Functional form of the superconducting critical temperature from machine learning

S. R. Xie, G. R. Stewart, J. J. Hamlin, P. J. Hirschfeld, and R. G. Hennig

Phys. Rev. B **100**, 174513 — Published 18 November 2019

DOI: 10.1103/PhysRevB.100.174513

Functional Form of the Superconducting Critical Temperature from Machine Learning

S. R. Xie,[1, 2] G. R. Stewart,[3] J. J. Hamlin,[3] P. J. Hirschfeld,[3] and R. G. Hennig[1, 2, *]

[1] *Department of Materials Science and Engineering,*
University of Florida, Gainesville FL 32611, USA
[2] *Quantum Theory Project, University of Florida, Gainesville FL 32611, USA*
[3] *Department of Physics, University of Florida, Gainesville FL, 32611 USA*

(Dated: September 13, 2019)

Predicting the critical temperature T_c of new superconductors is a notoriously difficult task, even for electron-phonon paired superconductors for which the theory is relatively well understood. Early attempts to obtain a simple T_c formula consistent with strong-coupling theory, by McMillan and Allen and Dynes, led to closed-form approximate relations between T_c and various measures of the phonon spectrum and the electron-phonon interaction appearing in Eliashberg theory. Here we propose that these approaches can be improved with the use of machine learning algorithms. As an initial test, we train a model for identifying low-dimensional descriptors using the $T_c < 10$ K data tested by Allen and Dynes, and show that a simple analytical expression thus obtained improves upon the Allen-Dynes fit. Furthermore, the prediction for the recently discovered high T_c material H_3S at high pressure is quite reasonable. Interestingly, T_c's for more recently discovered superconducting systems with a more two-dimensional electron-phonon coupling, which do not follow Allen and Dynes' expression, also do not follow our analytic expression. Thus, this machine learning approach appears to be a powerful method for highlighting the need for a new descriptor beyond those used by Allen and Dynes to describe their set of isotropic electron-phonon coupled superconductors. We argue that this machine learning method, and its implied need for a descriptor characterizing Fermi surface properties, represents a promising new approach to superconductor materials discovery which may eventually replace the serendipitous discovery paradigm begun by Kamerlingh Onnes.

Keywords: superconducitivity, machine learning

I. INTRODUCTION

Discovery of new superconductors has historically proceeded largely serendipitously, with guidance from rules of thumb (such as Matthias' e/a ratio) rather than many-body and ab-initio theory. The space of possible materials to search for new superconductors is vast, considering that many discoveries in the last thirty years are multinary compounds. Thus, it is desirable to appeal to recent computational developments, aided by theory, to assist this process. The history of ab-initio and materials-genome type approaches to superconducting materials discovery has recently been reviewed by Norman,[1] Pickett,[2] and Duan et al.[3]

While initially, success in prediction (as opposed to analysis after discovery, *i.e.*, postdiction) was rare to nonexistent, more recently the potential for theory to aid in the discovery of new high-temperature superconductors was dramatically demonstrated by the prediction and subsequent discovery, in 2015, of superconductivity at $T_c = 200$ K in H_3S at about 150 GPa pressure.[4] This experiment shattered the assumed ceiling for T_c in electron-phonon superconductors[5] and was followed by the recent discovery of superconductivity in compressed lanthanum hydride at 250 K,[6,7] also preceded by a theoretical prediction.[8,9] Recent computational approaches to hydride superconductivity have been reviewed in Refs. 10–12.

Despite these undeniable successes and the demonstration that the old assumed limit of 35-40 K for T_c due to the exchange of phonons, often quoted without proof in early cuprate debates, is incorrect, these experiments do not provide a clear strategy to optimize T_c in the vast phase space of materials. This is at least partially due to an inability to identify the correct materials descriptors, parameters directly reflecting the underlying mechanism of superconductivity. For some classes of materials, *e.g.*, thermoelectrics, considerable progress has been made in high-throughput approaches identifying simple observables recorded in databases that contribute to a material's figure of merit.[13] For superconductivity, however, such approaches[14] are considerably more difficult, both because the theory is more complex, and the figure of merit, T_c, depends extremely sensitively on the underlying interactions.

