Polymorphism in elemental silicon: Probabilistic interpretation of the realizability of metastable structures

Eric B. Jones and Vladan Stevanović

Polymorphism in elemental silicon: Probabilistic interpretation of the realizability of metastable structures

Eric Jones and Vladan Stevanović*

Colorado School of Mines, Golden, Colorado 80401, USA and
National Renewable Energy Laboratory, Golden, Colorado 80401, USA

With few systems of technological interest having been studied as extensively as elemental silicon, there currently exists a wide disparity between the number of predicted low-energy silicon polymorphs and those, which have been experimentally realized as metastable at ambient conditions. We put forward an explanation for this disparity wherein the likelihood of formation of a given polymorph under near-equilibrium conditions can be estimated on the basis of mean field isothermal-isobaric (N, p, T) ensemble statistics. The probability that a polymorph will be experimentally realized is shown to depend upon both the hypervolume of that structure's potential energy basin of attraction and a Boltzmann factor weight containing the polymorph's potential enthalpy per particle. Both attributes are calculated using density functional theory relaxations of randomly generated initial structures. We find that the metastable polymorphism displayed by silicon can be accounted for using this framework to the exclusion of a very large number of other low-energy structures.

Keywords: Silicon, Polymorph, Allotrope, Crystal Structure Prediction, Statistical Mechanics, Experimental Realizability

I. INTRODUCTION

Driven largely by the success of its monocrystalline ground state as the cornerstone of modern semiconductor technology, elemental silicon continues to see major efforts both to understand and to consider applications for the prolific polymorphism it displays[1]. While equilibrium methods for studying Si polymorphism, such as application of pressure via diamond anvil cell (DAC), have been employed for over 50 years[2], recent advances in discovering new polymorphs have relied upon unconventional synthesis methods such as ultrafast laser-induced confined microexplosion[3] and high pressure precursor routes[4]. Theoretical activity geared towards predicting new silicon polymorphs has kept pace, and as is summarized by Haberl et al.[1], predictions of hitherto un-sythesized metastable structures now numbers in the many dozens, with proximity in energy to the ground state as the main criterion used to judge experimental realizability. The experimental polymorph spectrum of silicon generated by equilibrium and nearer-to-equilibrium methods has been thoroughly characterized over a wide range of temperatures and pressures and is schematized in Fig. 1. Note that on first reference, equivalent naming conventions will be placed in round brackets next to the colloquial name. Thereafter, polymorphs will be referenced by their space group assignment. Upon compression in a DAC, the diamond ground state (dc-Si, Si-I, Fd$\bar{3}$m) metalizes into the beta tin phase (β-Sn-Si, Si-II, I4$_1$/amd) at 11.7 GPa[2] before passing through two closely related higher pressure phases: orthorhombic Imma (Si-XI, Imma) at 13.2 GPa[5,6] and simple hexagonal (sh-Si, Si-V, P6/mmm) at 15.3 GPa[5,7]. This series of phase transitions is reversible upon slow decompression back to the I4$_1$/amd phase, but upon full decompression at ambient temperatures, the body-centered structure with 8 atoms in its primitive unit cell (bc8-Si, Si-III, Ia$\bar{3}$) is recovered along with trace remnants of the rhombohedral 8 atom polymorph (r8-Si, Si-XII, R$\bar{3}$)[8,9]. Pressure release from an I4$_1$/amd phase obtained by indentation loading produces a higher fraction of R$\bar{3}$ to Ia$\bar{3}$ in the resulting metastable mixed phase. Hence, it appears that the formation of R$\bar{3}$ as a metastable state is concomitant to the formation of Ia$\bar{3}$ as well[10]. By annealing Ia$\bar{3}$ produced by DAC from room temperature to 473 K, it transforms into a third metastable polymorph, the hexagonal diamond, lonsdaleite analogue (hd-Si, Si-

FIG. 1. Polymorphs of silicon synthesized under equilibrium conditions. Bidirectional arrows denote reversible, pressure-induced phase transformations while a unidirectional arrow denotes an irreversible transformation. Polymorphs labeled with red boxes are unstable at ambient conditions, while those labeled with blue are metastable at ambient conditions, and the white box represents the cubic diamond ground state.

IV, $P6_3/mmc$)[11]. Meanwhile, *ex situ* annealing of the mixed Ia$\bar{3}$/R$\bar{3}$ phase produced by indentation loading to 473 K results in an as yet poorly characterized phase (Si-XIII, P4$_1$2$_1$2), which has only been reported to occur as a phase mixture[12].

Our emphasis here will be on metastable polymorphs of silicon. For completeness however, we mention that upon further compression, P6/mmm transforms reversibly into orthorhombic Cmca (Si-VI, Cmca) at 38 GPa[13], hexagonal close pack (hcp-Si, Si-VII, P6$_3$/mmc) at 42 GPa[7,13] (any future reference to P6$_3$/mmc will refer to the metastable lonsdaleite phase), and face-centered cubic (fcc-Si, Si-X, Fm$\bar{3}$m) at 79 GPa[14]. Two further metastable phases result exclusively upon *rapid* pressure release from 14.8 GPa (Si-VIII, P4$_1$2$_1$) and 12 GPa (Si-IX, P4$_2$2)[15], but neither has seen a full characterization of its crystal structure. Therefore, the polymorphs of silicon, which are both synthesized under near-equilibrium conditions and metastable at ambient conditions are plausibly considered to be those denoted by blue boxes in Fig. 1: R$\bar{3}$, Ia$\bar{3}$, and P6$_3$/mmc.

