ARTICLE   OPEN

Check for updates

The glassy solid as a statistical ensemble of crystalline microstates

Eric B. Jones [1,2] and Vladan Stevanović [1,2✉]

We present an alternative and, for the purpose of non-crystalline materials design, a more suitable description of covalent and ionic glassy solids as statistical ensembles of crystalline local minima on the potential energy surface. Motivated by the concept of partially broken ergodicity, we analytically formulate the set of approximations under which the structural features of ergodic systems such as the radial distribution function (RDF) and powder X-ray diffraction (XRD) intensity can be rigorously expressed as statistical ensemble averages over different local minima. Validation is carried out by evaluating these ensemble averages for elemental Si and SiO_2 over the local minima obtained through the first-principles random structure sampling that we performed using relatively small simulation cells, thereby restricting the sampling to a set of predominantly crystalline structures. The comparison with XRD and RDF from experiments (amorphous silicon) and molecular dynamics simulations (glassy SiO_2) shows excellent agreement, thus supporting the ensemble picture of glasses and opening the door to fully predictive description without the need for experimental inputs.

INTRODUCTION

Modern approaches to describing the atomic structure of non-crystalline solids are predicated upon the three classical models: (i) the continuous random network (CRN) for covalent and ionic systems, (ii) random close packing for metallic glasses, and (iii) the random coil model for polymeric glasses[1]. While the random close packing and random coil models are defined statistically over an ensemble of random packings and random conformations respectively, the CRN is usually conceptualized in terms of a single "optimal" structural model[2]. Depicted in the left-hand side of Fig. 1, the optimal structure is a single microstate that minimizes the total energy subject to a certain bonding topology constraint[3].

That the covalent (or ionic) non-crystalline macrostate is also identifiable with an *ensemble* of microstates can be understood by considering the instance in which a glassy state is obtained by supercooling from a liquid. To clarify, we use the terms "glass" or "glassy state" to refer to a non-crystalline state obtained specifically by supercooling a liquid and the term "amorphous" to discuss non-crystalline solids obtained by any other means[4]. A liquid can be described as a point in $3N$ dimensional configuration space thermally sampling different local potential energy minima and their attraction basins[5,6]. Liquids are ergodic. Time averages of liquid properties are equivalent to ensemble averages on the timescales accessible to experiment. Upon supercooling, the resulting glassy state is no longer fully ergodic; not every microstate strictly allowable by energy considerations is accessible over experimental timescales. However, given that the configuration point of the liquid roams freely over the potential energy surface, it has been pointed out that a realistic quenching scenario involves the configuration point becoming kinetically constrained to some smaller region of configuration space, partially breaking the full ergodicity but otherwise able to thermally sample and locally relax between different basins of attraction in the glassy state[7–9].

This picture of glassy state formation complements and supports the concept known as "confusion" as described by A. L. Greer, which explains the formation of glasses in metallic systems with complex alloy structure[10]. By creating an alloy with many elements of widely varying atomic radii—$Zr_{41.2}Ti_{13.8}Cu_{12.5}Ni_{10.0}Be_{22.5}$, for example—the potential energy surface remains so complex upon supercooling from the liquid there becomes a very low "chance that the alloy can select viable crystal structures, and the greater the chance of glass formation" there becomes. From the perspective of partially broken ergodicity, the alloy attempts to visit all viable crystalline local minima accessible to it and therefore winds up in a glassy state, which equates structurally to an average over that set of minima.

In the present work, we present a general theory for describing the atomic scale structure of glassy states, which is in harmony with the concept of only partially broken ergodicity and which allows for the covalent and ionic glassy solids to be treated in a statistical manner on equal footing with metallic and polymeric glasses. We emphasize that the ensemble picture described in this paper does not equate to the well-known crystallite hypothesis in which the glassy state is described as spatially local regions of crystalline material—typically the ground state crystal structure—embedded in a disordered matrix[11]. Rather, as we will demonstrate an accurate interpretation of our results is that the structural features such as the X-ray diffraction (XRD) pattern and the radial distribution function (RDF) of the glassy state, broadly characterized as a non-periodic arrangement of atoms (local minimum) in a macroscopically large unit cell, can be approximated as a linear combination of the same structural features of many other local minima, which do exhibit translational symmetry with relatively small unit cells. The fact that the linear combination is formulated as a thermodynamic average implies the concept of partially broken ergodicity and brings dynamics and fluctuations into the picture. We thus argue neither for the notion of spatially localized crystallinity nor the existence of an embedding disordered matrix.

