Predicting novel superconducting hydrides using machine learning approaches

Michael J. Hutcheon,* Alice M. Shipley,† and Richard J. Needs
Theory of Condensed Matter Group, Cavendish Laboratory, J. J. Thomson Avenue, Cambridge CB3 0HE, United Kingdom

(Received 27 January 2020; revised manuscript received 27 March 2020; accepted 27 March 2020; published 22 April 2020)

The search for superconducting hydrides has, so far, largely focused on finding materials exhibiting the highest possible critical temperatures (T_c). This has led to a bias toward materials stabilized at very high pressures, which introduces a number of technical difficulties in experiment. Here we apply machine learning methods in an effort to identify superconducting hydrides that can operate closer to ambient conditions. The output of these models informs subsequent crystal structure searches, from which we identify stable metallic candidates prior to performing electron-phonon calculations to obtain T_c. Hydrides of alkali and alkaline earth metals are identified as especially promising; of particular note, a T_c of up to 115 K is calculated for RbH_{12} at 50 GPa, which extends the operational pressure-temperature range occupied by hydride superconductors toward ambient conditions.

DOI: 10.1103/PhysRevB.101.144505

I. INTRODUCTION

While hydrogen is predicted to be a room-temperature superconductor at very high pressures [1], metal hydrides, in which the hydrogen atoms are "chemically precompressed," are predicted to exhibit similar behavior in experimentally accessible regimes [2,3]. In recent years, potential superconductivity has been investigated in many compressed hydrides, including scandium [4], sulfur [5–7], yttrium [8–15], calcium [16], actinium [17], thorium [18], pnictogen [19], praseodymium [20], cerium [21,22], neodymium [23], lanthanum [10,11,15,24–26], and iron hydrides [27–29]. Several reviews summarizing recent developments in the field are available [30–35]. Inspired by known superconductors, researchers have also attempted to increase T_c by chemical means: replacing atoms in known structures and assessing stability and superconductivity [36], doping known binaries with more electronegative elements to make ternary hydrides [37], and mapping alchemical phase diagrams [38].

Experimental measurements of superconductivity in high-pressure hydrides have helped to address several misconceptions about conventional superconductivity, fueling hope that it may be achieved at ambient temperature and waving a definitive farewell to the Cohen-Anderson limit [39]. The associated theoretical studies have demonstrated that the crystal structures and superconducting properties of real materials can now be accurately predicted from first principles.

In this work, we train machine learning models on a set of literature data for superconducting binary hydrides. Machine learning has previously been used in modeling hydride superconductors, with a focus on predicting the maximum obtainable critical temperature for a given composition [40]. However, on examination of the literature (see Fig. 1), it becomes apparent that the pursuit of superconductivity close to ambient conditions is as much about reducing the required pressure as it is about increasing the critical temperature. This is especially important given that working at high pressure can often present a far greater experimental challenge than working at low temperature. In this work, we therefore model critical temperature and operational pressure on an equal footing. Our models are used to inform the choice of composition for crystal structure searches and subsequent electron-phonon calculations, with the aim of extending the operation of hydride superconductors toward ambient conditions.

II. TRENDS IN HYDRIDES

A large amount of computational—and some experimental—data for the binary hydrides is available in the literature [4,9–14,16–18,20,24,25,28,40–85] (values from these references form our dataset, shown in Fig. 1). In some subsets of hydrides, certain material properties show a simple dependence on the properties of the nonhydrogen element. For example, in the alkaline earth hydrides, the van der Waals radius of the ion is well correlated with the metallization pressure [86]. However, obtaining strong electron-phonon coupling at low pressures is, in general, a more complicated process; simple correlations between composition and operational pressure or critical temperature are therefore absent in the dataset as a whole. We look at more complicated trends by constructing machine learning models of critical temperature and operational pressure that take as input a set of easily obtained material descriptors. For a particular element E and corresponding binary hydride EH_n these descriptors are as follows:

(i) Hydrogen content (n).
(ii) Van der Waals radius of E.
(iii) Atomic number of E.
(iv) Mass number of E.
(v) Numbers of s, p, d, and f electrons in the (atomic) electron configuration of E.

