FORWARD AND INVERSE DESIGN OF HIGH T_C SUPERCONDUCTORS WITH DFT AND DEEP LEARNING

Daniel Wines, Kevin F. Garrity & Kamal Choudhary*
Materials Science and Engineering Division
National Institute of Standards and Technology (NIST)
Gaithersburg, MD 20899, USA
{daniel.wines,kevin.garrity,kamal.choudhary}@nist.gov

Tian Xie
Microsoft Research AI4Science
Cambridge, United Kingdom CB1 2FB
tianxie@microsoft.com

ABSTRACT

We developed a multi-step workflow for the discovery of next-generation conventional superconductors. 1) We started with a Bardeen–Cooper–Schrieffer (BCS) inspired pre-screening of 55000 materials in the Joint Automated Repository for Various Integrated Simulations (JARVIS) density functional theory (DFT) database resulting in 1736 materials with high Debye temperature and electronic density of states at the Fermi-level. 2) Then, we performed DFT based electron-phonon coupling calculations for 1058 materials to establish a systematic database of superconducting properties. 3) Further, we applied forward deep-learning (DL) using atomistic line graph neural network (ALIGNN) models to predict properties faster than direct first-principles computations. Notably, we find that by predicting the Eliashberg function as an intermediate quantity, we can improve the model performance versus a direct DL prediction of T_C. Finally, 4) we used an inverse deep-learning method with a crystal diffusion variational autoencoder (CDVAE) model to generate thousands of new superconductors with high chemical and structural diversity. 5) We screened these CDVAE-generated structures using ALIGNN to identify candidates that are stable with high T_C. 6) We verified the top superconducting candidates with DFT.

1 INTRODUCTION

Since the discovery of superconductivity in 1911 by Onnes Kamerlingh Onnes (1911), the identification of novel superconducting materials with high transition temperatures (T_C) has been an active area of research in condensed matter physics Poole et al. (2013); Rogalla & Kes (2011). A data-driven search can assist in expediting the discovery of potentially high-T_C superconductors. There are two key ingredients required to computationally identify Bardeen–Cooper–Schrieffer (BCS) conventional superconductors Cooper & Feldman (2010); Giustino (2017) with high-T_C: 1) a robust computational workflow, and 2) a database of curated materials with prior knowledge such as elastic constants and electronic density of states. Using density functional theory perturbation theory (DFT-PT), the electron-phonon coupling (EPC) can be calculated and used to predict T_C with reasonable accuracy for a variety of materials Giustino (2017); Kawamura et al. (2020).

In this work, we developed such a computational approach to discover BCS superconductors, combining several methods at various levels of computational expense and accuracy. We start with a BCS-inspired pre-screening based on materials with high Debye temperature (θ_D) and high electron density of states (DOS) at Fermi-level ($N(0)$), using the existing Joint Automated Repository for Various Integrated Simulations (JARVIS) DFT database Choudhary et al. (2020c). We then develop

*Theiss Research, La Jolla, CA, 92037, USA

and apply a DFT-PT workflow to compute T_C using electron-phonon coupling and the McMillan-Allen-Dynes formula McMillan (1968), with initially low k-point and q-point convergence settings. For the best candidates, we perform additional convergence.

Going a step further, we used the data obtained from 1058 DFT-PT calculations of EPC properties to train deep learning models for the forward and inverse design of new superconductors. For forward design, we use the recently developed atomistic line graph neural network (ALIGNN) Choudhary & DeCost (2021) (`https://github.com/usnistgov/alignn`). We used ALIGNN to train models for θ_D and the DOS at the Fermi-level using datasets from the JARVIS-DFT database. We also train models to predict the DFT-based T_C and the underlying EPC parameters (ω_{log} and λ), using the smaller datasets computed in this work. ALIGNN and the hand-crafted (Classical force-field descriptor (CFID) Choudhary et al. (2018b))-based approaches are used for comparison. For inverse design, we used a crystal diffusion variational autoencoder (CDVAE) model (`https://github.com/txie-93/cdvae`) developed by Xie et al. (2021) which is trained on these 1058 DFT-PT calculations. The CDVAE combines a variational autoencoder Kingma & Welling (2013) and a diffusion model Song & Ermon (2019); Sohl-Dickstein et al. (2015) to generate new stable periodic structures. After using CDVAE to generate new superconductors, we screened these candidates with ALIGNN and performed additional DFT calculations on the best candidates.

2 METHODS

2.1 DFT CALCULATIONS

To investigate new superconductors and generate training data, we performed EPC calculations using DFT-PT Baroni et al. (1987); Gonze (1995) (using the interpolated/Gaussian broadening method Wierzbowska et al. (2005)) with the Quantum Espresso (QE) software package Giannozzi et al. (2009), PBEsol functional Perdew et al. (2008), and the GBRV Garrity et al. (2014) pseudopotentials. We begin with structures from the JARVIS-DFT database, and perform the full workflow in QE. It is important to note that the robustness of this workflow was heavily benchmarked against experimental data and higher levels of theory in ref. Choudhary & Garrity (2022) and Wines et al. (2023a), which indicates that the training data for these deep learning models is of high quality, given the level of theory used to produce the data.