[1]Colorado School of Mines, Golden, CO 80401, USA. [2]National Renewable Energy Laboratory, Golden, CO 80401, USA. ✉email: vstevano@mines.edu
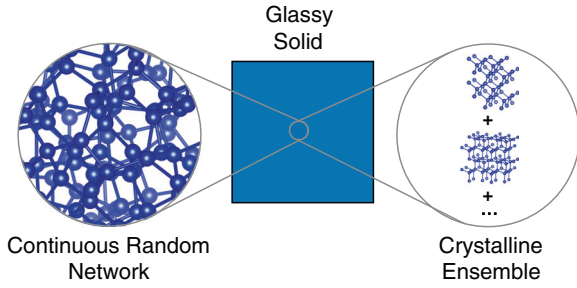
Fig. 1 Two descriptions of the covalent and ionic glassy solids. The widely accepted continuous random network model describes the glassy state as a single microstate. Meanwhile, our findings speak to a statistical ensemble of crystalline microstate description that features remnant, but reduced, ergodicity from the liquid.

The idea of approximating non-periodic local minima as a linear combination of finite number of periodic ones can be found in lattice models such as the cluster expansion[12,13]. Cluster expansion is a model Hamiltonian, a generalization of the Ising model, often used when modeling alloys. It contains a finite number of interaction terms (clusters) with the parameters describing interaction strengths (the Js) typically obtained by fitting to the first-principles computed total energies of a finite set of ordered (small-cell, periodic) structures[14]. Thus, when computing the energy of a large-cell, disordered alloy configuration using a cluster expansion, one effectively uses an implicit linear combination of a finite number of ordered structures. What is new in our work is a generalization of these ideas to systems that do not contain an underlying lattice. Our results show that these kinds of systems can also be represented as linear combinations of a finite number of small-cell, periodic structures, each with their own underlying lattice.

We proceed to show analytically how important structural descriptors of the glassy state can be calculated as averages within such an ensemble framework. The two quantities we choose to calculate are the RDF, which captures the local order of the state, and the XRD intensity, which captures the long-range order of the state. For our theoretical development, we make no assumptions regarding the structural periodicity of the ensemble microstates. However, we will find that evaluating the RDF and XRD expectation values for structural ensembles with a periodic unit cell as small as 24 atoms reproduces experimentally measured RDF and XRD patterns remarkably well for amorphous silicon (a-Si). This result therefore establishes the "crystalline ensemble," shown on the right-hand side of Fig. 1 as an accurate, approximate view of the glassy state. Similar ideas have been introduced previously by Curtarolo and co-workers who proposed a spectral descriptor for predicting glass-forming metallic alloys based on enthalpy distribution of ordered structures: the more dissimilar the structures with similar energies, the higher is the probability of formation of the glassy state[15].

RESULTS

Ensemble construction

In order to show the decomposition of the glassy RDF into a thermal average, we begin with the expression for the two-point density function for our glassy solid in the isothermal–isobaric (N, p, T) ensemble[16]. We choose this ensemble because it is under constant number of particles, temperature, and pressure (rather than volume) that many practical experiments take place, and which therefore need to accommodate microstates of varying volume. We will hence consider the enthalpy as the relevant thermodynamic potential. Later, we will set $p = 0$ for simplicity. However, for the present theoretical development, it is important

to keep pressure arbitrary so as to allow for microstates of differing volume, as mentioned above, since the volume of the glassy state should be derivable as an expectation value, not taken as an empirical input. The ensemble expectation value for the two-point density function of an ergodic system reads:

$$n^{(2)}(\mathbf{r}',\mathbf{r}'') = \left\langle \sum_{i=1}^{N} \sum_{j\neq i} \delta(\mathbf{r}_i - \mathbf{r}')\delta(\mathbf{r}_j - \mathbf{r}'') \right\rangle_{(N,p,T)} =$$

$$= \frac{1}{\Xi} \int dV \int d\mathbf{r}_1...d\mathbf{r}_N e^{-\beta(U(\mathbf{r}_1,...,\mathbf{r}_N,V)+pV)}$$

$$\times \sum_{i=1}^{N} \sum_{j\neq i} \delta(\mathbf{r}_i - \mathbf{r}')\delta(\mathbf{r}_j - \mathbf{r}''), \qquad (1)$$

where Ξ is the partition function, U potential energy, and $U + pV$ enthalpy of the system of N particles. The double summation in the integrand runs over all particle pairs, and thus $n^{(2)}(\mathbf{r}',\mathbf{r}'')$ describes the probability that any two particles occupy simultaneously the positions $\mathbf{r}'$ and $\mathbf{r}''$. Next, we split the configurational integrals in Eq. (1) into a sum of integrals over the basins of attraction B_a of various potential enthalpy minima (labeled a). In this summation, we multiply and divide by the intra-basin partition function $\Xi_a \equiv \int_{B_a} dV d\mathbf{r}_1...d\mathbf{r}_N e^{-\beta(U(\mathbf{r}_1,...,\mathbf{r}_N,V)+pV)}$. This allows us to express the two-point density function exactly as a weighted sum over two-point density functions constrained to reside within the basins $\{B_a\}$.

$$n^{(2)}(\mathbf{r}',\mathbf{r}'') = \sum_a P_a\, n_a^{(2)}(\mathbf{r}',\mathbf{r}'') \qquad (2)$$