This last difficulty is clear already from the Bardeen-Cooper-Schrieffer (BCS) theory of superconductivity,[15] among whose great successes was the proof that for weak attractive interactions, fermions pair with an instability that corresponds to an essential singularity in the dimensionless coupling constant λ, leading to the well-known expression,

$$T_c \simeq 1.14\, \omega_D\, e^{-\frac{1}{\lambda}}, \tag{1}$$

where ω_D is the Debye frequency. BCS theory is successful because it predicts superconducting properties accurately in terms of measured T_c's, but the essential singularity alone suggests that accurate calculations will be difficult. Besides, Eq. (1) is strictly valid only in the weak coupling limit $\lambda \ll 1$ and if the Coulomb interaction is neglected.

The inadequacy of the BCS expression for T_c was already clear by the late 1960's, when McMillan[5] introduced an improved formula based on Eliashberg theory,[16] relating T_c to a small number of physical quantities calculated from the effective electron-phonon interaction $\alpha^2F(\omega)$ that could in principle be extracted from tunneling data,[17]

$$T_c \simeq \frac{\omega_D}{1.45} \exp\left(-\frac{1.04(1+\lambda)}{\lambda - \mu^*(1+0.62\lambda)}\right), \tag{2}$$

where μ^* is the Coulomb pseudopotential. This expression, although it was probably only meant to apply to a finite range of λ, predicts a saturation of T_c in the strong-coupling limit for fixed ω_D. Dynes[18] later replaced the prefactor $\omega_D/1.45$ of the McMillan equation (2) with $\langle\omega\rangle/1.20$, where $\langle\omega\rangle$ is the first moment of the distribution $g(\omega) = 2/(\lambda\omega)\alpha^2F(\omega)$.

Based on a reanalysis of Eliashberg theory and newly available computational checks in special cases, Allen and Dynes[19] proposed an alternate approximate formula,

$$T_c = \frac{f_1 f_2 \omega_{\log}}{1.20} \exp\left(-\frac{1.04(1+\lambda)}{\lambda - \mu^*(1+0.62\lambda)}\right), \tag{3}$$

where f_1 and f_2 are correction factors that depend on $\lambda, \mu^*, \omega_{\log}$, and $\bar{\omega}_2$. The frequencies $\bar{\omega}_n$ are the n^{th} root of the n^{th} moment of $g(\omega)$. The additional tunnelling-derived parameters ω_{ph}, defined as the high-frequency cutoff in $\alpha^2F(\omega)$, and η, defined as McMillan-Hopfield parameter, also appear in their discussion. They showed that the expression (3) fit the T_c of a variety of superconductors known at the time, using data derived from tunneling, and that it implied the absence of any maximum T_c, except that caused by the competition between λ and $\omega_{\log} \equiv \exp\langle\ln\omega\rangle$, where the average is taken over $g(\omega)$. Unlike the McMillan expression, which saturates to a constant value as $\lambda \to \infty$, the Allen-Dynes equation obeys an asymptotic result of Eliashberg theory, that $T_c \sim \sqrt{\lambda}$ as $\lambda \to \infty$ with other parameters fixed.

The Allen-Dynes equation has played a crucial role in the discussion of high-temperature superconductivity and indeed is often used to extract quoted values of λ in the literature for materials where tunneling data is not available. Nevertheless, it is important to recall that it has been derived from Eliashberg theory, which itself is implemented with various approximations, *e.g.*, the momentum dependence of the electron-phonon interaction was often neglected in early studies. The full evaluation of the Eliashberg equation is computationally expensive and not currently suitable for high-throughput superconductor discovery. It would be highly desirable to develop an expression for T_c that generalizes the Allen-Dynes equation and is applicable over a large range of parameters that are cheap to compute to guide such searches.

In this letter, we use modern machine learning techniques to critically examine the Allen-Dynes equation to demonstrate that similar analytic expressions can be obtained from relatively small experimental datasets. These symbolic regression techniques are *analytical* in nature, meaning they search for analytical relations between a minimal set of features, *i.e.*, physical parameters, and the desired properties.[20–22] Specifically, we apply the Sure-Independence Screening and Sparsifying Operator (SISSO) method[22] to estimate T_c from λ, μ^*, and $\omega_{\log}$ with the goal to obtain an equation of similar or enhanced performance to the one proposed by Allen and Dynes.[19] We find that we can improve on the Allen-Dynes fit to strong-coupling superconductors, with a smaller set of descriptors. More interestingly, the approach identifies outliers like MgB_2, T_c=39 K, which suggests the importance of new physics essential to high T_c that needs to be incorporated in an improved formula to guide the search for new electron-phonon superconductors in materials space.

