HOKKAIDO UNIVERSITY

Title	Bridging a mesoscopic inhomogeneity to macroscopic performance of amorphous materials in the framework of the phase field modeling
Author(s)	Avalos, Edgar; Xie, Shuangquan; Akagi, Kazuto et al.
Citation	Physica. D, Nonlinear phenomena, 409, 132470 https://doi.org/10.1016/j.physd.2020.132470
Issue Date	2020-08-01
Doc URL	https://hdl.handle.net/2115/86489
Rights	© 2020. This manuscript version is made available under the CC-BY-NC-ND 4.0 license http://creativecommons.org/licenses/by-nc-nd/4.0/
Rights(URL)	https://creativecommons.org/licenses/by-nc-nd/4.0/
Type	journal article
File Information	PHYSD_2020_61_Revision 1_V0.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

Bridging a mesoscopic inhomogeneity to macroscopic performance of amorphous materials in the framework of the phase field modeling

Edgar Avalos, Shuangquan Xie, Kazuto Akagi and Yasumasa Nishiura[1]

Mathematical Science Group, WPI-Advanced Institute for Materials Research (AIMR), Tohoku University, Japan
[1]*Research Institute for Electronic Sciences, Hokkaido University and MathAM-OIL, Tohoku University and AIST, Japan*

Abstract

One of the big challenges in materials science is to bridge microscopic or mesoscopic properties to macroscopic performance such as fracture toughness. This is particularly interesting for the amorphous materials such as epoxy resins because their micro/meso structures are difficult to characterize so that any information connecting different scales would be extremely useful. At the process level, the polymerization rate can be changed experimentally that influences a lot over the performance of materials, however, it is known that the maximum toughness does not always appear at the maximum polymerization rate, which suggests that some differences in the micro/meso-scopic structure affect the macroscopic property behind. The goal of this article is to present a framework to bridge between a mesoscopic observation of X-ray CT images and the macroscopic criterion of fracture toughness computed via phase field modeling. First we classify the data of the X-ray image with different polymerization rates by using two different methods: one is SVD and the other is persistent homology. Secondly we compute a crack propagation of each sample and evaluate a scalar value called the effective toughness (ET) via J-integral, which is one of the good candidates indicating a toughness of materials. It turns out that ET reflects the performance of each sample and consistent with the experimental results.

Keywords: SVD, persistent homology, denoising, amorphous materials, phase field model, fracture toughness, effective toughness, crack propagation

1. Introduction

High fracture toughness, light weight and easy manufacturing are some of the aspects that make epoxy resins very attractive in applications ranging from micro electronics and housewares

to aerospace components [1], however the Process-Structure-Properties-Performance (PSPP) relation is still unclear even for the simplest family that uses polyfunctional primary amines as a hardener. It is well known that the polymerization rate of epoxy resins is affected by the heating temperature and time. The state of these resins changes from liquid to gel, and then to solid when the polymerization rate increases. Recently the fact that actively and scarcely reacted domains are formed during the process [2] has become convincing with the development of advanced measurement. This finding brings a new insight into the origin of the structural heterogeneity in epoxy resins, and the heterogeneity at the micro/meso-level has a profound relation with the macroscopic performance [3, 4, 5]. It is therefore reasonable to assume that the development of the three-dimensional (3D) network of covalent bonds leads to an increase in the fracture toughness. However, the experimental results show that the maximum toughness does not always appear at the maximum polymerization rate, which suggests that differences in the microscopic structure are hidden behind the macroscopic polymerization rate. Actually the microstructure of epoxy resins has been recently unveiled by the advances of scanning electron microscope (SEM) and atomic force microscope (AFM). These techniques allow us to resolve submicrometer-sized granular structures and micrometer-sized inhomogeneous structures [6, 7, 8, 9]. While SEM and AFM provide 2D information, X-ray computed tomography (CT) is a powerful tool to visualize the inner architecture of 3D objects based on their electron density. The most familiar application is for medical uses, but industrial applications have also become popular for internal inspection of materials. The 3D images obtained using a commercial apparatus, typically have a resolution of 1-2 μm, which can resolve micro-cracks or voids in various materials. This technique is well established to observe the microstructure of polymeric materials made of heterogeneous components with different densities such as block copolymers, however detecting inhomogeneities in single amorphous materials is still challenging. In fact, X-ray images of thermosetting resins are almost undistinguishable to the naked eye as shown in Fig. 1. More importantly, the size of the structural inhomogeneities seems to be comparable to the resolution of CT images so that it is not a priori clear to identify the "background" image shared by all sample data. The information obtained from the CT observations is convoluted onto the background part (i.e., spotty patterns in the image) originated from the fluctuation of the X-ray intensity. Nevertheless we are able to present a reasonable definition of background in Fourier space, which allows us to classify the samples with different processes and extract the inherent inhomogeneities after removing the background image. For this purpose, we introduce two

statistical approaches to classify the samples in section 2: one is a well-known SVD (singular value decomposition) method, and the other is a topological data analysis, namely persistent homology [10, 11]. Both methods allow us to categorize the samples depending on the polymerization process, however they extract different information on the inhomogeneities hidden in each sample data. In particular, the topological approach so called the persistent diagram (PD) tells us appropriately reduced but quantitative information based on the "hole" structure in a generalized sense embedded in the samples. More precisely see the Ref.[10, 11, 12]. Although some of the conventional techniques use X-rays to identify crystal structures, to the best our knowledge, the use of X-ray images to correctly identify epoxy resins has not been appropriately addressed. The structural information itself is important, however our goal is how we can make a link micro/meso-scopic information to the macroscopic performance via computable criterion based on the classification, namely, what kind of micro/meso-scopic property has a strong influence over the dynamic performance of materials. The statistical methods in section 2 suggest a couple of important facts: firstly all the samples have the same dominant frequency, secondly they also show significant differences, for instance, the intensity of gradient or total variation of the inhomogeneity, which will be explained in section 2.4. This strongly motivates us to see how such a difference affects the performance of materials.

In order to study the macroscopic performance such as fracture toughness theoretically, one of the powerful methods is to apply the phase field approach and derive a criterion by observing the time evolution of the materials under some external stress. The scalar quantity effective toughness (ET) introduced by Ref.[13] is one of the good macroscopic indicators for materials performance, which is defined by the maximum value of the J-integral during the time course of crack propagation. Here a natural question arises "how do we make a bridging between the two different scales?". Namely, our task is to implement mesoscopic information based on X-ray CT images to the macroscopic phase field model, which eventually leads to the indicator ET for dynamic performance at macroscopic level. Here note that the bright part of a X-ray CT image corresponds to high electron density ($\coloneqq$ high atomic density), and that the three-dimensional network of covalent bonds hinders the packing of molecules so that highly polymerized part corresponds to lower (dark) density. On the other hand, it is known that highly polymerized part has larger Young's modulus due to the strength of network. Therefore dark part is considered to have larger Young's modulus locally, which leads to the following assumption that inhomogeneity of X-ray image is inherited to that of Young's modulus distribution of epoxy samples. Since all of the epoxy samples are built

using blocks of the same monomer but different polymerization rates, it is reasonable to assume that the local toughness of each sample is uniform, but its elasticity field (Young's modulus) is heterogeneous depending on the curing process. This motivates us to make an attempt to simulate the phase field model with the parameter of the Young's modulus being imported from the X-ray CT data. More precisely, the Young's modulus is proportional to the gray-scale of the X-ray image. Under this framework, we numerically study the crack propagation in all samples and compute their ETs. The result of this naive approach nevertheless agrees with the experimental observations of fracture toughness, which suggests that the above assumption about the linkage between microscopic information and macroscopic performance is a reasonable one.

As a summary, we propose a new framework to utilize X-ray CT images to predict the mechanical toughness of epoxy resins by analyzing microscopic variations of the structural inhomogeneities in the samples which are related to the curing process. Such a connection has not been studied as far as we know and this study provides a better understanding of the relation between the micro-structure of epoxy resins observed by X-ray images and the material performance.

The rest of this work is organized as follows. We start off with a brief description of the experimental setup to probe the samples of materials. In section 2.2, we use SVD to discover a couple of important features encoded in the X-ray images. We employ basic statistical reasoning to infer the meaning of the learned features. Several methods are presented for the removal of background image. We illustrate how the topological approach PD can be used in the clustering process and to extract a geometrical information from the X-ray data. Then in section 2.4 we show that the Euclidean norm of the gradient of the intensity in the X-ray images is a quantity that characterizes the material with the best mechanical performance. Section 3 introduces the theory of phase field modeling of crack propagation and presents an indicator of macroscopic toughness called the effective toughness (ET). Then we present results of the numerical simulation of crack propagation and show that ETs of four epoxy samples are consistent with the experiments. In Section 4, we conclude the paper and discuss about the future perspectives.

2. Characterization of X-ray images of thermosetting resins

In this study, we consider X-ray CT images of four different samples of thermosetting resins as shown in Fig. 1. Note that the heating process is different each other, so is the polymerization rate. The samples and original experimental images were prepared and characterized by NIPPON STEEL

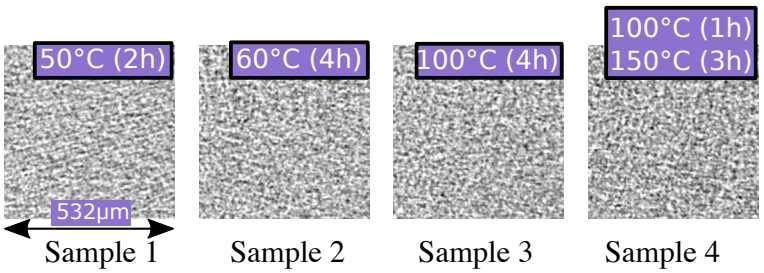


Figure 1: X-ray CT images of four samples of epoxy resins cut out from the left cube in Fig. 3. Although each sample was produced by the different process, it is difficult to distinguish what is the essential difference among them by our naked eyes.

Chemical & Material CO., LTD. The physical properties of the samples are shown in Table 1 and Fig. 2. Overall, the sample 3 shows the best performance with respect to the absorbed energy, e.g., the largest area under the s-s curve, although the fracture toughness of samples 2 and 3 are more or less the same. Here a natural question arises "what is the geometrical feature that makes sample 3 tougher than others?" Such a characterization may be hidden in the heterogeneous patterns of X-ray images, however it is not so clear how we can extract such a feature from more or less the same random morphologies. Moreover various types of noise and fluctuation are unavoidable in those experimental data so that we have to remove them in an appropriate way. We try to resolve these issues step by step in the following subsections. We begin this section by describing the experimental observations and then we will present the two methods to classify the X-ray CT images.

Table 1: Properties of the samples (provided by NIPPON STEEL Chemical & Material CO., LTD.)

	Sample 1	Sample 2	Sample 3	Sample 4
Polymerization Rate [%]	13.7	45.9	60.3	90.0
Density $[g/cm^3]$	1.151	1.157	1.145	1.143
Fracture Toughness [MPa·m$^{1/2}$]	0.16	1.03	0.99	0.83

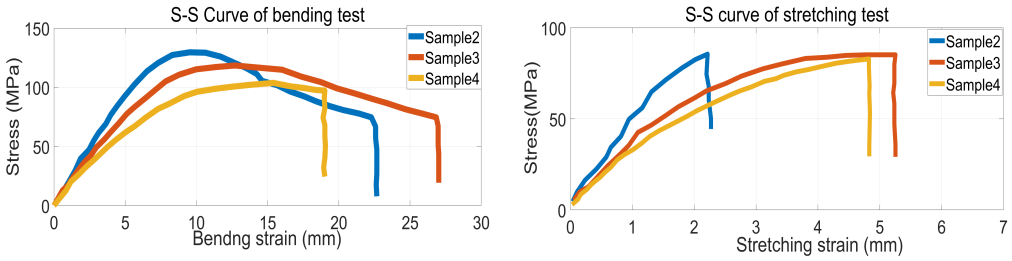


Figure 2: Strain-stress curve of three samples under bending and stretching test, respectively. In both cases, sample 3 has the highest total absorbed energy (the area under s-s curve). Note that sample 1 is too fragile and not adequate for mechanical test.

2.1. Preparation of imaging data

The samples 1 to 4 used in our analysis have the same chemical composition of bisphenol A type epoxy molecule and hardener molecule (primary diamine) in the ratio of 3:1, and only the heating temperature and time are different. They were prepared as a plate with a thickness of 1mm and their mechanical properties were measured by the standard means of evaluation. The polymerization rate, density and fracture toughness are summarized in Table 1. The X-ray CT images were obtained using a commercial apparatus with the resolution of 2 μm/pixel by NIPPON STEEL Chemical & Material CO., LTD. Avoiding the distorted central area, two 266×266 pixels squares were cut out from left and right sides, and two cubes (left cube, right cube) were constructed using 266 slices every 2 μm (Fig. 3). In a X-ray CT image, the bright part corresponds to high electron density (≒ high atomic density) part. Since the three-dimensional network of covalent bonds hinders the packing of molecules, we assume that highly polymerized part corresponds to low density (dark) part. On the other hand, highly polymerized part has larger Young's modulus, therefore, dark part is considered to have larger Young's modulus than bright one. This correspondence becomes a key when we compute crack propagation in the framework of phase field modeling in section 3.