The $P_a = \Xi_a/\Xi$ in Eq. (2) are the ensemble probabilities of individual local minima, which we showed previously to correlate with the experimental realizability of metastable polymorphs[17,18]. To evaluate these probabilities, we will assume the "flat basin approximation," i.e., that the potential enthalpy adopts a square well shape in every basin of attraction. Under this approximation, $P_a \approx f_a \exp(-\frac{h_a}{\tau})/\Xi$, where h_a is the enthalpy per particle at the local minimum of the basin B_a, $\tau = k_B T/N$ is an effective (scaled) temperature, and f_a is the hypervolume of the basin B_a. We also showed that f_a can be estimated from the relative frequency of occurrence of a local minimum a in the first-principles random structure sampling (explained below). Meanwhile, $n_a^{(2)}(\mathbf{r}',\mathbf{r}'')$ is the intra-basin expectation value of the two-point density operator.

We will approximate the intra-basin average by the two-point density function found at the minimum of the basin B_a. Given the relationship between the two-point density function and the two-point correlation function $n^{(2)}(\mathbf{r}',\mathbf{r}'') = \langle n \rangle^2 g^{(2)}(\mathbf{r}',\mathbf{r}'')$ with $\langle n \rangle$ the ensemble averaged global number density, we express $g^{(2)}(\mathbf{r}',\mathbf{r}'')$ also as a weighted sum using Eq. (2)

$$g^{(2)}(\mathbf{r}',\mathbf{r}'') = \sum_a \left(\frac{n_a}{\langle n \rangle}\right)^2 P_a\, g_a^{(2)}(\mathbf{r}',\mathbf{r}''). \qquad (3)$$

where n_a is the global number density of the minimum of the basin B_a. In order to derive the RDF, we perform three more steps. First, make the coordinate transformation to a "displacement" coordinate $\mathbf{r} = \mathbf{r}' - \mathbf{r}''$ and a "center of mass" coordinate, $\mathbf{R} = \frac{1}{2}(\mathbf{r}' + \mathbf{r}'')$. Second, volume average $g^{(2)}(\mathbf{r}, \mathbf{R})$ over the center of mass coordinate. And third, average the resulting pair correlation function $g^{(2)}(\mathbf{r})$ over the polar coordinates of $\theta_\mathbf{r}$ and $\phi_\mathbf{r}$. The final expression for the RDF as a weighted sum is:

$$g^{(2)}(r) = \sum_a \left(\frac{n_a}{\langle n \rangle}\right) P_a\, g_a^{(2)}(r). \qquad (4)$$

And using the relationship between the structure factor and the pair correlation function $S(\mathbf{q}) = 1 + \langle n \rangle \int_{(V)} d\mathbf{r} e^{-i\mathbf{q}\cdot\mathbf{r}} g^{(2)}(\mathbf{r})$, one can show that the powder diffraction intensity $I(2\theta)$ can be similarly

expressed (see Supplementary Material)[16] as:

$$I(2\theta) = \sum_{\alpha} P_{\alpha} I_{\alpha}(2\theta). \qquad (5)$$

Here we pause to comment on the relationship between the glass-forming ability (GFA) descriptor introduced by Perim et al., the derivations in this section, and Eq. (2) in particular[15]. For a fixed stoichiometry $\{x\}$, the GFA descriptor reads $\chi_{GFA}(\{x\}) = \sum_i f(H_i)g(|\psi_i\rangle)/h(\{x\})$. Each $f(H_i)$ plays a role equivalent to P_{α} in Eq. (2), the difference being that each $f(H_i)$ takes into account the Boltzmann-type weighting of a local minimum while each P_{α} additionally integrates the Boltzmann contribution of the basin of attraction to the local minimum. Meanwhile, $g(|\psi_i\rangle)$ gives a numerical factor that assesses the structural similarity of the local minimum $|\psi_i\rangle$ to the ground state crystal structure $|\psi_0\rangle$, and $h\{x\}$ is a normalization constant. Equivalently, each $n_{\alpha}^{(2)}(\mathbf{r}', \mathbf{r}'')$ in Eq. (2) constitutes a normalized measure of structural similarity to the ground state insofar as similar crystal structures will share similar two-point density functions. One could therefore imagine calculating any of the expectation values in this section through a range of compositions at some fixed τ. At some compositions, the RDF, for example, might contain sharp coordination shell peaks, while at another composition, such peaks might become more broadened. We expect that such correspondences would occur at compositions where the GFA descriptor displays little prominence and pronounced peaks, respectively. Such tandem calculations could be an interesting direction for future work. Having formulated the conditions under which the glassy RDF and XRD can be evaluated by proper statistical averages over sets of potential enthalpy minima, we would like to emphasize that these derivations do not suggest whether the minima are ordered (i.e., representable within a small size unit cell) or not. What we will show next in the numerical evaluation of XRDs and RDFs for silicon and silica glass, the details of which are outlined in the "Methods" section, is that an accurate description of these quantities can be obtained if one restricts the set of local minima only to small-cell, periodic structures.

