

Superconductor discovery in the emerging paradigm of Materials Informatics

Huan Tran,^{1,*} Hieu-Chi Dam,² Christopher Kuenneth,³ Tuoc N. Vu,⁴ and Hiori Kino⁵

¹*School of Materials Science & Engineering, Georgia Institute of Technology, 771 Ferst Dr. NW, Atlanta, GA 30332, USA*

²*Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Nomi, Ishikawa 923-1292, Japan*

³*Faculty of Engineering Science, University of Bayreuth, Universitätsstr. 30, 95447 Bayreuth, Germany*

⁴*Institute of Engineering Physics, Hanoi University of Science & Technology, 1 Dai Co Viet Rd., Hanoi 10000, Vietnam*

⁵*Research Center for Electronic and Optical Materials,*

National Institute for Materials Science, 1-1 Namiki, Tsukuba, Ibaraki 305-0044, Japan

(Dated: October 31, 2025)

The last two decades have witnessed a tremendous number of computational predictions of hydride-based (phonon-mediated) superconductors, mostly at extremely high pressures, i.e., hundreds of GPa. These discoveries were heavily driven by Migdal-Éliashberg theory (and its first-principles computational implementations) for electron-phonon interactions, the key concept of phonon-mediated superconductivity. Dozens of predictions were experimentally synthesized and characterized, triggering not only enormous excitement in the community but also some debates. In this Article, we review the computational-driven discoveries and the recent developments in the field from various essential aspects, including the theoretical, computational, and, specifically, artificial intelligence (AI)/machine learning (ML) based approaches emerging within the paradigm of materials informatics. While challenges and critical gaps can be found in all of these approaches, AI/ML efforts specifically remain in its infant stage for good reasons. However, opportunities exist when these approaches can be further developed and integrated in concerted efforts, in which AI/ML approaches could play more important roles.

I. INTRODUCTION

Superconductivity was discovered by H. Kamerlingh Onnes in 1911 [1]. When cooling some ordinary substances such as mercury and lead down to a critical temperature T_c near absolute zero, Onnes found that their electrical resistance R completely disappears, signaling that they can carry a current indefinitely without losing any energy. In this state, a superconducting material is a perfect diamagnet, i.e., it completely repels external magnetic fields in the phenomenon known as the Meissner effect [2]. Although T_c discovered for the first superconductors is generally very low, e.g., just a few K [3–5], the discovery of superconductivity triggered not only the development of a new branch of condensed matter physics [6–30] but also significant interest from society [31–33] in the century to come. The ultimate research goal is a material that “superconducts” electricity under ambient conditions, i.e., at the room temperature and above ($T_c \geq 300$ K) and at atmospheric pressure ($P \simeq 0.1$ GPa). If such a material could be discovered, it may unlock many technologies previously confined to science fiction, potentially transforming human civilization as we know it [34, 35]. Before the 21st century, numerous new superconductors were discovered [36–51], pushing the record of T_c to $\simeq 150$ K [45, 46] under (near-)ambient pressure (see Fig. 1). This period was largely propelled by experimental efforts, often performed in combinations with profound physics expertise and intuition.

Concurrently with experimental advancements, a new branch of theoretical condensed matter physics emerged,

dedicating to the understanding of superconductivity, “a manifestation of the quantum world on a macroscopic scale” [22], from the atomic level [6–23, 25–27]. The first solid microscopic theory of superconductivity was developed by Bardeen, Cooper, and Schrieffer (BCS) in 1957 [12–14]. Long story short, lattice phonons can mediate a net attractive interaction between certain pairs of electrons in the neighborhood of the Fermi surface, fusing them into a bound state called Cooper pair [12]. Below T_c , the Cooper pairs can form a robust quantum condensation that can flow without dissipations. Some years later, Migdal-Éliashberg theory was developed [19–24], providing a more complete, truly many-body approach for the simplified model of instantaneous electron-phonon (EP) interactions in the BCS theory. From another end, theoretical endeavors [25–30] have expanded far beyond the phonon-mediated pairing mechanism, which applies to about one third of known (conventional) superconductors [52]. For the remaining two-thirds of (unconventional) superconductors, where this pairing mechanism does not apply, dozens of new mechanisms were suggested, involving possible roles of screened Coulomb interactions [53], polarons [54], anyons [55], Majorana fermions [56], topology [57], spin fluctuations [58], resonating valence-bond states [59], and more. Unconventional superconductors are arranged into numerous classes, e.g., cuprates [38], iron-based [48–51], heavy fermion materials [60, 61], and organic materials [62–65], each of them may be associated with one or more pairing mechanisms. Theoretically, the proposed mechanisms impose no inherent limits on T_c – in fact, the Kohn-Luttinger theorem [53] allows for arbitrarily small T_c . After all, superconductivity remains one of the most challenging and enigmatic topics of physics, with much

* huan.tran@mse.gatech.edu

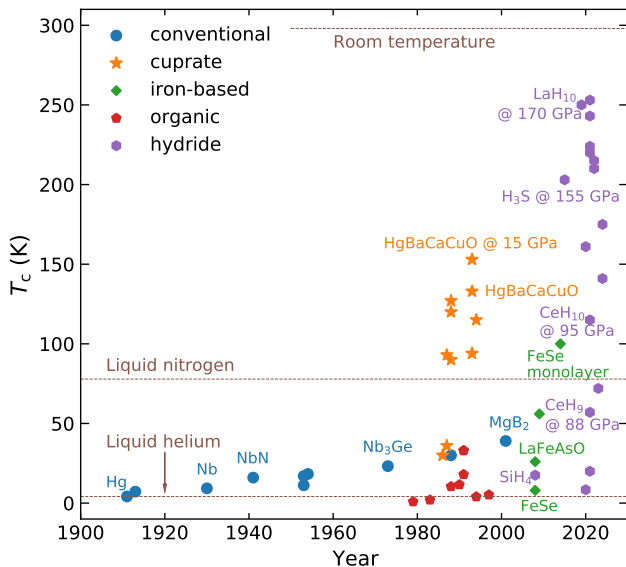


FIG. 1. The time evolution of T_c for some classes of superconductors.

yet to be understood.

Starting from the 2000s, first-principles computational approaches and the required infrastructure have progressed [66–72] to the point where they can provide, within the framework of the Migdal-Éliashberg theory, some valuable guidance for the searches of new phonon-mediated superconductors [73–97]. Two classes of first-principles-based methods that are critical for superconductor discovery are structure prediction methods [66, 67] and density functional perturbation theory (DFPT) [71, 72]. The former is used to explore the configuration space, while the latter is used to approximate the EP interactions. Equipped with these tools, one may start from a hypothetical chemical formula, predicting the lowest-enthalpy atomic structures at specific pressures, examining their thermodynamic and dynamical stability, evaluating the EP interactions, and, finally, estimating their T_c . This generic workflow (shown in Fig. 2 (a)) was used predominantly during the last two decades, leading to thousands of predictions for high- T_c materials [73–92]. Dozens of them, including SiH_4 [98], H_3S [99], LaH_{10} [100, 101], ThH_{10} [102], BaH_{12} [103], YH_6 [104, 105], YH_9 [105], CeH_9 and CeH_{10} [106], CaH_6 [107, 108], and LaBeH_8 [109] were synthesized and tested experimentally (see Table I for a summary), elevating T_c to a record of $\simeq 250$ K, but at the cost of ultra-high pressures of $P \simeq 100 - 200$ GPa. Despite these achievements, critical gaps remain in various stages of the workflow, possibly impacting its predictive power, which is one of the points raised [52] in the on-going debates in the field [52, 110–116].

Fig. 1 provides a snapshot of the T_c evolution during the last century. New discoveries, labeled as “hydride superconductors” in Fig. 1 and summarized in Table I, were synthesized and tested experimentally while some

of them were reproduced independently. The highest- T_c conventional superconductor is MgB_2 with $T_c \simeq 39$ K [47], while unconventional superconductors raise T_c to $\simeq 150$ K [45, 46]. The discoveries of hydride superconductors push the boundary to $\simeq 250$ K at hundreds of GPa. Some points should be noted on them. First, the discoveries are heavily driven by the Migdal-Éliashberg theory-based first-principles computational workflow, suggesting the raising role of this non-experimental method. Second, hydride superconductors are predominant [80], perhaps because the field is strongly inspired by Ashcroft, who, in 2004, predicted [117, 118] that high- T_c superconductivity may be found in hydrogen-dominant alloys, probably at high pressure P . The main rationale is that such materials may feature high vibrational frequencies involving the hydrogen atoms, thus enhancing the electron-phonon interactions. Finally, some extraordinary experimental claims were retracted in the last several years after facing valid concerns from the community [52, 110–116]. We believe that transparency and healthy debates are important for the scientific integrity of all fields, including superconductivity.

The upsurge of computation-inspired superconductor discoveries [73–92] occurs almost simultaneously with the emergence of materials informatics [119–126], a new frontier of materials research. Within this paradigm, artificial intelligence (AI)/machine-learning (ML) based methods are developed to learn past data, ultimately accelerating the understanding, discovering, and designing of new materials. Given their nature, AI/ML approaches may, in principle, complement any physics-based experimental and computational methods that can produce reliable data. Despite considerable efforts [91, 92, 127–165], the development of materials informatics approaches for superconductor discovery remains in an early stage [142, 157]. Looking at other branches of materials informatics where concerted efforts with traditional physics-based methods are blossoming, it is obvious that superconductor discovery may be greatly benefited when the informatics-related developments could gain more momentum.

In this Article, we will review the recent computation-driven discoveries of the hydride superconductors with an eye to the emerging paradigm of materials informatics, which can, in principle, be useful for the future discovery of superconductors in any classes. For this goal, Sec. II is devoted to the essential background of the hydride superconductor discovery, including the Migdal-Éliashberg theory, the first-principles computations of phonon-mediated superconductivity-related properties, and the computational workflow used widely for the discoveries. Sec. II also includes a minimal coverage of some experimental methods needed for the discussions, although the main theme of this Article is non-experimental. Then, in Sec. III, major computation-driven experimental discoveries that were reproduced and confirmed at some levels are reviewed. Next, a simplified introduction of materials informatics is given in Sec. IV,

followed by a review on the materials informatics works that have been completed in this field. Finally, we offer in Sec. V some opinions on the remaining challenges and some forward-looking next steps, specifically in terms of what materials informatics can do to further contribute to the future of superconductor discovery.

II. FUNDAMENTALS OF SUPERCONDUCTOR DISCOVERY

A. Migdal-Éliashberg theory

In the BCS theory [8, 13], the phonon-mediated interaction between certain pairs of electrons residing within an energy cutoff ω_c from the Fermi surface is assumed to be instantaneous and constant, while vanishing beyond ω_c . This simplified picture does not include enough the physics of interacting electron-phonon systems. Migdal-Éliashberg theory is more completed [19–24], taking into account the retarded nature of the EP interactions, considering the first order of vertex corrections [21], while relying on Migdal theorem [20] to treat the damping of the excitations. Central to this theory are the anisotropic non-linear integral Éliashberg equations involving the Matsubara gap $\Delta(\mathbf{k}, i\omega_n)$ and the renormalization factor $Z(\mathbf{k}, i\omega_n)$. On the imaginary axis, they are [166]

$$Z(\mathbf{k}, i\omega_n) = 1 + \frac{\pi T}{N_F \omega_n} \sum_{\mathbf{k}', n'} \frac{\omega_{n'}}{\sqrt{\omega_{n'}^2 + \Delta^2(\mathbf{k}, i\omega_{n'})}} \times \lambda(\mathbf{k}, \mathbf{k}', n - n'), \quad (1)$$

and

$$Z(\mathbf{k}, i\omega_n) \Delta(\mathbf{k}, i\omega_n) = \frac{\pi T}{N_F} \sum_{\mathbf{k}', n'} \frac{\Delta(i\omega_{n'})}{\sqrt{\omega_{n'}^2 + \Delta^2(\mathbf{k}, i\omega_{n'})}} \times [\lambda(\mathbf{k}, \mathbf{k}', n - n') - N_F V(\mathbf{k} - \mathbf{k}')]. \quad (2)$$

Here, N_F is the density of state (DOS) at the Fermi level ϵ_F , $i\omega_n = i\pi T(2n - 1)$ is the n^{th} Matsubara frequency where $n = 0, \pm 1, \pm 2, \dots$, T is the temperature, and $V(\mathbf{k} - \mathbf{k}')$ is the screened Coulomb interaction between the electronic states $\mathbf{k}$ and $\mathbf{k}'$. Furthermore,

$$\lambda(\mathbf{k}, \mathbf{k}', n - n') = 2 \int_0^\infty d\omega \frac{\omega}{\omega^2 + (\omega_n - \omega_{n'})^2} \times \alpha^2 F(\mathbf{k}, \mathbf{k}', \omega) \quad (3)$$

is the EP coupling, determined from the EP spectral function

$$\alpha^2 F(\mathbf{k}, \mathbf{k}', \omega) = N_F \sum_\nu |g_{\mathbf{k}, \mathbf{k}'}^\nu|^2 \delta(\omega - \omega_{\mathbf{k} - \mathbf{k}'}^\nu). \quad (4)$$

In Eq. 4, $g_{\mathbf{k}, \mathbf{k}'}^\nu$ is the EP matrix elements, ν is the polarization index of the phonon with frequency ω , and δ is the delta-Dirac function. When the anisotropy of the Fermi surface is weak, the isotropic spectral function

$$\alpha^2 F(\omega) = \frac{1}{N_F^2} \sum_{\mathbf{k}, \mathbf{k}'} \alpha^2 F(\mathbf{k}, \mathbf{k}', \omega) \delta(\epsilon_{\mathbf{k}}) \delta(\epsilon_{\mathbf{k}'}) \quad (5)$$

can be used to suppress the $\mathbf{k}$ dependence in Eqs. (1), (2), and (3), yielding two isotropic Éliashberg equations and an isotropic EP coupling $\lambda(n - n')$. Within this approximation, the BCS gap equation can be reproduced by setting $\lambda(n - n') = \lambda$ for both $|\omega_n|, |\omega_{n'}| < \omega_c$, and 0 otherwise [18], where

$$\lambda = 2 \int_0^\infty d\omega \frac{\alpha^2 F(\omega)}{\omega} \quad (6)$$

is the dimensionless isotropic EP coupling $\lambda(n - n')$ at $n = n'$ that will be used extensively in the literature.

One way to examine the superconductivity of a material is to solve the Éliashberg equations (1) and (2) self-consistently [166–168] at multiple values of T for a non-trivial solution $\Delta(\mathbf{k}, i\omega_n)$, and the highest T at which $\Delta(\mathbf{k}, i\omega_n) \neq 0$ defines T_c . In this procedure, two main inputs are the spectral function $\alpha^2 F(\mathbf{k}, \mathbf{k}', \omega)$ and a dimensionless Coulomb pseudopotential μ^* , introduced as a treatment for the last term of Eq. (2). It is defined by [166, 167, 169]

$$\mu^* = \frac{N_F \langle\langle V(\mathbf{k} - \mathbf{k}') \rangle\rangle}{1 + N_F \langle\langle V(\mathbf{k} - \mathbf{k}') \rangle\rangle \ln(\epsilon_F / \omega_c)} \quad (7)$$

where $\langle\langle \dots \rangle\rangle$ stands for the double average of the screened Coulomb interaction over $\mathbf{k}$ and $\mathbf{k}'$ on the Fermi surface. With μ^* defined in Eq. 7 and precomputed, the factor of $[\lambda(\mathbf{k}, \mathbf{k}', n - n') - N_F V(\mathbf{k} - \mathbf{k}')] \rightarrow [\lambda(\mathbf{k}, \mathbf{k}', n - n') - \mu^*]$ before Eqs. (1) and (2) are solved numerically [166, 167].

To bypass the cumbersome step of solving the Éliashberg equations (1) and (2), in 1968, McMillan [170] started from some solutions of these equations to develop a direct, empirical formula for T_c , given that the isotropic spectral function $\alpha^2 F(\omega)$ is known. It was then modified by Allen and Dynes [171, 172] in 1975 to be

$$T_c = \frac{\omega_{\log}}{1.2} \exp \left[-\frac{1.04(1 + \lambda)}{\lambda - \mu^*(1 + 0.62\lambda)} \right] \quad (8)$$

where

$$\omega_{\log} = \exp \left[\frac{2}{\lambda} \int_0^\infty d\omega \ln(\omega) \frac{\alpha^2 F(\omega)}{\omega} \right] \quad (9)$$

is the first logarithmic moment of $\alpha^2 F(\omega)$. In the McMillan approach, μ^* is an empirical parameter typically ranging from 0.10 to 0.20 [117, 118]. The empirical formula (8) is good for $\lambda \leq 1.5$ while additional empirical parameters are needed for larger values [172]. Recently, there are some efforts [133, 134] to derive alternative formulae of (8), and they will be reviewed in Sec. IV E. Thanks to its simplicity, the McMillan formula (8) and some related versions [172] were used extremely widely in the literature to estimate T_c , given computed $\alpha^2 F(\omega)$.

The domain of applicability is an essential aspect of a mathematical model. According to Migdal theorem [20], the Migdal-Éliashberg theory is believed [19] to be

valid if the phonon energy scale is much smaller than the electronic energy scale, i.e., $\lambda\omega_c/\epsilon_F \ll 1$, even in the strong coupling regime, i.e., $\lambda \geq 1$. Subsequent examinations [173–178] suggest a more complex picture. Among others, a common conclusion of these works is that the Migdal-Éliashberg theory may be inaccurate when λ exceeds a certain value, which can be 0.4 [173], 1.0 [174, 175], 1.3 [176], or 3.7 [177]. Given that $\lambda \geq 1.0$ in most of the published reports employing this theory to predict T_c [73–92], we believe that extra care are needed for the predictions. Readers, who are interested in the validity and the possible breakdown of the Migdal-Éliashberg theory, are referred to some recent beautiful reviews [21, 22] and original articles [173–177].

B. First-principles computations

Computing the EP matrix element $g'_{\mathbf{k},\mathbf{k}'}$ from first principles, and thus $\alpha^2 F(\mathbf{k}, \mathbf{k}', \omega)$ and $\alpha^2 F(\omega)$, is crucial for using the Migdal-Éliashberg theory in practice.[24] The standard method for such computations is DFPT [71, 72], a perturbation treatment for the response of quantum systems described at the level of density functional theory (DFT) [179, 180]. Notable implementations of DFPT can be found in QUANTUM ESPRESSO [181, 182] and ABINIT [183–185], two major DFT tools for calculating phonon-related properties of solids. Using these packages, thousands of predictions of superconductors have been reported since the 2000s [73–95]. While the calculations of $\alpha^2 F(\mathbf{k}, \mathbf{k}', \omega)$ and $\alpha^2 F(\omega)$ are computationally demanding, some high-throughput efforts [140, 141, 148–150] have made significant strides in this area.

In practice, $\alpha^2 F(\mathbf{k}, \mathbf{k}', \omega)$ and its isotropic version $\alpha^2 F(\omega)$ are computed on two finite-size Γ -centered grids, one for the (electronic) $\mathbf{k}$ points, and one for the (phononic) $\mathbf{q}$ points defined as $\mathbf{q} = \mathbf{k} - \mathbf{k}'$, while the $\mathbf{q}$ grid must be a subgrid of the $\mathbf{k}$ grid. Once $\alpha^2 F(\mathbf{k}, \mathbf{k}', \omega)$ and $\alpha^2 F(\omega)$ are computed, there are two ways to estimate T_c , as discussed in Sec. II A. In the first approach, Eqs. (1) and (2) are solved iteratively, e.g., on the imaginary axis, for a non-trivial solution of $\Delta(\mathbf{k}, i\omega_n)$ using a designated code like Electron-Phonon Wannier (EPW) [166–168]. Then, the real-axis superconducting gap function $\Delta_0(T)$ is approximated from $\Delta(\mathbf{k}, i\omega_n)$, e.g., using Páde continuation [186]. Finally, T_c is determined as the maximum temperature at which the order parameter $\Delta_0(T)$ remains non-zero. The second, and more commonly employed, method uses empirical formulae for T_c from prior solutions to the Éliashberg equations, such as the McMillan formula (8). Having $\alpha^2 F(\omega)$, λ and $\omega_{\log}$ can be computed using some post-processing tools [187] of QUANTUM ESPRESSO suite [181, 182] and ABINIT [183–185] for Eqs. (6) and (9). Unfortunately, both approaches need an empirical value of μ^* for some reasons [167, 181–185], while efforts to derive more rational values of μ^* , e.g., using Eq. (7), are limited [188].

Although the DFT-based calculations may be exact

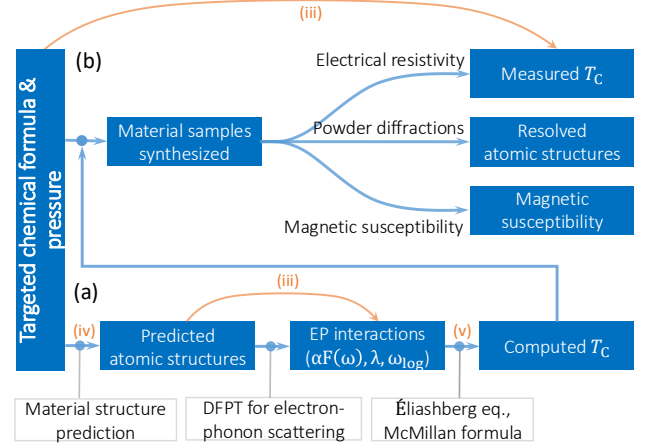


FIG. 2. Typical computational (a) and experimental (b) workflows for superconductor discovery. ML efforts are categorized in 5 groups (see Sec. IV), three of them are (iii) predicting superconductivity-related properties, (iv) accelerate the structure predictions by ML potentials, and (v) deriving new formulae of T_c . Some computational predictions were advanced to experimental synthesis and characterization.

in principle, practical calculations of superconductivity-related quantities require certain finite-size $\mathbf{k}$ - and $\mathbf{q}$ -point grids, pseudopotentials, exchange-correlation functionals, finite-energy cutoffs, finite smearing widths for the δ functions in Eqs. (4) and (5), empirical values of μ^* , whether to include anharmonic effects[104] and/or spin-orbit couplings[189] or not, and more [142, 190, 191]. In many cases, the desired convergence, e.g., wrt the $\mathbf{k}$ - and $\mathbf{q}$ -point grids, may be computationally prohibitive, and some affordable parameters must be assumed. In fact, controlling these factors is challenging in phonon-related calculations, because acceptable numerical errors in energy and force evaluations are often orders of magnitude smaller than those deemed sufficient for standard DFT calculations. Despite the established reliability of DFT in regular electronic-structure problems [192], the above factors present potential sources of numerical error when computing $\alpha^2 F(\mathbf{k}, \mathbf{k}', \omega)$, $\alpha^2 F(\omega)$, and ultimately, T_c . A more detailed discussion on these challenges is presented in Sec. V A.

