

RESEARCH ARTICLE | JANUARY 21 2026

# Multiscale structural analysis of defective graphene in transmission electron microscopy images using persistent homology

Ryuto Eguchi  ; Ayako Hashimoto 

APL Mater. 14, 011108 (2026)

<https://doi.org/10.1063/5.0305461>

## Articles You May Be Interested In

Abelian link invariants and homology

*J. Math. Phys.* (June 2010)

Topological trends in ionic transport through metal-oxide composites

*Appl. Phys. Lett.* (February 2021)

Zero-dipole Schottky contact: Homologous metal contact to 2D semiconductor



Trusted in Research  
for over 40 years

## Materials & Thin Films: Composition, Interfaces and Depth Profiling

Surface analysis and plasma/process diagnostics for materials R&D

Find Solutions for Your Research

# Multiscale structural analysis of defective graphene in transmission electron microscopy images using persistent homology

Cite as: APL Mater. 14, 011108 (2026); doi: 10.1063/5.0305461

Submitted: 6 October 2025 • Accepted: 15 December 2025 •

Published Online: 21 January 2026



Ryuto Eguchi<sup>1,2,a)</sup>  and Ayako Hashimoto<sup>1,2,a)</sup> 

## AFFILIATIONS

<sup>1</sup> National Institute for Materials Science, 1-2-1 Sengen, Tsukuba, Ibaraki 305-0047, Japan

<sup>2</sup> Subprogram in Materials Science and Engineering, Degree Programs in Pure and Applied Sciences, Graduate School of Science and Technology, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8577, Japan

<sup>a)</sup> Authors to whom correspondence should be addressed: [eguchi.Ryuto.tkb\\_kb@u.tsukuba.ac.jp](mailto:eguchi.Ryuto.tkb_kb@u.tsukuba.ac.jp) and [HASHIMOTO.Ayako@nims.go.jp](mailto:HASHIMOTO.Ayako@nims.go.jp)

## ABSTRACT

We developed a persistent-homology-based strategy that converts transmission electron microscopy images of defective graphene into quantitative, multiscale descriptors spanning local carbon-atom polygons and their global connectivity. Zero-dimensional persistence diagrams for grayscale images extracted carbon-ring interiors as local structures, allowing identification of defect structures. Roundness as a measure of the distortion of carbon rings and polygon number were quantitatively calculated. It was found that electron irradiation increased both the fraction of non-hexagonal polygons and distortion of carbon-atom polygons in graphene. One-dimensional persistence diagrams for polygon-center point clouds after downsampling the center of non-hexagonal polygons yielded large rings that reliably enclosed connected defect clusters ranging from simple defects to complex clusters. These rings also enclosed non-crystalline regions composed of hexagons, which were difficult to capture considering only polygon number. The large rings enable automated counting of the number of each type of polygon within each defect cluster. The number of polygons within defect clusters thereby quantified the connectivity of carbon-atom polygons and revealed that electron-beam irradiation induced the growth of defect clusters in an interconnected manner.

© 2026 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/5.0305461>

## I. INTRODUCTION

A fundamental understanding of disordered structure-property relationships is essential for studies on glassy and amorphous materials.<sup>1,2</sup> Although complete determination of the three-dimensional (3D) atomic arrangements in amorphous materials is still challenging,<sup>3</sup> as for two-dimensional (2D) materials, transmission electron microscopy (TEM) enables direct visualization of 2D atomic arrangement.<sup>4</sup> Among 2D materials, graphene has been extensively studied because of its exceptional properties, including high carrier mobility<sup>5,6</sup> and high thermal conductivity.<sup>7,8</sup> Ideally, graphene is a perfect 2D crystal of tessellation by six-membered carbon rings; however, lattice defects are ubiquitous. Lattice defects alter the electronic transport and magnetic properties of materials.<sup>9–11</sup> Consequently, for defect characterization, atomic-scale observation has often been performed using high-resolution transmission electron microscopy (HR-TEM)

and scanning transmission electron microscopy.<sup>12,13</sup> In particular, HR-TEM has been widely used to probe dynamic behaviors of defects because it provides not only high spatial but also high temporal resolution.<sup>14</sup>

Disordered graphene including lattice defects can be realized by electron-beam irradiation using an electron microscope. In the 2000s, individual point defects were first identified in irradiated graphene and related materials using microscopy.<sup>15–17</sup> These primitive defects were manually labeled by the size and arrangement of the carbon-atom rings (polygons) involved, e.g., 55–77, 5–9, and 5–8–5. Subsequently, Eder *et al.*<sup>18</sup> produced amorphous 2D carbon through high-dose electron-beam irradiation and then developed an automated pipeline that extracted atomic positions from HR-TEM images and derived ring-size distributions. Graphene is often characterized by intricately interconnected non-hexagonal rings (mostly  $\leq$ nine-membered) extending over large lateral regions,

unlike the other well-known 2D materials, such as hexagonal boron nitride, which has large prism-shaped pores,<sup>19</sup> and transition-metal dichalcogenides with  $\geq$ ten-membered rings that assemble into even larger rings.<sup>20,21</sup> Quantitative characterization of nearly amorphous 2D carbon has been largely confined to ring-size distributions, although disorder metrics and detection methods of isolated mesoscopic defect clusters have been reported for crystals with low defect density.<sup>22,23</sup> In addition to the numerical distribution of ring sizes, the spatial distribution of the carbon rings is another degree of freedom that needs to be quantified. For studies on property measurement and predictive modeling in disordered structures, multiscale descriptors are required, including the nanometer-scale connectivity between the carbon-ring polygons. For example, Lahiri and colleagues showed that an extended line defect composed of alternating five- and eight-membered rings exhibited enhanced local carrier density and metallic properties.<sup>24</sup>

Persistent homology (PH) is a multiscale analysis method that is widely used in materials science.<sup>25–28</sup> PH represents the shape of data through  $k$ -dimensional holes ( $N = 0$ : connected components;  $N = 1$ : loops/tunnels). In PH, the scales at which  $k$ -dimensional holes appear and disappear under continuous expansion are tracked by a filtration process leading to the PH's multiscale nature.<sup>29</sup> Various filtration processes allow PH to be used with different datasets, including point clouds, binary images, and grayscale images.<sup>30–33</sup> In the era of rapidly advancing machine learning, PH is an unsupervised learning approach that requires no labeled data and operates effectively on small datasets. Moreover, it is characterized by minimal hyperparameters and low computational cost, distinguishing it from convolutional neural network (CNN)-based encoder-decoder semantic segmentation architectures and from point-based deep neural network architectures for 3D point clouds. Most studies using PH have been applied to atomic arrangements, i.e., point clouds, obtained by molecular dynamics (MD) simulations to find the hidden medium-range order, which is represented by the minimum unit such as the shortest-path ring defined by Si–O,<sup>34,35</sup> C–C,<sup>36,37</sup> and Si–Si<sup>38</sup> bonds. When the minimum unit includes different elements, higher-order holes generated by the assembled units can be automatically revealed by selecting a target element.<sup>39</sup> However, this strategy is unsuitable for single-component materials and multiple components with low elemental contrast. Consequently, the connectivity of minimum units has remained largely unexplored.