That this number of metastable, equilibrium-formed polymorphs is so few compared to the multitude, which are energetically competitive with the ground state in the silicon potential energy surface (PES)[16], motivates the development of some organizing principle by which candidate metastable polymorphs can be assessed for their likelihood of experimental realization. As recognized by authors such as Stillinger[17], a natural choice of formalism is statistical mechanics, since it simultaneously allows for a landscape of microstates to be meaningfully related to the macroscopic properties of a system and explicitly takes thermodynamic inputs such as temperature and pressure as parameters of the theory. In addition, it has recently been shown that the experimental polymorphism displayed by MgO, ZnO, and SnO$_2$ can be largely accounted for by considering the hypervolume of a polymorph's basin of attraction[18]. Since these three systems also possess very large numbers of potential polymorphs uncovered by computation, the introduction of basin hypervolume as an extra constraint in the prediction process was shown to be necessary in order to accurately select those polymorphs, which appear experimentally. This basin hypervolume constraint was given a probabilistic interpretation, which further suggests a fully probabilistic framework within which to discuss the experimental realizability of polymorphs. While the primary accomplishment of Ref. 18 was to illuminate *that* basin hypervolume plays a role in the experimental realizability of metastable structures, the conclusions made there are to a large extent qualitative and observational. The main thrust of the present work is to explain the underlying physics, bring the energy (enthalpy) into the picture as well as to formulate *how* the competition between basin hypervolume and enthalpy minimization work to attract ions into certain configurations. We choose enthalpy rather than energy above the ground state since the adoption of a thermodynamic potential
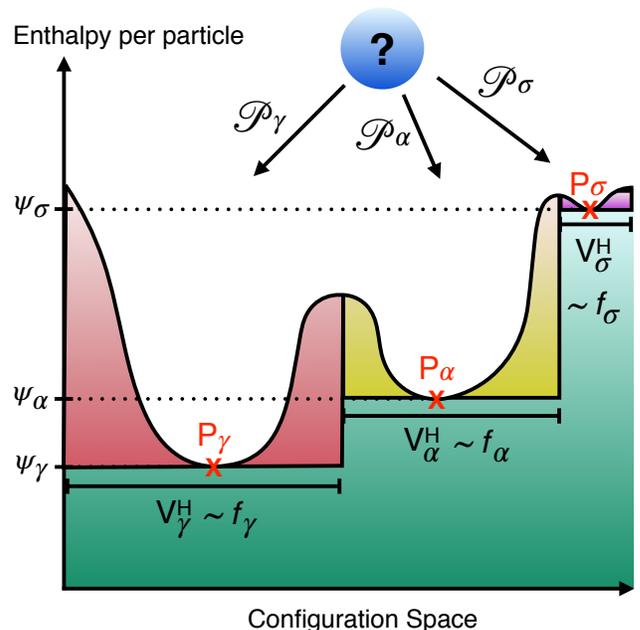
FIG. 2. Cartoon model of how an ion in a bulk material decides which structural configuration to become a part of. The ion is more likely to find minima of basins with large hypervolume (V_γ^H), which is proportional to a structure's frequency of occurrence (f_γ). Of the minima it finds, it will try to join the corresponding structure, which minimizes its enthalpy (ψ_γ). The interplay between these two effects is neatly accounted for by the probability of occurrence $\mathcal{P}_\gamma = f_\gamma e^{-\beta\psi_\gamma}$.

suited to the experimental setup under question (such as enthalpy for diamond anvil cell or annealing experiments) should bring predictions of metastable structures in closer agreement to the experimental reality.

Of course, any complete analysis of metastable polymorphism for a given compound must also involve determination of kinetic barriers between candidate polymorphs. Our model serves as an initial screening process so that the number of candidate polymorphs between which kinetic barriers must be calculated can be drastically reduced in order to make more efficient the selection of realizable, kinetically stable polymorphs from a multitude of possible low-energy structures. True metastable polymorphs from the model-selected set will then be those that are also kinetically stable. It is to the construction of this model that we now turn.

II. POLYMORPH REALIZABILITY FROM ISOTHERMAL-ISOBARIC ENSEMBLE STATISTICS

In order to develop a model, which accurately accounts for silicon's experimental behavior, we consider the perspective of an ion deciding which of a number of nucleated structures it would most like to join, represented by the matte blue circle in Fig. 2. Note that this mean field,

single-particle picture is already tacitly assumed in much of the polymorph prediction community since activation barriers and total energies are rarely cited as a function of all N particles simulated[19] but rather as energy per atom. We therefore begin our analysis with the partition function for a single particle immersed in an N-particle ionic configuration (corresponding to the N particles in a simulated unit cell) with 3N coordinates $\boldsymbol{r}$, periodic boundary conditions, average volume per atom v, mean field potential energy per particle $\phi(\boldsymbol{r}, v)$, and held under constant temperature (T) and pressure (p).

$$\Xi_1(N, p, T) = \frac{1}{V_0 \Lambda_T^3} \int dv \int d\boldsymbol{r} e^{-\beta(\phi(\boldsymbol{r},v)+pv)} \quad (1)$$

$\Lambda_T = h/\sqrt{2\pi M k_B T}$ is the mean thermal wavelength produced by the separable momentum integral, V_0 is a constant with units of volume to make the right-hand side dimensionless, and $\beta = 1/k_B T$. The mean-field treatment of the interactions between the atoms in a solid we find appropriate because it allows elimination of the dependence on the number of atoms N from the partition function. Also, it is important to note that the Born–von Karman boundary conditions are assumed, that is, the finite size effects such as the surface and interface energies that are relevant for the nucleation and growth of different phases are not accounted for in this discussion. Meanwhile, the integrand in Eq. 1 can be interpreted as the probability density for a single ion to find itself as part of an ionic configuration $\boldsymbol{R} = (\boldsymbol{r}, v)$ with mean field potential enthalpy per particle $\psi(\boldsymbol{R}) = \phi(\boldsymbol{r}, v) + pv$.

