Thermal Conductivity Analysis of Polymer-Derived Nanocomposite via Image-Based Structure Reconstruction, Computational Homogenization, and Machine Learning

Mozhdeh Fathidoost,* Yangyiwei Yang,* Nathalie Thor, Jan Bernauer, Astrid Pundt, Ralf Riedel, and Bai-Xiang Xu*

Macroscopic thermal properties of engineered or inherent composites depend substantially on the composite structure and the interface characteristics. While it is acknowledged that unveiling such dependency relation is essential for materials design, the complexity involved in, e.g., microstructure representation and limited data impedes the research progress. Herein, this issue is tackled by machine learning techniques on image-based microstructure and property data predicted from physics simulations, along with experimental validation. The methodology is demonstrated for the model system $(Hf_{0.7}Ta_{0.3})C/SiC$ ultrahigh-temperature ceramic nanocomposite. The structure is reconstructed from scanning electron microscope images, and is resolved by a diffuse-interface representation, which is advantageous in handling complicated structure and interface properties. Subsequently, hierarchical finite element homogenization is carried out to evaluate the effective thermal conductivity. A thorough comparison between the computed results and experimentally measured data, conducted across diverse temperatures and varying interface thermal resistances, reveals a high level of agreement. The observed agreement allows for the inverse estimation of the interface thermal resistance, a parameter typically challenging to ascertain directly through experimental means. Utilizing comprehensive data, a machine learning surrogate model has been meticulously trained to accurately predict the effective thermal conductivity of composite structures with exceptional performance.

1. Introduction

The macroscopic thermal properties of inherent or engineered composites are strongly dependent on the composite structure and interface characteristics. Recognizing the importance of unraveling this dependency for materials design, we acknowledge the hindrances posed by complexities such as challenges in microstructure representation and limited data availability that impede research progress. $(Hf_{0.7}Ta_{0.3})C/SiC$ is a model system of an inherent nanocomposite (NC) ultra-high-temperature ceramic (UHTC), where the thermal conductivity depends on the composite structure, especially the interfaces or interphases.

UHTCs with melting points exceeding 3000 °C have attracted significant interest for their potential application as protective coatings on metal substrates in extreme environments.[1–3] They are well suited for use in thermal protection structures for gas turbine engines, chemical plants, or aerospace applications, where they can withstand temperatures above approximately 1600 °C and harsh environmental conditions.[1–4] UHTCs are primarily binary compounds composed of elements such as boron, carbon, or nitrogen in conjunction with early transition metals such as Zr, Hf, Ti, Nb,

M. Fathidoost, Y. Yang, B.-X. Xu
Mechanics of Functional Materials Division
Technical University Darmstadt
Institute of Materials Science
Otto-Berndt-Straße 3, 64287 Darmstadt, Germany
E-mail: mozhdeh.fathidoost@tu-darmstadt.de;
yangyiwei.yang@mfm.tu-darmstadt.de; xu@mfm.tu-darmstadt.de

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/adem.202302021.

N. Thor
Institute of Applied Geoscience
Technical University Darmstadt
Schnittspahnstraße 9, D-64287 Darmstadt, Germany

J. Bernauer, R. Riedel
Department of Dispersive Solids
Technical University Darmstadt
Institute of Materials Science
Otto-Berndt-Straße 3, 64287 Darmstadt, Germany

A. Pundt
Insitute for Applied Materials - Materials Science and Engineering (IAM)-WK
Karlsruhe Institute of Technology
Engelbert-Arnold-Str. 4, 76131 Karlsruhe, Germany

or Ta.[1–3,5] UHTCs have a distinctive mixture of metallic and covalent bonds, resulting in exceptional properties, including high refractoriness, extremely high melting points, high hardness, high elastic moduli, and outstanding chemical stability. Additionally, their metallic bonding grants UHTCs enhanced thermal and electrical conductivity when compared to oxide ceramics.[1,2,4,6,7]