In this study, we aimed to automate the extraction of multiscale structural information from 2D disordered graphene, including the local structure of carbon-atom polygons (mostly  $\leq$ nine-membered rings) as the minimum unit, and the global network information on larger scales than the primitive polygons. The developed PH strategy is used to extract and quantitatively evaluate local defects in graphene. The global network analysis allows us to extract long-range information beyond the scale of the minimum structural units by using the developed pipeline based on point cloud analysis.

## II. EXPERIMENTAL PROCEDURE AND ANALYTICAL METHOD

A graphene specimen was prepared by transferring commercial graphene (Trivial Transfer Graphene, ACS Materials, USA) onto a molybdenum-supported carbon film (Quantifoil, Quantifoil Micro Tools GmbH, Germany). The sample was heated in a

microscope column to clean the graphene surface, followed by observation with an aberration-corrected transmission electron microscope (JEM-ARM200F, JEOL, Japan). Details of the cleaning process and observation conditions are provided in [supplementary material S1](#).

The as-prepared graphene sample was irradiated using a 200-keV electron beam to introduce lattice defects. Initial irradiation was conducted for 20 s at an electron dose rate of  $1.5 \times 10^5 \text{ e}^- \text{ nm}^{-2} \text{ s}^{-1}$ . To avoid irradiation damage during observation, the accelerating voltage was decreased to 80 keV and HR-TEM images were acquired for 2 s. The accelerating voltage was raised again to 200 keV to conduct irradiation. Defect evolution was recorded after the voltage was lowered to 80 keV.

The first step of PH is recording the scales at which  $k$ -dimensional holes appear (birth) and disappear (death) along filtration. A persistence diagram (PD) is a 2D histogram of the birth and death pairs (b–d pairs). The procedures for PD calculation and inverse analysis of PDs were conducted using the data analysis software “HomCloud.”<sup>40</sup> The obtained data were processed by different filtration schemes. For grayscale images, superlevel filtration<sup>41</sup> was used, where the evolution of topological features is defined by pixel intensity thresholds as shown in [supplementary material S2](#). The signed Euclidean distance transform was applied to binarized images. For point clouds, alpha-filtration was used.<sup>42</sup> PDs are denoted as g-PDs for grayscale images, b-PDs for binarized images, and p-PDs for point clouds. In PDs, a lifetime, defined as the difference between birth and death times, plays an essential role, and therefore, the details are provided in [supplementary material S2](#). [Figure 1](#) presents an overview of the PH strategy for local structure extraction and global network analysis, the details of which are described in Secs. III B–III D, respectively.

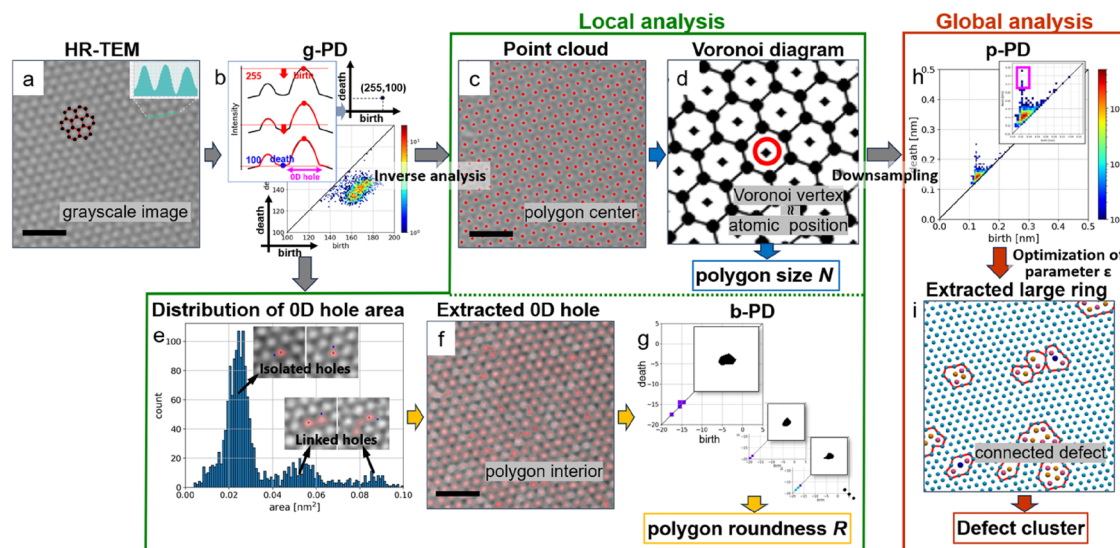
## III. RESULTS AND DISCUSSION

### A. HR-TEM images of defective graphene after electron-beam irradiation

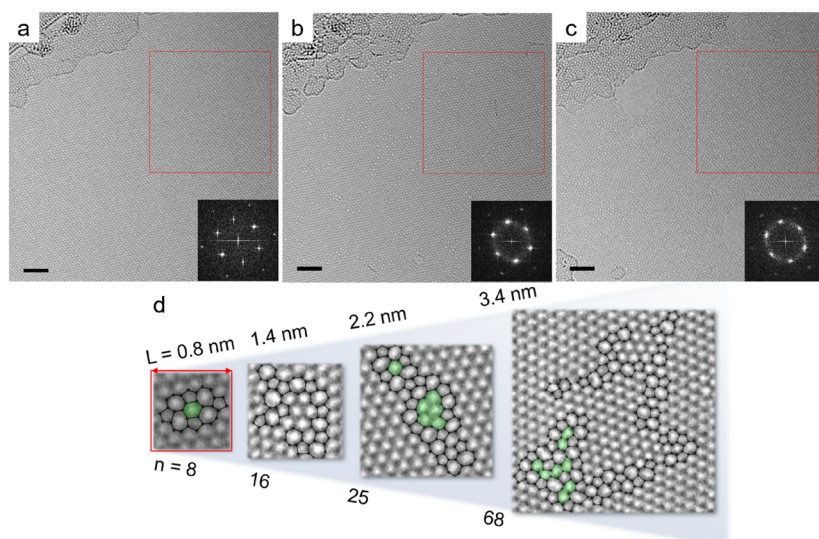
[Figures 2\(a\)–2\(c\)](#) show the time evolution of the graphene lattice under 200-keV electron-beam irradiation at  $t_0$  (0 s),  $t_1$  (20 s), and  $t_2$  (40 s), respectively. [Figure 2\(a\)](#) shows the hexagonal lattice of the as-prepared single-layer graphene specimen before irradiation at 200 keV. After electron irradiation, pores and defects including non-hexagonal polygons were observed, as illustrated in [Figs. 2\(b\) and 2\(c\)](#). Moreover, as presented in the enlarged view in [Fig. 2\(d\)](#), many defective regions did not consist of typical isolated defects, i.e., 55–77, 5–9, 5–8–5, 555–777, and 5555–6–7777,<sup>15</sup> but aggregated non-hexagonal polygons that formed complex clusters (further details are provided later). Although the corresponding FFT patterns exhibited sharp sixfold spots before irradiation [inset of [Fig. 2\(a\)](#)], they broadened markedly after electron-beam irradiation [insets of [Figs. 2\(b\) and 2\(c\)](#)]. The similar FFT patterns after irradiation make it difficult to evaluate structural differences from the FFT patterns even though they reflect the evolution of defect structures. Therefore, we analyzed these HR-TEM images using g-PD, which extracts hole information derived from local intensity contrasts.