Evaluation of observables

We note that, from hereon, all calculated observable quantities will be cited at $p \approx 0$ GPa in order to be able to reference them against experimental data. The terms "energy" and "enthalpy" will therefore be synonymous in the following. In order to understand the effective thermodynamics of our ensemble, we plot two important thermodynamic quantities in Fig. 2 as a function of the ensemble effective temperature $\tau = k_B T/N$ for elemental Si. The expectation value of the energy per particle relative to the diamond Si ground state, $\langle E \rangle - E_{GS}$, is shown in black, while the expectation value of the volume as a fraction of the ground state volume, $\langle V \rangle / V_{GS}$, is shown in gray (right axis). All of these expectation values are obtained using previously derived ensemble averaging formulae and the large number of local minima generated using the first-principles random structure sampling procedure explained in detail in the "Methods" section. Briefly, this procedure involves generating fully random unit cell vectors and atomic positions for N atoms of a predefined chemical identity. Subsequently, all of these random structures are fully relaxed (volume, cell shape, and positions) to the closest local minimum using density functional theory and the steepest descent algorithm. Solid and dashed lines in Fig. 2 correspond to the results with $N = 24$ atoms (~10,000 random structures) and $N = 16$ atoms (~15,000 random structures), respectively.

All curves shown display the sigmoid-type signature of an effective order–disorder phase transition, with increasing unit cell size $N = 16 \rightarrow 24$ rendering the character of the transition sharper. While the low τ regime, demarcated by shaded blue region, clearly describes the ground state properties of the system—zero energy
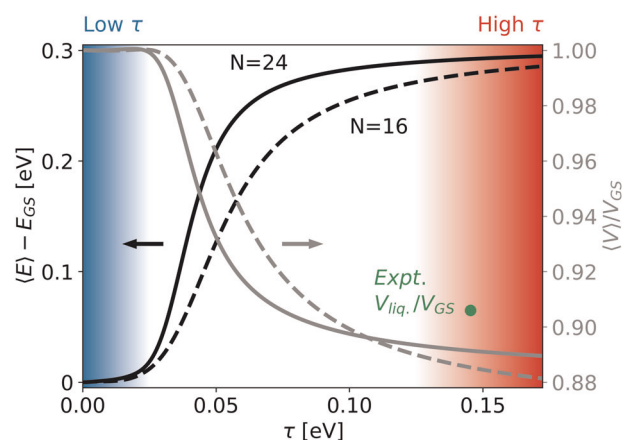
Fig. 2 Evidence of effective order–disorder transition. Energy expectation value and fractional volume per atom as a function of the effective temperature τ. A low τ and high τ regime are clearly discernible, which correspond to structurally ordered and disordered states, respectively.
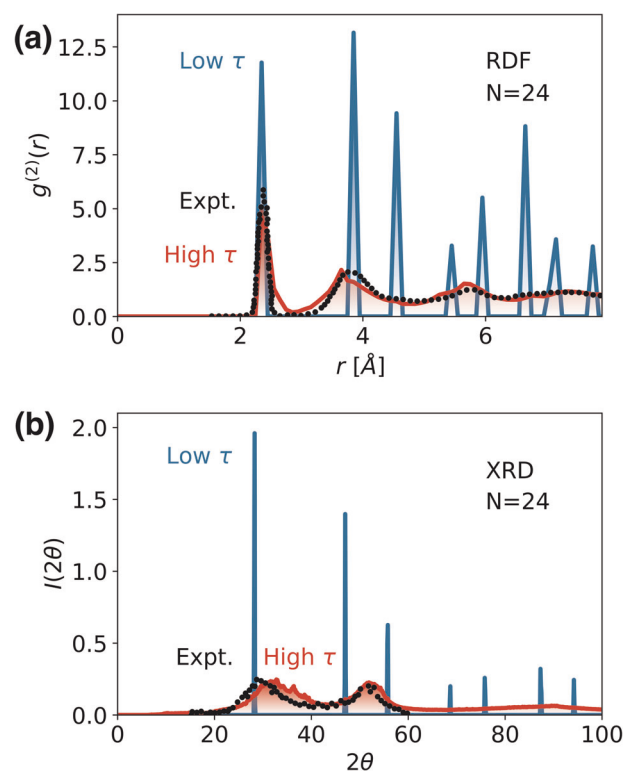
Fig. 3 Accurate prediction of short- and long-range order for ideal, glassy Si. In both the ensemble averaged **a** RDF and **b** XRD patterns, the low τ regime shows the characteristic peaks of diamond silicon while the high τ regime displays remarkable agreement with the experimental RDF and XRD for a-Si.

and unit fractional volume—the high τ regime (red shading) involves a large density of states with energy roughly clustered around ~0.3 eV/atom, leading to a high τ energy expectation, which asymptotes to that value. Meanwhile, the fractional volume in the high τ limit goes to ~0.89.

