

Fractality and percolation in the phase transition of polycrystalline graphite

Vilmos Neuman¹, Zuzanna Malinowska-Trzmielak¹, and Mark Wilson^{1*}

Department of Chemistry, Physical and Theoretical Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QZ, United Kingdom



(Received 28 April 2026; revised 1 July 2026; accepted 8 July 2026; published 10 August 2026)

The pressure- and temperature-driven phase transition from polycrystalline graphite to diamond is investigated using molecular dynamics computer simulation. We partition the atomic coordinates into two disjoint subsets corresponding to a diminishing three-coordinate and a developing four-coordinate network. It is shown that the fractal dimension and its time evolution, as calculated from the radial distribution function “colored” for four coordinates only, tracks the transformation mechanism, capturing the growth and spatial organization of the developing domains. In particular, it reflects the increasing space-filling character and homogenization of the emerging diamondlike structure. At sufficiently high pressures and temperatures, the system consistently transforms into polycrystalline diamond, whereas at lower pressures the final structures depend sensitively on both temperature and initial microstructure, leading to greater variability in the fractal dimension. Additionally, we demonstrate that the transformation to four-coordinate sites leads to a percolation transition, which, compared to a simple hard sphere null model, occurs at a significantly lower threshold, highlighting the role of the interatomic potential in correlating network connectivity. These results demonstrate that the fractal dimension may provide a simple and robust metric for tracking topological changes during phase transitions in polycrystalline solids or other phenomena that may be characterized as domain growth.

DOI: [10.1103/6zv5-drsn](https://doi.org/10.1103/6zv5-drsn)

I. INTRODUCTION

The phase behavior of carbon is a notoriously challenging problem and has been a longstanding area of research, both theoretical and experimental [1–4]. This difficulty stems from carbon’s exceptional bonding versatility across multiple hybridization states. As a result, carbon displays numerous crystalline allotropes, extended polyamorphism, and rich cluster chemistry [5–7]. In this contribution, we discuss a phase transition of polycrystalline graphite under high-pressure, high-temperature (HPHT) conditions, with a particular focus on the developing network of often diamondlike, four-coordinated carbon atoms. Phase transitions under these conditions are of particular interest since they can result in composite diamond-graphite structures, termed diaphites, which are predicted to have valuable, technologically relevant properties [8,9]. However, the mechanisms underlying their formation are not yet well understood, and synthesis remains difficult to control. We have explored elsewhere the detailed formation mechanism for diaphites from polycrystalline graphite [10].

In what follows, we discuss results obtained from molecular dynamics (MD) simulations using an atomic cluster expansion (ACE) [11] machine-learned interatomic potential

[12] that has previously been benchmarked against *ab initio* calculations for diaphites [13]. In particular, we highlight the relationship between an apparent power law tail of the partial radial distribution function “colored” for four-coordinate sites only and the box counting fractal dimension, which measures how space filling the developing network is. We demonstrate that the fractal dimension is a useful structural metric that provides insight into the structure and dynamics of the four-coordinate network, and we compare it with simpler independent measures, such as spatial homogeneity and density. We then show that the domain growth is also reflected in the corresponding partial powder average structure factors. We discuss the relationship between the fractal dimension and the temperatures and pressures achieved in the simulations and relate it directly to the formation mechanism of the diaphitic phase. Finally, we show that the four-coordinated carbon network undergoes a percolation transition at a characteristic threshold, which we compare with a hard sphere null model.

II. THEORY

The radial distribution function, $g(r)$, is computed as a spherical average over reference particles

$$g(r) = \frac{2}{\rho N} \left\langle \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{\delta(r - |\mathbf{r}_{ij}|)}{4\pi r^2} \right\rangle, \quad (1)$$

where ρ is the number density of the particles considered, N is the number of such particles, and $\mathbf{r}_{ij}$ is the vector connecting particles i and j the modulus of which corresponds to the interparticle separation. This immediately gives the number of particles, $dN(r)$, within a spherical shell of radius r and

*Contact author: mark.wilson@chem.ox.ac.uk

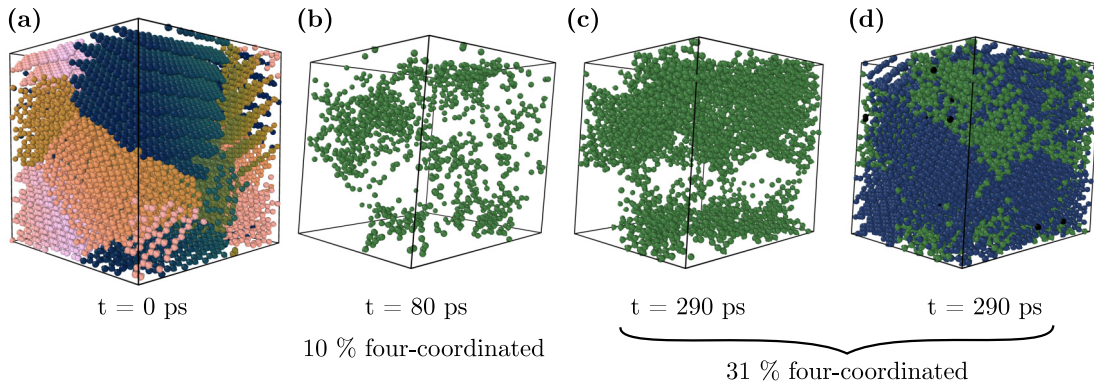


FIG. 1. Time ordered snapshots from a representative molecular dynamics trajectory of the phase transition in a graphitic polycrystal containing five grains heated to 3000 K and pressurized to 27.5 GPa. Panel (a) shows the initial fully three-coordinate polycrystalline configuration with atoms coloured by grain identity whereas (b) and (c) show the evolution of the network of four-coordinate carbons. Panel (d) shows the final equilibrated configuration with three- and four-coordinates colored blue and green, respectively.

infinitesimal thickness dr as

$$dN(r) = 4\pi\rho \int_r^{r+dr} dr' g(r') \cdot r'^2 = 4\pi r^2 g(r) \rho dr. \quad (2)$$

For a set of points embedded in two or three dimensional space subdivided by a grid of boxes of side length ε , the box-counting fractal dimension, d_f , is defined in terms of the minimum number of such boxes, $N(\varepsilon)$, required to fully cover the set. If, for sufficiently small ε , the scaling

$$N(\varepsilon) \sim \left(\frac{\ell}{\varepsilon}\right)^{d_f} \quad (3)$$

holds with ℓ an observation length over which the fractal scaling is measured, then d_f , the fractal dimension, is defined by

$$d_f = \lim_{\varepsilon \rightarrow 0} \frac{\ln N(\varepsilon)}{\ln(\ell/\varepsilon)}. \quad (4)$$

Simple objects' fractal dimensions are familiar ones. For a square of side length ℓ , $N(\varepsilon) \approx (\ell/\varepsilon)^2$, so $d_f = 2$; for a cube, $N(\varepsilon) \approx (\ell/\varepsilon)^3$ and $d_f = 3$. More generally, d_f need not be an integer. Fractional values indicate a set whose scaling behavior lies between integer dimensions: for example, one might expect a collection of nonparallel sheets of finite thickness, or a three-dimensional structure containing voids [see Figs. 1(b) and 1(c)], to have a dimension between two and three. Therefore, d_f quantifies a degree of space filling.