### B. Extraction of interior regions and maximum-intensity pixels of carbon-atom polygons

[Figure 1](#) presents the PH-based strategy used to extract local structures in defective graphene from an HR-TEM image. In this



**FIG. 1.** Overview of the PH strategy for local structure extraction and global network analysis. (a) Gaussian-filtered HR-TEM image and intensity profile across three hexagons (inset). The filtering details are provided in [supplementary material S3](#). (b) Construction of a 0D g-PD (right) by superlevel filtration (left) of the filtered image in (a). (c) Birth point distribution (red points) obtained by inverse analysis of the 0D g-PD in (b) for b–d pairs with lifetime > 5. (d) Voronoi diagram calculated from birth point distribution in (c). (e) Area distribution of the extracted 0D holes. In the filtered images (inset), the red regions indicate 0D holes and the red and blue points represent birth and death points, respectively. (f) Overlay of the extracted 0D holes on the filtered image, including both isolated holes and holes separated from linked holes through the watershed algorithm. (g) Binarized images of the 0D holes extracted from (e) and their zero-dimensional b-PD. The scale bars in (a), (c), and (e) indicate 1 nm. (h) One-dimensional (1D) p-PD of the downsampled point cloud in (c). (i) Large ring obtained by the inverse analysis of the 1D p-PD in (h).



**FIG. 2.** HR-TEM images and corresponding fast Fourier transform (FFT) patterns (inset) of single-layer graphene after electron-beam irradiation at 200 keV for (a) 0 s, (b) 20 s, and (c) 40 s. The regions indicated by red squares without any pores were selected for FFT patterns and further analysis. The scale bars represent 2 nm. (d) TEM images of various typical defect clusters obtained from different regions in (a)–(c). The defect clusters have different patch sizes ( $L$ , indicated by red lines) and numbers of non-hexagonal polygons ( $n$ , with atomic models). The green regions consist of hexagons that do not form a crystalline domain; they are defective regions.

study, we treat a single polygon (carbon ring) and its surrounding region as one unit, as illustrated in [Fig. 1\(a\)](#). It should be noted that here we applied g-PD to grayscale experimental TEM images to analyze material structures, whereas PH applications usually employ p-PD and b-PD. The intensity profile across the carbon-atom polygon hole, i.e., vacuum region (inset), has its maximum close to the

polygon center. Accordingly, as shown on the left of [Fig. 1\(b\)](#), the vacuum region is treated as a domain, namely, a zero-dimensional (0D) hole, under the 0D superlevel filtration. The corresponding g-PD for the defective graphene in [Fig. 1\(a\)](#) is shown on the right of [Fig. 1\(b\)](#). The b–d pairs with short lifetimes correspond to holes with low contrast, which mostly arise from noise.<sup>43</sup> The average



lifetime of large vacuum regions without a graphene layer was 2.44, mainly originating from background noise, and is extremely small compared with the intensity range of 0–255. Here, b–d pairs with lifetimes exceeding 5, sufficiently larger than the background lifetime in the vacuum region, were considered as significant holes corresponding to interior regions of carbon-atom polygons. The inverse analysis results for the b–d pairs with lifetimes >5 are shown in Fig. 1(c). The birth points marked by red circles indicate the pixel of maximum intensity for each hole and closely coincide with the center of the carbon-atom polygons. As shown in Fig. 1(d), using the obtained birth points, we calculated carbon-atom positions via a Voronoi diagram (supplementary material S4). The validity of the atomic model was confirmed by comparing corresponding simulated and experimental HR-TEM images (supplementary material S4). Figure 1(e) shows the area distribution of the extracted holes, which contained distinct peaks. The first peak was attributable to one carbon-atom polygon, and the second and subsequent peaks were ascribed to linked carbon-atom polygons (insets). Then, we applied watershed segmentation to the linked holes formed by connecting carbon-atom polygons and extracted the one isolated domain containing the birth point (see supplementary material S5). Finally, we extracted one isolated hole per birth point and mapped all holes, including those separated from linked clusters, onto the raw TEM image, as shown in Fig. 1(f). Thus, we extracted regions enclosed by a single carbon-atom polygon using the g-PD. Although similar extraction has been often attempted via binary segmentation (e.g., adaptive thresholding), our PH-based approach is less susceptible to point-like noise but can still separate local high-intensity regions (see supplementary material S6). Subsequently, as shown in Fig. 1(g), the 0D holes in Fig. 1(f) were binarized and then their b-PDs were calculated. We calculated the roundness ( $R$ ), which quantifies the deviation of a hole from a perfect circle (supplementary material S7; cf. Fig. 3), from the b-PD. The point cloud of the obtained birth points was used to calculate the polygon number  $N$  (Fig. 3) and analyze the global network (Fig. 4).

### C. Local analysis of roundness and polygon number

To analyze the carbon-atom polygons following the process in Fig. 1, we selected regions free of adsorbates and pores [indicated by red squares in Figs. 2(a)–2(c)] in the HR-TEM images. Figures 3(a)–3(c) show enlarged HR-TEM images of defective graphene after electron irradiation at  $t_0$ ,  $t_1$ , and  $t_2$ , respectively, with color maps of  $R$ .