Given evidence of an effective order–disorder transition, we now proceed to show that the ground (low τ) state corresponds to diamond silicon (d-Si) and that the high τ state corresponds to an ideal silicon glass (g-Si). In Fig. 3a, we show ensemble RDFs for the low τ regime (blue) and the high τ regime (red), averaged

according to Eq. (4) with a unit cell of $N = 24$ atoms. We also plot the experimental RDF for a-Si prepared by ion implantation (black dots)[19]. The distinct crystalline peaks of the low τ RDF confirm that our ground state structure indeed corresponds to d-Si[20]. At high τ, we find that the RDF broadens and overlays the experimental RDF to a rather remarkable degree of accuracy. This agreement in the local ordering of a-Si demonstrates that averaging the local order parameter (RDF) of an ensemble of crystalline arrangements of atoms can be described according to the principle of partial, remnant ergodicity.

Figure 3b contains powder XRD patterns, again for the ensemble low and high τ regimes averaged according to Eq. (5) and for an experimentally prepared a-Si[21]. The low τ peaks again clearly demonstrate the crystallinity of the d-Si ground state[22]. Meanwhile, the high τ averaged XRD faithfully reproduces the two broad experimental humps shown in black. Given that each individual microstate in our ensemble possesses sharp diffraction peaks—due to a long-range order on the periodic 24 atom cell—the two broad humps in the high τ regime must result from the thermal superposition destroying the long-range order found in all of the constituent microstates. Therefore, within the context of our present framework, one can effectively describe the lack of long-range order in a glassy solid as resulting from an incommensurateness in the long-range order of different (periodic) microstates.

Having established that the crystalline ensemble correctly reproduces both the short- and long-range order parameters of a-Si, we turn to a discussion of two other important measured properties of disordered silicon states, the excess enthalpy and the density, in order to elucidate the high τ state as describing g-Si. The excess enthalpy of liquid silicon over that of the crystal has been measured to be ~0.47 eV/atom[23]. Recall that the high τ state has an excess energy of ~0.3 eV. From the perspective that a liquid shares many of the same configurational states with its corresponding glass and differs mostly with respect to some added kinetic energy, the high τ state sits at an energy value identifiable with glassy states potentially accessible by very fast melt quench. Indeed, if we approximate the kinetic energy of an atom in liquid silicon at temperature T as $(3/2)k_BT$ and use the melting temperature of silicon (1687 K), then the kinetic energy per Si atom is 0.22 eV. The estimate of the excess enthalpy of liquid silicon with respect to the crystalline phase is then $0.22 + 0.3 = 0.52$ eV. The measurement of 0.47 eV/atom constitutes a lower bound for the excess enthalpy per atom of liquid Si, with 0.47–0.52 eV/atom the full range cited in ref. [23]. Our calculation of configurational excess enthalpy, coupled with a simple estimate of kinetic excess enthalpy from the equipartition theorem, is therefore able to accurately replicate the excess enthalpy of liquid Si with respect to the crystalline ground state. This view is supported by the fact that the fractional volume asymptotes in the high τ limit to ~0.89 at $N = 24$, representing a 2% error relative to the experimental value of the volume of liquid Si as a fraction of the volume of the diamond Si ($V_{liq.}/V_{GS} \approx 0.91$), shown as a green point in Fig. 2. Note that the effective temperature at which the experimental $V_{liq.}/V_{GS}$ is plotted corresponds to the physical melting temperature of d-Si, $T = 1687$ K[24].

We therefore predict that the ideal g-Si should be more dense than d-Si, consistent with the fact that liquid Si is also more dense. The ideal CRN model also predicts that g-Si should be more dense since any variation in bond angles from perfect tetrahedral coordination densifies the structure[25]. By contrast, a-Si has been measured to be 1.8% less dense than d-Si[26]. This density deficit is attributable to a large concentration of coordination defects in the glassy structure due to the ion implantation[2,19,27,28], which also lower the excess energy of a-Si to around ~0.07–0.15 eV/atom as measured by differential scanning calorimetry[29,30]. A quenching method that is fast enough to produce the true glass transition

should, according to our results and the discussion from ref. [25], create the high τ g-Si.

Cell size and the number of structures

Given all this, one might ask about the dependence of these results on both the number of atoms in the simulation cell and the number of structures needed to accurately represent XRD and RDF of a-Si. To answer these questions, we performed the following exercise. We treated the RDF and XRD from Fig. 3, computed with $N = 24$ over approximately 10,000 random structures, as a reference, and asked how many structures at fixed N are needed to reproduce those curves to within a somewhat arbitrary 0.03 average absolute deviation. The deviation is calculated as an integral $\int |RDF - RDF_{ref.}| dr$ divided by $\int RDF_{ref.} dr$ where $RDF_{ref.}$ is the reference RDF. Analogous equation is used for the XRD.