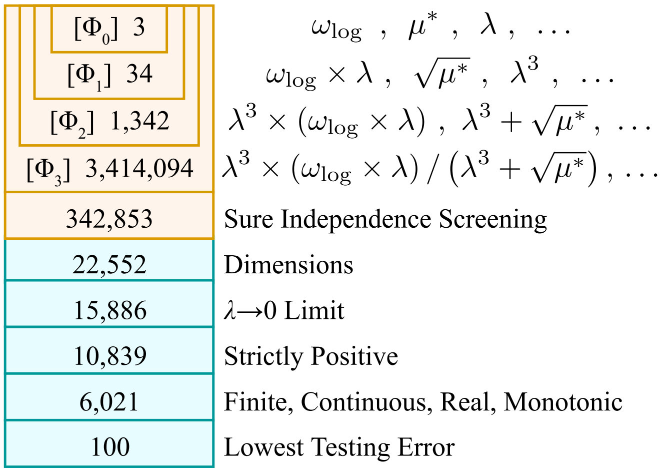
[Φ_0] 3
$\omega_{\log}$, μ^* , λ , ...
[Φ_1] 34
$\omega_{\log} \times \lambda$, $\sqrt{\mu^*}$, λ^3 , ...
[Φ_2] 1,342
$\lambda^3 \times (\omega_{\log} \times \lambda)$, $\lambda^3 + \sqrt{\mu^*}$, ...
[Φ_3] 3,414,094
$\lambda^3 \times (\omega_{\log} \times \lambda) / (\lambda^3 + \sqrt{\mu^*})$, ...
342,853
Sure Independence Screening
22,552
Dimensions
15,886
$\lambda \to 0$ Limit
10,839
Strictly Positive
6,021
Finite, Continuous, Real, Monotonic
100
Lowest Testing Error

FIG. 1. Beginning with feature space Φ_0, consisting of $\omega_{\log}$, λ and μ^*, each additional tier Φ_i is constructed by applying 4 binary operators (+, −, ×, /) and 7 unary operators (exp, log, $\sqrt{\ }$, $\sqrt[3]{\ }$, $^{-1}$, ^{2}, ^{3}) to features from preceding tiers. This procedure is applied up to level Φ_3, after which sure-independence screening is applied to eliminate features with correlation factors (inner product) below 0.5 with respect to T_c. Physical constraints as listed are then applied to further reduce the feature space. We fit coefficients to the 6,021 features and obtain the 100 models with lowest root-mean-square error in predictions on the testing set.

II. METHODS

To generate models for predicting T_c, we apply recently developed methods of equation-based machine learning, subject to physical constraints. In the SISSO approach, the predictive models are expressed as analytical formulas relating physical quantities with algebraic operations such as addition and exponentiation. Given a tabulated set of scalar-valued physical quantities, or features, the SISSO method constructs additional features by iteratively applying operations from a specified set, *e.g.*, +, ×, exp, $\sqrt{\ }$, ^{2}.

