Applied Physics Reviews REVIEW scitation.org/journal/are

Metastable materials discovery in the age of large-scale computation

Cite as: Appl. Phys. Rev. 8, 031310 (2021); doi: 10.1063/5.0049453
Submitted: 4 March 2021 · Accepted: 22 July 2021 ·
Published Online: 25 August 2021

Félix Therrien,[1,2] Eric B. Jones,[1,2] and Vladan Stevanović[1,2,a)]

AFFILIATIONS

[1]Colorado School of Mines, Golden, Colorado 80401, USA
[2]National Renewable Energy Laboratory, Golden, Colorado 80401, USA

[a)]Author to whom correspondence should be addressed: vstevano@mines.edu

ABSTRACT

Computational materials discovery has been successful in predicting novel, technologically relevant materials. However, it has remained focused almost exclusively on finding ground-state structures. Now that the lower-hanging fruit has been found in many fields of application, materials exploration is moving toward metastable materials: higher energy phases that are stable at practical time scales. Because of the challenges associated with predicting which phases are realistic, this class of materials has remained relatively unexplored, despite numerous examples of metastable structures with unmatched properties (e.g., diamond). This article highlights recent advances in developing computational and theoretical methods for predicting useful and realizable metastable materials. Topics discussed cover (1) the latest strategies for identifying potential metastable phases, (2) methodologies for assessing which phases can be realized experimentally, and (3) current approaches to estimate the lifetime of metastable materials.

TABLE OF CONTENTS

I. INTRODUCTION 1
II. EXPLORATION: FINDING CANDIDATE METASTABLE STATES 2
III. REALIZABILITY: IDENTIFYING SYNTHESIZABLE METASTABLE STRUCTURES 3
A. Thermodynamic limits 4
B. Size of attraction basins 5
C. Application to noncrystalline (glassy, amorphous) materials 7
IV. LIFETIME: UNDERSTANDING THE TRANSITIONS BETWEEN METASTABLE MATERIALS 9
A. Lifetime assessment 9
B. Matching crystal structures 11
C. Minimal energy pathways 13
V. CONCLUSION AND CURRENT CHALLENGES 14

I. INTRODUCTION

Metastable forms of matter have the particularity of being durable in environments and conditions at which they do not represent the global minimum of the free energy or some other relevant thermodynamic potential. It is fairly easy to make an argument on how invaluable they are to our daily lives. Take diamond for example, a crystalline form of elemental carbon that exists at ambient conditions despite not being the lowest energy state.[1] Diamond is sufficiently long-lived not only for wedding rings to outlast any terrestrial marriage but also for its mechanical, optical, and electronic properties to be used in a number of technologically relevant applications. Another well-known example of a long-lived metastable state is "ordinary" glass, a noncrystalline form of SiO_2 that is critically important in a number of ways.[2,3] The "sweetest" example is certainly solid chocolate. The most utilized form for its desired taste and texture is the so-called polymorph-V that is also metastable at room temperature and ambient pressure.[4,5]

The beginnings of research into the phenomenon of metastability and the metastable states of matter, polymorphs in particular, can be traced back to the late 18th and early 19th centuries. The discovery that calcite and aragonite, two minerals with very different physical properties, share the same $CaCO_3$ chemical composition was most likely the initial motivator.[6,7] Numerous studies have been conducted since, and a significant body of knowledge and literature has been developed.[8–10] Nonetheless, if one were to ask which metastable states exhibiting a given set of target properties exist within a given chemistry but are also experimentally synthesizable and sufficiently long-lived so that their properties can be utilized, they would quickly realize that

critical knowledge gaps still remain for these questions to be answered with a high degree of certainty. As will be elaborated later, the main reason is that current theories and computational tools lack sufficient predictive capabilities. On the flip side, the space of metastable states represents a virtually uncharted territory for finding new and potentially game-changing materials as the aforementioned examples (and many others) clearly illustrate. Consequently, improving the predictive power of theory, while without a doubt representing an important and intellectually very appealing endeavor, is also of substantial practical significance. In this paper, we review recent work aimed to advance metastable materials theory and enable development of robust computational methodologies for metastable materials discovery and design. We also discuss their strengths and weaknesses as well as the remaining outstanding challenges standing in the way of reliable predictions of potentially useful, synthesizable, and long-lived metastable states of solid matter.

As depicted in Fig. 1, the starting point for any metastable materials discovery is the exploration step, which addresses the first part of the problem: What is there to be found? It can be argued that presently, theorists are well equipped to predict which potentially useful structures might be present in a given chemistry. The electronic structure methods have matured to the point where the properties (functional or otherwise) of large swaths of different chemistries can be reliably predicted for any given crystal structure.[11] When combined with structure prediction methods, which themselves have seen substantial advancements over the past several decades,[12,13] property predictions can be utilized quite effectively to screen various ground-state and higher-lying structures for useful properties. Depending on the desired accuracy, this might require significant computational resources, of course, but those will become more and more available with time.

The second and third parts of the problem are where the limits of our current predictive power lie. Step 2 of the metastable materials discovery process consists of filtering out from all possible structures those that are more accessible in terms of their experimental realizability (or synthesizability). In Fig. 1, two of the five structures identified in step 1 are not realizable experimentally. They could be identified, for example, by measuring the relative size of their basin of attraction, which is an indicator of their feasibility. This is one of the approaches discussed in Sec. III.

In step 3, only structures that will "live" sufficiently long are selected as plausible metastable materials. In Fig. 1, evaluating the energy along the minimal energy pathway, using tools described in Sec. IV, could reveal the remaining unstable structure. In the hypothetical example of Fig. 1, there is, in the end, only one plausible metastable structure (the other one being the—stable—ground state).

Because of challenges associated with predicting lifetime and realizability, recent work mainly concentrates on steps 2 and 3, as does this review/discussion. This paper will, however, start with a brief overview of exploration methodologies because they naturally provide the context within which synthesizability/realizability and lifetimes questions emerge. Afterward, it dives deeper into the ideas and methodologies that address realization and lifetimes of metastable phases.

II. EXPLORATION: FINDING CANDIDATE METASTABLE STATES

As already noted, modern (first-principles) electronic structure methods are presently more than capable of providing reasonably accurate predictions of many physical properties of materials spanning a range of different chemistries and virtually any crystal (or molecular) structure. There are limits in the number of atoms, but those are being pushed continuously as computer power increases and numerical implementations improve. The key then really is the input crystal structure. How can one know, in the absence of experimental knowledge, of which structure to compute the properties? Which ones out of infinitely many possibilities warrant our attention? These are the central questions of materials discovery and the ones that will be discussed in this section.

One can go about answering these questions in two different, complementary, and arguably equally interesting and important ways. First, one can focus on finding the structure that maximizes (or minimizes) a target property and leave the question of experimental realizability to experiments. Second, one can try to identify structures that

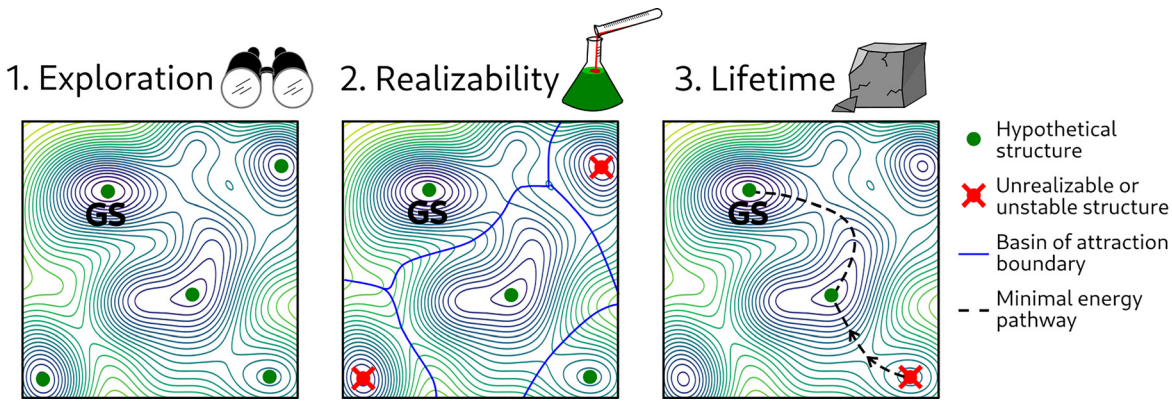

FIG. 1. A visual representation of the process of finding useful metastable materials. The three square panels are contour plots of the potential energy surface. Blue lines in panel 2 outline the basins of attraction of various local minima, while the dashed black line in panel 3 represents the minimal energy pathway between the structures. In this example, among the five hypothetical phases of a material found at step 1, three are realizable (step 2) out of which two are stable on a practical timescale (step 3). Each of the three steps narrows down the search.

fulfill some predefined set of synthesis constraints and then ask the question of whether any one of those exhibits the desired properties. There are numerous examples of both approaches in the literature.

Predictions of structures that are harder than diamond[14] or specific atomic configurations in an alloy that achieve a certain feature in the electronic structure, such as maximizing the electronic bandgap,[15] making the bandgap direct and optically allowed,[16] etc., belong to the former category. While extremely important in revealing what a given chemistry has to offer, the primary weakness of these property-driven predictive approaches is that they lack any quantifiable measure of experimental realizability of the candidate structures and/or particular arrangements of atoms in an alloy. It is not surprising then that the experimental validation of these kinds of predictions is somewhat lagging.

Alternatively, one can take some aspects of synthesizability into account by invoking the thermodynamic energy minimization principle and by limiting the search to structures that represent global minima of the total energy (or some other thermodynamic potential) and ask which ones of those possess the desired properties. This approach is typically adopted in what is broadly referred to as "structure prediction," which typically focuses (but not exclusively) on finding the ground-state structures, that is, the global minima on the potential energy (only the structure with the lowest energy in Fig. 2) or in the case of high-pressure synthesis, the global minima on the potential enthalpy surface. Again, it is the development of first-principles methods, which provide the means to efficiently and accurately evaluate total energies and their gradients (forces), that enabled the emergence of contemporary structure predictions. Incidentally, the problem of predicting the ground-state structure of solids was initially thought of as unlikely to ever be solved.[17]

A number of methods have been developed and applied for the purpose of predicting the ground-state structure of various chemistries. More prominent ones include simulated annealing,[18,19] methods based on evolutionary algorithms,[20–22] data mining and machine learning approaches,[23,24] and structure prototyping.[25,26] Less biased toward the global minimum are sampling methods such as metadynamics,[27,28] basin and minima hopping,[29,30] and *ab initio* random structure search.[31]

Structure predictions, when integrated with property prediction, have led to much more successful experimental validation and discoveries of truly exciting new materials and/or high-pressure phases. These include new phosphors for white light-emitting diodes,[32] new phosphates for lithium-ion battery cathodes,[33] an unexpected transparent high-pressure phase of sodium,[34] and new conventional superconductors with the critical temperature approaching 200 K in the high-pressure phase of $(H_2S)_2H_2$,[35] to name a few of many (further examples can be found in Refs. 36 and 37) These methods have been so successful that it has been argued, and rightfully so, that the structure predictions are literally the key driver of materials discovery.[37]

It is important to note that very few, if any, of the theory-predicted and experimentally realized materials are actually metastable. Most of them are in their ground states either at ambient conditions or elevated pressure. While structure predictions at high pressure achieved a significant number of very important discoveries, only a small fraction of these were successfully quenched to ambient pressures,[38] and their ability to be quenched (metastability) was not predicted; it was largely serendipitous. Furthermore, in virtually all these works, the discovery of metastable states, crystalline typically, has largely been reduced to identifying low-energy local minima (inherent structures) and computing their properties. (There are numerous references in the literature; see, for example, Refs. 31 and 39.) As will be discussed in the Sec. III, energy alone is hardly enough to serve as a good predictor or descriptor of the experimental realizability of metastable structures.

Finally, it is also important to note that predicting the structure of noncrystalline states, both glasses and amorphous structures, which are usually thought of as a single microstate (inherent structure) in an infinitely large unit cell,[2,3] falls beyond the reach of present structure prediction methods. The amorphous structures are typically modeled using sampling methods (molecular dynamics, Monte Carlo, etc.) in an attempt to meet some *a priori* requirements that are known from experiments, for example, the radial distribution function, distribution of bonding environments, and/or density.[40] Hence, with regard to noncrystalline phases, research is presently explanatory (i.e., fitting models to experiments) rather than predictive, which is a serious impediment to the discovery and design of functional, noncrystalline materials.

III. REALIZABILITY: IDENTIFYING SYNTHESIZABLE METASTABLE STRUCTURES

Once candidate metastable materials have been identified, i.e., local minima of the potential energy surface have been found, the question that arises is: How does one identify, among those candidates, the ones that are realistic? Indeed, the mere predicted existence of a minimum on the PES provides no guarantee of the realizability of its corresponding phase; getting to the minimum and staying there may be exceedingly hard or even impossible. Two methods have emerged to assess how realistic a potential metastable material is. They are presented in this section.