2.2 ALIGNN MODEL

In ALIGNN, a crystal structure is represented as a graph using atomic elements as nodes and atomic bonds as edges. Each node in the atomistic graph is assigned 9 input node features based on its atomic species: electronegativity, group number, covalent radius, valence electrons, first ionization energy, electron affinity, block and atomic volume. The inter-atomic bond distances are used as edge features with radial basis function up to 8 Å cut-off. We use a periodic 12-nearest-neighbor (N) graph construction. This atomistic graph is then used for constructing the corresponding line graph using interatomic bond-distances as nodes and bond-angles as edge features. ALIGNN uses edge-gated graph convolution for updating nodes as well as edge features using a propagation function (f) for layer (l), atom features (h), and node (i):

$$h_i^{(l+1)} = f(h_i^l \{h_j^l\}_i) \tag{1}$$

ALIGNN uses bond-distances as well as bond-angles to distinguish atomic structures. One ALIGNN layer composes an edge-gated graph convolution on the bond graph with an edge-gated graph convolution on the line graph. The line graph convolution produces bond messages that are propagated to the atomistic graph, which further updates the bond features in combination with atom features. The hyperparameters for ALIGNN are kept same as the original paper. For the Debye temperature and electronic density of states at the Fermi-level, we use a batch size of 64 for 500 epochs and 80:10:10 training-validation-testing data split, while for predicting the EPC parameters and Eliashberg functions, we use a batch size of 16, 90:5:5 split and training for 300 epochs. The test set was never used during the training procedure.

2.3 CDVAE MODEL

In this work, we used a CDVAE model developed by Xie et al. (2021) The CDVAE combines a variational autoencoder Kingma & Welling (2013) and a diffusion model to generate new periodic structures, where crystals are represented by a tuple containing: atom types, atomic coordinates, and the basis vectors of the unit cell. The CDVAE consists of three networks that are trained jointly: 1) the encoder, 2) the property predictor, and 3) the decoder. The encoder is a SE(3) equivariant graph neural network, that encodes onto a lower dimensional latent space. From here, the property predictor predicts the number of atoms N, lattice vectors, and the chemical composition. The decoder, which is a noise conditional score network Song & Ermon (2019); Ho et al. (2020), then takes a structure with noise added to the coordinates and type of atoms and learns to denoise it into the original stable structure. This noise added to the type of atoms updates the element type for each atom into another element within the predicted composition with a certain probability (given by the noise level). Coordinate noise is Gaussian noise added to the coordinates of each atom of the crystal. The score of the conditional score network diffusion model is an estimate of the gradient of the probability distribution of the materials and is predicted by another SE(3) equivariant GNN. The decoder being an equivariant diffusion model makes it possible to work with atomic positions directly, without any need for intermediate representations such as graphs or descriptors. As a result, this makes the CDVAE model general to the kind of structures and chemical environments it is used for. More details of the CDVAE method can be found in Xie et al. (2021).

3 RESULTS AND DISCUSSION

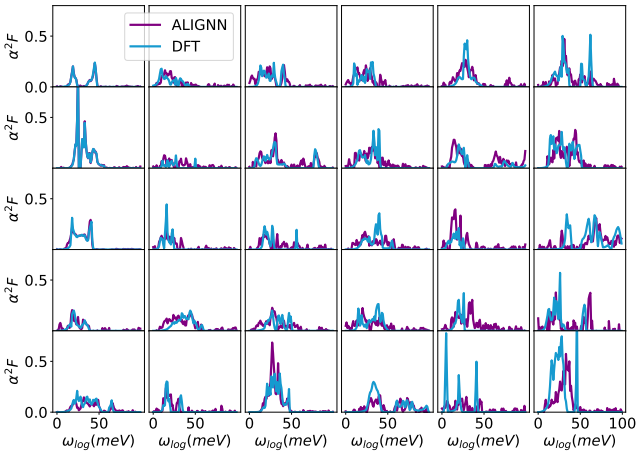

Figure 1: Prediction of Eliashberg function with ALIGNN for the 5 % test set. ALIGNN can capture peak-positions and heights reasonably well.

For the forward design of new superconductors, we developed deep-learning models to accelerate both our initial BCS-inspired screening and our calculation of the EPC parameters. While the BCS pre-screening step is less expensive than a full EPC calculation, it still requires the DOS at the Fermi level and θ_D, which still requires substantial computation. Therefore, we developed regression models to predict these properties directly from an arbitrary crystal structure, using the large datasets available in the JARVIS-DFT database. In this work we used the ALIGNN model, which has been shown to outperform many well-known benchmarks for solids and explicitly capture chemical and many-body physical interactions. Our results of these models on 5 % held test sets are shown in Fig. A.2. The baseline model MAE was computed by using mean of the target values in the training dataset and using it as predictions for all the materials in the 5% test data. We observe that the mean absolute error for the Debye temperature is 49 K while that for DOS is 1.5 states/eV/Nelect. The baseline model MAE for the Debye temperature and DOS are 145.5 K and 3.62 states/eV/Nelect

respectively. In addition, we developed machine-learning models to directly predict EPC properties using our database of 626 dynamically stable data-points from the explicit DFT calculations. Two methods were used: hand-craft descriptors (CFID) and a deep-learning approach (ALIGNN). Specifically, we trained models for the McMillan-Allen-Dynes transition temperature (T_C) and the EPC parameters (ω_{log} and λ). We note that usually ML models require larger datasets, but we display preliminary useful results with our current smaller dataset, which will continue to grow.