C. A practical computational discovery workflow

Targeted materials discovery, often referred to as “inverse design”, involves the guided exploration of the vast materials space to find synthesizable materials possessing desired properties [193–198]. In a broad sense, the materials space is truly infinite, containing all possible chemical compositions, external conditions, microscopic details, macroscopic morphologies, additives, and processing parameters needed to describe and fabricate a material. Due to this immense complexity and space, a *trial-and-error* approach is impractical, even with highly

efficient evaluation and decision-making processes. Instead, an effective inverse-design strategy must involve an “intelligent” protocol that rationally directs the exploration toward the targeted properties. Limited versions of this protocol were developed [193–198] actively contributing to the recent discovery of various functional materials [199–206]. One strategy, which is widely used for discovering superconductors [73–95], is depicted in Fig. 2 (a). In this strategy, computational structure prediction methods for exploring the atomic configuration space reached an advanced level of maturity [66–68].

The computational workflow for discovering new superconductors, as shown in Fig. 2 (a), consists of several steps. First, usually guided by chemical intuition and expertise, a targeted chemical formula and external conditions (e.g., a range of pressures) are selected. Subsequently, employing methods like DFT, the atomic structures that are both thermodynamically and dynamically stable are predicted. In this step, a pressure-temperature phase diagram of the targeted chemical formula are usually needed.[85, 207, 208] For this purpose, one may need very expensive computations for the Gibbs free energy of different atomic phases, e.g., using advanced tools like PHONOPY[209] and SSCHA[210]. Next, the superconducting properties of these atomic structures are computed, by either solving the Eliashberg equations or using the McMillan formula, to identify those with respectable predicted T_c . While most of the computational works conclude at this juncture, some proceed to experimental synthesis and testing of these new materials, confirming the predicted superconductivity, as sketched in Fig. 2 (b).

Putting the computations of phonon-mediated superconducting properties aside, atomic structure prediction is a very time consuming step. The goal is to find a thermodynamically stable arrangement of atoms for a given chemical formula at a given pressure, considering its potential decomposition into all possible related materials. This requires, at least, a convex hull analysis [84–89, 208, 211], which necessitates the prediction of the lowest-enthalpy atomic structure, i.e., the global minimum of the multi-dimension potential energy surface (PES), for each related formula. Given the exponential increase in the number of local minima on a PES with the system size [212], the global optimization problem for each formula presents a formidable challenge. Despite its computational costs, an advanced first-principles based materials structure search is highly desired to explore *unknown domains* of the materials space in which *unknown chemistries and atomic structures* are expected, and relying on the known prototype structures is insufficient. In fact, unconstrained structure prediction endeavors have uncovered numerous novel (non-existing) atomic structures [207, 213–215] that were then confirmed experimentally [216–218]. For superconductors discovery, the structure prediction step is critical, especially when the searches extend to the territories no one has ever explored, e.g., at hundreds of GPa. Current *state-of-the-art* structure prediction methods, which were used widely

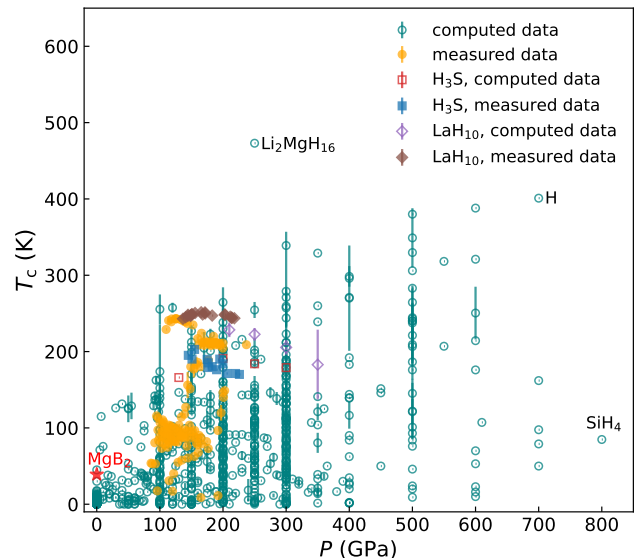


FIG. 3. An *incompleted* snapshot of the phonon-mediated superconductors discovered computationally. For some of them, e.g., H_3S and LaH_{10} , experimental data are available and also shown. Highlighted in the figure, data for MgB_2 at 0 GPa, $\text{Li}_2\text{MgH}_{16}$ at 250 GPa, H at 700 GPa, and SiH_4 at 800 GPa are taken from Refs. 47, 227, 228, and 229, respectively. Part of the data used for this Fig. has been used for Fig. 2c of Ref. 142.

in superconductors discovery, are USPEX [219, 220], CALYPSO [221], *ab initio* random structure search (AIRSS) [222], XtalOpt [223], MAISE [224], and minima-hopping [225, 226].

D. Experimental synthesis and characterizations

Some computationally discovered hydride superconductors were synthesized and tested experimentally [98–109]. For those advanced to this step, the targeted formulations are typically synthesized by reacting pure metals with excess hydrogen or hydrogen-rich gases, e.g., ammonia borane and hydrocarbons, often using laser heating in diamond anvil cells (DAC) under the desired pressure.

Having the samples, the *electrical resistance* R was always examined, typically using the four-point probe technique [250, 251]. By definition, the presence of a superconducting transition is implied when R drops sharply to zero at the critical temperature T_c . The isotope effect [252, 253], a footprint of the lattice dynamics on the phonon-mediated superconductivity can then be observed in the dependence of T_c on the average isotope mass of the material. *Magnetic susceptibility measurements* are more challenging but critical [254, 255] because they can help probing the Meissner effect [2], a hallmark of superconductivity, while providing other essential information such as the critical field, penetration depth, and critical current density. Four methods that may be used are alternating-current (AC) sus-

TABLE I. Notable hydride superconductors, given in terms of chemical formula, that were discovered computationally and then confirmed experimentally from the 2000s. For each of them, maximum observed critical temperature $T_c^{\max}$, the pressure P @ $T_c^{\max}$ at which $T_c^{\max}$ was observed, years and references of (experimental) discoveries and (computational) predictions, and a short summary of the discoveries, are provided.

Formula	$T_c^{\max}$ (K)	P @ $T_c^{\max}$ (GPa)	Exper. discovery	Comput. prediction	Short summary
SiH ₄	17.5	90	2008 [98]	2006, [230, 231] 2007 [232]	Electrical resistance measured, $P6_3$ structure resolved, not matched with predictions, multiple structure searches follow, refining the structure
H ₃ S	203	155	2015 [99]	2014 [82]	Electrical resistance & magnetic susceptibility measured, $Im\bar{3}m$ structure resolved & matched with predictions, superconductivity replicated [233–237]
LaH ₁₀	250	170	2019 [100, 101] 2020 [102]	2017 [87, 88]	Electrical resistance measured, $Fm\bar{3}m$ structure resolved & matched with predictions, superconductivity replicated [238]
ThH ₉	146	170	2020 [102]	N/A	Electrical resistance measured, $P6_3/mmc$ structure resolved
ThH ₁₀	161	174	2020 [102]	2018 [89]	Electrical resistance measured, $Fm\bar{3}m$ structure resolved & matched with predictions
PrH ₉	8.4	120	2020 [239]	2020 [239]	Electrical resistance measured, $P6_3/mmc$ structure resolved & matched with predictions
BaH ₁₂	20	140	2021 [103]	2021 [103]	Electrical resistance measured, $Cmc2_1$ structure resolved & matched with predictions
(La,Y)H ₁₀	253	199	2021 [211]	2021 [211]	Electrical resistance measured, $Fm\bar{3}m$ structure resolved & matched with predictions
YH ₆	224	168	2021 [104]	2015 [86, 88]	Electrical resistance measured, $Im\bar{3}m$ structure resolved
YH ₉	220	183	2021 [105]	2015 [86, 88]	& matched with predictions, superconductivity replicated [105]
	243	201	2021 [105]	2017 [88]	Electrical resistance measured, $P6_3/mmc$ structure resolved & matched with predictions, superconductivity replicated [240]
CeH ₉	57	88	2021 [106]	2017 [88]	Electrical resistance measured, $P6_3/mmc$ structure resolved & matched with predictions, T_c replicated & diamagnetism of Meissner effect probed [241]
CeH ₁₀	115	95	2021 [106]	2017 [88]	Electrical resistance measured, $Fm\bar{3}m$ structure resolved & matched with predictions
ScH ₃	18.5	131	2021 [242]	2010, [243] 2016 [244]	Electrical resistance measured, $Fm\bar{3}m$ structure resolved & matched with predictions
LuH ₃	12.4	122	2021 [242]	N/A	Electrical resistance measured, $Fm\bar{3}m$ structure resolved
CaH ₆	215/210	172/160	2022 [107, 108]	2012 [81, 148]	Electrical resistance measured, $Im\bar{3}m$ structure resolved & matched with predictions, superconductivity replicated [108]
SnH ₄	72	180	2023 [245]	2023 [245]	Electrical resistance measured, $Fm\bar{3}m$ structure resolved & matched with predictions
LaBeH ₈	110	80	2023 [109]		Electrical resistance measured, $Fm\bar{3}m$ structure resolved
NbH ₃	42	187	2024 [246]	N/A	Electrical resistance measured, $Fm\bar{3}m$ structure resolved
La _{0.5} Ce _{0.5} H ₁₀	175	155	2024 [247]	2024 [247]	Electrical resistance measured, $Fm\bar{3}m$ structure resolved
Y _{0.5} Ce _{0.5} H ₉	141	130	2024 [248]	N/A	Electrical resistance measured, $P6_3/mmc$ structure resolved
La ₄ H ₂₃	90	95	2024 [249]	NA	Electrical resistance measured, $Pm\bar{3}m$ structure resolved

ceptibility measurements using a pickup/compensating coil architecture [234], superconducting quantum interference device (SQUID) magnetometer [99], synchrotron Mössbauer technique [256], and a quantum sensing approach involving Nitrogen-vacancy centers implanted in DAC [241, 257]. As operating these techniques under hundreds of GPa in a DAC is non-trivial, only a few magnetic susceptibility measurements were reported [234, 238]. Instead, the most widely-used approach is to measure the dependence of T_c on the external magnetic field, the behavior that could reveal the upper critical field, ultimately linking to the coherence length of the superconducting state [7, 258]. A new method [259], measuring the magnetic flux trapped inside the superconducting samples, i.e., the incomplete Meissner effect, seems to be useful. Finally, *resolving the atomic structure* of the synthesized samples is desirable. For this goal,

x-ray diffraction (XRD) pattern and sometimes Raman scattering are powerful methods. Overall, tremendous challenges remain for the experimental techniques in superconductor discovery, specifically magnetic susceptibility measurements at extremely high pressure in a DAC, and readers are referred to Refs. 241, 257, 259 for more information.

Fig. 2 (b) summarizes the aforementioned measurements, among others, desired and used for probing the possible superconductivity. All the discovered hydride superconductors discussed in Sec. III have been experimentally synthesized and at least one of the three characteristics are examined.

III. COMPUTATION-DRIVEN SUPERCONDUCTOR DISCOVERIES

We now turn to the discoveries of hydride superconductors in the last two decades, mostly driven by the first-principles computational methods and workflow. It is interesting to note that Ashcroft's prediction [117, 118] on the possible superconductivity in hydrides came more than three decades after his similar prediction on compressed hydrogen [260]. One way to view the difference between the two predictions is that hydrogen atoms in solid materials are “chemically precompressed” [230], thus realizing the prediction in these materials could be technically more feasible. This may be a reason why the later [117, 118] was soon followed by thousands of computational and dozens of experimental discoveries.

Table I and Fig. 3 provide a (likely incompleting) summary of the computational and experimental discoveries reported from the 2000s. Not surprisingly, most of the materials are hydrogen-rich, predicted to be superconductors at very high pressures, ranging up to $\simeq 800$ GPa. One of the most striking computational predictions is the superconductivity of $\text{Li}_2\text{MgH}_{16}$ at 250 GPa with $T_c \simeq 473\text{K}$, or 200°C [227], while the first major experimental discovery is the superconductivity of H_3S at 155 GPa with $T_c = 203\text{ K}$ [99]. This discovery is remarkable because it reproduced a majority of the computational predictions reported one year earlier [82].

The remaining part of this Section is devoted to a detailed discussion on some superconducting materials that were predicted computationally and then synthesized and tested. Given the scope and the limited length of this Article, and an extensive list of experimental discoveries shown in Table I, we will examine a few of them that were reproduced independently and/or sparked significant interest and follow-up works from the community. For other discoveries, readers are referred to Table I in which a short summary and relevant references are given for each of them.

A. Silane SiH_4

Silane SiH_4 is the first hydride superconductor computationally predicted in 2006 [230–232] and experimentally reported in 2008 [98], shortly after Ashcroft's later prediction [117, 118]. By examining multiple atomic structures of SiH_4 and computing the frequency cutoff ω_c at a first-principles level, the orthorhombic $Pman$ structure stands out with the predicted $T_c \simeq 1.13\hbar\omega_c/k_B \exp(-1/N_F V) \simeq 166\text{ K}$ at 202 GPa [230]. In the same year, after a search for high-pressure structures of SiH_4 , a qualitative assessment into the possible superconductivity, specifically involving a monoclinic $C2/c$ phase, was conducted with a preferable conclusion at pressures up to 50 GPa [231]. From 50 GPa to 250 GPa, an insulating tetragonal $I4_1/a$ phase was predicted [231] to be thermodynamically stable and then confirmed

experimentally [261, 262]. One year later, the computational workflow based on Migdal-Éliashberg theory was used [232] for another monoclinic $C2/c$ structure, visualized in Fig. 4 (a), predicting a $T_c \simeq 45 - 55\text{ K}$ at 90 and 125 GPa. The spectral function $\alpha^2F(\omega)$ and the accumulated $\lambda(\omega)$ of this phase are shown in Fig. 4 (b).

The computational predictions were followed by an experimental discovery [98], in which SiH_4 samples, shown in Fig. 4 (c), were fabricated. Electrical resistance measurements, one of them is given in Fig. 4 (b), clearly show the superconducting-like drop at a pressure-dependent critical temperature T_c . Between 65 GPa and 90 GPa, T_c grows from 7K to 17.5 K before going down and up again in the regime of 120 – 200 GPa, as shown in Fig. 4 (e). This behavior suggests the involvement of more than one phases of SiH_4 . Examining Raman scattering and XRD pattern, a superconducting hexagonal $P6_3$ phase was resolved in the first regime and the predicted $I4_1/a$ phase [231] was confirmed again in the second regime, in which both phases co-exist.

While the superconducting hexagonal $P6_3$ phase was not predicted computationally, the discovery of the predicted superconductivity in SiH_4 and the verification of the predicted $I4_1/a$ phase are intriguing. In fact, they motivated multiple computational structure searches for SiH_4 [229, 263, 264] and disilane, a related compound with formula Si_2H_6 [265, 266]. As the $P6_3$ phase was found [263] to be dynamically unstable, a related orthorhombic $Cmca$ phase was suggested as a candidate for the superconducting phase of SiH_4 . Among the structures subsequently found [264] for SiH_4 , an orthorhombic $Pbnc$ phase is not only thermodynamically more stable than the $Cmca$ phase but is also related to the $P6_3$ phase through its dynamical instability. Furthermore, computed λ and $\omega_{\log}$ of the $Pbnc$ phase lead to $T_c \simeq 16.5\text{ K}$ by using the McMillan formula [264]. To close this section, we refer readers to Ref. 262 for a summary of such SiH_4 -related efforts.

B. Sulfur hydride H_3S

In 2014, a structure prediction effort [82] was performed to search for possible stable atomic structures of $(\text{H}_2\text{S})_2\text{H}_2$ at elevated pressure P up to 300 GPa. This work was partially motivated by an early experimental report [233], showing that H_2S and H_2 , both are gas at ambient conditions, can form stoichiometric compound H_3S near 3.5 GPa. The computational examination [82] reveals that when P increases, H_3S compound undergoes a series of structural phase transitions from $P1$ to $Cccm$ at 37 GPa, then to $R3m$ at 111 GPa, and finally to $Im\bar{3}m$ at 180 GPa. By examining the simulated XRD pattern at 22 GPa, the stable $P1$ phase of H_3S was found [82] to match with that experimentally reported [233]. As the $R3m$ and $Im\bar{3}m$ phases are metallic at the pressures they are stable, superconducting-related calculations suggest that at 130 GPa, the $R3m$ phase has

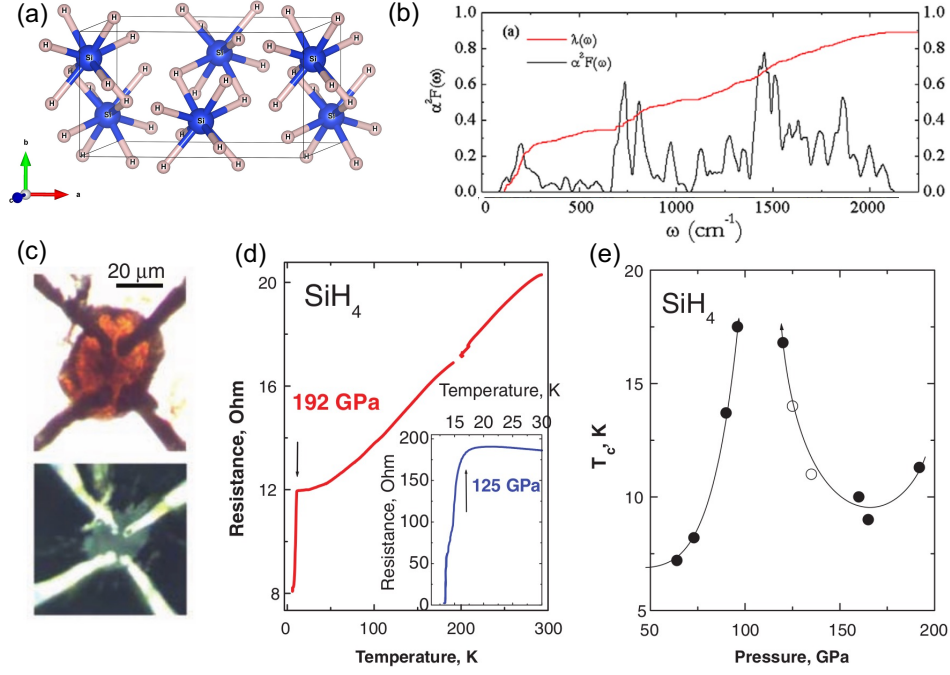


FIG. 4. (a) The predicted $C2/c$ structure of silane SiH_4 232, (b) the spectral function $\alpha^2 F(\omega)$ and accumulated $\lambda(\omega)$ of the $C2/c$ structure, (c) the fabricated sample placed within the four-point probe, (d) a representative superconducting step in measured resistance, and (e) pressure-dependent T_c measured for silane. Panels (b) and (c, d, e) were reprinted with permission from Ref. 232 (Copyright 2007, IOP Publishing) and Ref. 98 (Copyright 2008, The American Association for the Advancement of Science), respectively.

$\lambda = 2.07$, $\omega_{\text{log}} = 1125.1$ K, which can be translated to $T_c \simeq 155 - 166$ K using Eq. 8 with μ^* ranging from 0.10 to 0.13. Likewise, at 200 GPa, the $Im\bar{3}m$ phase has $\lambda = 2.19$, $\omega_{\text{log}} = 1334.6$ K, and finally $T_c \simeq 191 - 204$ K. The phonon band structure, the spectral function $\alpha^2 F(\omega)$, and the accumulated $\lambda(\omega)$ of the $Im\bar{3}m$ phase of H_3S at 200 GPa are reproduced in Fig. 5 (a).

In 2015, the superconductivity with a record-breaking $T_c = 203$ K at 155 GPa was observed on synthesized samples of H_3S (see Fig. 5 (c)) and reported [99]. Fig. 5 (d) shows the sharp drops of the electrical resistance to zero when T is lowered to a critical value that depends on the pressure. A strong isotope effect was also observed [99] and shown in Fig. 5(e), implying a link to the lattice phonons. Magnetic susceptibility measurements were performed using SQUID [99], revealing a transition from the diamagnetic to the paramagnetic state at $\simeq 203.5$ K, as shown in Fig. 5(f).

The superconductivity of superconductivity in H_3S was reproduced by some other groups [234–237, 256] In one effort [234], the superconductivity of H_3S was probed in a range of P using AC magnetic susceptibility measurements. As P increases, the superconductivity appears at 117 GPa with $T_c = 38$ K. The critical temperature then rises to 183 K at 149 GPa before lowering to 140 K at 171 GPa. This non-monotonic behavior may come from the stoichiometry change and the effects of P on λ .

At first, there is possibility of the H_2S compound to

involve in the observation [99], but the predicted $T_c \simeq 80$ K of H_2S at this pressure range [267] is far from the measured data. Although the reported pressure (155 GPa) is in the regime where the $R3m$ of H_3S was predicted [82], the $Im\bar{3}m$ phase was anticipated [99] and then experimentally confirmed [190, 235–237]. The values of T_c that were computed [82] and measured [99] for H_3S at different pressure P are summarized in Fig. 3, indicating an overall good agreement.

The discovery [99] and confirmation [234–237] of the superconductivity in H_3S with strikingly high T_c came soon after the prediction [82]. Moreover, the predicted superconducting $Im\bar{3}m$ phase was confirmed while the predicted T_c agrees well with the measured values. These factors make the superconductivity of H_3S a major discovery, sparking significant excitement not only in science community but also in media and society [268].