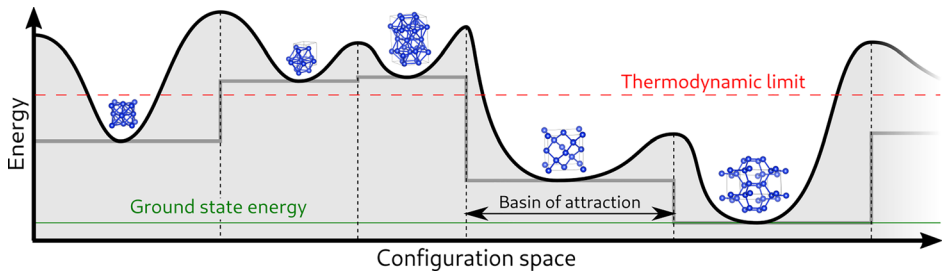

FIG. 2. One-dimensional representation of a hypothetical PES. Each crystal structure resides in a local minimum of the PES. The pale gray line is a representation of the "flat basin" approximation. The dashed red line illustrates the idea of a thermodynamic limit above which structures would not be realizable.

A. Thermodynamic limits

An intuitive way to resolve this problem is to use the total energy of the hypothetical material and compare it to its most stable configuration at that composition, including phase separated states. To do so, one uses the concept of convex hull. In a binary phase diagram (enthalpy of formation [ΔH_f] vs chemical composition [x]), the convex hull (as defined in materials science) is the set of straight lines connecting the "lowest energy" phases.[42] In other words, in a ΔH_f vs x binary phase diagram where all possible crystalline structures are represented as points, if one were to attach an elastic band to the two pure compounds, stretch it under all the points, and let it rest on them, it would have the shape of the convex hull. Any state on the convex hull is the most stable at the corresponding composition. Any point that is directly above the convex hull is a high-energy state; if it is also stable, it is a metastable state. In phase diagrams with several crystalline structures and/or multiple composition dimensions where the convex hull is a hyper surface, determining the convex hull can be difficult. Systematic approaches to solving this problem and evaluating phase stability of hypothetical phases are presented in Refs. 43 and 44. The difference in energy between the metastable state and the ground state, referred to as energy above the convex hull or simply energy above the hull, can be seen as a representation of the thermodynamic scale of metastability, i.e., it represents how metastable a material is. It is important to note that in many cases, the ground state could be the phase separation into competing phases, not necessarily the lower energy structure at the same chemical composition.

Sun *et al.*[41] analyzed about 30 000 materials in the Inorganic Crystal Structure Database (ICSD) to measure the scale of metastability. They took into account density functional theory (DFT) approximation errors and curated the ICSD data to include only well-described, experimentally realized materials. They found that 50.4 ± 4% of those materials are metastable and that their median energy above the hull is 15 ± 0.5 meV per atom with about 90% of the structures below 67 meV. When comparing the energy above the ground state for different classes of materials, nitrides were found to exhibit the largest scale of metastability on average and iodine the smallest. It is important to note that a large fraction of the systems in ICSD are metastable relative to chemical decomposition into competing phases. Nitrides in particular are largely unstable/metastable relative to elemental nitrogen due to strong bonding in the N_2 molecule. To explain this difference between nitrides and other chemistries, the authors suggested that the stronger bonds found in nitrides could prevent the long-range diffusion needed for phase separation and help to stabilize higher energy metastable phases that would otherwise decompose. More recently, in a study that included a larger number of compounds (≈37 000), it was found that the average energy above the hull of metastable compounds could be even higher and that hypothetical structures with energy above 200 meV should not automatically be discarded.[45]

The fact (statistically determined) that most observed metastable materials exist within a relatively small scale of metastability, as apparent in Fig. 3(a), does not imply that the converse is true: low energy above the hull does not imply that a hypothetical phase will be realizable. Indeed, looking at Fig. 3(b) where hypothetical polymorphs are shown in red and observed ones in blue, it is clear that several hypothetical polymorphs have lower energy than observed ones. The question then remains: Given a hypothetical structure, how does one *predict* whether it is realistic? In an attempt to answer this question, Sun *et al.*[41] introduced the concept of remnant metastability: A phase is realizable if and only if there exists a set of thermodynamic conditions (temperature, pressure, magnetic field, etc.) under which it is the lowest energy phase. In Fig. 4, polymorph C would be realizable

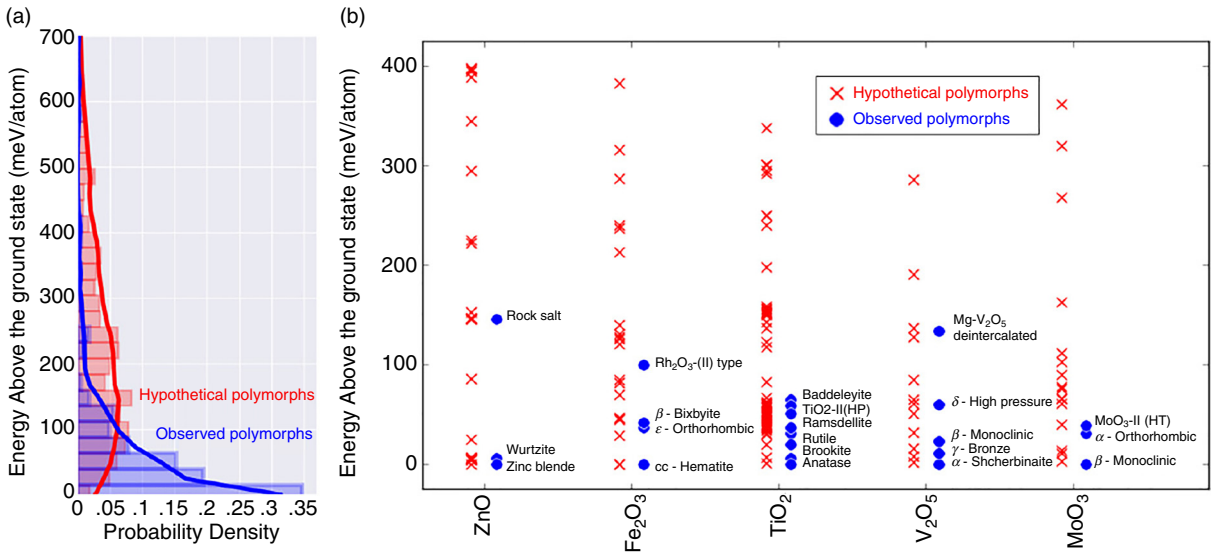

FIG. 3. Representations of the thermodynamic scale of metastability. (a) Distribution of the energy above the ground state for observed and hypothetical polymorphs (generated with a data-mined structure predictor). (b) Energy distribution of hypothetical (red) and observed (blue) structures for five common oxides. Reprinted with permission from Sun *et al.*, Science Adv. **2**, e1600225 (2016). Copyright 2016 Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. Licensed under a Creative Commons Attribution 4.0 Noncommercial (CC BY-NC) license.[41]

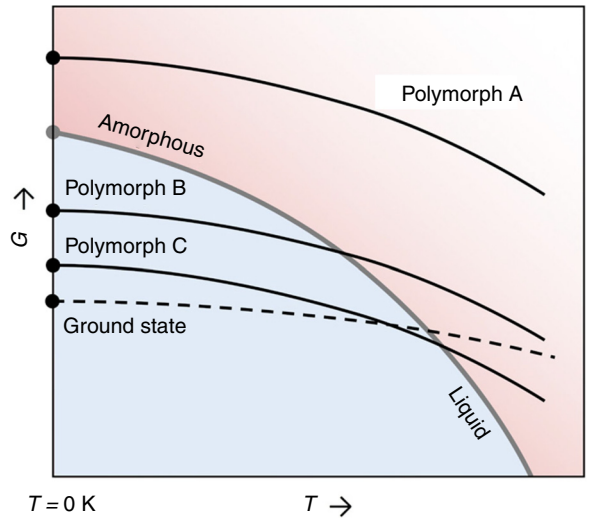

FIG. 4. Evolution of the free energy of different polymorphs as a function of the thermodynamic conditions. Reprinted with permission from Aykol *et al.*, Science Adv. **4**, eaaq0148 (2018). Copyright 2018 Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. Licensed under a Creative Commons Attribution 4.0 Noncommercial (CC BY-NC) license.[47]

because it is the lowest free energy state for a set of temperatures, but polymorph B would not. Since adjusting the thermodynamic conditions, temperature in particular, to change the free energy by more than about 100 meV/atom is difficult, remnant metastability could explain why the great majority of measurable states have energies below that threshold. However, the authors did not provide a way to estimate the free energy under different thermodynamic conditions, and there is, in principle, an infinity of such conditions, which makes the concept of remnant metastability difficult to verify. Furthermore, if one analyzes means other than temperature to deform the potential energy surface, pressure, for example, a much larger energy scale, becomes experimentally accessible. For example, just by taking the volume difference between diamond and graphite (≈ -3 Å^3/atom) and multiplying by a pressure of 100 GPa ≈ 0.6 eV/Å^3, which today is accessible, one estimates that by applying high pressure, the energy differences of approximately 1.6 eV/atom can be overcome. This is to say that while practically very useful, thermodynamic scales of accessible energies depend on the actual experimental situation and that every situation requires its corresponding scale.

Srinivasan *et al.* recently created a methodology to construct metastable phase diagrams, i.e., 3D phase diagrams where the axes are pressure, temperature, and energy above the ground state.[46] They used an evolutionary algorithm to find all the possible polymorphs at a given composition and used molecular dynamics to evaluate their free energy on a grid of thermodynamic conditions (P,T). They then used machine learning to estimate the exact phase boundaries within the 3D phase diagram. This promising methodology can directly produce plots, like the one in Fig. 4, and could represent a way to verify the concept of remnant metastability.

In a later study, Aykol *et al.*[47] showed that the energy of the amorphous state of a material can serve as an upper limit to the scale of metastability. Indeed, if a crystalline state has higher free energy than the amorphous state at $T = 0$ K (e.g., polymorph A in Fig. 4), it can *never* be the lowest energy state at higher temperatures because the entropy (S) of the amorphous state is always higher than the entropy of any crystalline state and $\frac{\partial G}{\partial T}|_P = S$. This concept is illustrated in Fig. 4. Furthermore, changing other thermodynamic properties, such as pressure or electric field, might move the system to higher energy crystalline phases, but these phases would likely not survive at ambient conditions because of spontaneous amorphization.[48] The authors found their hypothesis to be true for all 41 systems that they tested in the ICSD. Among these systems, no experimentally realized phase was found to have higher energy than the amorphous state. It was also found that the amorphous limit is highly variable from one system to the other. Setting a fixed energy limit instead would either yield more false negatives or, on average, a higher energy bound. In other words, the amorphous limit is not so high that it could simply be replaced by a higher energy cutoff.

In short, the energy above the ground state, i.e., the thermodynamic scale of metastability, can be used to exclude high-energy hypothetical structures provided that the energy of the amorphous state can be calculated. However, it cannot identify which of the low-energy hypothetical structures, e.g., Fig. 3(b), are unrealistic. Furthermore, the connection—or *causality*–between the scale of metastability and the physical processes that affect synthesizability has yet to be established clearly. In other words, the following question needs to be answered: At a given set of thermodynamic conditions, why would a high-energy state with a certain energy above the ground state be more or less synthesizable than another one with a different energy? For example, when using sputtering as a synthesis method where the entropy of the deposited material is high, it is not the formation energy that determines the phase that is realized but rather the capacity to "accommodate" entropy.[49,50]

B. Size of attraction basins

That the energy above the structural ground state alone is insufficient to identify experimentally realizable hypothetical structures, as observed in Fig. 3, indicates that there are additional ordering principles at work in their manifestation. The conceptual use of the *basin of attraction* has a storied history in describing the configurational energy landscapes within materials science as well as in other fields.[8,9] If a hypothetical structure (P_α) is regarded as a local minimum (α) on the energy landscape, then a basin of attraction (B_α) is the set of nuclear configurations obtained by small deformations away from the local minimum. Small in this context indicates that a local (gradient-based) search algorithm seeded at the deformed coordinates would yield P_α as the local minimum. Equivalently, deformed coordinates of a structure in Fig. 2 would not cross any dashed lines within its basin of attraction.

It was not until relatively recently, however, that basin hypervolume was quantitatively identified as a critical element in the formation of experimentally realizable metastable structures.[51,53] In the context of metallic clusters, Goedecker and colleagues noted how only those clusters experiencing a nonglassy PES landscape (few larger minima as opposed to many narrow ones) can be synthesized experimentally.[53] Subsequently, Stevanović[51] employed a modified version of the *ab initio* random structure search[31] and measured the frequencies of occurrence of different structures, which gave a measure of the relative hypervolumes occupied by the corresponding local minima. He demonstrated how the structures with wider local minima (larger

hypervolumes) are more likely to be experimentally synthesized.[51] Specifically, in a procedure called random superlattice (RSL) sampling, thousands of hypothetical unrelaxed structures are created by generating, for each one, a set of random lattice parameters (a, b, c, α, β, γ) and atomic positions. Atoms are carefully distributed inside the randomly generated cells such that cations and anions alternate and that ions have a low probability of being placed too close together. Then, the atomic positions and cell shapes of all the randomly generated structures are relaxed using density functional theory, and their (relaxed) total energy is computed. As depicted in Fig. 5(a), the compounds MgO, ZnO, and SnO_2 all have hundreds of metastable candidates that reside lower than about 400 meV/atom above the ground- state energy. In contrast, the ground-state rock salt phase is the only experimentally known structure of MgO, while the known structures of ZnO are wurtzite, the ground state, and its two metastable polymorphs, zinc blende and rock salt. Finally, SnO_2 displays only its rutile ground state and one metastable phase, α-PbO_2. Solely using proximity to the ground-state energy would therefore lead to very poor predictive power of the random structure sampler. The "frequency of occurrence" plotted in Fig. 5, which is proportional to the basin hypervolume of a given structure (designated as basin of attraction in Fig. 2), orders candidates from largest hypervolume (frequency $\rightarrow$ 1) to smallest (frequency $\rightarrow$ 0) and is defined as the relative number of randomly seeded structures that relax to a particular structural minimum in the context of RSL sampling. Intuitively, since a larger basin hypervolume corresponds to an increased likelihood that a given minimum is "found" by the thermal motion of ions in the first place, the frequency of occurrence reflects which candidate structures are more likely to be experimentally realized.