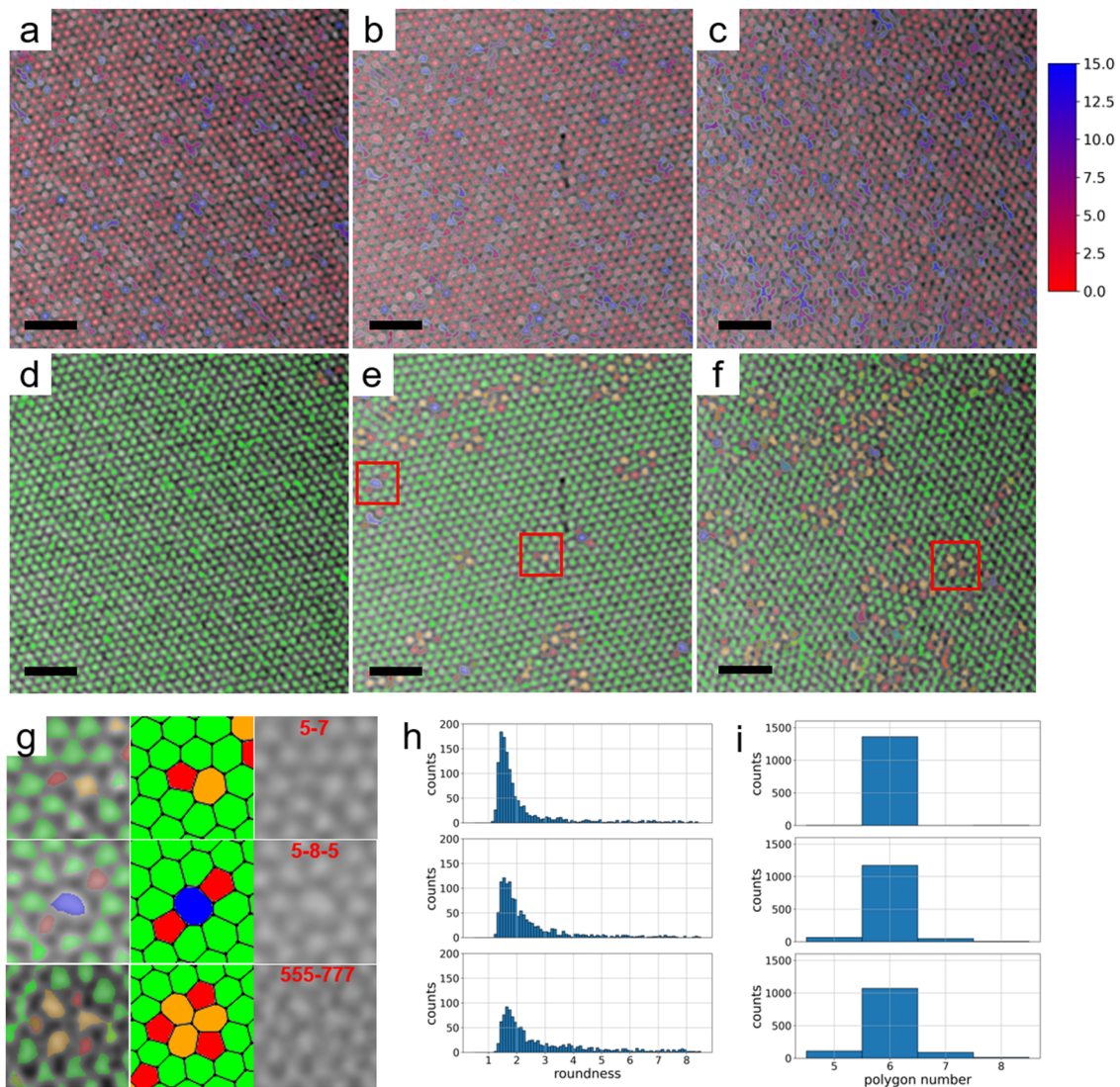
The colored areas represent the extracted 0D holes shown in Fig. 1(f), i.e., the interiors of the polygons.  $R$  is defined as  $R = A/(\pi b_{\text{ave}}^2)$  (see supplementary material S7), where  $b_{\text{ave}}$  is the average birth in the b-PD.  $R$  is one for a circle and increases as hole distortion increases. The blue regions (large  $R$ ) correspond to more distorted holes than red regions (small  $R$ ). In Figs. 3(d) and 3(e), the  $N$ -polygon holes in the same defective graphene are color-coded by  $N$ . Here,  $N$  was measured from the birth point distribution [Fig. 1(c)] without threshold tuning, which is different from applying a radius threshold to count nearest neighbors (supplementary material S8). As illustrated in Fig. 3(g) for the magnified views indicated by the red squares in Figs. 3(e) and 3(f), typical defects (e.g., 5–7 and 5–8–5) were readily identified using the calculated  $N$ .

Thus, PH-assisted analysis can automate the detection of defects and quantitative evaluation of their type and shape.

In addition to the visualization of the graphene defects in Figs. 3(a)–3(f), assessing electron-irradiation effects requires a quantitative analysis of defect shape. Figures 3(h) and 3(i) show the distributions of  $R$  and  $N$ , respectively, at  $t_0$ ,  $t_1$ , and  $t_2$ . For  $t_0$ ,  $t_1$ , and  $t_2$ , the  $R$  distributions range from 1 to more than 7 and show a peak at  $\sim 1.6$ . Note that these  $R$  values are much higher than 1.1 for an ideal regular hexagon. This is because the holes with a spur-like protrusion and residually linked holes markedly increase  $R$ , as discussed in supplementary material S9. In addition, experimental HR-TEM images are influenced by not only intrinsic structural changes in graphene but also extrinsic observation parameters (e.g., aberrations and focus shift) (see supplementary material S9). Therefore, the observation and further analysis were conducted under the same conditions and with the same algorithm across  $t_0$ – $t_2$ , to minimize external effects. With increasing irradiation time, the fraction of counts to the right of the peak increased; e.g., the proportion with  $R > 2$  was 34% at  $t_0$ , 48% at  $t_1$ , and 63% at  $t_2$ . The  $N$  distributions demonstrate that electron irradiation increased the number of defects: the non-hexagonal fraction rose from 0% at  $t_0$  to 9.5% at  $t_1$  to 16.4% at  $t_2$  [Fig. 3(i)]. That is, electron irradiation induced the structural change from hexagon to non-hexagon, leading to polygon distortion.

The use of both  $R$  and  $N$  values enables selective analysis of holes arising from hexagons vs defects. The average  $R$  value and its standard deviation for each  $N$ -polygon are summarized in supplementary material S10. The average  $R$  value for hexagons at  $t_0$  (defect-free) was 3.49, whereas the  $R$  values only from non-hexagonal polygons after irradiation were 4.49 at  $t_1$  and 4.98 at  $t_2$ . This increase in  $R$  suggests that electron irradiation introduces defects with more distorted polygons than hexagons in the graphene specimen. Furthermore, the standard deviation increased with  $N$ , indicating that heptagons are more susceptible to geometric distortion than pentagons. Thus, we can simultaneously investigate the spatial and numerical distributions of geometric features of interest by exploiting the quantitative nature of PH together with inverse analysis.

So far, we have analyzed the local features of graphene defects; that is, the 0D holes formed by nearest-neighbor carbon atoms. However, emergent properties depend on the connectivity of various polygons.<sup>44–46</sup> As illustrated in Fig. 2(d), for various typical defect clusters, cluster-enclosing patch sizes range from 0.8 to 3.4 nm, and the number of non-hexagonal polygons per patch reaches 67. When the cluster sizes are widely connected and lack characteristic scale, conventional patch-based CNN (e.g., U-Net and SegNet) are ill-suited for their analysis.<sup>47,48</sup> Even multiscale designs (e.g., DeepLabv3+ and FPN) typically require substantial architectural tuning and sizable annotated datasets.<sup>49,50</sup> By contrast, PH is inherently multiscale and does not require labeled data. This lightweight unsupervised strategy is well suited to cases such as this study, where no characteristic object size exists and the amount of experimental data is limited. Moreover, such clusters sometimes contain both isolated hexagons and aggregates of hexagons that do not form crystalline domains [green regions in Fig. 2(d)]. This, in turn, implies that the  $N$  value alone is insufficient to identify defective regions because hexagonal areas within the clusters could still



**FIG. 3.** Analysis of the 0D holes in single-layer graphene extracted by the process outlined in Fig. 2. Roundness maps overlaid on HR-TEM images acquired at (a)  $t_0$ , (b)  $t_1$ , and (c)  $t_2$ . The color bar shows the  $R$  value for the extracted 0D hole.  $N$ -polygon maps overlaid on HR-TEM images acquired at (d)  $t_0$ , (e)  $t_1$ , and (f)  $t_2$ . The scale bars represent 1.5 nm. The holes of the five-, six-, seven-, and eight-membered polygons are red, green, orange, and blue, respectively. (g) Typical defects found in the red rectangles in (e) and (f). Left: magnified images with color mapping. Middle: corresponding atomic structures. Right: filtered TEM images. Frequency distributions of (h) roundness and (i)  $N$ . Top:  $t_0$ , middle:  $t_1$ , and bottom:  $t_2$ .