Polymer-derived ceramic (PDC) technology is an emerging and promising approach for producing ceramics such as UHTCs. It involves the production of ceramics from preceramic polymer precursors that undergo cross-linking and pyrolysis. This method enables the fabrication of high-performance nonoxide ceramics, including a variety of binary, ternary, and quaternary materials, in various forms. PDC technology offers advantages over conventional ceramic synthesis techniques, such as sol–gel and solid-state reactions, in terms of energy efficiency and formability.[8–11]

The effective thermal conductivity (k_{eff}) of composite materials is influenced by several factors, including the properties of the individual components, microstructures, and interfaces.[12–14] Modeling and simulation methods are valuable for understanding these effects, especially when experiments are costly or challenging. Simulation data can complement experimental results for composite design and guide the direction of the material microstructure design. In addition, they can help to gain a deeper understanding of the characteristics of heterogeneous materials.[12,15–17]

Computational thermal homogenization using the finite element method (FEM) is widely used to numerically determine the thermal conductivity of composites.[18,19] Sharp-interface and diffuse-interface approaches are used to describe complex microstructures, while the latter one allows more flexibility and consideration of thermal resistance and thermal flow effects at the interfaces.[20–22]

In this investigation, the novel diffuse-interface thermal homogenization model is applied to the real microstructure of the $(Hf, Ta)C/SiC$ material system obtained from the scanning electron microscope (SEM) image. To address the issue of varying microstructure scales across different regions, a hierarchical homogenization scheme is planned to be implemented across different scales. Notably, in this material system the microstructure in one region is significantly finer compared to the other region. Subsequently, we intend to validate the calculated effective thermal conductivity by comparing it with experimentally measured values at various temperatures and different normalized interface thermal resistances (R_s) within the material.

As part of the methodological investigation, it is intended to construct a surrogate model based on machine learning (ML). This model aims to receive the microstructural features of a sample and predict its effective thermal conductivity as an output. There are numerous techniques for representing microstructures that can be incorporated into ML models to predict properties and develop the structure–property relationship in composites. These techniques are generally categorized into three primary groups: descriptor-based, spatial correlation-based, and image-based.[23] In the case of descriptor-based approaches, the challenge is to identify appropriate descriptors that have a significant impact on material properties. The use of image-based descriptors and convolutional neural networks is computationally expensive. Consequently, in recent years, numerous studies have shifted their focus to the use of statistical features

such as the n-point correlation function to represent microstructures.[24,25] The n-point correlation functions capture the probability of specific n-point arrangements occurring within the microstructure.[26] As the higher order correlations are computationally expensive, the two-point statistics of the microstructures are fed into the model as representative microstructural input.

There are several analytical and numerical methods for predicting the effective thermal conductivity of composites. Analytical approaches such as the Maxwell[27] and Bruggeman models[28] are straightforward but do not consider the intricate details of material distribution within composites. There are also some other analytical homogenization methods for random heterogeneous materials, such as the contrast expansion methods. As these methods can directly derive the effective thermal conductivity tensor via provided morphological information in the form of correlation functions (such as the n-point correlation mentioned by the reviewer) without considering the interface thermal resistance, they can be readily integrated into our proposed two-level homogenization scheme by replacing the diffuse-interface-based computational thermal homogenization.[29,30] However, the interface thermal resistance is a crucial factor to be considered for the Si(Hf,Ta)C material system in this study, especially for the accurate characterization of complex microstructures using the diffuse-interface model.

The comprehensive approach in this work includes both the application of a thermal homogenization model and the development of a ML-based structure–property relationship. This allows a thorough investigation of the thermal properties of the $(Hf, Ta)C/SiC$ material system.

2. Material System

The material system investigated in this study is a SiC-based PDC NC (PDC-NC) containing the two refractory metals hafnium (Hf) and tantalum (Ta). The polymer-to-ceramic route was used to prepare a single-source precursor with a tailored Hf : Ta ratio of 0.7 : 0.3. Subsequent sintering via the field-assisted sintering (FAST) technique resulted in an UHTC NC (UHTC-NC), namely, $(Hf_{0.7}Ta_{0.3})C/SiC$. Further information regarding the synthesis route and a detailed description of the microstructure development upon high-temperature sintering can be found in ref. [2]. The mentioned UHTC-NC is intended to be used as a top coat on an intermetallic alloy substrate, to improve the thermal and corrosion protection properties of the compound material. Therefore, it should have low k_{eff}.