The X-ray CT images are two-dimensional projections of a three-dimensional object. Projecting a 3D object onto a plane may lead to ambiguities in the identification of the structural features. The spotty pattern is noise originating from fluctuation of X-ray intensity, and there still remain some streaks and ring artifacts. Ideally we want to remove these ambiguities that collectively constitute a "background". There are many different ways to characterize the background. We present some

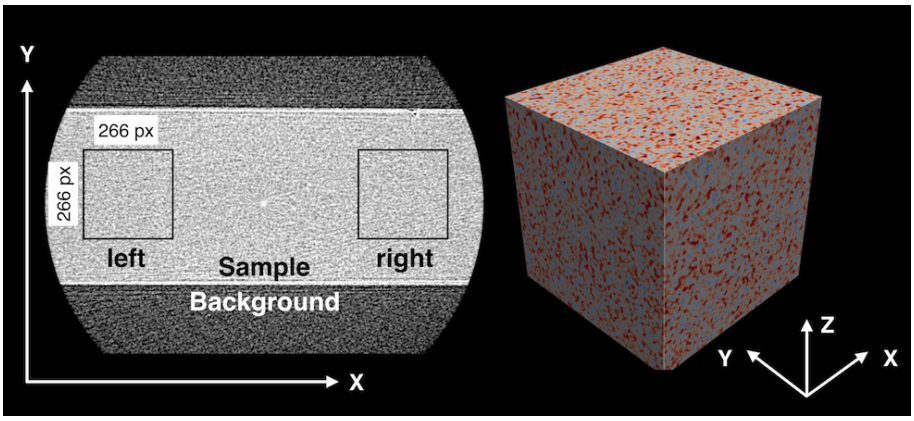


Figure 3: (Left) An example of X-ray CT slice image. The middle bright area corresponds to the sample. The upper and lower dark background area shows a noisy image originated from the fluctuation of X-ray intensity. (Right) An example of cubic data constructed with 266 slices. The part brighter (darker) than average is colored by red (blue).

of them in section 2.2.2 and section 2.3.1, and enhance the distinction among these four samples by removing it. A good news is that we are able to classify the samples more or less independently of the methods and to extract a key geometrical feature hidden in the image.

2.2. Singular value decomposition (SVD)

Machine learning algorithms are widely used to identify and categorize different samples of materials and to predict their properties [14, 15]. Among some few applications, we can mention the use of principal component analysis to categorize data from a superconductor database to identify compounds according to their valency changes [16]. A notable instance is in the case of crystals, in which the crystallographic information can be utilized to predict the bulk modulus via a decision trees-based algorithm [17]. In another case also related to crystals, a convolutional neural network is trained to classify crystals with astonishing accuracy [18]. Although some of techniques mentioned above use X-rays to identify crystal structures, to the best our knowledge the use of X-ray images to correctly identify epoxy resins has not been appropriately addressed.

We start by describing the SVD-based approach of clustering to map the X-ray images into different groups, which is a machine learning tool that has extraordinary applications [19, 20, 21, 22, 23].

A total of 1064 (= 266×4) slices of samples 1 to 4 are concatenated into a matrix X =

$(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m)$. Each vector $\mathbf{x}_k$ with $k = 1, 2, ..., m$ contains the intensity field, x_i of each X-ray image, with $i = 1, 2, ..., N$. Notice that each image consists of 266×266 pixels or $N = 70756$ and $m = 1064$. Because of the large size of X, we consider the singular value decomposition of $X = U\Sigma V^*$, in which the columns of U are eigenvectors for the matrix of images X and V^* indicates the complex conjugate of V. These eigenvectors define directions along which the images have the largest variance. In this context, the eigenvectors are the principal component directions PC_i. By computing the inner product of each slice with the eigenvectors U, we project each X-ray observation onto a space of reduced dimension. The first three coordinates of this space are the principal components PC_1, PC_2 and PC_3. We study the clusters obtained from different projections in this eigenspace and as shorthand for the projection onto the plane defined by the directions PC_1 and PC_2, we simply refer to it as the plane x-y. As an example, the resulting clusters in this eigenspace are shown in Fig. 4c. The ordering of the clusters shown in Fig. 4c on the x-y plane will be explained in section 2.2.1.

Fig. 5 shows the distribution of loadings scores for the first two modes for the X-ray images. The distribution of loadings of the first mode Fig. 5a shows that sample 1 appears at the rightmost location on the horizontal axis, suggesting that the X-ray images of this sample have the highest average brightness. The distribution of loadings of the second mode Fig. 5b shows that X-ray images of sample 2 have the highest deviation from the average brightness. This interpretation will be further clarified by the statistical results presented in section 2.2.1.

2.2.1. Basic statistical quantities

In this section, we look into some basic statistical parameters in the set of the X-ray images. The aim is to explain the meaning of the eigendirections obtained from the SVD above. We start with an analogy. It has been shown that one approach to recognize handwriting digits consists in relating the the size of the characters to the value along the principal components [19]. In other words, the mean value of the images seems to be related to ordering of the center of mass of clusters in an eigenspace. This shows that the principal components can be related to specific features of the samples.

To analyze the distribution of the pixel values in the X-ray images, we proceed as follows. For each grayscale X-ray image with pixels taking values between 0 and 255, we compute the mean value in a fixed squared domain (266×266). The total intensity of the images divided by the pixel

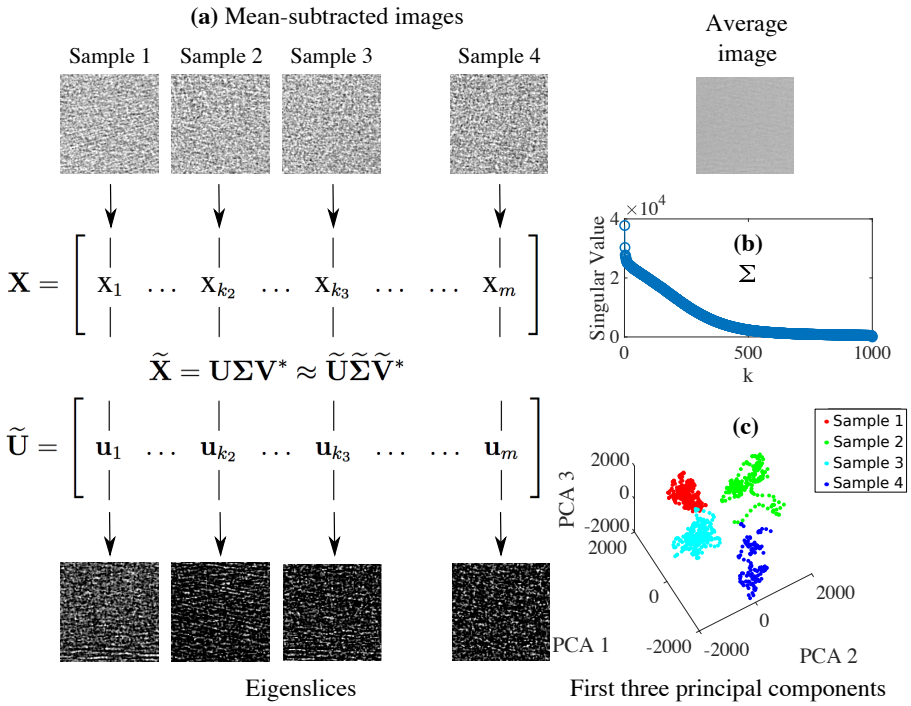


Figure 4: SVD and clustering of samples. (a) X-ray images are arranged in tandem as columns of a matrix X and subsequently the SVD is computed. The eigenvectors u_k span a space of X-ray images. (b) Singular values (σ_i) show that the first few eigenvectors can be used to summarize the most prominent features contained in the original matrix X. (c) Clusters of different samples of materials. The clusters are projections onto the three first eigenvectors obtained from the SVD of matrix X containing 1064 columns.

count corresponds to the mean value of the images. Fig. 6a shows the mean value of all the slices of samples 1 to 4. The ordering of the mean values corresponds to that of the loadings in the histograms seen in Fig. 5a, according to the sequence 1,2,3,4. Fig. 6b shows detrended values of Fig. 6a. To obtain these values, we compute the least-square fit to a straight line to the data and subtract the resulting function from the original distribution. Similarly, we compute the standard deviations of the mean value, $\sigma(\overline{X_{Ti}})$, and the ordering of these standard deviations corresponds to that of the loadings in the histograms seen in Fig. 5b, according to the sequence 2,4,3,1.

We have shown how basic statistical analysis of the original X-ray CT images suggests that the quantities $\overline{X}$ and $\sigma(\overline{X_{Ti}})$ explain the ordering of the clusters along PC_1 and PC_2, respectively.

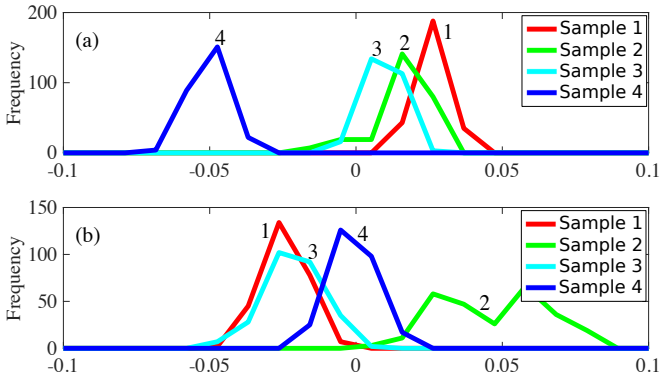


Figure 5: Histograms of the distribution of loadings for samples 1, 2, 3 and 4 on the first (a) and second (b) dominant SVD modes. The loadings come from the columns of the V matrix of the SVD. The ordering of loadings along the first mode shown in (a) suggests that a prominent feature along PC_1 orders the samples according to the mean value of the intensity of the slices following the sequence, from right to left as: 1, 2, 3, 4 (Fig. 6a). The ordering of loadings along the second mode shown in (b) suggests that a prominent feature along PC_2 orders the samples according to the variance of the average intensity of the slices following the sequence, from right to left as: 2, 4, 3, 1 (Fig. 6b).

2.2.2. Removal of the background of the images

SVD was used to group unprocessed X-ray images into different categories above. This is possible because the SVD captures the features with the largest variance. However, in order to extract intrinsic nature of inhomogeneity, it is necessary to isolate the features contained in the images. To do this, we propose some techniques to remove the background of the images. This will be a basic step to link two different scales in the following sections and to find a macroscopic criterion in section 3.

As it can be seen in Fig. 1 the X-ray images are almost undistinguishable from one another. The intensity of these images has several contributions and the hope is to discover what are the most relevant features to look at and use these features to express the mechanical properties of the materials. In what follows, we employ the techniques to isolate the features encoded in the X-ray images.

ROLLING BALL ALGORITHM. First we start with the Sternberg's method also known as rolling ball algorithm [24]. This technique removes smooth continuous backgrounds from gels and other images and we use this method as a simple approximation to compute a background to be subtracted

from the original images, hopefully removing large spatial variations of the background intensities. One of the advantages of this method is its simplicity. It defines a background as the opening of the image based on a structural element. This background is then subtracted from the original image. A drawback of implementing this method is that it requires a user-defined size and shape of the structural element. The Sternberg's method preserves the original clustering as shown in Fig. 8f.

HIGH-PASS FILTER. Another simple method to construct a background is to consider a high-pass filter to remove the the low-frequency features in the images [25]. Fig. 7 shows the details of the high-pass filter used to construct Fig. 8c. After computing the power spectrum of each image, we suppress all frequencies lower than a cutoff of appropriate size. As it can be seen, the cutoff is small enough such that the dominant frequency corresponding to the characteristic wavelength remains unchanged. The strong point of this method is that it preserves the characteristic length, which has a value of about 28.6 μm. However one disadvantage of this method is that it is necessary to supply a parameter, which in this case is the threshold of frequencies to be suppressed. Because of this, we essentially subtract the low-frequency features, then the clusters will arrange the high-frequency features along the second PC, instead of the third PC as shown in Fig. 8g.

EIGENBACKGROUD REMOVAL. While traditional tools such as the Sternberg's method or the Fourier high-pass filter preserve the ordering of the clusters, these methods require at least one free parameter. This parameter is the radius and shape of the Sternberg's ball or the threshold of the high-pass filter. Here we propose a parameter-free method to subtract a background from the X-ray images.

