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
(Dated: January 28, 2020)

Searching 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 towards materials stabilised 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 which can operate closer to ambient conditions. The output of these models informs structure searches, from which we identify and screen stable candidates before performing electron-phonon calculations to obtain $T_c$. Hydrides of alkali and alkaline earth metals are identified as particularly promising; a $T_c$ of up to 115 K is calculated for RbH$_{12}$ at 50 GPa and a $T_c$ of up to 90 K is calculated for CsH$_7$ at 100 GPa.

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 pre-compressed”, are predicted to exhibit similar behaviour in experimentally-accessible regimes [2]. In recent years, potential superconductivity has been investigated in many compressed hydrides, including scandium [3], sulfur [4, 5], yttrium [6, 7], calcium [8], thorium [9], pnictogen [10, 11], praseodymium [12, 13], cerium [14, 15], neodymium [16] and iron hydrides [17–19]. Several reviews summarising recent developments in the field are available [20–22]. 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 [23], doping known binaries with more electronegative elements to make ternary hydrides [24, 25] and mapping out alchemical phase diagrams [26].

Experimental measurements of superconductivity in high-pressure hydrides have helped to address several misconceptions about conventional superconductivity, fuelling hope that it may be achieved at ambient temperature and waving a definitive farewell to the Cohen-Anderson limit [27]. Theoretical studies have also demonstrated that the structures and superconducting properties of real materials can now be accurately predicted from first principles.

Machine learning has previously been used in modelling superconductors, with a focus on predicting the maximum obtainable critical temperature for a given composition [28]. However, looking at a collection of data from 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 structure searches and subsequent electron-phonon calculations, with the aim of extending the operation of hydride superconductors towards ambient conditions.

II. TRENDS IN HYDRIDES

A large amount of computational - and some experimental - data for the binary hydrides is available in the literature [3, 8–13, 15–17, 19, 23–25, 39–85] (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 non-hydrogen 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 critical temperature and operational pressure models based on a set of easily-obtained material descriptors. For a particular element E and corresponding binary hydride EH$_n$ these descriptors are

- Hydrogen content ($n$)
- Van der Waals radius of E
- Atomic number of E
- Mass number of E
- Numbers of $s$, $p$, $d$ and $f$ electrons in the electron configuration of E

We construct models using two different methods: stochastic optimization of a sequential neural network and a generalized linear regression. Once we have constructed a model, we apply it to all materials with the chemical composition EH$_n$, where E is any element in the periodic table and $n \in [1, 2, \ldots, 32]$. Selecting those
FIG. 1. The critical temperatures of binary hydrides at various pressures found in the literature are shown as circles. Materials on the frontier towards ambient conditions are labeled. New phases found in this work (as reported in Tab. I) are shown as translucent squares; of note is the Immm phase of RbH_{12} (labeled, see also Fig. 6), which extends the frontier significantly.

that are predicted to exhibit superconductivity closest to ambient conditions serves to guide our searches for new binary hydrides.

A. Sequential neural network

We train a sequential neural network, 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 [87]. The input (and expected output) data is positive definite and therefore not normally distributed and has a non-zero mean, prompting the use of self-normalizing activation functions [88, 89] to improve training behaviour. Since the number of data points is comparable to the number of parameters in our network, the risk of overfitting 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 increasing, allowing us to choose the model parameters from the training epoch for which the validation set error is minimal. This process is repeated several times and the predictions cross-validated. We also apply $L_2$ regularization to the parameters in the intermediate dense nodes to decrease the propensity towards over-fitting.
Predicted pressure (GPa) correlation = 0.88
Observed pressure (GPa) Observed (K) Predicted (K)

Predicted pressure (GPa) correlation = 0.76
Observed pressure (GPa) Observed (K) Predicted (K)

FIG. 3. Behaviour of our machine learning model of superconducting pressures and temperatures 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.

B. Generalized linear regression

To provide a baseline against which to test the neural network, we also carry out a generalized linear regression on the data in Fig. 1, constructing a model of the following form

\[ P(d) = \sum_{ij} c_{ij}(P) d_i^p d_j^p \]  

\[ T_c(d) = \sum_{ij} c_{ij}(T) d_i^p d_j^p \]

essentially building a linear combination of powers \( p_j \in \{1, 2, 1/2, 1/4\} \) of the basic descriptors \( d_i \) (introduced in section II). The derived regression model coefficients \( c_{ij} \) are provided in the Supplementary Information [90].

C. Model behaviour

The basic behaviour of the machine learning model is shown in Fig. 3. The generalised linear regression behaves similarly but, unsurprisingly, its predictions correlate less well with the dataset. To gain further insight we define a measure of distance \( D = |(P, T_c - 293)| \). This distance decreases as we move towards ambient conditions from the pressure-temperature region containing the known hydrides (c.f Fig. 1). In Fig. 4 we plot the distribution of material properties for hydrides predicted to exhibit superconductivity closest to ambient conditions. Both the machine learning model and the generalized linear regression predict the heavy alkali and alkaline earth metal hydrides to be the best candidates. Indeed, in Fig. 4 the number of close-to-ambient materials decreases as we go across the periodic table, until we hit the next alkali metal. 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. The predicted optimal (minimum \( D \)) hydride compositions from the machine learning model are shown across the whole periodic table in Fig. 5.

III. STRUCTURE SEARCHING

The models constructed in the previous section point towards the alkali and alkaline earth metal hydrides as being some of the best candidates. From these, we studied caesium and rubidium hydrides; these systems were chosen due to their predicted proximity to superconductivity at ambient conditions (see Fig. 5) and the fact that they have not been studied extensively in the past. Our structure searching calculations were performed using \textit{ab initio} random structure searching (AIRSS) [91, 92] and the plane-wave pseudopotential code castep [93]. Since our models suggest that a wide range of stoichiometries should be considered, convex hulls were constructed using AIRSS and qhull [94] in order to identify those which are stable at 50, 100 and 200 GPa [95]. The Perdew-Burke-Ernzerhof (PBE) generalised gradient approximation [96], CASTEP QC5 pseudopotentials, a 400 eV plane-wave cut-off and a \( k \)-point spacing of \( 2\pi \times 0.05 \text{ Å}^{-1} \) were
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 is shown for each element. Elements are colored according to the predicted distance from ambient-condition-superconductivity $D = |(P, T_c - 293)|$.

used in all searches.

Caesium polyhydrides have been studied previously using structure searching methods in Ref. [96], though potential superconductivity was not investigated. At 150 GPa, CsH$_3$, CsH$_4$, CsH$_5$, CsH$_7$ and CsH$_{14}$ were found to lie on the convex hull [96], which partially agrees with our hulls calculated at 100 and 200 GPa [90]. The structures of rubidium polyhydrides under pressure were studied in Ref. [97], which found RbH$_5$ to be stable across a large pressure range and also found ranges of thermodynamic stability for RbH$_3$ and RbH$_9$.

IV. SCREENING CANDIDATES

A. Pre-screening: enthalpy and metallicity

Once stable stoichiometries were identified, additional AIRSS searches using the same parameters and pseudopotentials were performed at 100 and 200 GPa; RbH$_3$, RbH$_5$, RbH$_9$, RbH$_{11}$, RbH$_{12}$, CsH$_5$, CsH$_7$, CsH$_{13}$ and CsH$_{15}$ were investigated [90]. Selecting the low-enthalpy structures for each of these stoichiometries, we performed geometry optimisations using QUANTUM ESPRESSO [98, 99] in order to obtain plots of enthalpy as a function of pressure. These calculations used a PBE functional, a 950 eV cut-off, ultrasoft pseudopotentials [99] and a $k$-point spacing of $\sim 2\pi \times 0.02$ Å$^{-1}$. We also calculated the electronic density of states (DOS) at 50 GPa and 150 GPa and evaluated this quantity at the Fermi energy in order to identify metallic structures.