The relationship between d_f and the radial distribution function has been discussed in the aerogel and scattering literature [14–16], and the derivation presented below proceeds in a manner similar to that of Kieffer and Angell in Ref. [17]. The ratio ℓ/ε is a dimensionless measure of scale. It is the length over which we observe the fractal scaling expressed in units of the resolution ε , which we shall choose to be physically meaningful, on the order of the atomic separation. Therefore, in the fractal regime, the scaling of Eq. (3) may be rewritten equivalently as

$$N(r) \sim A r^{d_f}, \quad (5)$$

where A is a constant of dimensions length^{-d_f} . Taking the logarithm and differentiating with respect to $\ln r$ gives

$$\frac{d \ln N(r)}{d \ln r} = d_f. \quad (6)$$

Using Eq. (2), $dN(r)/dr = 4\pi r^2 \rho g(r)$, and hence by the chain rule,

$$\frac{dN(r)}{d \ln r} = r \frac{dN(r)}{dr} = 4\pi r^3 \rho g(r). \quad (7)$$

Using the chain rule again and combining Eqs. (6) and (7) yields

$$d_f N(r) = 4\pi r^3 \rho g(r). \quad (8)$$

Now we make the important assumption that there exists a range of r over which d_f is constant. Differentiating over this range gives

$$d_f \frac{dN(r)}{dr} = 4\pi \rho \left(3r^2 g(r) + r^3 \frac{dg(r)}{dr} \right). \quad (9)$$

Substituting $\frac{dN}{dr} = 4\pi r^2 \rho g(r)$ and dividing by $4\pi r^2 \rho g(r)$ yields

$$d_f = 3 + \frac{r}{g(r)} \frac{dg(r)}{dr} = 3 + \frac{d \ln g(r)}{d \ln r}. \quad (10)$$

Since a straight line in a logarithmic plot corresponds to a power law scaling, the assumption of a range of r over which the fractal dimension is constant is equivalent to assuming that the radial distribution function scales as

$$g(r) \sim r^{-\chi}, \quad \text{with} \quad \chi = 3 - d_f \quad (11)$$

over this range. Crucially, the power law regime must be limited at both small and large length scales, being constrained at small scales by the finite atomic radii and at large scales by the attainment of bulk density, where pair correlations vanish. Consequently, $g(r)$ must converge to unity at large length scales, which is inconsistent with a power law form which can therefore only be physically meaningful over an intermediate range of r [14,15]. This problem is particularly serious for quantities that are to be experimentally observable since Eq. (11) has the unfortunate consequence of making the

isotropically averaged structure factor diverge in the small-angle limit, implying correlations that extend to arbitrarily large length scales. A convenient way to treat the existence of this finite r range in the structure factor, is to introduce an ansatz for the total correlation function, $h(r) = g(r) - 1$, in which a finite correlation length ξ , cuts off the power law tail by an exponential decay [14,15,18] in analogy with the Ornstein-Zernike correlation of critical phenomena [19],

$$h(r) \sim r^{d_f-3} e^{-r/\xi}. \quad (12)$$

This yields an analytical form [14] for the powder average structure factor,

$$S(k) = 1 + 4\pi\rho \int_0^\infty dr h(r) r^2 \frac{\sin(kr)}{kr}, \quad (13)$$

where k is the magnitude of the wave vector. This yields the small angle behavior of a mass fractal with an exponential cut off

$$S(k) = 1 + \frac{d_f \Gamma(d_f - 1)}{(kr_0)^{d_f}} \frac{\sin((d_f - 1) \arctan(k\xi))}{\left(1 + \frac{1}{(k\xi)^2}\right)^{(d_f-1)/2}}. \quad (14)$$

Here, Γ is the gamma function and r_0 is the lower cut off of the fractal scaling. Equation (14) now shows the fractal power-law scattering $S(k) \sim k^{-d_f}$ that is expected based on scaling arguments for mass fractals [15] for $k\xi \gg 1$, whereas if $k\xi \ll 1$ the structure factor approaches a constant, reflecting the restoration of a uniform bulk density beyond the length scale ξ . The structures discussed below also display coordination structure at short length scales and effective fractality on an intermediate length scale, extending beyond the range of short-range coordination correlations but not to distances sufficiently large for pair correlations to have decayed.

III. COMPUTATIONAL DETAILS

Polycrystalline graphite models were constructed by Voronoi tessellation as implemented in ATOMSK [20–22] using randomly placed seeds in a periodic cubic cell of side length 45 Å. From an initial ensemble of 50 configurations, the ten most structurally diverse ones were chosen based on Wasserstein distances between persistence diagrams within the HomCloud [23] package. Five structures with five grains and five with ten grains were chosen, each containing $\sim 10^4$ (10 016–10 166) atoms.

Molecular dynamics simulations were performed in LAMMPS [24] using an ACE [11] potential for carbon. Structures were first relaxed by conjugate gradient minimization converged with thresholds in energy $< 10^{-12}$. All simulations used periodic boundary conditions, a 1 fs time step, and the NPT ensemble with a Nosé-Hoover thermostat [25,26] (1 ps damping) and fully anisotropic Nosé-Hoover barostat [27,28] (10 ps damping). Each trajectory comprised 20 ps equilibration at 1 bar and 300 K, 50 ps pressurization, 50 ps heating, 50 ps at target (T, p), 100 ps depressurization/cooling to ambient conditions, and a final 20 ps equilibration at 1 bar and 300 K. Simulations were performed for $T \in \{1500, 2000, 2500, 3000, 3500\}$ K and $p \in \{25, 27.5, 30, 32.5, 35, 40\}$ GPa for all ten polycrystals yielding 300 simulations total. The temperature and pressure regime was chosen to be similar to experimentally reported

values [1] although are somewhat harsher to account for the short timescale of the simulations.

Structural analysis was performed using OVITO [29] and diamondlike environments were identified via a common neighbor analysis [30].

IV. RESULTS AND DISCUSSION

A. The fractal dimension

In the following, we shall define a bond between any two atoms if they are separated by at most 1.7 Å. This cutoff was chosen as the minimum distance that reproduces the expected coordination numbers, consistent with common neighbor analysis [30] used to identify diamondlike domains and also happens to encompass the entirety of the first coordination peaks observed in the radial distribution functions. Upon pressurization and heating, consistent with previous reports [31], the graphitic polycrystal initially forms four-coordinated carbon atoms at the grain boundaries. These act as seeds and, with further pressurization and heating, promote domain growth of the four-coordinated networks. Interestingly, the grain structure remains essentially unchanged during the transformation, such that if graphitic grains do become diamondlike, whole grains tend to do so thereby leaving the positions of the grain boundaries basically unchanged. This mechanism is explored elsewhere in detail [10]. We distinguish here between four-coordinated atoms in crystalline, diamondlike environments and those in amorphous environments lacking long-range order. Within our simulations, most equilibrated diaphitlike systems contain both types, typically with a smaller diamondlike domain surrounded by amorphous four-coordinated carbon. The structure of this amorphous four-coordinated phase is itself of interest, but is not explored here and related structures have been discussed elsewhere in the literature [32]. Figure 1 shows snapshots from a representative MD trajectory, illustrating the time evolution of domains of four-coordinated carbon (atoms colored in green). Panel (a) shows an example initial polycrystalline graphitic structure containing five domains. Panels (b)–(d) highlight how the number of four-coordinate atom sites increases with time (here at 80 ps and 290 ps, respectively). Figure 2 shows that the total radial distribution function changes from that of an initially highly crystalline configuration [Fig. 2(a)], with pronounced (δ -like) peaks, to one more characteristic of an amorphous material [Fig. 2(b)]. In particular, a clear shoulder develops on the first coordination peak, consistent with contributions from both diamond and graphite C–C bond lengths. Additional useful information can be extracted by defining radial distribution functions “colored” in terms of the local coordination numbers. In the present work, the most useful such function defines the spatial correlation between pairs of four-coordinate atoms, $g_{44}(r)$. Figure 2(c) shows $g_{44}(r)$, emphasizing its apparent power law tail between 7.5 and 14 Å. We take this interval as the relevant length scale over which we analyze fractality. In the logarithmic plot, the putative power law regime appears as a straight line, up to numerical noise.