C. Lanthanum hydride LaH_{10}

The discovery of the superconductivity in lanthanum hydride LaH_{10} was also initiated by computational structure prediction endeavors. In 2017, dozens of clathrate-like structures of hydrides were discovered computationally [87, 88] at high pressures with very high predicted T_c . In these prototype structures, rare-earth elements like La and Y are located at the center of hydrogen cavities, linked together throughout the space (see Figs. 6(a)

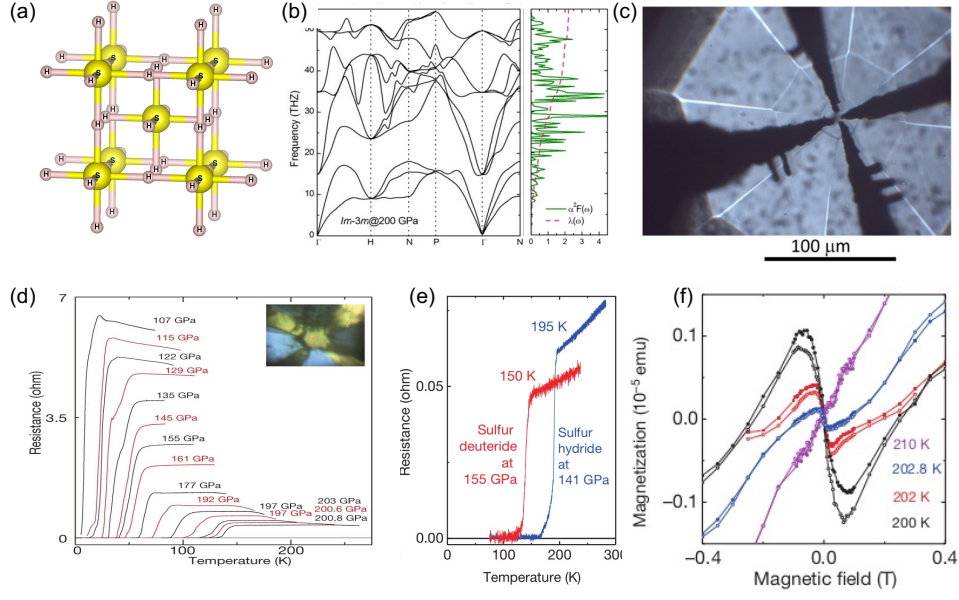


FIG. 5. (a) The predicted $Im\bar{3}m$ structure of H_3S , (b) phonon band structure, the spectral function $\alpha^2 F(\omega)$, and accumulated $\lambda(\omega)$ of the $Im\bar{3}m$ structure at 200 GPa, (c) four Ti electrodes sputtered on a diamond anvil, i.e., a four-point probe, at the center of them is the sample, (d) R - T dependence measured for H_3S at different pressures, (e) R - T dependence measured for H_3S and H_3D , which reveals the isotope effect, and (f) the magnetization measured as functions of external field, showing the diamagnet and paramagnet characteristics below and above $T_c \simeq 203$ K, respectively. Panels (b) and (c, d, e, f) were reprinted with permission from Ref. 82 (under a Creative Commons license) and Ref. 99 (Copyright 2015, Springer Nature), respectively.

and 7 (d) for visualizations). Among the examined rare-earth hydrides, LaH_{10} in its cubic $Fm\bar{3}m$ clathrate-like structure was predicted to have an estimated $T_c \simeq 288$ K at 200 GPa [88] and $T_c \simeq 274 - 286$ K at 210 GPa [87]. Quickly, LaH_{10} was synthesized [269] in the predicted $Fm\bar{3}m$ clathrate structure at 170 GPa, the pressure it was predicted computationally to be dynamically unstable [87]. The small discrepancy between computations and experiments was then attributed to lattice vibrations [270], neglected in early computations [87]. Nevertheless, the experimental synthesis of LaH_{10} in the predicted structure is a success.

The superconductivity of LaH_{10} was observed in 2019 [100, 101]. Superconducting-like drops in the measured electrical resistance R of the synthesized samples were reported to be at $T_c \simeq 260$ K under 180 – 200 GPa [100] and $T_c \simeq 250$ K under 170 GPa [101], as shown in Fig. 6 (c). In Ref. 101, the cubic $Fm\bar{3}m$ structure of LaH_{10} was confirmed while other characteristics of superconductivity, including the isotope effect, shown in the inset of Fig. 6 (c), and the reduce of T_c under increasing magnetic field, shown in Fig. 6 (d), were also observed. By obtaining the upper critical field as a function of T and fitting the data into the Ginzburg-Landau model [7], a coherence length of about 1.56 – 1.86 nm was extracted [101]. However, magnetization measurements from the DAC cannot be performed using SQUID because of the small volume and size (10 – 20 μm) of the samples.

One year later, the synthesis of LaH_{10} was repeated [238]. The samples are not larger, but using the

pickup/compensating coil technique, weak but measurable signals from magnetic susceptibility measurements, as shown in Fig. 6(e), were obtained [238]. After removing the background, the final data, shown in Fig. 6(f), point to superconducting transitions at $T_c \simeq 250 - 280$ K and $P \simeq 170 - 180$ GPa [238]. In 2021, the cubic $Fm\bar{3}m$ phase was confirmed again while some other key superconducting characteristics such as the upper critical fields and coherence lengths was established [271].

The clathrate structures realized in LaH_{10} , in which the hydrogen cavities are interconnected and distributed continuously throughout the space (see Figs. 6(a)), allows for unusually high hydrogen content. For that reason, it can be viewed a close realization of metallic hydrogen, for which high- T_c values were predicted in 1968, also by Ashcroft [260]. Remarkably, this host-atoms-in-hollows motif has been predicted computationally [75, 81, 86, 88] and then realized experimentally [104, 105] for many superconducting hydrides such as YH_6 , YH_9 , CeH_9 and CeH_{10} , as discussed later. Some computational works follow this promising lead, suggesting even hydrogen-rich clathrate structures of superhydrides, e.g., LaH_{16} [272], LaH_{18} , YH_{18} , AcH_{18} , CeH_{18} , and ThH_{18} [228, 273] with predicted $T_c \gtrsim 200$ K.

D. Yttrium hydrides YH_6 and YH_9

Yttrium hydrides YH_6 and YH_9 are in the list of ten high- T_c superconducting rare-earth hydrides, including

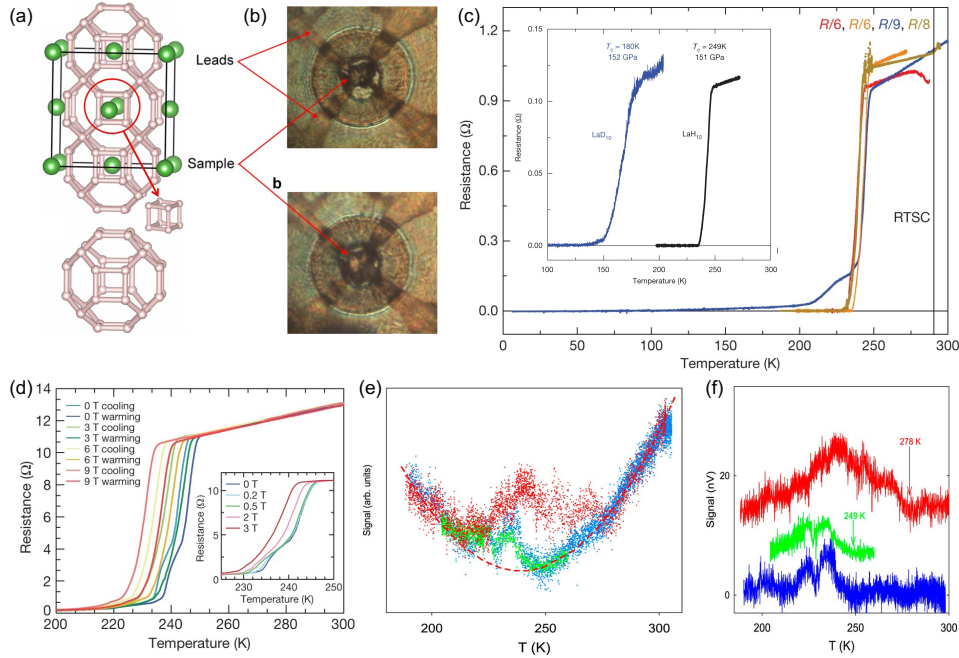


FIG. 6. (a) The $Fm\bar{3}m$ clathrate-like structure predicted for LaH_{10} , (b) LaH_{10} samples fabricated and placed between four probes, (c) R - T dependence measured for LaH_{10} samples at zero external magnetic field, (d) R - T dependence measured for LaH_{10} samples at varying external magnetic field up to 9 T, (e) magnetic response signals from LaH_{10} samples in a DAC, whose the background is shown in dashed line, and (f) magnetic response after removing the background signal. The inset of panel (c) show the R - T curves measured for LaH_{10} and LaD_{10} , which reveal the isotope effect. Panels (a), (b, c, d), and (e, f) were reprinted with permission from Ref. 87 (Copyright 2017, National Academy of Sciences of the United States of America), Ref. 101 (Copyright 2019, Springer Nature), and Ref. 238 (under the Creative Commons CC BY license), respectively.

LaH_{10} , computationally predicted in 2017 [88]. The common feature of this series is their clathrate-like structure, which, as mentioned above, is interesting in the context of high- T_c superconductivity [75]. The predicted T_c of YH_6 (in its cubic $Im\bar{3}m$ phase) and YH_9 (in its hexagonal $P6_3/mmc$ phase) are $T_c \simeq 250$ K at 120 GPa and $T_c \simeq 260$ K at 150 GPa, respectively [88]. In fact, the cubic $Im\bar{3}m$ phase of YH_6 was also predicted to be a superconductor with $T_c \simeq 251 - 264$ K at 164 GPa in 2015 in a structure prediction campaign [86].

The experimentally observed superconductivity of YH_6 and YH_9 was reported in Refs. 104 and 105, both in 2021. First, YH_6 was synthesized and observed in the predicted $Im\bar{3}m$ phase [104]. The measured electrical resistance R , shown in Fig. 7 (a), has a superconducting-like drop at $T_c \simeq 224$ K at 168 GPa. The isotope effect was also observed, suggesting a link between the superconductivity and the lattice phonons. This discovery was reproduced in Ref. 105 where a $T_c \simeq 220$ K at 183 GPa was reported for the $Im\bar{3}m$ phase of YH_6 , as shown in Figs. 7 (b), 7 (c), and 7 (e). The superconductivity of YH_9 in its predicted hexagonal $P6_3/mmc$ phase (see Fig. 7 (d)) was also reported in Ref. 105, exhibiting a pressure-dependent T_c peaking $\simeq 243$ K at 201 GPa (see Fig. 7 (e)). The dependence of T_c on the external magnetic field is shown in Fig. 7 (f) and Fig. 7 (g), pointing to an upper critical field and a coherence

length. Fitting the measured T -dependent upper critical fields to the Ginzburg-Landau [7] and Werthamer-Helfand-Hohenberg [258] models, the coherence length at 0 K of YH_6 and YH_9 is 1.45 – 1.75 nm and 2.3 – 2.7 nm, respectively.

The superconducting-like drops in the electrical resistance R , the dependence of T_c on external magnetic field, and the atomic structure of both YH_6 and YH_9 were reproduced recently [240]. Nevertheless, magnetic susceptibility measurements have not been reported thus far, perhaps because of known technical challenges. In this context, emerging techniques to visualize local domain of diamagnetism [241] or measuring the trapped flux [259] could be useful, as demonstrated [241] for CeH_9 and mentioned in Sec. III E.

E. Cerium hydrides CeH_9 and CeH_{10}

Similar to LaH_{10} , YH_6 , and YH_9 , CeH_9 and CeH_{10} are also in the series of rare-earth hydrides predicted in Ref. 88. The predicted structures of CeH_9 and CeH_{10} , visualized in Fig. 8(a), are both clathrate-like, which belong to the $P6_3/mmc$ and $Fm\bar{3}m$ space groups, respectively. Although the predicted T_c of CeH_9 and CeH_{10} is lower than others, i.e., $\simeq 50 - 60$ K, the pressure at which the superconductivity was predicted is relatively

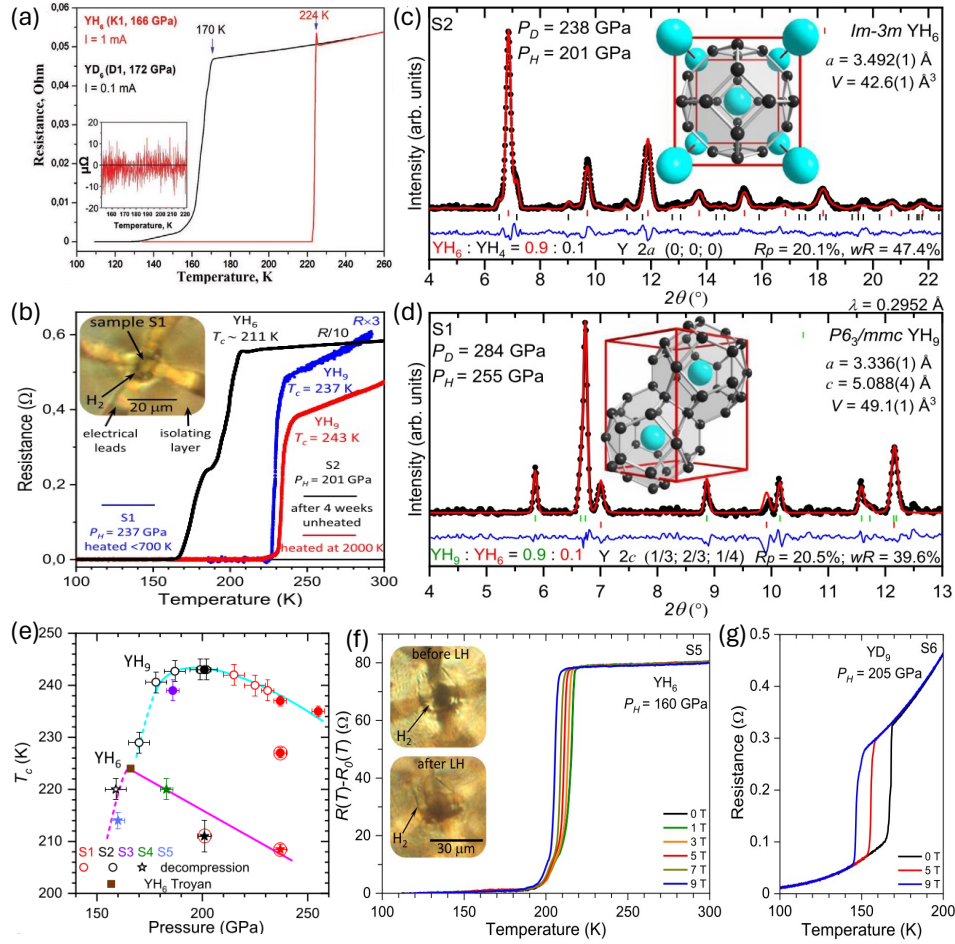


FIG. 7. (a) R - T dependence measured for YH₆ (166 GPa) and YD₆ (172 GPa), (b) R - T dependence measured for YH₆ (237 GPa) and YH₉ (201 GPa), (c) XRD-resolved $Im\bar{3}m$ clathrate-like structure of YH₆, (d) XRD-resolved $P6_3/mmc$ clathrate-like structure of YH₉, (e) P -dependent T_c measured for YH₆ and YH₉, and R - T dependence measured for YH₆ (f) and YH₉ (g) at varying external magnetic field. Panels (a) and (b, c, d, e, f, g) were reprinted with permission from Ref. 104 (Copyright 2021, John Wiley and Sons) and Ref. 105 (under the Creative Commons CC BY license), respectively.

lower, i.e., 100 GPa for CeH₉ and 200 GPa for CeH₁₀ [88]. Given that the superconductivity of H₃S, LaH₁₀, YH₆, and YH₉ was reported at close to 200 GPa, this prediction stimulates some interest. Amid the search, CeH₉ was the synthesized [274], confirming the predicted $P6_3/mmc$ phase. A structure prediction campaign was then launched [274], recovering the predicted $P6_3/mmc$ and $Fm\bar{3}m$ phases of CeH₉ and CeH₁₀, respectively. Calculations performed return $\lambda = 2.3$ and $\omega_{\log} = 740$ K for the $P6_3/mmc$ phase of CeH₉ at 200 GPa, which were translated to $T_c = 105 - 117$ K using the McMillan formula with $\mu^* = 0.10 - 0.13$ [274].

In 2021, the superconductivity of CeH₉ and CeH₁₀ was reported at 88 GPa and 95 GPa, respectively [106]. The critical temperature, extracted from measured electrical resistance R (shown in Fig. 8(b)), is 57 K for CeH₉ and 115 K for CeH₁₀ [106]. The $P6_3/mmc$ phase of CeH₉ is predominant in the synthesized samples while the CeH₁₀ in its $Fm\bar{3}m$ phase plays a minor role, as shown in Fig. 8(c). Using the Ginzburg-Landau [7] and Werthamer-

Helfand-Hohenberg [258] models for the magnetic field-dependent T_c measured data (see Fig. 8(d)), the upper critical field at 0 K of CeH₉ was estimated to be 17.7 – 22.9 T at 140 GPa, while the coherence length is 3.4 nm at 120 GPa, 4.3 nm at 150 GPa, and 5.1 nm at 150 GPa [106]. An isotope effect was observed for CeH₉ [106], but no magnetic susceptibility measurements were reported.

Recently, local diamagnetic domains of CeH₉ were observed [241] using a new technique that is capable of performing local magnetometry with sub-micron spatial resolution inside a DAC [241]. These domains, shown in bright colors in Fig. 8(e), are about 10 μm in size and small for SQUID. This result is intriguing, suggesting that the new technique would be useful for probing the superconductivity in small samples, the scenario that is common in superconductor discovery [101].

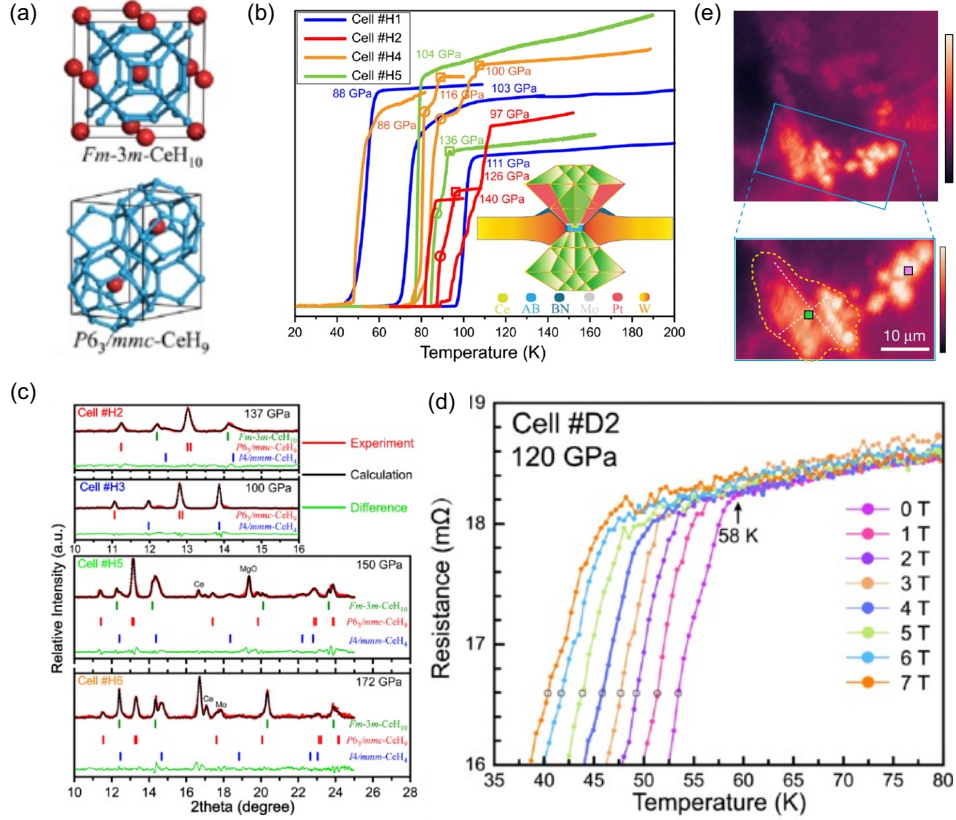


FIG. 8. (a) Predicted $Fm\bar{3}m$ structure of CeH_{10} and $P6_3/mmc$ structure of CeH_9 , (b) R - T dependence measured for CeH_9 and CeH_{10} , (c) XRD-resolved $Fm\bar{3}m$ structure of CeH_{10} and $P6_3/mmc$ structure of CeH_9 , (d) R - T dependence measured for CeH_9 and CeH_{10} at varying external magnetic field, and (e) confocal fluorescence image of CeH_9 in which the diamagnet domains are shown in bright color. Panels (a, b, c, d) and (e) were reprinted with permission from Ref. 106 (Copyright 2021, American Physical Society) and Ref. 241 (Copyright 2024, Springer Nature), respectively.

IV. MATERIALS INFORMATICS IN SUPERCONDUCTOR DISCOVERY

Emerging in the early 2010s and partly propelled by the Materials Genome Initiative [275], materials informatics has rapidly developed into a widely-used tool in materials research [119–126]. This approach employs AI/ML techniques to learn materials data, creating models that can generate rapid predictions and complement traditional approaches in accelerating materials discovery. Synergetic approaches involving materials informatics methods, simulations, and physical experiments have proactively driven numerous recent materials discoveries. Many of these newly discovered materials have been synthesized and tested. Notable examples include battery materials [199], green energy materials [200], functional and sustainable polymers [201], alloys [202, 203], and more.

Data is the cornerstone of materials informatics. In this field, data can originate from different sources such as experimental measurements and/or physics-based computations, be represented in different formats, contain different amounts and details of information, and

possess different levels of fidelity. An ideal dataset for ML model training should be sufficiently big in size, diverse in the chemical, configuration, and parameter spaces, and complete in the relevant information. The development of major DFT-based materials databases like Materials Project [276], OQMD [277], AFLOWLIB [278], and NOMAD [122, 123] is important for future the progress of materials informatics.

For problems related to superconductivity, a typical model development procedure, sketched in Fig. 9, starts with collecting and preparing training data for related materials. This data includes both the structural infor-

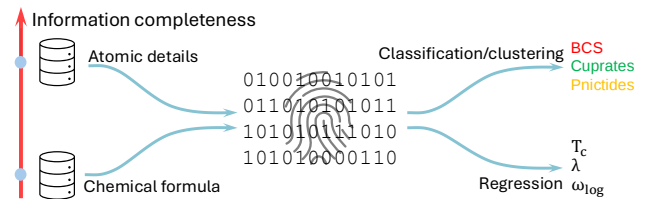


FIG. 9. A widely adopted schematic pipeline for ML predictive models of superconductivity.

mation of the materials and the target labels or properties of interest. In a next step, the materials, represented either by their chemical formulae or atomic structures, are converted into numerical “hand-crafted” features that can be readily interpreted and utilized by ML algorithms [119, 125, 126]. This step is required from the early days, but recent advances in deep-learning techniques [138, 139, 279–285] signal that the materials features may also be “learned”. Subsequently, a suitable ML algorithm learns the features, establishing a clustering, classification, or regression model capable of rapidly predicting or evaluating new materials.

As the model input, a chemical formula inherently provides less detailed information than a fully specified atomic structure. On the spectrum of information completeness, as previously discussed [286] and sketched in Fig. 9, the information content of a chemical formula is lower than that of an atomic structure. From the physics standpoint, each formula corresponds to an infinite number of metastable atomic structures (local minima on the PES) [212]. Many of these structures are energetically close to the global minimum, implying that they can be stabilized in experiments or appear in computational models. Importantly, these low-energy structures have completely different atomic arrangements and properties, with one being insulating while another is conducting [287]. Superconductivity, as discussed in Sec. II A, is highly sensitive to the atomic arrangements. Thus, describing a superconductor only by its chemical formula entails the risk of inaccurate results. Eventually, models relying only on formulae as inputs inevitably encounter a degree of irreducible (aleatoric) uncertainty in their predictions that cannot be simply reduced by increasing the amount of data [286, 288]. As early superconductor databases provide only chemical formulae, recent efforts to bring atomic details in, as discussed in Sec. IV A, have become a focal point of research.