Consideration of the metallicity of structures lead us to focus our attention on RbH$_{12}$, RbH$_3$, CsH$_7$, CsH$_{15}$ and RbH$_{11}$. Study of the enthalpy plots for these stoichiometries then allowed us to select the subset of structures which are competitive in the low-pressure region (25-125 GPa).

B. Testing further screening techniques

Superconductivity in hydrides generally requires hydrogenic states close to the Fermi level. Intuitively, this often means avoiding structures with molecular-character H$_2$ units - an idea backed up by the fact that cage-like structures exhibit some of the highest predicted (and recorded) $T_c$s. Therefore, a simple way of screening for potential high-temperature hydride superconductors could involve calculating the hydrogen-derived DOS normalised by the total DOS at the Fermi energy, $N_H(E_F)/N(E_F)$. Here, we also calculate hydrogen-derived electron-phonon coupling estimates ($\eta_H$) using Gaspari-Gyorffy theory [100] and test whether these two quantities could act as future screening methods to be applied before performing expensive electron-phonon calculations.
We implemented Gaspari-Gyorffy theory within the ELK code\textsuperscript{101}. The basics of this theory and its use here are explained in Appendix \textsuperscript{I}. Our tests on various hydride systems show that $N_H(E_F)/N(E_F)$ and $\eta_H$ may give an indication of whether a given hydride will exhibit high-temperature superconductivity \textsuperscript{90}. The two quantities often predict the same general trends, but the DOS ratio is cheaper to calculate since it can be obtained using a pseudopotential code. A full evaluation of these potential screening methods is given in the supplementary information.

V. ELECTRON-PHONON COUPLING AND SUPERCONDUCTIVITY

The Hamiltonian of a coupled electron-phonon system is given by \textsuperscript{90}

$$H = \sum_{kn} \epsilon_{nk} c_{nk}^\dagger c_{nk} + \sum_{q\nu} \omega_{q\nu} (a_{q\nu}^\dagger a_{q\nu} + \frac{1}{2}) + \frac{1}{\sqrt{N_p}} \sum_{kqmn\nu} g_{mn\nu}(k,q)c_{m,k+q}^\dagger c_{nk}(a_{q\nu} + a_{q\nu}^\dagger)$$ \hspace{1cm} (3)

In this work, we calculate the electronic Kohn-Sham eigenvalues $\epsilon_{nk}$, phonon frequencies $\omega_{q\nu}$, and electron-phonon coupling constants $g_{mn\nu}(k,q)$ appearing in $H$ from first-principles using density functional perturbation theory (DFPT) as implemented in the QUANTUM ESPRESSO code \textsuperscript{98, 99}. The resulting Hamiltonian is then treated within Migdal-Eliashberg theory \textsuperscript{102–104} where we solve the Eliashberg equations using the ELK code \textsuperscript{101}. 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, an 820 eV plane-wave cut-off, and a $q$-point grid with a spacing of $\approx 2\pi \times 0.1 \, \text{Å}^{-1}$ (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-delta smearing width necessary to calculate the critical temperature \textsuperscript{105}. Full electron-phonon calculations were performed for a range of competitive RbH$_{12}$, CsH$_7$ and RbH$_3$ structures. The results of these calculations are shown in Tab. I. A few of the structures found have low $T_c$ values resulting from unfavorable hydrogen arrangements; an example is the layered-hydrogen structure of RbH$_3$ shown in Fig. 6. It is perhaps unsurprising that our machine learning model suggests such compositions, despite their resulting unfavorable structures, as it is trained mostly on cage-like materials. As a result, the model may implicitly assume that compositions it is given will behave as if they adopt cage-like arrangements, leading to an overestimation of $T_c$. Despite this, most of the structures found are high-$T_c$ cage-like superconductors, of which the $Immm$ phase of RbH$_{12}$ (see Fig. 6) is particularly interesting due to its location in Fig. 1. Supplementing structure searching techniques with predictions from machine learning have allowed us to find these phases much more efficiently than would have otherwise been possible. It can be seen from Fig. 1 that the hydrides resulting from this work are biased towards ambient conditions when compared to the dataset as a whole.

Our calculated $T_c$ values also allow us to further assess the potential screening methods tested in this work; we observe that $\eta_H$ correctly predicts $T_c$ ordering for the RbH$_{12}$ structures at fixed pressure, as was the case for LaH$_{10}$ and YH$_{10}$ \textsuperscript{90}. $N_H(E_F)/N(E_F)$ appears to be much less predictive for the CsH$_7$ structures and the performance of $\eta_H$ is also mixed \textsuperscript{90}. The use of these quantities for screening applications may therefore require fur-

| 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$_7$     | Pmmm       | 100           | 33        |
| CsH$_7$     | Pmmm       | 100           | 10        |
| CsH$_7$     | Cm          | 100           | 5         |
| CsH$_7$     | Cm21       | 100           | 89        |
| RbH$_3$     | Pmmm       | 100           | 0         |
| RbH$_3$     | Cmmm       | 100           | 0         |

TABLE I. Critical temperatures calculated for promising hydride compositions. This data is also shown in Fig. 1 for comparison with previous literature. The crystal structures for these compositions are available \textsuperscript{106}.

FIG. 6. Left: The crystal structure of the $Immm$ phase of RbH$_{12}$ found in our AIRSS search, showing the hydrogen cage structure. Right: the crystal structure of the $Pmmm$ phase of RbH$_3$ found in our AIRSS search, showing the layered hydrogen structure.
ACKNOWLEDGEMENTS

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 number 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 (www.csd3.cam.ac.uk), provided by Dell EMC and Intel using Tier-2 funding from the EPSRC (capital grant EP/P020259/1), and DiRAC funding from the STFC (www.dirac.ac.uk).

APPENDIX

I. GASPARI-GYORFFY THEORY

McMillan \cite{107} showed that for strong-coupled superconductors the electron-phonon coupling constant, \( \lambda \), can be expressed as

\[
\lambda = 2 \int \frac{d\omega \alpha^2(\omega)F(\omega)}{\omega} = \frac{N(E_F)}{M} \langle I^2 \rangle
\]

where \( \eta \) is the so-called Hopfield parameter. Hopfield was one of the first to stress the importance of the local environment in determining \( \lambda \) \cite{108}. In situations where we have nearly perfect separation of vibrational modes into those of different atomic character (such as we may see in hydrides) we can write

\[
\lambda = \sum_j \lambda_j = \sum_j \frac{\eta_j}{M_j \langle \alpha_j^2 \rangle}
\]

where \( j \) is the atom type.

The quantity \( \langle I^2 \rangle \) appearing in Eq. 4 can be approximated using Gaspari-Gyorffy (GG) theory \cite{100}. Recent work has emerged using this theory for metal hydrides under high pressure \cite{35} \cite{109} despite it originally being designed for elemental transition metals. The theory, based on the rigid muffin-tin approximation (RMTA), relies on several approximations \cite{109} 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 calculated \( \langle I^2 \rangle \) for each atom type and hence obtain \( \eta_j \). The GG equation is

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

where \( N_{l+1}^{(1)} \) is the free-scatterer DOS given by

\[
N_{l+1}^{(1)} = \sqrt{\frac{E_F}{\pi}} (2l+1) \int_0^{R_{MT}} R_l^2(r,E_F)r^2dr
\]

and \( \delta_l \) 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 characterise the long-distance behaviour of the wavefunction, can be written in terms of the logarithmic derivative of the radial wavefunction

\[
tan\delta_{l}(R_{MT},E_F) = \frac{j_l'(kR_{MT}) - j_l(kR_{MT})L_l(R_{MT},E_F)}{n_l(kR_{MT}) - n_l(kR_{MT})L_l(R_{MT},E_F)}
\]
where $k = \sqrt{E_F}$, $L_j = \frac{n_j^2}{\eta}$ 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. [8] to obtain the phase shifts [10].