CFID based performances are shown in Fig. A.2c), Fig. A.2d) and Fig. A.2e) for T_C, ω_{log} and λ. Similarly, ALIGNN based performances are shown in Fig. Fig. A.2f), Fig. A.2g) and Fig. A.2h). We observe that the MAE using the CFID approach for T_C, ω_{log} and λ are 1.84 K, 53.48 K and 0.19. Additionally, the MAEs for the ALIGNN approach are 1.84 K, 37.43 K and 0.14 respectively. ALIGNN outperforms CFID in predicting ω_{log}, but the performances for the other quantities are similar. Moreover, we notice that it is significantly easier to learn ω_{log} than T_C and λ. We observe that the model for λ is only slightly better than the baseline model. Using ALIGNN for ω_{log} and λ, and Equation A.7, we predict T_C with an MAE of 1.77 K.

We attempt an alternate method to directly predict the Eliashberg function using the ALIGNN model. We choose an energy range of 0 meV to 100 meV with 1 meV bin size and predict the Eliashberg functions. In Fig. 1, we show the DFT and ALIGNN based Eliashberg functions for samples in the test set. We find that the ALIGNN model does a good job of capturing most of the peaks. We calculate the T_C using the ALIGNN-based Eliashberg function predictions, and find the MAE to be 1.39 K, which improves on our direct prediction method above by 24%. This implies that learning more fundamental and information-rich quantities such as Eliashberg functions can be useful for ML approaches with limited data, as compared to direct predictions of integrated quantities.

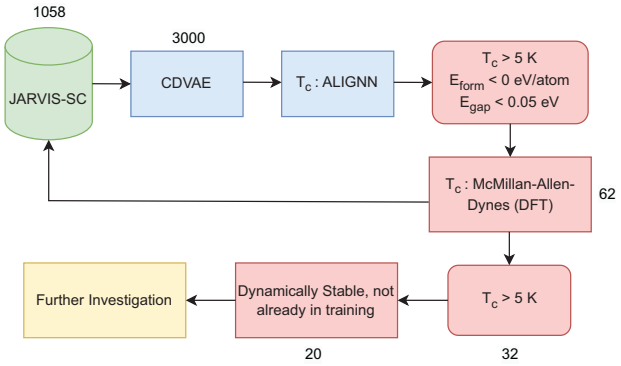

Figure 2: The full inverse design workflow for new superconductors using DFT, ALIGNN and the CDVAE generative model.

For the inverse design of new superconductors, we used our database of superconducting properties along with a CDVAE generative model to predict unknown stable superconductors with high T_C. A full schematic of the inverse design workflow is depicted in Fig. 2. After training on the 1058 structures in the JARVIS superconductor database (JARVIS-SC), 3000 structures are created using the CDVAE model, with the target property being a high T_C. An analysis of the distribution of the 3000 CDVAE created structures, including prototypes (chemical formula), densities and packing fractions, indicates that CDVAE can generate a structurally diverse set of structures from the relatively small amount of training data. To screen these 3000 structures, we use the previously mentioned ALIGNN model to predict the T_C, formation energy, and band gap with the intention of finding materials with high T_C, negative formation energy and low band gap (high DOS at the Fermi level). Fig. A.4 depicts a comparison of the distribution of the T_C obtained from the 1058 structures in the JARVIS-DFT database the 3000 CDVAE created structures with a T_C predicted from ALIGNN, respectively. After carefully screening the ALIGNN computed properties of the CDVAE structures, we found 62 materials that met the criteria (shown in Fig. 2), and performed our DFT workflow on them. Out of these 62 structures, we found 20 to have a DFT computed T_C above 5 K, to be dynamically stable and do not have a chemical formula already in the training data (the 1058

structures from JARVIS-SC). In addition, we computed the formation energy and energy above the convex hull of these 20 structures. Full details of these 20 new candidate superconductors (which all have a P1 spacegroup) can be found in Table 1. The structures that are closest to the convex hull are depicted in Fig. A.5.

4 Conclusion

In this work, we have used deep learning methods trained with DFT data for forward and inverse design of high T_C superconductors. Using the ALIGNN model for property prediction (forward design), we trained models for the Debye temperature and DOS at the Fermi-level (for BCS screening) and trained models to predict the T_C and the underlying EPC parameters, using the smaller DFT datasets computed in this work. Using the CDVAE model (inverse design), we generated thousands of candidate superconductors, which we screened with ALIGNN and verified with additional DFT calculations. This work emphasizes the importance of deep learning for the discovery of exotic materials. Unfortunately, a drawback of deep-learning models is that it is difficult to extract physical insight from their internal parameters, but we hope to investigate these ideas further in future works. In addition, the small datasets used in this work limit our deep learning models, but we are making an ongoing effort to expand our database of superconducting properties.