We define an "eigenbackground" and then remove it from the X-ray images. The idea behind this method is that a background is a field that is shared by all the images. We proceed as follows. Firstly, we fourier-transform a matrix containing all the slices, X, and then from the the SVD of the transformed matrix, $X = U\Sigma V^*$, we obtain the eigenvector matrix, U, which has columns that are orthogonal with one another. Because the columns of U are uncorrelated, we can remove any column in U without affecting the remaining columns. In particular, we remove the first column, which represents the 'background' field. This is reasonable because statistically this is the meaning of the first eigenvector. It is noteworthy to mention that the first eigenvector contains about five times more energy than the following eigenvector, as shown in Fig. 9.

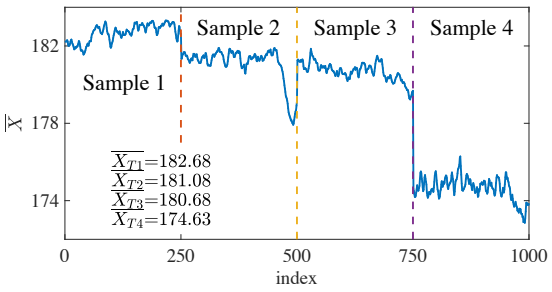
Secondly, we proceed to reconstruct all the slices but now we push aside the eigen-background

vector, as shown in equation 1.

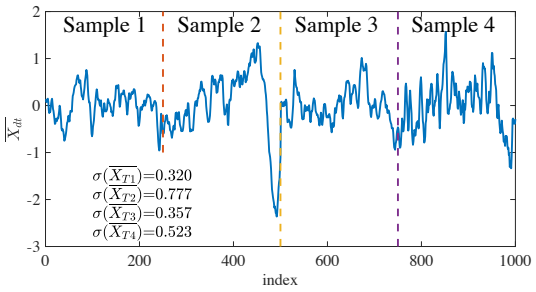
$$X_{new} = \sum_{j=2} \langle U_j, X_j \rangle U_j \tag{1}$$

Once we have the new set of slices, X_{new}, then again we perform a SVD as $X_{new} = U_{new}\Sigma_{new}V^*_{new}$, followed by clustering. Fig. 8h shows that same clustering is preserved and now the eigenvector containing features with large wave numbers is the second column of U_{new}. The interpretation of this is that (i) we can safely remove the eigenbackground vector to reconstruct all the X-ray images and (ii) all the clusters will be preserved after removing such eigenbackground vector. Fig. 8d shows the result of subtracting the eigenbackground from a sample image.

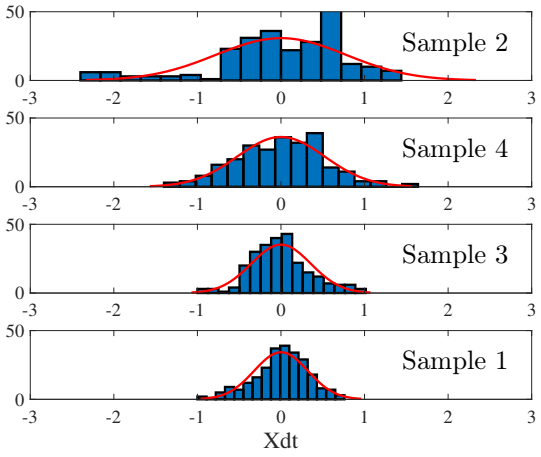
Unlike the two other methods mentioned above, the eigenbackground removal is optimal in the sense that the columns of the eigenvector U are hierarchically ordered by how much correlation they capture in the X-ray images. So, the first column captures the largest correlation variations and this column happens to be the mean value in the space of frequencies. In summary, this method does not need any free parameter.



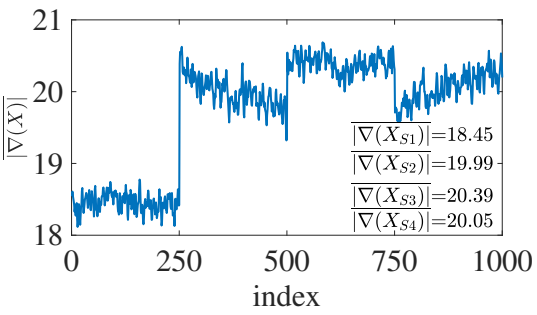
(a) Mean of each slice.



(b) Mean of each slice (Detrended).



(c) Standard deviation of the mean



(d) Module of the gradient

Figure 6: (a) Mean of the columns (slices) of the matrix X. From left to right, the vertical lines separate samples 1 to 4. Mean value of each sample is shown as $\overline{X_{Si}}$. The loadings in Fig. 5a are ordered, from left to right, in increasing values of $\overline{X_{Si}}$. (b) Detrended value of the mean of the matrix X to remove the linear trend from each set of slices (sample 1 to 4.). (c) Standard deviation of the mean. See numerical values in Fig. 6b. (d) Mean of the module of the gradient of each slice, $\frac{1}{N}\int(f_x^2+f_y^2)^{1/2}dxdy$, with N being the pixel count. Notice that sample 3 has the largest average value of the module of the gradient.

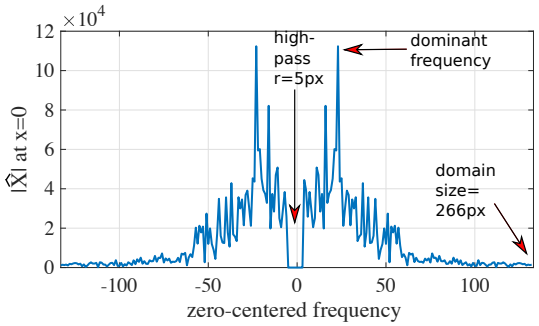


Figure 7: Power spectrum of a typical X-ray image of sample 1 with a high-pass filter of size 5pixels. The average dominant frequency is common to all samples.

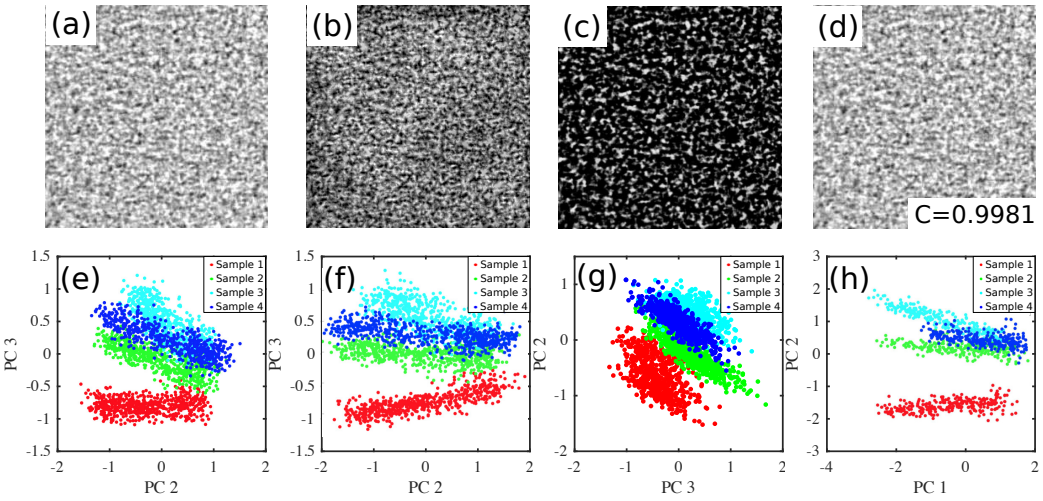


Figure 8: Different techniques to remove background. (a) Original image of a typical sample. Three techniques of background removal are considered. (b) Sternberg method of background removal using a ball of size 50 pixels. (c) High-pass filter with a circle of radius 5 pixels. (d) Eigenbackground removal. The first column of the left singular vector of the SVD was suppressed. Although the new image looks quite similar to the original, its correlation with the original image is slightly smaller than 1. (e-h) Clustering of the X-ray images in the frequency domain. For each image we compute the two-dimensional Fourier transform as $F(u,v) = \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f(x,y) \exp\{-(i\frac{2\pi ux}{N} + i\frac{2\pi vy}{M})\}$, with $N = M = 266$. (e) Original clustering obtained before background removal. (f) Clustering after the Sternberg method of background removal. (g) Clustering after the high-pass filter. (h) Clustering after the eigenbackground removal.

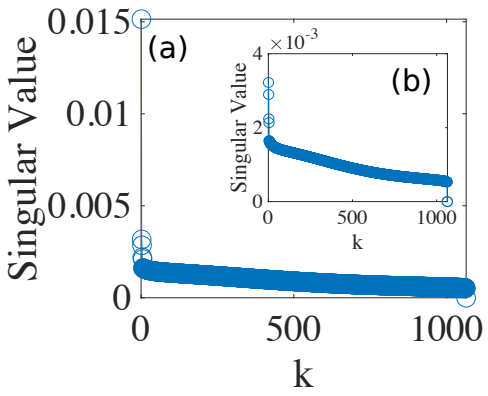


Figure 9: (a) Singular values of the SVD of the Fourier-transformed matrix, X. (b) Singular values of the SVD of the reconstructed matrix, X_{new} as it is shown in equation 1. The energy content of the 1st eigenvalue in (a) is about five times larger than that of the 1st eigenvalue in (b).

2.3. Topological data analysis

Persistent homology (PH) is a new mathematical framework that characterizes the "shape of data" based on the information of "holes". The holes in PH_0 (degree 0 PH), PH_1, PH_2 correspond to connectivity, ring and cavity, respectively. These holes are recorded as a set of birth and death events (generator) by the operation called the filtration. Then a complex image data is converted into a two-dimensional histogram called the persistence diagram (PD). The filtration for gray-scale images is given by the levelset method, where birth and death of the holes by scanning the threshold of brightness are recorded. For 8-bit grayscale images the corresponding PD is spanned by 256×256 two-dimensional brightness space. A large number of generators appear near the diagonal of the PD, but most of them are "noise", as revealed by their short lifetime. By appropriate PD-vector translation reducing their weight, useful classification by machine learning or characterization becomes possible. For the computation of PDs, we used the software HomCloud [26].

2.3.1. Pre-conditioning of the X-ray CT images

The structure of epoxy resin in a 266 pixel (= 532 μm) cube is expected to be isotropic, however, the distribution of averaged brightness of each slice in the given X-ray CT images appears to be anisotropic as shown in Fig. 10. Such an anisotropy should be removed because it affects the center of mass of PDs and disturbs the classification by "shape of data". Therefore, we introduce another pre-conditioning which normalize the average brightness of all the slices cut out along x-, y- and z-axis. We assume that such anisotropy is due to disturbances in the apparatus. To overcome this issue, we applied the series of operations described in Eqns 2-4 to average the level of brightness along multiple axis. The outcome is that the experimental observations appear uniform and isotropic. In other words, all the slices along x-, y- and z-axis appear to have the same mean brightness.

$$\overline{\rho}^{(1)}(x,y,z) = \rho(x,y,z) - \overline{\rho}_z(x,y) + \rho_0 \tag{2}$$

$$\overline{\rho}^{(2)}(x,y,z) = \overline{\rho}^{(1)}(x,y,z) - \overline{\rho}_y^{(1)}(x,z) + \rho_0 \tag{3}$$

$$\overline{\rho}(x,y,z) = \overline{\rho}^{(2)}(x,y,z) - \overline{\rho}_x^{(2)}(y,z) \tag{4}$$

In Eqns 2-4, the function $\rho(x,y,z)$ represents a 3D X-ray CT image, with mean brightness, ρ_0, and $\overline{\rho}_z(x,y)$ is the mean brightness along z-axis. Since the four samples have different densities

according to their polymerization rate, then the corresponding X-ray CT image should have different ρ_0 for each sample. However, while the density of sample 2 is the largest from Table 1, ρ_0 of sample 1 is the brightest among all four samples. Because of this inconsistency, we removed the information of ρ_0 through the operations shown in Eqns. 2-4 (we did not add ρ_0 at the end of Eq 4). Then we perform a topological data analysis based on persistent homology to analyze the microscopic inhomogeneities from non-trivial information in the X-ray images. Notice that the fact that ρ_0 and density align in different order, shows that the electric current of the source of the X-ray light is different for each sample. This means that the standard deviation of the fluctuations of the X-ray intensity, which we consider to be the source of spotty noise pattern, is also different for each sample.