To pinpoint the best equations, the SISSO method em-

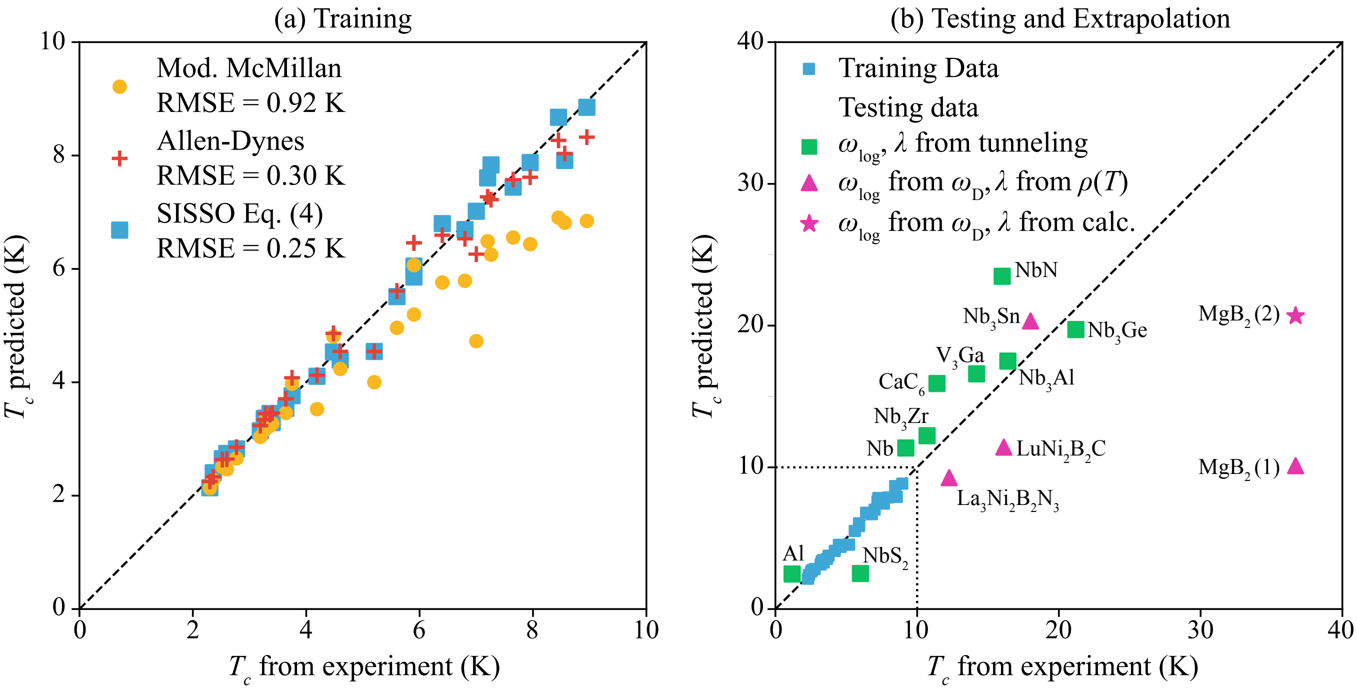

FIG. 2. Machine learning of optimal *analytical* expression for T_c as a function of three parameters ($\omega_{\log}$, λ, and μ^*) trained on the low-T_c dataset of Allen and Dynes[19] using the SISSO algorithm.[22] (a) The 3-parameter machine-learned equation results in a smaller RMSE than the 4-parameter Allen-Dynes or the 3-parameter McMillan equation (b) The testing of the machine-learned equation using nine different superconductors assumes that $\mu^* = 0.1$ and takes $\omega_{\log}$ and λ from tunneling measurements.[23–31] This extrapolation shows larger deviations with a testing RMSE = 3.4 K or 17%. To compare, we also show four materials (Nb_3Sn, MgB_2(1), $La_3Ni_2B_2N_3$, and $LuNi_2B_2C$) for which $\omega_{\log}$ is obtained from low-temperature specific heat measurements and λ from high-temperature resistivity[32] and MgB_2(2), for which λ is from density-functional calculations.[33] The extrapolation reveals two outliers, NbS_2 at low temperatures and MgB_2 at high temperatures.

ploys the sure-independence screening (SIS) method and the sparse-solution algorithm using sparsifying operators (SO) in tandem. After constructing the feature space, the SIS method selects a subspace of features with the largest linear correlation with the target property (T_c), *i.e.*, the largest absolute value of their dot product. The SO step then evaluates all possible combinations of features from the SIS subspace, yielding the optimal least-squares solution and residual. With such a vast feature space, the combinatorial optimization in each SO step relies on L_0 regularization, which penalizes the number of non-zero coefficients. Combined with one numerical prefactor, fit from available data, each feature is used to generate one predictive model.

We benchmark the performance of different models identified by SISSO using leave-one-out cross-validation. Given N available data points, each model is repeatedly fit using $N-1$ points and evaluated with the excluded point. The average evaluation error across N iterations, where each point is tested once, is the leave-one-out cross-validation error. This method can help to maximize the transferability of a model by reducing "overfitting", *i.e.*, models that exhibit low root-mean-square error in predictions on the training data but very high root-mean-square error in the testing data.