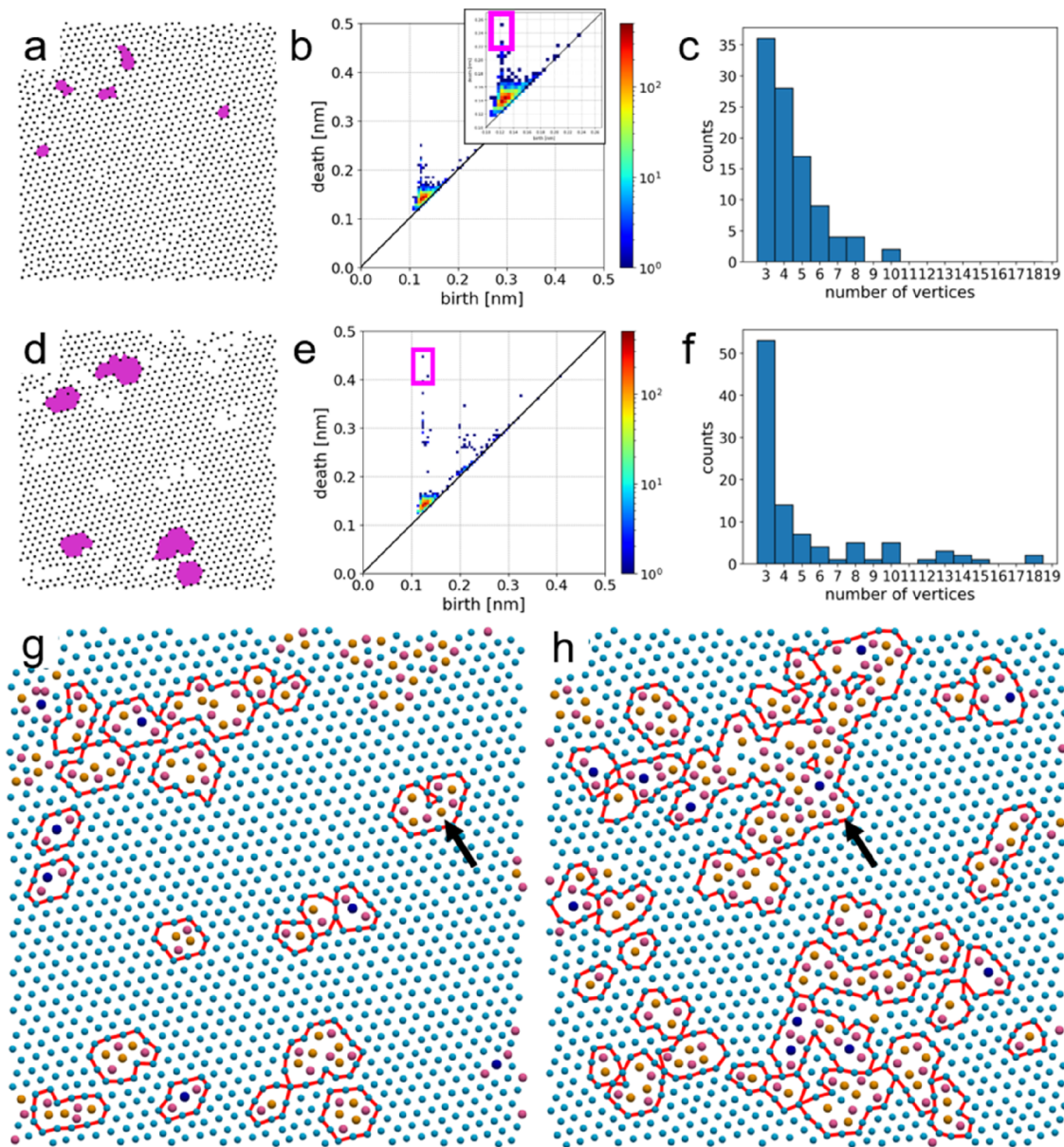
contain defects. Accordingly, we analyzed the structure on global scales—i.e., length scales exceeding the minimum unit defined by the shortest-path C–C ring—to quantify the spatial distribution of these units.

#### D. Global analysis of defect clusters

Figure 4(a) shows the birth point distribution (hereafter denoted as point cloud) corresponding to the centers of carbon-atom polygons obtained in Fig. 1(c), and Fig. 4(b) shows its one-dimensional (1D) p-PD. In Fig. 4(d), we visualize only centers

from hexagons by removing the point cloud associated with non-hexagonal polygons from Fig. 4(a). After this removal (hereafter called “downsampling”), we calculated the 1D p-PD of the downsampled point cloud. These operations—downsampling and calculation of 1D p-PD—correspond to the step illustrated in Fig. 1(h) in the overview of PH-based strategy. The b–d pairs in Fig. 4(e) are more dispersed than those in Fig. 4(b). Such b–d pairs with longer lifetimes are known to indicate larger rings in p-PD.<sup>51</sup> We conducted inverse analysis of the five b–d pairs with the longest lifetimes for both p-PDs [Figs. 4(b) and 4(e)] and visualize corresponding rings by color in Figs. 4(a) and 4(d). To obtain stable rings robust to





**FIG. 4.** (a) Birth point distribution (point cloud) and (b) its 1D p-PD of all polygon positions in the defective graphene specimen at  $t_1$ . The five purple regions in (a) were obtained from the colored rectangles with a long lifetime in the p-PD [(b), inset] by inverse analysis. (c) Frequency distribution of the number of vertices for 100 rings with a long lifetime in (b). (d) Point cloud and (e) corresponding 1D p-PD of only hexagonal polygon positions obtained by removing the non-hexagonal polygon positions from (a). (f) Frequency distribution of the number of vertices of 100 rings with long lifetimes in (e). Ring structures (red lines) surrounding defective regions with non-hexagonal polygons obtained at (g)  $t_1$  and (h)  $t_2$ . The light blue points correspond to hexagon centers, and the other colored points represent non-hexagonal polygons (red: pentagons, orange: heptagons, and deep blue: octagons) that were removed during computations.

minute perturbations in point positions, we used stable-volume calculation in the inverse analysis.<sup>52</sup> Stable-volume calculation requires a noise bandwidth parameter ( $\epsilon$ ). Here,  $\epsilon$  was fixed at 0.02 for all b–d pairs to reveal ring-size differences via downsampling. Inverse mapping of the rings revealed that the p-PD in Fig. 4(e) contained more large-ring features than that in Fig. 4(b). Furthermore, Figs. 4(c)

and 4(f) display the calculated ring-size (vertex-count) distributions for the 100 longest-lifetime b–d pairs for the p-PDs in Figs. 4(b) and 4(e), respectively. The p-PD in Fig. 4(e) contains more large-ring features than that in Fig. 4(b). Thus, downsampling of the point cloud is an effective strategy to extract long-range geometric information from the p-PD. Here, although we removed the points derived

from non-hexagonal polygons to focus on defect connectivity, the point cloud from pentagons and octagons should be removed when focusing on line defects composed of pentagons and octagons.