In the instance of MgO shown in the left panel of Fig. 5(a), the $Fm\bar{3}m$ rock salt structure occurs more than an order of magnitude more frequently than the next most frequent structure $F\bar{4}3m$, offering an explanation as to why rock salt is the only observed experimental phase of the compound. Similarly, in the case of ZnO (middle panel), there is a cluster of structures, with space groups $F\bar{4}3m$, $I4mm$, $Fm\bar{3}m$, and $P6_3mc$, that occurs significantly more frequently than the rest. The $I4mm$ structure was shown subsequently to possess unstable phonon modes, and the remaining three correspond exactly to the three experimentally realized metastable polymorphs of ZnO. Moreover, those three structures are persistently the most frequently occurring as a function of increased unit cell size. Finally, while there is a crowd of likely low-energy structures for SnO_2 (right panel), it was shown that the only persistently frequent structures as a function of increased unit cell size were the experimentally known rutile ($P4_2/mnm$) and α-PbO_2 ($Pbcn$) phases.[51] Basin hypervolume as captured by frequency of occurrence is therefore a powerful ordering principle in discriminating between hypothetical structures that are more likely to be synthesizable experimentally and those that are less.

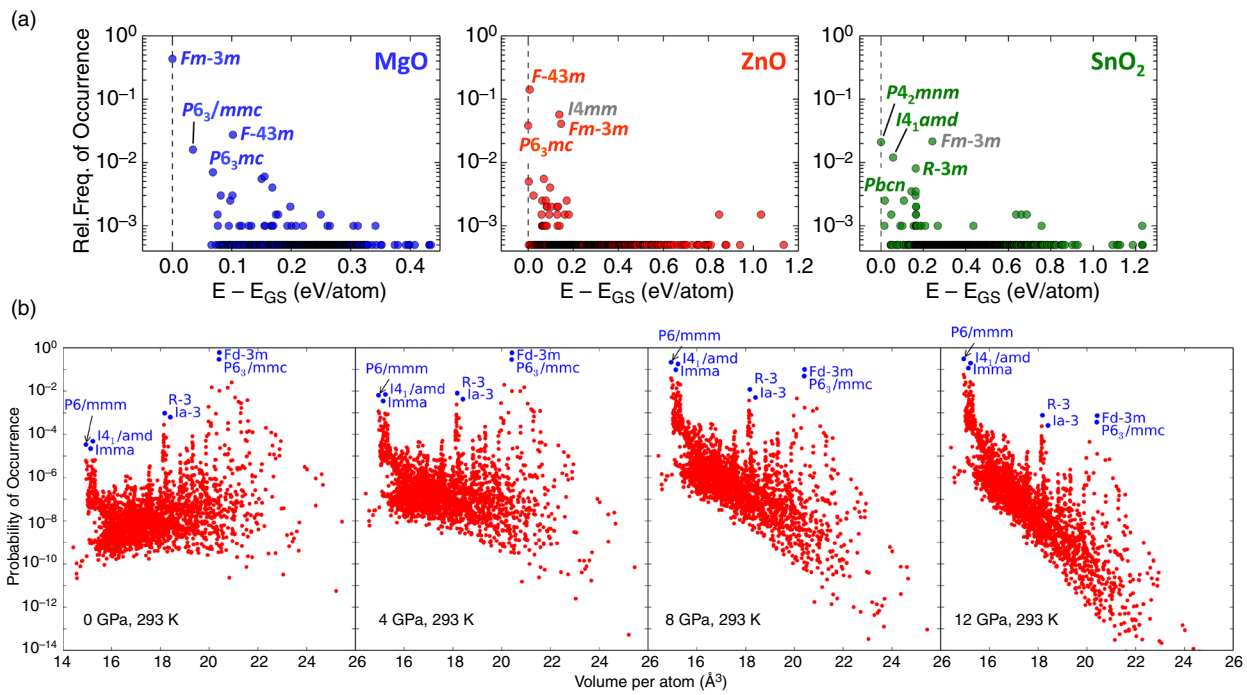

FIG. 5. (a) Relative frequencies (Rel. Freq.) of the occurrence of structures resulting from the random superlattice sampling shown against the energy above the ground state. Only the top occurring structures have their space groups explicitly marked. Space groups shown in gray represent the structures that are predicted to be dynamically unstable. Reprinted with permission from Stevanović, Phys. Rev. Lett. **116**, 075503 (2016). Copyright 2016 American Physical Society.[51] (b) Probability of occurrence at room temperature (≈293 K). Blue points refer to experimentally realized polymorphs, while red points refer to potential polymorphs hitherto not synthesized. Reprinted with permission from Jones and Stevanović, Phys. Rev. B **96**, 184101 (2017). Copyright 2017 American Physical Society.[52]

In order to provide a more formal foundation for this discussion, consider the configurational partition function of a generic and ergodic thermodynamic ensemble

$$\Xi = \prod_{j=1}^{M} \int dy_j \prod_{i=1}^{N} \int d\mathbf{r}_i e^{-\beta(\epsilon(\{\mathbf{r}_i\},\{x_j\})+x_j\cdot y_j)}, \tag{1}$$

where i indexes each of the N nuclei with position $\mathbf{r}_i$, j indexes each pair of M conjugate variables $x_j \cdot y_j$ involved in specifying the relevant thermodynamic potential (and thus ensemble), $\beta = 1/k_B T$ is the inverse temperature, and ϵ is the total energy per atom as a function of the variable sets $\{\mathbf{r}_i\}$ and $\{\mathbf{x}_j\}$. For instance, in the isothermal–isobaric, or (N, p, T), ensemble, $x_j \cdot y_j = pv\delta_{j,1}$, where p is external pressure and v is volume per atom, while in the grand canonical ensemble, $x_j \cdot y_j = -\mu\delta_{j,1}$, where μ is the chemical potential and the number of nuclei has been normalized to 1. One may even deal with magnetic ensembles with $x_j \cdot y_j = -\mathbf{B} \cdot \mathbf{m}_j$, where $\mathbf{m}_j$ is the local magnetic moment per atom and $\mathbf{B}$ is an applied (uniform) magnetic field. Quantities that are typically extensive in statistical mechanics, such as energy, volume, and particle number, are scaled to be intensive in Eq. (1) following the notion of the thermodynamic limit. Moreover, from a practical standpoint, in computer simulations, N is typically regarded as the number of atoms in the computational unit cell. Since different metastable structures may appear at different N, citing extensive, rather than intensive, thermodynamic quantities would bias sampling toward hypothetical metastable structures that exist in smaller unit cells.

On general grounds, one can partition the integrals in Eq. (1) into a sum over basins of attraction indexed by α. In the following, we collapse positional and conjugate variable coordinates into a single $(3N + M)$-dimensional "degree of freedom" vector where $\prod_j \int dy_j \prod_i \int d\mathbf{r}_i \equiv \int d\mathbf{R}$, and the attendant thermodynamic potential is defined as $\phi(\mathbf{R}) \equiv \epsilon(\{\mathbf{r}_i\}, \{x_j\}) + x_j \cdot y_j$. Equation (1) then reads

$$\Xi = \sum_{\alpha} \int_{B_\alpha} d\mathbf{R}_\alpha e^{-\beta\phi(\mathbf{R}_\alpha)}. \tag{2}$$

Each term in the summand of Eq. (2) is an intra-basin configurational partition function and may be denoted Ξ_α. In order to evaluate the probability that ions come to occupy a particular basin and, thus, to be assigned to a particular hypothetical structure P_α, we make the following "flat basin" approximation (represented as a gray line in Fig. 2) $\phi(\mathbf{R}_\alpha) \to \phi_\alpha$, the potential at the minimum of the basin B_α. Under this approximation, which is well justified at ambient temperatures, the "probability of occurrence" becomes

$$\mathcal{P}_\alpha \equiv \frac{\Xi_\alpha}{\Xi} \approx \frac{V_\alpha^H e^{-\beta\phi_\alpha}}{\Xi}, \tag{3}$$

where $V_\alpha^H = \int_{B_\alpha} d\mathbf{R}_\alpha$ is the hypervolume of basin B_α, justifying theoretically its use as an ordering principle in indicating experimentally realistic metastable structures. In moving from basin hypervolume to frequency of occurrence (f_α), one notes that

$$\frac{V_\alpha^H}{V_\beta^H} = \lim_{N_{RS}\to\infty} \frac{N_\alpha/N_{RS}}{N_\beta/N_{RS}} \equiv \frac{f_\alpha}{f_\beta}, \tag{4}$$

where computationally, N_α is the number of initial random structures N_{RS} that map to a given local minimum α. Given the limit in Eq. (4), practically, one must ensure convergence of the frequencies as a function of N_{RS}.

Jones and Stevanović introduced and demonstrated that the formalism in Eqs. (1)–(4) could be applied in order to predict the rich polymorphism of elemental silicon (Si).[52] Figure 5(b) shows the probabilities of occurrence $\mathcal{P}_\alpha = f_\alpha e^{-\beta(\epsilon_\alpha + pv_\alpha)}/\Xi$ of a little over 2500 hypothetical structures discovered via random structure sampling on an eight-atom unit cell of silicon in the (N, p, T) ensemble as a function of volume. All distributions are plotted at room temperature, and the various panels correspond to increasing pressure values from left to right. Importantly, all experimental metastable polymorphs of Si realized under equilibrium conditions are identified as local peaks in the probability distributions (shown in blue), namely, the body-centered eight-atom polymorph ($Ia\bar{3}$), the rhombohedral eight-atom polymorph ($R\bar{3}$), and the lonsdaleite analog ($P6_3/mmc$), in addition to the diamond ground state ($Fd\bar{3}m$). As pressure is increased, the first three of the high-pressure, reversible polymorphs of Si, space groups $I4_1/amd$, $Imma$, and $P6/mmm$, become the most probable structures. Further confirmation of this methodology can be found in the work of Yang *et al.*[54] dedicated to various carbon allotropes.

Hence, employing the probability of occurrence to order results from random structure sampling has strong predictive power; accommodates thermodynamic parameters; and, perhaps most importantly, is *exclusionary* toward the vast majority of structures uncovered during such a search.

C. Application to noncrystalline (glassy, amorphous) materials

In addition to its utility in describing the experimental realizability of metastable structures, the statistical formalism developed in Sec. III B also holds predictive power toward describing structural descriptors of glassy solids, as demonstrated by Jones and Stevanović.[55] In particular, they showed that starting from the two-point density function

$$n^{(2)}(\mathbf{r}', \mathbf{r}'') = \left\langle \sum_{i=1}^{N} \sum_{j\neq i} \delta(\mathbf{r}_i - \mathbf{r}')\delta(\mathbf{r}_j - \mathbf{r}'') \right\rangle_{\Xi}, \tag{5}$$

where the expectation value is taken in the relevant ensemble Ξ, and using an approach similar to the derivation in Eqs. (1)–(4), one can express the radial distribution function (RDF) of species in the material as

$$g^{(2)}(r) = \sum_{\alpha} \left(\frac{n_\alpha}{\langle n \rangle_\Xi} \right) \mathcal{P}_\alpha g_\alpha^{(2)}(r), \tag{6}$$

where n_α and $g_\alpha^{(2)}(r)$ are the number density and RDF of the local minimum P_α, respectively; $\langle n \rangle_\Xi$ is the number density ensemble expectation value; and $\mathcal{P}_\alpha = f_\alpha e^{-\beta\phi_\alpha}/\Xi$ is a structure's probability of occurrence, as before. Similarly, the powder diffraction intensity can be expressed as

$$I(2\theta) = \sum_{\alpha} \mathcal{P}_\alpha I_\alpha(2\theta). \tag{7}$$

Equations (6) and (7) were developed primarily to describe ionic and covalent glasses but, in principle, could be applied to any glassy system. It is perhaps unsurprising that the description of such systems involves thermal averages given that polymeric and metallic glasses

have traditionally been modeled statistically through the random coil and random close packing models, respectively.[56] What is particularly surprising, however, is that one can evaluate Eqs. (6) and (7) using computational unit cells as small as a few dozen formula units and still obtain experimentally accurate structural descriptors.

Figure 6 shows the ensemble-averaged RDF (top panel) and XRD (bottom panel) patterns evaluated by performing random structure sampling on a 24-atom unit cell of Si. In the low-temperature ($\tau = k_B T/N$ is an effective, scaled temperature) limit shown in blue, both the RDF and XRD display the sharp crystalline peaks characteristic of ground-state, diamond Si. As τ is raised, the peaks broaden until in the high-temperature limit (red curve) they overlay the corresponding experimental RDF and XRD of amorphous Si to a remarkable degree of accuracy. In addition, it was demonstrated that partial RDFs of vitreous SiO_2, calculated using molecular dynamics on a 3000-atom supercell, could also be replicated using this ensemble method, establishing such "crystalline ensembles" as a complementary statistical model to the continuous random network (CRN) model for calculating properties of glassy materials. Moreover, the ensemble models are devoid of experimental inputs and *ad hoc* parameters often required by CRN models.