Let α index structures P_α, which are local potential enthalpy minima, indicated in bright red in Fig. 2. We also introduce B_α as the basin of attraction leading to P_α with attendant intra-basin coordinates $\boldsymbol{R}_\alpha$. The configurational integral can then be carved up into a sum of integrals over basins

$$\Xi_1(N, p, T) = \frac{1}{V_0 \Lambda_T^3} \sum_\alpha \int_{B_\alpha} d\boldsymbol{R}_\alpha e^{-\beta \psi(\boldsymbol{R}_\alpha)} . \quad (2)$$

To lowest order, we approximate each basin's contribution to the integrand by $\psi(\boldsymbol{R}_\alpha) \to \psi_\alpha$, the value of the potential enthalpy per particle evaluated at the basin minimum. As is represented in Fig. 2 by the matte red, yellow, and purple coloration, this approximation amounts to a flattening of each basin into a "square" well. The integral then yields the hypervolume (V_α^H) of the square well that maps to P_α, the well's local minimum. With these modifications, an estimate for the relative probability for an ion to choose between two nascent polymorphs P_α and P_γ results.

$$\frac{\mathcal{P}_\alpha}{\mathcal{P}_\gamma} = \left(\frac{V_\alpha^H}{V_\gamma^H}\right) e^{-\beta(\psi_\alpha - \psi_\gamma)} \quad (3)$$

The Boltzmann factor in Eq. 3 exponentially suppresses states, which are far away from the thermodynamically most favorable state at a given (T, p), while the ratio V_α^H / V_γ^H selects which states around the global enthalpy minimum are more likely to form.

In practice, interest in metastable states, which are energetically competitive with the ground state has typically dictated that consideration of the $(T = 0, p = 0)$ potential energy surface is sufficient in order to identify candidate polymorphs for synthesis[1]. Under these circumstances, $\psi_\alpha \to \phi_\alpha = \Phi_\alpha/N$, the potential energy per particle. However, simply because a state is metastable at ambient conditions does not mean that it necessarily forms initially under those conditions. We therefore would like an estimate for the potential enthalpy proper (ψ_α) such as to allow for varying synthesis conditions, but one that is also not too dissimilar to the PES at $(T = 0, p = 0)$. One can recover such an estimate for the enthalpy per particle by using the energy and volume per ion evaluated at $(T = 0, p = 0)$, $\psi_\alpha^{approx.} = \phi_\alpha + pv_\alpha^{approx.}$ (again, p is the pressure to be inserted as a parameter). A relevant assumption for our analysis then becomes that the ratios V_α^H / V_γ^H remain mostly unaltered for states not too far away from the ground state. This assumption is justified on the basis that the potential enthalpy surface at low to moderate pressures is expected to be an adiabatic deformation of the potential energy surface at zero pressure[20]. $\psi_\alpha^{approx.}$ allows us to evaluate the Boltmann factor in Eq. 3.

Meanwhile, one can estimate the prefactor V_α^H / V_γ^H by populating the PES with N^{RS} randomly generated structures. If N_α is the number of those structures, which are assignable to the local minimum P_α by virtue of being in the basin of attraction B_α then

$$\frac{V_\alpha^H}{V_\gamma^H} = \lim_{N^{RS} \to \infty} \left(\frac{N_\alpha}{N^{RS}}\right) \bigg/ \left(\frac{N_\gamma}{N^{RS}}\right) \equiv \frac{f_\alpha}{f_\gamma}, \quad (4)$$

where the f_α are termed "frequencies of occurrence"[18]. The $\mathcal{P}_\alpha = f_\alpha \exp(-\beta \psi_\alpha^{approx.})$ are termed "probabilities of occurrence" and are the main objects of interest in this paper.

Finally, the connection between single particle statistics and the thermodynamics of a bulk solid can be made through the Gibbs free energy per particle.

$$g = -\frac{1}{\beta} \ln\left(\sum_\alpha f_\alpha e^{-\beta \psi_\alpha}\right) \quad (5)$$

If a given state $\gamma \in \{\alpha\}$ is heavily favored in the above summation then to lowest order Eq. 5 can be written

$$g \approx \psi_\gamma - \frac{1}{\beta} \ln f_\gamma. \quad (6)$$

Hence under a given (T, p), a resonant probability of occurrence of a given state drives the system to both minimize its enthalpy per particle and maximize its entropy per particle to $s = k_B \ln f_\gamma$, by choosing that state.

Since there are a number of approximations, which have been made in this section in order to estimate the contributions to the partition function in Eq. 1, we will

look for qualitative features in the probability distribution $\mathcal{P}_\alpha$ that might, under certain experimental conditions, lead to resonant probabilities of the type in Eq. 6. To this end, we detail in Sec. III how random structures are initialized on the PES, mapped to local minima, and sorted into structural equivalence classes.

III. COMPUTATIONAL EVALUATION OF POTENTIAL ENTHALPY SURFACE STATISTICS

A. Random Structure Sampling

In order to initialize random structures on the silicon PES, we utilize the previously developed procedure for doing so[18]. Each unit cell is specified by six parameters chosen randomly, the lengths of the unit cell's lattice vectors a, b, c bounded between 0.6 and 1.4 in scaled units, and the angles between them α, β, γ bounded between $60°$ and $140°$. In fractional coordinates, ions are randomly distributed within the unit cell by first constructing a reciprocal lattice vector $\mathbf{G} = n_1\mathbf{g}_1 + n_2\mathbf{g}_2 + n_3\mathbf{g}_3$ with n_1, n_2, n_3 all between 4 and 7. This defines a plane wave $\cos(\mathbf{G} \cdot \mathbf{r})$. At each plane wave minimum, silicon ions are randomly distributed into planes defined by $\mathbf{G}$. In order to minimize instances where ions are randomly placed too close together, Gaussian probability distributions are centered on atoms, which have already been placed and subsequent ions are preferentially located in low probability regions on each superlattice plane. Once all of this has been done, a coordinate transformation is performed back into cartesian coordinates and the random structure is scaled such that the minimum distance between any two ions is no shorter than 1.8 Å, which helps with convergence of subsequent DFT relaxations and sets the scale of the dynamics to the physical scale at which they occur. In order to ensure that the resulting statistics were robust under changes in this minimal bond length, sets of relaxations were performed where the minimal bond length was adjusted to as small as 1.4 Å and as large as 2.2 Å. No subsequent dependence of the sampling statistics was found.