Extending some previous discussions [142, 157], materials informatics efforts in superconductor discovery can be categorized into five groups. The first group (i) includes those fully devoted to [160–165] or partially involved in [136, 137, 140–142] data generation, extraction, and dissemination. The other groups are (ii) classifying superconductors [135, 137, 159], (iii) predicting superconducting-related properties [136, 137, 139–148], (iv) accelerating the structure prediction step with ML potentials [127–129], and (v) refining empirical formulae for T_c [133, 134, 148]. Among these groups, (iii), (iv), and (v) are sketched in Fig. 2. Due to the integrated nature of the materials informatics approaches, the availability of the data, and specifically the forward-looking aspiration of the field, subsequent discussions in this Section will involve all classes of superconductors, including those mediated by phonons. The current status, challenges, and critical next steps of the materials informatics efforts in superconductor discovery will be discussed in Sec. V.

A. Data generation, extraction, and dissemination

The Superconducting Material Database maintained by Japan’s National Institute for Materials Science (NIMS) provides data for many efforts in the field [135–137, 159]. This database, referred to as NIMS SuperCon, records the chemical formula of about 32,000 known conventional (phonon-mediated) and unconventional superconductors, not all of them have (measured) T_c included. Recently, NIMS SuperCon was cleaned, re-edited, and released as MDR SuperCon, containing 26,323 records. Parallel cleaning and curation efforts [136, 137] resulted in another version, presently known as SuperCon, containing about 16,400 records, among them 4,000 records have no T_c . Details on NIMS SuperCon, MDR SuperCon, and SuperCon are given in Table II.

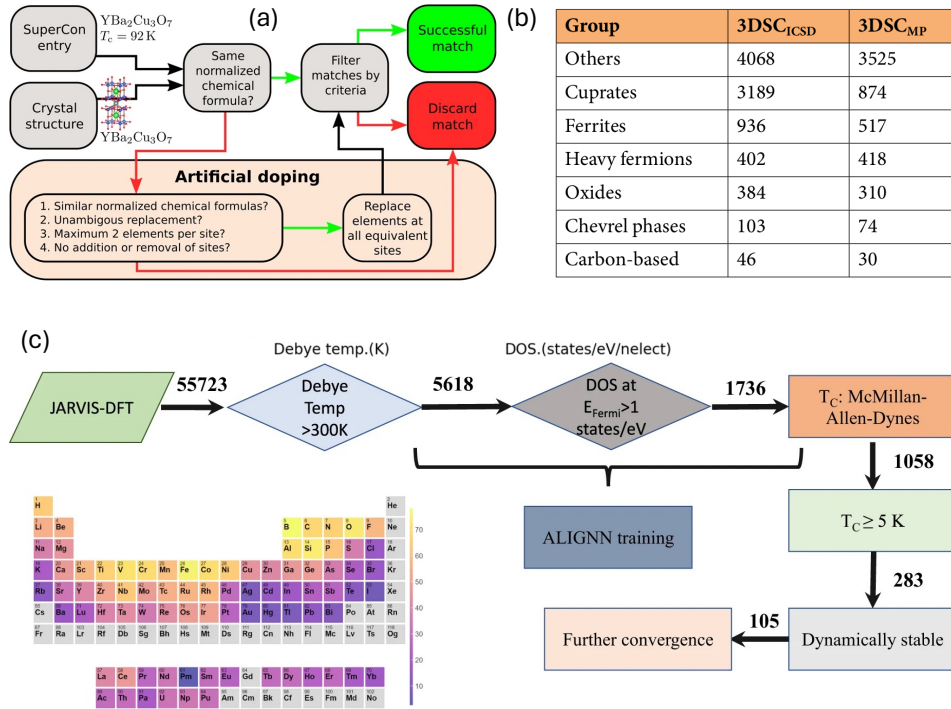
Some approaches were used to bring in atomic-level information. First, the chemical formula of a superconductors are looked up in the available databases like ICSD [289] and Materials Project [276] for the most-reasonable structures [135, 137, 143]. Fig. 10 (a) details the procedure [165] used to create 3DSC_{ICSD} and 3DSC_{MP}, two superconductor datasets in which T_c from NIMS SuperCon is paired with the atomic structures from ICSD and Materials Project by matching the chemical formula. In this procedure, an “artificial doping” step was used to obtain complete matches in the chemical formula from “nearly-complete” matches. A summary of 3DSC_{ICSD} and 3DSC_{MP} is given in Fig. 10 (b).

In the second approach, superconductivity-related parameters such as $\alpha^2F(\omega)$, λ , $\omega_{\log}$, and T_c are computed (see Sec. II B) for the atomic structures obtained from existing databases and/or predicted computationally [140, 141, 148]. JARVIS-EPC is a dataset generated [140] using such a computationally demanding approach with technical details shown in Fig. 10 (c). Starting from JARVIS-DFT, a database of 55,645 materials, screening steps involving some accessible data, e.g., Debye temperature and N_F , were used [140]. Then, $\alpha^2F(\omega)$, λ , and $\omega_{\log}$ were computed for 1,058 materials, identifying 626 dynamically stable structures, 105 of them have McMillan $T_c \geq 5$ K. More recently, a dataset of more than 7,000 records of atomic structures and their computed λ , $\omega_{\log}$, and T_c was obtained in a ML-driven computational discovery effort, whose workflow is shown in Fig. 14 (e) [141]. The two computational datasets are summarized in Table II.

The third approach is inspired by the presence of thousands of computational reports for possible superconductivity at multiple ranges of pressure. Most of them start from the atomic structures predicted computationally, thus being highly expensive and trustworthy. The main challenge in this approach is how to collect and validate the literature data in a reliable and scalable manner. In an initial effort, a few hundred atomic structures and their λ and $\omega_{\log}$, computed at pressures up to 500 GPa, were manually collected [142]. The curation involves constructing the reported atomic structures, uniformly op-

TABLE II. A summary of the currently available databases of superconductors that have been/can be used for machine-learning techniques.

Name	Description	URL (https://+)	References
NIMS SuperCon	$\simeq 31,000$ records of chemical formula & T_c @ 0 GPa	supercon.nims.go.jp/index_en.html	
MDR SuperCon	Originated from NIMS SuperCon, 26,323 records of chemical formula & T_c @ 0 GPa	mdr.nims.go.jp/collections/5712mb227	
SuperCon	Originated from NIMS SuperCon, 16,400 records of chemical formula & T_c @ 0 GPa	github.com/vstanev1/Supercon	136, 137
3DSC _{ICSD}	9,150 records of atomic structure @ 0 GPa (from ICSD) “matched” with experimental T_c (from NIMS SuperCon) via chemical formula. License needed for ICSD.	github.com/aimat-lab/3DSC	165
3DSC _{MP}	5,759 records of atomic structure @ 0 GPa (from Materials Project) “matched” with experimental T_c (from NIMS SuperCon) via chemical formula	github.com/aimat-lab/3DSC	165
Jarvis_EPC	626 records of atomic structure and λ & $\omega_{\log}$, computed @ 0 GPa using DFPT	doi.org/10.6084/m9.figshare.21370572	140
N/A	$\simeq 7,000$ records of atomic structure and λ & $\omega_{\log}$, computed @ 0 GPa using DFPT		141
CompSC	587 atomic structures for which 584 values of λ & 567 values of $\omega_{\log}$ were computed @ up to 500 GPa. Data from literature, reoptimized using DFT, & validated visually	github.com/huand/matsdata	142
SC-CoMics	1,000 annotated abstracts, developed & tailored for extracting superconductivity-related information, e.g., using NLP	data.mendeley.com/datasets/x9fjz2p3h/2 github.com/tti-coin/sc-comics	160, 161
SuperMat	142 articles, 16,052 entities, & 1,398 links, for NLP	github.com/lfoppiano/SuperMat	163
SuperCon ²	40,324 records of superconductors, T_c , applied pressure, measurement method	github.com/lfoppiano/supercon	164

FIG. 10. Workflows used to develop (a) 3DSC_{ICSD}, 3DSC_{MP} and (c) Jarvis_EPC, three datasets of superconductivity-related parameters and atomic structures. A summary of 3DSC_{ICSD} and 3DSC_{MP} is given in (b). Panels (a, b) and (c) were taken from Ref. 165 (under the Creative Commons CC BY license) and Ref. 140 (under the Creative Commons CC BY license), respectively.

timizing them using DFT, and inspecting them visually. The resulted dataset CompSC, summarized in Table II, contains 587 atomic structures for which 584 values of computed λ and 567 values of computed $\omega_{\log}$ are available [142]. This approach can create a reliable and highly

diverse data, but laborious and obviously unscalable.

Using Natural Language Processing (NLP) tools to automatically extract superconductor-related data from scientific literature, a largely unexplored data reservoir, is more scalable and sustainable. This approach has

emerged and recently gained some momentum [160–164]. SC-CoMics is a corpus of 1,000 annotated abstracts, created [160, 161] for extracting superconductivity-related information using NLP-based tools like Named Entity Recognition (NER). Going beyond abstracts, SuperMat is an annotated corpus, supplying 142 full-texts, which contain 16,052 entities and 1,398 links [163]. Such efforts were further levitated to Grobid-superconductors, a module designed to automatically extract superconductor names and properties from text, and finally to SuperCon², a database containing 40,324 records of superconductors chemical formulae, T_c , applied pressure, and measurement method [164].

B. Categorizing superconductors from data

The developed databases could be useful for two categorical questions. First, whether or not a material is a superconductor, and second, if yes, what class, e.g., cuprates and iron-based, it belongs to. In an endeavor to address the first question, SuperCon (with 16,400 records) was augmented by 300 materials found [50] to be non-superconducting. For these materials, T_c was set to zero [137]. Then, an adjustable parameter T_{sep} was introduced to separate the combined dataset into two groups. The “below- T_{sep} ” group includes the non-superconductors ($T_c = 0$ K), $\simeq 4,000$ records without T_c , and those with $T_c < T_{sep}$, while the “above- T_{sep} ” group hosts records with $T_c \geq T_{sep}$. By using the Magpie features [290] for chemical formulae, setting $T_{sep} = 10$ K, and employing Random Forest algorithm [291], the classification model developed, shown in Fig. 11 (a), can reach an accuracy of $\simeq 92\%$ [137].

By including non-superconducting materials in the “below- T_{sep} ” group, the problem of superconductor recognition was just partially addressed [137]. The reason is that T_c is not the only measurable characteristics of a superconductor but the only superconductivity-related property available in NIMS SuperCon dataset. Moreover, as T_c may be arbitrarily low [53], some materials currently classified as non-superconducting may be discovered to exhibit superconductivity as technological advancements enable us to probe lower temperature regimes. An example of such a (very rare) finding is the discovery of superconductivity in elemental Li below 4×10^{-3} K [292], despite a long-standing belief that Li would not exhibit superconductivity [293, 294].

In fact, the first attempt to categorize superconductors based on T_c emerged slightly earlier, starting from $\simeq 700$ chemical formula- T_c records collected from literature, handbooks, and NIMS SuperCon [135]. The chemical formulae were then matched with suitable atomic structures in AFLOWLIB [278], creating a dataset of 464 structure- T_c records. The data were featurized by SiRMS, a fragment-based Simplex representation [295] before learned using the Random Forest algorithm. By setting the temperature separator $T_{sep} = 20$ K, an accuracy of $\simeq 0.97$ was

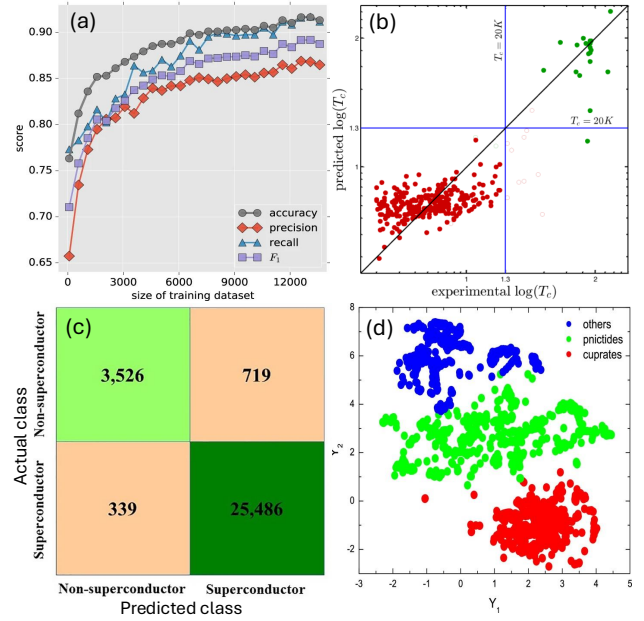


FIG. 11. (a) Four accuracy-related score of the superconductor/non-superconductor classification model, (b) predictions of a classification model (given in red and green labels) and a regression model (given on the y axis) on a dataset of 464 structure- T_c records, (c) a confusion matrix for the classification model for superconductors that can reach an accuracy of 96.5%, and (d) t-SNE plot of 4,500 randomly selected superconductors from the SuperCon, showing distinct clusters for different classes of superconductors. Panels (a), (b), (c), and (d) were reprinted with permission from Ref. 137 (under the Creative Commons CC BY license), Ref. 135 (Copyright 2015, American Chemical Society), Ref. 158 (Copyright 2020, Elsevier), and Ref. 159 (Copyright 2022, Elsevier), respectively.

search by the obtained classification model, visualized in Fig. 11 (b).

In an effort to distinguish superconductors from non-superconductors, NIMS SuperCon was cleaned and augmented by 3,000 non-superconducting materials, including insulators, semiconductors, and metals [158]. Then, each chemical formula was represented by a row vector formed by the composition (contribution) of the constituent species. These vectors were aggregated in a chemical composition matrix that has 96 columns for 96 available species and $\simeq 30,000$ rows for the dataset size. Fig. 11 (c) shows that the classification model trained using the k-Nearest Neighbors algorithm performs very well [158] with an overall accuracy of $\simeq 96\%$. It is worth noting that the augmented 3,000 materials were assumed [158] to be non-superconductors, skipping the aforementioned small possibility that some (metals) of them may be superconductors at very low temperatures.

The second question involves recognizing superconductors of different classes, i.e., those governed by different pairing mechanisms. Addressing this problem, NIMS SuperCon was cleaned and represented [159] by the chem-

ical composition matrix [158]. Several clustering algorithms were tested before Self-Organizing Map [296] was selected. Fig. 11 (d) shows the t -distributed Stochastic Neighbor Embedding (t -SNE) [297] of 4,500 superconductors randomly selected from SuperCon, in which iron-based, cuprates, and those in other classes are distinguished [159]. It seems that data, when curated and learned properly, could be useful to recognize a superconductor and the governing mechanism, if applicable.

C. Predicting superconducting-related properties

ML efforts in this class aim at accelerating the predictions of superconducting-related properties such as T_c , traditionally obtained by expensive computations (Sec. II B) and/or physical measurements (Sec. II D). The critical role of NIMS SuperCon and its descendants, evidenced in Sec. IV B, is also visible here. In fact, most of the works aiming at predicting T_c from the chemical formula rely on NIMS SuperCon. Efforts to introduce atomic-level information in the development of ML models for λ and $\omega_{\log}$ emerged recently. This Section is devoted to not only the ML efforts in the two subcategories but also the ML-driven searches for superconductors. An in-depth discussion on the remaining challenges and opportunities of ML efforts in this class is given in Sec. V B 2.

1. Predictions from chemical formula

In an early ML work starting from NIMS SuperCon, a dataset of 21,263 records was curated, containing multiple classes of superconductors [136]. For each material, the chemical formula was featurized by some simple functions, e.g., mean, weighted mean, and entropy, etc., of the basic properties of the constituent species, e.g., atomic mass and electron affinity, etc. Some algorithms were tested and XGBoost [298] was selected. The developed ML model for T_c with an averaged out-of-sample error of $\simeq 9.5$ K is visualized in Fig. 12 (a) [136].

Following the classification between low- T_c and high- T_c superconductors, T_c predictive models were developed for the high- T_c group [137]. The materials data were represented by Magpie (composition) features [290] and learned by the Random Forest (regression) algorithm. One of the models developed is visualized in Fig. 12 (b), showing the performance on low- T_c , Fe-based, and cuprate superconductors with a coefficient of determination $R^2 \simeq 0.88$ [137]. Screening ICSD database by a combination of a classification and a regression model, $\simeq 2,000$ materials with predicted $T_c > 20$ K at 0 GPa were identified [137]. While most of them contain copper and oxygen, i.e., they may be related to cuprates, a subset of 35 materials without obvious connection to known high- T_c families was compiled and reported [137].

Roughly 2,000 records of AlB_2 -, Chevrel-, Cr_3Si -, spinel-, NaCl -, and skutterudite-type superconductors

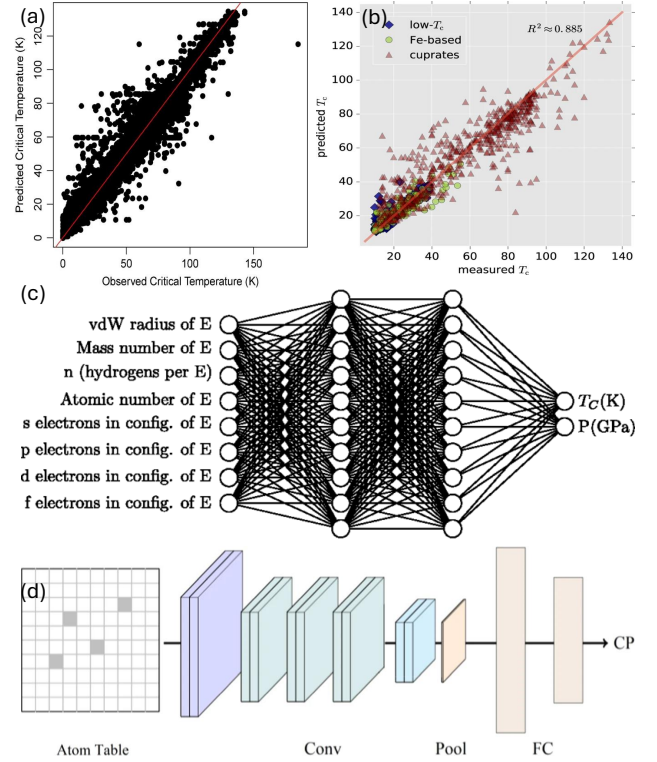


FIG. 12. (a, b) Two ML models for predicting T_c from chemical formula, (c) a Fully-Connected Neural Network created to predict T_c and P from chemical formula, and (d) an atom table Convolutional Neural Network used to train ML predictive model for T_c on the datasets whose chemical formulae are represented as 10×10 images. Panels (a), (b), (c), and (d) were reprinted with permission from Ref. 136 (Copyright 2018, Elsevier), Ref. 137 (under the Creative Commons CC BY license), Ref. 147 (Copyright 2020, American Physical Society), and Ref. 138 (under the Creative Commons CC BY license), respectively.

were extracted from NIMS SuperCon and augmented with elemental superconductors [144]. This dataset was believed to include phonon-mediated superconductors, for which MgB_2 [47] is the highest- T_c material – the averaged T_c of MgB_2 in this dataset is 38.6 K. Featurizing the data by some simple functions of the fundamental attributes of existing species, the trained Random Forest model could reach an $R^2 = 0.98$ on the training data (80% of the dataset) and $R^2 = 0.92$ on the test data (the remaining 20% of the dataset). This model was used to create predicted T_c map for three family of ternary materials, namely Mg-B-Ti, Fe-Te-Se, and Ca-B-C. In the Mg-B-Ti map, the region containing MgB_2 was predicted to be highest in T_c while in the Fe-Te-Se map, the high T_c was predicted along the line connecting FeTe and FeSe [144]. For the Ca-B-C system, materials with formula close to CaB_6 and B_{13}C_2 were also predicted [144] to be high in T_c .

Focusing on the recent discoveries, a few hundreds of binary hydride superconductors EH_n were collected from

literature [147]. For each material, its chemical formula, T_c , and the pressure at which the superconductivity was predicted, are available. A set of features, including the hydrogen content n and some fundamental attributes of the species E such as the atomic number, the van der Waals radius, and the electron configurations of E was used to describe EH_n . The featurized data was then fed into a Fully-Connected Neural Network, visualized in Fig. 12 (c), whose output layer has two nodes, one for T_c and the other for P . This network is an example of a multi-task learning architecture, which can be trained multiple datasets to exploit the hidden correlations among them and leverage the performance [288, 299, 300]. The T_c predictivity of the model reaches $R^2 = 0.88$ with root mean square error (RMSE) $\simeq 33.7$ K for T_c predictions.

Nevertheless, the main objective of the developed Neural Network model is to screen over the Periodic Table for the species E that minimize the “distance” from the predicted T_c and P to the ambient conditions, i.e., $P = 0$ GPa and 293 K [147]. The analysis suggests that alkali and alkaline-earth metal hydrides could be the best candidates for superconductivity near ambient conditions. Next, the AIRSS method [222] was used, identifying dozens of atomic structures of alkali and alkaline earth metal hydrides with respectable computed T_c . Specifically for the predicted $C2/m$, $Cmcm$, and $Immm$ phases of RbH_{12} , the computed T_c could be as high as 126 K at 100 GPa and below [147].

Materials informatics endeavors relying on NIMS SuperCon were extended into the deep-learning territory, where high-level features may be learned directly from raw data. In the atom table Convolutional Neural Network (ATCNN) [138], each chemical formula was represented by an image of 10×10 pixels, visualized at the left end of Fig. 12 (d). Each pixel corresponds to a species, and its value is the composition (contribution) of this species in the chemical formula. As there are 86 species that appear in the dataset, thus $10 \times 10 = 100$ pixels is sufficient for the representation. Then, the “atom tables” are accepted by an architecture, sketched in Fig. 12 (d), which consists several convolutional layers to process the images [138]. Two models for T_c , i.e., ATCNN-I and ATCNN-II, were developed, one trained on a cleaned version of NIMS SuperCon containing 13,598 records and the other trained on the same dataset after being augmented by 9,399 energetically stable insulators, the obvious non-superconductors. Both models show good performance with a mean absolute error (MAE) of $\simeq 4.2$ K, RMSE of $\simeq 8.2$ K, and $R^2 \simeq 0.97$. Compared to the measured T_c of some well-known superconductors such as Hg, MgB_2 , and $\text{YBa}_2\text{Cu}_3\text{O}_7$, the predictions of these two models are accurate [138].