Since $M_j(\omega_j^2)$ is often considerably smaller for H than for the other components, it is clear from Eq. [5] that the hydrogen atoms can provide a considerable fraction of $\lambda$ even if the Hopfield parameter of the other atom type is similar in magnitude. Calculating $\eta_H$ 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. It is in this context that we assess GG theory in this work.

[1] N. W. Ashcroft, Phys. Rev. Lett. 21, 1748 (1968).
[2] N. W. Ashcroft, Phys. Rev. Lett. 92, 187002 (2004).
[3] A. P. Durajski and R. Szczesniak, Supercond. Sci. and Tech. 27, 115012 (2014).
[4] D. Duan, Y. Liu, F. Tian, D. Li, X. Huang, Z. Zhao, H. Yu, B. Liu, W. Tian, and T. Cui, Nature - Scientific Reports 4, 6968 (2014).
[5] A. P. Drozdov, M. I. Eremets, I. A. Troyan, V. Ksenofontov, and S. I. Shylin, Nature 525, 73 (2015).
[6] 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).
[7] D. Y. Kim, R. H. Scheicher, and R. Aluja, Phys. Rev. Lett. 103, 077002 (2009).
[8] Y. Li, J. Hao, H. Liu, J. S. Tse, Y. Wang, and Y. Ma, Scientific Reports 5, 9948 (2015).
[9] H. Liu, I. I. Naumov, R. Hoffmann, N. W. Ashcroft, and R. J. Hemley, Proc. Nat. Acad. Sci. 114, 6990 (2017).
[10] F. Peng, Y. Sun, C. J. Pickard, R. J. Needs, Q. Wu, and Y. Ma, Phys. Rev. Lett. 119, 107001 (2017).
[11] C. Heil, S. di Cataldo, G. B. Bachelet, and L. Boeri, Phys. Rev. B 99, 220502(R) (2019).
[12] I. A. Troyan, D. V. Semenok, A. G. Kvashnin, A. G. Ivanova, V. B. Prakapenka, E. Greenberg, A. G. Gavrilik, I. S. Lyubutin, V. V. Struzhkin, and A. R. Oganov, arXiv preprint arXiv:1908.05334 (2019).
[13] P. P. Kong, V. S. Minkov, M. A. Kuzovnikov, S. P. Besedin, A. P. Drozdov, 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 569, 528 (2019).
[14] 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, A. Majumdar, J. S. Tse, M. Wu, and Y. Yao, Phys. Rev. B 96, 201107 (2017).
[15] A. G. Kvashnin, I. A. Kruglov, D. V. Semenok, and A. R. Oganov, J. Phys. Chem. C 122, 4731 (2018).
[16] C. Heil, G. B. Bachelet, and L. Boeri, Physical Review B 97, 214510 (2018).
[17] D. Duan, Y. Liu, Y. Ma, Z. Shao, B. Liu, and T. Cui, Natl. Sci. Rev. 4, 121 (2017).
[18] E. Zurek and T. Bi, J. Chem. Phys. 150, 050901 (2019).
[19] J. A. Flores-Livas, L. Boeri, A. Sanna, G. Profeta, R. Arita, and M. I. Eremets, arXiv preprint arXiv:1905.06693 (2019).
[20] L. Boeri and G. B. Bachelet, J. Phys: Condens. Matt. 31, 234002 (2019).
[21] C. J. Pickard, I. Errea, and M. I. Eremets, Annual Review of Condensed Matter Physics 11, 57 (2020).
[22] T. Bi, N. Zarifi, T. Terpstra, and E. Zurek (Elsevier, 2019).
[23] P.-H. Chang, S. Silayi, D. Papaconstantopoulos, and M. Mehl, Journal of Physics and Chemistry of Solids , 109315 (2019).
[24] Y. Sun, J. Lv, X. Xie, H. Liu, and Y. Ma, Phys. Rev. Lett. 123, 097001 (2019).
[25] C. Heil and L. Boeri, Phys. Rev. B 92, 060508 (2015).
[26] M. L. Cohen and P. W. Anderson, in AIP Conference Proceedings, Vol. 4 (AIP, 1972) pp. 17–27.
[90] See supplementary information.

[91] C. J. Pickard and R. J. Needs, J. Phys: Condens. Matt. 23, 053201 (2011).

[92] R. J. Needs and C. J. Pickard, APL Materials 4, 053210 (2016).

[93] S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. I. J. Probert, K. Reñon, and M. C. Payne, Zeitschrift für Kristallographie-Crystalline Materials 220, 567 (2005).

[94] C. B. Barber, D. P. Dobkin, and H. T. Huhdanpaa, ACM Trans. on Mathematical Software 22, 469 (1996).

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

[96] A. Shamp, J. Hooper, and E. Zurek, Inorganic chemistry 51, 9333 (2012).

[97] J. Hooper and E. Zurek, Chemistry–A European Journal 18, 5013 (2012).

[98] P. Giannozzi, S. 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, Journal of Physics: Condensed Matter 21, 395502 (19pp) (2009).

[99] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Curnimeo, A. D. Corso, S. de Gironcoli, P. Delugas, R. A. D. Jr, 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. Kicikbenli, 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. Xu, and S. Baroni, Journal of Physics: Condensed Matter 29, 465901 (2017).

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

[101] http://elk.sourceforge.net/ the ELK FP-LAPW code.

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

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

[104] A. Sanna, J. A. Flores-Livas, A. Davydov, G. Profeta, K. Dewhurst, S. Sharma, and E. K. U. Gross, Journal of the Physical Society of Japan 87, 041012 (2018).

[105] M. Wierzbowska, S. de Gironcoli, and P. Giannozzi, arXiv:cond-mat/0504077 (2005).

[106] Crystal structure data can be found at https://doi.org/10.17863/CAM.48347.

[107] W. L. McMillan, Physical Review 167, 331 (1968).

[108] J. J. Hopfield, Physical Review 180, 443 (1969).

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

[110] J. J. Sakurai and J. Napolitano, “Scattering theory,” in Modern Quantum Mechanics (Cambridge University Press, 2017) p. 386–445, 2nd ed.

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

[112] P. Morel and P. W. Anderson, Phys. Rev. 125, 1263 (1962).
SUPPLEMENTARY INFORMATION

I. GENERALIZED LINEAR REGRESSION

The goal of generalized linear regression is to make a linear model of the relationship

\[ O_n = O_n(f_1,n, f_2,n, \ldots, f_N,n) \]  

(9)

The linear model looks like

\[ P_n = \sum_i c_i f_{i,n} + \epsilon_n \]  

(10)

where \( c_i \) are the linear coefficients of the features. In matrix notation this looks like

\[ P_n = f c + \epsilon_n \]

where we call \( f \) the feature matrix. We pick the coefficient vector \( c \) by minimizing the modulus of the residual vector; \( c = \text{arg min}_c |\epsilon|^2 \). In order to simplify the model, we may bias the coefficient vector using Tikhonov regularization. This involves adding a cost function, \( |T c|^2 \), which is large when the coefficient vector has many significant entries. The matrix \( T \) is known as the Tikhonov matrix. This results in the minimization

\[ c = \text{arg min}_c |\epsilon|^2 + |T c|^2 \]

(11)

Minimizing with respect to \( c_n \) we require

\[ \frac{\partial L}{\partial c_n} = 2[f_{in}(f_{ik} c_k - P_i) + T_{im} T_{ij} c_j] = 0 \]  

(13)

In matrix notation this reads

\[ f^T (f c - p) + T^T T c = 0 \implies c = (f^T f + T^T T)^{-1} f^T p \]  

(14)

which gives us the optimal feature coefficients \( c \) for the model in Eq. (10). The coefficients for the linear regression model discussed in the main text are given in table II.