B. DFT Relaxations

The mapping $\psi(\boldsymbol{R}_\alpha) \to \psi_\alpha$ depicted in Fig. 2 is implemented by relaxing initialized random structures to nearby local minima via Density Functional Theory. We note that using this method, DFT relaxations can push a random structure initialized in one basin over a small energy barrier into an adjacent basin. Thus, it will be more accurate in the following to say that the α index funnels of attraction[18]. At each relaxation step the electronic ground state is computed for the ionic configuration at that step using the Perdew, Burke, Ernzerhof (PBE) form for the exchange correlation functional[21]. Valence electron behavior near ionic cores is accounted for using

the projector augmented wave (PAW) method[22], and the simulation package used for calculations is the Vienna Ab initio simulation package (VASP)[23]. At each step, ionic positions, unit cell volume and unit cell shape are all relaxed using the conjugate gradient algorithm[24]. This process is continued until the total energy is converged to within 3 meV/atom between successive iterations. In order to ensure correct numerical convergence, both volume and ionic relaxations are restarted four times and complemented by a self consistent DFT run afterwards. In addition, relaxations resulting in structures with residual forces above 10^{-4} eV/Å and/or pressures above 3 kbar are restarted in order to alleviate these residual forces and pressures. The pylada wrapper for high-throughput calculations was used in order to manage the computations and workflows[25].

C. Structure Comparison

Once initialized random structures are successfully relaxed to local minima, the resulting structures are compared and sorted into classes of structural equivalence according to four attributes: (*i*) common total energy to within 10 meV/atom, (*ii*) identical space groups to within a tolerance of 0.3 Å, (*iii*) volume per atom match to within a tolerance of 1%, and (*iv*) coordination of up to the fourth nearest neighbor to within a tolerance of 0.2 Å. The total energy per atom and volume per atom of each equivalence class is then used to compute the approximate enthalpy per particle of that equivalence class ($\psi_\alpha^{approx.}$), and the ratio of the equivalence class magnitude to the total number of relaxed structures results in the frequency of occurrence ($f_\alpha = N_\alpha/N^{RS}$).

IV. RESULTS AND DISCUSSION

While the analysis of Sec. II postulates a correlation between features of the single particle partition function in Eq. 1 and the tendency of ions in a crystal to join particular structures, a full accounting of the statistics of nucleation and growth would involve calculations of interfacial and/or surface energies[19]. Alternatively, given that nucleation of crystals in silicon is thought to occur on the order of a few nanometers, or roughly a few dozen atoms[26], we reason that features in the PES, which display scale invariance with respect to unit cell size, will be more likely to persist at the nucleation scale. For this reason, we calculated probabilities of occurrence over unit cell sizes $N = 8, 12, 16,$ and 24. This choice of sampling allows for the formation of polymorphs with primitive unit cell size commensurate to any of those discussed in Sec. I. However, due to the exponential growth of the number of local minima, which enter into the PES as a function of system size[17,27], and hence the large number of relaxations required to reach convergence of the $\mathcal{P}_\alpha$, $N = 24$ represents a current rough upper limit on the
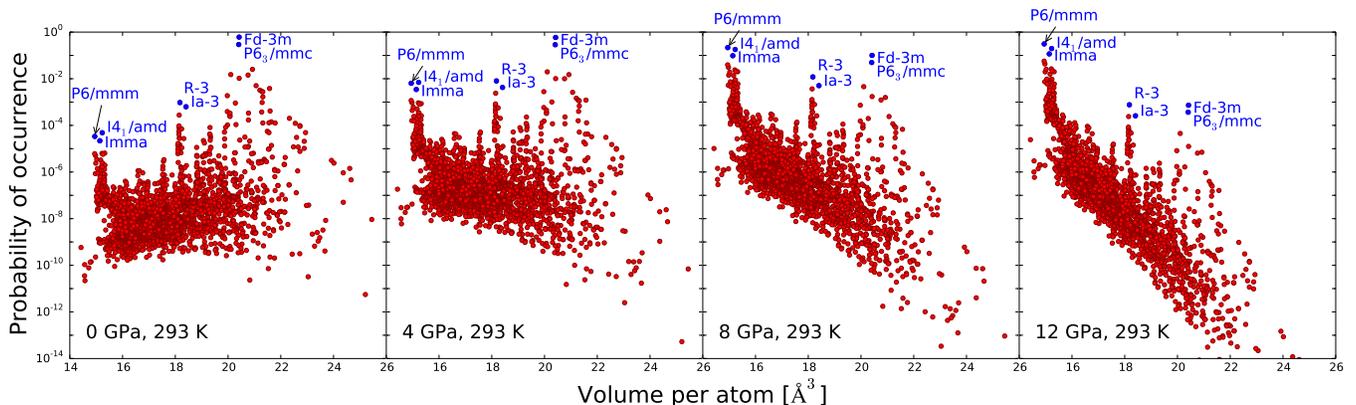
FIG. 3. Plot of the terms $\mathcal{P}_\alpha$ in the $N = 8$ mean-field partition function at room temperature ($\sim$ 293 K). Starting from the left panel, the pressures plotted are 0, 4, 8, and 12 GPa. Blue points refer to experimentally realized polymorphs while red points refer to potential polymorphs hitherto not synthesized.

unit cell size for which good statistics can be generated for silicon.

A total of 5,845 random structures were relaxed with a unit cell size of N=8, which produced 2,597 structural equivalence classes, each corresponding to a unique structure. The choice of N=8 as our initial unit cell size to consider stems from the fact that the primitive unit cell sizes of the Fd$\bar{3}$m, P6$_3$/mmc, R$\bar{3}$, and Ia$\bar{3}$ phases are N=2, 4, 8, and 8 respectively, with N=8 being the least common multiple of those known primitive unit cell sizes. Note that N=8 also accommodates I4$_1$/amd, Imma, and P6/mmm with primitive unit cells of N=2, 2, and 1 respectively. Fig. 3 shows the dependence of the N=8 probability distribution on pressure plotted against the natural order parameter for pressure- and temperature-induced phase transitions: volume per atom. First, it is interesting to note that of the volume range accessed, not all structural volumes are equally probable. A common feature to all panels in Fig. 3 (and also Figs. 4 and 5), there are regions of comparatively high probability, which occur at roughly 15 Å^3 per atom, from 18 to 18.5 Å^3 per atom, and again at around 20.5 Å^3 per atom. We will call these regions "jets", and in contrast to the regions at around 15.5 - 18 Å^3 per atom, 18.5 - 20 Å^3 per atom, and greater than about 21 Å^3 per atom (which we term "hulls"), these high probabilities contribute the most strongly to the partition function. Therefore, it is significant that the most probably occurring structures, which cap each "jet" in Fig. 3 are experimentally realized polymorphs as will be discussed more extensively below. Looking forward, discovery of these features in a new compound's PES would point towards where in parameter space various high-probability structures might likely become accessible to experiment.