In addition to predicting structural features, which are prerequisite for any property predictions, another important property, in particular

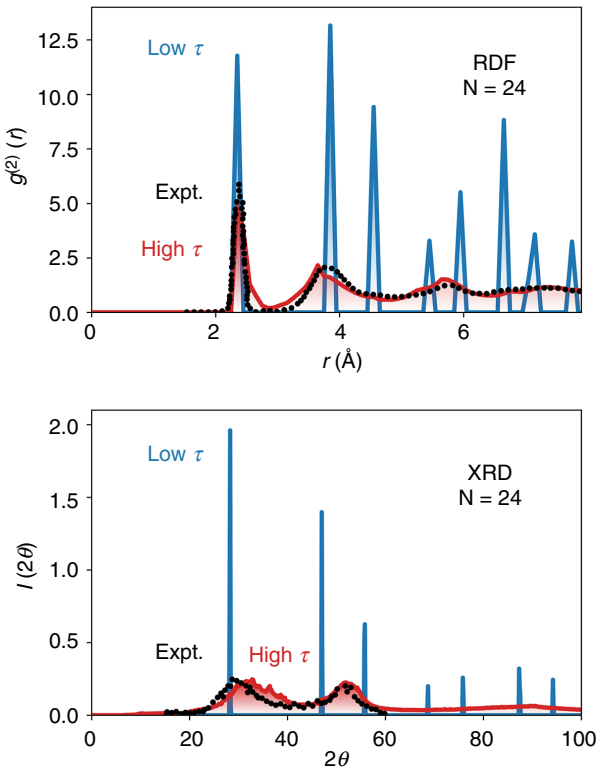

FIG. 6. In both the ensemble-averaged RDF (top panel) and XRD (bottom panel) patterns, the low-temperature (τ) regime shows the characteristic peaks of diamond silicon, while the high-temperature regime displays remarkable agreement with the experimental (Expt.) RDF and XRD for a-Si. Reprinted with permission from Jones and Stevanović, Npj Comput. Mater. **6**, 1–6 (2020). Copyright 2020 Authors, licensed under a Creative Commons Attribution (CC BY) license.[55]

for glassy systems, is the ability to undergo the glass transition. Not all chemistries can be quenched from the melt fast enough so as to form the glassy state. To predict the realizability of the glassy state, which is missing in the above discussion, Curtarolo and colleagues devised a spectral descriptor of the glass-forming ability (GFA) for metallic systems[57] (Fig. 7). In short, their work is founded in the concept of "confusion" proposed by Greer[58] in which the glass-forming ability is a consequence of the existence of a relatively large number of low-energy local minima that differ in their main structural characteristics. These structures would "confuse" the system in an attempt to crystallize upon cooling, leading to low crystallization rates and allowing for quenching into the glassy state. Hence, the higher the number of dissimilar structures at similar energy, the higher the glass-forming ability. It has been found recently that including slightly off-stoichiometric structures to this number improves the GFA predictions.[59]

In combination with the ensemble description of structural features and possibly other properties, this completes the set of ideas that will likely result in a robust methodology for discovering and designing useful and realizable glassy solids.

At the end of this discussion, it is important to mention another class of metastable systems related to glasses in the origin of their metastability: those exhibiting significant levels of lattice disorder. Examples include not only various alloy systems but also multinary compounds in which different atom types can exchange their lattice sites with a relatively low energy cost and higher entropy gain.[60] Like glasses, their metastability at ambient or near-ambient conditions stems from the suppressed long-range diffusion of atoms that would lead to ordering and/or phase separation, which are the lower free energy situations at lower temperatures. This type of metastability has been known for a long time. In some metal alloys, for example, it is possible to quench a single-phase mixture formed above a certain temperature due to the stabilizing influence of configurational entropy to conditions where, according to the phase diagrams, a two or more phase coexistence is the thermodynamically preferred state (miscibility gap).

This is a well-known and well-studied behavior in alloys, an important case being the metastable martensite phase in the Fe–C system.[61] Martensite is actually a metastable Fe–C phase at all temperatures and is a consequence of a rapid transformation of the entropy-stabilized austenite or the γ phase into a body-centered tetragonal Fe–C solid solution rather than the phase separation into Fe-rich body-centered cubic phase and the compound cementite (Fe_3C). Similarly, in some classes of multinary compounds, the low energy cost to creating the lattice disorder leads to low values of the order–disorder transition temperatures. This, in turn, allows the disordered state to be quenched rather easily if the diffusivity of atoms is also low, below the transition temperatures. Examples of these types of compounds that typically exhibit more significant levels of lattice disorder at almost any temperature include the entire family of spinel oxides,[62,63] various ternary nitride compounds such as $ZnSnN_2$[64] or $ZnGeN_2$,[65] etc. While all these effects have been known for a while, somewhat surprising recent work has shown entropy stabilization and quenching to lower temperature in manifestly immiscible systems, achieved by increasing the number of components both in metallic[66] and nonmetallic systems (ceramics).[67] These high-entropy alloys and ceramics are presently a very active area of research, and rather than reviewing here, we point the interested reader to already-existing

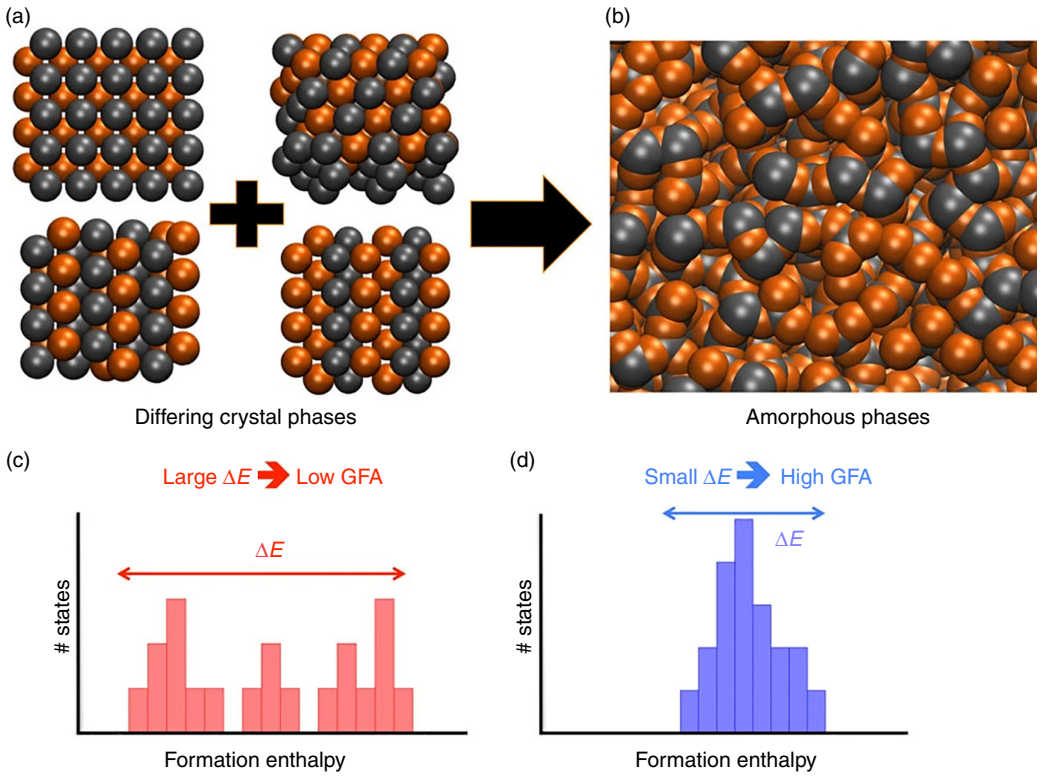

FIG. 7. (a)–(d) Illustration of the concept of "confusion"[58] and its utility in predicting the glass-forming ability (GFA) of metallic systems demonstrated by Curtarolo.[57] The larger the number of structurally dissimilar states at low energy, the more confused the system that phases to crystallize. This slows down the crystallization rate and allows for the glass transition to occur. Reprinted with permission from Perim *et al.*, Nat. Commun. **7**, 12315 (2016). Copyright 2016 Authors, licensed under a Creative Commons Attribution (CC BY) license.[57]

reviews.[66,67] The question most pertinent to the metastability topic covered here is under which circumstances will the disorder be quenchable? The previously discussed concept of "confusion" could also be helpful here. Basically, if a relatively large number of competitive structures or ordering patterns, similar in energy but dissimilar in structure, exist, the likelihood for the quenchable disorder could be expected to increase. As already noted, this concept has been utilized by Curtarolo and colleagues to suggest novel glass formers,[57] but more work is needed to move to a truly quantitative description of the quenchable disorder, including predicting the disorder-freezing (glass transition) temperatures.

IV. LIFETIME: UNDERSTANDING THE TRANSITIONS BETWEEN METASTABLE MATERIALS

This section discusses the dynamic aspect of metastability. For a material to be functional, it must exist on a practical timescale. To determine that timescale—its lifetime—the processes of transformation between the relevant phase and other existing phases must be understood both crystallographically and energetically.

A. Lifetime assessment

The lifetime of a metastable state, what makes it *metastable*, is the average time it takes for it to relax to a lower energy state. It can also be seen as the amount of time a material stays in its functional state; in that sense, it is closely related to chemical degradation.[68] Lifetime is highly variable because phase change is a Markov process,[10] i.e., a probabilistic process where past events have no influence on future events. Imagine a board game where a player has landed on a "jail" square. The jail square is analogous to a metastable state of a material. The player rolls two dice every turn, and the only way for them to get out of jail is to roll a double. Rolling close to a double, getting 5 and 6, for example, has no effect; the *only* way to get out is to roll a double. The trapped player may roll a double on their first turn, and their time in jail would be limited to one turn, but there is no way to predict exactly how long they would stay in jail. Only the average time a player spends in jail can be estimated. Since the number of turns it takes to get out of jail follows a geometric distribution, its expectation value is $1/p$ where p is the probability to roll a double. Therefore, at this game, the average time a player spends in jail is six turns. The same goes for materials: there is no way to predict the exact lifetime of a specific phase; only its expectation value can be estimated as the inverse of the transition probability. In chemistry and materials science, it is common to talk about transition rate, the amount of material that is transformed per unit of time, because the transition rate does not depend on the size of the system. In general, this rate is given by the Arrhenius equation:

$$R = R_0 e^{\frac{-\Delta E}{k_B T}}, \tag{8}$$

where R_0, referred to as the "attempt frequency," is proportional to the frequency at which an event could trigger a phase transition occurs. In the board game analogy, it would be equivalent to the number of times a player can roll the dice at each turn. The rest of the expression represents the probability for each event to trigger a phase transition where ΔE is the activation energy of the phase transition (analogous to the score necessary to get out of jail), k_B is the Boltzmann constant, and T is the temperature. This probability is proportional to the number of particles that have an energy higher than ΔE, hence the Boltzmann term. The Arrhenius equation was obtained empirically, however, by replacing the activation energy ΔE with an activation *free* energy ΔG, and by making R_0 dependent on the temperature, it is equivalent to the Eyring–Polanyi equation, which can be formally derived in the context on transition state theory (TST).[69–71]

This review focuses on condensed matter phase transitions that typically involve nucleation. In these cases, the difference in free energy between a growth nucleus and its untransformed state is the result of two competing terms: the negative difference in free energy of the material as it transforms to the lower energy state and the positive interface energy at the interface between the nucleus and the bulk. This relation is represented in Fig. 8(a). The first term is directly proportional to the number of particles in the nucleus, or its volume, and the second term is proportional to the area of the nucleus:

$$\Delta G(n) = \Delta G_{\text{bulk}}(n) + \Delta G_{\text{surf.}}(n) \tag{9}$$

$$= n\,\Delta g + n^{\frac{2}{3}} B\,\sigma, \tag{10}$$

where n is the number of particles in the growth nucleus; Δg is the difference in Gibbs free energy between the initial phase (α) and the final phase (β); B is a constant independent of n, which depends on the shape of the growth nucleus; and σ is the effective interface energy (different interfaces in different orientation may have different interface energies). This function has a maximum at $n^* = -\left(\frac{2B\sigma}{3\Delta g}\right)^3$, which corresponds to $\Delta G^* = \frac{4B^3\sigma^3}{27\Delta g^2}$ [see Fig. 8(a)]. Because of thermal fluctuations within the material, nuclei of the most stable phase will randomly form. If a nucleus has more than n^* atoms, even if it initially represents an increase in free energy, the nucleus will naturally grow to minimize ΔG. Therefore, for nucleation to start and for the transformation to occur, the growth nucleus must reach this critical size, and thermal fluctuations must be higher than ΔG^*. This is why for condensed matter systems the Eyring–Polanyi equation [analogous to Eq. (8)] becomes

$$R = R_0(T) e^{\frac{-(\Delta G^* + \Delta g^*)}{k_B T}}, \tag{11}$$

referred to as the Turnbull and Fisher equation,[72] where ΔG^* represents the energy barrier for nucleation and Δg^* the activation free energy of the steady-state transformation. In other words, Δg^* is the activation energy for transforming one atom to the β phase once the nucleus size is much larger than n^* [see Fig. 8(b)]. It is defined as the difference between the activated complex and the mean free energy between states α and β:

$$\Delta g^* = \Delta g_{\alpha\to\beta} + \frac{\Delta g}{2}. \tag{12}$$

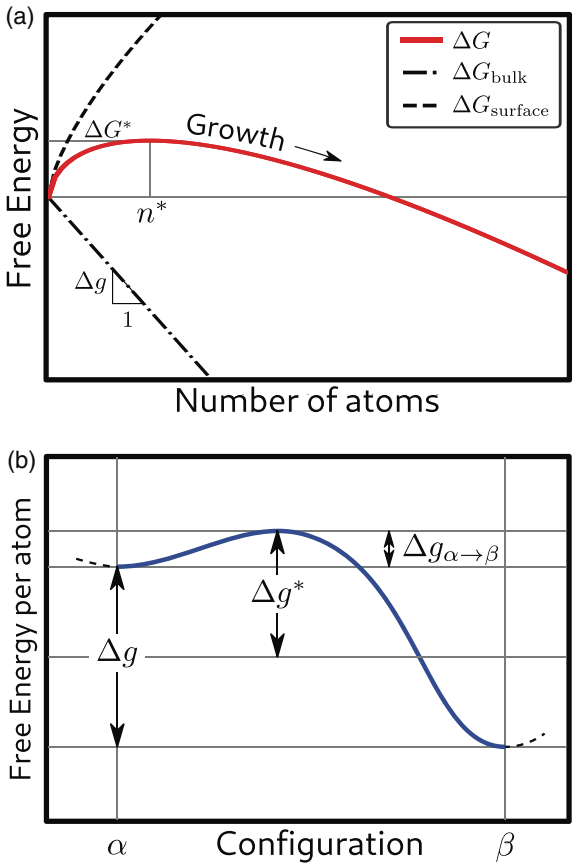

FIG. 8. Energy barriers in condensed matter phase transitions. (a) Difference in free energy as a function of the number of atoms in the nucleus. (b) Free energy along the transformation between phase α and β as a function of the reaction coordinate.