To explore the physical meaning of the four-coordinated fractal dimension, d_f , for the systems studied here, we compare it with two structural parameters independent from how

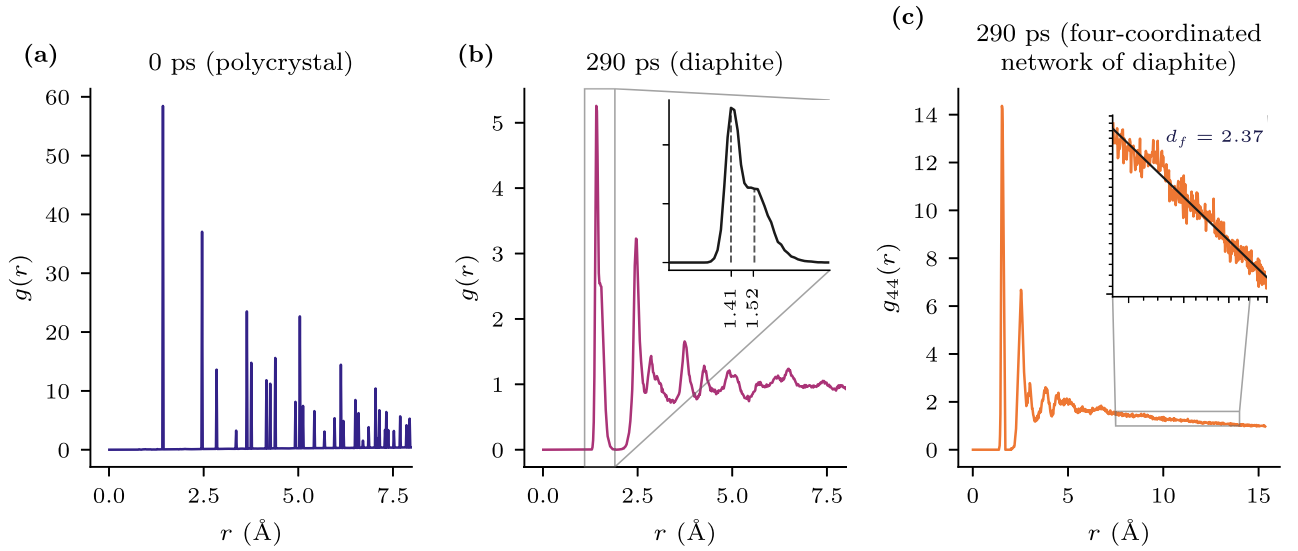


FIG. 2. (a) Total radial distribution function of the initial polycrystalline graphite configuration shown in Fig. 1(a). (b) Total radial distribution function of the same system after pressurization, heating, and equilibration, with the shoulder on the first coordination peak shown in the inset. The structure corresponding to this $g(r)$ is displayed in Fig. 1(d). (c) Radial distribution function of the four-coordinated subnetwork for the configuration in (b) shown pictorially in Fig. 1(c). The inset shows a log-log plot with a linear fit through the power-law regime, corresponding to a fractal dimension of $d_f = 2.37$.

we compute d_f . A set of points embedded in three dimensional space, here a proposed mass fractal, cannot have a fractal dimension greater than the embedding dimension. The fractal dimension measures how rapidly the mass of the set scales with an observation distance r , and therefore quantifies how completely the set fills the space available to it. For example, a close packed arrangement of hard spheres should exhibit $d_f = 3$ over all physically relevant length scales. Since HPHT treatment of the graphitic polycrystal leads to domain growth, the fractal dimension of the four-coordinated network should also reflect how homogeneously the four-coordinated atoms are distributed within the simulation cell. As coarsening proceeds, the volume occupied by three coordinated atoms decreases, d_f approaches three, and the distribution of four-coordinates becomes essentially homogeneous throughout the cell. To assess the above intuition, we compare the fractal dimension with a simple measure of spatial homogeneity. The simulation cell is partitioned into 125 subcells each with volume v . These are parallelepipeds in general for simulations with six cell degrees of freedom. The particle count $n_{i,4}$ in each bin is recorded giving a local number density of four-coordinated atoms as

$$\rho_{i,4} = \frac{n_{i,4}}{v}, \quad \bar{\rho}_4 = \frac{1}{125} \sum_{i=1}^{125} \rho_{i,4}. \quad (15)$$

As a dimensionless measure of nonuniformity we use the squared coefficient of variation,

$$U_{CV}^2 = \frac{\text{Var}(\rho_4)}{\bar{\rho}_4^2} = \frac{\sum_{i=1}^{125} (\rho_{i,4} - \bar{\rho}_4)^2}{125 \cdot \bar{\rho}_4^2}, \quad (16)$$

for which $U_{CV} = 0$ corresponds to a perfectly uniform density field. Even for a uniform perfect gas, finite sampling produces counting noise that we approximate as Poisson with

$$\text{Var}(n_{i,4}) \approx \langle n_{i,4} \rangle, \text{ giving}$$

$$\text{Var}(\rho_{i,4}) \approx \frac{\bar{\rho}_4}{v}. \quad (17)$$

We therefore define an excess inhomogeneity, U^* , of the four-coordinated subnetwork, corrected for shot noise, as

$$(U^*)^2 = \frac{\text{Var}(\rho_4)}{\bar{\rho}_4^2} - \frac{1}{\bar{\rho}_4 v}, \quad U^* = \sqrt{\max(0, (U^*)^2)}. \quad (18)$$

We also compare the fractal dimension with a dimensionless number density, defined as

$$\eta_4 = \frac{n_4}{V} \cdot V_{\text{exc}}, \quad (19)$$

where n_4 is the number of four-coordinated atoms in a simulation cell of volume V , and V_{exc} is the volume of a sphere of radius $1.2 \text{ \AA}$, taken as an excluded volume length scale inferred from the total radial distribution function. Note that η_4 is a dimensionless number density rather than a packing fraction: excluded volumes may overlap, so η_4 can exceed unity.

Figure 3 confirms the expected correlation between the fractal dimension, U^* , and η_4 . It also shows that simulations of five grained structures sample configurations with lower fractal dimensions than their ten grain counterparts. This likely reflects the fact that, under pressurization, four-coordinated networks nucleate at grain boundaries. Ten grain systems contain more grain boundaries, making the initiation of four-coordination more likely and leading to a more spatially uniform distribution of the initial four-coordinated regions. We note that, both here and throughout the following discussion, fractal dimensions slightly exceeding three are simply indicative of the difficulty in fitting the near horizontal line in $g_{44}(r)$.