*mjh261@cam.ac.uk
†ams277@cam.ac.uk

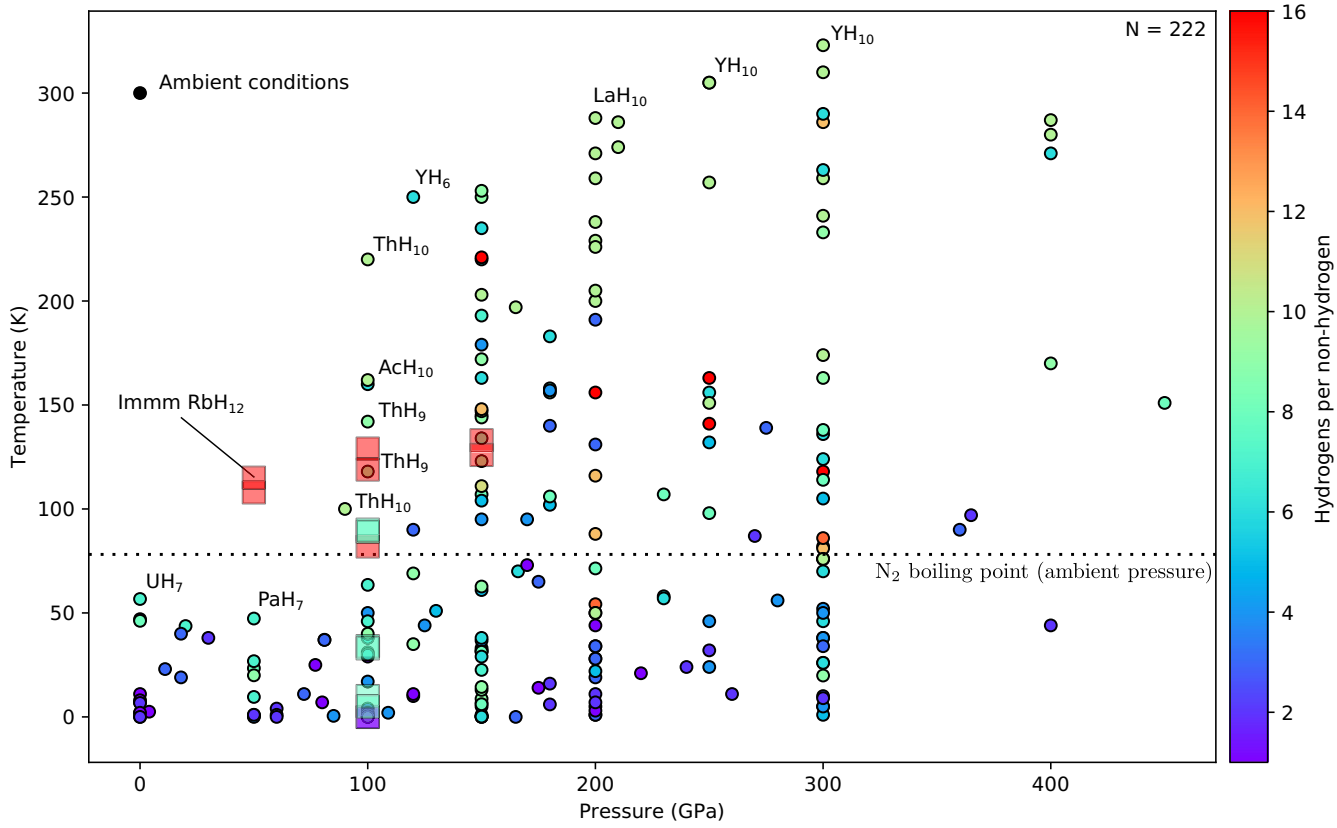

FIG. 1. The critical temperatures of binary hydrides at various pressures found in the literature are shown as circles. Materials on the frontier toward ambient conditions are labeled. Multiple points with the same stoichiometry arise from T_c for a particular phase calculated at different pressures, or from different structural phases of the same material. New structures found in this work, with T_c calculated using DFPT (as reported in Table I), are shown as translucent squares; of note is *Immm*-RbH_{12} (labeled, see also Fig. 6), which extends the frontier significantly.

Once constructed, we apply the model to all materials with the chemical composition EH_n, where E is any element in the Periodic Table and $n \in [1, 2, \ldots, 32)$ [87]. From these, the materials that are predicted to exhibit superconductivity closest to ambient conditions serve as a guide for searches for new binary hydrides.

A. Neural network

We train a fully connected neural network (using the Keras frontend to the Tensorflow machine-learning library [88,89]), with the topology shown in Fig. 2, on the dataset shown in Fig. 1. The squared absolute error $|(\Delta T_c, \Delta P)|^2$ between the predicted and literature values serves as our cost function, which we minimize using the *Adam* stochastic optimizer [90]. The input (and expected output) data are positive-definite and therefore have a nonzero mean and are not normally distributed, prompting the use of self-normalizing activation functions [91,92] to improve training behavior. Since the number of data points is comparable to the number of parameters in our network, the risk of over-fitting becomes significant. To mitigate this, we split the data into a randomly selected validation set (consisting of 25% of the initial data points) and a training set (consisting of the other 75%). Once the model starts overfitting to the training data, the validation set error starts to increase, allowing us to choose the model parameters from the training epoch for which the validation set error is minimal. We cross-validate the results by repeating this process 64 times and averaging the predictions—this is an approximation of leave-p-out cross-validation with $p = 25\%$ of the dataset. We also apply L_2 regularization to the parameters in the intermediate dense nodes to decrease the propensity for overfitting, improving the convergence of this cross-validation scheme.

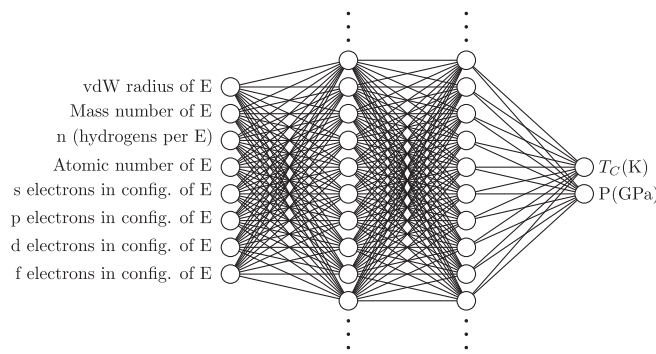

FIG. 2. Topology of our neural network model. An input layer is fed to the material descriptors for the hydride EH_n, one per input node. This layer then feeds two densely connected intermediate layers (of 32 nodes each), the last of which feeds the output layer with one temperature node and one pressure node.

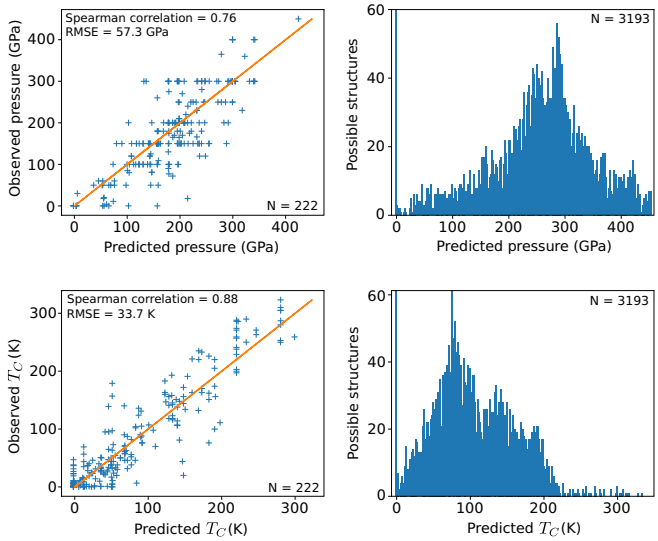

FIG. 3. Behavior of our machine learning model of critical temperatures and associated operational pressures for binary hydrides. The correlation between the predicted and observed values for the data in the literature is shown, as well as the resulting distribution of pressures and temperatures when the model is applied to the set of all possible binary hydrides as defined in Sec. II.

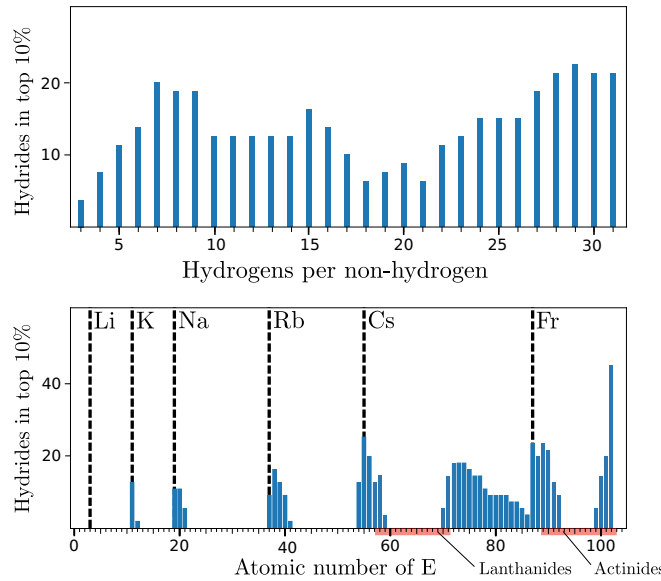

FIG. 4. Distribution of hydrogen atoms per nonhydrogen atom and atomic number of the nonhydrogen element for the 10% of hydrides that our machine learning model predicted to exhibit superconductivity closest to ambient conditions (i.e., the 10% with lowest D). Black dashed lines indicate the atomic numbers of alkali metals.