The idea of using a Convolutional Neural Network (CNN) architecture in superconductor discovery has evolved from recognizing “atom tables” [138] to “reading periodic tables” [139]. In the latter, the species composition of each superconductor in NIMS SuperCon is “written” directly to the Periodic Table. Then, the

table was separated into four “channels” for recording those with s , p , d , and f valence electrons. A CNN was trained on 95% of NIMS SuperCon and tested on the remaining 5% of the data, yielding $R^2 = 0.92$ in T_c predictions [139]. This model predicts about 17,000 materials (out of $\sim 48,000$ records) in Crystallography Open Database (COD) [301], to have $T_c > 10$ K. The obtained result is unreasonable, perhaps because the training data have almost no non-superconductors. After augmenting it by a synthetic dataset of non-superconductors (assumed $T_c = 0$ K), the final (new) model becomes more reliable. Testing on 400 materials (including 330 non-superconductors) reported in Ref. 50, the model reaches a precision of 62%, an accuracy of 76%, a recall of 67 %, and an f1 score of 63% in predicting materials with $T_c > 0$ K. Using it for COD, 70 materials, including CaBi_2 and $\text{Hf}_{0.5}\text{Nb}_{0.2}\text{V}_2\text{Zr}_{0.3}$ (both are not in NIMS SuperCon) were predicted to have $T_c > 10$ K. One of them, CaBi_2 is indeed a superconductor [302].

We close this Section by noting the connection from the “atom table” and the “periodic table” in these CNN approaches to the chemical composition row vector [158] discussed in Sec. IV B. In particular, the former is the latter rearranged in a two-dimensional image so that it can be used in a CNN architecture. The main information encoded in these images is the composition of the constituent species, the highest level of information that can be extrated from a chemical formula.

2. Predictions from atomic structure

In fact, attempts to bring atomic-level details to ML models for T_c emerged quite early [135]. By matching $\simeq 700$ formulae collected for superconductors with AFLOWLIB and excluding those with $T_c < 2$ K, a dataset of 295 atomic structure- T_c records was obtained. The atomic structures were represented using SiRMS [295] before being mapped onto T_c by Random Forest and Partial Least Squares [303] algorithms. The obtained models for T_c could reach $R^2 \simeq 0.66$, one of them is shown in Fig. 11 (b). As SiRMS is a fragment-based representation [295], an analysis was performed, compiling a catalog of fragments that may likely present in materials with low and high values of T_c [135].

An atomic structure matching procedure was also used in Ref. 143. By looking at ICSD database for which atomic structures are available, $\simeq 1,700$ superconductors were obtained. Next, by identifying suitable (close-matched) structures and “doping” them to recover the chemical formulae in NIMS SuperCon, a dataset of 5,713 atomic structure- T_c records was finally obtained [143]. Then, the atomic structures were featurized using the “Smooth Overlap of Atomic Position” (SOAP) scheme [304]. Finally, three models were trained used Random Forest, XGBoost, and Support Vector Regression [305]. One of them, the Random Forest regression model with $R^2 \simeq 0.92$, is visualized in Fig. 13 (a) [143]. These mod-

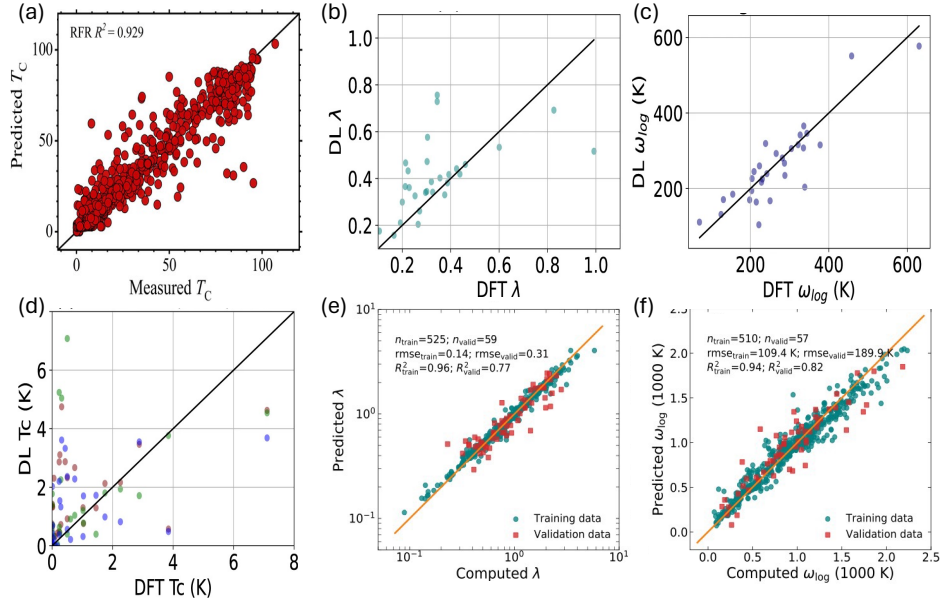


FIG. 13. Nine ML predictive models starting from atomic structures to predict (a, d) T_c , (b, e) λ , and (c, f) $\omega_{\log}$. Panels (a), (b, c, d), and (e, f) were reprinted with permission from Ref. 143 (Copyright 2019, American Chemical Society), Ref. 140 (under the Creative Commons CC BY license), and 142 (Copyright 2023, American Physical Society), respectively.

els were used to screen over ICSD database, compiling 10 most promising superconductors whose averaged T_c (predicted by 3 models) is at least ≈ 20 K at 0 GPa [143]. Among them, $\text{Ba}_4\text{Ca}_4\text{Cu}_6\text{O}_{19}\text{Ti}_3$ was predicted to have $T_c \approx 103$ K [143].

Starting from JARVIS-EPC, two learning methods were used to develop predictive models for λ , $\omega_{\log}$, and T_c [140]. The first relies on some force-field-inspired features and the Gradient Boosting Decision tree algorithm [306]. The second, referred to as “Atomistic Line Graph Neural Network” (ALIGNN), is a deep-learning architecture, in which a graph convolutional layer was designed to describe two- and three-body interactions among the atoms of an atomic structure [307]. Figs. 13 (b), (c), and (d) visualize three ALIGNN models [140] trained on JARVIS-EPC to predict λ , $\omega_{\log}$, and T_c , respectively.

Two ML models trained [142] on CompSC to predict λ and $\omega_{\log}$ are shown in Figs. 13 (e) and (f). In this work, the atomic structures predicted at pressures up to 500 GPa and reported in literature were collected, uniformly reoptimized, visually validated, represented using MATMINER [308], and learned using Gaussian Process Regression (GPR) algorithm [309, 310]. The training data are highly diverse, containing numerous (“unusual”) atomic details realized at different P and computationally linked to the values of λ and $\omega_{\log}$ that lead to high values of T_c . Therefore, the models are expected to be capable of recognizing at any P , including 0 GPa, the atomic structures that resemble the “unusual” atomic-level details they were exposed to [142]. Using these models to screen the hydrides from Materials Project database, an $Fm\bar{3}m$ structure of CrH and another $Fm\bar{3}m$ structure of CrH_2 were identified with computed $T_c = 15.7$ K and

10.7 K, respectively, at 0 GPa [142].

3. ML-driven search for high- T_c superconductors

An efficient materials discovery strategy, even powered by ML models, should be target-driven, extending beyond a brute-force screening, as discussed in Sec. II C. One such workflow was developed [146], utilizing an evolutionary algorithm and a GPR model to discover hydrogen-containing superconductors. In this ML-driven strategy, visualized in Fig. 14 (a), mating and mutation operate directly on the atomic structures while space group, hydrogen concentration, atomic mass, pressure, and μ^* were used as descriptors to train the GPR model [146]. Having new candidates, new computed data come, the ML model is retrained, and the workflow cycles. Some hydride superconductors were discovered, including a $C2/m$ structure of KScH_{12} with computed $T_c = 122$ K at 300 GPa and a $Pm\bar{3}$ structure of GaAsH_6 , shown in Fig. 14 (b), with computed $T_c = 98$ K at 180 GPa.

Another ML-powered discovery strategy for hydride superconductors was recently developed [148] and visualized in Fig. 14 (c). In this procedure, the GPR model was trained on some formula-based features like the Gaspari-Gyorffy electron-phonon coupling estimates of available species, the atomic mass of non-hydrogen species, and N_F ; all of them were normalized appropriately. In this strategy, the atomic structures of the formula candidates were searched using AIRSS [222], and as the workflow cycles, 27 new superconductors were discovered at pressures up to 500 GPa. Among them, NaH_6 , shown in Fig. 14 (d), was predicted to have T_c in the range from 228 K

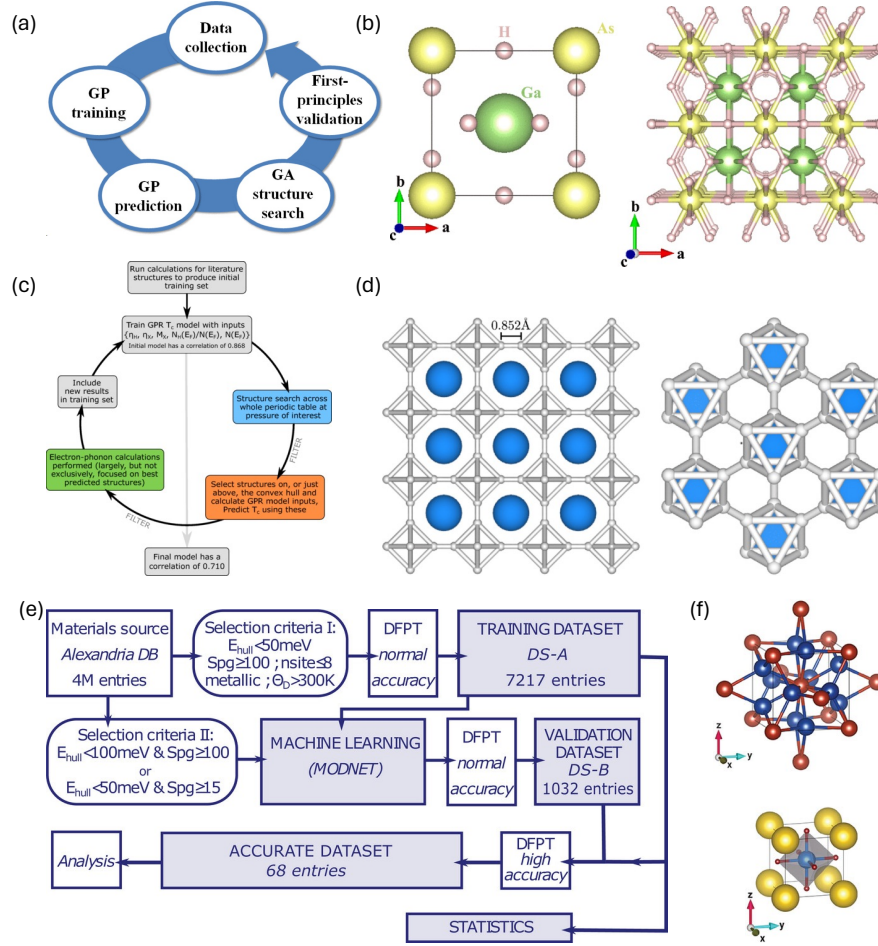


FIG. 14. (c) A discovery loop using a ML predictive model to search for high- T_c superconductors, (d) the atomic structure of discovered GaAsH_6 , (e) the atomic structure of discovered NaH_6 , and (f) the Éliashberg spectral function $\alpha^2 F(\omega)$ of NaH_6 and some related materials. Panels (a, b), (c, d), and (e, f) were taken from Ref. 146 (under the Creative Commons Attribution 4.0 International license), Ref. 148 (Copyright 2021, American Physical Society), and Ref. 141 (Copyright 2024, John Wiley and Sons), respectively.

to 279 K at 100 GPa [148].

D. Accelerated structure prediction with ML potentials

Fig. 14 (e) visualizes the workflow used to create the dataset of $\simeq 7,000$ records discussed in Sec. IV A and summarized in Table II [141]. In this workflow, a deep-learning model trained on a “Material Optimal Descriptor Network”, or MODNet [311], was used to select candidates from Alexandria database [312] for λ , $\omega_{\log}$, and T_c computations. MODNet accepts features computed using MATMINER [308] for the atomic structures before passing them through a series of successive blocks for feature selection, encoding, decoding, and splitting to learn multiple properties concurrently. Among $\simeq 7,000$ data records produced using this workflow, 541 materials with $T_c > 10$ K (at 0 GPa) were identified. Two of them, a $Pm\bar{3}n$ structure of Ti_3Te (computed $T_c = 16.3$ K) and a $Pm\bar{3}m$ structure of KCdH_3 (computed $T_c = 12.3$ K) are visualized in Fig. 14 (f) [141].

In the computational discovery workflow discussed in Sec. II C, the structure prediction step is computationally very demanding. For each chemical formula, hundreds of thousands of energy evaluations are typically needed to search for its stable atomic structures, and they must be performed at a first-principles level. The development of ML potentials [313], mostly in the context of molecular dynamics (MD) simulations, offer an approach to accelerate this step. A ML potential is a ML model, generally based on a Neural Network, accepting an atomic structure and returning its potential energy faster than a normal DFT calculation by multiple orders of magnitude. *State-of-the-art* ML potentials such as Gaussian Approximation Potential (GAP) [314, 315], Spectral Neighbor Analysis Potential (SNAP) [316, 317], Moment Tensor Potentials (MTP) [318, 319],

and Deep Potential [320], trained on millions of atomic environments, are expected to be comparable with DFT in accuracy, i.e., a few meV/atom. However, like any ML models, out-of-domain energy predictions (on unseen atomic environments) should be treated with care. In superconductor discovery, ML potentials started leaving discernible trails [127–129, 131, 132].

A reliable atomic structure predicted computationally should be thermodynamically, dynamically, and kinetically stable. In the context of superconductor discovery, the dynamical stability of a structure, i.e., whether it is a local minimum, not a saddle point of the PES, is accessible during the calculations of the spectral function $\alpha^2 F(\omega)$ using DFPT. Examining the kinetic stability of a structure, i.e., if the kinetic (energy) barrier protecting this local minimum is high enough so its lifetime could be reasonable, is much harder than inspecting its thermodynamic stability. The reason is that the former is non-local in nature, especially when the configuration space is extremely high in dimensionality, which is roughly 3 times of the system size, i.e., in orders of 10^2 .

In most of the cases, ML potentials were used to accelerate the energy evaluations during the search, i.e., their role is on the side of thermodynamic stability. Some results of these efforts are the discoveries at $P = 0$ GPa of C_4K (in $P4/mmm$ symmetry with predicted $T_c = 30.4$ K) [127], $c-B_{24}$ (in $Pm\bar{3}$ symmetry with predicted $T_c = 13.8$ K) [128], and Mg_2IrH_6 (in $Fm\bar{3}m$ symmetry with predicted $T_c = 160$ K) [129]. At a slightly higher pressure ($P = 20$ GPa), a series of 14 La-N-H trinary materials was predicted with support from the universal Neural Network potential developed by MATLANTIS [321]. Examining their possible superconductivity by DFPT calculations [130], the predicted T_c of this series spans from 0.49 K to 14.41 K. In another interesting effort [131], ephemeral data-derived (ML) potentials [322] were used in a large-scale structure prediction campaign, ending up disproving an early claim of ambient-condition superconductivity in the Lu-N-H systems. Compared with positive conclusions, such a negative conclusion requires *much bigger* computational effort because all the attainable possibilities, e.g., chemical formulae and system sizes, should be considered. ML potentials are particularly useful for this purpose.

ML potentials have also been used to accelerate the examination of the kinetic stability, which connects with kinetic barriers. A typical approach is to perform long first-principles MD simulations at room temperatures and above to examine the stability of the long-range order against elevated thermal energies [127]. Accessing directly to the kinetic barriers, e.g., using stochastic self-consistent harmonic approximation, requires sufficiently large supercells and numerous randomly displaced structures [132]. With help from MTP, such a computationally expensive analysis was completed, suggesting the stability of the $Fm\bar{3}m$ phase of $BaSiH_8$, whose predicted T_c is about 90 K [132].

E. Refining empirical formulae for T_c

The empirical McMillan formula (8) of T_c was developed [170] and refined [171, 172] by “manually learning” small datasets generated from the Éliashberg equations. This formula, and some other variants, are particularly useful in the search for new superconductors. Recent ideas emerged [133, 134], suggesting that these formulae may be refined further using advanced symbolic ML techniques and bigger datasets of the solutions of the Éliashberg equations, which become available thanks to new generations of computational infrastructures.

One such symbolic ML techniques is “Sure Independence Screening and Sparsifying Operator” (SISSO) [323]. By defining some physically meaningful operators and functions of the primary variables, which are λ , $\omega_{\log}$, and μ^* , millions of features (expressions) were generated. Training a linear regression model using these features with L_0 regularization, some expressions for T_c were obtained, among them the simplest version is [133]

$$T_c = 0.0953 \frac{\lambda^4 \omega_{\log}}{\lambda^3 + \sqrt{\mu^*}}. \quad (10)$$

In a subsequent effort [134], a modified version of the McMillan formula (8) was assumed as

$$T_c = \frac{f_\omega f_\mu \omega_{\log}}{1.2} \exp \left[-\frac{1.04(1 + \lambda)}{\lambda - \mu^*(1 + 0.62\lambda)} \right] \quad (11)$$

in which two prefactors f_ω and f_μ are learned from data to be

$$f_\omega = 1.92 \frac{\lambda + \omega_{\log}/\bar{\omega}_2 - \sqrt[3]{\mu^*}}{\sqrt{\lambda} \exp(\omega_{\log}/\bar{\omega}_2)} - 0.08 \quad (12)$$

and

$$f_\mu = \frac{6.86 \exp(-\lambda/\mu^*)}{1/\lambda - \mu^* - \omega_{\log}/\bar{\omega}_2} + 1 \quad (13)$$

where [172]

$$\bar{\omega}_2 = \left[\frac{2}{\lambda} \int_0^\infty d\omega \omega \alpha^2 F(\omega) \right]^{1/2}. \quad (14)$$

Likewise, data obtained from the ML-powered discovery strategy developed [148] for hydride superconductors offer an opportunity to test and improve the McMillan formula (8). Assuming

$$T_c = \frac{\omega_{\log}}{1.2} \exp \left[-\frac{1.04(1 + \lambda)}{\lambda - \mu^*(1 + 0.62\lambda)} \right] (a + b\lambda), \quad (15)$$

data fitting yields $a = 1.0061$ and $b = 0.0663$ [148], slightly modifying the McMillan formula (8).

V. CHALLENGES, OPPORTUNITIES, AND CRITICAL NEXT STEPS

A. Theoretical and computational methods

Despite thousands of computational discoveries reported in the last two decades, only about 20 – 30 materials were synthesized and reported. This small fraction (about 1% or below) may suggest that we simply do not have a “silver bullet” for such a tough problem like superconductivity understanding, prediction, and discovery. Looking forward, this situation may change, but some major challenges must be resolved. Apparently, phonon-mediated pairing is one of many mechanisms proposed for superconductivity, and an exact and/or well-controlled theory for this mechanism is not readily available. It is known that Migdal-Éliashberg theory may become inaccurate [173–178], and one reason could be the fact that it accounts only for the first order of vertex corrections [21]. Challenges in pushing the theoretical front in superconductivity understanding are obviously enormous.

In the computational front, two major steps of the workflow discussed in Sec. II C, namely the atomic structure prediction and the computations of $\alpha^2 F(\omega)$, λ , and $\omega_{\log}$, are highly non-trivial. The former faces truly infinite, extremely high-dimensional configuration spaces. For the latter, implications of the possible breakdown of Migdal-Éliashberg theory have not been well understood while the desirable accuracy and convergence are hard to control and obtain [140–142, 148]. Technically, even a numerical error that is comparable with the best attainable level of accuracy, i.e., $\sim 10^{-4} - 10^{-3}$ eV/Å in the atomic force calculations, could be translated into a sizeable change in the computed phonon frequencies, and thus, in the superconductivity-related properties.

For a more quantitative assessment, we show in Fig. 15 some computed and measured data of T_c of some notable superconductors, obtained from the same atomic structure under the same pressure. Four materials with highest- T_c shown in this Figure are MgB₂ at ambient pressure,[47, 166, 324–328] H₃S at 200 GPa,[99, 190] YH₆ at 166 GPa,[104] and LaYH₂₀ at 180 GPa.[211] For each of them, multiple computational schemes were used, involving different values of μ^* and whether or not the anharmonicity is included. These empirical treatments allow the computed T_c to spread over a sizable range, i.e., $\simeq 29$ K for MgB₂, $\simeq 125$ K for H₃S, $\simeq 100$ K for YH₆, and $\simeq 70$ K for LaYH₂₀. Considering the significant challenges in both experimental and computational techniques, the moderate agreement is reasonable and understandable, highlighting important cautions for future works.

Some emerging techniques may be useful for these challenges. Given that the atomic structure predictions can be effectively coupled with powerful ML potentials [127–132], they are expected to be further accelerated by the next developments of this fast-evolving field [313–320]. Further into the future, when deep-learning can be used

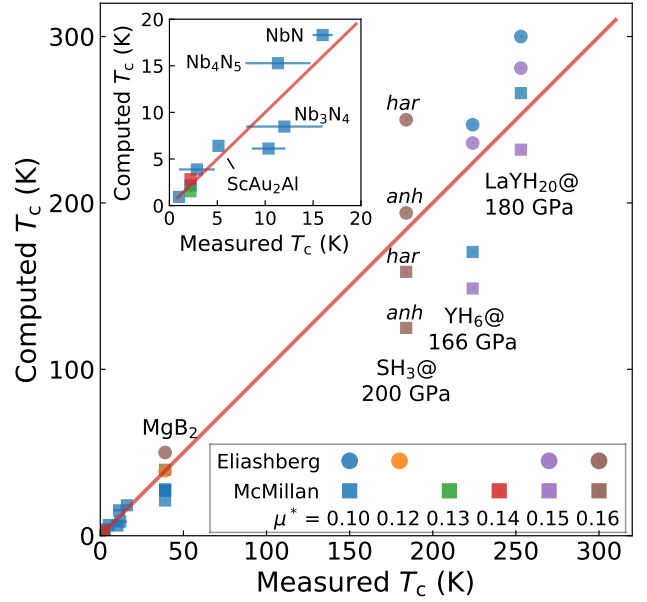


FIG. 15. The critical temperature T_c , computed for some notable superconductors using McMillan formula and solving the Éliashberg equations, given in comparison with measured T_c . Value of μ^* and whether or not anharmonicity, labeled by **anh** and **har**, is used, are shown. The inset is used for the small- T_c regime. The x -axis error bars were obtained from different measured values of T_c while different computed values of T_c are shown separately for further discussions.

to accelerate the essential quantum mechanics-based calculations like DFT [329–333] and DFPT [334] in practice, significant advancements may be envisioned. On the other hand, some advanced physics-based computational approaches are also under developments. Among them are SuperConducting DFT [335–337], a fully *ab-initio* approach for T_c , and a GW perturbation theory based scheme for computing the electron-phonon coupling [338]. Once these methods can be fully demonstrated, further advances in this field may be anticipated.