II. CONVEX HULLS

We constructed (static-lattice) convex hulls using AIRSS [91] and qhull [94] for various hydride systems (at 50, 100 and 200 GPa). As detailed in the main text, we produced convex hulls for systems that were indicated as promising for superconductivity closer to ambient conditions by our machine learning model and generalised linear regression.

The systems we focus on here are Cs-H, Ra-H and Rb-H and the hulls produced inform our further structure searches.
### a) Pressure model

| Term in expansion | Coefficient | Norm |
|-------------------|-------------|------|
| Hydrogens/ion     | 16.91408    | 6.24775 |
| Hydrogens/ion\(^{0.5}\) | 16.43228    | 2.49955 |
| Hydrogens/ion\(^{0.25}\) | 15.24157    | 1.58100 |
| S electrons\(^{0.25}\) | 14.13379    | 1.16718 |
| S electrons\(^{0.5}\) | 13.85894    | 1.36230 |
| S electrons\(^{2}\) | 13.51666    | 1.85586 |
| Hydrogens/ion\(^{2}\) | 12.43521    | 39.03435 |
| VdW radius\(^{0.25}\) | 12.23301    | 3.87678 |
| VdW radius\(^{0.5}\) | 11.49370    | 15.02940 |
| VdW radius\(^{2}\) | 10.18405    | 225.88288 |
| Ion Mass\(^{2}\) | -10.09502   | 15057.39153 |
| P electrons\(^{2}\) | 9.17555     | 0.39189 |
| Ion atomic number\(^{2}\) | -8.16066    | 2560.72470 |
| VdW radius\(^{2}\) | 8.12073     | 51023.07678 |
| F electrons\(^{2}\) | -7.88385    | 1.76577 |
| Ion atomic number\(^{0.25}\) | 7.45944     | 2.66714 |
| Ion Mass\(^{0.25}\) | 7.03580     | 3.32827 |
| F electrons\(^{0.5}\) | -6.88904    | 1.32882 |
| F electrons\(^{0.25}\) | -6.27614    | 1.15275 |
| P electrons\(^{0.5}\) | 4.73478     | 0.62601 |
| D electrons\(^{0.25}\) | 4.25008     | 1.26306 |
| D electrons\(^{0.5}\) | 3.91072     | 1.59532 |
| Ion atomic number\(^{0.5}\) | 3.72079     | 7.11362 |
| D electrons\(^{2}\) | 3.46107     | 2.54505 |
| P electrons\(^{2}\) | 3.33830     | 0.79121 |
| Ion Mass\(^{0.5}\) | 3.03423     | 11.07739 |
| Ion Mass\(^{2}\) | -2.21899    | 122.70856 |
| F electrons\(^{2}\) | 2.17818     | 3.11793 |
| Ion atomic number\(^{2}\) | -1.16444    | 50.60360 |
| D electrons\(^{4}\) | 1.06367     | 6.47725 |
| P electrons\(^{4}\) | -0.34647    | 0.15358 |

### b) Temperature model

| Term in expansion | Coefficient | Norm |
|-------------------|-------------|------|
| Hydrogens/ion     | 16.34886    | 6.24775 |
| Hydrogens/ion\(^{2}\) | 15.72061    | 39.03435 |
| F electrons\(^{0.5}\) | -14.62359   | 1.76577 |
| F electrons\(^{0.25}\) | -12.34196   | 1.32882 |
| Hydrogens/ion\(^{0.5}\) | 11.42645    | 2.49955 |
| F electrons\(^{0.25}\) | -11.15791   | 1.15275 |
| S electrons\(^{2}\) | 8.18184     | 3.44420 |
| Hydrogens/ion\(^{0.25}\) | 7.70346     | 1.58100 |
| D electrons\(^{2}\) | -6.27340    | 2.54505 |
| S electrons\(^{0.25}\) | 5.97147     | 1.85586 |
| D electrons\(^{0.5}\) | 5.01341     | 1.26306 |
| S electrons\(^{0.5}\) | 4.45893     | 1.36230 |
| VdW radius\(^{0.25}\) | 4.12143     | 51023.07678 |
| VdW radius\(^{0.5}\) | 3.58428     | 225.88288 |
| S electrons\(^{0.25}\) | 3.55632     | 1.16718 |
| Ion Mass\(^{0.25}\) | 3.32475     | 3.32827 |
| Ion atomic number\(^{0.5}\) | 3.28585     | 7.11362 |
| VdW radius\(^{0.5}\) | 3.26998     | 15.02940 |
| Ion Mass\(^{0.5}\) | 3.25892     | 11.07739 |
| Ion Mass\(^{2}\) | -3.24468    | 15057.39153 |
| VdW radius\(^{0.25}\) | 3.10714     | 3.87678 |
| P electrons\(^{2}\) | 2.72494     | 0.39189 |
| Ion atomic number\(^{2}\) | 2.35759     | 50.60360 |
| Ion atomic number\(^{2}\) | -2.20360    | 2560.72470 |
| Ion Mass\(^{2}\) | 2.15283     | 122.70856 |
| F electrons\(^{2}\) | 1.90288     | 3.11793 |
| D electrons\(^{2}\) | -0.87843    | 6.47725 |
| P electrons\(^{0.25}\) | -0.39708    | 0.79121 |
| D electrons\(^{0.5}\) | 0.25192     | 1.59532 |
| P electrons\(^{0.5}\) | 0.13546     | 0.62601 |
| P electrons\(^{2}\) | -0.00791    | 0.15358 |

**TABLE II.** Generalized linear regression model for the critical temperature and corresponding operational pressure of hydrides. Each parameter is normalized before the regression by dividing by the average value for the dataset (right-hand column).

As shown in Fig. 8, CsH, CsH\(_3\), CsH\(_3\)CsH, CsH\(_7\), CsH\(_{13}\) and CsH\(_{15}\) are on the hull at 100 GPa. We can see from Fig. 9 that CsH\(_7\) and CsH\(_{15}\) remain on the hull at 200 GPa, while CsH\(_{13}\) is 14 meV/atom above it. We also note that CsH\(_5\) is on the hull at 200 GPa. We chose to investigate CsH\(_3\), CsH\(_7\), CsH\(_{13}\) and CsH\(_{15}\) further.

**FIG. 9.** Convex hull for Cs-H system at 200 GPa.
As shown in Fig. 10, RbH, RbH$_5$, RbH$_{10}$ and RbH$_{21}$ are on the hull at 50 GPa. There are many compositions that are very close to the hull at this pressure, including: RbH$_{16}$ (<1 meV/atom), RbH$_6$ (1 meV/atom), RbH$_{23}$ (4 meV/atom), RbH$_{19}$ (7 meV/atom), RbH$_{17}$ (8 meV/atom), RbH$_8$ (9 meV/atom), RbH$_7$ (9 meV/atom), RbH$_3$ (9 meV/atom), RbH$_{14}$ (11 meV/atom), RbH$_{12}$ (12 meV/atom), RbH$_{18}$ (13 meV/atom), RbH$_{15}$ (14 meV/atom) and RbH$_{11}$ (14 meV/atom), where the number in brackets indicates distance above the hull.