Acknowledgments and Disclosure of Funding

All authors thank the National Institute of Standards and Technology for funding, computational, and data-management resources, specifically the NIST Nisaba and Raritan HPC clusters. K.C. thanks the computational support from XSEDE (Extreme Science and Engineering Discovery Environment) computational resources under allocation number TG-DMR 190095. Contributions from K.C. were supported by the financial assistance award 70NANB19H117 from the U.S. Department of Commerce, National Institute of Standards and Technology.

References

Orson L Anderson. A simplified method for calculating the debye temperature from elastic constants. *Journal of Physics and Chemistry of Solids*, 24(7):909–917, 1963.

John Bardeen, Leon N Cooper, and John Robert Schrieffer. Theory of superconductivity. *Physical review*, 108(5):1175, 1957.

Stefano Baroni, Paolo Giannozzi, and Andrea Testa. Green's-function approach to linear response in solids. *Physical review letters*, 58(18):1861, 1987.

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

Kamal Choudhary and Kevin Garrity. Designing high-tc superconductors with bcs-inspired screening, density functional theory, and deep-learning. *npj Computational Materials*, 8(1): 244, 2022. doi: 10.1038/s41524-022-00933-1. URL https://doi.org/10.1038/s41524-022-00933-1.

Kamal Choudhary and Francesca Tavazza. Convergence and machine learning predictions of monkhorst-pack k-points and plane-wave cut-off in high-throughput dft calculations. *Computational materials science*, 161:300–308, 2019.

Kamal Choudhary, Irina Kalish, Ryan Beams, and Francesca Tavazza. High-throughput identification and characterization of two-dimensional materials using density functional theory. *Scientific reports*, 7(1):1–16, 2017.

Kamal Choudhary, Gowoon Cheon, Evan Reed, and Francesca Tavazza. Elastic properties of bulk and low-dimensional materials using van der waals density functional. *Physical Review B*, 98(1): 014107, 2018a.

Kamal Choudhary, Brian DeCost, and Francesca Tavazza. Machine learning with force-field-inspired descriptors for materials: Fast screening and mapping energy landscape. *Physical review materials*, 2(8):083801, 2018b.

Kamal Choudhary, Qin Zhang, Andrew CE Reid, Sugata Chowdhury, Nhan Van Nguyen, Zachary Trautt, Marcus W Newrock, Faical Yannick Congo, and Francesca Tavazza. Computational screening of high-performance optoelectronic materials using optb88vdw and tb-mbj formalisms. *Scientific data*, 5(1):1–12, 2018c.

Kamal Choudhary, Marnik Bercx, Jie Jiang, Ruth Pachter, Dirk Lamoen, and Francesca Tavazza. Accelerated discovery of efficient solar cell materials using quantum and machine-learning methods. *Chemistry of Materials*, 31(15):5900–5908, 2019a.

Kamal Choudhary, Kevin F Garrity, and Francesca Tavazza. High-throughput discovery of topologically non-trivial materials using spin-orbit spillage. *Scientific reports*, 9(1):1–8, 2019b.

Kamal Choudhary, Jaafar N Ansari, Igor I Mazin, and Karen L Sauer. Density functional theory-based electric field gradient database. *Scientific Data*, 7(1):1–10, 2020a.

Kamal Choudhary, Kevin F Garrity, Jie Jiang, Ruth Pachter, and Francesca Tavazza. Computational search for magnetic and non-magnetic 2d topological materials using unified spin–orbit spillage screening. *NPJ Computational Materials*, 6(1):1–8, 2020b.

Kamal Choudhary, Kevin F Garrity, Andrew CE Reid, Brian DeCost, Adam J Biacchi, Angela R Hight Walker, Zachary Trautt, Jason Hattrick-Simpers, A Gilad Kusne, Andrea Centrone, et al. The joint automated repository for various integrated simulations (jarvis) for data-driven materials design. *npj Computational Materials*, 6(1):1–13, 2020c.

Kamal Choudhary, Kevin F Garrity, Vinit Sharma, Adam J Biacchi, Angela R Hight Walker, and Francesca Tavazza. High-throughput density functional perturbation theory and machine learning predictions of infrared, piezoelectric, and dielectric responses. *NPJ Computational Materials*, 6 (1):1–13, 2020d.

Kamal Choudhary, Kevin F Garrity, Nirmal J Ghimire, Naween Anand, and Francesca Tavazza. High-throughput search for magnetic topological materials using spin-orbit spillage, machine learning, and experiments. *Physical Review B*, 103(15):155131, 2021.

Leon N Cooper and Dmitri Feldman. *BCS: 50 years*. World scientific, 2010.