We apply the SISSO method to estimate T_c from λ, μ^*, and $\omega_{\log}$ to obtain an equation of similar performance to the one proposed by Allen and Dynes.[19] We use the values of λ, μ^*, and $\omega_{\log}$, and the target property, T_c, from the data for 29 superconducting materials provided by Allen and Dynes (Table I in Ref. 19). Next, we apply the SISSO method with 4 binary operators ($+$, $-$, $\times$, $/$) and 7 unary operators (exp, log, $\sqrt{\ }$, $\sqrt[3]{\ }$, $^{-1}$, ^{2}, ^{3}) three times to generate 3,414,094 features. Fig. 1 shows the rapid growth of the feature space with the number of iterations. Of the initial feature space, we select the equations with the highest linear correlation to T_c using sure-independence screening with a minimum correlation magnitude (inner product) of 0.5. To further reduce the number of features and eliminate unphysical equations, we apply constraints. We select equations that are linearly proportional to $\omega_{\log}$ and obey the proper $\lambda \to 0$ limiting behavior. Additionally, we filter for equations that are strictly positive, real, finite, continuous, and monotonic across the relevant training and testing feature spaces. To evaluate the generalizability and performance of these equations, we compute the error against a testing set of 9 superconductors,[23–31] shown in green in Fig. 2.

Our software that processes the SISSO equations to enforce physical constraints and proper physical dimensions as well as perform linear regression with additional additive and multiplicative numerical coefficients is freely available at Github.[34]

III. RESULTS

A. Optimal T_c Expression

Fig. 2 illustrates the main proof-of-principle result that machine learning can provide an analytic equation of similar performance to the Allen-Dynes equation. The equation-based machine learning uses the values of $\lambda, \omega_{\log}$, and μ^* of the 29 materials in Table I of Allen and Dynes,[19] and neglects the average frequency $\bar{\omega}_2$ that is also used in the Allen-Dynes equation. The SISSO method and subsequent physical constraints lead to the optimal equation,

$$T_c{}^{\mathrm{SISSO}} = 0.0953 \frac{\lambda^4 \omega_{\log}}{\lambda^3 + \sqrt{\mu^*}}. \tag{4}$$

Importantly, Eq. (4) emerged from our approach with the smallest root-mean-square error (RMSE) even before any of the physical constraints summarized in Fig. 1 were applied. Fig. 2(a) compares the performance of this equation with the modified McMillan and Allen-Dynes equations for the measured T_c's of the 29 materials that train the model. The leave-one-out cross-validation RMSE (LOOCV-RMSE) is 0.26 K, which is very similar to the RMSE of this equation evaluated on the training data of 0.25 K. This indicates that the model is not overfit to the training set. The RMSE of 0.25 K of the learned analytic equation is significantly smaller than the RMSE of 0.92 K for the modified McMillan equation, and also slightly lower than the RMSE of 0.30 K for the Allen-Dynes equation. This result is impressive given the use of only 3 parameters and a single numerical coefficient compared to 3 parameters and 4 coefficients for the modified McMillan and 4 parameters and 7 coefficients for the Allen-Dynes equation.

Figure 2(b) shows the testing of Eq. (4) for a variety of other superconductors, mostly of higher T_c. Because μ^* data were not available for these materials, we adopt a constant value of $\mu^* = 0.1$. This procedure introduces some unknown error into the analysis, but despite this, the fit to the new materials is rather good, with a RMSE of only 3.4 K (17%) on the testing set.

It is important to note that Eq. (4) is not derived from any physical theory and therefore may contain some terms that may make no physical sense, *e.g.*, the appearance of the $\sqrt{\mu^*}$ term, which may be a proxy for a constant term due to the small range of data and the paucity of features at this level of learning. The limit $T_c \to 0$ as $\lambda \to 0$ in Eq. (4) even at nonzero μ^* may reflect the lack of data at small coupling. Also, Eq. (4) increases monotonically with λ, with linear behavior at very high couplings. This behavior violates the asymptotic limit of Eliashberg theory, $T_c \sim \sqrt{\lambda}$, built into the Allen-Dynes equation.[19] Again, this disagreement with physics is due to the absence of data points, either in the training or the testing set, which deviate significantly from the linear behavior predicted by Eq. (4).