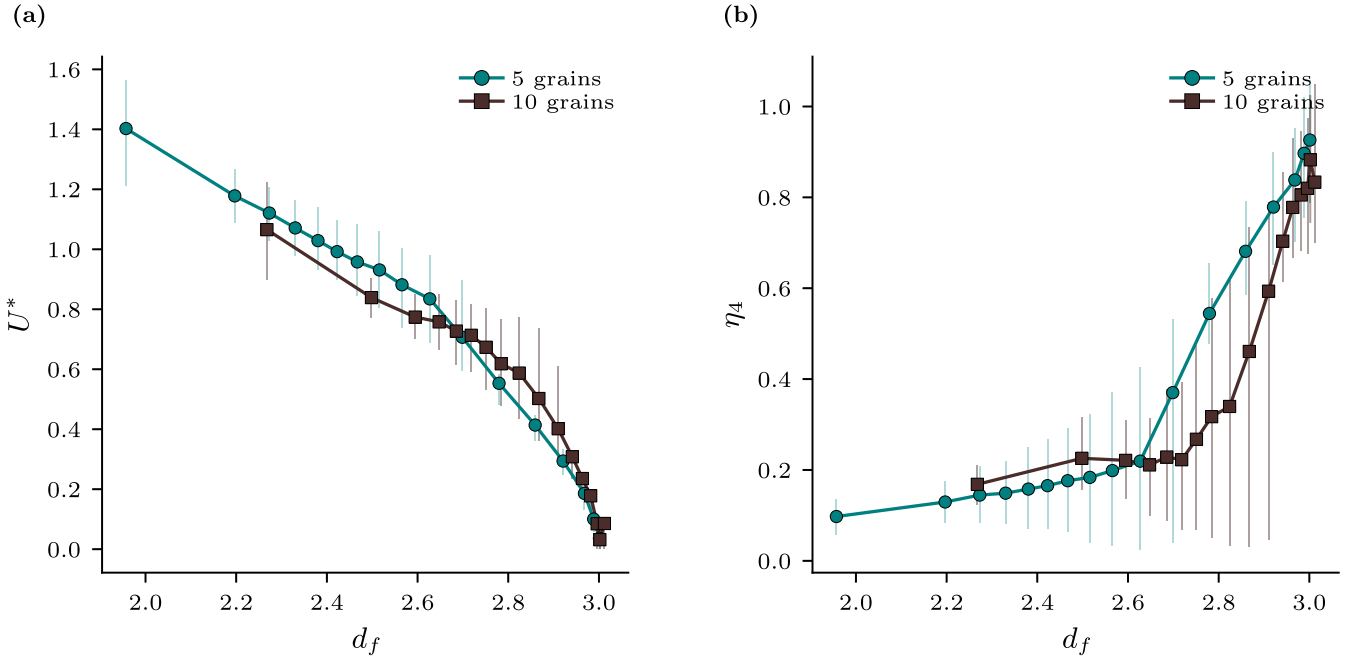


FIG. 3. (a) Spatial homogeneity metric of the four-coordinated subnetwork, U^* , plotted against the fractal dimension, averaged over all sampled simulation frames. (b) Reduced density of four-coordinated atoms, η_4 , against the fractal dimension averaged in the same manner. Error bars indicate the spread of data around each mean average value.

The temperature and pressure dependence of the fractal dimension is also largely as expected, as shown in Fig. 4. At higher temperatures and pressures, the four-coordinated network becomes more space filling (i.e., less domainlike), driving d_f toward three and causing U^* to vanish. The less domainlike character of the four-coordinated network at high temperatures is also reflected in the four-coordinated structure factors shown in Fig. 5, which exhibit increased small angle scattering at lower temperatures. This indicates stronger correlations (density fluctuations) over long length scales, consistent with well defined, and relatively large, domains. The four-coordinated structure factor is computed from the Fourier components as

$$S_{44}(\mathbf{k}) = \frac{1}{N_4} \left\langle \left| \sum_{j=1}^{N_4} e^{i\mathbf{k} \cdot \mathbf{r}_j} \right|^2 \right\rangle, \quad (20)$$

where $\mathbf{r}_j$ are the positions of the four-coordinated atoms. The powder average $S_{44}(k)$ is then obtained by averaging $S_{44}(\mathbf{k})$ over all reciprocal vectors with the same magnitude $k = |\mathbf{k}|$. In a computer simulation in a periodic box, allowed wave vectors of the density fluctuations are given by the reciprocal lattice of the simulation cell. In general for a triclinic cell, writing the cell vectors as columns of the matrix $\mathbf{H} = [\mathbf{a} \ \mathbf{b} \ \mathbf{c}]$, gives

$$\mathbf{k} = 2\pi \mathbf{H}^{-T} \mathbf{n}, \quad \mathbf{n} = (n_1, n_2, n_3)^T \in \mathbb{Z}^3. \quad (21)$$

This highlights the well known limitation of only a small number of wave vectors being available in the small angle regime. As a result, although Eq. (14) suggests that the fractal dimension can, in principle, be extracted from both the structure factor and its real space counterpart $g(r)$, we find the small

angle data too noisy to support a reliable fit. We therefore propose that, in simulations, the radial distribution function is the more robust route for estimating the fractal dimension, since it does not suffer from wave vector “quantization” (i.e., the limitation by the cell geometry) nor does it require an ansatz for the total correlation function, for example, as in Eq. (12).

To conclude this section, we discuss a somewhat unexpected aspect of the domain growth process. Figures 4(a) and 4(b) show that, at low pressures, increasing the temperature decreases the fractal dimension and increases the nonuniformity of the four-coordinated network. Figure 4(c) highlights that this is not due to fewer four-coordinated atoms forming at higher temperature: the fractal dimension decreases even as the number of four-coordinated atoms increases. We attribute this behavior to the fact that it is in this regime that diamondlike, four-coordinated, domains begin to form and persist in the equilibrated phases. At this low pressure, these domains tend to be small and appear only in a subset of simulation trajectories. Compared with the four-coordinated sites that initially form on the grain boundaries upon pressurization and remain only transiently at low pressure, these domains are spatially confined and are therefore characterized by greater nonuniformity and a lower fractal dimension. As the temperature increases at low pressure, a larger fraction of equilibrated structures contain domains, albeit small ones, rather than only a few scattered four-coordinated sites on the grain boundaries, leading to a clear downward trend in d_f . This effect is further enhanced by the fact that, in the highest-temperature simulations at 25 GPa, if domains do form, they tend to be predominantly diamondlike, which is a highly dense phase. At high pressures, even low-temperature conditions tend to produce significant domains; consequently,

