mayaffre, M-H Julien, et al. Superconductivity in the crystallogenide lafesio1- δ with squeezed fesi layers. *npj Quantum Materials*, 7(1):1–8, 2022.

H Takahashi, N Môri, M Azuma, Z Hiroi, and M Takano. Effect of pressure on tc of hole-and electron-doped infinite-layer compounds up to 8 gpa. *Physica C: Superconductivity*, 227(3-4): 395–398, 1994.

JH Schön, M Dorget, FC Beuran, XZ Zu, E Arushanov, C Deville Cavellin, and M Lagues. Super-conductivity in cacuo2 as a result of field-effect doping. *Nature*, 414(6862):434–436, 2001.

WM Li, JF Zhao, LP Cao, Z Hu, QZ Huang, XC Wang, Y Liu, GQ Zhao, J Zhang, QQ Liu, et al. Superconductivity in a unique type of copper oxide. *Proceedings of the National Academy of Sciences*, 116(25):12156–12160, 2019.

Tian Xie, Xiang Fu, Octavian-Eugen Ganea, Regina Barzilay, and Tommi S. Jaakkola. Crystal diffusion variational autoencoder for periodic material generation. In *The Tenth International Conference on Learning Representations, ICLR 2022, Virtual Event, April 25-29, 2022*. OpenReview.net, 2022. URL `https://openreview.net/forum?id=03RLpj-tc_`.

Youzhi Luo, Chengkai Liu, and Shuiwang Ji. Towards symmetry-aware generation of periodic materials. *Advances in Neural Information Processing Systems*, 36, 2024.

Johannes Gasteiger, Florian Becker, and Stephan Günnemann. Gemnet: Universal directional graph neural networks for molecules. In *Conference on Neural Information Processing Systems (NeurIPS)*, 2021.

Simon Peschke, Tobias Stürzer, and Dirk Johrendt. Ba1–xrbxfe2as2 and generic phase behavior of hole-doped 122-type superconductors. *Zeitschrift für anorganische und allgemeine Chemie*, 640 (5):830–835, 2014.

Yu-Long Hai, Ning Lu, Hui-Li Tian, Meng-Jing Jiang, Wei Yang, Wen-Jie Li, Xun-Wang Yan, Chao Zhang, Xiao-Jia Chen, and Guo-Hua Zhong. Cage structure and near room-temperature superconductivity in tbh n (n= 1–12). *The Journal of Physical Chemistry C*, 125(6):3640–3649, 2021.

Lazar Novakovic, Ashkan Salamat, and Keith V Lawler. Machine learning using structural repre-sentations for discovery of high temperature superconductors. *arXiv preprint arXiv:2301.10474*, 2023.

Yuhao Fu, Xiangpo Du, Lijun Zhang, Feng Peng, Miao Zhang, Chris J Pickard, Richard J Needs, David J Singh, Weitao Zheng, and Yanming Ma. High-pressure phase stability and superconductivity of pnictogen hydrides and chemical trends for compressed hydrides. *Chemistry of Materials*, 28(6):1746–1755, 2016.

Shoutao Zhang, Yanchao Wang, Jurong Zhang, Hanyu Liu, Xin Zhong, Hai-Feng Song, Guochun Yang, Lijun Zhang, and Yanming Ma. Phase diagram and high-temperature superconductivity of compressed selenium hydrides. *Scientific reports*, 5(1):15433, 2015.

Andrzej Ptok, Konrad Jerzy Kapcia, Małgorzata Sternik, and Przemysław Piekarz. Superconductivity of kfe2as2 under pressure: Ab initio study of tetragonal and collapsed tetragonal phases. *Journal of Superconductivity and Novel Magnetism*, 33(8):2347–2354, 2020.

Mohammad Azam, Manasa Manasa, Tatiana Zajarniuk, Ryszard Diduszko, Tomasz Cetner, Andrzej Morawski, Andrzej Wiśniewski, and Shiv J Singh. Antimony doping effect on the superconducting properties of smfeas (o, f). *IEEE Transactions on Applied Superconductivity*, 2023.

CS Wang and WE Pickett. Density-functional theory of excitation spectra of semiconductors: application to si. *Physical review letters*, 51(7):597, 1983.

MKY Chan and Gerbrand Ceder. Efficient band gap prediction for solids. *Physical review letters*, 105(19):196403, 2010.

John P Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Physical review letters*, 77(18):3865, 1996.

Peter E Blöchl. Projector augmented-wave method. *Physical review B*, 50(24):17953, 1994.

Hendrik J Monkhorst and James D Pack. Special points for brillouin-zone integrations. *Physical review B*, 13(12):5188, 1976.