B. Model behavior

The basic behavior of the machine learning model is shown in Fig. 3. We see that it achieves a reasonable correlation with the literature values and predicts sensible pressures and temperatures for unseen materials. To gain insight into properties that favor ambient-condition superconductivity, we define a measure of distance $D = |(P, T_c - 293)|_1$. This distance decreases as we move toward ambient conditions from the pressure-temperature region containing the known hydrides (see Fig. 1). In Fig. 4 we plot the distribution of material properties for the 10% of hydrides predicted to exhibit superconductivity closest to ambient conditions (i.e., the 10% with lowest D). We can see that the model predicts the heavy alkali and alkaline earth metal hydrides to be the best candidates, with the number of close-to-ambient materials then decreasing as we go across each period. The distribution of the number of hydrogen atoms is more uniform, suggesting it is necessary to consider a range of different stoichiometries for each composition. These conclusions are reinforced by the construction of a simple linear regression model [93], which reproduces the general trends exhibited by the machine learning model (but, unsurprisingly, exhibits worse correlation with the literature values). The predicted optimal (minimum D) hydride compositions from the machine learning model are shown for each element of the Periodic Table in Fig. 5.

We note that the points included in our dataset will be of varying quality, come from different research groups, and are of both experimental and theoretical origin. The majority are theoretical and calculated within the harmonic approximation. Although it has been shown that anharmonicity can affect the calculated critical temperature for hydrides [7,94], there are insufficient data in the literature to build a model exclusively from anharmonic results. However, since we only seek to extract general trends, which will serve simply to inform areas of focus for structure searching, the dataset is sufficient for our purposes.

III. STRUCTURE SEARCHING

The models constructed in the previous section point towards the alkali and alkaline earth metal hydrides as some of the best candidates for superconductivity near ambient conditions. From these, we studied caesium and rubidium hydrides; these systems were chosen due to their predicted proximity to superconductivity at ambient conditions (see Figs. 4 and 5) and the fact that they have not been studied extensively in the past, unlike the hydrides of other elements in these two groups. Caesium and rubidium polyhydrides have been studied previously using structure searching methods in Refs. [95] and [96], respectively, although potential superconductivity was not investigated in either case.

Our structure searching calculations were performed using *ab initio* random structure searching (AIRSS) [97,98] and the plane-wave pseudopotential code CASTEP [99]. Since our models suggest that a wide range of stoichiometries should be considered, convex hulls were constructed using AIRSS and qhull [100] in order to identify those which are stable at 50, 100, and 200 GPa [93]. The Perdew-Burke-Ernzerhof (PBE) generalized gradient approximation [101], CASTEP QC5 pseudopotentials, a 400 eV plane-wave cutoff and a **k**-point spacing of $2\pi \times 0.05$ Å^{-1} were used in all searches. The Cs-H convex hulls calculated in this work at 100 and 200 GPa both partially agree with the hull calculated at 150 GPa in Ref. [95]. Once stable stoichiometries had been identified, additional AIRSS searches for RbH_3, RbH_5, RbH_9, RbH_{11}, RbH_{12}, CsH_5, CsH_7, CsH_{13}, and CsH_{15} using the same parameters and pseudopotentials were performed at 100 and 200 GPa.

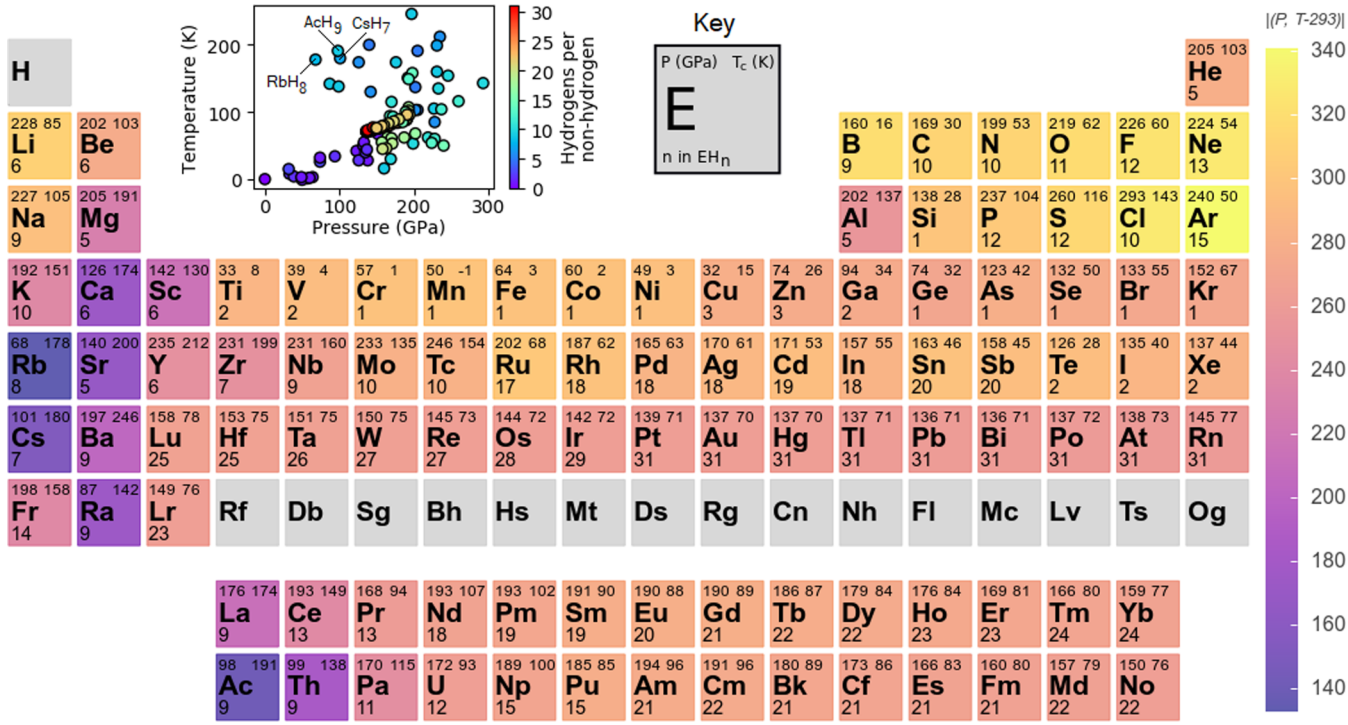

FIG. 5. The Periodic Table of optimal binary hydrides according to our machine learning model. The predicted critical temperature, corresponding pressure, and optimal hydrogen content are shown for each element. Elements are colored according to the predicted distance from ambient-condition superconductivity $D = |(P, T_c - 293)|$. Inset: the distribution in pressure-temperature space of these predictions. We note that we did not explicitly prevent the neural network from predicting negative critical temperatures and it does so for MnH. However, in general, the machine learning model has learned that critical temperatures should be positive (see the lower right panel of Fig. 3).