As expected, larger the cell size used in the random structure sampling less structures one needs to get within a certain tolerance and this is true for both RDF and XRD. However, one does not achieve a given tolerance with the same number of structures. Again, not surprisingly, it takes more structures to converge XRD than RDF, as XRD is generally more sensitive to long-range correlations that are present in finite size structures. From our calculations, it requires only 450 structures with $N = 24$ to get within 0.03 of the reference RDF and about 1300 structures to get converged XRD. With smaller number of atoms, however, it becomes much harder to reach this accuracy. With $N = 16$, one needs about 4300 structures to converge RDF, and even with ~15,000 random structures, the average absolute deviation is hardly <0.05 in XRD. As shown in Fig. 4, this is mainly because of the relatively sharp features in the ensemble averaged XRD patterns that remain for $N = 16$ as well as $N = 12$ implying a level of long-range correlations not present in the glassy state. Further decrease in the cell size to $N = 12$ render achieving the target accuracy unattainable for both XRD and RDF with the total of nearly 10,000 structures we calculated. This result indicates that for a given accuracy there exists a lower limit in the number of atoms needed to describe the disorder in the glassy state. Furthermore, there are also trade-offs between the size and the required number of structures that one needs to be mindful of when employing this methodology to model glassy solids.

Beyond Si

Finally, in order to demonstrate that our results are generally applicable, we show three partial RDFs for vitreous SiO_2 in Fig. 5. The red curve in each panel corresponds to the high τ state while
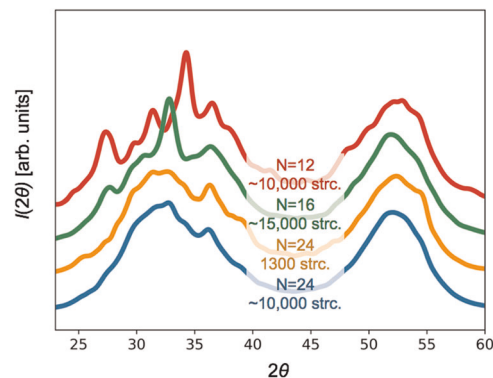
Fig. 4 The ensemble averaged high-τ X-ray diffraction patterns calculated for different unit cell sizes and different number of structures. The curves are manually shifted along the y-axis for clarity. The blue curve ($N = 24$ with ~10,000 structures) is used as a reference while all the others illustrate the size and number of structure dependence (see text for details).
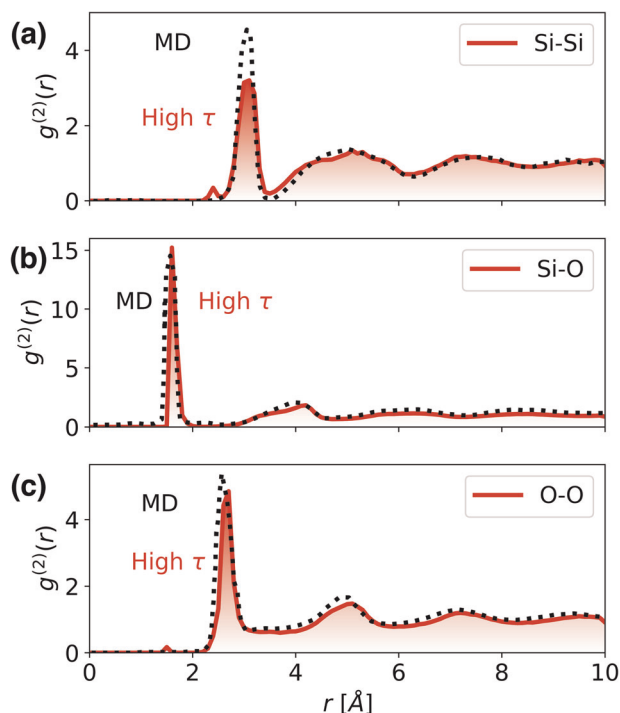
Fig. 5 Partial radial distribution functions for the components of vitreous SiO₂. The three partial RDFs considered are **a** Si–Si, **b** Si–O, and **c** O–O. At high τ, the RDF expectation values agree well with previously calculated molecular dynamics results from Hoang[31].

the dotted black line represents the RDF from a structural model of amorphous silica (a-SiO₂) obtained by a molecular dynamics calculation by Hoang on a 3000-atom unit cell[31]. Here again, the high degree of agreement between the partial RDFs calculated as a single microstate on a large supercell (black) and our high τ RDFs calculated as ensemble averages of ~4000, 24-atom crystalline microstates (red) establishes the mutual compatibility of these viewpoints in describing glassy structure. Furthermore, the first coordination shells of Si with O and O with Si are computed to be ~4.01 and ~2.01 respectively, consistent with known experimental measurements[32]. The small discrepancy in peak heights in the Si–Si and O–O components of Fig. 5a, c are a result of a splitting of the first coordination shells and attendant creation of small peaks below the hard-core radius due to a small set of crystalline microstates with substitutional disorder between Si and O sites. This splitting is expected to wash out in a larger sampling of the potential energy surface. Finally, we mention here that, although our methodology performs well in characterizing structural descriptors of glassy states due to its present form as a configurational theory, dynamically dependent quantities such as specific heat remain most easily and accurately calculated via supercell-based methods such as parallel tempering molecular dynamics[33]. Efficient formulations of ensemble expectations for dynamical observables is an interesting direction for future work.