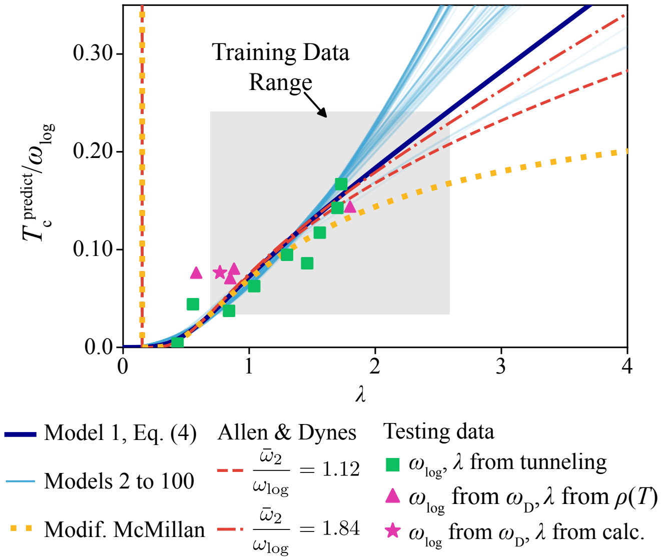

FIG. 3. λ dependence of T_c in the top 100 models, ranked by testing error assuming $\mu^* = 0.1$. Two red curves correspond to the Allen-Dynes equation with the minimum and maximum values of $\bar{\omega}_2/\omega_{\log}$ in the training set. The modified McMillan equation systematically predicts smaller T_c's and over the range of available λ values, the simple machine-learned model closely matches the more complex Allen-Dynes equation.

Fig. 3 shows the functional behavior $T_c(\lambda)$ of the 100 highest-scored equations discovered by SISSO; it is clear that almost all of these equations are equally valid over the range of λ values where data exist. This highlights the need for measurements to determine the materials parameters λ, $\omega_{\log}$, and μ^* reliably for both very low T_c materials, as well as for some of the recently discovered higher-T_c systems.

Fig. 2(b) also shows some dramatic failures of the learned equation, namely for MgB_2 and NbS_2. The probable reasons for these failures are both revealing and reassuring. The point labeled MgB_2(1) with a predicted T_c of 10 K is one where $\omega_{\log}$, a logarithmic average of the electron-phonon interaction function α^2F/ω, was determined from a specific heat measurement of the Debye frequency ω_D, which depends only on the phonon density of states $F(\omega)$. Relating the Debye frequency with $\omega_{\log}$ neglects the difference between the two distributions.[19] This assumption is particularly poor in MgB_2, where high-frequency phonons couple anomalously strongly. In addition, λ was determined from standard expressions for the high-temperature resistivity of a 3D metal. It is well known that MgB_2 has strong 2D character, and that the full momentum and band dependence of the Eliashberg function $\lambda_{n\mathbf{k},n\mathbf{k}'}$ must be accounted for to obtain reasonable values for T_c from first principles.[33] It is interesting to note that if one uses the higher value of λ obtained from Ref. 33 in Eq. (4), one obtains data point $MgB_2(2)$, with the significantly enhanced predicted T_c of 20 K, but still far from the measured value of 40 K and even further from the full Eliashberg calculation of 50 K.[33]

These discrepancies indicate, not surprisingly, that a machine trained on a database of nearly isotropic low-T_c superconductors cannot capture the physics of highly anisotropic higher-T_c materials using the simple averaged descriptors chosen by Allen and Dynes. The same principle apparently applies to NbS_2, which while having a low-T_c is quite 2-dimensional. Nevertheless, Eq. (4) may have significant predictive power extrapolated to higher-T_c 3D systems. To illustrate this extrapolation, we apply Eq. (4) to the two high-pressure hydrides, LaH_{10} and H_3S, taking the values of λ and $\omega_{\log}$ calculated from first principles and $\mu^* = 0.1$. For LaH_{10} at 210 GPa[35] we obtain $T_c = 273$ K, compared to 286 K for the Eliashberg calculation[35] and about 250 K for the experiment at 170 GPa.[6] For H_3S at 140 GPa pressure,[36,37] the predicted T_c from Eq. (4) is 262 K, compared to the measured value of 203 K. This result is similar to the result obtained from the Allen-Dynes equation, but substantially higher than the modified McMillan equation used in Refs. 36 and 37.

B. Dimensionality and Complexity

Despite using one fewer feature, the performance of our machine-learned Eq. (4) is comparable in performance to the Allen-Dynes expression, Eq. (3). We next investigate if increasing the dimensionality and complexity can further increase the performance of the machine-learned expressions.