At 0 GPa (left-most panel of Fig. 3), the two clearly resonant probabilities are those at the tip of the 20.5 Å^3 per atom jet, the experimentally realized Fd-$\bar{3}$m and P6$_3$/mmc phases. By contrast, the next most probable structure is an order of magnitude less probable, and the

two other resonant jets at 15 and 18-18.5 Å^3 per atom fall well below this. At 12 GPa (right-most panel), the 15 Å^3/atom jet is clearly the most prevalent, capped by I4$_1$/amd, Imma, and P6/mmm, which are the first three experimental phases found upon static compression in a DAC. The intermediate pressures, 4 and 8 GPa (center-left and center-right panels), show situations where all three jets compete probabilistically with 8 GPa being an approximate coexistence point. That is, at 8 GPa the experimentally realized metastable R$\bar{3}$ and Ia$\bar{3}$ polymorphs, which are the two most probable structures around 18-18.5 Å^3 per atom nearly approach P6$_3$/mmc in probability. We note that each of these structures (blue points in Fig. 3) was verified against its corresponding entry in the Inorganic Crystal Structure Database (ICSD)[28]. Upon identifying high probability polymorphs, structures were brought into a standard primitive cell form using AFLOW online[29], and subsequently checked for dynamical stability by running phonon calculations with Quantum Espresso[30]. Fd$\bar{3}$m, P6$_3$/mmc, Ia$\bar{3}$, and I4$_1$/amd were all determined to be stable, while R$\bar{3}$, Imma and P6/mmm were not. This is consistent with the experimental observations that P6/mmm and Imma transform reversibly under slow pressure release back to I4$_1$/amd and that R$\bar{3}$ typically only forms as a small fraction of a mixed phase with Ia$\bar{3}$. Interestingly, the latter fact seems to suggest that ions find their way into the Ia$\bar{3}$ structure by first identifying R$\bar{3}$ and then further proceeding to Ia$\bar{3}$.

The pressure dependence of our model gives us a good first-pass estimate of which structures will be able to compete thermodynamically under various pressurized synthesis conditions. In a similar manner, the temperature dependence of the probability distribution shown in Fig. 4 shows qualitatively how certain states can become more competitive with increasing temperature. It demonstrates a vertical squeeze in the distribution at elevated temperatures, which results in less probable structures becoming more competitive with more probable
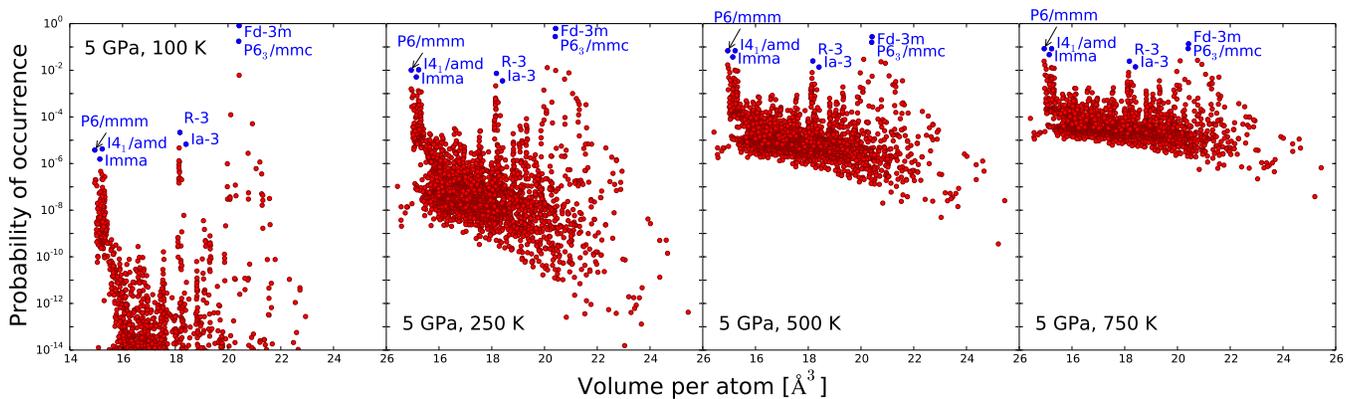
FIG. 4. Plot of the terms $\mathcal{P}_\alpha$ in the $N = 8$ mean-field partition function at a fixed pressure of 5 GPa. Starting from the left panel, the temperatures plotted are 100, 250, 500, and 750 K. Blue points refer to experimentally realized polymorphs while red points refer to potential polymorphs hitherto not synthesized.

structures. The real upshot of both Figs. 3 and 4 however, is that the ordering principle for experimental realizability is rather insensitive to model parameters. So long as jets and hulls can be identified in the probability distribution, then experimental realizability can be inferred from which structures are most probable within each jet. From those results, one can then proceed to assess dynamical stability, and calculate transition pressures and kinetic barriers.