DISCUSSION

We have shown that both the short- and long-range order of covalent and ionic glassy solids can be accounted for by taking thermal averages over an ensemble of crystalline microstates, thus validating the concept of remnant partial ergodicity in these systems. We have also laid out the framework and the approximations within which the ensemble picture rigorously follows from the statistical treatment of fully ergodic systems. The partial ergodicity is taken into account by considering the potential enthalpy of the system and including only low-energy

local minima obtained through the first-principles random structure sampling. Such an approach is consistent with the principle of remnant ergodicity within a region of configuration space and affords a fully first principles (no fitting to experiments) accounting of the structural features of glassy Si and SiO₂. At the same time, the crystalline ensemble model opens the door to calculating functional properties of glasses by averaging over the contributions to those properties carried by each microstate in the ensemble.

METHODS

Computational random structure sampling

The computational construction of an ensemble of structural microstates and attendant evaluation of ensemble averages utilizes the first-principles random structure sampling to generate the atomic configurations that reside at local minima, calculate their enthalpy h_α, and assess their associated basin hypervolumes f_α relative to one another. We remark that these three steps have been thoroughly detailed in our previous work[17,18], so we simply sketch them here.

The first step is to generate large number of random initial structures with a fixed number of atoms in a unit cell N. For Si, we will show results for $N = 16$ and 24, while for SiO₂ we will cite results just for $N = 24$. We find that $N = 24$ is sufficient in each case to attain converged RDF and XRD patterns via ensemble averaging. Once N has been chosen, a random unit cell geometry is specified by choosing randomly three lattice constants, a, b, c and the corresponding angles α, β, γ. The unit cell is then populated (quasi)randomly with N number of atoms. In case of Si, the population is truly random, while for SiO₂ we utilize the random superlattice construction[17] that biases the structures toward predominant Si–O coordination. For Si, 15,000 such random structures were generated with $N = 16$ atoms in the unit cell and 10,000 structures were generated for $N = 24$. For SiO₂, 4000 random structures were generated with $N = 24$. Volume, ionic, and cell-shape degrees of freedom were relaxed for all initialized random structures using the VASP implementation of density functional theory (DFT) with the projector augmented wave method and Perdew, Burke, Ernzerhof exchange correlation functional[34–36]. Calculations were restarted until total energy converged to within 3 meV/atom between successive ionic steps and until residual forces and pressures were below 10^{-4} eV/atom and 3 kbar, respectively. Once relaxed, the enthalpy was then evaluated at each minimum using the DFT-calculated total energy and volume per atom $h_\alpha = E_\alpha + pV_\alpha$.

Evaluation of ensemble averages

Finally, note that, in the evaluation of an ensemble average, atomic structures that show up more than once will naturally generate a multiplicity f_α relative to the total number of structures relaxed owing to the additivity of equivalent Boltzmann factors, i.e., for some observable A:

$$\langle A \rangle = \sum_\alpha A_\alpha \frac{e^{-h_\alpha/\tau}}{\Xi} = \sum_{\alpha'} A_{\alpha'} f_{\alpha'} \frac{e^{-h_{\alpha'}/\tau}}{\Xi}, \qquad (6)$$

where α here indexes the crystal structure at each local minimum just as it is found after the initial random structure sampling process and α' indexes equivalence classes of local minimum such that crystal structures found within the equivalence class are the same. The multiplicity f_α therefore quantifies the fraction of configuration space, which leads to the structure α' as the stable local minimum. Therefore, we do not need to sort structures into equivalence classes before taking the average and the relative size of different basins of attraction is naturally included.