As shown in Fig. 11, RbH, RbH$_3$, RbH$_5$, RbH$_9$ and RbH$_{12}$ are on the hull at 100 GPa. At 200 GPa (Fig. 12), RbH$_{12}$ is just 10 meV/atom above the hull, while RbH$_5$ and RbH$_9$ are 27 meV/atom and 30 meV/atom above the hull, respectively. RbH$_3$ and RbH$_{11}$ are on the hull at 200 GPa. We chose to investigate RbH$_3$, RbH$_5$, RbH$_9$, RbH$_{11}$ and RbH$_{12}$ further.

Radium hydrides were also indicated as promising candidates by the models and a Ra-H convex hull at 100 GPa was initially calculated. However, these hydrides were eliminated from further calculations as all isotopes of radium are radioactive. The Ra-H convex hull at 100 GPa is shown in Fig. 13, although radium hydrides were
not investigated any further in this work.

III. ENTHALPY PLOTS AND ELECTRONIC DENSITY OF STATES

After constructing convex hulls and deciding which stoichiometries to investigate, we performed further AIRSS searches. From these searches we then selected the lowest enthalpy candidates and relaxed these structures within QUANTUM ESPRESSO [98, 99] using more stringent parameters (as detailed in the main text) at a range of pressures. These calculations (and the following electron-phonon calculations) all used scalar-relativistic, ultrasoft PBE pseudopotentials downloaded from https://www.quantum-espresso.org/pseudopotentials/ps-library. The enthalpy plots are displayed below. Before performing expensive electron-phonon calculations to assess potential superconductivity, we also wanted to ensure that the materials were metallic at the pressures of interest. In the tables below we therefore present the electronic density of states (DOS) at the Fermi level ($E_F$) for each of the structures at 50 GPa and 150 GPa, calculated using QUANTUM ESPRESSO.

| Symmetry | Formula units | DOS($E_F$)(states/eV/f.u.) at 50GPa | DOS($E_F$)(states/eV/f.u.) at 150GPa |
|----------|---------------|----------------------------------|------------------------------------|
| Amm2     | 2             | 0.5977E-03                       | 0.2622                             |
| C2/m     | 2             | 0.1092E-07                       | 0.1559E-03                        |
| Cmcm     | 2             | 0.8085E-10                       | 0.2435                             |
| Cmm2     | 2             | 0.4350E-04                       | 0.1667E-01                        |
| P21/m    | 2             | 0.6636E-09                       | 0.1352                             |

| Symmetry | Formula units | DOS($E_F$)(states/eV/f.u.) at 50GPa | DOS($E_F$)(states/eV/f.u.) at 150GPa |
|----------|---------------|----------------------------------|------------------------------------|
| Cm       | 3             | 0.100005                          | 0.41645                            |
| Cmc21    | 2             | 0.1763                            | 0.3571                             |
| I4mmm    | 2             | 0.7785E-01                        | 0.1357                             |
| I4mmmm   | 2             | 0.6736E-01                        | 0.2685                             |
| Imm2     | 2             | 0.8825E-02                        | 0.29315                            |
| P4mm     | 2             | 0.016045                          | 0.1736                             |
| C2/c     | 2             | 0.3339E-07                        | 0.3146                             |
| I4mm     | 1             | 0.7988E-01                        | 0.1812                             |
| P1       | 3             | 0.1898                            | 0.6102                             |
| P31/m    | 3             | 0.6961E-07                        | 0.7867E-02                        |

| Symmetry | Formula units | DOS($E_F$)(states/eV/f.u.) at 50GPa | DOS($E_F$)(states/eV/f.u.) at 150GPa |
|----------|---------------|----------------------------------|------------------------------------|
| C2/c     | 2             | 3.7610312E-28                    | 0.00013395                         |
| C2/m     | 1             | 1.13605E-06                      | 0.1712                             |
| C2/m     | 2             | 0.3994E-21                       | 0.00002664                        |
| Cm       | 1             | 0.5642E-06                       | 0.1891                             |
| Cmcm     | 2             | 2.1535181E-27                   | 0.030615                           |
| P6m2     | 2             | 0.7362E-05                       | 0.1161                             |
| P21/m    | 2             | 0.7736E-04                       | 0.5526                             |
| P63/mmc  | 2             | 0.7946E-15                       | 0.1680                             |
| R3m      | 1             | 0.49935E-03                     | 0.4264                             |
| Cmce21   | 2             | –                                | 0.6071                             |
FIG. 14. Enthalpy vs pressure plot for CsH$_5$ structures from AIRSS

FIG. 15. Enthalpy vs pressure plot for CsH$_7$ structures from AIRSS

FIG. 16. Enthalpy vs pressure plot for CsH$_{13}$ structures from AIRSS
FIG. 17. Enthalpy vs pressure plot for CsH$_{15}$ structures from AIRSS

FIG. 18. Enthalpy vs pressure plot for RbH$_3$ structures from AIRSS

FIG. 19. Enthalpy vs pressure plot for RbH$_2$ structures from AIRSS
FIG. 20. Enthalpy vs pressure plot for RbH$_9$ structures from AIRSS

FIG. 21. Enthalpy vs pressure plot for RbH$_{11}$ structures from AIRSS

FIG. 22. Enthalpy vs pressure plot for RbH$_{12}$ structures from AIRSS
### CsH15 structures

| Symmetry       | Formula units | DOS($E_F$)(states/eV/f.u.) at 50GPa | DOS($E_F$)(states/eV/f.u.) at 150GPa |
|----------------|---------------|------------------------------------|-------------------------------------|
| C2/m           | 2             | 0.4644E-07                         | 0.031865                            |
| Cm             | 1             | 2.9626484E-21                      | 0.3559E-02                          |
| Cm             | 2             | 2.5831273E-21                      | 0.6947E-02                          |
| Cmcm           | 2             | 0.9725E-10                         | 0.02838                             |
| Ima2           | 2             | 0.1350E-19                         | 0.2176                              |
| Immm           | 1             | 7.3203181E-22                      | 0.136455E-04                        |
| P1             | 1             | 0.7504E-26                         | 0.7516E-02                          |
| P21            | 2             | 0.3003E-04                         | 0.7784                              |
| P2/c           | 2             | 0.7705E-08                         | 0.1093E-05                          |
| C2221          | 2             | 0.1044E-15                         | 0.0058975                           |
| C/c            | 2             | 0.5056E-06                         | 0.2833E-02                          |
| P2/m           | 2             | 0.2747E-26                         | 0.0029778                           |
| Pm             | 2             | 0.1263E-21                         | 0.10651                             |
| Pm21           | 2             | 0.109812E-05                       | 0.4108                              |
| Cmm2           | 2             | 0.8539E-53                         | 5.1819705E-14                       |
| P1             | 2             | 0.3265E-04                         | 0.9962                              |
| P2/m           | 2             | 0.1520E-30                         | 0.2716E-18                          |
| Pm             | 2             | 0.7671E-04                         | 0.7977E-14                          |
| Pc             | 2             | 0.1534E-16                         | 0.5661E-07                          |