In order to check for unit cell size dependence of the probability distribution, the $\mathcal{P}_\alpha$ for N=8 are again plotted at 5 GPa and 293 K in the left panel of Fig. 5 against volume per atom. 5 GPa and 293 K are chosen so as to be able to easily compare distribution features across the various unit cell sizes without biasing the distributions towards any one volume regime. With the three jets discernible at around 15, 18-18.5, and 20.5 $\mathring{A}^3$ per atom, the Fd$\bar{3}$m and P6$_3$/mmc phases are the most probable at large volumes, R$\bar{3}$ and Ia$\bar{3}$ the most probable at intermediate volumes, and I4$_1$/amd, Imma, and the simple hexagonal phase the most probable at small volumes. Among the structures we found at N=8 with probabilities lower than those, which have been experimentally realized, we were able to identify the C222$_1$ structure predicted by Botti et al. to be a quasi-direct gap semiconductor suitable for photovoltaic applications[16]. And while we do find a family of dissimilar I4$_1$/a structures beneath R$\bar{3}$ and Ia$\bar{3}$ at around 18 $\mathring{A}^3$/atom, none of the more probably occurring structures could be identified with the (bt8, I4$_1$/a) structure synthesized by Rapp et al.[3] This is unsurprising since the development of isothermal-isobaric statistics takes thermodynamic equilibrium as an axiom, while ultrafast laser-induced confined microexplosion is a decidedly non-equilibrium technique.

Upon doubling the unit cell to N=16, 14,813 random structures were relaxed, generating 12,854 structural equivalence classes. The resulting probability distribution is plotted in the center-right panel of Fig. 5.

The only major discrepancy among the most probably occurring structures in each volume regime is that Ia$\bar{3}$ falls below a pair of P-1 structures. The P-1 structure with smaller volume per atom and larger probability of occurrence was found to be dynamically stable while the one with larger volume per atom and lower probability of occurrence was found to be dynamically unstable. Calculation of kinetic barriers is beyond the scope of this paper, but it could be the case that the dynamically stable P-1 structure is in kinetically unstable, explaining its absence in nature. A similar argument might be applied to the C2/c outlier found in the larger-volume hull of the center-right panel of Fig. 5. In any case, since R$\bar{3}$ constitutes a simple distortion of Ia$\bar{3}$[1], it is likely appropriate to consider the two polymorphs as belonging to the same funnel of attraction. Such a renormalization of the Ia$\bar{3}$ probability would render it as jointly the most probably occurring structure around the $\sim$ 18-18.5 $\mathring{A}^3$ per atom jet and more probably occurring than the C2/c structure.

For good measure, we generated probabilities of occurrence for N=12, since at least one polymorph has been both theoretically predicted and experimentally synthesized by non-equilibrium methods with a 12 atom primitive unit cell[3]. For N=12, 9,919 random structures were relaxed, which resulted in 6,655 equivalence classes. Note that while Fd-$\bar{3}$m, P6$_3$/mmc, I4$_1$/amd, Imma, and P6/mmm are compatible with N=12 since their primitive unit cell sizes divide 12 evenly, R$\bar{3}$ and Ia$\bar{3}$ are incompatible with N=12 since they both have a primitive unit cell of N=8. As such, we should not expect the appearance of either R$\bar{3}$ or Ia$\bar{3}$ at N=12. At small volumes, I4$_1$/amd, Imma, and P6/mmm still squarely dominate the probability distribution. Meanwhile, the top two most probably occurring polymorphs at moderate volumes have space groups P4$_1$2$_1$2 and P4$_3$2$_1$2 followed by a family of inequivalent P2$_1$/c structures. Phonon calculations were carried out for P4$_1$2$_1$2 and P4$_3$2$_1$2, and both are confirmed to be dynamically stable. Surprisingly, the P4$_3$2$_1$2 structure is exactly the phase recently synthe-
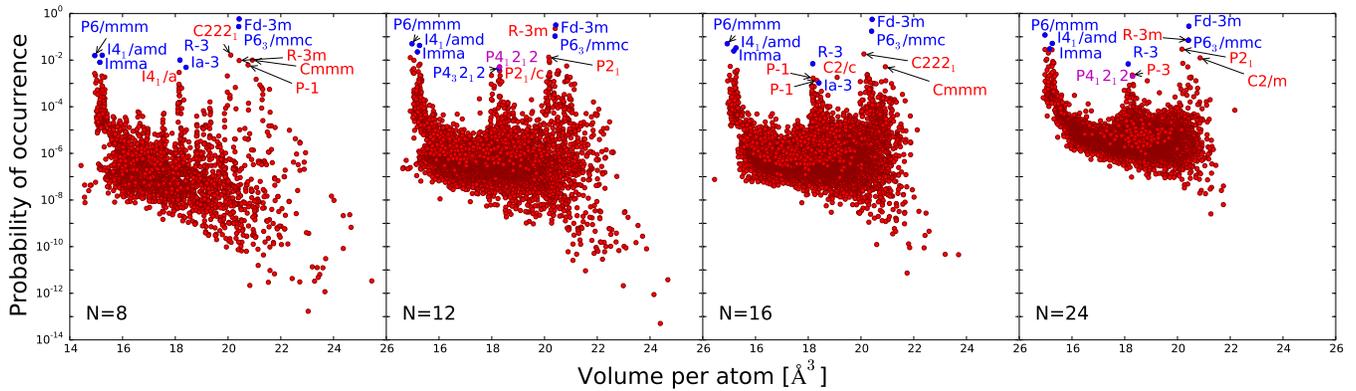
FIG. 5. Plot of the terms $\mathcal{P}_\alpha$ in the mean-field partition function at a fixed pressure of 5 GPa and room temperature ($\sim$ 293 K). Starting from the left panel, the unit cell sizes plotted are N=8, 12, 16, and 24. Blue points refer to experimentally realized polymorphs while red points refer to potential polymorphs hitherto not synthesized. The magenta point denotes ambiguity regarding synthesis.

sized under distinctly far-from-equilibrium conditions by Rapp et al.[3] (st12, $P4_32_12$). This stands in contrast to the same authors' $I4_1/a$ structure, which does not appear with high probability at N=8. More generally, other than $P4_32_12$ we do not recover as probably occurring any of the other polymorphs, which have successfully been synthesized under non-equilibrium or exclusively chemically altered conditions. $P4_12_1$ (Si-VIII) and $P4_22$ (Si-IX) are included in this classification since they are formed via rapid pressure release, and as regards chemical precursor synthesis routes, one would need to explore the PES of the modified chemistry rather than that of pure silicon in order to derive accurate statistics. On the one hand, we would not necessarily expect an equilibrium model to capture non-equilibrium phenomena (although it is not expressly forbidden). But in the case of $P4_32_12$, it could be that the frequency of occurrence for this state drives formation even when the Boltzmann factor becomes inapplicable due to the presence of the microscopic temperature, pressure, and electric field inhomogeneities likely introduced through ultrafast laser-induced confined microexplosion. Alternatively, there could be some semblance of local thermodynamic equilibrium even under these extreme conditions such that the Boltzmann factor still contributes locally to structure formation.