2.3.2. Classification of sample data

First, we divide each 266 pixels cube into eight 133 pixel sub-cubes and then calculate the persistent homology of each cube. We vectorize the degree 2 persistence diagrams (PD_2) obtained by sublevelset. The PD_2 contains information of dark "cavity" surrounded by bright pixels. Then we perform a principal component analysis (PCA) of 32 ($= 8\times4$) vectors. We construct a 3D plot with the axes being the first three principal components, each with a contribution rate of 0.62, 0.28 and 0.05, respectively. Four groups were obtained with eight points each, which correspond to the sub-cubes (Fig. 11). This suggests that the processed 3D data, $\overline{\rho}(x, y, z)$, still contains the structural information required to distinguish the samples. Furthermore, based on the spatial distribution of the eight points observed in Fig. 11, we notice that the sample 2 and the sample 3 have smaller inhomogeneities, meanwhile the sample1 possesses large inhomogeneities along the second principal component.

Next, we generate three sets of 266 two-dimensional images cut out from each of the processed three-dimensional images $\overline{\rho}(x, y, z)$ along x-, y- and z-axis, respectively. Then, they are translated into vectorized persistence diagrams PD_1 using sublevelset (information of dark "rings" surrounded by bright pixels). Applying PCA to the 266 sliced images along x-axis from the left cube, we obtain a 2D plot in which the four samples are classified into three groups using two principal components (Fig. 12a). The contribution rate of the first three principal components PC_1, PC_2, PC_3 is 0.26, 0.09 and 0.07, respectively. The component PC_3 weakly affects the distinction between sample 2 and sample 3, and for this reason PC_3 was omitted for clarity.

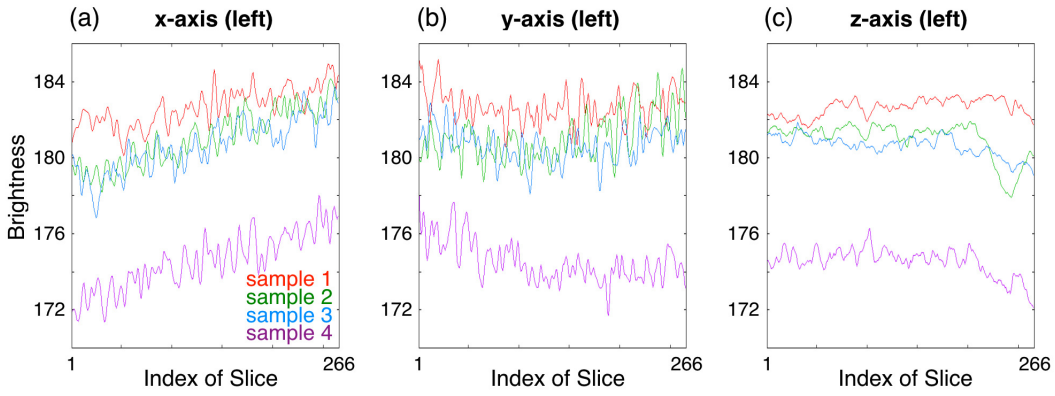


Figure 10: Mean brightness of the sliced image cut out from the left cube in Fig.3 along (a) x-, (b) y- and (c) z-axis. The order of brightness is "sample 1 $>$ sample 2 $\sim$ sample 3 $\gg$ sample 4" in all the cases, but the trend along each axis is clearly different among them. It implies unexpectedly anisotropic component is contained in the experimental raw data.

Using the PC_1 and PC_2 vectors as a ruler, we can "measure" the remaining five sets of sliced images from the left cube along y-, and z- axis, and those from the right cube along x-, y- and z-axis as well. The results are summarized in Fig. 12a-f. The results along the z-axis show relatively larger distribution along PC_2, but all of them give the similar results quantitatively. This is consistent with the assumption that microscopic structure of epoxy samples should be isotropic. Our analysis suggests that it is reasonable to employ the Eqns 2-4 to preprocess the X-ray images.

Figs. 13a-d show the persistent diagrams of PH_1 calculated based on the sublevelset of a sliced image along x-axis in the left cube of sample 1-4. Since they are similar to each other, it is difficult to identify the difference among them with our eyes. In order to know which information is used for grouping in Fig. 12, PC_1 and PC_2 vectors are visualized in the style of PD (Fig. 14). The 1st principal component value of the i-th slice is calculated as follows:

$$(\vec{v}_i - \vec{v}_{avg}) \cdot \vec{v}_{PC_1}$$

$$\vec{v}_{avg} = \frac{1}{N} \sum_{i}^{N} \vec{v}_i$$

Here, $\vec{v}_i$ is a vectorized PD of the i-th slice and $\vec{v}_{PC_1}$ is PC_1 vector. The vector $\vec{v}_{avg}$ represents the average PD. The number of generators in the red (blue) part is larger (smaller) than the average PD, it gives positive contribution to the corresponding principal value. The death values of the

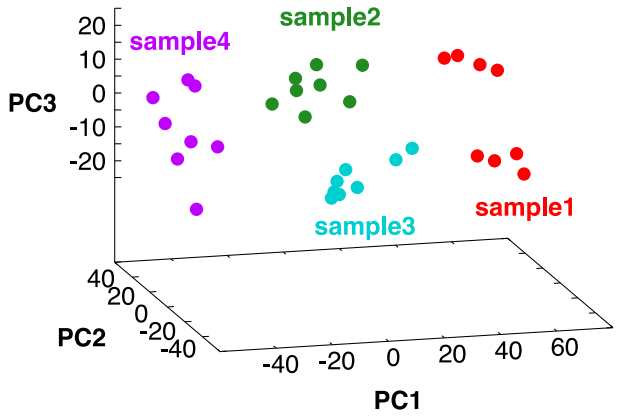


Figure 11: Classification of the four samples based on the left cube in Fig. 3. The cube is divided into eight smaller cubes and the second order persistent homology is calculated for sublevelset of each 3D grayscale image. Principal component analysis (PCA) is performed based on the vectorized persistence diagrams.

generator in the red part of PC_1 are smaller than that in the blue part, which means that the difference in brightness between the perimeter and center of rings (holes in PH_1) is small, i.e., difference in density is small. Figs.12a-f clearly show the first principal component PC1 has a negative correlation with the polymerization rate. Namely, PC1 value decreases with increase of the polymerization rate. Therefore, we can interpret the result of Fig.12 as follows: sample 1, sample 2, 3, and sample 4 have small, medium and large density fluctuation, respectively. Therefore, while sample 1, whose PC_1 is largely positive, has wide area with small density fluctuation, sample 4 has wide area with large density fluctuation.

2.3.3. Inverse analysis of the persistence diagrams

To check the spatial distribution of structural feature corresponding to red and blue parts in the first principal component vector in Fig. 14, which shows good correlation with polymerization rate, we made inverse analysis on the generators whose intensity is larger than 0.02. Fig. 15 is a set of examples representing inverse map of the 200th slices from sample 1-4. The structural feature originated from red part of Fig.14a (small difference in density) decreases with increase in polymerization rate from sample 1 to sample 4, while the structural feature from blue part of Fig.14a (large difference in density) increases. This trend supports the assumption that not small amount of low polymerization area remains even after the thermosetting. As mentioned above, we

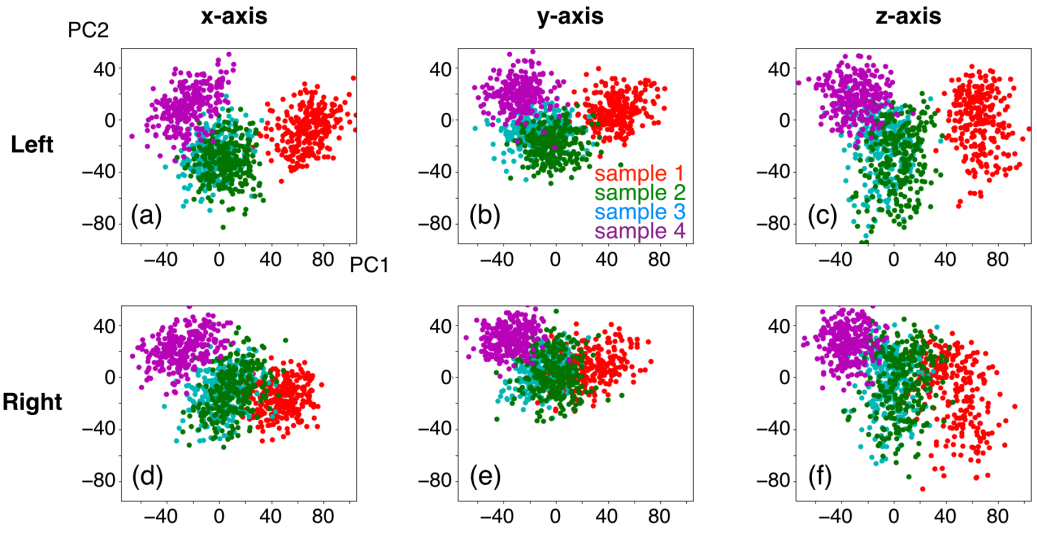


Figure 12: Classification of the four samples using the same "ruler". After the preconditioning, the left and right cubes were divided into 266 slices along each axis. First, principal component analysis was performed based on the vectorized persistence diagrams (PD_1) along x-axis in the left cube. Next, we measure the remaining five sets of data using the first and second principal component vectors as common rulers. Though the slices along z-axis spread more widely than the others, all the six data sets show the same trend quantitatively.

notice that the 1st principal component has a strong correlation with polymerization rate (Fig.12 and Fig.14) and its inverse analysis tells us the domain with large density difference grows from sample 1 to sample 4 (Fig.15). We will see that this observation is consistent with the discussion in the next subsection by using a computable quantity "total variation".

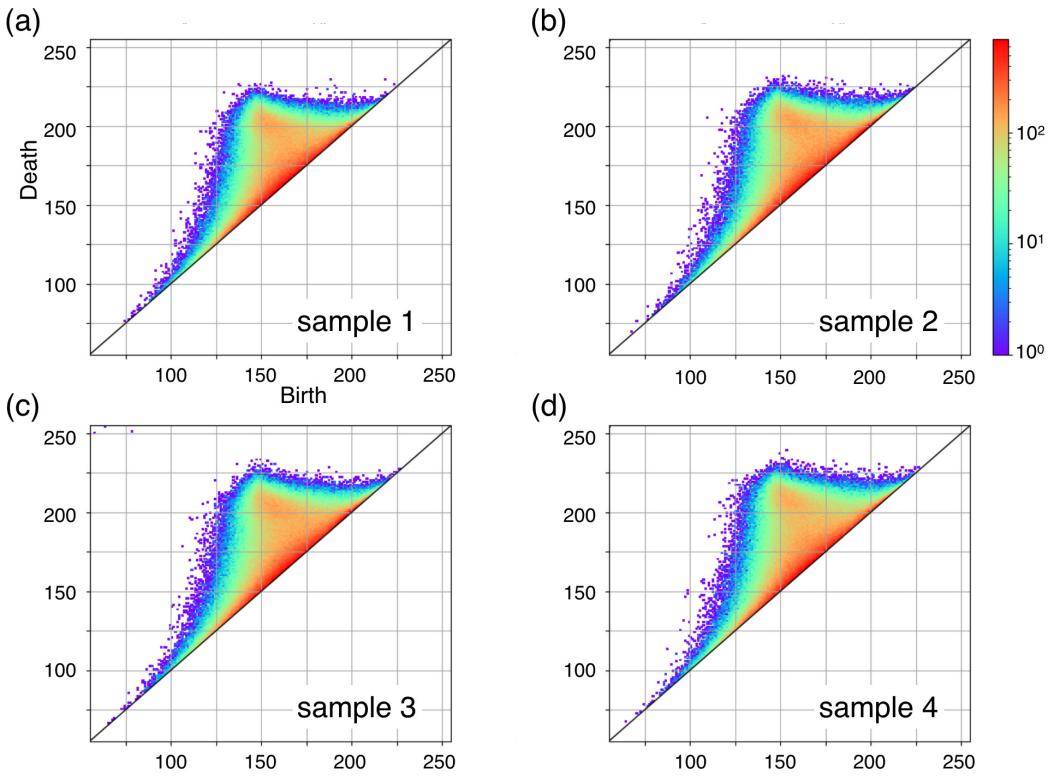


Figure 13: Persistence diagrams of PH_1 that has the information on ring structures. Since the sub-levelset is adopted as the filtration, they are spanned by the two-dimensional brightness space of 8-bit grayscale. (a)-(d) correspond to the sample 1-4, respectively. We can see there is a difference among them, but it is difficult to identify it without further analysis.