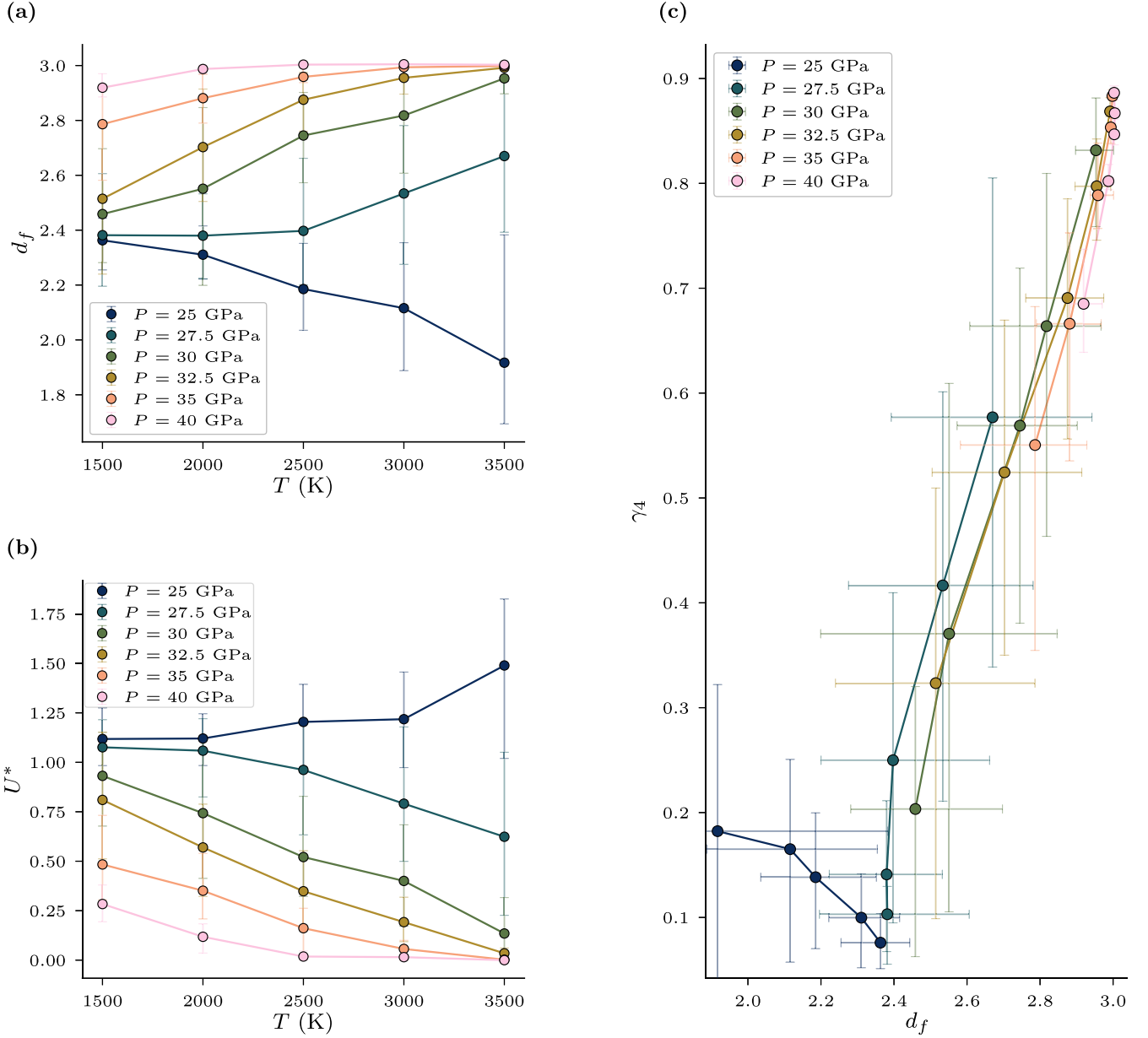


FIG. 4. (a), (b) Plots of the fractal dimension and the spatial inhomogeneity metric, U^* , respectively, as functions of temperature for given pressures. Values are averaged over the final 20 ambient equilibration frames of each simulation. (c) Data from the same configurations shown as a plot of the average fraction of four-coordinated atoms, γ_4 , against the four-coordinated fractal dimension. Error bars indicate the spread of data around each mean average value.

increasing the temperature simply promotes the coarsening of these domains or the formation of additional ones, enabling the four-coordinated network to fill the simulation cell more efficiently. The increase in temperature does increase the number of four-coordinate sites in the low-pressure regime, but it does so in a way that makes them distributed less uniformly compared to lower temperatures. The spread of the data around each point, represented by the error bars in Fig. 4, is also noteworthy and reinforces this interpretation. Very large spreads across all panels of Fig. 4 occur for the highest temperature simulations at the pressure of 25 GPa. This is because, under these conditions, some structures become diaphitic, while others remain almost entirely

graphitelike, leading to considerable structural diversity in the equilibrated systems. By contrast, the “minimum” in Fig. 4(c) shows small error bars, since it corresponds to conditions under which almost all simulations result in entirely graphitic structures, as the temperature and pressure applied to the system were not high enough to promote domain growth. Large error bars are also observed for intermediate values of d_f , since this regime is associated with a wide variety of diaphitic phases with markedly different topologies of the four-coordinated network. The harshest conditions, for which d_f converges to three, are associated with vanishing error bars, as all systems become almost entirely four coordinated. In essence, the spread in d_f reflects the degree to which the final

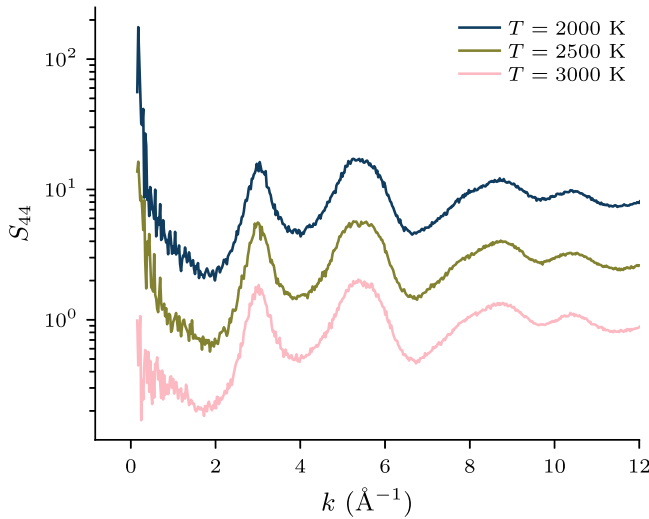


FIG. 5. Powder average four-coordinated structure factor shown on a vertical log scale, computed by averaging over the final 20 ambient equilibration configurations from three simulations of a representative system subjected to different temperatures.

configurations depend on the starting structure under a given set of temperature and pressure conditions.

Given that the four-coordinate fractal dimension proved useful for characterizing the equilibrated structures obtained from the simulations, we also analyzed its time series, $d_f(t)$, to elucidate the dynamics of the transformation mechanism. Figures 6(a) and 6(b) show the time evolution of the fractal dimension for both five and ten grain systems in our simulations. The four-coordinate fractal dimension is well defined from the onset of the pressurization ramp, since in the equilibrated polycrystalline structures, a fraction of atoms at the grain boundaries are already constrained to adopt four-coordinate environments. The initial fractal dimension is larger in the ten-grain systems, as these contain more grain boundary area, and hence atoms that are distributed more uniformly throughout the simulation cell. For both grain architectures, d_f initially increases as the system contracts under pressure, leading to an increase in the fraction of grain-boundary atoms that are four coordinate, resulting in a more homogeneously distributed four-coordinate network. The time series subsequently hits a maximum, which is attained earlier in simulations at higher pressures. This behavior is consistent with the emergence of four-coordinate seeds within highly confined regions that were formerly associated with grain boundaries. Afterwards, there is a competition between two effects that drive the fractal dimension in opposite directions. The initial formation of four-coordinate seeds renders the network less spatially homogeneous, thereby reducing d_f . However, the subsequent growth and the development of more of these seeds into domains that ultimately, in most simulations, occupy a large fraction of the supercell acts to increase d_f . Consequently, in most simulations, the initial maximum is followed by a minimum and then by an increase in d_f as the four-coordinate network progressively fills the simulation cell. Figure 6(c) shows snapshots from a simulation that displays both a maximum and a minimum in the d_f time series, highlighting seed

formation followed by coarsening. At high pressures and temperatures, the time series converge toward a fractal dimension of three, whereas under less extreme conditions the time series plateau at lower values. At the lowest applied pressure of 25 GPa, the time series does not exhibit a minimum in the four-coordinate fractal dimension. Following a plateau during the stage of the simulation in which the target pressure and temperature are maintained, most simulations instead show a decrease in d_f , followed by a second plateau at a lower value. This latter behavior reflects the graphitization of loosely scattered four-coordinate atoms that did not nucleate and grow into significant domains under the relatively low applied pressure. The absence of a minimum in the time series can be understood in terms of the limited stability and growth of four-coordinate seeds at low pressure. Under these conditions, only one or two seeds typically remain stable and develop into appreciable domains, but these do not become sufficiently large to drive an increase in d_f . Rather than multiple growing domains occupying a substantial fraction of the simulation cell, one or two moderately sized domains are formed, surrounded by loosely scattered transient four-coordinate atoms, resulting in a highly nonuniform four-coordinate network. It is these dispersed four-coordinate atoms that are lost upon depressurization, thereby leaving the remaining network even less spatially homogeneous and producing a lower value of d_f . By contrast, at higher pressures this effect becomes negligible, since the system contains large four-coordinate domains and the number of atoms that graphitize during depressurization and cooling is small in comparison. As a result, the value of d_f is largely retained under these high-pressure conditions following pressurization and heating.