Identification of the $P4_12_12$ phase is less straightforward, but it might be associated with the $P4_12_12$ (Si-XIII) phase formed experimentally by annealing an indentation loading sample of mixed $R\bar{3}/Ia\bar{3}$ at 473 K[12]. Multiple predictions have previously been put forward to solve the crystal structure of this phase[31,32], but its obstinacy towards coexistence with other phases has made any sort of definite characterization challenging. It is noteworthy that the experimental $P4_12_12$ phase only occurs upon annealing the particular mixture of $R\bar{3}/Ia\bar{3}$ produced by indentation loading, and has not to this point been obtained as a phase pure isolate. At large volumes, a new structure with space group $R\bar{3}m$ appears

between the $Fd\bar{3}m$ and $P6_3/mmc$ phases with 12 atoms in its primitive unit cell (inequivalent to the $R\bar{3}m$ structure found at N=8). Phonon calculations indicate that this phase is also dynamically stable, which would seem to suggest its viability as an experimentally realizable silicon polymorph. However, upon doubling the unit cell size from 12 to 24, a drop in the $R\bar{3}m$ probability occurs with respect to the $P6_3/mmc$ phase (the plot points for the two phases are overlaid). Extrapolating that relative drop to larger unit cell sizes would see a continued decrease in the probability of $R\bar{3}m$ and thus a restoration of $P6_3/mmc$ as the second most probably occurring at large volumes.

It should be noted that for N=24, the relative complexity and expansiveness of the PES prohibited us computationally from fully exploring funnel of attraction statistics, but we were able to extract some features of it from the 9,845 successfully relaxed random structures, sorted into 9,166 equivalence classes. The prevalence of $I4_1/amd$, Imma, and P6/mmm at small volumes is replicated at N=24, while $R\bar{3}$ remains the most probable intermediate-volume structure. As can be seen in the right-most panel of Fig. 5, we were unable to encounter the $Ia\bar{3}$ phase among any of the successfully relaxed and sorted structures. However, the argument that $Ia\bar{3}$ and $R\bar{3}$ share a common funnel of attraction would again place $Ia\bar{3}$ alongside $R\bar{3}$ as the most probable intermediate-volume structure. Below the $R\bar{3}/Ia\bar{3}$ probability, we find a structure with space group $P\bar{3}$ only slightly above the 12 atom $P4_12_12$ structure. This relative drop in probability could indicate either a reason for why phase-pure Si-XIII is difficult to achieve or alternatively, that our $P4_12_12$ structure is not assignable to Si-XIII and is some entirely different phase, which would be difficult to synthesize. Without further experimental input such as X-ray diffraction data on a phase-pure sample, it is challenging to determine which scenario reflects reality. Also missing from the N=24 probability distribution as being high-

probability is the $P4_32_12$ structure, potentially explaining why $R\bar{3}/Ia\bar{3}$ are easier to synthesize under equilibrium conditions, and why non-equilibrium methods are needed to realize $P4_32_12$. And conspicuously missing from any unit cell size are any phases with volume below about 14 $\mathring{A}^3$ per atom.[33] Figs. 3, 4, and 5 all show that no such structures were found with our sampling technique. One explanation for this is that at pressures far above 15 GPa (roughly where $P6/mmm$ is accessed), the potential enthalpy surface is so dissimilar to the zero pressure PES that high pressure local minima cease to be minima at all at ambient conditions.

There are three main approximations we have performed in order to be able to relate ab initio exploration of silicon's PES statistics to the experimental formation of its polymorphs. First, we have chosen to use a mean field, single particle model to represent the statistical mechanics of crystal formation in a real solid, a process that in reality is dynamical and proceeds by nucleation and growth. Second, we have replaced all intra-basin contributions to the configurational integral in Eq. 2 by those evaluated at basin minima, which corresponds to flattening each basin into a square well. And finally, we have approximated the true potential enthalpy per particle by $\psi^{approx.}$, calculated from potential energy and volume data generated at zero temperature and pressure. Despite these approximations, the qualitative trends in Fig. 5 are clear and indicate that of the thousands of potential polymorphs found, $Fd\bar{3}m$, $P6_3/mmc$, $R\bar{3}/Ia\bar{3}$, $I4_1/amd$, $Imma$, and $P6/mmm$ are the most persistently probable in their respective volume regimes and thus will be the most likely to be accessed under equilibrium conditions, corroborating the experimental reality. Of these, $R\bar{3}$, $Ia\bar{3}$, and $P6_3/mmc$ are known to be metastable. Thus, we have accurately reproduced the equilibrium-formed, metastable sector of silicon's polymorph spectrum using quantities readily produced by ab initio techniques (total energies and volumes) by gathering statistics on random structure relaxations, variants of which are also currently employed by the structure prediction community[34]. Our model therefore provides a tool by which candidate polymorphs of little-studied compounds can be easily assessed

for experimental realization in the future.