Kevin F Garrity, Joseph W Bennett, Karin M Rabe, and David Vanderbilt. Pseudopotentials for high-throughput dft calculations. *Computational Materials Science*, 81:446–452, 2014.

Paolo Giannozzi, Stefano Baroni, Nicola Bonini, Matteo Calandra, Roberto Car, Carlo Cavazzoni, Davide Ceresoli, Guido L Chiarotti, Matteo Cococcioni, Ismaila Dabo, et al. Quantum espresso: a modular and open-source software project for quantum simulations of materials. *Journal of physics: Condensed matter*, 21(39):395502, 2009.

Feliciano Giustino. Electron-phonon interactions from first principles. *Reviews of Modern Physics*, 89(1):015003, 2017.

Xavier Gonze. Perturbation expansion of variational principles at arbitrary order. *Physical Review A*, 52(2):1086, 1995.

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

H Kamerlingh Onnes. The resistance of pure mercury at helium temperatures. *Commun. Phys. Lab. Univ. Leiden, b*, 120, 1911.

Mitsuaki Kawamura, Yuma Hizume, and Taisuke Ozaki. Benchmark of density functional theory for superconductors in elemental materials. *Phys. Rev. B*, 101:134511, Apr 2020. doi: 10.1103/PhysRevB.101.134511. URL `https://link.aps.org/doi/10.1103/PhysRevB.101.134511`.

Diederik P Kingma and Max Welling. Auto-encoding variational bayes, 2013. URL https://arxiv.org/abs/1312.6114.

WL McMillan. Transition temperature of strong-coupled superconductors. *Physical Review*, 167 (2):331, 1968.

John P Perdew, Adrienn Ruzsinszky, Gábor I Csonka, Oleg A Vydrov, Gustavo E Scuseria, Lucian A Constantin, Xiaolan Zhou, and Kieron Burke. Restoring the density-gradient expansion for exchange in solids and surfaces. *Physical review letters*, 100(13):136406, 2008.

Charles P Poole, Horacio A Farach, and Richard J Creswick. *Superconductivity*. Academic press, 2013.

Horst Rogalla and Peter H Kes. *100 years of superconductivity*. Taylor & Francis, 2011.

Jascha Sohl-Dickstein, Eric A. Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. *CoRR*, abs/1503.03585, 2015. URL http://arxiv.org/abs/1503.03585.

Yang Song and Stefano Ermon. Generative modeling by estimating gradients of the data distribution, 2019. URL https://arxiv.org/abs/1907.05600.

Malgorzata Wierzbowska, Stefano de Gironcoli, and Paolo Giannozzi. Origins of low-and high-pressure discontinuities of $t_{-}\{c\}$ in niobium. *arXiv preprint cond-mat/0504077*, 2005.

Daniel Wines, Kamal Choudhary, Adam J. Biacchi, Kevin F. Garrity, and Francesca Tavazza. High-throughput dft-based discovery of next generation two-dimensional (2d) superconductors. *Nano Letters*, 23(3):969–978, 02 2023a. doi: 10.1021/acs.nanolett.2c04420. URL https://doi.org/10.1021/acs.nanolett.2c04420.

Daniel Wines, Kamal Choudhary, and Francesca Tavazza. Systematic dft+u and quantum monte carlo benchmark of magnetic two-dimensional (2d) crx3 (x = i, br, cl, f). *The Journal of Physical Chemistry C*, 127(2):1176–1188, 01 2023b. doi: 10.1021/acs.jpcc.2c06733. URL https://doi.org/10.1021/acs.jpcc.2c06733.

Tian Xie, Xiang Fu, Octavian-Eugen Ganea, Regina Barzilay, and Tommi Jaakkola. Crystal diffusion variational autoencoder for periodic material generation, 2021. URL https://arxiv.org/abs/2110.06197.

A Appendix

A.1 JARVIS Infrastructure

We utilize the publicly available JARVIS Choudhary et al. (2020c) infrastructure to for our DFT and deep learning goals mentioned above. JARVIS (https://jarvis.nist.gov/) is a collection of databases and tools to automate materials design using classical force-field, density functional theory, machine learning calculations and experiments. In particular, we obtain elastic tensor and DOS data from JARVIS-DFT database, establish the DFT workflow with JARVIS-Tools package and train the deep-learning model using ALIGNN. JARVIS-DFT is a density functional theory based database of over 70000 materials with several material properties such as formation energy, band gap with different level of theories Choudhary et al. (2018c), solar-cell efficiency Choudhary et al. (2019a), topological spin-orbit coupling spillage Choudhary et al. (2021; 2019b; 2020b), elastic tensors Choudhary et al. (2018a), dielectric tensors, piezoelectric tensors, infrared and Raman spectrum Choudhary et al. (2020d), electric field gradients Choudhary et al. (2020a), exfolation energies Choudhary et al. (2017), two-dimensional (2D) magnets Wines et al. (2023b), and bulk Choudhary & Garrity (2022) and 2D superconductors Wines et al. (2023a), all with stringent DFT-convergence setup Choudhary & Tavazza (2019).