IV. SELECTING CANDIDATE STRUCTURES

For each selected stoichiometry, the enthalpy was calculated as a function of pressure for the most stable structures arising from the AIRSS search. These geometry optimizations were performed using QUANTUM ESPRESSO [102,103], the PBE functional, a 950 eV cutoff, ultrasoft pseudopotentials [93], and a **k**-point spacing of $2\pi \times 0.02$ Å^{-1}. The electronic density of states (DOS) at the Fermi energy was also evaluated for each structure at 50 and 150 GPa in order to identify metallic structures. We were then able to limit our interest to structures that were both energetically competitive (according to the enthalpy plots) and had a considerable DOS at the Fermi energy in the low-pressure region (25–125 GPa). Full lists of the competitive structures predicted here are available, along with the enthalpy plots and DOS values, in the Supplemental Material [93]. The remaining candidates, for which electron-phonon coupling calculations were performed, include $C2/m$-RbH_{12}, $Immm$-RbH_{12}, and various CsH_7 and RbH_3 structures (see Table I).

TABLE I. Critical temperatures calculated using DFPT for promising hydride compositions. The structures listed here were found in this work using AIRSS and are available in an online repository [110]. The data in this table are also shown in Fig. 1 for comparison with previous results in the literature.

Stoichiometry	Space group	Pressure (GPa)	T_c (K)
RbH_{12}	$C2/m$	50	108
RbH_{12}	$C2/m$	100	129
RbH_{12}	$C2/m$	150	133
RbH_{12}	$Cmcm$	100	82
RbH_{12}	$Immm$	50	115
RbH_{12}	$Immm$	100	119
RbH_{12}	$Immm$	150	126
CsH_7	$P1$	100	90
CsH_7	$I4mm$	100	34
CsH_7	$P4mm$	100	33
CsH_7	$I4/mmm$	100	10
CsH_7	Cm	100	5
CsH_7	$Cmc2_1$	100	89
RbH_3	$Pmma$	100	0
RbH_3	$Cmmm$	100	0

V. ELECTRON-PHONON COUPLING AND SUPERCONDUCTIVITY

The Hamiltonian of a coupled electron-phonon system is given by

$$H = \sum_{kn} \epsilon_{nk} c^\dagger_{nk} c_{nk} + \sum_{q\nu} \omega_{q\nu} \left(a^\dagger_{q\nu} a_{q\nu} + \frac{1}{2} \right) + \frac{1}{\sqrt{N_p}} \sum_{kqmn\nu} g_{mn\nu}(k, q) c^\dagger_{m,k+q} c_{nk} (a_{q\nu} + a^\dagger_{-q\nu}). \quad (1)$$

In this work, we calculate the electronic Kohn-Sham eigenvalues ϵ_{nk}, phonon frequencies $\omega_{q,\nu}$, and electron-phonon

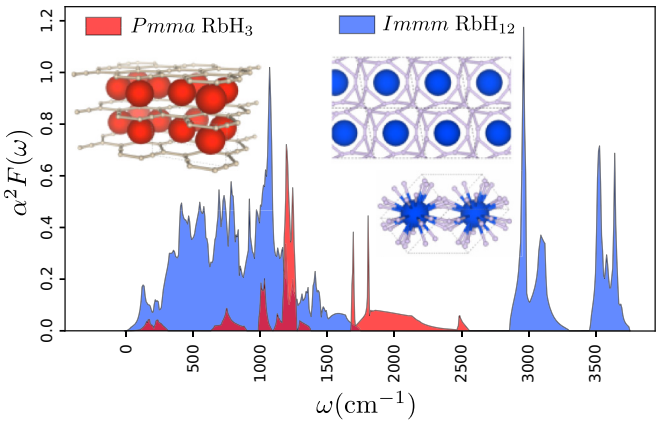

FIG. 6. The Eliashberg function for *Immm*-RbH_{12} and *Pmma*-RbH_3 found in our AIRSS searches (see Table I). It is clear to see the enhanced high-frequency part of the Eliashberg function for *Immm*-RbH_{12}, arising from the hydrogen cage. In contrast, the Eliashberg function for the layered RbH_3 structure does not extend to such high frequencies. This effect can also be seen from the phonon linewidths (plotted along with the phonon dispersion in the Supplemental Material [93]).

coupling constants $g_{mnv}(k, q)$ appearing in H from first principles using density functional perturbation theory (DFPT) as implemented in the QUANTUM ESPRESSO code [102,103]. The resulting Hamiltonian is then treated within Migdal-Eliashberg theory [104–106] where we solve the Eliashberg equations using the ELK code [107]. This gives us the superconducting gap as a function of temperature, from which we obtain a prediction for T_c.

To carry out these calculations, we use the PBE functional, the same ultrasoft pseudopotentials as in the geometry optimizations, an 820 eV plane-wave cutoff, and a **q**-point grid with a spacing of $\approx 2\pi \times 0.1$ Å^{-1} (e.g., a $2 \times 2 \times 2$ grid for a 26-atom unit cell of RbH_{12}). Two separate **k**-point grids are used (of 6^3 and 8^3 times the size of the **q**-point grid, respectively), allowing us to determine the optimal double-δ smearing width necessary to calculate the critical temperature [15,108].

Electron-phonon calculations were performed for a range of competitive RbH_{12}, CsH_7, and RbH_3 structures predicted in this work, and the results are shown in Table I. The highest-T_c results arise from structures with a cagelike arrangement of hydrogen atoms surrounding a central nonhydrogen element. The electronic states that originate from these cages are near the Fermi level, and they are strongly coupled together by cage vibrations. This provides the phonon-mediated pairing mechanism necessary for conventional superconductivity. Combined with a high average phonon frequency, due to the light mass of the hydrogen atoms, this results in a high critical temperature (cf. the Allen-Dynes equation [109]). This can be seen directly by looking at the Eliashberg function, shown in Fig. 6, for two illustrative structures from Table I. The enhanced high-frequency portion of the Eliashberg function for the high-T_c cagelike RbH_{12} structure is apparent. In contrast, strong electron-phonon coupling is absent at high phonon frequencies for states near the Fermi level in the layered RbH_3 structure, leading to a negligible T_c. It is perhaps unsurprising that our machine learning model suggests such compositions, despite their resulting unfavorable structures, as it is trained on mostly cagelike structures. As a result, the model may implicitly assume that compositions it is given will behave as if they adopt cagelike arrangements, leading to an overestimation of T_c. Despite this, most of the structures found are high-T_c cagelike superconductors, of which *Immm*-RbH_{12} is particularly interesting due to its location in Fig. 1.