Figures 4(g) and 4(h) present inverse analysis results of the b–d pairs in the p-PDs only from hexagons at  $t_1$  and  $t_2$ . As described in [supplementary material S11](#), the b–d pairs used for this inverse analysis were the death and lifetime that each exceeded the respective maxima recorded for pairs with a vertex of count of 3 ( $N = 3$ ) in the p-PD at  $t_0$ . Here,  $\epsilon$  was automatically determined for each b–d pair as described in [supplementary material S11](#). As a result, we obtained the large rings that are outlined by red lines in Figs. 4(g) and 4(h). This extraction of large rings is the final step depicted in Fig. 1(i) in the PH-based strategy. The light blue points correspond to hexagon centers used in the analysis, whereas the other points represent non-hexagonal polygons that were removed during computations but are overlaid in Figs. 4(g) and 4(h) for clarity. The red-outlined rings enclose a wide variety of connected defects, ranging from typical defects (5–7 and 5–8–5) to complex clusters including more than twenty non-hexagonal polygons. Note that some defect points at the image periphery are not enclosed by a ring, primarily because they percolate to the edge of the imaged area. It was possible to recover rings enclosing those points by expanding the image area (see [supplementary material S12](#)). The red-outlined rings achieve an appropriate balance of ring granularity—neither excessively complex nor overly simple—to accurately represent the defect-cluster shape, as illustrated by the regions indicated by black arrows in Figs. 4(g) and 4(h) (see [supplementary material S13](#) for details). For example, the ring indicated by an arrow in Fig. 4(g) might otherwise fragment into several small rings, yet it accurately traces the shape of the nearly closed hole (see [supplementary material S10](#)). In addition, the large ring indicated by an arrow in Fig. 4(h) is not overly complex even though it contains many hexagonal polygons. As a result, the large ring allows hexagons to be considered defective regions. This is important because such hexagonal points within the ring do not form a crystalline domain and should, therefore, be classified as defective.

In a quantitative analysis of global connectivity, the average number of defects inside each ring should be an appropriate indicator of defect connectivity. Here, a larger number of defects represent higher defect connection. In Figs. 4(g) and 4(h), the average number of defects inside each ring was 4.6 at  $t_0$  and 6.5 at  $t_1$ . Thus, we could quantitatively evaluate the defect clusters, which clearly revealed that the defects generated via electron-beam irradiation grow in an interconnected manner. Our method can extract global structural information at scales beyond the shortest-path rings of carbon atoms. In our graphene sample, downsampling of polygon centers was straightforward because each minimum unit consists of convex polygons no larger than an octagon, while, when the minimum unit is extremely large, complex, or non-convex, defining a robust unit center would be problematic, making downsampling difficult. When the minimum unit is simple and roughly convex, PD analysis with downsampling tailored to the minimum unit is, in principle, applicable in three dimensions, so it could be suitable for automated long-range structural analysis. Furthermore, it might also be possible to design structures exhibiting long-range features by including a machine-learning loss function for point-cloud optimization. Therefore, this analytical framework is expected to be useful for Monte

Carlo-based approaches for the generation and modeling of atomic structures in amorphous materials.<sup>53–55</sup>

#### IV. CONCLUSION

A PH-based framework for extracting both local and global structural information for defective graphene from experimental HR-TEM images was developed. Local descriptors obtained from g-PD and b-PD, i.e.,  $R$  and  $N$  values, indicated that electron irradiation increased both the fraction of non-hexagonal polygons and distortion of carbon-atom polygons. In the global analysis beyond the scale of minimal carbon-atom polygons, we obtained large rings that reliably enclosed defect clusters, covering cases from isolated defects to large complex aggregates. The key procedures to obtain the rings were downsampling of point clouds from non-hexagonal polygons and optimization of  $\epsilon$ . The number of polygons per ring increased with irradiation time, demonstrating that defect clusters grew in an interconnected manner. In conclusion, PH provides both local descriptors of minimum units and a global descriptor of their connectivity. The developed methodology is expected to be useful for designing structures with long-range features.

#### SUPPLEMENTARY MATERIAL

Additional information is provided in the [supplementary material](#): (S1) cleaning process of graphene sample and observation condition of HR-TEM images, (S2) superlevel filtration on grayscale images and representative examples of g-PDs, (S3) comparison between raw and Gaussian filtered HR-TEM images, (S4) computation of atomic carbon positions using birth points and comparison between experimental and simulated TEM images from the computed atomic model, (S5) separation of linked holes via the watershed method, (S6) comparison in extraction of zero-dimensional holes via the present PH pipeline with a representative binarization method, (S7) roundness calculation and correspondence between example shapes and their roundness values, (S8)  $N$  calculation in polygon number  $N$  using a Voronoi diagram constructed from the birth point distribution, (S9) representative holes obtained by using g-PD's inverse analysis and their  $R$  values, (S10) average  $R$  value and its standard deviation, (S11) determination of b–d pairs used for inverse analysis and optimization of noise bandwidth parameter  $\epsilon$  for calculating stable volume, (S12) edge effect on the inverse analysis results, and (S13) effect of optimization of noise bandwidth parameter  $\epsilon$  on inverse analysis results.

#### ACKNOWLEDGMENTS

This work was financially supported by JST FOREST Program (Grant No. JPMJFR213U, Japan), JST PRESTO (Grant No. JPMJPR17S7), and ARIM (Proposal Nos. JPMXP1225NM5101 and JPMXP1225NM5102). We thank M. Takeguchi for advice on sample preparation and M. Shimizu and E. Kano for advice on TEM observation. We also thank I. Obayashi for useful discussion on homology analysis. We thank Natasha Lundin, Ph.D., from Edanz (<https://jp.edanz.com/ac>) for editing a draft of this manuscript.



## AUTHOR DECLARATIONS

## Conflict of Interest

The authors have no conflicts to disclose.

## Author Contributions

**Ryuto Eguchi:** Conceptualization (equal); Methodology (equal); Software (equal); Validation (equal); Visualization (equal); Writing – original draft (equal). **Ayako Hashimoto:** Conceptualization (equal); Funding acquisition (equal); Project administration (equal); Resources (equal); Supervision (equal); Validation (equal).