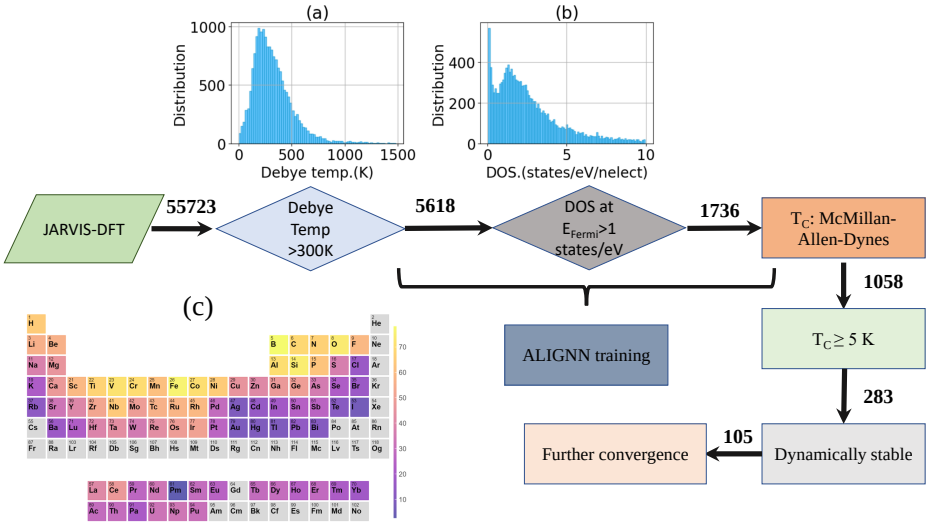

Figure A.1: Schematic showing the steps involved in identifying high-T_C superconductors. a) statistical distribution of Debye temperature (K) and b) statistical distribution of electronic density of states (states/eV/total number of electrons) at Fermi level from the JARVIS-DFT database, c) probability that compounds containing a given element have θ_D >300 K. The flow chart shows the application of BCS-inspired screening, density functional theory calculations and deep-learning training.

A.2 Further Details of BCS Screening and DFT Calculations

BCS-theory Bardeen et al. (1957) states that the attractive electron-electron interaction mediated by phonons results in Cooper pairs, which are bound states that are formed by two electrons with opposite spins and momenta. BCS-theory gives the relation between the Debye temperature (θ_D), electronic DOS at Fermi level $N(0)$, electron-phonon interaction (V) and the superconducting transition temperature (T_C):

$$T_C = 1.14\theta_D \exp\left(-\frac{1}{N(0)V}\right) \tag{A.1}$$

θ_D is defined as Anderson (1963):

$$\theta_D = \frac{h}{k_B}\left[\frac{3nN_a\rho}{4\pi M}\right]^{\frac{1}{3}} v_m \tag{A.2}$$

where h is Planck's constant, k_B is the Boltzmann constant, n is the number of atoms per formula unit, N_A is Avogadro constant, ρ is the crystal structure's density, M is the molar mass, and v_m is the average sound velocity obtained from the elastic tensor Anderson (1963).

The EPC parameter is derived from spectral function $\alpha^2F(\omega)$ which is calculated as follows:

$$\alpha^2F(\omega) = \frac{1}{2\pi N(\epsilon_F)} \sum_{qj} \frac{\gamma_{qj}}{\omega_{qj}} \delta(\omega - \omega_{qj}) w(q) \tag{A.3}$$

where ω_{qj} is the mode frequency, $N(\epsilon_F)$ is the DOS at the Fermi level ϵ_F, δ is the Dirac-delta function, $w(q)$ is the weight of the q point, γ_{qj} is the linewidth of a phonon mode j at wave vector q and is given by:

$$\gamma_{qj} = 2\pi\omega_{qj} \sum_{nm} \int \frac{d^3k}{\Omega_{BZ}} |g^j_{kn,k+qm}|^2 \delta(\epsilon_{kn} - \epsilon_F)\delta(\epsilon_{k+qm} - \epsilon_F) \tag{A.4}$$

Here, the integral is over the first Brillouin zone, ϵ_{kn} and ϵ_{k+qm} are the DFT eigenvalues with wavevector k and $k+q$ within the nth and m^{th} bands respectively, $g^{j}_{kn,k+qm}$ is the electron-phonon matrix element. γ_{qj} is related to the mode EPC parameter λ_{qj} by:

$$\lambda_{qj} = \frac{\gamma_{qj}}{\pi h N(\epsilon_F)\omega^2_{qj}} \tag{A.5}$$

Now, the EPC parameter is given by:

$$\lambda = 2\int \frac{\alpha^2 F(\omega)}{\omega} d\omega = \sum_{qj} \lambda_{qj} w(q) \tag{A.6}$$

with $w(q)$ as the weight of a q point.