Supplementing structure searching techniques with predictions from machine learning has allowed us to target novel regions of pressure-temperature space. We have therefore been able to identify low-pressure hydride superconductors without having to perform a large number of expensive electron-phonon calculations. It can be seen from Fig. 1 that the hydrides predicted in this work are biased towards ambient conditions when compared to the dataset as a whole.

VI. TESTING POTENTIAL SCREENING TECHNIQUES FOR HIGH-T_c CANDIDATES

In this work, we also tested two potential methods for cheaply estimating T_c ordering between structures. Good superconductivity in hydrides generally requires hydrogenic states close to the Fermi level, which (as exemplified by the findings of this work) often means favoring cagelike structures and avoiding structures with molecular-character H_2 units. It is therefore possible that the hydrogen-derived DOS normalized by the total DOS at the Fermi energy, $N_H(E_F)/N(E_F)$, may give some indication of whether a particular structure will exhibit high-T_c superconductivity. Here we also consider the hydrogen-derived electron-phonon coupling estimates (η_H) from Gaspari-Gyorffy theory [111] and test whether these two quantities could provide a method for ranking different structures (of the same stoichiometry and at the same pressure) before performing expensive electron-phonon calculations. We implemented Gaspari-Gyorffy theory within the ELK code [107]. The basics of this theory and its use here are explained in Appendix.

The calculated T_c values for the structures predicted and studied in this work allowed us to directly assess these potential screening methods. We observe that η_H correctly predicts the T_c ordering for the RbH_{12} structures at fixed pressure, as was the case for the LaH_{10} and YH_{10} systems on which preliminary tests were performed [93]. The two quantities tested here often predict the same general trends, but the DOS ratio is cheaper to calculate since it can be obtained using a pseudopotential code. Unfortunately, $N_H(E_F)/N(E_F)$ appears to be much less predictive for the CsH_7 structures, and the performance of η_H is also mixed [93]. The use of these quantities for screening applications therefore requires further investigation and testing in a wider variety of systems.

VII. CONCLUSIONS

Having identified the need to reduce the operational pressure of hydride superconductors, we searched for crystal structures that would exhibit superconductivity in novel regions of pressure-temperature space. We found that guiding structure searching techniques using a machine learning model allowed us to target the most promising regions.

Specifically, we constructed models of critical temperature and operational pressure trained on the available theoretical and experimental results for binary hydride superconductors. Several novel systems were identified as promising superconductors closer to ambient conditions; here we focused on Cs and Rb hydrides, using AIRSS to identify stable stoichiometries and predict crystal structures. Other promising candidates included Ca, Sr, Ba, Ra, Ac, Th, La, and Sc hydrides, most of which had already been theoretically studied to some extent [4,10,11,16–18,26,35,55,73,79,84]. Critical temperatures of energetically competitive candidate structures were then calculated from first principles using DFPT. A T_c of up to 115 K was calculated for RbH_{12} at 50 GPa, which represents a significant extension toward ambient-condition superconductivity from our dataset.

ACKNOWLEDGMENTS

We thank Po-Hao Chang for useful discussions regarding Gaspari-Gyorffy theory. M.J.H. acknowledges the EPSRC Centre for Doctoral Training in Computational Methods for Materials Science for funding under Grant No. EP/L015552/1. A.M.S. acknowledges funding through an EPSRC studentship. R.J.N. is supported by EPSRC under Critical Mass Grant EP/P034616/1 and the UKCP consortium grant EP/P022596/1. We are grateful for computational support from the UK national high performance computing service, ARCHER, for which access was obtained via the UKCP consortium and funded by EPSRC grant ref EP/P022561/1. This work was also performed using resources provided by the Cambridge Service for Data Driven Discovery (CSD3) operated by the University of Cambridge Research Computing Service [112], provided by Dell EMC and Intel using Tier-2 funding from the EPSRC (capital grant EP/P020259/1), and DiRAC funding from the STFC [113].

APPENDIX: GASPARI-GYORFFY THEORY

McMillan [114] showed that for strong-coupled superconductors, the electron-phonon coupling constant, λ, can be expressed as

$$\lambda = 2\int \frac{d\omega\alpha^2(\omega)F(\omega)}{\omega} = \frac{N(E_F)\langle I^2\rangle}{M\langle\omega^2\rangle}. \tag{A1}$$

λ can also be reformatted as

$$\lambda = \frac{\eta}{M\langle\omega^2\rangle},$$

where η is the so-called Hopfield parameter. Hopfield was one of the first to stress the importance of the local environment in determining λ [115]. In situations in which we have nearly perfect separation of vibrational modes into those of different atomic character (such as we may see in hydrides), we can write

$$\lambda = \sum_j \lambda_j = \sum_j \frac{\eta_j}{M_j\langle\omega_j^2\rangle}, \tag{A2}$$

where j is the atom type.

The quantity $\langle I^2\rangle$ appearing in Eq. (A1) can be approximated using Gaspari-Gyorffy (GG) theory [111]. Recent work has emerged using this theory for metal hydrides under high pressure [36,116] despite it originally being designed for elemental transition metals. The theory, based on the rigid muffin-tin approximation (RMTA), relies on several approximations [116] and allows us to reformulate the electron-phonon interaction in terms of phase shifts for a scattering potential. A self-consistent DOS calculation is thus all that is required to calculate $\langle I^2\rangle$ for each atom type and hence obtain η_j. The GG equation is

$$\langle I^2\rangle = \frac{E_F}{\pi^2 N^2(E_F)}\sum_l \frac{2(l+],1)\sin^2(\delta_{l+1}-\delta_l)N_l(E_F)N_{l+1}(E_F)}{N_l^{(1)}N_{l+1}^{(1)}}, \tag{A3}$$

where $N_l^{(1)}$ is the free-scatterer DOS given by

$$N_l^{(1)} = \frac{\sqrt{E_F}}{\pi}(2l+1)\int_0^{R_{\mathrm{MT}}} R_l^2(r,E_F)r^2dr \tag{A4}$$

and the δ_l are the scattering phase shifts. Here R_{MT} is the muffin-tin radius associated with atom type j, and R_l is the scattering solution of the Schrödinger equation. The phase shifts, which characterize the long-distance behavior of the wave function, can be written in terms of the logarithmic derivative of the radial wave function,

$$\tan(\delta_l(R_{\mathrm{MT}},E_F)) = \frac{j_l'(kR_{\mathrm{MT}}) - j_l(kR_{\mathrm{MT}})L_l(R_{\mathrm{MT}},E_F)}{n_l'(kR_{\mathrm{MT}}) - n_l(kR_{\mathrm{MT}})L_l(R_{\mathrm{MT}},E_F)}, \tag{A5}$$

where $k=\sqrt{E_F}$, $L_l = R_l'/R_l$ is the logarithmic derivative, j_l are spherical Bessel functions, and n_l are Neumann functions. We can therefore directly calculate the logarithmic derivative and use Eq. (A5) to obtain the phase shifts [117].