B. Materials informatics approaches

Different from the experimental and theoretical efforts with a century of history and the computational approaches with two decades of blossoming, materials informatics has less than a decade of development in superconductivity understanding and discovery. Compared to the DFPT-based computational approaches, the driving force behind *almost all of the discoveries* of hydride superconductors discussed in Sec. III, materials informatics remains in its early stage, both in terms of the methodology development and the impacts to the field. Nevertheless, opportunities exist on the horizon for the informatics-based methods to efficiently complement the traditional (computational and experimental) approaches. Toward this future, some critical challenges,

discussed below, are waiting to be resolved.

1. Data generation and accumulation

Data development is the most critical challenge of materials informatics approaches in superconductor discovery. In this area, available data are scarce and insufficient. The biggest superconductor dataset is **SuperCon**², developed in 2023 with $\simeq 40,000$ records. The most widely-used dataset, i.e., **NIMS SuperCon**, in its cleaned version (**SuperCon**), has $\simeq 16,000$ records [137]. For both of them, only the chemical formula is available for the description of the materials. Attempts to augment the chemical formula with atomic details using structure matching/lookup method produced datasets of up to $\simeq 5,700$ records in size [143], as they rely on **NIMS SuperCon** for the experimental values of T_c . Such datasets are small and less informative compared to others branches of materials informatics.

Beside a few high-throughput computational efforts that can generate hundreds to thousands of data records of λ , $\omega_{\log}$, and computed T_c [140, 141, 148–150], the vast majority of the literature can only report results from structure predictions works performed for a handful of materials. This is the main reason limiting the available computational data in the field. Initial attempt to collect atomic-level data from literature [142] returns nearly 600 records, but more efficient and scalable methods are needed. Given some known limits of the computational approaches, computed data, when developed, should be combined with (more trustworthy) experimental data and “machine” learned in a collective and complementing manner.

Scientific literature also hosts a huge, yet essentially untouched, experimental data reservoir. During the last decade, NLP-based approaches such as NER have demonstrated their power in creating **SC-CoMics** [160, 161], **SuperMat** [163], and **SuperCon**² [164]. Significant future developments are anticipated for these methods can be used to extract the crystallographic information, typically distributed throughout the full-text and in Tables, Supporting Information material, and, in few cases, in some atomic structure formats. In such a challenging endeavor, Large Language Models (LLM) like ChatGPT [339] and Meta Llama 2 and 3 [340] could be useful, as recently demonstrated [341, 342]. Nevertheless, enormous challenges are envisioned for this approach to become mature and efficient in literature data extraction.

2. Deep-learning for the physics of superconductivity

Superconductivity is highly collective and non-local in nature. The coherence length of the Cooper pairs can range from a few nanometers as in LaH_{10} [101], YH_6 , YH_9 [105] and CeH_9 [106] to several hundred nanometers as in many known superconductors [343]. Chemi-

cal formulae do not include such information, and thus, are insufficient for informatics techniques. Even when the atomic structures are used to describe the materials, recognizing such behaviors is impossible for local-environment- and fragment-based featurizing schemes like SOAP [304] and SiRMS [295]. In such the “hand-crafted” featurizing schemes, the typical “neighboring” distance cutoff of $\sim 10\text{\AA}$ is far too small for the physics of superconductivity. However, bringing this length scale to the order of hundreds of nanometers will certainly (and exponentially) blow up the size and the complexity of feature vectors, making them impractical.

In this context, deep-learning techniques that can directly accept the atomic structures (of course, when data are available) to understand both the local atomic environments and the long-range orders of the materials, could be useful. Deep Neural Network based architectures [139] like ATCNN [138] and ALIGNN [307] have been developed and used in some cases with encouraging results. Nevertheless, further developments for this challenging problem, going beyond these initial steps, are needed and anticipated. Equivariant Neural Networks, in which essential physics-inspired invariances are respected, starts emerging and shows its applicability in materials informatics [281–283].

C. ML-guided search for ambient pressure superconductors

After two decades being busy with discoveries at hundreds of GPa, recent attentions are gradually shifted to the search for possible high- T_c superconductivity at lower pressures [97, 129, 140–142, 147]. Given the infinite materials space and the complexity of the traditional physics-based approaches, rapid ML models for superconductivity predicting, as discussed in Sec. IV, could be useful, but not without challenges. Adding to the challenges outlined in Secs. VB1 and VB2 for developing the ML models, another notable problem is that inferences made outside, or far from, the scope of their experience, are generally not good [344]. Thus, if a ML model is trained on the superconductors whose $T_c \leq 150$ K (at ambient pressure), it may not be able to recognize superconductors with higher T_c .

One possible way to expand the “domain of applicability” is to include in the training data the atomic-level details of the (computational and experimental) discoveries at any pressures, for which T_c could be much higher (e.g., $\simeq 200$ K for H_3S and $\simeq 250$ K for LaH_{10}) [142]. Trained on such a dataset, the obtained models may recognize the atomic details related to high- T_c superconductivity at any pressures, including ambient pressure. In this approach, external pressure P is assumed to connect with the superconductivity indirectly, i.e., P determines the atomic structures, which, in turn, determines the superconductivity. This assumption has its roots in thermodynamics, where P , a macroscopic concept, de-

termines the atomic structures through $P = -\partial H/\partial V$, given that both the enthalpy H and the volume V of an unit cell are solely functions of the atomic structures. The current edition of such a training data, i.e., **CompSC**, is small, containing 587 atomic structures with 584 values of λ and 567 values of $\omega_{\log}$ [142]. When a substantially bigger and more diverse version is available and advanced deep-learning techniques are developed, the ML models trained on this dataset may be more effective and useful.

Along another dimension, potential non-hydrogen high- T_c superconductors have not been explored appropriately. Although hydrides may have favorable vibrational frequency spectrum for high- T_c superconductors, this field remains big for other enigmas and wonders. As **CompSC** currently contains mostly hydrides,[142] the inclusion of non-hydrogen superconductors would be critical towards an exploration beyond the hydride-related boundary.

VI. SUMMARY

Research efforts devoted to superconductor discovery in the last two decades are massive and incredible. Among thousands of superconductors predicted computationally at hundreds of GPa, a few dozens were synthesized and characterized experimentally. The inspiring results have somehow rekindled the dream of room-temperature superconductors. Nevertheless, this “holy grail” remains far from being attainable while technical

challenges are numerous and enormous. Many of them are related to conducting measurements at hundreds of GPa in a DAC, and then analyzing/interpreting the obtained data [52, 110–116]. Theoretical foundations and computational approaches, on the other hand, have their own inherent hard limits that are not easy to overcome.

Materials informatics approaches, emerging as a new frontier of materials research, could be useful to complement the traditional approaches. In the last decade, some essential components of materials informatics for superconductor discovery have been developed, reaching an inspiring level of maturity. Toward the future, tremendous opportunities exist when enormous challenges can be resolved. The most notable challenges are, but not limited to, those related to data generation and curation, deep learning techniques that can capture the physics of superconductivity, and a reliable ML-guided search protocol for high- T_c superconductors at ambient pressures. Looking at other branches of materials informatics where tremendous advancements have been made, we believe that superconductor discovery may be significantly advanced with the new developments in this new frontier of materials research.

ACKNOWLEDGEMENTS

Some data used for Figs. 3 were prepared with support from the San Diego Supercomputer Center (Expanse) within the XSEDE/ACCESS allocation number DMR170031. Work by VNT is supported by the Vietnam National Foundation for Science and Technology Development under Grant 103.01-2021.12.

-
- [1] H. K. Onnes, Commun. Phys. Lab. Univ. Leiden , 119, 120, 122 (1911).
 - [2] W. Meissner and R. Ochsenfeld, Naturwissenschaften **21**, 787 (1933).
 - [3] H. G. Smith and J. O. Wilhelm, Rev. Mod. Phys. **7**, 237 (1935).
 - [4] J. Eisenstein, Rev. Mod. Phys. **26**, 277 (1954).
 - [5] B. T. Matthias, T. H. Geballe, and V. B. Compton, Rev. Mod. Phys. **35**, 1 (1963).
 - [6] F. London, Phys. Rev. **74**, 562 (1948).
 - [7] L. D. Landau and V. L. Ginzburg, Zh. Eksp. Teor. Fiz. **20**, 1064 (1950).
 - [8] H. Fröhlich, Adv. Phys. **3**, 325 (1954).
 - [9] H. Fröhlich, Proc. R. Soc. London, Ser. A **223**, 296 (1954).
 - [10] J. Bardeen and D. Pines, Phys. Rev. **99**, 1140 (1955).
 - [11] J. Bardeen, Phys. Rev. **97**, 1724 (1955).
 - [12] L. N. Cooper, Phys. Rev. **104**, 1189 (1956).
 - [13] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **106**, 162 (1957).
 - [14] J. Bardeen, Science **181**, 1209 (1973).
 - [15] J. R. Schrieffer, *Theory of superconductivity* (CRC press, 2018).
 - [16] M. Tinkham, *Introduction to superconductivity* (Courier Corporation, 2004).
 - [17] V. L. Ginzburg and E. A. Andryushin, *Superconductivity (Revised Edition)* (World Scientific, 2004).
 - [18] J. Carbotte, Rev. Mod. Phys. **62**, 1027 (1990).
 - [19] G. Eliashberg, Sov. Phys. J. Exp. Theor. Phys. **11**, 696 (1960).
 - [20] A. B. Migdal, Sov. Phys. J. Exp. Theor. Phys. **7**, 996 (1958).
 - [21] A. V. Chubukov, A. Abanov, I. Esterlis, and S. A. Kivelson, Ann. Phys. **417**, 168190 (2020).
 - [22] F. Marsiglio, Ann. Phys. **417**, 168102 (2020).
 - [23] P. B. Allen and B. Mitrović, Solid State Phys. **37**, 1 (1983).
 - [24] C. Pellegrini and A. Sanna, Nat. Rev. Phys. **6**, 509 (2024).
 - [25] V. Z. Kresin, H. Morawitz, and S. A. Wolf, *Mechanisms of conventional and high T_c superconductivity*, Vol. 84 (Oxford University Press, 1993).
 - [26] G. Stewart, Adv. Phys. **66**, 75 (2017).
 - [27] P. W. Anderson, *The theory of superconductivity in the high- T_c cuprate superconductors* (Princeton University Press, 1997).
 - [28] G. Blatter, M. V. Feigel'man, V. B. Geshkenbein, A. I. Larkin, and V. M. Vinokur, Rev. Mod. Phys. **66**, 1125 (1994).

- [29] P. A. Lee, N. Nagaosa, and X.-G. Wen, *Rev. Mod. Phys.* **78**, 17 (2006).
- [30] J. Orenstein and A. Millis, *Science* **288**, 468 (2000).
- [31] C. Y. Johnson, *A superconductor claim blew up online. Science has punctured it.* (The Washington Post, 2023).
- [32] E. Cartlidge, *Nature* **524**, 277 (2015).
- [33] C. Jin and D. Ceperley, *Nature* **615**, 221 (2023).
- [34] W. Little, *Sci. Am.* **212**, 21 (1965).
- [35] A. W. Sleight, *Acc. Chem. Res.* **28**, 103 (1995).
- [36] B. Matthias and J. Huhl, *Phys. Rev.* **87**, 799 (1952).
- [37] K. A. Müller and J. G. Bednorz, *Science* **237**, 1133 (1987).
- [38] J. G. Bednorz and K. A. Müller, *Z. Phys. B* **64**, 189 (1986).
- [39] J. G. Bednorz and K. A. Müller, *Rev. Mod. Phys.* **60**, 585 (1988).
- [40] C. Chu, P. Hor, R. Meng, L. Gao, Z. Huang, Wang, and YQ, *Phys. Rev. Lett.* **58**, 405 (1987).
- [41] R. J. Cava, R. B. van Dover, B. Batlogg, and E. A. Rietman, *Phys. Rev. Lett.* **58**, 408 (1987).
- [42] M. K. Wu, J. R. Ashburn, C. J. Torng, P. H. Hor, R. L. Meng, L. Gao, Z. J. Huang, Y. Q. Wang, and C. W. Chu, *Phys. Rev. Lett.* **58**, 908 (1987).
- [43] R. J. Cava, B. Batlogg, R. B. van Dover, D. W. Murphy, S. Sunshine, T. Siegrist, J. P. Remeika, E. A. Rietman, S. Zahurak, and G. P. Espinosa, *Phys. Rev. Lett.* **58**, 1676 (1987).
- [44] R. M. Hazen, C. T. Prewitt, R. J. Angel, N. L. Ross, L. W. Finger, C. G. Hadjidakos, D. R. Veblen, P. J. Heaney, P. H. Hor, R. L. Meng, Y. Y. Sun, Y. Q. Wang, Y. Y. Xue, Z. J. Huang, L. Gao, J. Bechtold, and C. W. Chu, *Phys. Rev. Lett.* **60**, 1174 (1988).
- [45] A. Schilling, M. Cantoni, J. Guo, and H. Ott, *Nature* **363**, 56 (1993).
- [46] C. Chu, L. Gao, F. Chen, Z. Huang, R. Meng, and Y. Xue, *Nature* **365**, 323 (1993).
- [47] J. Nagamatsu, N. Nakagawa, T. Muranaka, Y. Zenitani, and J. Akimitsu, *Nature* **410**, 63 (2001).
- [48] Y. Kamihara, H. Hiramatsu, M. Hirano, R. Kawamura, H. Yanagi, T. Kamiya, and H. Hosono, *J. Am. Chem. Soc.* **128**, 10012 (2006).
- [49] Y. Kamihara, T. Watanabe, M. Hirano, and H. Hosono, *J. Am. Chem. Soc.* **130**, 3296 (2008).
- [50] H. Hosono, K. Tanabe, E. Takayama-Muromachi, H. Kageyama, S. Yamanaka, H. Kumakura, M. Nohara, H. Hiramatsu, and S. Fujitsu, *Sci. Technol. Adv. Mater.* **16**, 033503 (2015).
- [51] H. Hosono and K. Kuroki, *Physica C* **514**, 399 (2015).
- [52] J. Hirsch, *Appl. Phys. Lett.* **121**, 080501 (2022).
- [53] W. Kohn and J. Luttinger, *Phys. Rev. Lett.* **15**, 524 (1965).
- [54] N. Mott, *J. Phys.: Condens. Matter* **5**, 3487 (1993).
- [55] Y.-H. Chen, F. Wilczek, E. Witten, and B. I. Halperin, *Int. J. Mod. Phys. B* **3**, 1001 (1989).
- [56] C. Beenakker, *Annu. Rev. Condens. Matter Phys.* **4**, 113 (2013).
- [57] M. Leijnse and K. Flensberg, *Semicond. Sci. Technol.* **27**, 124003 (2012).
- [58] D. Scalapino, *J. Low Temp. Phys.* **117**, 179 (1999).
- [59] P. W. Anderson, *Science* **235**, 1196 (1987).
- [60] C. Petrovic, P. Pagliuso, M. Hundley, R. Movshovich, J. Sarrao, J. Thompson, Z. Fisk, and P. Monthoux, *J. Condens. Matter Phys.* **13**, L337 (2001).
- [61] F. Steglich, C. Bredl, W. Lieke, U. Rauchschwalbe, and G. Sporn, *Physica B+C* **126**, 82 (1984).
- [62] D. Jérôme, *Science* **252**, 1509 (1991).
- [63] G. Saito and Y. Yoshida, *Chem. Rec.* **11**, 124 (2011).
- [64] A. Ardavan, S. Brown, S. Kagoshima, K. Kanoda, K. Kuroki, H. Mori, M. Ogata, S. Uji, and J. Wosnitzer, *J. Phys. Soc. Jpn.* **81**, 011004 (2011).
- [65] A. G. Lebed, *The physics of organic superconductors and conductors*, Vol. 110 (Springer, 2008).
- [66] A. R. Oganov, ed., *Modern Methods of Crystal Structure Prediction* (Wiley-VCH, Weinheim, Germany, 2011).
- [67] A. R. Oganov, C. J. Pickard, Q. Zhu, and R. J. Needs, *Nat. Rev. Mater.* **4**, 331 (2019).
- [68] R. J. Needs and C. J. Pickard, *APL Mater.* **4**, 053210 (2016).
- [69] L. Zhang, Y. Wang, J. Lv, and Y. Ma, *Nat. Rev. Mater.* **2**, 17005 (2017).
- [70] M. Xu, Y. Li, and Y. Ma, *Chem. Sci.* **13**, 329 (2022).
- [71] S. Baroni, S. de Gironcoli, and A. Dal Corso, *Rev. Mod. Phys.* **73**, 515 (2001).
- [72] F. Giustino, *Rev. Mod. Phys.* **89**, 015003 (2017).
- [73] W. E. Pickett, *Rev. Mod. Phys.* **95**, 021001 (2023).
- [74] E. Zurek and T. Bi, *J. Chem. Phys.* **150**, 050901 (2019).
- [75] Y. Sun, X. Zhong, H. Liu, and Y. Ma, *Natl. Sci. Rev.*, nwad270 (2023).
- [76] K. P. Hilleke and E. Zurek, *J. Appl. Phys.* **131**, 070901 (2022).
- [77] G. Gao, L. Wang, M. Li, J. Zhang, R. T. Howie, E. Gregoryanz, V. V. Struzhkin, L. Wang, and S. T. John, *Mater. Today Phys.* **21**, 100546 (2021).
- [78] W. Zhao, X. Huang, Z. Zhang, S. Chen, M. Du, D. Duan, and T. Cui, *Natl. Sci. Rev.*, nwad307 (2023).
- [79] L. Boeri and G. B. Bachelet, *J. Condens. Matter Phys.* **31**, 234002 (2019).
- [80] C. J. Pickard, I. Errea, and M. I. Erements, *Annu. Rev. Condens. Matter Phys.* **11**, 57 (2020).
- [81] H. Wang, J. S. Tse, K. Tanaka, T. Iitaka, and Y. Ma, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 6463 (2012).
- [82] 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).
- [83] I. Errea, F. Belli, L. Monacelli, A. Sanna, T. Koretsune, T. Tadano, R. Bianco, M. Calandra, R. Arita, F. Mauri, and J. A. Flores-Livas, *Nature* **578**, 66 (2020).
- [84] S. Shah and A. N. Kolmogorov, *Phys. Rev. B* **88**, 014107 (2013).
- [85] T. D. Huan, *Phys. Rev. Mater.* **2**, 023803 (2018).
- [86] Y. Li, J. Hao, H. Liu, J. S. Tse, Y. Wang, and Y. Ma, *Sci. Rep.* **5**, 9948 (2015).
- [87] H. Liu, I. I. Naumov, R. Hoffmann, N. Ashcroft, and R. J. Hemley, *Proc. Natl. Acad. Sci. USA* **114**, 6990 (2017).
- [88] F. Peng, Y. Sun, C. J. Pickard, R. J. Needs, Q. Wu, and Y. Ma, *Phys. Rev. Lett.* **119**, 107001 (2017).
- [89] A. G. Kvashnin, D. V. Semenov, I. A. Kruglov, I. A. Wrona, and A. R. Oganov, *ACS Appl. Mater. Interfaces* **10**, 43809 (2018).
- [90] Z. Zhang, T. Cui, M. J. Hutcheon, A. M. Shipley, H. Song, M. Du, V. Z. Kresin, D. Duan, C. J. Pickard, and Y. Yao, *Phys. Rev. Lett.* **128**, 047001 (2022).
- [91] L. Boeri, R. G. Hennig, P. J. Hirschfeld, G. Profeta, A. Sanna, E. Zurek, W. E. Pickett, M. Amsler, R. Dias, M. Erements, C. Heil, R. Hemley, H. Liu, Y. Ma, C. Pierleoni, A. Kolmogorov, N. Rybin, D. Novoselov, V. I.