The microstructure of the monolithic $(Hf, Ta)C/SiC$ ceramic following FAST sintering was examined using SEM, as illustrated in **Figure 1** at different magnifications.

The FAST sintering process resulted in a crystalline ceramic material, as confirmed by XRD analysis.[31] The analysis of polished cross sections revealed three distinct microstructural regions: the former powder particles (FPPs), sinter necks (SNs) formed during the sintering process, and residual porosities within the ceramic. The microstructure consists mainly of FPPs and SNs. These two regions are labeled and outlined in light-blue in Figure 1b. Different types of porosity were observed, including small intragranular pores within the FPPs and porosity at the interface between the FPPs and the SNs. The dark regions
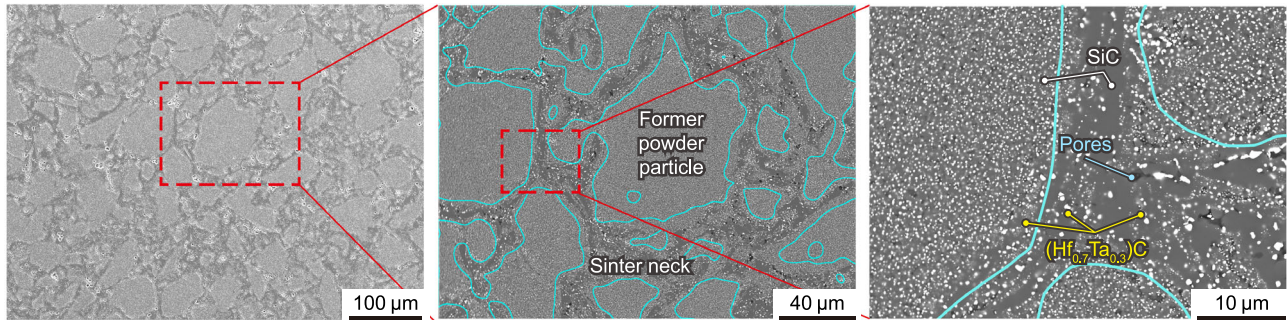
Figure 1. The different magnifications of the SEM images reveal significant variations in the morphological scales of different phases within the material system, suggesting a two-step hierarchical homogenization process.

in the microstructures represent pores, while the bright regions correspond to (Hf,Ta)C. The medium gray regions denote SiC, as shown in Figure 1c. Comparison of FPPs and SNs microstructures reveals a significant contrast in their scales, with FPPs microstructures being significantly finer than SNs microstructures.

To deal with this scale difference between the regions (as shown in Figure 1), a two-step homogenization process should be implemented in the calculations. In this study, the two-step homogenization is characterized as a two-step hierarchical thermal homogenization approach capable of calculating the effective thermal conductivity (k_{eff}) of the PDC-NC. In the first step, the nano- and microscale thermal homogenization in both FPPs and SNs is calculated separately. In the second step, the k_{eff} values obtained in the first step for FPPs and SNs are used as the reference thermal conductivity for these regions in the mesoscale computational thermal homogenization. In summary, the homogenization model is able to 1) address the nanosized microstructures within the FPPs, 2) analyze the microsized microstructures within the SNs; and 3) consider the mesosized microstructures including both regions.