In systems that involve phase separation, Δg^* is the activation energy of diffusion.[73]

In order for Eq. (11) to be useful for *ab initio* calculations, it needs to be rewritten in terms of the enthalpy, which can be readily evaluated with zero temperature methods, such as density functional theory. From the definition of the Gibbs free energy, one can write

$$\Delta g(T) = \Delta h - T\,\Delta s. \tag{13}$$

At the transition temperature T_m, the difference in Gibbs free energy between the two phases is 0 ($\Delta g(T_m) = 0$), and therefore, $\Delta s = \frac{\Delta h}{T_m}$. Equation (13) can be rewritten as

$$\Delta g(T) = \Delta h\left(1 - \frac{T}{T_m}\right). \tag{14}$$

For a metastable material, the temperature is far from the equilibrium temperature $T \ll T_m$, which means $\Delta g = \Delta h$. Finally, replacing Eq. (14) and Eq. (12) into Eq. (11), one gets

$$R = R_0(T)e^{\frac{-\left(\frac{\Delta h}{2}+\Delta g_{\alpha\to\beta}+\frac{4B^3\sigma^3}{27\Delta h^2}\right)}{k_B T}}. \quad (15)$$

This expression contains all the important contributors to the rate of transformation. It is clear that the rate of transformation is determined by three important factors: (1) Δh, the enthalpy difference between the two phases; (2) $g_{\alpha\to\beta}$, the activation energy of the transformation; and (3) σ, the interface energy of the growth nucleus. Factor 1 can easily be calculated; it is equivalent to the scale of metastability. The rest of this section will summarize recent methods for estimating factor 2. Significant efforts must be invested in order to develop approaches for estimating factor 3 in condensed matter systems, as it is lacking in current literature.

B. Matching crystal structures

The first step in evaluating the transformation barrier $\Delta g_{\alpha\to\beta}$ is to establish the correspondence between atoms of the initial and final states. In molecular dynamics simulations where atoms move freely between states, this step is not necessary, but it can be difficult (or impossible) to find the right conditions and potential to trigger a specific transition between two polymorphs; waiting for the transformation to occur is most likely not computationally affordable. Finding the correspondence between two structures, which we will also refer to as "matching crystal structures," is a necessary geometric first step for most methods that calculate the minimum energy path (MEP).

Matching crystal structures, in two dimensions, can be useful to describe interfacial structures, for example, to find a suitable substrate for epitaxial growth of a material.[75–77] Pioneering work on matching crystal structures was done within that area of materials science. Many models were developed to describe and predict interface structures: the O-lattice theory part of the CSL/DSC (Coincidental Site Lattice, Displacement of one crystal Lattice with respect to the second causes a pattern Shift which is Complete) Lattice Model,[78–81] the Edge to Edge Model,[82,83] the Coincidence of Reciprocal Lattice Points (CRLP) model,[84] and the work of Tkatchenko and Batina.[85–87] A notable mathematical method that comes out of interfacial structure theory is the Zur algorithm,[88] an algorithmic method to find the best-matching unit cells between two different 2D lattices. It is the basis of many more recent studies.[74,77,89]

Determining the correspondence between two structures is also very useful to measure similarity or closeness between crystal structures. Sadeghi *et al.*[90] defined a distance between molecules by finding a correspondence between each atom and the orientation that minimized the l_2-norm in configuration space. This method not only can measure the similarity between two molecules but also can establish the correspondence between their atoms, which constitutes a first approximation of the transformation mechanism. Although this method works very well for finite molecular systems, it is not suitable for crystals where the number of atoms is very large ($\approx N_A$). In order to measure similarity (or closeness) between crystals, many methods based on holistic descriptions of the structures (fingerprints) have been developed.[91–95] However, these approaches are not useful to determine the phase transformation mechanism and its associated energy barrier.

Capillas *et al.*[96] were the first to define a systematic procedure for determining transition pathways between arbitrary polymorphs. Their methodology is based on finding the maximal symmetry transition paths. First, they listed the symmetry subgroups $\{\mathcal{H}_{1,i}, i = 1, \ldots, n\}$ and $\{\mathcal{H}_{2,j}, j = 1, \ldots, n\}$ of the space groups of the initial and final structures $\mathcal{G}_1$ and $\mathcal{G}_2$. Then, for each pair of subgroups $(\mathcal{H}_{1,i}, \mathcal{H}_{2,j})$, they verified that (1) the two groups are isomorphic, (2) they meet the condition of a continuous transformation (the number of formula units per unit cell is the same when described by their subgroup), and (3) the occupied Wyckoff positions are the same under the symmetry-allowed permutations. If all three conditions are met and "there exists no pair of intermediate subgroups $(\mathcal{Z}_1, \mathcal{Z}_2)$, such that $\mathcal{H}_i < \mathcal{Z}_i < \mathcal{G}_1$[97] with i=1, 2," that also meet the three conditions, then the pair of subgroups is said to be of *maximal symmetry*. Finally, the unit cells of the two subgroups are matched together in order to minimize a metric distance based on their parameters (a,b,c, α,β,γ), and the mapping that minimizes the atomic displacements within the cells is chosen. This method produces multiple plausible paths for each transformation. Those paths have been shown to include previous theoretical descriptions of the same systems.

Stevanović *et al.*[74] took a slightly different approach to solving this problem that did not rely on symmetry. It comprises two major steps. In the first step, all possible supercells of the two end structures that meet the condition for a continuous transformation are enumerated using Hart–Forcade theory,[98] and the pair of supercells that minimizes a metric distance similar to the one used by Capillas *et al.*[96] is chosen. This first step is, in some way, a generalization of the Zur algorithm in three dimensions. The next step is to map the atoms inside the unit cell using the same method as Sadeghi *et al.*[90] The key idea is to use the Kuhn–Munkres[99] algorithm to find the optimal map between the two sets of atoms (initial and final). At the end of the two steps, the periodic movement of the atoms between the two polymorphs—the transformation pathway—is fully described. From there, an estimate of the activation energy can easily be calculated using density functional theory, computing the total energy at each step along the transformation. Figure 9 shows the computed mechanism and corresponding energy along the path for four transitions between polymorphs of SnO_2. Pathways obtained from the algorithm produce the red energy profiles, whereas pathways obtained from subsequent energy minimization using solid-state nudged elastic band (ss-NEB) are shown in gray (more details in Sec. IV C).

It is important to note that the energy barrier given by direct DFT calculation is an upper bound to that value; further minimization using NEB is necessary to obtain a more accurate estimate of $\Delta g_{\alpha\to\beta}$. However, for a more qualitative assessment of the transformation kinetics, inspired by Buerger's work,[100] Stevanović *et al.*[74] found that pathways where the coordination is never lower than in the two end phases lead to fast kinetics. In other words, when the number of bonds per atom increases or decreases monotonically during the transformation, as seen at the bottom of Fig. 6(a), the rate of transformation is high. In the same vein, a recent study[101] treated structures as networks to account for the topological transformations during the phase transition. This method naturally identifies transition pathways that minimize the number of broken bonds.

Both Refs. 74 and 96 are hindered by the fact that it is difficult to compare the effect of strain and atomic displacements within the unit cell; the quantity to minimize is ill-defined. Moreover, since an infinite number of unit cells (with an infinite number of sizes) can represent a single system, matching the unit cells—or their associated space groups for that matter—is a difficult problem. In fact, even identifying

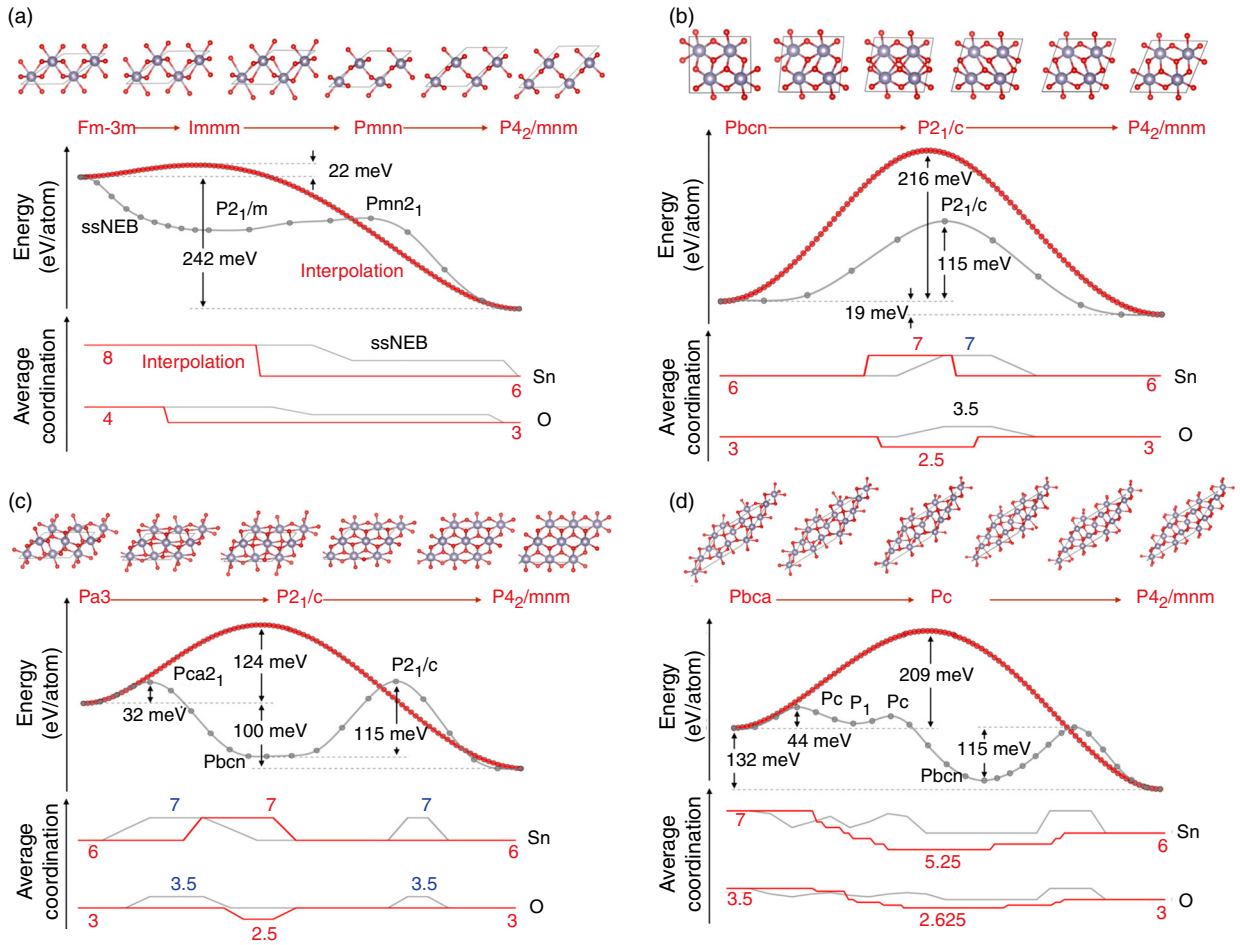

FIG. 9. (a)–(d) Calculated energy profiles (shown in red) for the four selected polymorphic transformations between Sn and O together with the space group symmetry, six crystal structure snapshots along the pathway, and the average coordination of Sn and O. The corresponding solid-state nudged elastic band results are shown in gray. The x axis represents the normalized reaction coordinate. Reprinted with permission from Stevanović *et al.*, Phys. Rev. Mater. **2**, 033802 (2018). Copyright 2018 American Physical Society.[74]

identical structures requires complex algorithms.[103] Figure 10 uses a two-dimensional example to demonstrate how matching based on the unit cell can lead to results with a substantially larger total distance traveled by the atoms.