### RbH3 structures

| Symmetry       | Formula units | DOS($E_F$)(states/eV/f.u.) at 50GPa | DOS($E_F$)(states/eV/f.u.) at 150GPa |
|----------------|---------------|------------------------------------|-------------------------------------|
| C2/c           | 2             | 0.1127E-08                         | 0.3533E-41                          |
| Ccccm          | 2             | 0.8055E-01                         | 0.2819                              |
| Cmmm           | 1             | 0.3865E-01                         | 0.1284                              |
| I41/amd        | 2             | 0.4432E-01                         | 0.3181E-03                          |
| Imma           | 2             | 0.9423E-01                         | 0.2798                              |
| P2/m           | 3             | 0.1244                              | 0.4055                              |
| P6222          | 3             | 0.1189                              | 0.3862                              |
| P6mmmm         | 3             | 0.2180                              | 0.5852                              |
| Pmma           | 2             | 0.6661E-01                         | 0.2889                              |
| R3m            | 3             | 0.1720                              |                                     |

### RbH5 structures

| Symmetry       | Formula units | DOS($E_F$)(states/eV/f.u.) at 50GPa | DOS($E_F$)(states/eV/f.u.) at 150GPa |
|----------------|---------------|------------------------------------|-------------------------------------|
| C2             | 2             | 2.1730038E-35                      | 0.1004E-13                          |
| C2/m           | 2             | 0.1588E-25                         | 0.1780                              |
| Cm             | 3             | 2.926505E-35                       | 0.2993E-04                          |
| Cmcm           | 2             | 9.9900005E-57                      | 1.347721E-16                        |
| Cmmm           | 2             | 0.1123E-14                         | 0.1138                              |
| Fmmm           | 3             | 9.8400152E-36                      | –                                   |
| P1             | 3             | 0.6238E-01                         | 0.2623                              |
| P2/c           | 2             | 0.216509E-04                       | 0.1228                              |
| P31            | 3             | 0.3055E-19                         | 0.1736E-01                          |
| P32            | 3             | 2.8353161E-17                      | 0.1875E-01                          |
### RbH₉ structures

| Symmetry | Formula units | DOS($E_F$)(states/eV/f.u.) at 50GPa | DOS($E_F$)(states/eV/f.u.) at 150GPa |
|----------|---------------|-------------------------------------|-------------------------------------|
| C2       | 3             | 2.4139685E-11                       | 0.1720E-03                          |
| C2/c     | 2             | 0.2987E-76                          | 0.1975E-01                          |
| C2/m     | 1             | 2.129423E-11                        | 0.2724                               |
| Cc       | 2             | 0.4171E-07                          | 0.2632E-01                          |
| Cccm     | 2             | 0.7734E-40                          | 0.6590E-01                          |
| Cm       | 2             | 1.759505E-08                        | 0.2133E-01                          |
| Cmcm     | 2             | 3.0176185E-12                       | 0.2983                               |
| Cmmm     | 2             | 0.3413E-04                          | 0.3413E-04                          |
| Cmc      | 2             | 0.91805                              | 0.7946                               |
| Cmc1     | 2             | 1.153                                | 0.95985                              |
| Cmc2     | 2             | 0.5849                               | 0.4330                               |
| Cmcm     | 2             | 0.90505                              | 0.6612                               |
| Fmmm     | 1             | 0.5736                               | 0.4078                               |
| Imm      | 1             | 0.5564                               | 0.3666                               |
| Imm1     | 1             | 0.5515                               | 0.3664                               |
| Imm2     | 2             | 0.8955                               | 0.8697                               |
| R3m      | 2             | 0.6024                               | 0.40845                              |
| P6₃/mmc  | 2             | –                                    | 0.7904                               |
| Cmc2     | 2             | –                                    | 1.00595                              |

### RbH₁₁ structures

| Symmetry | Formula units | DOS($E_F$)(states/eV/f.u.) at 50GPa | DOS($E_F$)(states/eV/f.u.) at 150GPa |
|----------|---------------|-------------------------------------|-------------------------------------|
| Cmc2₁    | 2             | 0.3969E-22                          | 0.1478                              |
| Imm2     | 3             | 0.1788                               | 1.0960                              |
| Imm      | 1             | 0.3101                               | 0.2932                              |
| Imm      | 2             | 0.3102                               | 0.2911                              |
| Pmnn     | 2             | 0.6679E-10                           | 0.3645                               |

### RbH₁₂ structures

| Symmetry | Formula units | DOS($E_F$)(states/eV/f.u.) at 50GPa | DOS($E_F$)(states/eV/f.u.) at 150GPa |
|----------|---------------|-------------------------------------|-------------------------------------|
| C2       | 2             | 0.91805                              | 0.7946                               |
| C2/c     | 2             | 1.153                                | 0.95985                              |
| C2/m     | 1             | 0.5849                               | 0.4330                               |
| Cmc      | 2             | 0.90505                              | 0.6612                               |
| Fmmm     | 1             | 0.5736                               | 0.4078                               |
| Imm      | 1             | 0.5564                               | 0.3666                               |
| Imm      | 2             | 0.5515                               | 0.3664                               |
| P2/c     | 2             | 0.8955                               | 0.8697                               |
| R3m      | 2             | 0.6024                               | 0.40845                              |
| P6₃/mmc  | 2             | –                                    | 0.7904                               |
| Cmc2₁    | 2             | –                                    | 1.00595                              |
IV. TESTING FURTHER SCREENING METHODS

In this section we compare $\eta_j$ values calculated using Gaspari-Gyorffy, (GG) theory [100] and DOS ratios ($N^H(E_F)/N(E_F)$) for various hydrides to calculated $T_c$s in order to assess these as potential screening methods. The preliminary test cases chosen are:

- H$_3$X (X=S, As, Se, Br, Sb, Te and I) at a range of pressures, so as to compare to Ref. [35] (which used GG theory implemented in the NRL code)
- various structures of LaH$_{10}$ and YH$_{10}$, so as to compare to Ref. [14]
- $I4/mmm$-FeH$_5$, so as to compare to the work of Ref. [17] which reported superconductivity in this system and Ref. [28] which (correctly) disputed this claim
- $Im3m$-YH$_6$, so as to compare to Ref. [14]
- The ternary hydride $Fd3m$-Li$_2$MgH$_{16}$ to compare to Ref. [16]
- $Im3m$-CaH$_6$ to compare to Ref. [15]
- $Im3m$-H$_3$S to compare to Ref. [4]
- C$\text{mcm}$-AsH, $C2/c$-AsH$_8$, $Pmmn$-SbH$_3$ and $P6_3/mmc$-SbH$_4$ to compare to Ref. [18]
- $I4/mmm$-NdH$_4$, $C2/c$-NdH$_7$ and $P6_3/mmc$-NdH$_9$ to compare to Ref. [22]
### Test cases