B. A percolation transition

Here we discuss aspects of a geometric phase transition in which the developing network of four-coordinated atoms first percolates. We define percolation as configurations with a four-coordinated cluster that spans the simulation cell along any of the three Cartesian directions. We note that this percolation transition also, in a sense, marks the breakdown of a polycrystalline simulation model with periodic boundary conditions. It indicates the point at which there exists a direction along which the developing phase no longer pays an interfacial free energy “cost.” This situation is unlikely to occur in real polycrystalline materials. The growth of the four-coordinated clusters is a highly correlated process, in which coupling between four-coordinated atomic positions arises through the interatomic potential. To assess the role of the interatomic potential in generating correlated network connectivity, we introduce a minimal null model for comparison. Configurations are generated using a simple algorithm. A cubic supercell of side length 45 Å is constructed under periodic boundary conditions, and 10 100 atoms are placed at random using a random sequential adsorption procedure [33], enforcing a minimum separation of 1.2 Å. This process is repeated 500 times for the given box size yielding 500 random configurations. The box is then uniformly rescaled by reducing its side length by 0.05 Å, and the procedure is repeated until the box reaches a side length of 33 Å at which the system was found to be well above the percolation threshold. Bonds

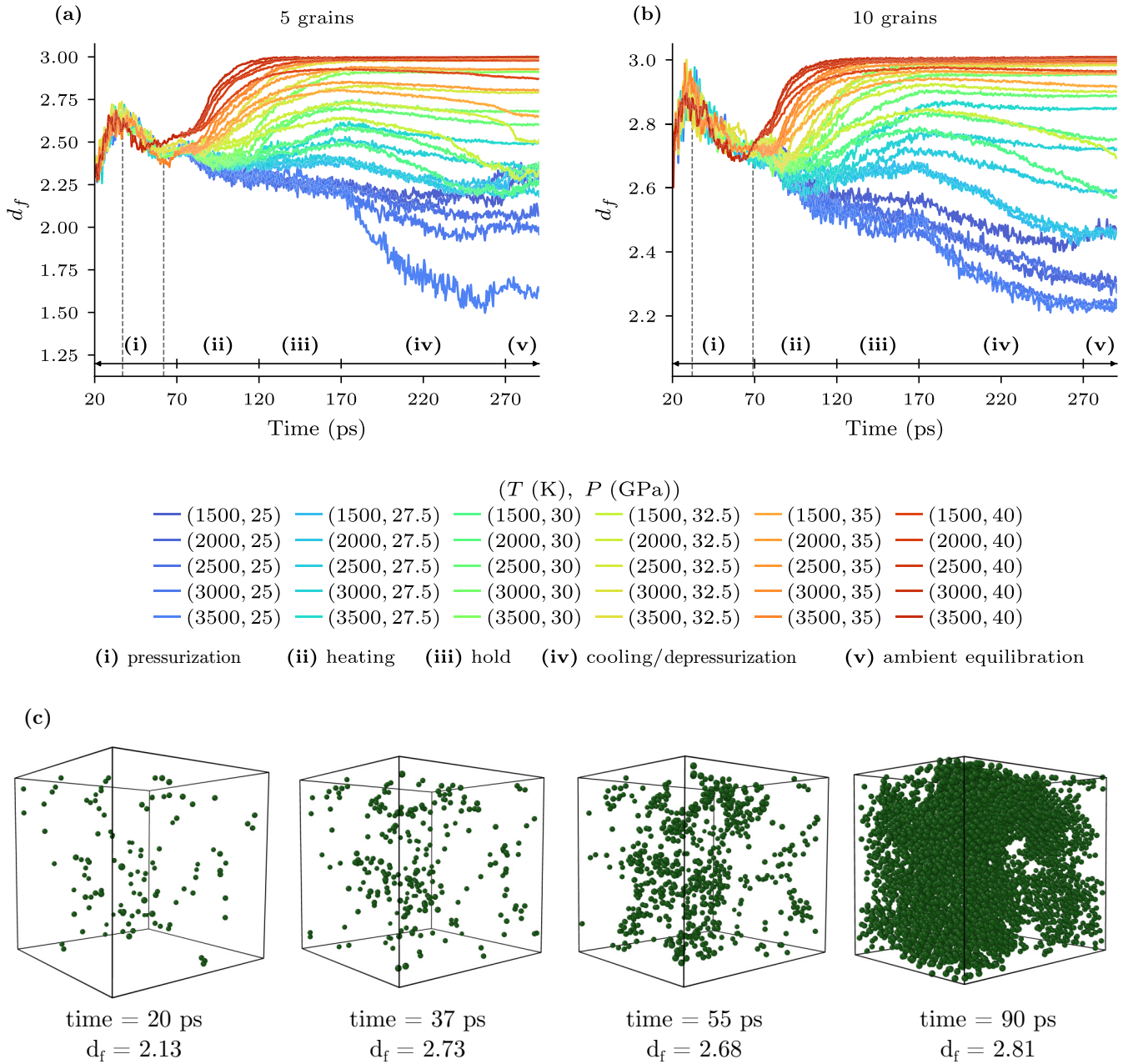


FIG. 6. Panels (a) and (b) show the fractal dimension time series for all pressure temperature combinations explored. Each time evolution is averaged over all polycrystal structures (five times five grain and five times ten-grain). The horizontal axis begins at the start of the pressurization ramp. Panel (c) shows the time evolution of a representative four-coordinated network as snapshots from an MD simulation with both a maximum and a minimum in the time series.

are defined using the same $1.7 \text{ \AA}$ criterion as above. Rather than attempting to remove finite size effects in percolation, this null model is designed to sample configurations in the absence of an interatomic potential. We note that the sampled configurations do not correspond to an equilibrium hard sphere liquid or solid (since the algorithm does not sample the Boltzmann distribution [33,34]). However, a well-defined percolation problem does not require equilibrium configurations and the finite-time MD trajectories themselves sample far from ergodic convergence [35] so they do not represent the equilibrium phases at the relevant temperature and pressure, and the results therefore depend strongly on the starting

polycrystalline configuration. One additional adjustment is required for the null model. In a simple hard sphere system there is nothing to suppress coordination numbers greater than four. At high packing fractions, higher coordinations will dominate, unlike in carbon where a well parametrized potential favors tetrahedral environments. To enable a comparison, we therefore restrict the hard sphere analysis to clusters composed of particles that are at least four coordinated, and we examine the percolation of such clusters. Note that, since this null model is an explicit hard sphere system, it has a well defined packing fraction. However, for consistency we retain the definition of η in Eq. (19). Using the packing fraction would simply