Since $M_j\langle\omega_j^2\rangle$ is often considerably smaller for hydrogen than for the other components, it is clear from Eq. (A2) that the hydrogen atoms can provide a considerable fraction of λ even if the Hopfield parameter of the other atom type is similar in magnitude. Calculating η_H can therefore, in some cases, provide a cheap screening method for identifying potential high-T_c hydrides. In particular, the average phonon frequencies for different structures are often similar when considering the same stoichiometry at the same pressure. If the average phonon frequencies are assumed to be exactly equivalent in such cases, we then arrive at a potential way of estimating T_c ordering between structures, simply by considering η_H. It is in this context that we assess the utility of GG theory in this work.

[1] N. W. Ashcroft, Phys. Rev. Lett. **21**, 1748 (1968).
[2] J. Gilman, Phys. Rev. Lett. **26**, 546 (1971).
[3] N. W. Ashcroft, Phys. Rev. Lett. **92**, 187002 (2004).
[4] A. P. Durajski and R. Szczesniak, Supercond. Sci. Technol. **27**, 115012 (2014).
[5] D. Duan, Y. Liu, F. Tian, D. Li, X. Huang, Z. Zhao, H. Yu, B. Liu, W. Tian, and T. Cui, Nat. Sci. Rep. **4**, 6968 (2014).

[6] A. P. Drozdov, M. I. Eremets, I. A. Troyan, V. Ksenofontov, and S. I. Shylin, Nature (London) **525**, 73 (2015).

[7] I. Errea, M. Calandra, C. J. Pickard, J. Nelson, R. J. Needs, Y. Li, H. Liu, Y. Zhang, Y. Ma, and F. Mauri, Phys. Rev. Lett. **114**, 157004 (2015).

[8] D. Y. Kim, R. H. Scheicher, and R. Ahuja, Phys. Rev. Lett. **103**, 077002 (2009).

[9] Y. Li, J. Hao, H. Liu, J. S. Tse, Y. Wang, and Y. Ma, Sci. Rep. **5**, 9948 (2015).

[10] H. Liu, I. I. Naumov, R. Hoffmann, N. W. Ashcroft, and R. J. Hemley, Proc. Natl. Acad. Sci. (USA) **114**, 6990 (2017).

[11] F. Peng, Y. Sun, C. J. Pickard, R. J. Needs, Q. Wu, and Y. Ma, Phys. Rev. Lett. **119**, 107001 (2017).

[12] C. Heil, S. di Cataldo, G. B. Bachelet, and L. Boeri, Phys. Rev. B **99**, 220502(R) (2019).

[13] I. A. Troyan, D. V. Semenok, A. G. Kvashnin, A. G. Ivanova, V. B. Prakapenka, E. Greenberg, A. G. Gavriliuk, I. S. Lyubutin, V. V. Struzhkin, and A. R. Oganov, arXiv:1908.01534.

[14] P. P. Kong, V. S. Minkov, M. A. Kuzovnikov, S. P. Besedin, A. P. Drozdov, S. Mozaffari, L. Balicas, F. F. Balakirev, V. B. Prakapenka, E. Greenberg, D. A. Knyazev, and M. I. Eremets, arXiv:1909.10482.

[15] A. M. Shipley, M. J. Hutcheon, M. S. Johnson, C. J. Pickard, and R. J. Needs, arXiv:2001.05305.

[16] H. Wang, J. S. Tse, K. Tanaka, T. Iitaka, and Y. Ma, Proc. Natl. Acad. Sci. (USA) **109**, 6463 (2012).

[17] D. V. Semenok, A. G. Kvashnin, I. A. Kruglov, and A. R. Oganov, J. Phys. Chem. Lett. **9**, 1920 (2018).

[18] A. G. Kvashnin, D. V. Semenok, I. A. Kruglov, I. A. Wrona, and A. R. Oganov, ACS Appl. Mater. Interfaces **10**, 43809 (2018).

[19] Y. Fu, X. Du, L. Zhang, F. Peng, M. Zhang, C. J. Pickard, R. J. Needs, D. J. Singh, W. Zheng, and Y. Ma, Chem. Mater. **28**, 1746 (2016).

[20] D. Zhou, D. V. Semenok, D. Duan, H. Xie, W. Chen, X. Huang, X. Li, B. Liu, A. R. Oganov, and T. Cui, Sci. Adv. **6**, eaax6849 (2020).

[21] N. P. Salke, M. M. D. Esfahani, Y. Zhang, I. A. Kruglov, J. Zhou, Y. Wang, E. Greenberg, V. B. Prakapenka, J. Liu, A. R. Oganov, and J.-F. Lin, Nat. Commun. **10**, 1 (2019).

[22] X. Li, X. Huang, D. Duan, C. J. Pickard, D. Zhou, H. Xie, Q. Zhuang, Y. Huang, Q. Zhou, B. Liu, and T. Cui, Nat. Commun. **10**, 3461 (2019).

[23] D. Zhou, D. V. Semenok, H. Xie, X. Huang, D. Duan, A. Aperis, P. M. Oppeneer, M. Galasso, A. I. Kartsev, A. G. Kvashnin, A. R. Oganov, and T. Cui, J. Am. Chem. Soc. **142**, 2803 (2020).

[24] M. Somayazulu, M. Ahart, A. K. Mishra, Z. M. Geballe, M. Baldini, Y. Meng, V. V. Struzhkin, and R. J. Hemley, Phys. Rev. Lett. **122**, 027001 (2019).

[25] A. P. Drozdov, P. P. Kong, V. S. Minkov, S. P. Besedin, M. A. Kuzovnikov, S. Mozaffari, L. Balicas, F. F. Balakirev, D. E. Graf, V. B. Prakapenka, E. Greenberg, D. A. Knyazev, M. Tkacz, and M. I. Eremets, Nature (London) **569**, 528 (2019).

[26] I. A. Kruglov, D. V. Semenok, H. Song, R. Szczesniak, I. A. Wrona, R. Akashi, M. M. D. Esfahani, D. Duan, T. Cui, A. G. Kvashnin, and A. R. Oganov, Phys. Rev. B **101**, 024508 (2020).

[27] A. Majumdar, J. S. Tse, M. Wu, and Y. Yao, Phys. Rev. B **96**, 201107(R) (2017).

[28] A. G. Kvashnin, I. A. Kruglov, D. V. Semenok, and A. R. Oganov, J. Phys. Chem. C **122**, 4731 (2018).

[29] C. Heil, G. B. Bachelet, and L. Boeri, Phys. Rev. B **97**, 214510 (2018).

[30] D. Duan, Y. Liu, Y. Ma, Z. Shao, B. Liu, and T. Cui, Natl. Sci. Rev **4**, 121 (2017).