| Structure       | Pressure/GPa | Calculated $T_c$ (K) | $\eta_H$ (eV/A^2) | $\eta_X$ (eV/A^2) | $X^H(E_F)/N(E_F)$ |
|-----------------|--------------|----------------------|-------------------|-------------------|-------------------|
| $Fm\bar{3}m$-LaH$_{10}$ | 250          | 234-259              | 13                | 5.58906          | 0.28996          |
| $P6_3/mmc$-LaH$_{10}$    | 250          | 224-250              | 13                | 5.45099          | 0.29676          |
| $C2/m$-LaH$_{10}$       | 250          | 205-228              | 13                | 4.48316          | 0.28758          |
| $Fm\bar{3}m$-YH$_{10}$   | 400          | 247-282              | 13                | 11.45718         | 0.32971          |
| $Cmcm$-YH$_{10}$         | 400          | 233-270              | 13                | 10.49255         | 0.32986          |
| $P6_3/mmc$-YH$_{10}$     | 400          | 223-262              | 13                | 9.85728          | 0.30800          |
| $Im\bar{3}m$-YH$_6$      | 160          | 223-247              | 13                | 4.44873          | 0.34787          |
| $I4/mmm$-FeH$_5$         | 150          | 33.6-45.8            | ≤1                | 0.03795          | 0.02791          |
| $I4/mmm$-NdH$_4$         | 100          | 31-51                | 22                | 0.09756          | 0.01437          |
| $C2/c$-NdH$_7$           | 100          | ~0.01                | 22                | 0.16644          | 0.01006          |
| $P6_3/mmc$-NdH$_9$       | 120          | 15-25                | 22                | 0.16994          | 0.01995          |
| $Im\bar{3}m$-CaH$_6$     | 150          | 220-235              | 15                | 2.67506          | 0.51126          |
| $Fd\bar{3}m$-Li$_2$MgH$_{16}$ | 250          | 473 [36]             |                  | 5.02036          | 0.33024          |
| $Im\bar{3}m$-H$_3$S      | 200          | 191-204              | 4                 | 5.21134          | 0.25999          |
| $Cmcm$-AsH               | 350          | 20.2-21.2            | 15                | 2.19678          | 0.06730          |
| $C2/c$-AsH$_8$           | 350          | 141 [18]             |                  | 4.51675          | 0.17811          |
| $Pmmm$-SbH$_3$           | 300          | 25.9 [18]            |                  | 3.91360          | 0.11186          |
| $P6_3/mmc$-SbH$_4$       | 300          | 93.9 [18]            |                  | 3.69085          | 0.11451          |

Comparing our calculated $\eta_j$ values for $H_2X$ (X=S, As, Se, Br, Sb, Te and I) to those in Ref. 35 allowed a rough validation of our implementation of GG theory against the NRL version. An extremely low $\eta_H$ is calculated for FeH$_5$, indicating that this material will not be a good superconductor - this result reflects the full electron-phonon calculations of Ref. 28. YH$_6$ is predicted to be a good superconductor at 160 GPa, in line with recent theoretical and experimental papers.

The average phonon frequencies for different structures are often very similar when we are considering the same stoichiometry at the same pressure. If we assume that the average phonon frequencies in such cases are exactly equivalent, we arrive at a potential way of cheaply estimating the $T_c$ order between structures. Using our results to rank the various LaH$_{10}$ and YH$_{10}$ structures (considering the fact that the $\eta_H$ term has the most impact on the total $\lambda$), we arrive at $Fm\bar{3}m > P6_3/mmc > C2/m$ for LaH$_{10}$ and $Fm\bar{3}m > Cmcm > P6_3/mmc$ for YH$_{10}$. Agreement with the calculated $T_c$ order is seen in both cases (while the $N^H(E_F)/N(E_F)$ predicts slightly incorrect ordering for LaH$_{10}$). In cases where we want to compare different stoichiometries/pressures or we want to approximate $T_c$ magnitudes, we can combine the calculation of $\eta_j$ with phonon calculations. Phonon calculations are still considerably cheaper than full electron-phonon calculations.

### V. SELECTING PROMISING CANDIDATES

Based on the initial DOS calculations, RbH$_{12}$ was chosen as one of the most promising compositions as it has several metallic structures down to 50 GPa. Looking at the RbH$_{12}$ enthalpy plot, the lowest enthalpy structures across the whole pressure range are $Cmcm$ (2 formula units), $Cmcm$ (3 formula units), $C2/m$ (1 formula unit), $Imm$ (1 formula unit) and $Imm$ (2 formula units). On refining the structures, $Cmcm$-3 reduced to the structure with 2 formula units; similar behaviour is seen for the two $Imm$ structures. All of these structures are metallic at 50 GPa according to the initial DOS calculations. We calculate the $N^H(E_F)/N(E_F)$ ratios and $\eta_H$ values for these structures at 50, 100 and 150 GPa in order to test these screening methods - these values are displayed in the table below. C2/m-1 and Imm-1 have similar (and fairly high) DOS ratios and $\eta_H$ even at these low pressures. $Cmcm$-2 also appears promising. We perform full electron-phonon superconductivity calculations on these 3 structures at various pressures.

According to the initial DOS calculations, CsH$_7$ is also a promising stoichiometry. The CsH$_7$ enthalpy plot shows a group of structures that are competitive at $P < 90$ GPa ($Imm$, $Imm$, $Cm$, $P4mm$) and then another competitive group at slightly higher pressures ($Imm$, $P1$, $Cmc$, $Imm$). The table below shows DOS ratios and $\eta_H$ values for these structures at 100 GPa. Many of the structures look fairly promising and we proceed with running full electron-phonon super...
conductivity calculations. The test screening methods described here suggest that Cmc21-2, Imm-1 and P1-3 may give the highest Tc at a given pressure.

We also consider CsH15. None of the structures are particularly metallic at 50 GPa according to initial DOS calculations, with a few becoming metallic at 150 GPa - these include Ima2-2, P21-2, P2/m-2, Pm-2 and Cmc21-2. Looking at the enthalpy plot, the most competitive structures are Cm-1, Cm-2, P1-1, P1-2, Pmn21-2, Imm-1, Pm-2, P21/m-2, P21-2 and Cmc21-2. These results show that Pm-2, Cmc21-2, P21/m-2 and P21-2 are the only structures which are both competitive and metallic. We calculate the DOS ratio and scattering theory parameters for some of the competitive structures at 150 GPa, but do not go on to perform electron-phonon calculations.

We also consider RbH3. We first refine the set of structures we will look at by considering the enthalpy plots. Ignoring the most energetically unfavourable structures at the static-lattice level (here, the highest 3) and noting that Cmmm-2 and Cmmm-1 correspond to the same structure, we are left with I41/amd-2, C2/c-2, Cmmm-1, Cccm-2, P2/m-3, P6222-3 and Pmma-2. However, I41/amd-2 and C2/c-2 were not metallic enough at 100 GPa to be able to calculate η values or DOS ratios; this reflects the results shown in the initial DOS table and explains why they aren’t included in the table below.

We performed electron-phonon calculations for two of the predicted best structures of RbH3, but found their superconducting temperatures to be extremely low; this can be explained by their layered nature as addressed in the main text. We do not consider the other structures any further.

We also consider RbH11 - for this stoichiometry, Cmc21-2 is most stable at lower pressures and Pmmm-2 and Imm-1 are competitive at higher pressures. Cmc21 and Pmmm were not metallic enough at the pressure considered in order to be included in the table below. We did not perform electron-phonon calculations for any RbH11 structures - Imm-1 could be promising, but since the enthalpy plot suggests it becomes competitive only above ∼125 GPa, we do not consider it further.