## DATA AVAILABILITY

The data that support the findings of this study are available within the article and its [supplementary material](#).

## REFERENCES

- <sup>1</sup>H. Tanaka, H. Tong, R. Shi, and J. Russo, "Revealing key structural features hidden in liquids and glasses," *Nat. Rev. Phys.* **1**, 333–348 (2019).
- <sup>2</sup>J. Kang, X. Yang, Q. Hu, Z. Cai, L.-M. Liu, and L. Guo, "Recent progress of amorphous nanomaterials," *Chem. Rev.* **123**, 8859–8941 (2023).
- <sup>3</sup>Z. Xie, Y. Zhang, S. Huang, Z. Li, Q. Cheng, and J. Zhou, "Towards quantitative determination of atomic structures of amorphous materials in three dimensions," *Nat. Sci. Open* **2**, 20220048 (2023).
- <sup>4</sup>J. C. Meyer, C. Kisielowski, R. Erni, M. D. Rossell, M. F. Crommie, and A. Zettl, "Direct imaging of lattice atoms and topological defects in graphene membranes," *Nano Lett.* **8**, 3582–3586 (2008).
- <sup>5</sup>K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, "Electric field effect in atomically thin carbon films," *Science* **306**, 666–669 (2004).
- <sup>6</sup>K. M. Borysenko, J. T. Mullen, E. A. Barry, S. Paul, Y. G. Semenov, J. M. Zavada, M. B. Nardelli, and K. W. Kim, "First-principles analysis of electron-phonon interactions in graphene," *Phys. Rev. B* **81**, 121412(R) (2010).
- <sup>7</sup>A. A. Balandin, S. Ghosh, W. Bao, I. Calizo, D. Teweldebrhan, F. Miao, and C. N. Lau, "Superior thermal conductivity of single-layer graphene," *Nano Lett.* **8**, 902–907 (2008).
- <sup>8</sup>D. L. Nika and A. A. Balandin, "Two-dimensional phonon transport in graphene," *J. Phys.: Condens. Matter* **24**, 233203 (2012).
- <sup>9</sup>Y. Anno, Y. Imakita, K. Takei, S. Akita, and T. Arie, "Enhancement of graphene thermoelectric performance through defect engineering," *2D Mater.* **4**, 025019 (2017).
- <sup>10</sup>M. Choi, T. G. Novak, J. Byen, H. Lee, J. Baek, S. Hong, K. Kim, J. Song, H. Shin, and S. Jeon, "Significantly enhanced thermoelectric performance of graphene through atomic-scale defect engineering via mobile hot-wire chemical vapor deposition systems," *ACS Appl. Mater. Interfaces* **13**, 24304–24313 (2021).
- <sup>11</sup>R. R. Nair, M. Sepioni, I.-L. Tsai, O. Lehtinen, J. Keinonen, A. V. Krashenninnikov, T. Thomson, A. K. Geim, and I. V. Grigorieva, "Spin-half paramagnetism in graphene induced by point defects," *Nat. Phys.* **8**, 199–202 (2012).
- <sup>12</sup>A. W. Robertson and J. H. Warner, "Atomic resolution imaging of graphene by transmission electron microscopy," *Nanoscale* **5**, 4079–4093 (2013).
- <sup>13</sup>C.-T. Pan, J. A. Hinks, Q. M. Ramasse, G. Greaves, U. Bangert, S. E. Donnelly, and S. J. Haigh, "In-situ observation and atomic resolution imaging of the ion irradiation induced amorphisation of graphene," *Sci. Rep.* **4**, 6334 (2014).
- <sup>14</sup>J. Kotakoski, A. V. Krashenninnikov, U. Kaiser, and J. C. Meyer, "From point defects in graphene to two-dimensional amorphous carbon," *Phys. Rev. Lett.* **106**, 105505 (2011).
- <sup>15</sup>F. Banhart, J. Kotakoski, and A. V. Krashenninnikov, "Structural defects in graphene," *ACS Nano* **5**, 26–41 (2011).
- <sup>16</sup>A. Hashimoto, K. Suenaga, A. Gloter, K. Urita, and S. Iijima, "Direct evidence for atomic defects in graphene layers," *Nature* **430**, 870–873 (2004).
- <sup>17</sup>E. Kano, A. Hashimoto, T. Kaneko, N. Tajima, T. Ohno, and M. Takeguchi, "Interactions between C and Cu atoms in single-layer graphene: Direct observation and modelling," *Nanoscale* **8**, 529–535 (2016).
- <sup>18</sup>F. R. Eder, J. Kotakoski, U. Kaiser, and J. C. Meyer, "A journey from order to disorder—Atom by atom transformation from graphene to a 2D carbon glass," *Sci. Rep.* **4**, 4060 (2014).
- <sup>19</sup>J. Kotakoski, C. H. Jin, O. Lehtinen, K. Suenaga, and A. V. Krashenninnikov, "Electron knock-on damage in hexagonal boron nitride monolayers," *Phys. Rev. B* **82**, 113404 (2010).
- <sup>20</sup>T. Lehnert, M. Ghorbani-Asl, J. Köster, Z. Lee, A. V. Krashenninnikov, and U. Kaiser, "Electron-beam-driven structure evolution of single-layer MoTe<sub>2</sub> for quantum devices," *ACS Appl. Nano Mater.* **2**, 3262–3270 (2019).
- <sup>21</sup>D. Shin, G. Wang, M. Han, Z. Lin, A. O'Hara, F. Chen, J. Lin, and S. T. Pantelides, "Preferential hole defect formation in monolayer WSe<sub>2</sub> by electron-beam irradiation," *Phys. Rev. Mater.* **5**, 044002 (2021).
- <sup>22</sup>A. W. Robertson, C. S. Allen, Y. A. Wu, K. He, J. Olivier, J. Neethling, A. I. Kirkland, and J. H. Warner, "Spatial control of defect creation in graphene at the nanoscale," *Nat. Commun.* **3**, 1144 (2012).
- <sup>23</sup>A. Trentino, J. Madsen, A. Mittelberger, C. Mangler, T. Susi, K. Mustonen, and J. Kotakoski, "Atomic-level structural engineering of graphene on a mesoscopic scale," *Nano Lett.* **21**, 5179–5185 (2021).
- <sup>24</sup>J. Lahiri, Y. Lin, P. Bozkurt, I. I. Oleynik, and M. Batzill, "An extended defect in graphene as a metallic wire," *Nat. Nanotechnol.* **5**, 326–329 (2010).
- <sup>25</sup>Y. Hiraoka, T. Nakamura, A. Hirata, E. G. Escobar, K. Matsue, and Y. Nishiyama, "Hierarchical structures of amorphous solids characterized by persistent homology," *Proc. Natl. Acad. Sci. U. S. A.* **113**, 7035–7040 (2016).
- <sup>26</sup>A. F. Firooz, R. Christensen, C. A. N. Biscio, and M. M. Smedskjaer, "Characterizing medium-range order structure of binary silicate glasses using ring analysis and persistent homology," *J. Am. Ceram. Soc.* **107**, 7739–7750 (2024).
- <sup>27</sup>D. Chen, C.-L. Chen, and G.-W. Wei, "Category-specific topological learning of metal-organic frameworks," *J. Mater. Chem. A* **13**, 9292–9303 (2025).
- <sup>28</sup>Y. Lee, S. D. Barthel, P. Dlotko, S. M. Moosavi, K. Hess, and B. Smit, "High-throughput screening approach for nanoporous materials genome using topological data analysis: Application to zeolites," *J. Chem. Theory Comput.* **14**, 4427–4437 (2018).
- <sup>29</sup>S. S. Sørensen, T. Du, C. A. N. Biscio, L. Fajstrup, and M. M. Smedskjaer, "Persistent homology: A tool to understand medium-range order glass structure," *J. Non-Cryst. Solids: X* **16**, 100123 (2022).
- <sup>30</sup>Y. Xiao, T. Du, S. S. Sørensen, Z. Chen, C. A. N. Biscio, L. Fajstrup, M. Bauchy, and M. M. Smedskjaer, "Deciphering the hierarchical structure of phosphate glasses using persistent homology with optimized input radii," *Phys. Rev. Mater.* **7**, 065602 (2023).
- <sup>31</sup>R. Eguchi, Y. Wen, H. Abe, and A. Hashimoto, "Interpretable structural evaluation of metal-oxide nanostructures in scanning transmission electron microscopy (STEM) images via persistent homology," *Nanomaterials* **14**, 1413 (2024).
- <sup>32</sup>M. Uy, D. Kido, and M. Kimura, "A topological approach to understanding crack initiation and propagation in carbon fiber reinforced polymer composites under an opening load," *Sci. Rep.* **15**, 19441 (2025).
- <sup>33</sup>A. N. Duman, "Grain analysis of atomic force microscopy images via persistent homology," *Ultramicroscopy* **220**, 113176 (2021).
- <sup>34</sup>Y. Onodera, S. Kohara, P. S. Salmon, A. Hirata, N. Nishiyama, S. Kitani, A. Zeidler, M. Shiga, A. Masuno, H. Inoue, S. Tahara, A. Polidori, H. E. Fischer, T. Mori, S. Kojima, H. Kawaji, A. I. Kolesnikov, M. B. Stone, M. G. Tucker, M. T. McDonnell, A. C. Hannon, Y. Hiraoka, I. Obayashi, T. Nakamura, J. Akola, Y. Fujii, K. Ohara, T. Taniguchi, and O. Sakata, "Structure and properties of densified silica glass: Characterizing the order within disorder," *NPG Asia Mater.* **12**, 85 (2020).
- <sup>35</sup>A. Tirelli and K. Nakano, "Topological data analysis for revealing the structural origin of density anomalies in silica glass," *J. Phys. Chem. B* **127**, 3302–3311 (2023).
- <sup>36</sup>E. Minamitani, I. Obayashi, K. Shimizu, and S. Watanabe, "Persistent homology-based descriptor for machine-learning potential of amorphous structures," *J. Chem. Phys.* **159**, 084101 (2023).
- <sup>37</sup>M. Yoshimoto, K. Ito, and K. Omote, "Local structure of amorphous carbon investigated by x-ray total scattering and RMC modeling," *Sci. Rep.* **14**, 25298 (2024).