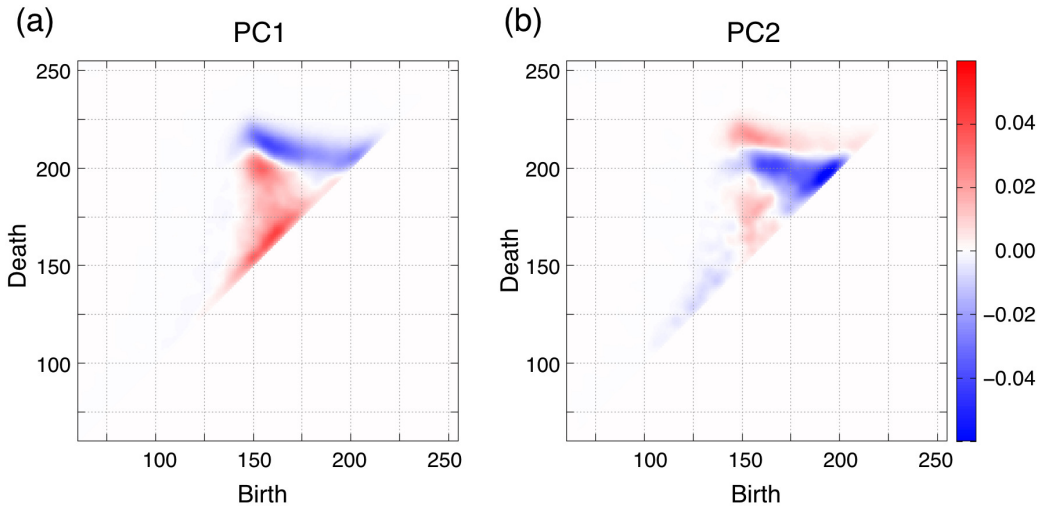


Figure 14: The first and second principal component vectors visualized in the same style with the persistence diagrams (PD) in Fig. 13. They are obtained by PCA on the 266×4 PDs and the red (blue) part shows the area where the number of generators is larger (smaller) than the averaged PD, respectively.

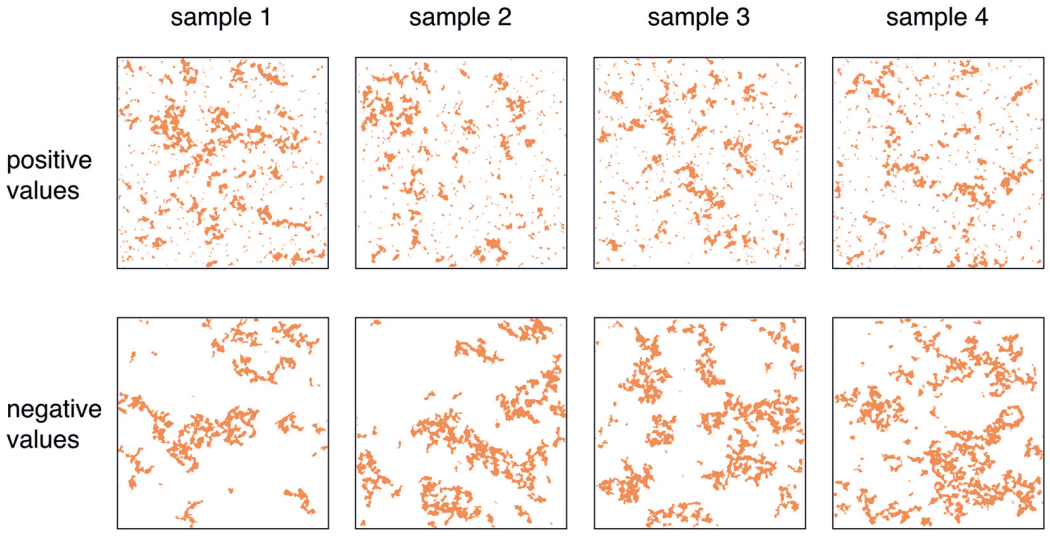


Figure 15: A set of images obtained by the inverse analysis based on PC_1 in Fig. 14 of the 200th slice along x-axis. The four images in the upper half and lower half correspond to the red and blue area of PC_1, respectively. The inverse analysis suggests the importance of the global geometrical network for the classification of four samples.

2.4. Characterization by gradient of the intensity

In computer vision field, regions in the image with large variations in intensity are important features [27]. To locate these features, it is better to study the absolute gradient of the image. In this section, we turn our attention to study the clusters of the projections of the X-ray images in a dimensionally-reduced absolute gradient space. We aim at finding some important features that are interpretable in application. By conducting PCA on the absolute gradient space, we extract a quantity, the total variation of the intensity, as an appropriate feature to distinguish the X-ray images and facilitates the identification of clusters.

We start by looking at the modulus of gradient of the X-ray images, $\sqrt{f_x^2 + f_y^2}$, where f_x and f_y represent discrete approximations of the central difference for interior data points in the horizontal and vertical direction, respectively. For example, for a matrix with unit-spaced data, f, that has horizontal gradient, f_x, the interior gradient values are computed as $\frac{1}{2}(f_{i,j+1} - f_{i,j-1})$. The vertical gradient is computed similarly. We then perform the SVD of a matrix containing column vectors of the absolute gradient of the X-ray images. The next step is to find clusters of different materials with similar modulus of gradient. To do this we project the modulus of gradient of the image onto the first two eigenvectors, PC_1 and PC_2. We are concerned with the interpretation of PC_1 in application. In general, there are no universal rule to interpret PC_1, however, in this case, we can see a direct relation between PC_1 and the mean of the absolute gradient by observation (Fig. 16g-i). The projection of the modulus of gradient of the image onto the first principal component is approximately proportional to the mean of the absolute gradient by observation. For a better understanding, we use the terminology "the total variation", which is widely used in mathematics and signal processing, as the interpretation of PC_1. The total variation of a image is defined as $\int \sqrt{f_x^2 + f_y^2}\ dxdy$ and it is equivalent to the mean of the modulus of gradient up to a constant factor.

Fig. 16a-c show clusters of all samples in all three planes XY, YZ and ZX on the gradient eigenspace spanned by PC_1 and PC_2. In this case, projection to PC_1 corresponds to the total variation. The cluster of sample 3 appears on the rightmost location along the PC_1, indicating that sample 3 contains the largest value of the total variation. Similarly, the ordering of the loadings of the histograms shown in Fig. 16d-f situate sample 3 at the rightmost location in all three planes. This further confirms that sample 3 contains a higher total variation. The claim is that the high total variation of the intensity could be related to the relatively high mechanical performance of

sample 3 . We will verify this claim in section 3.

A word of caution is in order here, though. Table 1 shows that the sample 2 and 3 possess values of fracture toughness that seem to be close to one another, the experimental error in the measurements is unclear. Moreover, the location of the clusters for sample 2 and 3 seems to overlap along the horizontal direction, which makes difficult to distinguish which sample has better mechanical properties. We will see this difficulty in the next section.

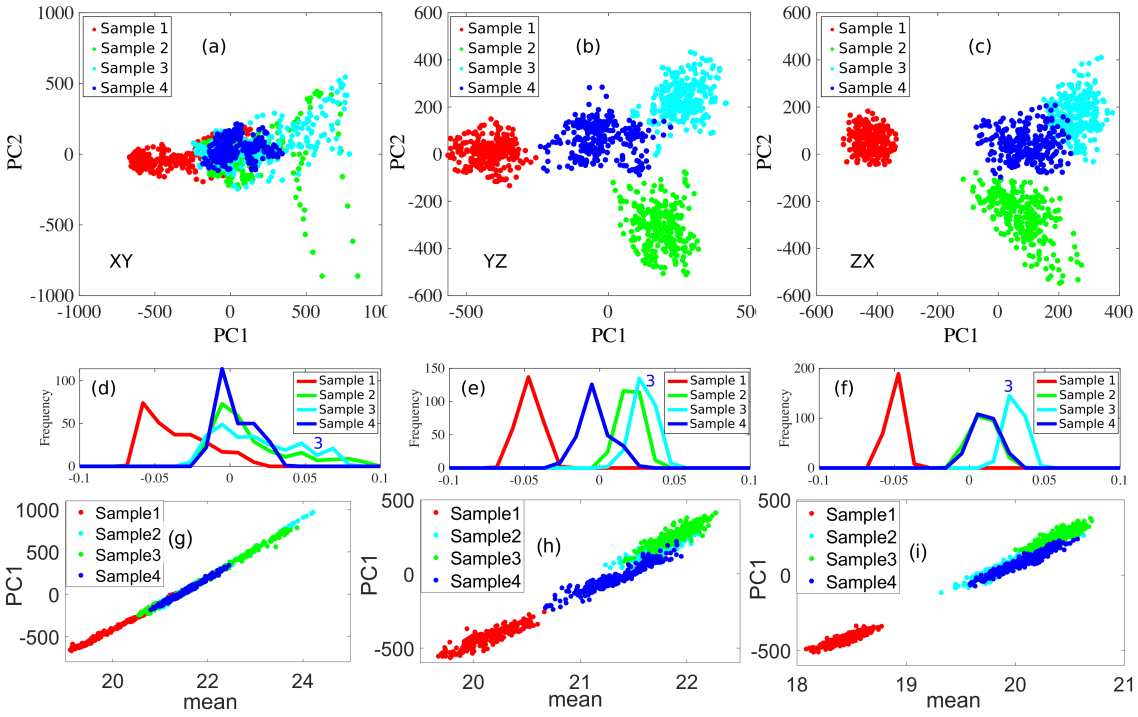


Figure 16: (a-c) Projection of the transformed images in the gradient space onto PC_1 and PC_2 for slices in planes XY, YZ and ZX, respectively. (d-f) Histograms of the distribution of loadings for samples 1, 2, 3 and 4 on the first dominant SVD modes in the gradient eigen-space for slices in planes XY, YZ and ZX, respectively. The loadings come from the column of the V matrix of the SVD. The ordering of the clusters along the first dominant SVD mode suggests that sample 3 contains the largest value of the total variation. Notice that cluster of sample 3 appears on the rightmost location. (g-i) Projection of the image onto PC_1 vs mean of the modulus of gradient in planes XY, YZ and ZX. Projection of the image onto PC_1 is approximately proportional to the mean of the modulus of gradient by observation.

3. Phase field modeling and the effective toughness

Due to the wide range of spatial and temporal scales involved in the characterization of polymeric materials, establishing a direct link between the X-ray image and macroscopic mechanical properties is not a simple task. Before going on, we briefly summarize what we discussed in previous sections.

1. The average dominant frequency of X-ray images is common to all samples.
2. The X-ray images suggest heterogeneous distributions of the Young's modulus. The darker one is stiffer than brighter one.
3. The first principal component PH_1 of the vectorized X-ray image has a negative correlation to the polymerization rate and its inverse analysis suggests that the domain with large density difference grows from sample 1 to sample 4.
4. One of the interesting differences among the four X-ray images is the total variation. Sample 3 has the largest total variation (TV) and sample 2 is the next one.

Recalling the results of mechanical experiments (see Table 1 and Fig. 2), sample 3 shows the best performance, then sample 2 follows. This implies that intermediate level of polymerization rate seems appropriate to show a good performance, which indicates that an appropriate balance of stiff and soft parts is necessary to make tough materials. A remarkable thing is that this order perfectly matches that of total variation (see Fig. 16). Is this just a coincidence or can we justify that the total variation can enhance the material performance.? For this purpose, we set up a macroscopic criterion based on the crack propagation in the framework of the phase field model developed by [28, 29], and then compute the effective toughness (ET) [13] defined later as a toughness indicator.

The crucial step is how we implement the mesoscopic information of X-ray data to the macroscopic model. In view of Eqs. 5-6, two basic parameters the Young's modulus E and the critical energy release rate γ are required in the phase field model. The item 2 above suggests that it is natural to assume that the heterogeneity at mesoscopic level can be introduced to the Young's modulus E and set the critical energy release rate (= fracture toughness of Table 1) γ to be a constant in space. Under this framework, we can see clearly how elastic heterogeneity affects the effective toughness, in fact, we will see that, for the class of heterogeneity with the same dominant frequency, a higher total variation in the elastic modulus will produce a higher effective toughness (Fig. 20, Fig. 22).