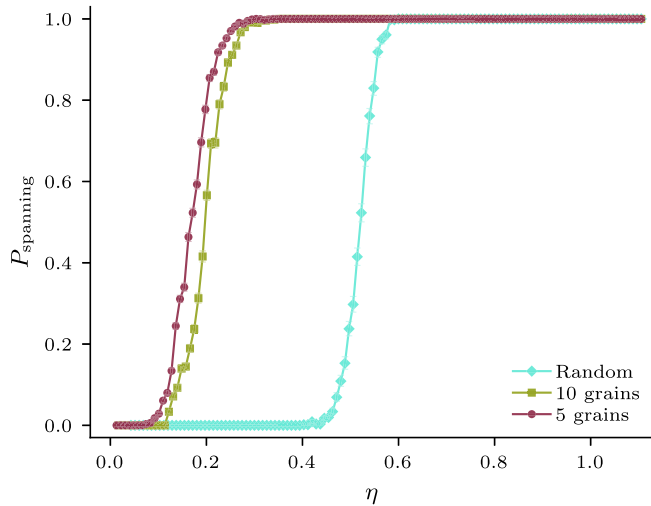


FIG. 7. The spanning probability, P_{spanning} , against η . For the five and ten grain systems, η denotes the reduced density of exactly four-coordinated atoms, whereas for the random model it denotes the reduced density of atoms that are at least four coordinated.

amount to a constant rescaling of η , which is arbitrary for our purposes. Let us now define a spanning probability, P_{spanning} , as the fraction of configurations which contain a percolating cluster relative to the total number of configurations at a given value of η . Figure 7 shows that the four-coordinated carbon network percolates at a significantly lower threshold than the random null model, providing strong evidence for correlated growth during the development of the network. In a finite-size system, as shown in Fig. 7, the transition in P_{spanning} is rounded, so there is no single sharply defined critical density, η_c . A common finite-size convention [36,37] is to define the percolation threshold as the value of η for which the spanning probability is $P_{\text{spanning}} = 0.5$, since this criterion often shows relatively weak finite-size dependence. This continuous variation of P_{spanning} should not be confused with the continuous onset of the order parameter P_∞ . The latter is defined as the fraction of sites belonging to the spanning cluster, relative to the total number of sites satisfying the required coordination number. Unlike the finite-size rounding of P_{spanning} , the continuous onset of P_∞ persists in the thermodynamic limit and reflects that the phase transition is second order. The five grain systems percolate at $\eta_c = 0.17$, the ten-grain systems at $\eta_c = 0.20$, whereas the random model percolates at $\eta_c = 0.52$. The lower threshold in the five grain systems may be attributed to the fact that fewer grains span a given dimension, so percolation requires the transformation of fewer graphitic grains into the four-coordinated phase.

Percolation theory [38] requires that, at the percolation threshold, the distribution of finite cluster sizes is scale free in the thermodynamic limit,

$$n_s(\eta_c) \sim s^{-\tau}, \quad (22)$$

where n_s is the number of clusters of size s , η_c is the percolation threshold, and τ is a critical exponent (known as the Fisher exponent) determined by the universality class. To test whether the percolating four-coordinated network exhibits this scale free regime and if it reproduces the expected

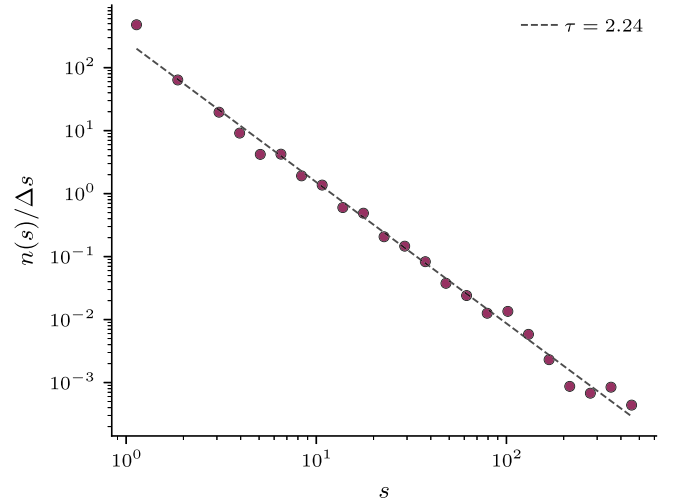


FIG. 8. Log-log plot of the cluster size distribution (number of clusters per bin, normalized by bin width) against cluster size. The dashed line shows a linear regression fit with gradient -2.24 .

exponent, we plot the cluster-size distribution on a log-log scale, as shown in Fig. 8. The cluster size distribution shows the expected approximately linear behavior on the logarithmic plot. A simple fit yields an apparent Fisher exponent of $\tau = 2.24$. Given that finite size effects are not corrected for, the number of data points is limited, and most configurations are not sampled exactly at η_c , this is in reasonably good agreement with previous reports which predict $\tau \simeq 2.19$ for three-dimensional percolation [39].

Finally, we comment on a correction, specifically, why the use of one is not needed here, to the fractal dimension evaluated near percolation. In systems such as branched polymers, where percolation occurs, extracting their fractal dimension from the structure factor (and hence from the radial distribution function) often requires accounting for the scale free cluster size distribution at the percolation transition [14,40]. The resulting extreme polydispersity means that larger clusters dominate the scattering, and the Fisher exponent enters the relation between the measured power law scattering and the underlying single cluster fractal dimension. In the present work, such a correction is not required because our quantity of interest is the fractal dimension of the entire set of four-coordinated carbon atoms, i.e., a measure of how space filling this set is within the simulation cell, rather than the fractal dimension of individual clusters. Individual clusters are typically dense domains and are therefore expected to have a fractal dimension close to three. We can therefore neglect any polydispersity related correction.

V. CONCLUSIONS

In this contribution, the phase transition of polycrystalline graphite to diamond when subjected to high-pressure, high-temperature conditions was investigated using molecular dynamics simulations. We showed that the developing network of four-coordinate carbon atoms may be characterized by its fractal dimension which we calculate from the radial distribution function “colored” for four-coordinate sites

only. The fractal dimension proved to be a useful structural metric that describes the diamondlike networks well and provides clues to the transformation pathway. Specifically, an unexpected trend in the fractal dimension at sufficiently low pressures give strong evidence for the formation of four-coordinate seeds which then tend to coarsen. However, in the low pressure regime, coarsening turned out to be dependent on the initial structure. Meanwhile, at high pressures and temperatures, all fractal dimensions converge toward three with little to no structure dependence. The growth of the four-coordinate domains eventually leads to a percolation transition which shows criticality with the expected approximate three-dimensional universality. We compared the percolation of the developing four-coordinate carbon network with a significantly simpler hard sphere random sequential adsorption model which, as expected, was found to percolate at a markedly higher threshold. Overall, we propose that the fractal dimension, and its time evolution, may be a useful

structural metric for domain growth processes such as those observed in polycrystalline materials.

ACKNOWLEDGMENTS

V.N. is grateful to Dr. Camilla Di Mino for helpful discussions regarding the structure factor. Z.M.T. acknowledges support from the EPSRC Centre for Doctoral Training in Inorganic Chemistry for Future Manufacturing (OxICFM), EP/S023828/1.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available upon publication because it is not technically feasible and/or the cost of preparing, depositing, and hosting the data would be prohibitive within the terms of this research project. The data are available from the authors upon reasonable request.