3. Theory and Method

3.1. FE-Based Computational Thermal Homogenization

The study by Yang et al. verified the diffuse-interface microstructure with anisotropic interface thermal conductivity using FE-based computational thermal homogenization models.[20] The binarized microstructure images (see **Figure 2**) served as initial conditions for the variable ϕ, which denotes the phase of the microstructure (e.g., $\phi = 1$ for particles; $\phi = 0$ for the matrix). In particular, the profile of ϕ across the interface must exhibit central symmetry around its inflection point, which is the expected position of the corresponding sharp interface. In this study, the Allen–Cahn equation was used to numerically construct the necessary diffuse interface as follows:

$$\frac{\partial \phi}{\partial t} = \nabla^2 \phi - \frac{4\alpha^2}{l^2} \phi(1 - 3\phi + 2\phi^2) \qquad (1)$$

where l denotes the diffuse-interface width which is adjusted by α parameter.[20,32] Here, we set α to 2.94 for all microstructures to replicate the setup of our previous work. Each microstructure was generated using Equation (1) for 1–3 time steps with a time interval of 0.1 unit to achieve a diffuse interface without changing the morphology, as shown in the mesh image in **Figure 3**. The continuity of the thermal flux along both the normal and tangential directions of the surface is considered via phase-dependent thermal conductivity tensor, as formulated in refs. [20,33].

$$\mathbf{k}(\phi) = [h_{(i)}k_{(i)} + h_{(m)}k_{(m)}]\mathbf{T} + \left[\frac{k_{(i)}k_{(m)}}{h_{(i)}k_{(m)} + h_{(m)}k_{(i)} + R_s|\nabla\phi|} \right]\mathbf{N} \qquad (2)$$
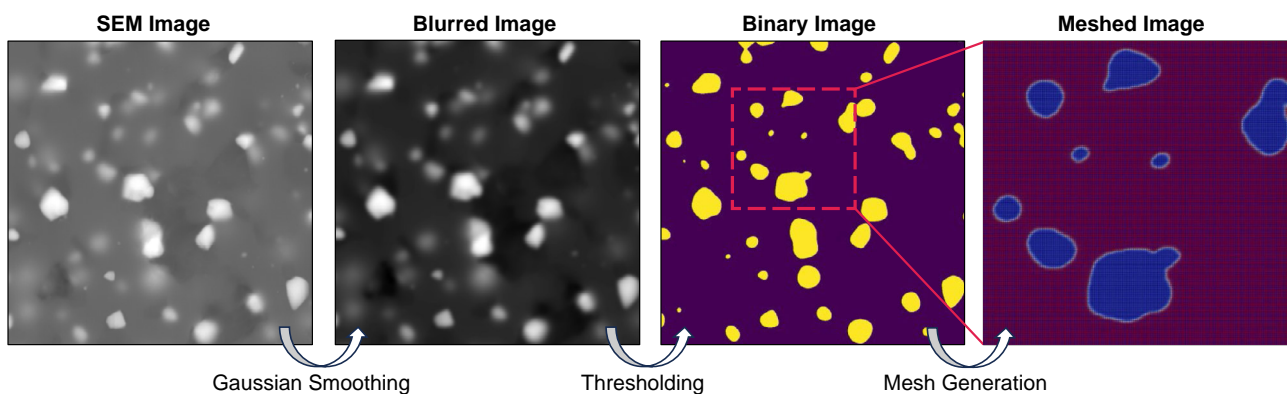