3.1. Phase field modeling of fracture

The phase-field method has been established as one of the basic tools for the description of crack propagation [30, 31, 32]. A big advantage of this method is that a whole problem is reduced to solving a well-defined PDEs derived from a free energy rather than tracking free boundaries like cracks and solidification fronts. Here we briefly explain our phase field model for crack propagation (see also [28, 29]) and derive a criterion for materials toughness. A crack does propagate when the energy release rate at the crack front exceeds the surface energy, which is gained during crack growth. We denote the total energy of the system as

$$\mathcal{E} = \int_{\Omega} (1-v)^2 W(\mathbf{x}, \varepsilon_{ij}) + \frac{3\gamma}{8}\left(l|\nabla v|^2 + \frac{v}{l}\right) d\mathbf{x} - \int_{\partial\Omega} \mathbf{T}\cdot\mathbf{u} ds \tag{5}$$

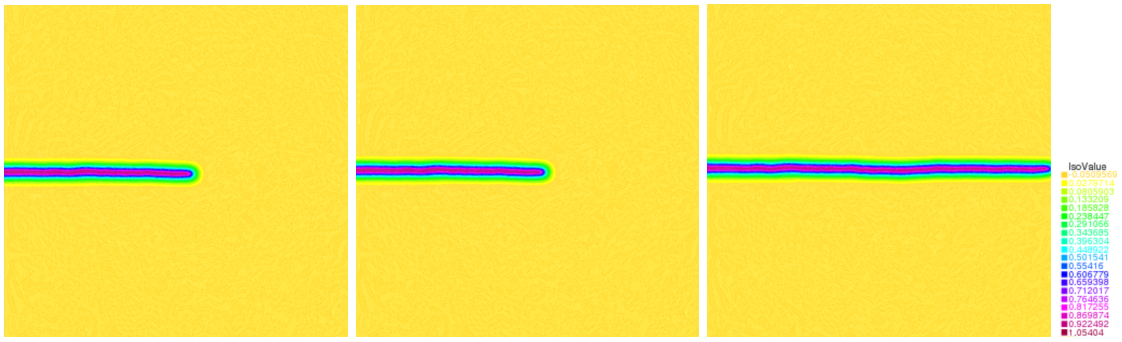
- $\mathbf{u}$ is the displacement vector.
- $\varepsilon_{ij} = \frac{1}{2}\left(u_{i,j} + u_{j,i}\right)$ is the strain tensor component.
- $W(\mathbf{x}, \varepsilon_{ij}) = \frac{1}{2}\sigma_{ij}\varepsilon_{ij}$ is the strain energy.
- σ_{ij} is the stress tensor component and

$$\begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} = \frac{E}{1-\nu^2}\left((1-\nu)\begin{pmatrix} \varepsilon_{11} & \varepsilon_{12} \\ \varepsilon_{21} & \varepsilon_{22} \end{pmatrix} + \nu\mathbf{I}(\varepsilon_{11}+\varepsilon_{22})\right), \tag{6}$$

  where E is the Young's modulus, ν is the Poisson's ratio and $\mathbf{I}$ is the identity matrix.
- v is the phase field: $v = 1$ indicate damaged region and $v = 0$ is the undamaged region.
- $\mathbf{T}$ is the boundary loading force.
- γ is the the critical energy release rate (how much energy is needed for creating new cracks).
- $\lim_{l\to 0} \int_{\Omega} \frac{3}{8}\left(l|\nabla v|^2 + \frac{v}{l}\right)$ represents the crack length.
- Ω is the whole domain and $\partial\Omega$ is its boundary.

The classical brittle fracture theory can be viewed as minimizing energy with respect to displacement field $\mathbf{u}$ and the crack sets $v = 1$ or a competition between energy release rate and the material toughness. Therefore, we obtain the $\mathbf{u}$ and v for each time step by minimizing the total energy subjected to the irreversible constraint $\dot{v} \geq 0$.

$$[\mathbf{u}, v] = \underset{\dot{v} \geq 0}{\operatorname{argmin}} \quad \mathcal{E} \tag{7}$$

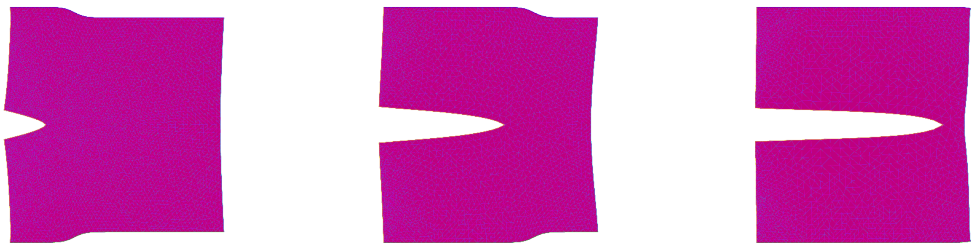


(a) $t = 0$ (b) $t = 0.36$ (c) $t = 0.72$

Figure 17: Profile of phase variable at different time. The value of phase variable v is 0 in the yellow part and 1 in the red part.

At the boundary, we assign, what is called, the surfing boundary condition (see [13]), which can guarantee the propagation of the crack along the x-direction.

$$u(x, y, t) = (0, \quad \frac{A}{2}\left(1 - \tanh(\frac{x - t}{d})\right) \operatorname{sign}(y - 0.5)) \tag{8}$$



(a) $t = 0$ (b) $t = 0.4$ (c) $t = 0.8$

Figure 18: Deformation of material under surfing boundary condition.

3.2. Effective Toughness

We are ready to define a scalar criterion called the effective toughness (ET) to evaluate macroscopic performance for the heterogeneous materials like epoxy resins. Such a quantity is independent

of macroscopic loading and without a prior assumption or restriction about the evolution of the crack set at the microscopic scale. The effective toughness is first proposed by [13].

Definition 3.1. *The effective toughness is defined as*

$$J_{eff} = \max_{0<t<t_b} (J) \tag{9}$$

where t_b is the time when material breaks down completely, namely the crack reaches the other end. J is the celebrated J-integral

$$J = \int_{\Gamma} \left(\mathbf{W} dx_2 - \mathbf{T} \cdot \frac{\partial \mathbf{u}}{\partial x_1} ds \right) \tag{10}$$

where W is the elastic strain energy density, x_1, x_2 are the coordinate directions, $\mathbf{T} = \sigma \mathbf{n}$ is the surface traction vector, $\mathbf{n}$ is the unit normal vector of the curve Γ.

The definition of J_{eff} implies that the applied J has to reach a critical value J_{eff} for the crack to propagate through a macroscopic distance, which is analogous to the maximum force in the peeling problem [33]. We use the surfing boundary condition, to ensure the macroscopic steady crack growth. If the elastic field is homogeneous, J_{eff} is the same as the critical energy release rate.

Note that J-integral is path-independent in homogeneous materials, but no longer path-independent in heterogeneous materials. In fact, we have the following relations.

- $J_{tip} = \gamma$
- $J_{\partial\Omega} = J_{tip} + \int_{\Omega/tip} \frac{\partial^* W}{\partial^* x_1} = \gamma + \int_{\Omega} \frac{\partial^* W}{\partial^* x_1}$.

where J_{tip} means the J-integral around crack tip. Since we are concerned with the macroscopic material performance, we use $\max_t J_{\partial\Omega}$ to represent the effective toughness here. Fig. 19 presents the time-evolution of J-integral for several different types of heterogeneities to show how the geometry affects the value of effective toughness. See supplement material Appendix C for more details.

3.3. The numerical procedure

Numerically, we let the crack set evolve as it chooses and compute the state of strain at each time. The macroscopic energy release rate J at each time is computed over the whole domain. The

domain is normalized to be $[0,1] \times [0,1]$ and the following non-dimensional values are used unless otherwise stated.

$$\gamma = 1, \quad \nu = 0.2, \quad l = 0.02, \quad h_{\min} = 0.005, \quad h_{\max} = 0.01. \tag{11}$$

Here we assume that γ is a constant independent of samples. This might depend on the process of each sample as shown in Table 1, however our results are robust to such alteration. The numerical discretization $h \in [h_{\min}, h_{\max}]$ has to be much smaller than l for convergence. The algorithm to realize the phase filed modeling is as follows:

Algorithm 1: Quasi dynamics of crack propagation

```
Set initial condition;
while i < N do
    Set err=1;
    while err>Tol do
        Build the adapted mesh based on v_i;
        Set a temporary variable v̂_i = v_i;
        u_i = arg min_{u_i ∈ W_2} E(u_i, v̂_i) ;
        v_i = arg min_{v_i ∈ W, v̂_i ≤ v_i ≤ 1} E(u_i, v_i);
        err = ||v_i − v̂_i||_{L^∞};
    end
    v_{i+1} = v_i;
    i ← i + 1;
    Update boundary condition;
end
```

Since the functional $\mathcal{E}$ is not convex in the pair $(\mathbf{u}, v)$, it is impractical to minimize the functional directly. We performed an alternate minimization algorithm employed in the paper [34]. The minimization of the functional with respect to $\mathbf{u}_i$ and v_i is performed by solving the corresponding Euler-Lagrange equations for $\mathbf{u}_i$, namely, a standard linear elliptic problem, whereas for v_i we employ the function "IPOPT" in Freefem [35], based on the interior point method. This allows us to explicitly enforce the irreversibility constraint (fracture can not recover, $\frac{\partial v}{\partial t} > 0$).

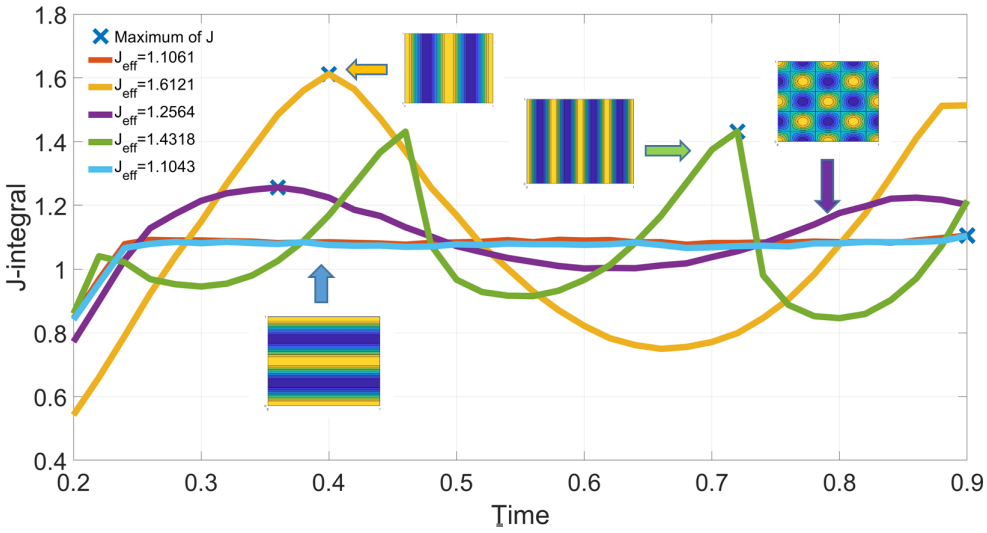


Figure 19: J integrals under different setting of inhomogeneities. In homogeneous material (see the red curve) crack start to propagate when J reaches a certain constant, which is exactly the critical energy release rate. We compare four different elastic heterogeneities with the homogeneous case, they are $E = 2(1 + 0.25\cos(4\pi x))$, $E = 2(1 + 0.25\cos(8\pi x))$, $E = 2(1 + 0.25\cos(4\pi x)\cos(4\pi y)$, and $E = 2(1 + 0.25\cos(4\pi y))$ respectively. Note that ET coincides with that of the homogeneous case when the pattern of heterogeneity is parallel to the crack propagation (see light blue case). Higher frequency oscillations along the direction of crack propagation result in lower effective toughness.

3.4. Mesoscopic inhomogeneity and macroscopic toughness

We are ready to see how effective toughness (ET) changes with different heterogeneities obtained from X-ray images. To gain some intuitive ideas of effective toughness, we start off a few simple but suggestive examples.

- **Periodic heterogeneity with different frequencies**: We take the elastic modulus as $E = 2(1+0.25\cos(a\pi x)\cos(b\pi x))$ and change the frequency by changing the parameter a and b. The simulation, see Fig. 19, reveals that higher frequency oscillations along the direction of crack propagation result in lower ET.

- **Periodic heterogeneity with different amplitudes**: We take the elastic modulus as $E = 2(1 - a\cos 4\pi x)$ and change the amplitude by changing the parameter a. Fig. 20 shows that higher amplitude oscillations along the direction of crack propagation yield higher ET.

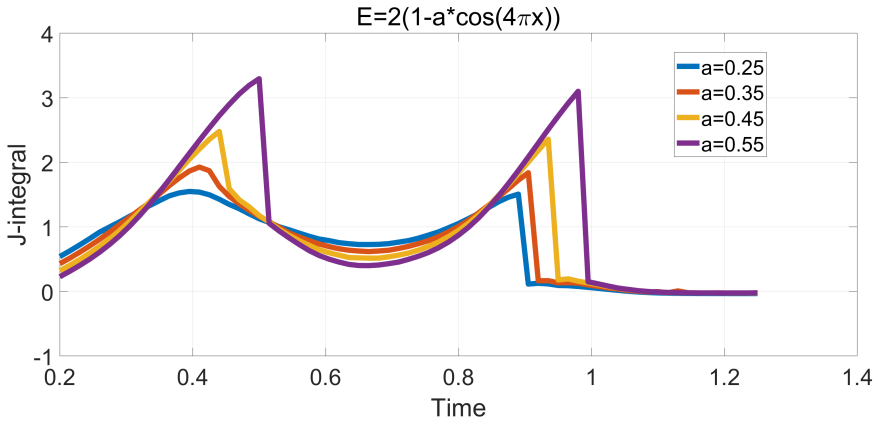
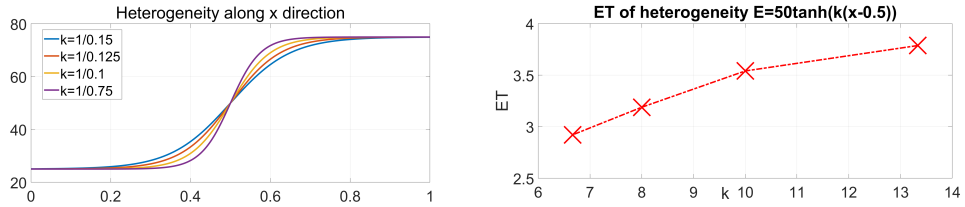


Figure 20: J integrals under the same shape of heterogeneity but with different amplitudes. Higher amplitude oscillations along the direction of crack propagation lead to higher effective toughness. In these cases, amplitude is proportional to the total variation.