To circumvent these issues, Stevanović *et al.*[102] mathematically stated the problem of finding an optimal correspondence between two crystal structures: Given the positions of all the atoms in the two crystals $\{\vec{a}_i | i = 1, \ldots, N\}$ and $\{\vec{b}_j | i = 1, \ldots, N\}$, the optimal match is the best atom-to-atom mapping p_{min} (permutation of atomic indices) and the best alignment of the two structures (linear transformation R_{min} and translation $\vec{t}_{min}$) that minimize a given distance (cost) function d. This is equivalent to the following equation when $N \to \infty$:

$$p_{min}, Q_{min}, \vec{t}_{min} = \underset{p,Q,\vec{t}}{\operatorname{argmin}}\ d_1(p, Q, \vec{t}), \tag{16}$$

with

$$d_1(p, Q, \vec{t}) = \sum_i^N d(\ \vec{a}_i\ ,\ Q\vec{b}_{p(i)} + \vec{t}). \tag{17}$$

In the case of phase transitions, Q_{min} is a rotation matrix and d is the Euclidean distance. The authors proposed an improved approach to finding transformation pathways based on numerically solving Eq. (16) for a finite, but large N. In their new method, the first step is to cut a large finite portion of both the initial and final crystal structures to obtain two sets of points. Then, an initial mapping (i.e., the correspondence between each atom) is established between the two sets using the same Kuhn–Munkres algorithm. After the correspondence is established, since the sum of distances between each pair of atoms (d_1) is now well defined, it is minimized with respect to the position of the final structure. The mapping and distance are updated simultaneously until they both converge to a minimal d_1. Since the initial alignment between the structures is random, the mapping and distance minimization are repeated several times such that the global minimum can be found. Finally, once the minimal distance is found, the algorithm retrieves the periodicity in the final mapping to obtain the unit cell of the "displacement crystal." This unit cell contains all the necessary information to fully describe the transformation, and energy profiles can be calculated as in Fig. 9. This method has been

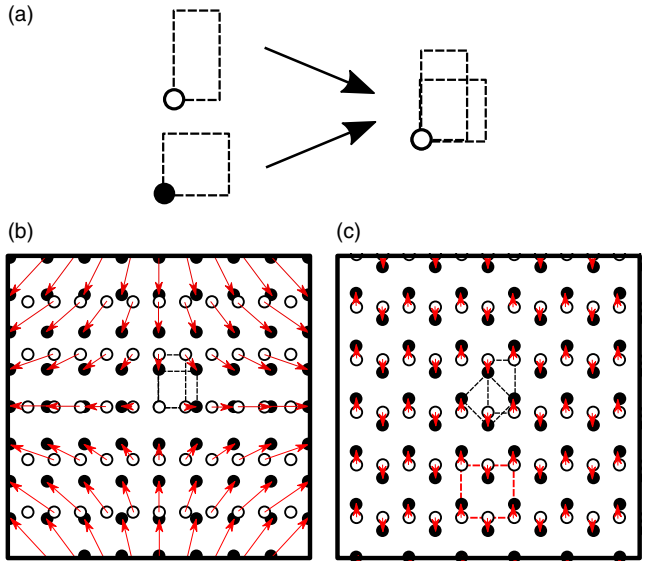

FIG. 10. Drawbacks of a method that relies on matching some choice of the unit cells of the two crystals. (a) Visual representation of the matching inside one cell for a simple 2D example where the cells have the same area. (b) The two overlaid crystal structures using the same matching. (c) The same crystal structures overlaid such that the distance (red arrows) does not increase away from the center; the red dashed rectangle shows the scale (periodic unit) of the match. Reprinted with permission from Therrien *et al.*, J. Chem. Phys. **152**, 074106 (2020). Copyright 2020 AIP Publishing.[102]

shown to produce a realistic transformation mechanism[104] without the need for energy-based *ab initio* methods.

C. Minimal energy pathways

Once the correspondence between the atoms has been established, the MEP can be determined. Most methods to find the MEP focus on finite systems, and they cannot be applied to solid-state systems with periodic boundary conditions. For detailed accounts of these methods, see Refs. 105 and 106.

The most widely used method for solid-solid transformation is the NEB method and its variants.[105,108–110] The concept behind NEB is relatively simple. Given two structures (initial and final) that are represented as points in configuration space as in Fig. 11, the initial path between the two structures forms a straight line in configuration space. Each point along that line represents an intermediate structure at a different stage of the transformation. One can evaluate the energy of each of these intermediate structures, using density functional theory, for example, and obtain an energy profile much like the one on Fig. 9. If each of these structures were simply relaxed to its closest minimum in energy, it would relax to either the initial or the final state (at least in Fig. 11 where they are separated by one saddle point). The NEB key idea is to separate each point by an "elastic band" that will keep them from collapsing onto each other but will let them relax toward the MEP.

In practice, the force F_i acting on one structure $\mathbf{R}_i$ (point in N-dimensional configuration space referred to as "image") is

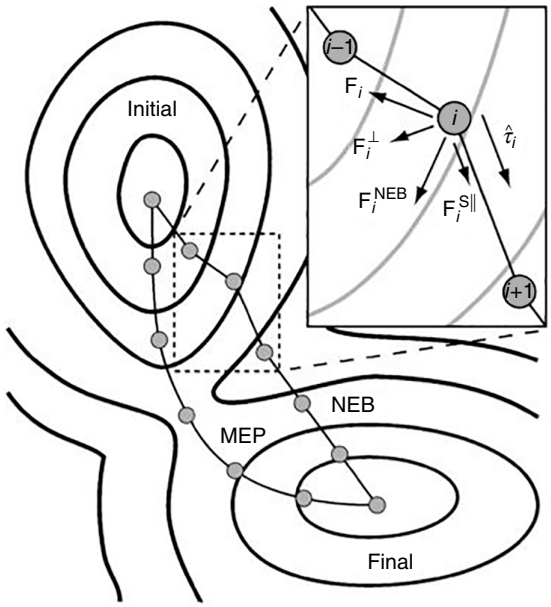

FIG. 11. Schematic representation of the nudged elastic band method. Reprinted with permission from Sheppard *et al.*, J. Chem. Phys. **128**, 134106 (2008). Copyright 2008 AIP Publishing.[107]

$$\mathbf{F}_i = -\nabla E(\mathbf{R}_i)_\perp + \mathbf{F}^s_{i\parallel}, \tag{18}$$

with

$$\nabla E(\mathbf{R}_i)_\perp = \nabla E(\mathbf{R}_i) - (\nabla E(\mathbf{R}_i)\cdot\hat{\boldsymbol{\tau}}_i)\hat{\boldsymbol{\tau}}_i \tag{19}$$

and

$$\mathbf{F}^s_{i\parallel} = k(|\mathbf{R}_{i+1}-\mathbf{R}_i| - |\mathbf{R}_i-\mathbf{R}_{i-1}|)\hat{\boldsymbol{\tau}}_i. \tag{20}$$

Notice that the total spring force $\mathbf{F}^s_{i\parallel}$ is parallel to the tangent vector $\hat{\boldsymbol{\tau}}_i$ and that the potential force $\nabla E(\mathbf{R}_i)_\perp$ is always perpendicular to it. This is the case to make sure that the elastic bands do not have an influence on the final result (it avoids corner cutting). The tangent direction $\hat{\boldsymbol{\tau}}_i$ is in the direction of the highest energy. To find the lowest energy path, one must simply minimize the forces using any minimization method, such as the Leapfrog or Verlet algorithm.

In climbing image nudged elastic band (CI-NEB),[105] which is included in most modern implementations of NEB, the highest energy image is set to move toward higher energies along the path such that there is always an image exactly at a maximum in energy along the MEP. The force for that image is given by the following equation:

$$\mathbf{F}_i = -\nabla E(\mathbf{R}_i)_\perp + (\nabla E(\mathbf{R}_i)\cdot\hat{\boldsymbol{\tau}}_i)\hat{\boldsymbol{\tau}}_i. \tag{21}$$

Not only does this provide a better estimate of the activation energy but also the frequencies of the normal vibrational modes (in configuration space) of the saddle point can be used to evaluate R_0 in Eq. (8).[111,112]

These methods (CI-NEB and NEB) cannot be used directly to find the MEP in solid–solid transformation as these systems involve an infinite number of atoms. In fact, because crystals are periodic, they can be represented either as an infinite set of coordinates (which is

obviously impractical) or as a finite set of coordinates along with three unit cell vectors that define periodic boundary conditions. The generalized ss-NEB[108] takes advantage of the latter representation to extend the capabilities of NEB to periodic systems. In that implementation of NEB, the unit cell vectors are treated as additional degrees of freedom, and the quantity to minimize is the stress ($\underline{\sigma}$) instead of force. All degrees of freedom can be optimized simultaneously using the following definitions:

$$\Delta\mathcal{R}_i = \left\{J_i\underline{\varepsilon_i}, \mathbf{R}_i - \mathbf{R}_{i-1}\right\} \quad \text{and} \quad \mathcal{F}_i = \left\{\frac{-\Omega_i\underline{\sigma_i}}{J_i}, \mathbf{F}_i\right\}, \tag{22}$$

where $J_i = \omega_i^{1/3}N_i^{1/6}$ is a Jacobian, Ω_i is the volume of the unit cell, N_i is the number of atoms per unit cell, and $\underline{\varepsilon_i}$ is the mechanical strain matrix. The role of the Jacobian is to make sure that all degrees of freedom have consistent units and can be simultaneously minimized. In Fig. 9, solid-state nudged elastic band has been used to compute the minimal energy path shown in gray. Note that the energy barrier is always lower than that of the initial pathway in red and that there is always a point (image) at maxima along the path.

Other methods have been developed for systems with periodic boundary conditions that aim to directly find saddle points on the potential energy surface of periodic systems.[113,114] These methods do not require prior knowledge of the transition path—crystal structure matching is not necessary—but they cannot be used to study specific phase transitions.

Phase transition sampling via swarm intelligence is a promising way to find minimum energy pathways that combine climbing image, fingerprinting, graph theory, and particle swarm optimization to directly find minimum energy pathways.[115] It does not explicitly require structure matching, although it internally performs structure matching using a fingerprint distance between saddle points of the PES.

Transition path sampling (TPS) is a way to probe the PES in molecular dynamics that circumvents the rare event problem mentioned previously by creating a statistical path ensemble through a Monte Carlo exploration.[116,117] This methodology requires an initial path that is modified to obtain the most statistically relevant paths. It has been shown to be efficient to model nanocrystalline systems,[118] yet little work has been done to make the connection between structure matching methods and TPS in crystals.

V. CONCLUSION AND CURRENT CHALLENGES

This review has discussed the three major steps of metastable materials discovery as illustrated in Fig. 1: exploration, realizability, and lifetime. As explained in Sec. II, the exploration step is by far the one for which the theory is the most advanced; an abundance of computational methods to explore the potential energy surface have been shown to reliably predict material properties and successfully identify candidate materials. Even though these methods can still be improved, they do not constitute a bottleneck for metastable materials discovery. Readers interested in the remaining challenges of structure prediction may refer to Refs. 20, 89, and 120.

Section III described two approaches to predict the realizability of hypothetical structures: (1) computing the scale of metastability (the energy above the ground state) or (2) calculating probabilities of occurrence through random structure sampling. The former is efficient at excluding higher energy structures but is unable to identify realizable hypothetical structures within the energy range of observed polymorphs. The concept of remnant metastability might be able to solve this problem, but work is yet to be done to tie this concept to existing theories that can predict the free energy under various thermodynamic conditions. The question as to *why* the energy above the ground state is a physically relevant scale of metastability also needs to be clearly answered. The second methodology is able to pinpoint which structures are most likely to be realized and even *how* likely they are to exist. However, because it combines the PES exploration and the realizability assessment (steps 1 and 2 in Fig. 1), it remains relatively computationally expensive and cannot be used in concert with other PES exploration methods. The probability of occurrence of a hypothetical structure resulting from, say, an evolutionary algorithm cannot be evaluated if it does not naturally occur during random sampling. Combining the probability of occurrence formalism presented in Sec. III B with other PES exploration schemes might be an interesting research avenue. When it comes to exploring metastable phases where the composition is not fixed, the probability of occurrence formalism would be required to perform random structure sampling at each possible composition and would not be able to take into account the likelihood of phase separation. A theory that could predict realizability across compositions would be an invaluable tool for materials discovery.