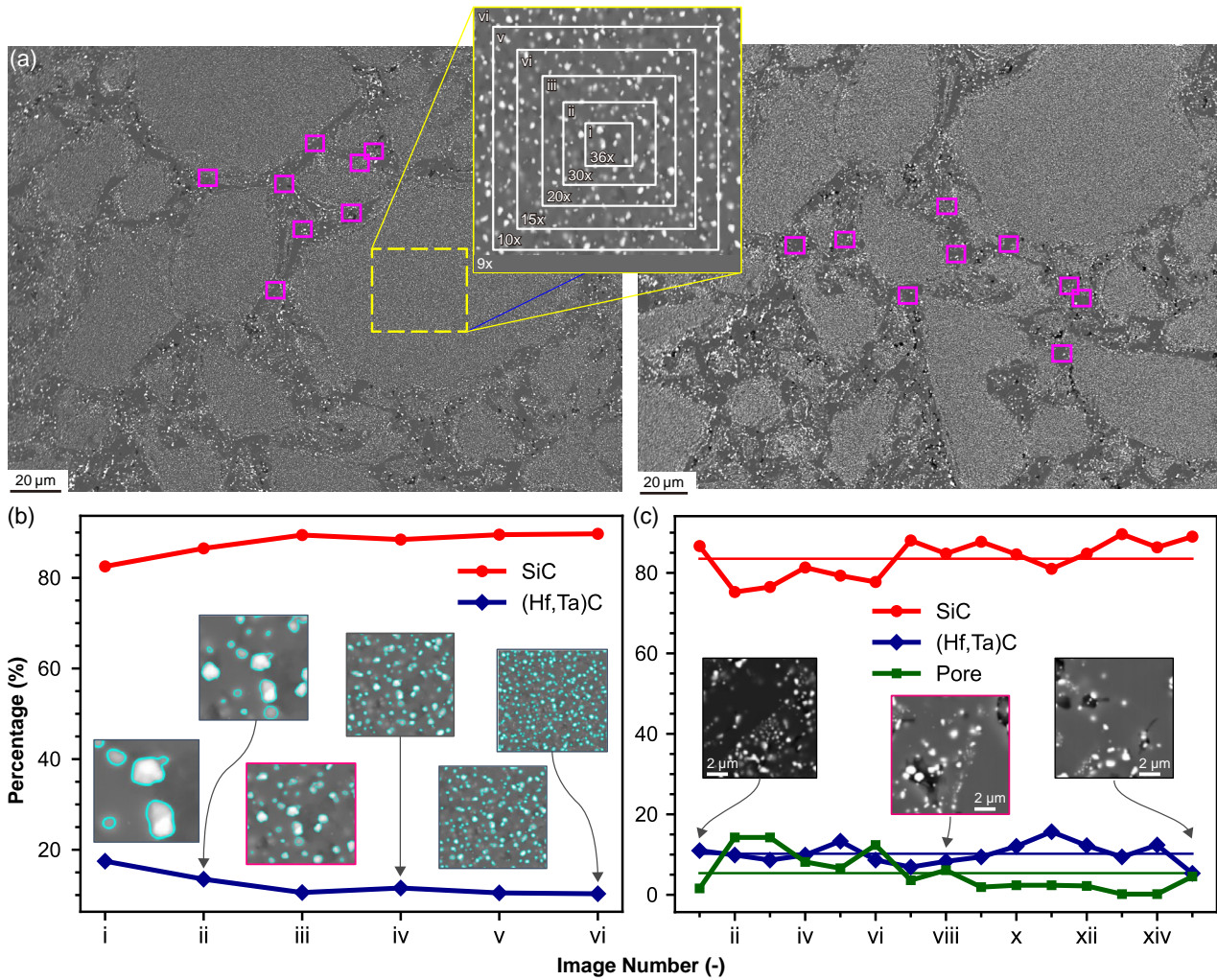**Figure 2.** Image processing from SEM image to binary and mesh images.

Figure 3. a) Selected ROIs within the different phases, i.e., former powder particles and sinter necks. The representative SEM images for b) former powder particles and c) sinter necks phase are highlighted by pink boxes.

where $k_{(i)}$ and $k_{(m)}$ are the thermal conductivities of the inclusion (i) and matrix (m), respectively. $h_{(i)}$ and $h_{(m)}$ are correspondingly the monotonic interpolation functions. **T** and **N** represent the tangential and normal tensors of the surface, respectively. R_s denotes the interface thermal resistance between the phases.

The diffuse-interface approach successfully addresses numerical and geometric challenges associated with the analysis of complex microstructures and interpolating the properties of the interface. In this research, this diffuse-interface-based computational thermal homogenization model is conducted to evaluate the k_{eff} of the complex microstructure of the sintered $(Hf_{0.7}Ta_{0.3})C/SiC$ PDC in a hierarchical manner. Because of the two-level hierarchical thermal homogenization used here, the terms $k_{(i)}$ and $k_{(m)}$ represent various phases at different scales. For instance, in the thermal homogenization within the FPPs, $k_{(m)}$ and $k_{(i)}$ refer to the thermal conductivity of SiC and (Hf,Ta)C, respectively. The governing equations included Fourier's law and the conservation law for temperature and thermal flux. The thermal homogenization problem was defined for both microscale and macroscale regions as below:

$$\begin{cases} \mathbf{j} = \mathbf{k}(\phi) \cdot \nabla T & \text{on microscale} \\ \langle \mathbf{j} \rangle = \mathbf{k}_{eff} \cdot \langle \nabla T \rangle & \text{on macroscale} \end{cases} \tag{3}$$

where T represents the temperature flux, $\mathbf{j}$ denotes the thermal flux, and $\mathbf{k}$ signifies the specified local conductivity. Here, $\mathbf{k}_{eff}$ represents the homogenized effective thermal conductivity tensor, with an effective value of $k_{eff} = \text{tr}(\mathbf{k}_{eff})/3$, which is evaluated in this investigation; $\langle \cdot \rangle = \int_\Omega (\cdot)d\Omega/\Omega$ is the mean operator.

The Hill–Mandel condition was used to ensure equality of thermal dissipation at both scales. A linear temperature boundary condition was used to satisfy this condition, and the mean temperature gradient and $\mathbf{k}_{eff}$ tensor components were determined based on calculated thermal flux components under applied gradients. More details on the homogenization formulation can be found in refs. [20,34]. The authors used numerical models implemented by the custom simulator NIsoS, which was developed based on the open-source finite element framework MOOSE.[35,36]

3.2. Two-Point Spatial Correlations

To study the microstructure of a material, which refers to its microscopic internal arrangement and characteristics, the challenge is to identify a reduced-form feature that effectively represents the complex microstructure. In other words, it should retain all essential structural information to serve as a substitute for detailed microstructure images. Various types of features, such as physical and statistical features, are employed for this purpose. Examples include the volume fraction percentage of the phases, as well as the size and aspect ratio of the particles.[23,37] A broader form of representation involves employing statistical correlation functions. The n-point correlation, along with its simplified variations, is the most widely used of various types. Due to the significant computational cost associated with the n-point correlation, the two-point correlation functions are often employed as an alternative. The two-point correlation function, often denoted as $S_2(r)$, is a mathematical function used in the analysis of spatial patterns such as microstructure, which describes the structural features and topological arrangements of a material at a particular length scale.

The two-point correlation function $S_2(r)$ is derived by repeatedly placing a randomly selected line segment of length (r) within the sample and noting how often its endpoints fall within a particular phase of the material system. By conducting this procedure for various lengths (r), a graph of S_2 versus (r) can be constructed. Consequently, $S_2(r)$ represents the probability that the endpoints of a line segment of length (r) are in a particular phase of the microstructure. It is obvious that fluctuations in $S_2(r)$ indicate the degree of correlation between the two points within the system.[26] $S_2(r)$ provides statistical information about the correlation between points in a system or distribution as a function of their separation distance (r). In summary, it gives the probability of finding two different points at a certain distance (r) from each other in the same phase, where this probability varies as the distance changes. Namely, it helps to quantify the correlation between different points at different locations in heterogeneous materials.[23,26,37–40] The two-point correlation function was computed for all images within the dataset to assess the variability in the microstructure images using the PoreSpy Python toolkit.[41] Further details regarding the application of the two-point correlation function in our study can be found in Section 5.1 and Figure 6b.

To establish the ML-based structure–property relation in our study, it was necessary to provide the ML model with microstructure features as input. For this aim, the two-point statistics[42] are employed. The two-point statistics characterize the spatial correlations between distinct local states within the internal structure of a material similar to the two-point correlation function. However, these correlations are determined by randomly inserting vectors of varying sizes and orientations into the material microstructure. This means that the orientation of the vectors will be taken into account in the correlation as well. Recent studies[25,42] have shown that the full set of two-point correlations contains all the necessary information to accurately reconstruct a representative microstructure, with the exception of translation and/or inversion, as demonstrated in previous research. The two-point statistics for a material system can be defined as[42]