To assess the performance of descriptors with increased dimensionality and complexity, we first use leave-one-out cross validation using all seven primary features reported by Allen and Dynes in Table I for 29 materials.[19] We note that while Allen and Dynes report these seven parameters for each of the 29 materials, values such as ω_1 and ω_2 were not reported in the literature for most materials in our testing set. When including all seven properties as primary features, Eq. (4) is the equation with the lowest LOOCV-RMSE. The next best SISSO equation satisfying physical constraints is

$$T_c = -0.0591\left(\bar{\omega}_2 - \bar{\omega}_1 - \frac{\bar{\omega}_2}{\lambda}\right)\frac{\lambda^3}{\sqrt[3]{\lambda}}. \tag{5}$$

with a RMSE of 0.27 K and a LOOCV-RMSE of 0.28 K. Among other equations with higher LOOCV-RMSE values, ω_{ph} and η occasionally appear. The observation that Eq. (4) provides the lowest LOOCV-RMSE demonstrates that the machine-learning of analytic relations can select the optimal primary features from a large list of plausible materials parameters.

Next, to assess the utility of increased complexity through additional fitting coefficients, we first followed the approach described in[22] to identify equations with increased descriptor dimensionality n, where n is the number of expressions from Φ_3 used to construct an equation by linear combination. When n is greater than 1, additional terms are iteratively selected using SIS based on the largest correlation with the residual error from each preceding iteration rather than the correlation with the target property. The SO step then pinpoints the best linear combination of terms, optimizing the n fit coefficients. The best two-term and three-term equations identified by SISSO are

$$T_c = 0.0983\,\frac{\lambda^4\omega_{\log}}{\lambda^3+\sqrt{\mu^*}} - 0.0148\,\lambda^2\omega_{\log}^3 e^{-\frac{1}{\mu^*}} \tag{6}$$

and

$$T_c = 0.248\,\frac{\lambda^{\frac{3}{2}}\omega_{\log}}{\lambda+\frac{1}{\lambda}} - 0.0264\,\lambda^2\omega_{\log}^3 e^{-\frac{1}{\mu^*}} + 0.0513\left(\lambda^3\mu^*\omega_{\log} - \lambda^{\frac{4}{3}}\omega_{\log}\right) \tag{7}$$

with RMSEs of 0.21 K and 0.19 K (LOOCV-RMSEs of 0.23 K and 0.20 K), respectively. On the testing set, the equations yield RMSEs of 4.0 K and 7.5 K, respectively. While the training errors are slightly lower than that of Eq. (4), we note that Eqs. (6) and (7) have even more terms with little physical meaning and significantly higher testing RMSEs. Moreover, none of the two- or three-term equations among the best 5,000 identified by SISSO satisfy our desired physical constraints.

As an alternative to a linear combination of models, we also investigate the inclusion of additional fit coefficients beyond the slope and intercept described in.[22] We inserted one additive and one multiplicative coefficient to each occurrence of a primary feature in a SISSO equation. After combining like terms, all remaining coefficients are optimized simultaneously using the Levenberg-Marquardt algorithm. When applied to Eq. (4), the reoptimized equation becomes

$$T_c = 0.715\,\omega_{\log}\frac{(0.507\lambda+0.0436)^4}{(0.828\lambda+0.00637)^3+\sqrt{1.85\mu^*-0.0743}}. \tag{8}$$

with a RMSE of 0.24 K and a LOOCV-RMSE of 0.29 K. On the testing set of 9 materials, this reoptimized equation has a RMSE of 3.5 K. Despite the increase in model complexity from a single numerical coefficient to seven numerical coefficients, which increases the risk of overfitting and reduces the physical interpretability of the equation, the model performance is nearly the same as that of Eq. (4).

IV. CONCLUSION

We have demonstrated that machine learning can discover equations and the relevant physical parameters that describe the dependence of superconducting T_c's on moments of distributions of phonon frequencies and electron-phonon couplings, as used originally by Allen and Dynes in their attempt to understand the systematics of T_c in the framework of Eliashberg theory. While the method is quite successful in predicting known superconductors of the same general type as the original Allen-Dynes dataset, with fewer parameters and only a single numerical coefficient, the existence of a few anomalous outliers suggests that the use of such methods for high-throughput materials discovery will require new descriptors that capture anomalous features, *e.g.*, the anisotropy of the electron-phonon interactions and unusual electronic states that take advantage of them. A natural modern extension of the philosophy of Allen and Dynes is then to calculate from first principles a few key measures of electronic structure crucial for superconductivity, together with the moments discussed above, and apply machine-learning methods as described here. We anticipate that this approach will allow a much more efficient and thorough investigation of materials space than current approaches that rely on fully anisotropic Eliashberg calculations for each material.