The superconducting transition temperature, T_C can then be approximated using McMillan-Allen-Dynes McMillan (1968) equation as follows:

$$T_C = \frac{\omega_{log}}{1.2} \exp\left[-\frac{1.04(1+\lambda)}{\lambda - \mu^*(1+0.62\lambda)}\right] \tag{A.7}$$

where

$$\omega_{log} = \exp\left[\frac{\int d\omega \frac{\alpha^2 F(\omega)}{\omega} \ln \omega}{\int d\omega \frac{\alpha^2 F(\omega)}{\omega}}\right] \tag{A.8}$$

In Eq. A.7, the parameter μ^* is the effective Coulomb potential parameter, which we take as 0.09.

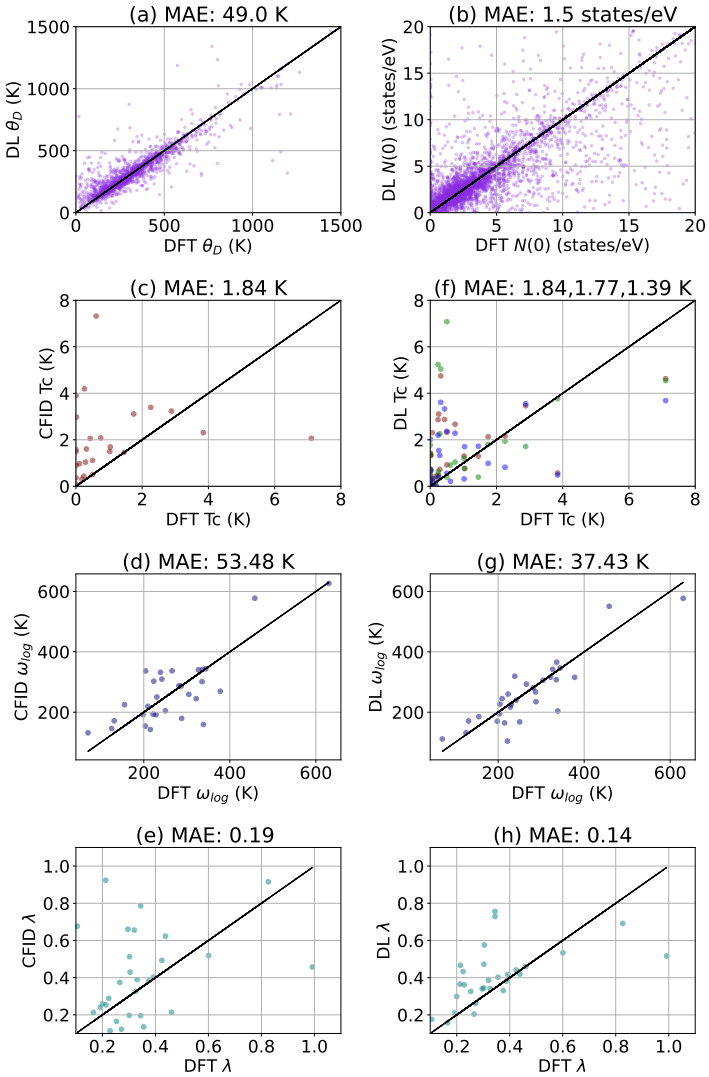

Figure A.2: Atomistic line graph neural network based deep-learning (DL) regression model performance on 5 % test set for a) Debye temperature and b) DOS. Classical force-field descriptor (CFID) (c,d,e) and DL (f,g,h) based regression model performance on 5 % test set for DFT calculated transition temperature (T_C), EPC parameter ω_{log}, and EPC parameter λ. In (f), we show performance with direct T_C prediction (red color), T_C prediction with direct prediction of ω_{log} and λ and then using eq. A.7 (green color) and T_C prediction with Eliashberg function (black color).

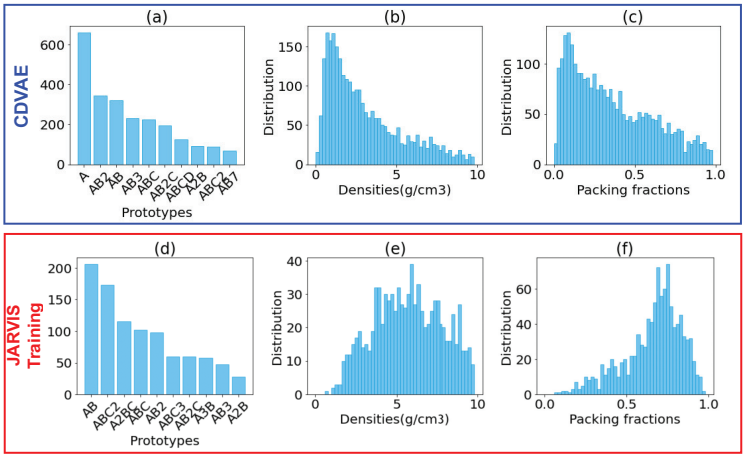

Figure A.3: The number of prototypes (chemical formula), distribution of densities, and distribution of packing fractions for a) - c) the 3000 CDVAE generated structures and for d) - f) the 1058 structures used for training from the JARVIS-DFT database.