V. CONCLUSION

In addition to their novel properties, metastable polymorphs such as the diamond phase of carbon see wide industrial application because of their amenability to being artificially synthesized in bulk quantities[35]. Given that equilibrium materials synthesis techniques have historically been more extensively studied and hence, have been used to greater effect to synthesize materials at the industrial scale, it is important to be able to understand which of the many candidate polymorphs produced by an ab initio search are likely to be synthesized under equilibrium conditions. We have therefore presented a mean-field model of experimental polymorph realizability predicated on the isothermal-isobaric statistics of a single, interacting ion. Using this formalism, we demonstrated that the polymorphs of silicon synthesizable under near-equilibrium conditions can be separated from the many other candidate polymorphs produced by an ab initio search by selection of resonant probabilities in the resulting distribution. By way of interpretation, we have shown that in addition to global enthalpic favorability, ions are driven into specific configurations by a measure of entropy, which is related to the hypervolume of a structure's funnel of attraction. This approach provides a prescription whereby candidate polymorphs uncovered for compounds, which have previously seen little experimental work, can be ranked for experimental consideration by their probabilities of occurrence.

ACKNOWLEDGMENTS

[*] vstevano@mines.edu

[1] B. Haberl, T. A. Strobel, and J. E. Bradby, Applied Physics Reviews **3**, 040808 (2016), http://dx.doi.org/10.1063/1.4962984.

[2] J. C. Jamieson, Science **139**, 762 (1963), http://science.sciencemag.org/content/139/3556/762.full.pdf.

[3] L. Rapp, B. Haberl, C. Pickard, J. Bradby, E. G. Gamaly, J. Williams, and A. V. Rode, Nature communications **6** (2015).

[4] D. Y. Kim, S. Stefanoski, O. O. Kurakevych, and T. A. Strobel, Nat Mater **14**, 169 (2015), letter.

[5] M. I. McMahon and R. J. Nelmes, Phys. Rev. B **47**, 8337 (1993).

[6] M. I. McMahon, R. J. Nelmes, N. G. Wright, and D. R. Allan, Phys. Rev. B **50**, 739 (1994).

[7] H. Olijnyk, S. Sikka, and W. Holzapfel, Physics Letters A **103**, 137 (1984).

[8] S. Minomura and H. Drickamer, Journal of Physics and Chemistry of Solids **23**, 451 (1962).

[9] R. H. Wentorf and J. S. Kasper, Science **139**, 338 (1963), http://science.sciencemag.org/content/139/3552/338.2.full.pdf.

[10] S. Ruffell, K. Sears, A. P. Knights, J. E. Bradby, and J. S. Williams, Phys. Rev. B **83**, 075316 (2011).

[11] G. Weill, J. L. Mansot, G. Sagon, C. Carlone, and J. M. Besson, Semiconductor Science and Technology **4**, 280 (1989).

[12] S. Ruffell, B. Haberl, S. Koenig, J. E. Bradby, and J. S. Williams, Journal of Applied Physics **105**, 093513 (2009), http://dx.doi.org/10.1063/1.3124366.

[13] M. Hanfland, U. Schwarz, K. Syassen, and K. Takemura, Phys. Rev. Lett. **82**, 1197 (1999).

[14] S. J. Duclos, Y. K. Vohra, and A. L. Ruoff, Phys. Rev. B **41**, 12021 (1990).

[15] Y.-X. Zhao, F. Buehler, J. R. Sites, and I. L. Spain, Solid State Communications **59**, 679 (1986).

[16] S. Botti, J. A. Flores-Livas, M. Amsler, S. Goedecker, and M. A. L. Marques, Phys. Rev. B **86**, 121204 (2012).

[17] F. H. Stillinger, *Energy Landscapes, Inherent Structures, and Condensed-Matter Phenomena* (Princeton University Press, 2015).

[18] V. Stevanović, Phys. Rev. Lett. **116**, 075503 (2016).

[19] R. Z. Khaliullin, H. Eshet, T. D. Kuhne, J. Behler, and M. Parrinello, Nat Mater **10**, 693 (2011).

[20] A. Verma and P. Krishna, *Polymorphism and polytypism in crystals*, Wiley monographs in crystallography (Wiley, 1966).

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

[22] P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994).

[23] G. Kresse and J. Furthmller, Comput. Mater. Sci. **6**, 15 (1996).

[24] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes 3rd Edition: The Art of Scientific Computing*, 3rd ed. (Cambridge University Press, New York, NY, USA, 2007).

[25] "Pylada Computational Framework," https://github.com/pylada/pylada-light (2014), accessed: 2016-12-08.

[26] M. K. Sunkara, S. Sharma, R. Miranda, G. Lian, and E. C. Dickey, Applied Physics Letters **79**, 1546 (2001).

[27] F. H. Stillinger, Phys. Rev. E **59**, 48 (1999).

[28] M. Hellenbrandt, Crystallography Reviews **10**, 17 (2004), http://dx.doi.org/10.1080/08893110410001664882.

[29] S. Curtarolo, W. Setyawan, G. L. Hart, M. Jahnatek, R. V. Chepulskii, R. H. Taylor, S. Wang, J. Xue, K. Yang, O. Levy, M. J. Mehl, H. T. Stokes, D. O. Demchenko, and D. Morgan, Computational Materials Science **58**, 218 (2012).

[30] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, Journal of Physics: Condensed Matter **21**, 395502 (19pp) (2009).

[31] A. Mujica, C. J. Pickard, and R. J. Needs, Physical Review B **91**, 214104 (2015).

[32] Z. Zhao, F. Tian, X. Dong, Q. Li, Q. Wang, H. Wang, X. Zhong, B. Xu, D. Yu, J. He, H.-T. Wang, Y. Ma, and Y. Tian, Journal of the American Chemical Society **134**, 12362 (2012), pMID: 22803841, http://dx.doi.org/10.1021/ja304380p.

[33] This observation might suggest that our calculations overestimate the volume of P6/mmm, which is typically cited to be between 13 to 14 Å^3 per atom. However, that is the equilibrium volume under pressure, which will generally be smaller than the volume at ambient conditions.

[34] C. J. Pickard and R. J. Needs, Journal of Physics: Condensed Matter **23**, 053201 (2011).

[35] "U.S. Geological Survey Mineral Commodity Summaries, 2017," https://minerals.usgs.gov/minerals/pubs/commodity/diamond/mcs-2017-diamo.pdf (2017), accessed: 2017-07-18.