- <sup>38</sup>E. Minamitani, T. Nakamura, I. Obayashi, and H. Mizuno, "Persistent homology elucidates hierarchical structures responsible for mechanical properties in covalent amorphous solids," *Nat. Commun.* **16**, 8226 (2025).
- <sup>39</sup>A. Hirata, T. Wada, I. Obayashi, and Y. Hiraoka, "Structural changes during glass formation extracted by computational homology with machine learning," *Commun. Mater.* **1**, 98 (2020).
- <sup>40</sup>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).
- <sup>41</sup>H. Edelsbrunner, D. Letscher, and A. Zomorodian, "Topological persistence and simplification," *Discrete Comput. Geom.* **28**, 511–533 (2002).
- <sup>42</sup>H. Edelsbrunner and E. P. Mücke, "Three-dimensional alpha shapes," *ACM Trans. Graphics* **13**, 43–72 (1994).
- <sup>43</sup>G. Kusano, Y. Hiraoka, and K. Fukumizu, "Persistence weighted Gaussian kernel for topological data analysis," *Proc. Mach. Learn. Res.* **48**, 2004–2018 (2016).
- <sup>44</sup>Z. Zhu, Z. -G. Fthenakis, and D. Tománek, "Electronic structure and transport in graphene/haeckelite hybrids: an *ab initio* study," *2D. Mater.* **2**, 035001 (2015).
- <sup>45</sup>O. V. Yazyev and S. G. Louie, "Electronic transport in polycrystalline graphene," *Nat. Mater.* **9**, 806–809 (2010).
- <sup>46</sup>A. Y. Serov, Z.-Y. Ong, and E. Pop, "Effect of grain boundaries on thermal transport in graphene," *Appl. Phys. Lett.* **102**, 033104 (2013).
- <sup>47</sup>O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," *Lect. Notes Comput. Sci.* **9351**, 234–241 (2015).
- <sup>48</sup>V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder–decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.* **39**, 2481–2495 (2017).
- <sup>49</sup>L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder–decoder with Atrous separable convolution for semantic image segmentation," *Lect. Notes Comput. Sci.* **11211**, 833–851 (2018).
- <sup>50</sup>T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proceedings of the IEEE Conference Comput. Vis. Pattern Recognit. (CVPR)* (IEEE, 2017), pp. 936–944.
- <sup>51</sup>I. Obayashi, Y. Hiraoka, and M. Kimura, "Persistence diagrams with linear machine learning models," *J. Appl. Comput. Topol.* **1**, 421–449 (2018).
- <sup>52</sup>I. Obayashi, "Stable volumes for persistent homology," *J. Appl. Comput. Topol.* **7**, 671–706 (2023).
- <sup>53</sup>Y.-T. Zhang, Y.-P. Wang, X. Zhang, Y.-Y. Zhang, S. Du, and S. T. Pantelides, "Structure of amorphous two-dimensional materials: Elemental monolayer amorphous carbon versus binary monolayer amorphous boron nitride," *Nano Lett.* **22**, 8018–8024 (2022).
- <sup>54</sup>Z. El-Machachi, M. Wilson, and V. L. Deringer, "Exploring the configurational space of amorphous graphene with machine-learned atomic energies," *Chem. Sci.* **13**, 13720–13731 (2022).
- <sup>55</sup>X. Zhang, Y.-T. Zhang, Y.-P. Wang, S. Li, S. Du, Y.-Y. Zhang, and S. T. Pantelides, "Structural and mechanical properties of monolayer amorphous carbon and boron nitride," *Phys. Rev. B* **109**, 174106 (2024).