DATA AVAILABILITY

CODE AVAILABILITY

REFERENCES

1. Zallen, R. *The Physics of Amorphous Solids* (John Wiley & Sons, 2008).
2. Pedersen, A., Pizzagalli, L. & Jónsson, H. Optimal atomic structure of amorphous silicon obtained from density functional theory calculations. *New J. Phys.* **19**, 063018 (2017).
3. Tu, Y., Tersoff, J., Grinstein, G. & Vanderbilt, D. Properties of a continuous-random-network model for amorphous systems. *Phys. Rev. Lett.* **81**, 4899 (1998).
4. Gupta, P. K. Non-crystalline solids: glasses and amorphous solids. *J. Non-Cryst. Solids* **195**, 158–164 (1996).
5. Dorner, F., Sukurma, Z., Dellago, C. & Kresse, G. Melting Si: beyond density functional theory. *Phys. Rev. Lett.* **121**, 195701 (2018).
6. Sjostrom, T., De Lorenzi-Venneri, G. & Wallace, D. C. Potential energy surface of monatomic liquids. *Phys. Rev. B* **98**, 054201 (2018).
7. DeGiuli, E. Field theory for amorphous solids. *Phys. Rev. Lett.* **121**, 118001 (2018).
8. Ozawa, M., Ikeda, A., Miyazaki, K. & Kob, W. Ideal glass states are not purely vibrational: insight from randomly pinned glasses. *Phys. Rev. Lett.* **121**, 205501 (2018).
9. Mauro, J. C. & Smedskjaer, M. M. Statistical mechanics of glass. *J. Non-Cryst. Solids* **396**, 41–53 (2014).
10. Greer, A. L. Confusion by design. *Nature* **366**, 303 (1993).
11. Porai-Koshits, E. Genesis of concepts on structure of inorganic glasses. *J. Non-Cryst. Solids* **123**, 1–13 (1990).
12. Sanchez, J. M. & de Fontaine, D. Ising model phase-diagram calculations in the fcc lattice with first- and second-neighbor interactions. *Phys. Rev. B* **25**, 1759–1765 (1982).
13. Sanchez, J., Ducastelle, F. & Gratias, D. Generalized cluster description of multi-component systems. *Phys. A Stat. Mech. Appl.* **128**, 334–350 (1984).
14. Nelson, L. J., Ozoliņš, V., Reese, C. S., Zhou, F. & Hart, G. L. W. Cluster expansion made easy with bayesian compressive sensing. *Phys. Rev. B* **88**, 155105 (2013).
15. Perim, E. et al. Spectral descriptors for bulk metallic glasses based on the thermodynamics of competing crystalline phases. *Nat. Commun.* **7**, 12315 (2016).
16. Ziman, J. M. *Models of Disorder: The Theoretical Physics of Homogeneously Disordered Systems* (CUP Archive, 1979).
17. Stevanović, V. Sampling polymorphs of ionic solids using random superlattices. *Phys. Rev. Lett.* **116**, 075503 (2016).
18. Jones, E. B. & Stevanović, V. Polymorphism in elemental silicon: probabilistic interpretation of the realizability of metastable structures. *Phys. Rev. B* **96**, 184101 (2017).
19. Laaziri, K. et al. High resolution radial distribution function of pure amorphous silicon. *Phys. Rev. Lett.* **82**, 3460 (1999).
20. Laaziri, K. et al. High-energy x-ray diffraction study of pure amorphous silicon. *Phys. Rev. B* **60**, 13520 (1999).
21. Hatchard, T. & Dahn, J. In situ XRD and electrochemical study of the reaction of lithium with amorphous silicon. *J. Electrochem. Soc.* **151**, A838–A842 (2004).
22. Westra, J., Vavruňková, V., Šutta, P., Van Swaaij, R. & Zeman, M. Formation of thin-film crystalline silicon on glass observed by in-situ XRD. *Energy Procedia* **2**, 235–241 (2010).
23. Veryatin, U. et al. Thermodynamic properties of inorganic substances. *Manual Atomizdat* 118 (1965).
24. Ubbelohde, A. R. *The Molten State of Matter: Melting and Crystal Structure* (John Wiley & Sons, 1978).
25. Popescu, M., Sava, F. & Hoyer, W. Amorphous semiconductors are denser? *Dig. J. Nanomater. Biostruct.* **1**, 49–54 (2006).
26. Custer, J. et al. Density of amorphous Si. *Appl. Phys. Lett.* **64**, 437–439 (1994).
27. Brodsky, M., Kaplan, D. & Ziegler, J. Densities of amorphous Si films by nuclear backscattering. *Appl. Phys. Lett.* **21**, 305–307 (1972).
28. Williamson, D. et al. On the nanostructure of pure amorphous silicon. *Appl. Phys. Lett.* **67**, 226–228 (1995).
29. Roorda, S., Doorn, S., Sinke, W., Scholte, P. & Van Loenen, E. Calorimetric evidence for structural relaxation in amorphous silicon. *Phys. Rev. Lett.* **62**, 1880 (1989).
30. Kail, F. et al. The configurational energy gap between amorphous and crystalline silicon. *Phys. Status Solidi Rapid Res. Lett.* **5**, 361–363 (2011).
31. Hoang, V. V. Molecular dynamics simulation of amorphous SiO_2 nanoparticles. *J. Phys. Chem. B* **111**, 12649–12656 (2007).
32. Prescher, C. et al. Beyond sixfold coordinated Si in SiO_2 glass at ultrahigh pressures. *Proc. Natl Acad. Sci. USA* **114**, 10041–10046 (2017).
33. Rathore, N., Chopra, M. & de Pablo, J. J. Optimal allocation of replicas in parallel tempering simulations. *J. Chem. Phys.* **122**, 024111 (2005).
34. Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).
35. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979 (1994).
36. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).

AUTHOR CONTRIBUTIONS
E.B.J. provided the analytic formulation and numerical computations. V.S. designed, oversaw, and verified the results of the research. Both authors contributed to the manuscript.

COMPETING INTERESTS
The authors declare no competing interests.

ADDITIONAL INFORMATION
Supplementary information is available for this paper at https://doi.org/10.1038/s41524-020-0329-2.

Correspondence and requests for materials should be addressed to V.S.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.