V. ACKNOWLEDGEMENTS

We are grateful to L. Boeri, P. Allen, and W. Pickett for valuable discussions. The work by SRX and RGH was supported by the National Science Foundation under grant Nos. PHY-1549132 and DMR-1609306. The work by GRS was performed under the auspice of Basic Energy Sciences, United States Department of Energy, contract number DE-FG02-86ER45268. Partial funding was also provided by the University of Florida Informatics Institute.

[*] rhennig@ufl.edu

[1] M. R. Norman, Rep. Prog. Phys. **79**, 074502 (2016).

[2] W. E. Pickett, arXiv preprint arXiv:1801.00165 (2017).

[3] D. Duan, H. Yu, H. Xie, and T. Cui, J. Supercond. Novel Mag. **32**, 53 (2019).

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

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

[6] A. Drozdov, P. Kong, V. Minkov, S. Besedin, M. Kuzovnikov, S. Mozaffari, L. Balicas, F. Balakirev, D. Graf, V. Prakapenka, *et al.*, Nature **569**, 528 (2019).

[7] 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).

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

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

[10] L. Boeri, arXiv:1903.05708 (2019).

[11] W. Pickett and M. Eremets, Physics Today **72**, 52 (2019).

[12] J. A. Flores-Livas, L. Boeri, A. Sanna, G. Profeta, R. Arita, and M. Eremets, arXiv:1905.06693 (2019).

[13] S. Curtarolo, G. L. Hart, M. B. Nardelli, N. Mingo, S. Sanvito, and O. Levy, Nat. Mater. **12**, 191 (2013).

[14] V. Stanev, C. Oses, A. G. Kusne, E. Rodriguez, J. Paglione, S. Curtarolo, and I. Takeuchi, npj Comput. Mater. **4**, 29 (2018).

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

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

[17] W. L. McMillan and J. M. Rowell, Phys. Rev. Lett. **14**, 108 (1965).

[18] R. Dynes, Solid State Commun. **10**, 615 (1972).

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

[20] M. Schmidt and H. Lipson, Science **324**, 81 (2009).

[21] C. Kim, G. Pilania, and R. Ramprasad, Chem. Mater. **28**, 1304 (2016).

[22] R. Ouyang, S. Curtarolo, E. Ahmetcik, M. Scheffler, and L. M. Ghiringhelli, Phys. Rev. Materials **2**, 083802 (2018).

[23] G. B. Arnold, J. Zasadzinski, J. W. Osmun, and E. L. Wolf, J. Low Temp. Phys. **40**, 225 (1980).

[24] B. Mitrović, H. G. Zarate, and J. P. Carbotte, Phys. Rev. B **29**, 184 (1984).

[25] J. S. Kim, L. Boeri, R. K. Kremer, and F. S. Razavi, Phys. Rev. B **74**, 214513 (2006).

[26] J. Kwo and T. H. Geballe, Phys. Rev. B **23**, 3230 (1981).

[27] S. J. Bending, M. R. Beasley, and E. L. Wolf, Phys. Rev. B **35**, 115 (1987).

[28] K. E. Kihlstrom, D. Mael, and T. H. Geballe, Phys. Rev. B **29**, 150 (1984).

[29] Y. Nishio, M. Shirai, N. Suzuki, and K. Motizuki, Int. J. Mod. Phys. B **7**, 188 (1993).

[30] E. L. Wolf and R. J. Noer, Solid State Commun. **30**, 391 (1979).

[31] K. E. Kihlstrom, R. W. Simon, and S. A. Wolf, Physica B+C **1**, 198 (1985).

[32] A. Junod, Y. Wang, F. Bouquet, and P. Toulemonde, arXiv:cond-mat/0106394 (2001).

[33] E. R. Margine and F. Giustino, Phys. Rev. B **87**, 024505 (2013).

[34] "Symbolic regression utilities," `https://github.com/henniggroup/symbolic-regression-utilities` (2019), accessed: 9/10/2019.

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

[36] D. Duan, Y. Liu, F. Tian, D. Li, X. Huang, Z. Zhao, H. Yu, B. Liu, W. Tian, and T. Cui, Sci. Rep. **4**, 6968 (2014).

[37] M. Komelj and H. Krakauer, Phys. Rev. B **92**, 205125 (2015).