- Anisimov, A. R. Oganov, C. J. Pickard, T. Bi, R. Arita, I. Errea, C. Pellegrini, R. Requist, E. Gross, E. R. Margine, S. R. Xie, Y. Quan, A. Hire, L. Fanfarillo, G. R. Stewart, J. J. Hamlin, V. Stanev, R. S. Gonnelli, E. Piatti, D. Romanin, D. Daghero, and R. Valenti, *J. Phys. Condens. Matter* **34**, 183002 (2021).
- [92] M. Yazdani-Asrami, A. Sadeghi, W. Song, A. Madureira, J. Pina, A. Morandi, and M. Parizh, *Supercond. Sci. Technol.* **35**, 123001 (2022).
- [93] F. J. Ribeiro and M. L. Cohen, *Phys. Rev. B* **69**, 212507 (2004).
- [94] R. Saniz, J. E. Medvedeva, L.-H. Ye, T. Shishidou, and A. J. Freeman, *Phys. Rev. B* **70**, 100505 (2004).
- [95] S. Di Cataldo, C. Heil, W. von der Linden, and L. Boeri, *Phys. Rev. B* **104**, L020511 (2021).
- [96] S. Zhang, H. Yu, J. Wei, T. Zhong, J. Sun, Q. Wang, L. Liu, H. Xu, J. Ma, and H. Liu, *Phys. Rev. B* **109**, 174507 (2024).
- [97] A. Sanna, T. F. Cerqueira, Y.-W. Fang, I. Errea, A. Ludwig, and M. A. Marques, *npj Comput. Mater.* **10**, 44 (2024).
- [98] M. Eremets, I. Trojan, S. Medvedev, J. Tse, and Y. Yao, *Science* **319**, 1506 (2008).
- [99] A. Drozdov, M. Eremets, I. Troyan, V. Ksenofontov, and S. Shylin, *Nature* **525**, 73 (2015).
- [100] 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).
- [101] A. P. Drozdov, P. P. Kong, V. S. Minkov, S. P. Besedin, M. A. Kuzovnikov, S. Mozaffari, L. Balicas, F. F. Balakirev, D. E. Graf, V. B. Prakapenka, E. Greenberg, D. A. Knyazev, T. M., and M. I. Eremets, *Nature* **569**, 528 (2019).
- [102] D. V. Semenok, A. G. Kvashnin, A. G. Ivanova, V. Svitlyk, V. Y. Fominski, A. V. Sadakov, O. A. Sobolevskiy, V. M. Pudalov, I. A. Troyan, and A. R. Oganov, *Mater. Today* **33**, 36 (2020).
- [103] W. Chen, D. V. Semenok, A. G. Kvashnin, X. Huang, I. A. Kruglov, M. Galasso, H. Song, D. Duan, A. F. Goncharov, V. B. Prakapenka, A. R. Oganov, and T. Cui, *Nat. Commun.* **12**, 273 (2021).
- [104] I. A. Troyan, D. V. Semenok, A. G. Kvashnin, A. V. Sadakov, O. A. Sobolevskiy, V. M. Pudalov, A. G. Ivanova, V. B. Prakapenka, E. Greenberg, A. G. Gavriliuk, I. S. Lyubutin, V. V. Struzhkin, A. Bergara, I. Errea, R. Bianco, M. Calandra, F. Mauri, L. Monacelli, R. Akashi, and A. R. Oganov, *Adv. Mater.* **33**, 2006832 (2021).
- [105] P. Kong, V. S. Minkov, M. A. Kuzovnikov, A. P. Drozdov, S. P. Besedin, S. Mozaffari, L. Balicas, F. F. Balakirev, V. B. Prakapenka, S. Chariton, D. A. Knyazev, E. Greenberg, and M. I. Eremets, *Nat. Commun.* **12**, 5075 (2021).
- [106] W. Chen, D. V. Semenok, X. Huang, H. Shu, X. Li, D. Duan, T. Cui, and A. R. Oganov, *Phys. Rev. Lett.* **127**, 117001 (2021).
- [107] L. Ma, K. Wang, Y. Xie, X. Yang, Y. Wang, M. Zhou, H. Liu, X. Yu, Y. Zhao, H. Wang, G. Liu, and Y. Ma, *Phys. Rev. Lett.* **128**, 167001 (2022).
- [108] Z. Li, X. He, C. Zhang, X. Wang, S. Zhang, Y. Jia, S. Feng, K. Lu, J. Zhao, J. Zhang, B. Min, Y. Long, R. Yu, L. Wang, M. Ye, Z. Zhang, V. Prakapenka, S. Chariton, P. A. Ginsberg, J. Bass, S. Yuan, H. Liu, and C. Jin, *Nat. Commun.* **13**, 2863 (2022).
- [109] Y. Song, J. Bi, Y. Nakamoto, K. Shimizu, H. Liu, B. Zou, G. Liu, H. Wang, and Y. Ma, *Phys. Rev. Lett.* **130**, 266001 (2023).
- [110] J. Hirsch and F. Marsiglio, *Nature* **596**, E9 (2021).
- [111] T. Wang, M. Hirayama, T. Nomoto, T. Koretsune, R. Arita, and J. A. Flores-Livas, *Phys. Rev. B* **104**, 064510 (2021).
- [112] J. Hirsch and F. Marsiglio, *Phys. Rev. B* **103**, 134505 (2021).
- [113] M. Gubler, J. A. Flores-Livas, A. Kozhevnikov, and S. Goedecker, *Phys. Rev. Mater.* **6**, 014801 (2022).
- [114] M. Eremets, V. Minkov, A. Drozdov, P. Kong, V. Ksenofontov, S. Shylin, R. Prozorov, F. Balakirev, D. Sun, S. Mozaffari, and L. Balicas, *J. Supercond. Nov. Magn.* **35**, 965 (2022).
- [115] X. Xing, C. Wang, L. Yu, J. Xu, C. Zhang, M. Zhang, S. Huang, X. Zhang, Y. Liu, B. Yang, X. Chen, Y. Zhang, J. Guo, Z. Shi, Y. Ma, C. Chen, and X. Liu, *Nat. Commun.* **14**, 5991 (2023).
- [116] S. Chen, *Physics* **16**, 39 (2023).
- [117] N. W. Ashcroft, *Phys. Rev. Lett.* **92**, 187002 (2004).
- [118] N. W. Ashcroft, *J. Phys. Condens. Matter* **16**, S945 (2004).
- [119] R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakithodi, and C. Kim, *npj Comput. Mater.* **3**, 54 (2017).
- [120] A. Agrawal and A. Choudhary, *APL Mater.* **4**, 053208 (2016).
- [121] A. Jain, G. Hautier, S. P. Ong, and K. Persson, *J. Mater. Res.* **31**, 977 (2016).
- [122] C. Draxl and M. Scheffler, *MRS Bull.* **43**, 676 (2018).
- [123] C. Draxl and M. Scheffler, *J. Phys.: Mater.* **2**, 036001 (2019).
- [124] C. Draxl and M. Scheffler, *Handbook of Materials Modeling: Methods: Theory and Modeling*, 49 (2020).
- [125] L. M. Ghiringhelli, J. Vybiral, S. V. Levchenko, C. Draxl, and M. Scheffler, *Phys. Rev. Lett.* **114**, 105503 (2015).
- [126] J. Damewood, J. Karagunesian, J. R. Langer, A. R. Tan, M. Xie, J. Peng, and R. Gómez-Bombarelli, *Annu. Rev. Mater. Res.* **53**, 399 (2023).
- [127] Q. Gu, D. Xing, and J. Sun, *Chin. Phys. Lett.* **36**, 097401 (2019).
- [128] Q. Yang, J. Lv, Q. Tong, X. Du, Y. Wang, S. Zhang, G. Yang, A. Bergara, and Y. Ma, *Phys. Rev. B* **103**, 024505 (2021).
- [129] K. Dolui, L. J. Conway, C. Heil, T. A. Strobel, R. P. Prasankumar, and C. J. Pickard, *Phys. Rev. Lett.* **132**, 166001 (2024).
- [130] T. Ishikawa, Y. Tanaka, and S. Tsuneyuki, *Phys. Rev. B* **109**, 094106 (2024).
- [131] P. P. Ferreira, L. J. Conway, A. Cucciari, S. Di Cataldo, F. Giannessi, E. Kogler, L. T. Eleno, C. J. Pickard, C. Heil, and L. Boeri, *Nat. Commun.* **14**, 5367 (2023).
- [132] R. Lucrezi, E. Kogler, S. Di Cataldo, M. Aichhorn, L. Boeri, and C. Heil, *Commun. Phys.* **6**, 298 (2023).
- [133] S. Xie, G. Stewart, J. Hamlin, P. Hirschfeld, and R. Hennig, *Phys. Rev. B* **100**, 174513 (2019).
- [134] S. Xie, Y. Quan, A. Hire, B. Deng, J. DeStefano, I. Salinas, U. Shah, L. Fanfarillo, J. Lim, J. Kim, G. R. Stewart, J. J. Hamlin, P. J. Hirschfeld, and R. G. Hennig, *npj Comput. Mater.* **8**, 14 (2022).
- [135] O. Isayev, D. Fourches, E. N. Muratov, C. Oses, K. Rasch, A. Tropsha, and S. Curtarolo, *Chem. Mater.*

- 27**, 735 (2015).
- [136] K. Hamidieh, *Comput. Mater. Sci.* **154**, 346 (2018).
 - [137] V. Stanev, C. Oses, A. G. Kusne, E. Rodriguez, J. Paglione, S. Curtarolo, and I. Takeuchi, *npj Comput. Mater.* **4**, 29 (2018).
 - [138] S. Zeng, Y. Zhao, G. Li, R. Wang, X. Wang, and J. Ni, *npj Comput. Mater.* **5**, 84 (2019).
 - [139] T. Konno, H. Kurokawa, F. Nabeshima, Y. Sakishita, R. Ogawa, I. Hosako, and A. Maeda, *Phys. Rev. B* **103**, 014509 (2021).
 - [140] K. Choudhary and K. Garrity, *npj Comput. Mater.* **8**, 244 (2022).
 - [141] T. F. Cerqueira, A. Sanna, and M. A. Marques, *Adv. Mater.* **36**, 2307085 (2024).
 - [142] H. Tran and T. N. Vu, *Phys. Rev. Mater.* **7**, 054805 (2023).
 - [143] J. Zhang, Z. Zhu, X.-D. Xiang, K. Zhang, S. Huang, C. Zhong, H.-J. Qiu, K. Hu, and X. Lin, *J. Phys. Chem. C* **126**, 8922 (2022).
 - [144] K. Matsumoto and T. Horide, *Appl. Phys. Express* **12**, 073003 (2019).
 - [145] C. Pereti, K. Bernot, T. Guizouarn, F. Laufek, A. Vymazalová, L. Bindi, R. Sessoli, and D. Fanelli, *npj Comput. Mater.* **9**, 71 (2023).
 - [146] T. Ishikawa, T. Miyake, and K. Shimizu, *Phys. Rev. B* **100**, 174506 (2019).
 - [147] M. J. Hutcheon, A. M. Shipley, and R. J. Needs, *Phys. Rev. B* **101**, 144505 (2020).
 - [148] A. M. Shipley, M. J. Hutcheon, R. J. Needs, and C. J. Pickard, *Phys. Rev. B* **104**, 054501 (2021).
 - [149] D. Wines, K. Choudhary, A. J. Bicchieri, K. F. Garrity, and F. Tavazza, *Nano Lett.* **23**, 969 (2023).
 - [150] Y.-L. Hai, M.-J. Jiang, H.-L. Tian, G.-H. Zhong, W.-J. Li, C.-L. Yang, X.-J. Chen, and H.-Q. Lin, *Adv. Sci.* **10**, 2303639 (2023).
 - [151] P. Song, Z. Hou, P. B. de Castro, K. Nakano, K. Hongo, Y. Takano, and R. Maezono, *arXiv preprint arXiv:2103.00193* (2021).
 - [152] T. D. Le, R. Noumeir, H. L. Quach, J. H. Kim, J. H. Kim, and H. M. Kim, *IEEE Trans. Appl. Supercond.* **30**, 1 (2020).
 - [153] P. J. García-Nieto, E. García-Gonzalo, and J. P. Paredes-Sánchez, *Neural. Comput. Appl.* **33**, 17131 (2021).
 - [154] V. Stanev, K. Choudhary, A. G. Kusne, J. Paglione, and I. Takeuchi, *Commun. Mater.* **2**, 1 (2021).
 - [155] S. Raviprasad, N. A. Angadi, and M. Kothari, in *2022 3rd International Conference for Emerging Technology (INCET)* (IEEE, 2022) pp. 1–5.
 - [156] G. Revathy, V. Rajendran, B. Rashmika, P. S. Kumar, P. Parkavi, and J. Shynisha, *Materials Today: Proceedings* (2022).
 - [157] Z. Bai, M. Bhullar, A. Akinpelu, and Y. Yao, *Mater. Today Phys.* **43**, 101384 (2024).
 - [158] B. Roter and S. V. Dordevic, *Physica C: Supercond. Appl.* **575**, 1353689 (2020).
 - [159] B. Roter, N. Ninkovic, and S. V. Dordevic, *Physica C: Supercond. Appl.* **598**, 1354078 (2022).
 - [160] K. Yamaguchi, R. Asahi, and Y. Sasaki, in *Proceedings of the Twelfth Language Resources and Evaluation Conference* (2020) pp. 6753–6760.
 - [161] K. Yamaguchi, R. Asahi, and Y. Sasaki, *Adv. Eng. Inform.* **54**, 101768 (2022).
 - [162] L. Foppiano, T. M. Dieb, A. Suzuki, and M. Ishii, in *IEICE Tech. Rep.*, Vol. 119 (2019) pp. 1–5.
 - [163] L. Foppiano, S. Dieb, A. Suzuki, P. Baptista de Castro, S. Iwasaki, A. Uzuki, M. G. E. Echevarria, Y. Meng, K. Terashima, L. Romary, Y. Takano, and M. Ishii, *Sci. Technol. Adv. Mater.: Methods* **1**, 34 (2021).
 - [164] L. Foppiano, P. B. Castro, P. Ortiz Suarez, K. Terashima, Y. Takano, and M. Ishii, *Sci. Technol. Adv. Mater.: Methods* **3**, 2153633 (2023).
 - [165] T. Sommer, R. Willa, J. Schmalian, and P. Friederich, *Sci. Data* **10**, 816 (2023).
 - [166] E. R. Margine and F. Giustino, *Phys. Rev. B* **87**, 024505 (2013).
 - [167] S. Poncé, E. R. Margine, C. Verdi, and F. Giustino, *Comput. Phys. Commun.* **209**, 116 (2016).
 - [168] F. Giustino, M. L. Cohen, and S. G. Louie, *Phys. Rev. B* **76**, 165108 (2007).
 - [169] P. Morel and P. W. Anderson, *Phys. Rev.* **125**, 1263 (1962).
 - [170] W. L. McMillan, *Phys. Rev.* **167**, 331 (1968).
 - [171] R. Dynes, *Solid State Commun.* **10**, 615 (1972).
 - [172] P. B. Allen and R. C. Dynes, *Phys. Rev. B* **12**, 905 (1975).
 - [173] I. Esterlis, B. Nosarzewski, E. W. Huang, B. Moritz, T. P. Devereaux, D. J. Scalapino, and S. A. Kivelson, *Phys. Rev. B* **97**, 140501 (2018).
 - [174] A. Alexandrov, *Europhys. Lett.* **56**, 92 (2001).
 - [175] J. Hague and N. d’Ambrumenil, *J. Low. Temp. Phys.* **151**, 1149 (2008).
 - [176] P. Benedetti and R. Zeyher, *Phys. Rev. B* **58**, 14320 (1998).
 - [177] E. A. Yuzbashyan and B. L. Altshuler, *Phys. Rev. B* **106**, 054518 (2022).
 - [178] R. Dynes, A. White, J. Graybeal, and J. Garno, *Phys. Rev. Lett.* **57**, 2195 (1986).
 - [179] P. Hohenberg and W. Kohn, *Phys. Rev.* **136**, B864 (1964).
 - [180] W. Kohn and L. Sham, *Phys. Rev.* **140**, A1133 (1965).
 - [181] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. D. Corso, S. d. Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, *J. Phys.: Condens. Matter* **21**, 395502 (2009).
 - [182] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. D. Corso, S. de Gironcoli, P. Delugas, R. A. DiStasio, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia, M. Kawamura, H.-Y. Ko, A. Kokalj, E. Küçükbenli, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N. L. Nguyen, H.-V. Nguyen, A. O. de-la Roza, L. Paulatto, S. Poncé, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A. P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu, and S. Baroni, *J. Phys.: Condens. Matter* **29**, 465901 (2017).
 - [183] X. Gonze, B. Amadon, P.-M. Anglade, J.-M. Beuken, F. Bottin, P. Boulanger, F. Bruneval, D. Caliste,

- R. Caracas, M. Côté, T. Deutsch, L. Genovese, P. Ghosez, M. Giantomassi, S. Goedecker, D. Hamann, P. Hermet, F. Jollet, G. Jomard, S. Leroux, M. Mancini, S. Mazevet, M. Oliveira, G. Onida, Y. Pouillon, T. Rangel, G.-M. Rignanese, D. Sangalli, R. Shaltaf, M. Torrent, M. Verstraete, G. Zerah, and J. Zwanziger, *Comput. Phys. Commun.* **180**, 2582 (2009).
- [184] X. Gonze, G. M. Rignanese, M. Verstraete, J.-M. Beuken, Y. Pouillon, R. Caracas, F. Jollet, M. Torrent, G. Zerah, M. Mikami, P. Ghosez, M. Veithen, J.-Y. Raty, V. Olevano, F. Bruneval, L. Reining, R. Godby, G. Onida, D. R. Hamann, and D. C. Allan, *Zeit. Kristallogr.* **220**, 558 (2005).
- [185] X. Gonze, F. Jollet, F. A. Araujo, D. Adams, B. Amadon, T. Applencourt, C. Audouze, J.-M. Beuken, J. Bieder, A. Bokhanchuk, E. Bousquet, F. Bruneval, D. Caliste, M. Côté, F. Dahm, F. D. Pieve, M. Delaveau, M. D. Gennaro, B. Dorado, C. Espejo, G. Geneste, L. Genovese, A. Gerossier, M. Giantomassi, Y. Gillet, D. Hamann, L. He, G. Jomard, J. L. Janssen, S. L. Roux, A. Levitt, A. Lherbier, F. Liu, I. Lukačević, A. Martin, C. Martins, M. Oliveira, S. Poncé, Y. Pouillon, T. Rangel, G.-M. Rignanese, A. Romero, B. Rousseau, O. Rubel, A. Shukri, M. Stankovski, M. Torrent, M. V. Setten, B. V. Troeye, M. Verstraete, D. Waroquiers, J. Wiktoria, B. Xu, A. Zhou, and J. Zwanziger, *Comput. Phys. Commun.* **205**, 106 (2016).
- [186] F. Marsiglio, M. Schossmann, and J. Carbotte, *Phys. Rev. B* **37**, 4965 (1988).
- [187] G. Marini, G. Marchese, G. Profeta, J. Sjakste, F. Macheda, N. Vast, F. Mauri, and M. Calandra, *Comput. Phys. Commun.* **295**, 108950 (2024).
- [188] E. Margine, H. Lambert, and F. Giustino, *Sci. Rep.* **6**, 21414 (2016).
- [189] G. Kuderowicz and B. Wiendlocha, *Phys. Rev. B* **108**, 224501 (2023).
- [190] 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).
- [191] R. Akashi, M. Kawamura, S. Tsuneyuki, Y. Nomura, and R. Arita, *Phys. Rev. B* **91**, 224513 (2015).
- [192] K. Lejaeghere, G. Bihlmayer, T. Björkman, P. Blaha, S. Blügel, V. Blum, D. Caliste, I. E. Castelli, S. J. Clark, A. D. Corso, S. de Gironcoli, T. Deutsch, J. K. Dewhurst, I. D. Marco, C. Draxl, M. Dułak, O. Eriksson, J. A. Flores-Livas, K. F. Garrity, L. Genovese, P. Giannozzi, M. Giantomassi, S. Goedecker, X. Gonze, O. Grånäs, E. K. U. Gross, A. Gulans, F. Gygi, D. R. Hamann, P. J. Hasnip, N. A. W. Holzwarth, D. Iuşan, D. B. Jochym, F. Jollet, D. Jones, G. Kresse, K. Koepernik, E. Küçükbenli, Y. O. Kvashnin, I. L. M. Locht, S. Lubeck, M. Marsman, N. Marzari, U. Nitzsche, L. Nordström, T. Ozaki, L. Paulatto, C. J. Pickard, W. Poelmans, M. I. J. Probert, K. Refson, M. Richter, G.-M. Rignanese, S. Saha, M. Scheffler, M. Schlipf, K. Schwarz, S. Sharma, F. Tavazza, P. Thunström, A. Tkatchenko, M. Torrent, D. Vanderbilt, M. J. van Setten, V. V. Speybroeck, J. M. Wills, J. R. Yates, G.-X. Zhang, and S. Cottenier, *Science* **351**, aad3000 (2016).
- [193] A. Franceschetti and A. Zunger, *Nature* **402**, 60 (1999).
- [194] B. Sanchez-Lengeling and A. Aspuru-Guzik, *Science* **361**, 360 (2018).
- [195] R. Gómez-Bombarelli, J. N. Wei, D. Duvenaud, J. M. Hernández-Lobato, B. Sánchez-Lengeling, D. Sheberla, J. Aguilera-Iparraguirre, T. D. Hirzel, R. P. Adams, and A. Aspuru-Guzik, *ACS Cent. Sci.* **4**, 268 (2018).
- [196] T. Weymuth and M. Reiher, *Int. J. Quantum Chem.* **114**, 823 (2014).
- [197] M. d’Avezac, J.-W. Luo, T. Chanier, and A. Zunger, *Phys. Rev. Lett.* **108**, 027401 (2012).
- [198] H. J. Xiang, B. Huang, E. Kan, S.-H. Wei, and X. G. Gong, *Phys. Rev. Lett.* **110**, 118702 (2013).
- [199] T. Lombardo, M. Duquesnoy, H. El-Bouysidy, F. Årén, A. Gallo-Bueno, P. B. Jørgensen, A. Bhowmik, A. Demortière, E. Ayerbe, F. Alcaide, M. Reynaud, J. Carrasco, A. Grimaud, C. Zhang, T. Vegge, P. Johansson, and A. A. Franco, *Chem. Rev.* **122**, 10899 (2021).
- [200] D. P. Tabor, L. M. Roch, S. K. Saikin, C. Kreisbeck, D. Sheberla, J. H. Montoya, S. Dwaraknath, M. Aykol, C. Ortiz, H. Tribukait, C. Amador-Bedolla, C. J. Brabec, B. Maruyama, K. A. Persson, and A. Aspuru-Guzik, *Nat. Rev. Mater.* **3**, 5 (2018).
- [201] H. Tran, R. Gurnani, C. Kim, G. Pilania, H.-K. Kwon, R. Lively, and R. Ramprasad, *Nat. Rev. Mater.*, XXXX (2024).
- [202] M. Hu, Q. Tan, R. Knibbe, M. Xu, B. Jiang, S. Wang, X. Li, and M.-X. Zhang, *Mater. Sci. Eng. R Rep.* **155**, 100746 (2023).
- [203] X. Liu, J. Zhang, and Z. Pei, *Prog. Mater. Sci.* **131**, 101018 (2023).
- [204] A. Mannodi-Kanakkithodi, G. Pilania, T. D. Huan, T. Lookman, and R. Ramprasad, *Sci. Rep.* **6**, 20952 (2016).
- [205] C. Kim, R. Batra, L. Chen, H. Tran, and R. Ramprasad, *Comput. Mater. Sci.*, 110067 (2021).
- [206] R. Batra, H. Dai, T. D. Huan, L. Chen, C. Kim, W. R. Gutekunst, L. Song, and R. Ramprasad, *Chem. Matter.* **32**, 10489 (2020).
- [207] T. D. Huan, V. Sharma, G. A. Rossetti, and R. Ramprasad, *Phys. Rev. B* **90**, 064111 (2014).
- [208] V. N. Tuoc and T. D. Huan, *J. Phys. Chem. C* **122**, 17067 (2018).
- [209] A. Togo, F. Oba, and I. Tanaka, *Phys. Rev. B* **78**, 134106 (2008).
- [210] L. Monacelli, R. Bianco, M. Cherubini, M. Calandra, I. Errea, and F. Mauri, *J. Phys.: Condens. Matter.* **33**, 363001 (2021).
- [211] D. V. Semenov, I. A. Troyan, A. G. Ivanova, A. G. Kvashnin, I. A. Kruglov, M. Hanfland, A. V. Sadakov, O. A. Sobolevskiy, K. S. Pervakov, I. S. Lyubutin, K. V. Glazyrin, N. Giordano, D. N. Karimov, A. L. Vasiliev, R. Akashi, V. M. Pudalov, and A. R. Oganov, *Mater. Today* **48**, 18 (2021).
- [212] F. H. Stillinger, *Phys. Rev. E* **59**, 48 (1999).
- [213] M. Amsler, J. A. Flores-Livas, T. D. Huan, S. Botti, M. A. L. Marques, and S. Goedecker, *Phys. Rev. Lett.* **108**, 205505 (2012).
- [214] T. D. Huan, M. Amsler, M. A. L. Marques, S. Botti, A. Willand, and S. Goedecker, *Phys. Rev. Lett.* **110**, 135502 (2013).
- [215] T. D. Huan, V. N. Tuoc, and N. V. Minh, *Phys. Rev. B* **93**, 094105 (2016).
- [216] X. Huang, D. Duan, X. Li, F. Li, Y. Huang, G. Wu, Y. Liu, Q. Zhou, B. Liu, and T. Cui, *Appl. Phys. Lett.* **107**, 041906 (2015).