(a) Heterogeneity along x direction

(b) Effective toughness v.s slope

Figure 21: J integrals under the same shape of heterogeneity but with different slopes. Higher slope along the direction of crack propagation results in higher effective toughness. In these cases, their total variations are the same.

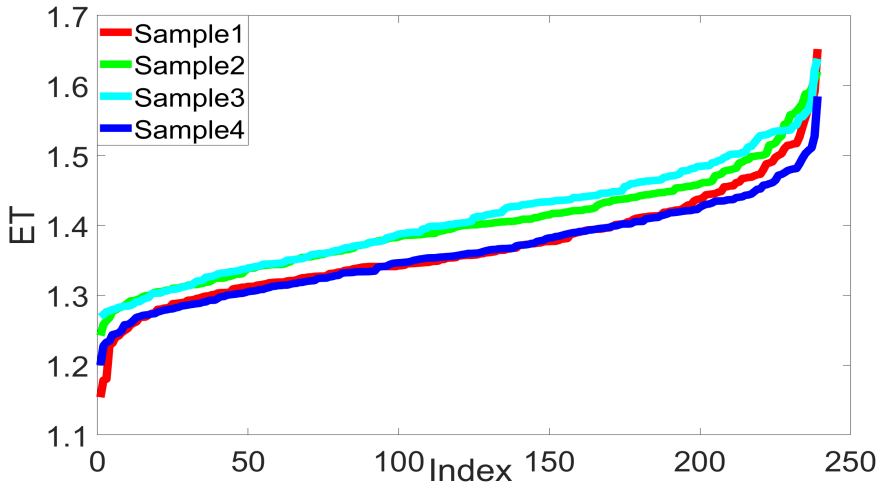
- **Heterogeneity with different steepness**: We take the elastic modulus as $E = 50(1 + 0.5\tanh(k(x - 0.5)))$ and change the slope by changing the parameter k. Fig. 21 shows that ET increases as k increases.

One property of the J integral worthy to note is that it increases as the strength of heterogeneity grows (Fig. 20). This property motivates us to connect ET to the feature "total variation" (TV) observed in image analysis (Fig. 16). We emphasize here that such a connection between ET and the total variation holds only under the condition that the samples have the same dominant frequency. This is because higher frequency yields larger TV when the amplitude is fixed, however ET does not

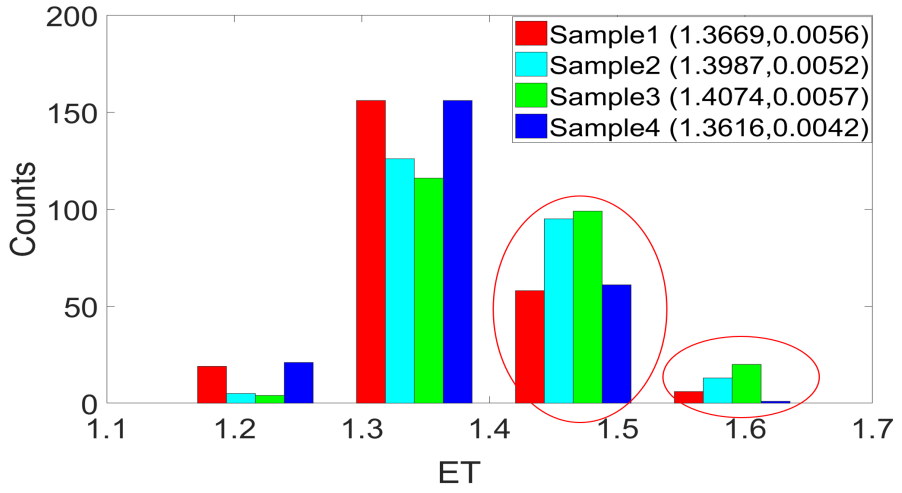
increase monotonically with TV as is shown in Fig. 19. On the other hand, Fig. 20 demonstrates in its simplest form that the total variation contributes to enhance the material performance for a fixed frequency. In our case for the epoxy resin samples, the image analysis shows that they have similar dominant frequencies (Fig. 7). Also note that Fig. 19 shows that steeper the slope along the crack direction, the larger the ET values with keeping the same TV. Since our samples have different total variations, we proceed to compute the ETs and check the consistency with the experimental results. Recall that our framework is that mesoscopic heterogeneity of X-ray data is inherited to that of the Young's modulus so that we define it in the following way. The Young's modulus is given by the gray scale distribution of X-ray image as follows:

$$E = 1 + 10(1 - F) \tag{12}$$

where $F = \frac{\text{gray value}}{255}$ is the normalized grey value. This means that darker parts have larger Young's modulus than brighter ones based on the item 2 at the beginning of this section. The coefficient 10 is chosen to enhance the difference among the samples. Fig. 22 shows the comparison among four samples. The index of Fig. 22(a) represents relabeling of the sample slices as the increasing order of ET. The ET values of samples 2 and 3 lie above the other two samples, in fact sample 3 shows slightly better performance than sample 2 in terms of the average value of ET, which is consistent with the experimental data of s-s curve in Fig. 2. Here we take the critical energy release rate γ to be equal to 1.0 and common to all samples to see the effect coming from the Young's modulus. Of course, one can take γ to be dependent on the sample as listed in Table 1, however the above order of ET values does not change, in fact the difference among them become even larger for that case.



(a) Ordered effective toughness vs. index



(b) Statistical summary

Figure 22: Statistical summary of effective toughness of four samples of epoxy. The upper figure shows the reordered effective toughness, where the index is reordered in a increasing way of ET. We can see that the toughness curves of samples 2 and 3 lie above those of samples 1 and 4, and that sample 3 has the highest ET on average. The bottom figure shows the histogram of the toughness. The number after the label is mean value and standard deviation respectively. From the red circle part, we can easily see that samples 3 contain the largest amount of high toughness samples.

4. Conclusions and perspectives

There are two big challenges for evaluating the performance of amorphous materials such as epoxy resins: one is the characterization of micro/mesoscopic structure depending on the curing process, and the other is how it influences the macroscopic performance such as fracture toughness. The latter problem, i.e., connecting two different scales, remains one of the central issues in materials science. We used the data of X-ray computed tomography (CT) for epoxy resins with four different curing processes as shown in Fig. 1. Those samples are almost undistinguishable to the naked eye, but it is expected that they inherit key information coming from the inhomogeneous structure depending on the curing process. As for the classification, we applied two statistical methods: one is a well-known SVD method, and the other is persistent homology (PH). In particular, the latter topological approach tells us about the generalized hole structure that is not captured by SVD method. Both methods allowed us to classify the four samples successfully by using the appropriate PCA-coordinate as in section 2. There are a couple of remarks here: firstly it is not a priori clear to identify the background image originating from the fluctuation of X-ray intensity shared by these samples. This is important because what we want to know is the inhomogeneous structure inherent to the curing process after removal of the background noise. This was done in section 2 and we classified four samples by using SVD and PH respectively. The structural information is important, however, our goal is to extract a mathematically well-defined feature characterizing four different samples and how it affects the material's performance. This is much more difficult and is still on the halfway, however, one interesting fact we found is that the total variation (TV) of the inhomogeneity is a key factor that increases the performance of materials, and the computations in section 3 support the fact that is consistent with the experimental results. To show this, we resorted to the phase field approach as in section 3. Introducing the phase variable indicating the level of damage, the crack propagation can be solved efficiently, because it has no more free boundary. The main issue is how we implement the mesoscopic information of heterogeneity to the macroscopic phase field model. Since all of the four types of epoxy resins are built using blocks of the same monomer with different polymerization rates, it is reasonable to assume that each sample has a homogeneous local toughness, but the different spatial distribution of elasticity (Young's modulus) reflecting their curing processes. Then, such information is transferred to the inhomogeneity of elastic modulus in the phase field model. In this context, we numerically study the crack propagation in materials and compute their effective toughness (ET), which is a scalar

indicator for the crack to reach the other end. Our results agree with the experimental observations, and confirm that the larger the total variation, the larger the ET value. This suggests that the above linkage between mesoscopic information and macroscopic performance is a reasonable one.

There remain many problems. For instance, as for the classification, PH approach turns out to be useful, however, we still don't know how to interpret the difference of heterogeneities geometrically. In other words, what is a good geometrical coordinate which is acceptable also from the physical viewpoint and allows us to identify the difference of inhomogeneity intuitively? We have made one step forward toward this direction in section 2.3 (see also 4 items at the beginning of section 3), however, to pursue it, we have to solve the inverse problem systematically for the associated persistent diagrams. Also such a characterization may not be just a local property but related to a long range 3D network of amorphous structure. It is possible to construct 3D network structure from 2D slices, however, there is still a gap between 2D and 3D, in fact, the optimization process shown in section 3.3 becomes quite expensive in 3D-case even in the framework of phase field modeling so that more efficient algorithm is needed to reach a statistical conclusion. Finally we assume that the mesoscopic inhomogeneity of X-ray data is inherited to the variation of the Young's modulus E, and that the critical energy release rate γ is a constant, but may depend on the curing process. Loosely speaking, this is equivalent to the hypothesis that E contains finer information than γ for the class of amorphous materials like epoxy resins. This is still an arguable issue and needs a systematic comparison among various mathematical models with different resolutions. See also the reference [6, 7, 8, 9].

As a conclusion, we proposed a new framework to utilize X-ray CT images to predict the mechanical toughness of epoxy resins by analyzing mesoscopic variations of the structural inhomogeneities in the samples that are related to the curing process. Such a connection has not been studied so far to the best of our knowledge. Once we could find topologically more efficient geometric coordinates reflecting 3D network of amorphous structure, and simultaneously reduce the effect coming from the "background image", this study would provide a better understanding of the relation between the micro-structure of epoxy resins observed by X-ray images and the material performance.

Appendix A. SVD calculation

The SVD consists of two parts. (i) First, diagonalize the matrix $X^T X$ in Eq A.1. This results in a set of eigenvectors, V, and eigenvalues, $\Sigma^T \Sigma$. (ii) The next step consists in using the result

from (i) to solve the Eq A.2 for U.

$$X^T X = V \Sigma^T \Sigma V^T \tag{A.1}$$

$$XV = U\Sigma \tag{A.2}$$

This operation allows to construct an eigen-slices space, which is a space spanned by eigen-slices.

Appendix B. Classification of samples

In section 2.2 we shown that the intensity of the X-ray images form four distinct clusters in an eigen-space spanned by PC_1 and PC_2 and this is illustrated in Fig.4c. In this section we compute decision boundaries for the clusters of samples 1 to 4. Such boundaries are necessary to predict the character of new X-ray images landing on the domain. To construct the decision boundaries we consider the projection of the clusters onto PC_1 and PC_2, as these components are related to two important features of the X-ray images, namely, the mean intensity and variance of the mean.

Firstly, we use k-means which is an unsupervised classification scheme consisting on iteratively (i) labeling the elements and (ii) computing the center of mass of the clusters [19, 36]. Secondly, we use a k-nearest neighbors scheme with $k = 17$ as supervised method to find the labels of the elements in each cluster. In each case we split the data using 80% for training and %20 for test. We perform a 10-fold cross-validation and randomize each realization. In all cases the misclassification error is less than 1%. Fig. B.23 shows the result of both methods of classification. Both methods are able to produce suitable decision boundaries.