$$f^{nn'} = \frac{1}{S} \sum_{s=0}^{S-1} m_s^n m_{s+t}^{n'} \tag{4}$$

where m_s^n is our digital microstructure array, n enumerates the local state, and s indicates the bin index. The subscripts t and S enumerate all vectors that occur randomly and all bins in the domain, respectively. More information and the mathematical relations can be found in refs. [25,43,44]. The PyMKS[43] framework was also employed for calculation of the two-point statistics of all the microstructure image and two examples of the two-point statistics on the microstructure are shown in Figure 7. The two-point statistics of a 2D image share the same dimension, resulting in significant information redundancy. To improve computational efficiency, dimensionality reduction techniques, such as principal component analysis (PCA), should be applied to minimize this redundancy.

3.3. PCA

PCA is a statistical technique used to reduce the dimensionality of data into its key components, known as principal components. These principal components are combinations of the original variables and are orthogonal to each other. The combinations are arranged to capture the maximum variance across all variables. Through this process, PCA provides an approximation of the original data table and reduces dimensionality by emphasizing only the crucial components.[45,46]

As discussed earlier, the two-point statistics provide a valuable representation of microstructure, but result in a significantly large feature space that may contain unnecessary information redundancies. Dimensionality reduction techniques, specifically PCA, can then be applied to generate low-dimensional microstructure descriptors from sets of spatial correlations. The reduced two-point correlation function obtained from PCA can be efficiently used as a ML input, reducing the time and computational cost of model training by using lower dimensional data.[43,47]

4. Image-Based FE Simulation

4.1. Identifying the Representative Area Element for Different Microstructure Scales

To implement FEM directly on a SEM image as the input geometry, several image processing steps are required to streamline the image. These processing steps are outlined in Figure 2. First, a Gaussian smoothing filter is applied to the SEM image, which reduces intricate details and results in a more opaque image. Next, thresholding is used to segment the image, resulting in a binary image. All of these steps are done using scikit-image package in Python.[48] The binary image is then utilized to generate a finite element mesh with an initial sharp interface. Interface relaxation is then conducted using the Allen–Cahn equation to obtain a diffuse-interface microstructure. More details on the diffuse-interface generation can be found in refs. [20,34].

To initiate the finite element simulation by hierarchical computational homogenization, it is crucial to establish the

representative images for all three scales discussed in Section 2. The workflow for exploring the representative region of interest (ROI) at different scales is illustrated in Figure 3.

Due to the limited number of variations in the mesoscale images of the sample, the ROI selection process focuses on identifying representative images specifically for FPPs and SNs. Considering that the microstructure within the FPPs is finer than that of SNs, representative testing involves selecting images of a single particle but at different image sizes (magnifications). These selected images are outlined by yellow boxes in Figure 3a. For SNs with a coarser microstructure, ROIs of identical size are selected from different locations. These selected regions are highlighted with pink boxes in Figure 3a.

Afterward, the selected ROI images undergo the same image processing steps discussed earlier to be converted into binary images. Subsequently, the proportions of each phase in these ROIs are calculated and marked on each image using the scikit-image package[48] in Python. The phase fractions versus image size for FPPs are then plotted in Figure 3b. It can be seen that the phase fractions stabilize after the image *iii*, indicating that the image size does not affect the phase fraction for the images larger than this limit. For SNs, the phase percentages for various images, all of the same size, are displayed, and the average values for different phases, including SiC, $(Hf_{0.7}Ta_{0.3})C$, and pores, are plotted in red, blue, and green solid lines, respectively, in Figure 3c. It can be seen that among all the images, the outlined image with the pink box is the closest to the average phase values. Therefore, it is selected as the representative image for SNs.