- [217] Y. Wei, P. Nukala, M. Salverda, S. Matzen, H. J. Zhao, J. Momand, A. S. Everhardt, G. Agnus, G. R. Blake, P. Lecoeur, B. J. Kooi, J. /' Iñiguez, B. Dkhil, and B. Noheda, *Nat. Mater.* **17**, 1095 (2018).
- [218] Y. Qi, S. Singh, C. Lau, F.-T. Huang, X. Xu, F. J. Walker, C. H. Ahn, S.-W. Cheong, and K. M. Rabe, *Phys. Rev. Lett.* **125**, 257603 (2020).
- [219] C. W. Glass, A. R. Oganov, and N. Hansen, *Comput. Phys. Commun.* **175**, 713 (2006).
- [220] A. R. Oganov and C. W. Glass, *J. Chem. Phys.* **124**, 244704 (2006).
- [221] Y. Wang, J. Lv, L. Zhu, and Y. Ma, *Comput. Phys. Commun.* **183**, 2063 (2012).
- [222] C. J. Pickard and R. J. Needs, *J. Phys. Condens. Matter.* **23**, 053201 (2011).
- [223] D. C. Lonie and E. Zurek, *Comput. Phys. Commun.* **182**, 372 (2011).
- [224] S. Hajinazar, A. Thorn, E. D. Sandoval, S. Kharabadze, and A. N. Kolmogorov, *Comput. Phys. Commun.* **259**, 107679 (2021).
- [225] S. Goedecker, *J. Chem. Phys.* **120**, 9911 (2004).
- [226] M. Amsler and S. Goedecker, *J. Chem. Phys.* **133**, 224104 (2010).
- [227] Y. Sun, J. Lv, Y. Xie, H. Liu, and Y. Ma, *Phys. Rev. Lett.* **123**, 097001 (2019).
- [228] X. Zhong, Y. Sun, T. Iitaka, M. Xu, H. Liu, R. J. Hemley, C. Chen, and Y. Ma, *J. Am. Chem. Soc.* **144**, 13394 (2022).
- [229] H. Zhang, X. Jin, Y. Lv, Q. Zhuang, Y. Liu, Q. Lv, K. Bao, D. Li, B. Liu, and T. Cui, *Sci. Rep.* **5**, 8845 (2015).
- [230] J. Feng, W. Grochala, T. Jaroń, R. Hoffmann, A. Bergara, and N. Ashcroft, *Phys. Rev. Lett.* **96**, 017006 (2006).
- [231] C. J. Pickard and R. Needs, *Phys. Rev. Lett.* **97**, 045504 (2006).
- [232] Y. Yao, J. Tse, Y. Ma, and K. Tanaka, *Europhys. Lett.* **78**, 37003 (2007).
- [233] T. A. Strobel, P. Ganesh, M. Somayazulu, P. Kent, and R. J. Hemley, *Phys. Rev. Lett.* **107**, 255503 (2011).
- [234] X. Huang, X. Wang, D. Duan, B. Sundqvist, X. Li, Y. Huang, H. Yu, F. Li, Q. Zhou, B. Liu, and T. Cui, *Natl. Sci. Rev.* **6**, 713 (2019).
- [235] M. Einaga, M. Sakata, T. Ishikawa, K. Shimizu, M. I. Eremets, A. P. Drozdov, I. A. Troyan, N. Hirao, and Y. Ohishi, *Nat. Phys.* **12**, 835 (2016).
- [236] H. Nakao, M. Einaga, M. Sakata, M. Kitagaki, K. Shimizu, S. Kawaguchi, N. Hirao, and Y. Ohishi, *J. Phys. Soc. Jpn.* **88**, 123701 (2019).
- [237] I. Osmond, O. Moulding, S. Cross, T. Muramatsu, A. Brooks, O. Lord, T. Fedotenko, J. Buhot, and S. Friedemann, *Phys. Rev. B* **105**, L220502 (2022).
- [238] V. Struzhkin, B. Li, C. Ji, X.-J. Chen, V. Prakapenka, E. Greenberg, I. Troyan, A. Gavriluk, and H.-k. Mao, *Matter Radiat. Extrem.* **5**, 028201 (2020).
- [239] D. Zhou, D. V. Semenov, D. Duan, H. Xie, W. Chen, X. Huang, X. Li, B. Liu, A. R. Oganov, and T. Cui, *Sci. Adv.* **6**, eaax6849 (2020).
- [240] Y. Wang, K. Wang, Y. Sun, L. Ma, Y. Wang, B. Zou, G. Liu, M. Zhou, and H. Wang, *Chin. Phys. B* **31**, 106201 (2022).
- [241] P. Bhattacharyya, W. Chen, X. Huang, S. Chatterjee, B. Huang, B. Kobrin, Y. Lyu, T. Smart, M. Block, E. Wang, Z. Wang, W. Wu, S. Hsieh, H. Ma, S. Mandyam, B. Chen, E. Davis, Z. M. Geballe, C. Zu, V. Struzhkin, R. Jeanloz, J. E. Moore, T. Cui, G. Galli, B. I. Halperin, C. R. Laumann, and N. Y. Yao, *Nature*, 73 (2024).
- [242] M. Shao, S. Chen, W. Chen, K. Zhang, X. Huang, and T. Cui, *Inorg. Chem.* **60**, 15330 (2021).
- [243] D. Y. Kim, R. H. Scheicher, H.-k. Mao, T. W. Kang, and R. Ahuja, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 2793 (2010).
- [244] Y.-K. Wei, J.-N. Yuan, F. I. Khan, G.-F. Ji, Z.-W. Gu, and D.-Q. Wei, *RSC Adv.* **6**, 81534 (2016).
- [245] I. A. Troyan, D. V. Semenov, A. G. Ivanova, A. V. Sadakov, D. Zhou, A. G. Kvashnin, I. A. Kruglov, O. A. Sobolevskiy, M. V. Lyubutina, D. S. Perekalin, T. Helm, S. W. Tozer, M. Bykov, A. F. Goncharov, V. M. Pudalov, and I. S. Lyubutin, *Adv. Sci.* **10**, 2303622 (2023).
- [246] X. He, C. Zhang, Z. Li, K. Lu, S. Zhang, B. Min, J. Zhang, L. Shi, S. Feng, Q. Liu, J. Song, X. Wang, Y. Peng, L. Wang, V. Prakapenka, S. Chariton, H. Liu, and C. Jin, *Mater. Today Phys.* **40**, 101298 (2024).
- [247] G. Huang, D. Peng, T. Luo, L.-C. Chen, P. Dalladay-Simpson, Z.-Y. Cao, F. A. Gorelli, G.-H. Zhong, H.-Q. Lin, and X.-J. Chen, *J. Phys.: Condens. Matter* **36**, 075702 (2023).
- [248] L.-C. Chen, T. Luo, Z.-Y. Cao, P. Dalladay-Simpson, G. Huang, D. Peng, L.-L. Zhang, F. A. Gorelli, G.-H. Zhong, H.-Q. Lin, and X.-J. Chen, *Nat. Commun.* **15**, 1809 (2024).
- [249] S. Cross, J. Buhot, A. Brooks, W. Thomas, A. Kleppe, O. Lord, and S. Friedemann, *Phys. Rev. B* **109**, L020503 (2024).
- [250] F. Smits, *Bell Syst. Tech. J.* **37**, 711 (1958).
- [251] I. Miccoli, F. Edler, H. Pfnür, and C. Tegenkamp, *J. Phys. Condens. Matter* **27**, 223201 (2015).
- [252] J. De Launay, *Phys. Rev.* **93**, 661 (1954).
- [253] J. W. Garland Jr, *Phys. Rev. Lett.* **11**, 114 (1963).
- [254] R. B. Goldfarb, M. Lelental, and C. Thompson, "Alternating-field susceptometry and magnetic susceptibility of superconductors," in *Magnetic Susceptibility of Superconductors and Other Spin Systems*, edited by R. A. Hein, T. L. Francavilla, and D. H. Liebenberg (Plenum Press, New York, 1991) Chap. 2, pp. 49–80.
- [255] V. V. Struzhkin and H.-k. Mao, *Natl. Sci. Rev.* **11**, nwae005 (2024).
- [256] I. Troyan, A. Gavriluk, R. Rüffer, A. Chumakov, A. Mironovich, I. Lyubutin, D. Perekalin, A. P. Drozdov, and M. I. Eremets, *Science* **351**, 1303 (2016).
- [257] K. O. Ho, K. C. Wong, M. Y. Leung, Y. Y. Pang, W. K. Leung, K. Y. Yip, W. Zhang, J. Xie, S. K. Goh, and S. Yang, *J. Appl. Phys.* **129**, 241101 (2021).
- [258] N. Werthamer, E. Helfand, and P. Hohenberg, *Phys. Rev.* **147**, 295 (1966).
- [259] V. Minkov, V. Ksenofontov, S. Bud'ko, E. Talantsev, and M. Eremets, *Nat. Phys.*, 1293 (2023).
- [260] N. W. Ashcroft, *Phys. Rev. Lett.* **21**, 1748 (1968).
- [261] X.-J. Chen, V. V. Struzhkin, Y. Song, A. F. Goncharov, M. Ahart, Z. Liu, H.-k. Mao, and R. J. Hemley, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 20 (2008).
- [262] T. A. Strobel, A. F. Goncharov, C. T. Seagle, Z. Liu, M. Somayazulu, V. V. Struzhkin, and R. J. Hemley, *Phys. Rev. B* **83**, 144102 (2011).
- [263] X.-J. Chen, J.-L. Wang, V. V. Struzhkin, H.-k. Mao, R. J. Hemley, and H.-Q. Lin, *Phys. Rev. Lett.* **101**, 077002 (2008).

- [264] M. Martinez-Canales, A. R. Oganov, Y. Ma, Y. Yan, A. O. Lyakhov, and A. Bergara, *Phys. Rev. Lett.* **102**, 087005 (2009).
- [265] X. Jin, X. Meng, Z. He, Y. Ma, B. Liu, T. Cui, G. Zou, and H.-k. Mao, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 9969 (2010).
- [266] J. A. Flores-Livas, M. Amsler, T. J. Lenosky, L. Lehtovaara, S. Botti, M. A. Marques, and S. Goedecker, *Phys. Rev. Lett.* **108**, 117004 (2012).
- [267] Y. Li, J. Hao, H. Liu, Y. Li, and Y. Ma, *J. Chem. Phys.* **140**, 174712 (2014).
- [268] A. Cho, “Stinky hydrogen sulfide smashes superconductivity record,” *Science* (2015-08-17).
- [269] Z. M. Geballe, H. Liu, A. K. Mishra, M. Ahart, M. Somayazulu, Y. Meng, M. Baldini, and R. J. Hemley, *Angew. Chem. Int. Ed.* **130**, 696 (2018).
- [270] H. Liu, I. I. Naumov, Z. M. Geballe, M. Somayazulu, S. T. John, and R. J. Hemley, *Phys. Rev. B* **98**, 100102 (2018).
- [271] D. Sun, V. S. Minkov, S. Mozaffari, Y. Sun, Y. Ma, S. Chariton, V. B. Prakapenka, M. I. Erements, L. Balicas, and F. F. Balakirev, *Nat. Commun.* **12**, 6863 (2021).
- [272] I. A. Kruglov, D. V. Semenok, H. Song, R. Szczesniak, I. A. Wrona, R. Akashi, M. M. D. Esfahani, D. Duan, T. Cui, and A. G. Kvashnin, *Phys. Rev. B* **101**, 024508 (2020).
- [273] P. Song, Z. Hou, K. Nakano, K. Hongo, and R. Maezono, *Mater. Today Phys.* **28**, 100873 (2022).
- [274] N. P. Salke, M. M. Davari Esfahani, Y. Zhang, I. A. Kruglov, J. Zhou, Y. Wang, E. Greenberg, V. B. Prakapenka, J. Liu, A. R. Oganov, and J.-F. Lin, *Nat. Commun.* **10**, 4453 (2019).
- [275] *Materials Genome Initiative* (2011 (accessed July 1, 2019)), <https://www.mgi.gov/>.
- [276] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson, *APL Materials* **1**, 011002 (2013).
- [277] J. E. Saal, S. Kirklin, M. Aykol, B. Meredig, and C. Wolverton, *JOM* **65**, 1501 (2013).
- [278] R. H. Taylor, F. Rose, C. Toher, O. Levy, K. Yang, M. B. Nardelli, and S. Curtarolo, *Comput. Mater. Sci.* **93**, 178 (2014).
- [279] A. Agrawal and A. Choudhary, *MRS Commun.* **9**, 779 (2019).
- [280] J. Lee and R. Asahi, *Comput. Mater. Sci.* **190**, 110314 (2021).
- [281] S. Batzner, A. Musaelian, L. Sun, M. Geiger, J. P. Mailoa, M. Kornbluth, N. Molinari, T. E. Smidt, and B. Kozinsky, *Nat. Commun.* **13**, 2453 (2022).
- [282] O. Kaba and S. Ravanbakhsh, *Advances in Neural Information Processing Systems* **35**, 4150 (2022).
- [283] X. Gong, H. Li, N. Zou, R. Xu, W. Duan, and Y. Xu, *Nat. Commun.* **14**, 2848 (2023).
- [284] C. Kuenneth and R. Ramprasad, *Nat. Commun.* **14**, 4099 (2023).
- [285] R. Gurnani, C. Kuenneth, A. Toland, and R. Ramprasad, *Chem. Mater.* **35**, 1560 (2023).
- [286] W. Fan and F. Geerts, “Information completeness,” in *Foundations of Data Quality Management* (Springer International Publishing, Cham, 2012) pp. 93–118.
- [287] T. N. Vu, S. K. Nayak, N. T. T. Nguyen, S. P. Alpay, and H. Tran, *AIP Adv.* **11**, 045120 (2021).
- [288] V. N. Tuoc, N. T. Nguyen, V. Sharma, and T. D. Huan, *Phys. Rev. Mater.* **5**, 125402 (2021).
- [289] G. Bergerhoff, R. Hundt, R. Sievers, and I. Brown, *J. Chem. Inf. Comput. Sci.* **23**, 66 (1983).
- [290] L. Ward, A. Agrawal, A. Choudhary, and C. Wolverton, *npj Comput. Mater.* **2**, 16028 (2016).
- [291] L. Breiman, *Mach. Learn.* **45**, 5 (2001).
- [292] J. Tuoriniemi, K. Juntunen-Nurmilaukas, J. Uusvuori, E. Pentti, A. Salmela, and A. Sebedash, *Nature* **447**, 187 (2007).
- [293] B. Matthias, in *Progress in Low Temperature Physics*, Vol. 2 (Elsevier, 1957) Chap. 5, pp. 138–150.
- [294] C. Buzea and K. Robbie, *Supercond. Sci. Technol.* **18**, R1 (2004).
- [295] V. Kuz'min, A. G. Artemenko, and E. N. Muratov, *J. Comp. Aided Mol. Des.* **22**, 403 (2008).
- [296] T. Kohonen, *Neural Netw.* **37**, 52 (2013).
- [297] L. Van der Maaten and G. Hinton, *J. Mach. Learn. Res.* **9**, 2579 (2008).
- [298] T. Chen and C. Guestrin, in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (2016) pp. 785–794.
- [299] C. Kuenneth, A. C. Rajan, H. Tran, L. Chen, C. Kim, and R. Ramprasad, *Patterns* **2**, 100238 (2021).
- [300] C. Kuenneth, J. Lalonde, B. L. Marrone, C. N. Iverson, R. Ramprasad, and G. Pilania, *Commun. Mater.* **3**, 96 (2022).
- [301] S. Gražulis, A. Daškevič, A. Merkys, D. Chateigner, L. Lutterotti, M. Quirós, N. R. Serebryanaya, P. Moeck, R. T. Downs, and A. Le Bail, *Nucleic Acids Res.* **40**, D420 (2012).
- [302] M. J. Winiarski, B. Wiendlocha, S. Gołąb, S. K. Kushwaha, P. Wiśniewski, D. Kaczorowski, J. D. Thompson, R. J. Cava, and T. Klimczuk, *Phys. Chem. Chem. Phys.* **18**, 21737 (2016).
- [303] S. Wold, M. Sjöström, and L. Eriksson, *Chemometr. Intell. Lab. Syst.* **58**, 109 (2001).
- [304] S. De, A. P. Bartók, G. Csányi, and M. Ceriotti, *Phys. Chem. Chem. Phys.* **18**, 13754 (2016).
- [305] A. J. Smola and B. Schölkopf, *Stat. Comput.* **14**, 199 (2004).
- [306] J. H. Friedman, *Ann. Stat.*, 1189 (2001).
- [307] K. Choudhary and B. DeCost, *npj Comput. Mater.* **7**, 185 (2021).
- [308] L. Ward, A. Dunn, A. Faghaninia, N. E. Zimmermann, S. Bajaj, Q. Wang, J. Montoya, J. Chen, K. Bystrom, M. Dylla, K. Chard, M. Asta, K. A. Persson, G. J. Snyder, I. Foster, and A. Jain, *Comput. Mater. Sci.* **152**, 60 (2018).
- [309] C. K. I. Williams and C. E. Rasmussen, in *Advances in Neural Information Processing Systems 8*, edited by D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo (MIT Press, 1995).
- [310] C. E. Rasmussen and C. K. I. Williams, eds., *Gaussian Processes for Machine Learning* (The MIT Press, Cambridge, MA, 2006).
- [311] P.-P. De Breuck, G. Hautier, and G.-M. Rignanese, *npj Comput. Mater.* **7**, 83 (2021).
- [312] “Alexandria database,” <https://alexandria.icams.rub.de>.
- [313] J. Behler and M. Parrinello, *Phys. Rev. Lett.* **98**, 146401 (2007).
- [314] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, *Phys. Rev. Lett.* **104**, 136403 (2010).

- [315] A. P. Bartók, S. De, C. Poelking, N. Bernstein, J. R. Kermode, G. Csányi, and M. Ceriotti, *Sci. Adv.* **3**, e1701816 (2017).
- [316] A. P. Thompson, L. P. Swiler, C. R. Trott, S. M. Foiles, and G. K. Tucker, *J. Comput. Phys.* **285**, 316 (2015).
- [317] C. Chen, Z. Deng, R. Tran, H. Tang, I.-H. Chu, and S. P. Ong, *Phys. Rev. Mater.* **1**, 043603 (2017).
- [318] A. V. Shapeev, *Multiscale Model. Simul.* **14**, 1153 (2016).
- [319] K. Gubaev, E. V. Podryabinkin, G. L. Hart, and A. V. Shapeev, *Comput. Mater. Sci.* **156**, 148 (2019).
- [320] L. Zhang, J. Han, H. Wang, R. Car, and E. Weinan, *Phys. Rev. Lett.* **120**, 143001 (2018).
- [321] “Matlantis,” <https://matlantis.com/>.
- [322] C. J. Pickard, *Phys. Rev. B* **106**, 014102 (2022).
- [323] R. Ouyang, S. Curtarolo, E. Ahmetcik, M. Scheffler, and L. M. Ghiringhelli, *Phys. Rev. Mater.* **2**, 083802 (2018).
- [324] H. J. Choi, S. G. Louie, and M. L. Cohen, *Phys. Rev. B* **79**, 094518 (2009).
- [325] M. Calandra, G. Profeta, and F. Mauri, *Phys. Rev. B* **82**, 165111 (2010).
- [326] Y. Kong, O. Dolgov, O. Jepsen, and O. Andersen, *Phys. Rev. B* **64**, 020501 (2001).
- [327] K.-P. Bohnen, R. Heid, and B. Renker, *Phys. Rev. Lett.* **86**, 5771 (2001).
- [328] H. J. Choi, D. Roundy, H. Sun, M. L. Cohen, and S. G. Louie, *Phys. Rev. B* **66**, 020513 (2002).
- [329] A. Chandrasekaran, D. Kamal, R. Batra, C. Kim, L. Chen, and R. Ramprasad, *npj Comput. Mater.* **5**, 22 (2019).
- [330] H. Li, Z. Wang, N. Zou, M. Ye, R. Xu, X. Gong, W. Duan, and Y. Xu, *Nat. Comput. Sci.* **2**, 367 (2022).
- [331] M. Scherbela, R. Reisenhofer, L. Gerard, P. Marquetand, and P. Grohs, *Nat. Comput. Sci.* **2**, 331 (2022).
- [332] B. G. del Rio, B. Phan, and R. Ramprasad, *npj Comput. Mater.* **9**, 158 (2023).
- [333] B. G. Del Rio, C. Kuenneth, H. D. Tran, and R. Ramprasad, *J. Phys. Chem. A* **124**, 9496 (2020).
- [334] H. Li, Z. Tang, J. Fu, W.-H. Dong, N. Zou, X. Gong, W. Duan, and Y. Xu, *Phys. Rev. Lett.* **132**, 096401 (2024).
- [335] M. Lüders, M. Marques, N. Lathiotakis, A. Floris, G. Profeta, L. Fast, A. Continenza, S. Massidda, and E. Gross, *Phys. Rev. B* **72**, 024545 (2005).
- [336] M. Marques, M. Lüders, N. Lathiotakis, G. Profeta, A. Floris, L. Fast, A. Continenza, E. Gross, and S. Massidda, *Phys. Rev. B* **72**, 024546 (2005).
- [337] A. Sanna, C. Pellegrini, and E. Gross, *Phys. Rev. Lett.* **125**, 057001 (2020).
- [338] Z. Li, G. Antonius, Y.-H. Chan, and S. G. Louie, *Comput. Phys. Commun.* **295**, 109003 (2024).
- [339] “ChatGPT,” <https://chat.openai.com/chat>.
- [340] “Llama,” <https://llama.meta.com/>.
- [341] J. Dagdelen, A. Dunn, S. Lee, N. Walker, A. S. Rosen, G. Ceder, K. A. Persson, and A. Jain, *Nat. Commun.* **15**, 1418 (2024).
- [342] Z. Zheng, O. Zhang, C. Borgs, J. T. Chayes, and O. M. Yaghi, *J. Am. Chem. Soc.* **145**, 18048 (2023).
- [343] P. Cadden-Zimansky, J. Wei, and V. Chandrasekhar, *Nat. Phys.* **5**, 393 (2009).
- [344] R. Balestrieri, J. Pesenti, and Y. LeCun, *arXiv preprint arXiv:2110.09485* (2021).