[31] E. Zurek and T. Bi, J. Chem. Phys. **150**, 050901 (2019).

[32] J. A. Flores-Livas, L. Boeri, A. Sanna, G. Profeta, R. Arita, and M. Eremets, Phys. Rep. (2020), doi: 10.1016/j.physrep.2020.02.003.

[33] L. Boeri and G. B. Bachelet, J. Phys.: Condens. Matter **31**, 234002 (2019).

[34] C. J. Pickard, I. Errea, and M. I. Eremets, Annu. Rev. Condens. Matter Phys. **11**, 57 (2020).

[35] T. Bi, N. Zarifi, T. Terpstra, and E. Zurek, The search for superconductivity in high pressure hydrides, in *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering* (Elsevier, 2019).

[36] P.-H. Chang, S. Silayi, D. Papaconstantopoulos, and M. J. Mehl, J. Phys. Chem. Solids **139**, 109315 (2019).

[37] Y. Sun, J. Lv, Y. Xie, H. Liu, and Y. Ma, Phys. Rev. Lett. **123**, 097001 (2019).

[38] C. Heil and L. Boeri, Phys. Rev. B **92**, 060508(R) (2015).

[39] M. L. Cohen and P. W. Anderson, Comments on the Maximum Superconducting Transition Temperature, in *Superconductivity in d and f Band Metals*, edited by Hugh C. Wolfe and D. H. Douglass, AIP Conf. Proc. No. 4 (AIP, New York, 1972), pp. 17–27.

[40] D. V. Semenok, I. A. Kruglov, I. A. Savkin, A. G. Kvashnin, and A. R. Oganov, Curr. Opin. Solid State Mater. Sci. 100808 (2020), doi: 10.1016/j.cossms.2020.100808.

[41] Z. M. Geballe, H. Liu, A. K. Mishra, M. Ahart, M. Somayazulu, Y. Meng, M. Baldini, and R. J. Hemley, Angew. Chem., Int. Ed. **57**, 688 (2018).

[42] I. Errea, F. Belli, L. Monacelli, A. Sanna, T. Koretsune, T. Tadano, R. Bianco, M. Calandra, R. Arita, F. Mauri, and J. A. Flores-Livas, Nature **578**, 66 (2020).

[43] Y. Xie, Q. Li, A. R. Oganov, and H. Wang, Acta Crystallogr. Sect. C **70**, 104 (2014).

[44] D. Zhou, X. Jin, X. Meng, G. Bao, Y. Ma, B. Liu, and T. Cui, Phys. Rev. B **86**, 014118 (2012).

[45] S. Yu, Q. Zeng, A. R. Oganov, C. Hu, G. Frapper, and L. Zhang, AIP Adv. **4**, 107118 (2014).

[46] D. C. Lonie, J. Hooper, B. Altintas, and E. Zurek, Phys. Rev. B **87**, 054107 (2013).

[47] J. S. Tse, Z. Song, Y. Yao, J. S. Smith, S. Desgreniers, and D. D. Klug, Solid State Commun. **149**, 1944 (2009).

[48] T. P. Zemła, K. M. Szcześniak, A. Z. Kaczmarek, and S. V. Turchuk, Mod. Phys. Lett. B **33**, 1950169 (2019).

[49] I. A. Kruglov, A. G. Kvashnin, A. F. Goncharov, A. R. Oganov, S. S. Lobanov, N. Holtgrewe, S. Jiang, V. B. Prakapenka, E. Greenberg, and A. V. Yanilkin, Sci. Adv. **4**, eaat9776 (2018).

[50] X. Feng, J. Zhang, G. Gao, H. Liu, and H. Wang, RSC Adv. **5**, 59292 (2015).

[51] Q. Gu, P. Lu, K. Xia, J. Sun, and D. Xing, Phys. Rev. B **96**, 064517 (2017).

[52] M. M. Davari Esfahani, A. R. Oganov, H. Niu, and J. Zhang, Phys. Rev. B **95**, 134506 (2017).

[53] Y.-K. Wei, J.-N. Yuan, F. I. Khan, G.-F. Ji, Z.-W. Gu, and D.-Q. Wei, RSC Adv. **6**, 81534 (2016).

[54] L.-L. Liu, H.-J. Sun, C. Z. Wang, and W.-C. Lu, J. Phys.: Condens. Matter **29**, 325401 (2017).

[55] D. V. Semenok, A. G. Kvashnin, A. G. Ivanova, V. Svitlyk, V. Y. Fominski, A. V. Sadakov, O. A. Sobolevskiy, V. M. Pudalov, I. A. Troyan, and A. R. Oganov, Mater. Today **33**, 36 (2020).

[56] N. Zarifi, T. Bi, H. Liu, and E. Zurek, J. Phys. Chem. C **122**, 24262 (2018).

[57] K. V. Shanavas, L. Lindsay, and D. S. Parker, Sci. Rep. **6**, 28102 (2016).

[58] X.-F. Li, Z.-Y. Hu, and B. Huang, Phys. Chem. Chem. Phys. **19**, 3538 (2017).

[59] V. Z. Kresin, J. Supercond. Novel Magn. **31**, 3391 (2018).

[60] Y. Liu, D. Duan, F. Tian, H. Liu, C. Wang, X. Huang, D. Li, Y. Ma, B. Liu, and T. Cui, Inorg. Chem. **54**, 9924 (2015).

[61] D. Ohlendorf and E. Wicke, J. Phys. Chem. Solids **40**, 721 (1979).

[62] C. Chen, F. Tian, D. Duan, K. Bao, X. Jin, B. Liu, and T. Cui, J. Chem. Phys. **140**, 114703 (2014).

[63] G. Gao, R. Hoffmann, N. W. Ashcroft, H. Liu, A. Bergara, and Y. Ma, Phys. Rev. B **88**, 184104 (2013).

[64] Q. Zhuang, X. Jin, T. Cui, Y. Ma, Q. Lv, Y. Li, H. Zhang, X. Meng, and K. Bao, Inorg. Chem. **56**, 3901 (2017).

[65] S. Yu, X. Jia, G. Frapper, D. Li, A. R. Oganov, Q. Zeng, and L. Zhang, Sci. Rep. **5**, 17764 (2015).

[66] V. I. Spitsyn, V. E. Antonov, O. A. Balakhovskii, I. T. Belash, E. G. Ponyatovskii, V. I. Rashchupkin, and V. S. Shekhtman, Dokl. Akad. Nauk SSSR **260**, 795 (1982).

[67] F. Li, D. Wang, H. Du, D. Zhou, Y. Ma, and Y. Liu, RSC Adv. **7**, 12570 (2017).

[68] Y. Liu, D. Duan, F. Tian, C. Wang, Y. Ma, D. Li, X. Huang, B. Liu, and T. Cui, Phys. Chem. Chem. Phys. **18**, 1516 (2016).