-
- [1] K. Luo, B. Liu, W. Hu, X. Dong, Y. Wang, Q. Huang, Y. Gao, L. Sun, Z. Zhao, Y. Wu, Y. Zhang, M. Ma, X.-F. Zhou, J. He, D. Yu, Z. Liu, B. Xu, and Y. Tian, Coherent interfaces govern direct transformation from graphite to diamond, *Nature (London)* **607**, 486 (2022).
 - [2] S. Srinivasan, R. Batra, D. Luo, T. Loeffler, S. Manna, H. Chan, L. Yang, W. Yang, J. Wen, P. Darancet, and S. K. R. S. Sankaranarayanan, Machine learning the metastable phase diagram of covalently bonded carbon, *Nat. Commun.* **13**, 3251 (2022).
 - [3] D. R. Robinson and M. Wilson, The liquid $\leftrightarrow$ amorphous transition and the high pressure phase diagram of carbon, *J. Phys.: Condens. Matter* **25**, 155101 (2013).
 - [4] F. Bundy, W. Bassett, M. Weathers, R. Hemley, H. Mao, and A. Goncharov, The pressure-temperature phase and transformation diagram for carbon; updated through 1994, *Carbon* **34**, 141 (1996).
 - [5] C. Hull, S. Raj, and R. Saykally, The liquid state of carbon, *Chem. Phys. Lett.* **749**, 137341 (2020).
 - [6] C. Lifshitz, Carbon clusters, *Int. J. Mass Spectrom.* **200**, 423 (2000).
 - [7] A. Barnard, Theory and modeling of nanocarbon phase stability, *Diam. Relat. Mater.* **15**, 285 (2006).
 - [8] P. Németh, K. McColl, L. A. J. Garvie, C. G. Salzmann, M. Murri, and P. F. McMillan, Complex nanostructures in diamond, *Nat. Mater.* **19**, 1126 (2020).
 - [9] Y. Ge, K. Luo, Y. Liu, G. Yang, W. Hu, B. Li, G. Gao, X.-F. Zhou, B. Xu, Z. Zhao, and Y. Tian, Superconductivity in graphite-diamond hybrid, *Mater. Today Phys.* **23**, 100630 (2022).
 - [10] Z. Malinowska-Trzmielak, V. Neuman, and M. Wilson, Grain-boundary-mediated kinetic arrest in graphite-to-diamond transformation, *arXiv:2607.08868*.
 - [11] R. Drautz, Atomic cluster expansion for accurate and transferable interatomic potentials, *Phys. Rev. B* **99**, 014104 (2019).
 - [12] M. Qamar, M. Mrovec, Y. Lysogorskiy, A. Bochkarev, and R. Drautz, Atomic cluster expansion for quantum-accurate large-scale simulations of carbon, *J. Chem. Theory Comput.* **19**, 5151 (2023).
 - [13] Z. Malinowska-Trzmielak, N. Grobert, and M. Wilson, Transferability of carbon potentials for novel carbon polymorphs, *Modelling Simul. Mater. Sci. Eng.* **34**, 025004 (2026).
 - [14] J. Teixeira, Small-angle scattering by fractal systems, *J. Appl. Crystallogr.* **21**, 781 (1988).
 - [15] T. Freltoft, J. K. Kjems, and S. K. Sinha, Power-law correlations and finite-size effects in silica particle aggregates studied by small-angle neutron scattering, *Phys. Rev. B* **33**, 269 (1986).
 - [16] D. L. Gurina, Y. A. Budkov, and M. G. Kiselev, Molecular aspects on impregnation of silica-based aerogels with mefenamic acid in supercritical carbon dioxide, *J. Mol. Liq.* **415**, 126358 (2024).
 - [17] J. Kieffer and C. Angell, Generation of fractal structures by negative pressure rupturing of SiO₂ glass, *J. Non-Cryst. Solids* **106**, 336 (1988).
 - [18] S. Sinha, T. Freltoft, and J. Kjems, Observation of power-law correlations in silica-particle aggregates by small-angle neutron scattering, in *Proceedings of the International Topical Conference on Kinetics of Aggregation and Gelation, April 2–4, 1984, Athens, GA, USA* (Taylor & Francis, London, 1984), pp. 87–90.
 - [19] Edited by J. Binney, *The Theory of Critical Phenomena: An Introduction to the Renormalization Group*, Oxford Science Publications (Clarendon Press, Oxford University Press, Oxford: Oxford, New York, 1992).
 - [20] W. Brostow, J.-P. Dussault, and B. L. Fox, Construction of Voronoi polyhedra, *J. Comput. Phys.* **29**, 81 (1978).
 - [21] J. L. Finney, A procedure for the construction of Voronoi polyhedra, *J. Comput. Phys.* **32**, 137 (1979).
 - [22] P. Hirel, AtomsK: A tool for manipulating and converting atomic data files, *Comput. Phys. Commun.* **197**, 212 (2015).
 - [23] I. Obayashi, T. Nakamura, and Y. Hiraoka, Persistent homology analysis for materials research and persistent homology software: HomCloud, *J. Phys. Soc. Jpn.* **91**, 091013 (2022).

- [24] A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in 't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, and S. J. Plimpton, LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales, *Comput. Phys. Commun.* **271**, 108171 (2022).
- [25] S. Nosé, A unified formulation of the constant temperature molecular dynamics methods, *J. Chem. Phys.* **81**, 511 (1984).
- [26] W. G. Hoover, Canonical dynamics: Equilibrium phase-space distributions, *Phys. Rev. A* **31**, 1695 (1985).
- [27] S. Nosé and M. Klein, Constant pressure molecular dynamics for molecular systems, *Mol. Phys.* **50**, 1055 (1983).
- [28] W. G. Hoover, Constant-pressure equations of motion, *Phys. Rev. A* **34**, 2499 (1986).
- [29] A. Stukowski, Visualization and analysis of atomistic simulation data with OVITO—the Open Visualization Tool, *Modelling Simul. Mater. Sci. Eng.* **18**, 015012 (2010).
- [30] E. Maras, O. Trushin, A. Stukowski, T. Ala-Nissila, and H. Jónsson, Global transition path search for dislocation formation in Ge on Si(001), *Comput. Phys. Commun.* **205**, 13 (2016).
- [31] S.-c. Zhu, X.-z. Yan, J. Liu, A. R. Oganov, and Q. Zhu, A revisited mechanism of the graphite-to-diamond transition at high temperature, *Matter* **3**, 864 (2020).
- [32] P. H. Gaskell, A. Saeed, P. Chieux, and D. R. McKenzie, Neutron-scattering studies of the structure of highly tetrahedral amorphous diamondlike carbon, *Phys. Rev. Lett.* **67**, 1286 (1991).
- [33] J. Talbot, G. Tarjus, P. Van Tassel, and P. Viot, From car parking to protein adsorption: An overview of sequential adsorption processes, *Colloids Surf. A* **165**, 287 (2000).
- [34] B. Widom, Random sequential addition of hard spheres to a volume, *J. Chem. Phys.* **44**, 3888 (1966).
- [35] V. Neuman, P. A. Wesołowski, K. K. Bojarski, D. Dewan, M. Schäffler, P. Smardz, and D. J. Wales, Visualizing the energy landscape for a molecular dynamics trajectory, *J. Chem. Phys.* **164**, 044112 (2026).
- [36] T. Harter, Finite-size scaling analysis of percolation in three-dimensional correlated binary Markov chain random fields, *Phys. Rev. E* **72**, 026120 (2005).
- [37] D. M. Heyes and J. R. Melrose, Percolation cluster statistics of Lennard-Jones fluids, *Mol. Phys.* **66**, 1057 (1989).
- [38] D. Stauffer and A. Aharony, *Introduction to Percolation Theory*, 2nd ed. (Taylor & Francis, London, 1994).
- [39] C. D. Lorenz and R. M. Ziff, Precise determination of the bond percolation thresholds and finite-size scaling corrections for the sc, fcc, and bcc lattices, *Phys. Rev. E* **57**, 230 (1998).
- [40] D. Stauffer, Gelation in concentrated critically branched polymer solutions. Percolation scaling theory of intramolecular bond cycles, *J. Chem. Soc. Faraday Trans. 2* **72**, 1354 (1976).