As addressed in the main text, our calculated Tc values allow us to assess the screening methods tested in this work. The ηH values correctly predict Tc ordering for the RbH12 structures at fixed pressure and NH(E_F)/N(E_F) also comes close to doing so. NH(E_F)/N(E_F) appears to be much less predictive for the CsH7 structures and the performance of ηH is also mixed. We note that in some cases the trends predicted by the DOS ratio and the scattering calculation disagree considerably (see I4/mmm-CsH7 and P1-CsH7). This can occur in cases where the material is not very metallic and the DOS ratio is therefore artificially inflated.
VI. ELECTRON-PHONON COUPLING IN DFT

Typically, within DFT the nuclear coordinates, \( R \), are treated as fixed and the electronic Kohn-Sham system is solved within the fixed nuclear potential. In order to calculate the effects of electron-phonon coupling from within the DFT formalism we must consider leading-order corrections to the Born-Oppenheimer approximation in nuclear displacements. Expanding our Kohn-Sham potential in terms of these displacements leads to

\[
V_{KS}(R + \delta R) = V_{KS}(R) + \sum_{\kappa,p} \frac{\partial V_{KS}}{\partial R_{\kappa,p}} \cdot \delta R_{\kappa,p} + O(\delta R^2).
\]

where \( R_{\kappa,p} \) is the position of atom \( \kappa \) in unit cell \( p \). An atomic displacement of an atom can be written in terms of phonon creation and annihilation operators as

\[
\delta R_{\kappa,p} = \frac{1}{\sqrt{N_p M_\kappa}} \sum_{q} e^{iq \cdot R_p} \frac{1}{\sqrt{2\omega_{q\kappa}}} \left( a_{q\kappa} + a_{q\kappa}^\dagger \right) e_{\kappa\nu}(q)
\]

where \( e_{\kappa\nu}(q) \) and \( \omega_{q\nu} \) are, respectively, the eigenvector and frequency of the phonon mode with creation operator \( a_{q\kappa} \). \( R_p \) is the position of the \( p \)th unit cell within the periodic cell, of which there are \( N_p \). \( M_\kappa \) is the mass of atom \( \kappa \). Substituting this into Eq. 15 we obtain

\[
V_{KS}(R + \delta R) = V_{KS}(R) + \frac{1}{\sqrt{N_p}} \sum_{q} G_{q\kappa}(a_{q\kappa} + a_{q\kappa}^\dagger)
\]

where

\[
G_{q\kappa} = \frac{1}{\sqrt{2\omega_{q\kappa}}} \sum_{\kappa} e_{\kappa\nu}(q) \cdot \sum_{p} e^{iq \cdot R_p} \frac{\partial V_{KS}}{\partial R_{\kappa,p}}
\]

This allows us to write down the resulting electron-phonon coupling Hamiltonian in second-quantized form as

\[
H_{ep}(\delta R) = \sum_{nkn'k'} \langle n,k|V_{KS}(R + \delta R) - V_{KS}(R)|n',k'\rangle c_{nk}^\dagger c_{n'k'}
\]

\[
= \frac{1}{\sqrt{N_p}} \sum_{q} \left[ \sum_{nkn'k'} \langle n,k|G_{q\kappa}|n',k'\rangle c_{nk}^\dagger c_{n'k'} \right] (a_{q\kappa} + a_{q\kappa}^\dagger)
\]

where \( c_{nk}^\dagger \) creates a Kohn-Sham electron in orbital \( n \), wavevector \( k \) (i.e. occupies the Bloch state \( u_{nk}(x) \exp(ik \cdot x)/\sqrt{N_p} \)). Substituting our definition of \( G_{q\kappa} \) we have

\[
\langle n,k|G_{q\kappa}|n',k'\rangle = \frac{1}{\sqrt{2\omega_{q\kappa}}} \sum_{\kappa} e_{\kappa\nu}(q) \cdot \sum_{p} e^{iq \cdot R_p} \langle n,k|\frac{\partial V_{KS}}{\partial R_{\kappa,p}}|n',k'\rangle
\]

Now

\[
\langle n,k|\frac{\partial V_{KS}}{\partial R_{\kappa,p}}|n',k'\rangle = \int N_p^{-1/2} u_{nk}^*(x)e^{-ik \cdot x} \frac{\partial V_{KS}}{\partial R_{\kappa,p}}(x)N_p^{-1/2} u_{n'k'}(x)e^{ik' \cdot x} \ dx
\]

\[
= \int N_p^{-1/2} u_{nk}^*(x-R_p)e^{-ik \cdot (x-R_p)} \frac{\partial V_{KS}}{\partial R_{\kappa,p}}(x-R_p)\times N_p^{-1/2} u_{n'k'}(x-R_p)e^{ik' \cdot (x-R_p)} \ dx
\]

\[
= \epsilon_R(p-k') \int_{1^{st} \text{ unit-cell}} u_{nk}^*(x)e^{-ik \cdot x} \frac{\partial V_{KS}}{\partial R_{\kappa,0}}(x)u_{n'k'}(x)e^{ik' \cdot x} \ dx
\]

where in the last line we have used Bloch’s theorem and the fact that

\[
\frac{\partial V_{KS}}{\partial R_{\kappa,0}}(x-R_p) = \frac{\partial V_{KS}}{\partial R_{\kappa,0}}(x)
\]

where \( R_{\kappa,0} \) is the position of atom \( \kappa \) in the first unit cell. We may now write Eq. 20 as

\[
\langle n,k|G_{q\kappa}|n',k'\rangle = \frac{1}{\sqrt{2\omega_{q\kappa}}} \sum_{\kappa} e_{\kappa\nu}(q) \cdot \sum_{p} \epsilon_{m+k,q,n,k}(a_{q\kappa} + a_{q\kappa}^\dagger)
\]

where we have defined

\[
G_{q\kappa,uc} = \frac{1}{\sqrt{2\omega_{q\kappa}}} \sum_{\kappa} e_{\kappa\nu}(q) \cdot \frac{\partial V_{KS}}{\partial R_{\kappa,0}}
\]

This allows us to write down the Hamiltonian for an interacting Kohn-Sham-electron-phonon system, correct to first order in electron-phonon coupling constants \( g_{mn\nu}(k,q) = \langle m,k+q|G_{q\kappa,uc}|n,k\rangle \):

\[
H = \sum_{kn} e_{kn} c_{nk}^\dagger c_{nk} + \sum_{q} \omega_{q} \left( \frac{1}{2} \epsilon_{q,k+q,n,k} a_{q\kappa}^\dagger a_{q\kappa} - \frac{1}{2} \right) + \frac{1}{\sqrt{N_p}} \sum_{kqmnu} g_{mn\nu}(k,q) c_{m,k+q,n,k}^\dagger a_{q\kappa} a_{q\kappa}^\dagger.
\]

This does not allow us to write down the Hamiltonian for an interacting Kohn-Sham-electron-phonon system, correct to first order in electron-phonon coupling constants \( g_{mn\nu}(k,q) = \langle m,k+q|G_{q\kappa,uc}|n,k\rangle \):
with each phonon mode, \( \lambda_{q\nu} \), and the isotropic Eliashberg spectral function, \( \alpha^2 F(\omega) \)

\[
\lambda_{q\nu} = \frac{1}{N(\epsilon_F)\omega_{q\nu}\Omega_{BZ}} \times \sum_{nm} \int_{BZ} |g_{mn\nu}(k, q)|^2 \delta(\epsilon_{n,k} - \epsilon_F)\delta(\epsilon_{m,k+q} - \epsilon_F)dk
\]

\[
\alpha^2 F(\omega) = \frac{1}{2\Omega_{BZ}} \sum_{\nu} \int_{BZ} \omega_{q\nu}\lambda_{q\nu}\delta(\omega - \omega_{q\nu})dq
\] (27)

from which we may approximate the critical temperature using the McMillan formula \[111\]

\[
T_c = \frac{\omega_{\log}}{1.2} \exp \left( \frac{-1.04(1 + \lambda)}{\lambda(1 - 0.62\mu^*) - \mu^*} \right)
\] (29)

where

\[
\lambda = \sum_{q\nu} \lambda_{q\nu} = 2 \int \alpha^2 F(\omega) \frac{d\omega}{\omega},
\] (30)

\[
\omega_{\log} = \exp \left( \frac{2}{\lambda} \int \alpha^2 F(\omega) \log(\omega) \frac{d\omega}{\omega} \right)
\] (31)

and \( \mu^* \) is the Morel-Anderson pseudopotential \[112\], which is typically treated as an empirical parameter with values between 0.1 and 0.2.