[69] X. Ye, N. Zarifi, E. Zurek, R. Hoffmann, and N. Ashcroft, J. Phys. Chem. C **122**, 6298 (2018).

[70] T. Skoskiewicz, A. W. Szafranski, W. Bujnowski, and B. Baranowski, J. Phys. C **7**, 2670 (1974).

[71] X. Li and F. Peng, Inorg. Chem. **56**, 13759 (2017).

[72] I. Errea, M. Calandra, and F. Mauri, Phys. Rev. B **89**, 064302 (2014).

[73] K. Tanaka, J. S. Tse, and H. Liu, Phys. Rev. B **96**, 100502(R) (2017).

[74] C.-H. Hu, A. R. Oganov, Q. Zhu, G.-R. Qian, G. Frapper, A. O. Lyakhov, and H.-Y. Zhou, Phys. Rev. Lett. **110**, 165504 (2013).

[75] Y.-K. Wei, N.-N. Ge, G.-F. Ji, X.-R. Chen, L.-C. Cai, S.-Q. Zhou, and D.-Q. Wei, J. Appl. Phys. **114**, 114905 (2013).

[76] P. Hou, X. Zhao, F. Tian, D. Li, D. Duan, Z. Zhao, B. Chu, B. Liu, and T. Cui, RSC Adv. **5**, 5096 (2015).

[77] X. Li, H. Liu, and F. Peng, Phys. Chem. Chem. Phys. **18**, 28791 (2016).

[78] R. Szczesniak and A. P. Durajski, Supercond. Sci. Technol. **27**, 015003 (2013).

[79] J. Hooper, B. Altintas, A. Shamp, and E. Zurek, J. Phys. Chem. C **117**, 2982 (2013).

[80] Y. Yao, J. S. Tse, Y. Ma, and K. Tanaka, Europhys. Lett. **78**, 37003 (2007).

[81] M. I. Eremets, I. A. Trojan, S. A. Medvedev, J. S. Tse, and Y. Yao, Science **319**, 1506 (2008).

[82] Y. Li, G. Gao, Y. Xie, Y. Ma, T. Cui, and G. Zou, Proc. Natl. Acad. Sci. (USA) **107**, 15708 (2010).

[83] X. Jin, X. Meng, Z. He, Y. Ma, B. Liu, T. Cui, G. Zou, and H.-k. Mao, Proc. Natl. Acad. Sci. (USA) **107**, 9969 (2010).

[84] H. Liu, I. I. Naumov, Z. M. Geballe, M. Somayazulu, J. S. Tse, and R. J. Hemley, Phys. Rev. B **98**, 100102(R) (2018).

[85] A. P. Drozdov, V. S. Minkov, S. P. Besedin, P. P. Kong, M. A. Kuzovnikov, D. A. Knyazev, and M. I. Eremets, arXiv:1808.07039.

[86] C. Zhang, X.-J. Chen, R.-Q. Zhang, and H.-Q. Lin, J. Phys. Chem. C **114**, 14614 (2010).

[87] A maximum of 31 hydrogens per atom was chosen to avoid overextrapolation from the dataset (where the maximum is 16).

[88] F. Chollet *et al.*, Keras, https://keras.io (2015).

[89] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, TensorFlow: Large-scale machine learning on heterogeneous systems, available at https://www.tensorflow.org/ (2015).

[90] D. P. Kingma and J. Ba, arXiv:1412.6980.

[91] G. Klambauer, T. Unterthiner, A. Mayr, and S. Hochreiter, in *Advances in Neural Information Processing Systems 30*, edited by I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Curran Associates, Inc., 2017), pp. 971–980.

[92] D. A. Clevert, T. Unterthiner, and S. Hochreiter, arXiv:1511.07289.

[93] See Supplemental Material http://link.aps.org/supplemental/10.1103/PhysRevB.101.144505 for full lists of the predicted competitive structures along with the enthalpy plots and DOS values.

[94] I. Errea, M. Calandra, and F. Mauri, Phys. Rev. Lett. **111**, 177002 (2013).

[95] A. Shamp, J. Hooper, and E. Zurek, Inorg. Chem. **51**, 9333 (2012).

[96] J. Hooper and E. Zurek, Chem.—Eur. J. **18**, 5013 (2012).

[97] C. J. Pickard and R. J. Needs, J. Phys.: Condens. Matter **23**, 053201 (2011).

[98] R. J. Needs and C. J. Pickard, APL Mater. **4**, 053210 (2016).

[99] S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. I. J. Probert, K. Refson, and M. C. Payne, Z. Krist.-Cryst. Mater. **220**, 567 (2005).

[100] C. B. Barber, D. P. Dobkin, and H. T. Huhdanpaa, ACM Trans. Math. Softw. **22**, 469 (1996).

[101] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[102] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal 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, J. Phys.: Condens. Matter **21**, 395502 (2009).

[103] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. D. Corso, S. de Gironcoli, P. Delugas, R. A. DiStasio, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia, M. Kawamura, H.-Y. Ko, A. Kokalj, E. Küçükbenli, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N. L. Nguyen, H.-V. Nguyen, A. O. de la Roza, L. Paulatto, S. Poncé, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A. P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu, and S. Baroni, J. Phys.: Condens. Matter **29**, 465901 (2017).

[104] A. B. Migdal, Sov. Phys. JETP **7**, 996 (1958).

[105] G. M. Eliashberg, Sov. Phys. JETP **11**, 696 (1960).

[106] A. Sanna, J. A. Flores-Livas, A. Davydov, G. Profeta, K. Dewhurst, S. Sharma, and E. K. U. Gross, J. Phys. Soc. Jpn. **87**, 041012 (2018).

[107] The ELK FP-LAPW code, http://elk.sourceforge.net/.

[108] M. Wierzbowska, S. de Gironcoli, and P. Giannozzi, arXiv:cond-mat/0504077.

[109] P. B. Allen and R. C. Dynes, Phys. Rev. B **12**, 905 (1975).

[110] Input files for the structures predicted in this work can be found at https://doi.org/10.17863/CAM.48347.

[111] G. D. Gaspari and B. L. Gyorffy, Phys. Rev. Lett. **28**, 801 (1972).

[112] www.csd3.cam.ac.uk

[113] www.dirac.ac.uk.

[114] W. L. McMillan, Phys. Rev. **167**, 331 (1968).

[115] J. J. Hopfield, Phys. Rev. **186**, 443 (1969).

[116] D. A. Papaconstantopoulos, B. Klein, M. J. Mehl, and W. E. Pickett, Phys. Rev. B **91**, 184511 (2015).

[117] J. J. Sakurai and J. Napolitano, Scattering theory, *Modern Quantum Mechanics*, 2nd ed. (Cambridge University Press, Cambridge, 2017), pp. 386–445.