In Sec. IV, several methods have been presented to find the transition mechanism and precisely calculate its associated activation energy Δg^*. However, in comparison, only a few studies have focused on estimating the critical nucleation energy ΔG^* from first principles, even though the growth conditions within the transformed lattice can have a significant impact on the actual mechanism of transformation, especially when Δg^* is small. Therrien and Stevanović[104] showed that when considering the bulk martensitic transformation, without any nucleation, the Bain path has the lowest activation energy Δg^*, despite not being experimentally observed. However, when considering the interface between the transformed and untransformed lattice, the Bain path becomes energetically unfavorable compared to an alternative path that is consistent with experimental observations. In the same vein, Khaliullin *et al.*[119] used molecular dynamics paired with machine learning to simulate the nucleation process of diamond in graphite. Figure 12 shows the total enthalpy that they obtained as a function of the size of the nucleus [analogous to Fig. 8(a)] at different pressures. The nucleation barriers ΔG^* at pressures where the transformation occurs are of the order of 100 eV, whereas for the same system, transformation free energies per atom Δg^* are of the order of 200 meV, a difference of nearly three orders of magnitude. This shows the importance of nucleation on metastable materials lifetime and stresses the need for a predictive *ab initio* nucleation theory. The effect of surfaces on phase transitions is particularly important in nanoscale materials, where phase diagrams have been shown to depend strongly on particle size[121] or where surface energy can stabilize phases that would not be stable otherwise.[122] For more details on nanoscale materials and the effect of interfaces on metastability, see Refs. 123 and 124. Another interesting avenue would be to apply TPS to crystal structures using transformation pathways from Sec. IV B as a starting point. This would be analogous to applying the ensemble-averaging approach presented in Sec. III B to the transition paths themselves. Once again, the methods presented in Sec. IV have mostly been applied to

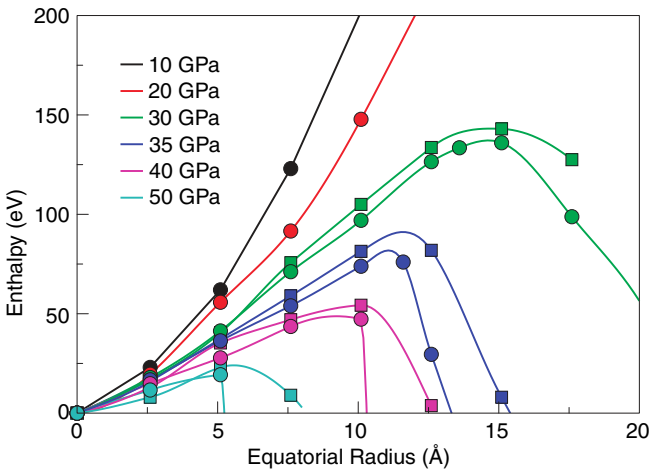

FIG. 12. Pressure dependence of the nucleation barriers for the rhombohedral graphite → cubic diamond (circles) and hexagonal graphite → hexagonal diamond (squares) transformations. Reprinted with permission from Khaliullin *et al.*, Nat. Mater. **10**, 693–697 (2011). Copyright 2011 Springer Nature.[119]

polymorphic transformations; evaluating the activation energy of diffusion when phase separation occurs remains a challenge.

In conclusion, remarkable progress has been made toward fully predictive metastable materials discovery. Challenges remain in assessing realizability and lifetime, but in many cases, the tools exist and if used in concert, could already lead to important discoveries.

ACKNOWLEDGMENTS

This work is supported by the National Science Foundation (grant no. DMR-1945010).

DATA AVAILABILITY

No new data were generated to support the findings.

REFERENCES

[1]F. P. Bundy, "The p, t phase and reaction diagram for elemental carbon, 1979," J. Geophysical Res. Solid Earth **85**, 6930–6936 (1980).
[2]J. E. Shelby, *Introduction to Glass Science and Technology* (Royal Society of Chemistry, 2005).
[3]I. Gutzow, *The Vitreous State: Thermodynamics, Structure, Rheology, and Crystallization* (Springer, 2013).
[4]N. Garti, *Cocoa Butter and Related Compounds* (AOCS Press, 2012).
[5]L. Bayés-García, T. Calvet, M. À. Cuevas-Diarte, E. Rovira, S. Ueno, and K. Sato, "New textures of chocolate are formed by polymorphic crystallization and template effects: Velvet chocolate," Cryst. Growth Des. **15**, 4045–4054 (2015).
[6]A. R. Verma and P. Krishna, *Polymorphism and Polytypism in Crystals* (John Wiley & Sons, Inc., 1966).
[7]A. R. Verma, P. Krishna, and H. M. Otte, "Polymorphism and polytypism in crystals," Phys. Today **20**(3), 111 (1967).
[8]D. Wales, R. Saykally, A. Zewail, and D. King, *Energy Landscapes With Applications to Clusters, Biomolecules and Glasses* (Cambridge University Press, 2003).
[9]F. Stillinger, *Energy Landscapes, Inherent Structures, and Condensed-Matter Phenomena* (Princeton University Press, 2015).
[10]A. Bovier and F. den Hollander, *Metastability: A Potential-Theoretic Approach* (Springer International Publishing, 2015).
[11]Systems with strongly correlated electrons usually found among transition metal and rare-earth chemistries are probably not yet there.
[12]A. E. Oganov, *Modern Methods of Crystal Structure Prediction* (Wiley-VCH John Wiley distributor, 2010).
[13]C. Evrenk and A. E. Aspuru-Guzik, *Prediction and Calculation of Crystal Structures: Methods and Applications* (Springer, 2014).
[14]A. Y. Liu and M. L. Cohen, "Prediction of new low compressibility solids," Science **245**, 841–842 (1989).
[15]A. Franceschetti and A. Zunger, "The inverse band-structure problem of finding an atomic configuration with given electronic properties," Nature **402**, 60–63 (1999).
[16]M. d'Avezac, J.-W. Luo, T. Chanier, and A. Zunger, "Genetic-algorithm discovery of a direct-gap and optically allowed superstructure from indirect-gap Si and Ge semiconductors," Phys. Rev. Lett. **108**, 027401 (2012).
[17]J. Maddox, "Crystals from first principles," Nature **335**, 201–201 (1988).
[18]J. Pannetier, J. Bassas-Alsina, J. Rodriguez-Carvajal, and V. Caignaert, "Prediction of crystal structures from crystal chemistry rules by simulated annealing," Nature **346**, 343–345 (1990).
[19]J. C. Schön and M. Jansen, "First step towards planning of syntheses in solid-state chemistry: Determination of promising structure candidates by global optimization," Angew. Chem. Int. Ed. English **35**, 1286–1304 (1996).
[20]A. R. Oganov and C. W. Glass, "Crystal structure prediction using ab initio evolutionary techniques: Principles and applications," J. Chem. Phys. **124**, 244704 (2006).
[21]G. Trimarchi and A. Zunger, "Global space-group optimization problem: Finding the stablest crystal structure without constraints," Phys. Rev. B **75**, 104113 (2007).
[22]B. Meredig and C. Wolverton, "A hybrid computational–experimental approach for automated crystal structure solution," Nat. Mater. **12**, 123–127 (2013).
[23]C. C. Fischer, K. J. Tibbetts, D. Morgan, and G. Ceder, "Predicting crystal structure by merging data mining with quantum mechanics," Nat. Mater. **5**, 641–646 (2006).
[24]G. Hautier, C. Fischer, V. Ehrlacher, A. Jain, and G. Ceder, "Data mined ionic substitutions for the discovery of new compounds," Inorg. Chem. **50**, 656–663 (2011).
[25]X. Zhang, V. Stevanović, M. d'Avezac, S. Lany, and A. Zunger, "Prediction of A_2bX_4 metal-chalcogenide compounds via first-principles thermodynamics," Phys. Rev. B **86**, 014109 (2012).
[26]R. Gautier, X. Zhang, L. Hu, L. Yu, Y. Lin, T. O. L. Sunde, D. Chon, K. R. Poeppelmeier, and A. Zunger, "Prediction and accelerated laboratory discovery of previously unknown 18-electron abx compounds," Nat. Chem. **7**, 308–316 (2015).
[27]A. Laio and M. Parrinello, "Escaping free-energy minima," Proc. Nat. Acad. Sci. **99**, 12562–12566 (2002).
[28]A. Barducci, M. Bonomi, and M. Parrinello, "Metadynamics," WIREs Comput. Mol. Sci. **1**, 826–843 (2011).
[29]D. J. Wales and J. P. K. Doye, "Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms," J. Phys. Chem. A **101**, 5111–5116 (1997).
[30]S. Goedecker, "Minima hopping: An efficient search method for the global minimum of the potential energy surface of complex molecular systems," J. Chem. Phys. **120**, 9911–9917 (2004).
[31]C. J. Pickard and R. J. Needs, "Ab initio random structure searching," J. Phys. Condens. Matter **23**, 053201 (2011).
[32]Z. Wang, J. Ha, Y. H. Kim, W. B. Im, J. McKittrick, and S. P. Ong, "Mining unexplored chemistries for phosphors for high-color-quality white-light-emitting diodes," Joule **2**, 914–926 (2018).
[33]G. Hautier, A. Jain, S. P. Ong, B. Kang, C. Moore, R. Doe, and G. Ceder, "Phosphates as lithium-ion battery cathodes: An evaluation based on high-throughput ab initio calculations," Chem. Mater. **23**, 3495–3508 (2011).
[34]Y. Ma, M. Eremets, A. R. Oganov, Y. Xie, I. Trojan, S. Medvedev, A. O. Lyakhov, M. Valle, and V. Prakapenka, "Transparent dense sodium," Nature **458**, 182–185 (2009).

[35]D. Duan, Y. Liu, F. Tian, D. Li, X. Huang, Z. Zhao, H. Yu, B. Liu, W. Tian, and T. Cui, "Pressure-induced metallization of dense $(H_2S)_2H_2$ with high-T_c superconductivity," Sci. Rep. **4**, 6968 (2014).

[36]A. Jain, Y. Shin, and K. A. Persson, "Computational predictions of energy materials using density functional theory," Nat. Rev. Mater. **1**, 15004 (2016).

[37]A. R. Oganov, C. J. Pickard, Q. Zhu, and R. J. Needs, "Structure prediction drives materials discovery," Nat. Rev. Mater. **4**, 331–348 (2019).

[38]A. R. Oganov, J. Chen, C. Gatti, Y. Ma, Y. Ma, C. W. Glass, Z. Liu, T. Yu, O. O. Kurakevych, and V. L. Solozhenko, "Ionic high-pressure form of elemental boron," Nature **457**, 863–867 (2009).

[39]S. Botti, J. A. Flores-Livas, M. Amsler, S. Goedecker, and M. A. L. Marques, "Low-energy silicon allotropes with strong absorption in the visible for photovoltaic applications," Phys. Rev. B **86**, 121204 (2012).

[40]P. Biswas, D. N. Tafen, F. Inam, B. Cai, and D. A. Drabold, "Materials modeling by design: Applications to amorphous solids," J. Phys.: Condens. Matter **21**, 084207 (2009).

[41]W. Sun, S. T. Dacek, S. P. Ong, G. Hautier, A. Jain, W. D. Richards, A. C. Gamst, K. A. Persson, and G. Ceder, "The thermodynamic scale of inorganic crystalline metastability," Sci. Adv. **2**, e1600225 (2016).

[42]Strictly speaking, it is the shortest set of connected straight lines that have an energy lower than or equal to all possible crystalline compounds along the composition axis and which forms a convex set together with the horizontal line joining the pure compounds.

[43]C. Oses, E. Gossett, D. Hicks, F. Rose, M. J. Mehl, E. Perim, I. Takeuchi, S. Sanvito, M. Scheffler, Y. Lederer *et al.*, "AFLOW-CHULL: Cloud-oriented platform for autonomous phase stability analysis," J. Chem. Inf. Model. **58**, 2477–2490 (2018).

[44]A. Anelli, E. A. Engel, C. J. Pickard, and M. Ceriotti, "Generalized convex hull construction for materials discovery," Phys. Rev. Mater. **2**, 103804 (2018).

[45]J.-H. Pöhls, M. Heyberger, and A. Mar, "Comparison of computational and experimental inorganic crystal structures," J. Solid State Chem. **290**, 121557 (2020).

[46]S. Srinivasan, R. Batra, D. Luo, T. Loeffler, S. Manna, H. Chan, L. Yang, W. Yang, J. Wen, P. Darancet *et al.*, "Machine learning the metastable phase diagram of materials," arXiv:2004.08753 (2020).

[47]M. Aykol, S. S. Dwaraknath, W. Sun, and K. A. Persson, "Thermodynamic limit for synthesis of metastable inorganic materials," Sci. Adv. **4**, eaaq0148 (2018).

[48]W. Johnson, "Crystal-to-glass transformation in metallic materials," Mater. Sci. Eng. **97**, 1–13 (1988).

[49]J. Yong, Y. Jiang, D. Usanmaz, S. Curtarolo, X. Zhang, L. Li, X. Pan, J. Shin, I. Takeuchi, and R. L. Greene, "Robust topological surface state in kondo insulator SmB_6 thin films," Appl. Phys. Lett. **105**, 222403 (2014).

[50]G. L. W. Hart, S. Curtarolo, T. B. Massalski, and O. Levy, "Comprehensive search for new phases and compounds in binary alloy systems based on platinum-group metals, using a computational first-principles approach," Phys. Rev. X **3**, 041035 (2013).

[51]V. Stevanović, "Sampling polymorphs of ionic solids using random superlattices," Phys. Rev. Lett. **116**, 075503 (2016).

[52]E. B. Jones and V. Stevanović, "Polymorphism in elemental silicon: Probabilistic interpretation of the realizability of metastable structures," Phys. Rev. B **96**, 184101 (2017).

[53]S. De, B. Schaefer, A. Sadeghi, M. Sicher, D. Kanhere, and S. Goedecker, "Relation between the dynamics of glassy clusters and characteristic features of their energy landscape," Phys. Rev. Lett. **112**, 083401 (2014).

[54]X. Yang, J. Wang, J. Zheng, M. Guo, and R.-Z. Zhang, "Screening for planar carbon allotropes using structure space sampling," J. Phys. Chem. C **124**, 6379–6384 (2020).