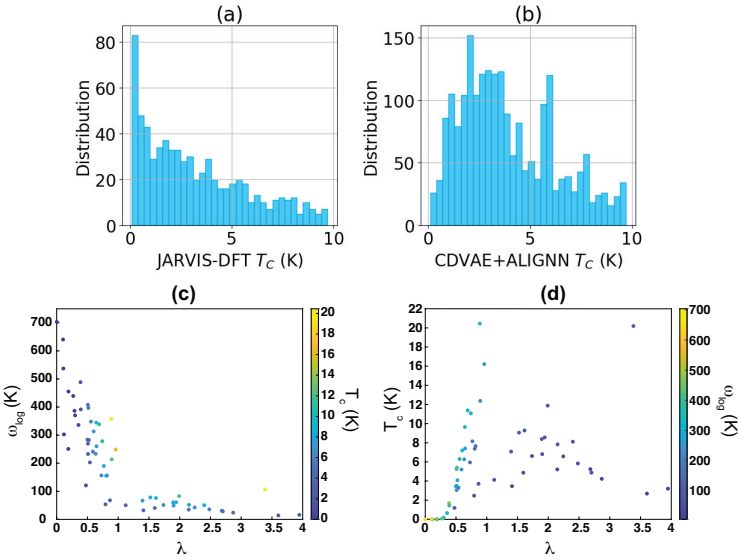

Figure A.4: The distribution of T_C for the a) 1058 JARVIS-DFT structures (T_C computed with DFT) and b) the 3000 CDVAE structures (T_C computed with ALIGNN). c) - d) the relation between EPC parameters for the CDVAE candidate materials verified with DFT.

Table 1: Chemical formula, JARVIS ID (JID), T_C, formation energy per atom and energy above the convex hull per atom of the 20 candidate superconductors from CDVAE verified by DFT calculations.

Structure	JID	T_C (K)	E_{form} (eV/atom)	E_{hull} (eV/atom)
VN_2	JVASP-161655	20.2	-0.56	0.32
NTaB (I)	JVASP-161630	16.2	-0.32	0.84
BORu	JVASP-161610	12.4	-0.72	0.75
BTa_2N	JVASP-161612	11.9	-0.37	0.67
NTaB (II)	JVASP-161624	11.1	-0.32	0.84
BN2Zr	JVASP-161608	9.3	-1.23	0.43
BTaNS	JVASP-161613	9.1	-0.33	0.88
NP_2Sr	JVASP-162663	8.4	-0.54	0.22
TaP_2	JVASP-161649	8.1	-0.45	0.25
$NPdTi_2$	JVASP-161629	7.7	-0.78	0.47
$PScSi_2$	JVASP-161644	7.4	-0.43	0.49
AlN_2V (I)	JVASP-161599	7.2	-1.14	0.34
TiO_2NbN	JVASP-161653	7.1	-2.15	0.30
NBaP	JVASP-161626	6.6	-0.59	0.22
NVBRu	JVASP-161631	6.3	-0.14	0.74
ScO_3Zr	JVASP-161647	6.3	-3.43	0.15
Al_2N	JVASP-161597	5.9	-0.79	0.31
AlN_2V (II)	JVASP-162662	5.4	-1.14	0.34
ScBORuCa	JVASP-161646	5.2	-1.12	0.58
B_2TaS	JVASP-161604	5.2	-0.07	0.58

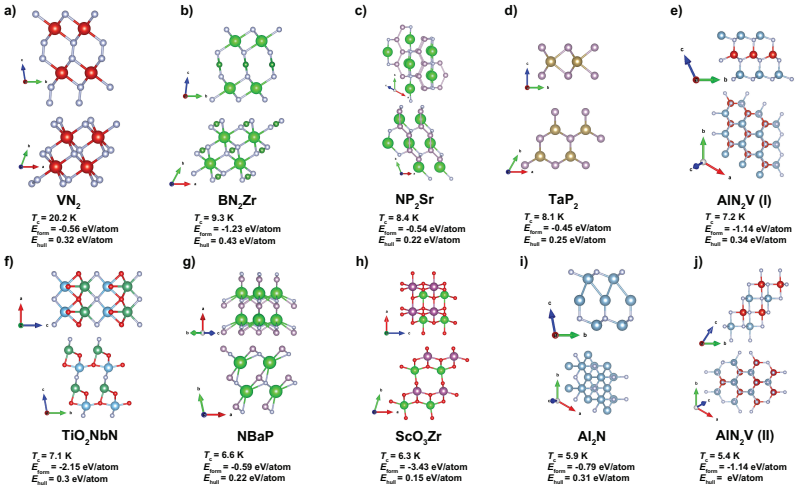

Figure A.5: Top and side view of the top superconductor candidates (closest to the convex hull) generated with CDVAE and verified with DFT: a) VN_2, b) BN_2Zr, c) NP_2S, d) TaP_2, e) AlN_2V (I), f) TiO_2NbN, g) NBaP, h) ScO_3Zr, i) Al_2N and j) AlN_2V (II). T_c, E_{form} and E_{hull} are also given for each material.