Appendix C. J-integral

The potential energy of an elastic body (the total enthalpy of the system) is given by

$$\mathcal{P} = \int_\Omega \mathbf{W} d\mathbf{x} - \int_{\partial\Omega} \mathbf{t} \cdot \mathbf{u} ds \tag{C.1}$$

Assume the change of total enthalpy changes $\delta\mathcal{P}$ when the crack tip moves by $\delta\xi_i$. The driving force on the singularity is defined as

$$f_i = -\frac{\delta\mathcal{P}}{\delta\xi_i} \tag{C.2}$$

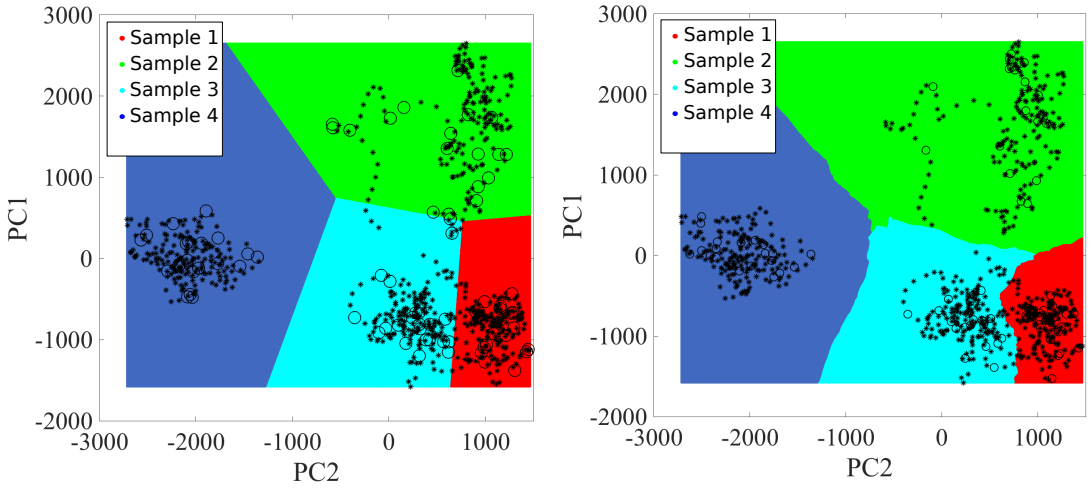


Figure B.23: Different methods of classification. (a) k-means consists of the iterative process of (i) labeling and (ii) computing the center of mass (unsupervised). These two steps were iterated seven times until convergence. (b) For the k-nearest neighbors we use $k = 17$. The process consists of finding the distances of a test set to the k-nearest neighbors and labeling it accordingly (supervised). The solid circles are the training set (80%) and the open circles are the test set (20%).

J-integral is path dependent in the heterogeneous material. Consider the J-integral taking two different integration paths Γ_1 and Γ_2. Without loss of generality, we assume that Γ_2 is inside Γ_1. The difference between the different integration paths is

$$\begin{aligned} J(\Gamma_1) - J(\Gamma_2) &= \int_{\Gamma_1 - \Gamma_2} \left(\mathbf{W} dx_2 - \mathbf{t} \cdot \frac{\partial \mathbf{u}}{\partial x_1} ds \right) \\ &= \int_{A(\Gamma_1,\Gamma_2)} \frac{\partial W}{\partial x_1} - \left(\sigma_{ij} \frac{\partial u_i}{\partial x_1} \right)_j d\mathbf{x} \\ &= \int_{A(\Gamma_1,\Gamma_2)} \frac{\partial^* W}{\partial^* x_1} + \sigma_{ij} \frac{\partial u_i}{\partial x_j \partial x_1} - \left(\sigma_{ij} \frac{\partial u_i}{\partial x_1} \right)_j d\mathbf{x} \\ &= \int_{A(\Gamma_1,\Gamma_2)} \frac{\partial^* W}{\partial^* x_1} - (\sigma_{ij})_j \frac{\partial u_i}{\partial x_1} \\ &= \int_{A(\Gamma_1,\Gamma_2)} \frac{\partial^* W}{\partial^* x_1} \end{aligned} \tag{C.3}$$

where $A(\Gamma_1, \Gamma_2)$is the area between the contours Γ_1 and Γ_2, ε_{ij} and $\sigma_{ij} = \frac{\partial W}{\partial \varepsilon_{ij}}$ are the strain and stress tensor components respectively, $\frac{\partial^* W}{\partial^* x_1}$ is the explicit derivative due to x_1. Note that we apply

the equilibrium balance equation $(\sigma_{ij})_j = 0$ in the last step.

- From Eq (C.3), we see that if the material between two contours is homogeneous, the J-integral evaluate on these two contours have the same value. Therefore, J-integral is path independent in homogeneous materials.

- When the material is heterogeneous in elasticity, J-integral is path dependent. We define J_{tip} and J_{domain} as the J-integral along crack tip and the whole domain, then $J_{tip} = G_c$ and $J_{domain} = J_{tip} + \int_\omega \frac{\partial^* W}{\partial^* x_1}$. Therefore, effective toughness we have defined reflect local fracture toughness and global stress distribution of material.

Acknowledgment. This work was partially supported by the Council for Science, Technology and Innovation (CSTI), Cross-ministerial Strategic Innovation Promotion Program (SIP), 'Structural Materials for Innovation' and 'Materials Integration' for Revolutionary Design System of Structural Materials. Also Y.N. gratefully acknowledges the support of KAKENHI Grants-in-Aid no.18H05482 and the Fusion Research and Advanced Target Project of WPIAIMR, Tohoku University. Special thanks go to Professor Takeshi Takaishi (Musashino University) for providing us a code of the phase field model.

References

[1] C. A. May, Resins for Aerospace, Vol. 285 of ACS Symposium Series, American Chemical Society, 1985, Ch. 24, pp. 557–580, 0. doi:10.1021/bk-1985-0285.ch024.
URL https://doi.org/10.1021/bk-1985-0285.ch024

[2] M. Aoki, A. Shundo, R. Kuwahara, S. S. Yamamoto, K. Tanaka, Mesoscopic heterogeneity in the curing process of an epoxy-amine system, Macromolecules 52 (2019) 2075–2082.

[3] S. Torquato, Random heterogeneous materials: microstructure and macroscopic properties, Vol. 16, Springer Science & Business Media, (2013).

[4] N. Tian, R. Ning, J. Kong, Self-toughening of epoxy resin through controlling topology of cross-linked networks, Polymer 99 (2016) 376–385.

[5] Z. J. Thompson, M. A. Hillmyer, J. Liu, H.-J. Sue, M. Dettloff, F. S. Bates, Block copolymer toughened epoxy: role of cross-link density, Macromolecules 42 (7) (2009) 2333–2335.

[6] H. Kishi, T. Naitou, S. Matsuda, A. Murakami, Y. Muraji, Y. Nakagawa, Mechanical properties and inhomogeneous nanostructures of dicyandiamide-cured epoxy resins, Journal of Polymer Science Part B-Polymer Physics 45 (2007) 1425–1434.

[7] A. Haris, T. Adachi, W. Araki, Nano-scale characterization of fracture surfaces of blended epoxy resins related to fracture properties, Materials Science and Engineering a-Structural Materials Properties Microstructure and Processing 496 (2008) 337–344.

[8] C. M. Sahagun, S. E. Morgan, Thermal control of nanostructure and molecular network development in epoxy-amine thermosets, Acs Applied Materials & Interfaces 4 (2012) 564–572.

[9] S. Morsch, Y. Liu, S. Lyon, S. Gibbon, Insights into epoxy network nanostructural heterogeneity using afm-ir, Acs Applied Materials & Interfaces 8 (2016) 959–966.

[10] H. Edelsbrunner, J. Harer, Persistent homology-a survey, Contemporary mathematics 453 (2008) 257–282.

[11] N. Otter, M. A. Porter, U. Tillmann, P. Grindrod, H. A. Harrington, A roadmap for the computation of persistent homology, EPJ Data Science 6 (1) (2017) 17.

[12] Y. Hiraoka, T. Nakamura, A. Hirata, E. Escolar, G, K. Matsue, Y. Nishiura, Hierarchical structures of amorphous solids characterized by persistent homology, PNAS 113 (23) (2016) 7035–7040.

[13] M. Hossain, C.-J. Hsueh, B. Bourdin, K. Bhattacharya, Effective toughness of heterogeneous media, Journal of the Mechanics and Physics of Solids 71 (2014) 15–32.

[14] J. Schmidt, M. R. G. Marques, S. Botti, M. A. L. Marques, Recent advances and applications of machine learning in solid-state materials science, npj Computational Materials 5 (1) (2019) 83. doi:10.1038/s41524-019-0221-0.
URL https://doi.org/10.1038/s41524-019-0221-0

[15] J. Wei, X. Chu, X.-Y. Sun, K. Xu, H.-X. Deng, J. Chen, Z. Wei, M. Lei, Machine learning in materials science, InfoMat 1 (3) (2019) 338–358. doi:10.1002/inf2.12028.
URL https://doi.org/10.1002/inf2.12028

[16] K. Rajan, Materials informatics, Materials Today 8 (10) (2005) 38–45. doi:10.1016/s1369-7021(05)71123-8.
URL https://doi.org/10.1016/s1369-7021(05)71123-8

[17] A. Furmanchuk, A. Agrawal, A. Choudhary, Predictive analytics for crystalline materials: bulk modulus, RSC Adv. 6 (2016) 95246–95251. doi:10.1039/C6RA19284J.
URL http://dx.doi.org/10.1039/C6RA19284J

[18] A. Ziletti, D. Kumar, M. Scheffler, L. M. Ghiringhelli, Insightful classification of crystal structures using deep learning, Nature Communications 9 (1) (2018) 2775. doi:10.1038/s41467-018-05169-6.
URL https://doi.org/10.1038/s41467-018-05169-6

[19] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning, Springer Series in Statistics, Springer New York Inc., New York, NY, USA, (2001).

[20] J. N. Kutz, Data-Driven Modeling & Scientific Computation: Methods for Complex Systems & Big Data, Oxford University Press, Inc., New York, NY, USA, (2013).

[21] O. Alter, P. O. Brown, D. Botstein, Singular value decomposition for genome-wide expression data processing and modeling, Proceedings of the National Academy of Sciences of the United States of America 97 (18) (2000) 10101–10106, 10963673[pmid]. doi:10.1073/pnas.97.18.10101.
URL https://www.ncbi.nlm.nih.gov/pubmed/10963673

[22] M. Kang, J.-M. Kim, Singular value decomposition based feature extraction approaches for classifying faults of induction motors, Mechanical Systems and Signal Processing 41 (1) (2013) 348 – 356. doi:https://doi.org/10.1016/j.ymssp.2013.08.002.
URL http://www.sciencedirect.com/science/article/pii/S0888327013003762

[23] F. Fioranelli, M. Ritchie, H. Griffiths, Classification of unarmed/armed personnel using the netrad multistatic radar for micro-doppler and singular value decomposition features, IEEE Geoscience and Remote Sensing Letters 12 (9) (2015) 1933–1937. doi:10.1109/LGRS.2015.2439393.

[24] Sternberg, Biomedical image processing, Computer 16 (1) (1983) 22–34. doi:10.1109/MC.1983.1654163.

[25] R. C. Gonzalez, R. E. Woods, Digital Image Processing (3rd Edition), Prentice-Hall, Inc., Upper Saddle River, NJ, USA, (2006).

[26] I. Obayashi, HomCloud (version 2.8.0) (2019).
URL http://www.wpi-aimr.tohoku.ac.jp/hiraoka_labo/homcloud/index.en.html

[27] C. G. Harris, M. Stephens, et al., A combined corner and edge detector., in: Alvey vision conference, Vol. 15, Citeseer, (1988), pp. 10–5244.

[28] T. Takaishi, M. Kimura, Phase field model for mode iii crack growth in two dimensional elasticity, Kybernetika 45 (4) (2009) 605–614.

[29] M. Kimura, T. Takaishi, Phase field models for crack propagation, Theoretical and Applied Mechanics Japan 59 (2011) 85–90.

[30] M. J. Borden, C. V. Verhoosel, M. A. Scott, T. J. Hughes, C. M. Landis, A phase-field description of dynamic brittle fracture, Computer Methods in Applied Mechanics and Engineering 217 (2012) 77–95.

[31] C. Miehe, L.-M. Schaenzel, H. Ulmer, Phase field modeling of fracture in multi-physics problems. part i. balance of crack surface and failure criteria for brittle crack propagation in thermoelastic solids, Computer Methods in Applied Mechanics and Engineering 294 (2015) 449–485.

[32] M. Ambati, T. Gerasimov, L. De Lorenzis, A review on phase-field models of brittle fracture and a new fast hybrid formulation, Computational Mechanics 55 (2) (2015) 383–405.

[33] S. Xia, L. Ponson, G. Ravichandran, K. Bhattacharya, Toughening and asymmetry in peeling of heterogeneous adhesives, Physical review letters 108 (19) (2012) 196101.

[34] B. Bourdin, G. A. Francfort, J.-J. Marigo, The variational approach to fracture, Journal of elasticity 91 (1-3) (2008) 5–148.

[35] F. Hecht, New development in freefem++, J. Numer. Math. 20 (3-4) (2012) 251–265.

[36] M. Sugiyama, Introduction to Statistical Machine Learning, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, (2016).