[55]E. B. Jones and V. Stevanović, "The glassy solid as a statistical ensemble of crystalline microstates," npj Comput. Mater. **6**, 1–6 (2020).

[56]R. Zallen, *The Physics of Amorphous Solids* (John Wiley & Sons, 2008).

[57]E. Perim, D. Lee, Y. Liu, C. Toher, P. Gong, Y. Li, W. N. Simmons, O. Levy, J. J. Vlassak, J. Schroers, and S. Curtarolo, "Spectral descriptors for bulk metallic glasses based on the thermodynamics of competing crystalline phases," Nat. Commun. **7**, 12315 (2016).

[58]A. L. Greer, "Confusion by design," Nature **366**, 303–304 (1993).

[59]D. C. Ford, D. Hicks, C. Oses, C. Toher, and S. Curtarolo, "Metallic glasses for biodegradable implants," Acta Mater. **176**, 297–305 (2019).

[60]C. Toher, C. Oses, D. Hicks, and S. Curtarolo, "Unavoidable disorder and entropy in multi-component systems," npj Comput. Mater. **5**, 69 (2019).

[61]Z. Nishiyama, *Martensitic Transformation* (Elsevier, 2012).

[62]V. Stevanović, M. d'Avezac, and A. Zunger, "Universal electrostatic origin of cation ordering in A_2BO_4 spinel oxides," J. Am. Chem. Soc. **133**, 11649–11654 (2011).

[63]V. Stevanović, M. d'Avezac, and A. Zunger, "Simple point-ion electrostatic model explains the cation distribution in spinel oxides," Phys. Rev. Lett. **105**, 075501 (2010).

[64]S. Lany, A. N. Fioretti, P. P. Zawadzki, L. T. Schelhas, E. S. Toberer, A. Zakutayev, and A. C. Tamboli, "Monte Carlo simulations of disorder in ZnSnN_2 and the effects on the electronic structure," Phys. Rev. Mater. **1**, 035401 (2017).

[65]J. J. Cordell, J. Pan, A. C. Tamboli, G. J. Tucker, and S. Lany, "Probing configurational disorder in $ZnGeN_2$ using cluster-based Monte Carlo," Phys. Rev. Mater. **5**, 024604 (2021).

[66]E. P. George, D. Raabe, and R. O. Ritchie, "High-entropy alloys," Nat. Rev. Mater. **4**, 515–534 (2019).

[67]C. Oses, C. Toher, and S. Curtarolo, "High-entropy ceramics," Nat. Rev. Mater. **5**, 295–309 (2020).

[68]A. K. Singh, L. Zhou, A. Shinde, S. K. Suram, J. H. Montoya, D. Winston, J. M. Gregoire, and K. A. Persson, "Electrochemical stability of metastable materials," Chem. Mater. **29**, 10159–10167 (2017).

[69]H. Eyring, "The activated complex in chemical reactions," J. Chem. Phys. **3**, 107–115 (1935).

[70]M. G. Evans and M. Polanyi, "Some applications of the transition state method to the calculation of reaction velocities, especially in solution," Trans. Faraday Soc. **31**, 875–894 (1935).

[71]D. G. Truhlar, B. C. Garrett, and S. J. Klippenstein, "Current status of transition-state theory," J. Phys. Chem. **100**, 12771–12800 (1996).

[72]D. Turnbull and J. C. Fisher, "Rate of nucleation in condensed systems," J. Chem. Phys. **17**, 71–73 (1949).

[73]S. Vyazovkin, "Modern isoconversional kinetics: From misconceptions to advances," in *Handbook of Thermal Analysis and Calorimetry* (Elsevier, 2018), Vol. 6, pp. 131–172.

[74]V. Stevanović, R. Trottier, C. Musgrave, F. Therrien, A. Holder, and P. Graf, "Predicting kinetics of polymorphic transformations from structure mapping and coordination analysis," Phys. Rev. Mater. **2**, 033802 (2018).

[75]H. Brune, *Surface and Interface Science* (John Wiley & Sons, 2014), pp. 421–492.

[76]K. R. Poeppelmeier and J. M. Rondinelli, "Mismatched lattices patched up," Nat. Chem. **8**, 292–294 (2016).

[77]H. Ding, S. S. Dwaraknath, L. Garten, P. Ndione, D. Ginley, and K. A. Persson, "Computational approach for epitaxial polymorph stabilization through substrate selection," ACS Appl. Mater. Interfaces **8**, 13086–13093 (2016).

[78]W. Bollmann, *Crystal Defects and Crystalline Interfaces*, 1st ed. (Springer-Verlag, 1970), pp. 83–97.

[79]W. Bollmann, "O-lattice calculation of an F.C.C.-B.C.C. interface," Phys. Status Solidi (a) **21**, 543–550 (1974).

[80]D. A. Smith and R. C. Pond, "Bollmann's 0-lattice theory; a geometrical approach to interface structure," Int. Met. Rev. **21**, 61–74 (1976).

[81]R. Balluffi, A. Brokman, and A. King, "CSL/DSC lattice model for general crystalcrystal boundaries and their line defects," Acta Metall. **30**, 1453–1470 (1982).

[82]M.-X. Zhang and P. Kelly, "Crystallography and morphology of Widmanstätten cementite in austenite," Acta Mater. **46**, 4617–4628 (1998).

[83]M. Zhang, P. Kelly, M. Easton, and J. Taylor, "Crystallographic study of grain refinement in aluminum alloys using the edge-to-edge matching model," Acta Mater. **53**, 1427–1438 (2005).

[84]Y. Ikuhara and P. Pirouz, "Orientation relationship in large mismatched bicrystals and coincidence of reciprocal lattice points (CRLP)," Mater. Sci. Forum **207-209**, 121–124 (1996).

[85]A. Tkatchenko and N. Batina, "Unequal-sphere packing model for the structural arrangement of the well-ordered adsorbate-substrate system," Phys. Rev. B **70**, 195403 (2004).

[86]A. Tkatchenko and N. Batina, "Unequal-sphere packing model for simulation of the uniaxially compressed iodine adlayer on Au," J. Phys. Chem. B **109**, 21710–21715 (2005).
[87]A. Tkatchenko and N. Batina, "Classification of hexagonal adlayer arrangements by means of collective geometrical properties," J. Chem. Phys. **125**, 164702 (2006).
[88]A. Zur and T. C. McGill, "Lattice match: An application to heteroepitaxy," J. Appl. Phys. **55**, 378–386 (1984).
[89]K. Mathew, A. K. Singh, J. J. Gabriel, K. Choudhary, S. B. Sinnott, A. V. Davydov, F. Tavazza, and R. G. Hennig, "MPInterfaces: A materials project based python tool for high-throughput computational screening of interfacial systems," Comput. Mater. Sci. **122**, 183–190 (2016).
[90]A. Sadeghi, S. A. Ghasemi, B. Schaefer, S. Mohr, M. A. Lill, and S. Goedecker, "Metrics for measuring distances in configuration spaces," J. Chem. Phys. **139**, 184118 (2013).
[91]A. R. Oganov and M. Valle, "How to quantify energy landscapes of solids," J. Chem. Phys. **130**, 104504 (2009).
[92]A. P. Bartók, R. Kondor, and G. Csányi, "On representing chemical environments," Phys. Rev. B **87**, 184115 (2013).
[93]L. Yang, S. Dacek, and G. Ceder, "Proposed definition of crystal substructure and substructural similarity," Phys. Rev. B **90**, 054102 (2014).
[94]S. De, A. P. Bartók, G. Csányi, and M. Ceriotti, "Comparing molecules and solids across structural and alchemical space," Phys. Chem. Chem. Phys. **18**, 13754–13769 (2016).
[95]L. Zhu, M. Amsler, T. Fuhrer, B. Schaefer, S. Faraji, S. Rostami, S. A. Ghasemi, A. Sadeghi, M. Grauzinyte, C. Wolverton, *et al.*, "A fingerprint based metric for measuring similarities of crystalline structures," J. Chem. Phys. **144**, 034203 (2016).
[96]C. Capillas, J. Perez-Mato, and M. Aroyo, "Maximal symmetry transition paths for reconstructive phase transitions," J. Phys.: Condens. Matter **19**, 275203 (2007).
[97]< reads "is a subgroup of."
[98]G. L. W. Hart and R. W. Forcade, "Generating derivative structures from multilattices: Algorithm and application to hcp alloys," Phys. Rev. B **80**, 014120 (2009).
[99]H. W. Kuhn, "The Hungarian method for the assignment problem," Naval Res. Logistics Quart. **2**, 83–97 (1955).
[100]M. J. Buerger, "Crystallographic aspects of phase transformations," in *Phase Transformations in Solids*, edited by J. E. M. R. Smoluchowski and W. A. Weyl (John Wiley and Sons, 1951), pp. 183–211.
[101]V. A. Blatov, A. A. Golov, C. Yang, Q. Zeng, and A. A. Kabanov, "Network topological model of reconstructive solid-state transformations," Sci. Rep. **9**, 6007 (2019).
[102]F. Therrien, P. Graf, and V. Stevanović, "Matching crystal structures atom-to-atom," J. Chem. Phys. **152**, 074106 (2020).
[103]D. C. Lonie and E. Zurek, "Identifying duplicate crystal structures: XtalComp, an open-source solution," Comput. Phys. Commun. **183**, 690–697 (2012).
[104]F. Therrien and V. Stevanović, "Minimization of atomic displacements as a guiding principle of the martensitic phase transformation," Phys. Rev. Lett. **125**, 125502 (2020).
[105]G. Henkelman, B. P. Uberuaga, and H. Jónsson, "A climbing image nudged elastic band method for finding saddle points and minimum energy paths," J. Chem. Phys. **113**, 9901–9904 (2000).
[106]V. Lasrado, D. Alhat, and Y. Wang, "A review of recent phase transition simulation methods: Transition path search," in *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (ASME, 2008), Vol. 43253, pp. 93–101.
[107]D. Sheppard, R. Terrell, and G. Henkelman, "Optimization methods for finding minimum energy paths," J. Chem. Phys. **128**, 134106 (2008).
[108]D. Sheppard, P. Xiao, W. Chemelewski, D. D. Johnson, and G. Henkelman, "A generalized solid-state nudged elastic band method," J. Chem. Phys. **136**, 074103 (2012).
[109]K. J. Caspersen and E. A. Carter, "Finding transition states for crystalline solid–solid phase transformations," Proc. Nat. Acad. Sci. **102**, 6738–6743 (2005).
[110]G.-R. Qian, X. Dong, X.-F. Zhou, Y. Tian, A. R. Oganov, and H.-T. Wang, "Variable cell nudged elastic band method for studying solid–solid structural phase transitions," Comput. Phys. Commun. **184**, 2111–2118 (2013).
[111]G. H. Vineyard, "Frequency factors and isotope effects in solid state rate processes," J. Phys. Chem. Solids **3**, 121–127 (1957).
[112]A. F. Voter and J. D. Doll, "Transition state theory description of surface self-diffusion: Comparison with classical trajectory results," J. Chem. Phys. **80**, 5832–5838 (1984).
[113]P. Xiao, D. Sheppard, J. Rogal, and G. Henkelman, "Solid-state dimer method for calculating solid-solid phase transitions," J. Chem. Phys. **140**, 174104 (2014).
[114]C. Shang, X.-J. Zhang, and Z.-P. Liu, "Stochastic surface walking method for crystal structure and phase transition pathway prediction," Phys. Chem. Chem. Phys. **16**, 17845–17856 (2014).
[115]L. Zhu, R. E. Cohen, and T. A. Strobel, "Phase transition pathway sampling via swarm intelligence and graph theory," J. Phys. Chem. Lett. **10**, 5019–5026 (2019).
[116]C. Dellago, P. G. Bolhuis, and D. Chandler, "Efficient transition path sampling: Application to Lennard-Jones cluster rearrangements," J. Chem. Phys. **108**, 9236–9245 (1998).
[117]P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, "Transition path sampling: Throwing ropes over rough mountain passes, in the dark," Annu. Rev. Phys. Chem. **53**, 291–318 (2002).
[118]M. Grünwald and C. Dellago, "Transition path sampling studies of solid-solid transformations in nanocrystals under pressure," in *Trends in Computational Nanomechanics* (Springer, 2010), pp. 61–84.
[119]R. Z. Khaliullin, H. Eshet, T. D. Kühne, J. Behler, and M. Parrinello, "Nucleation mechanism for the direct graphite-to-diamond phase transition," Nat. Mater. **10**, 693–697 (2011).
[120]S. M. Woodley and R. Catlow, "Crystal structure prediction from first principles," Nat. Mater. **7**, 937–946 (2008).
[121]S. Bajaj, M. G. Haverty, R. Arróyave, and S. Shankar, "Phase stability in nanoscale material systems: Extension from bulk phase diagrams," Nanoscale **7**, 9868–9877 (2015).
[122]C. Wang and G. Yang, "Thermodynamics of metastable phase nucleation at the nanoscale," Mater. Sci. Eng. R Rep. **49**, 157–202 (2005).
[123]B. Tappan and R. L. Brutchey, "Polymorphic metastability in colloidal semiconductor nanocrystals," ChemNanoMat **6**, 1567–1588 (2020).
[124]V. Esposito and I. E. Castelli, "Metastability at defective metal oxide interfaces and nanoconfined structures," Adv. Mater. Interfaces **7**, 1902090 (2020).