4.2. Validation of FE Results through Experimental Measurements

In this work, all detected pores in the microstructure are smeared out and merely the composite of FPPs and SNs is considered in the thermal homogenization, assuming the perfectly synthesized microstructure. Nonetheless, we recognize the potential effect of pores (blocking the heat flux) in reducing the measured thermal conductivity of the composite, which has a similar effect to increasing the interface thermal resistance. In other words, the value of the R_s already recapitulates the influence of pores. Initially, the results of the computational thermal homogenization of the material system are demonstrated on three different scales: the nanoscale within the FPPs, the microscale in the SNs, and the mesoscale through the combination of the two aforementioned regions, using the aforementioned diffuse-interface model. The results are shown visually in **Figure 4**. In this figure, the colored regions indicate the amount of normalized thermal flux passing through the microstructure, and the vertical lines indicate the normalized temperature in the microstructure. All homogenization processes were conducted at the temperature of 1600 °C, which is the expected operating temperature in high-temperature applications such as gas turbine engines. The thermal conductivity values for the constituent phases, i.e., SiC and (Hf,Ta)C, were obtained through an extensive literature review, and the corresponding plots are illustrated in **Figure 5**a,b.

The simulation results need to be validated by experimental measurements. Information on the equipment and methodology used for experimental thermal conductivity measurements can be found in the study by Bernauer et al.[49] It is important to note that the validation simulation is performed using the aforementioned two-step homogenization approach described in Section 2 and 3.1. In order to simulate the thermal conductivity using the specified model, certain input parameters are required, namely, the thermal conductivity of the constituent phases and the geometry or microstructure of the material system at different scales. The former comes from the literature, specifically the thermal conductivity of different phases at different temperatures. This information is presented as plots of thermal conductivity versus temperature for SiC and (Hf,Ta)C in Figure 5a,b. The latter, which is the ROI at different scales, is explained in Figure 3. The only unknown input pertains to the R_s between the phases. In nanoscale (FPPs) and microscale (SNs) microstructures, the R_s between SiC and (Ta,Hf)C is taken into account. However, no R_s is considered between FPPs and SNs at the mesoscale. In fact, FPPs and SNs are treated as separate regions rather than distinct phases. Therefore, incorporating R_s in this context would have no physical meaning.
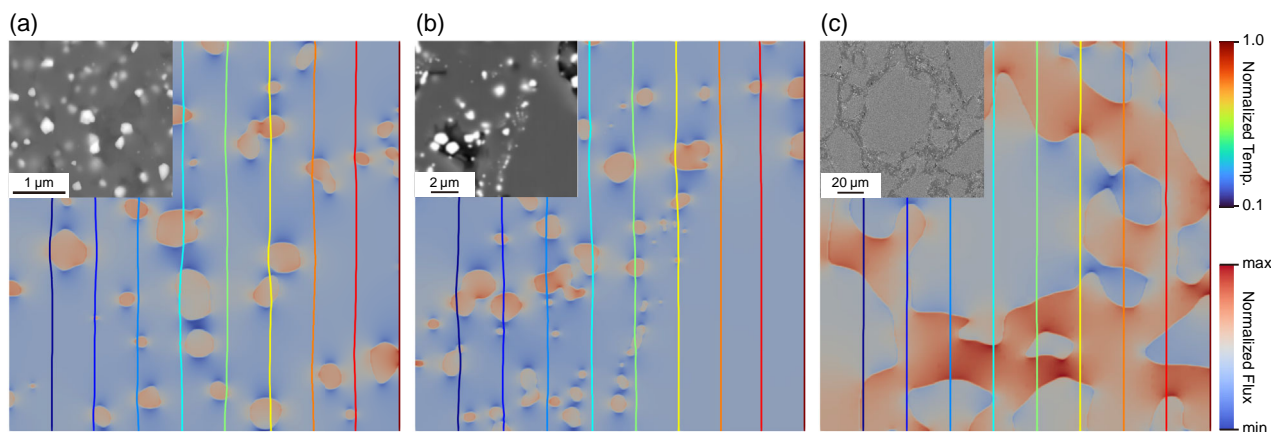

Figure 4. In the hierarchical computational thermal homogenization process, the effective thermal conductivities calculated for different regions at the lower scale, i.e., a) former powder particle at the nanoscale and b) sintered neck at the microscale, serve as the reference thermal conductivities of the corresponding regions in c) mesoscale microstructure.
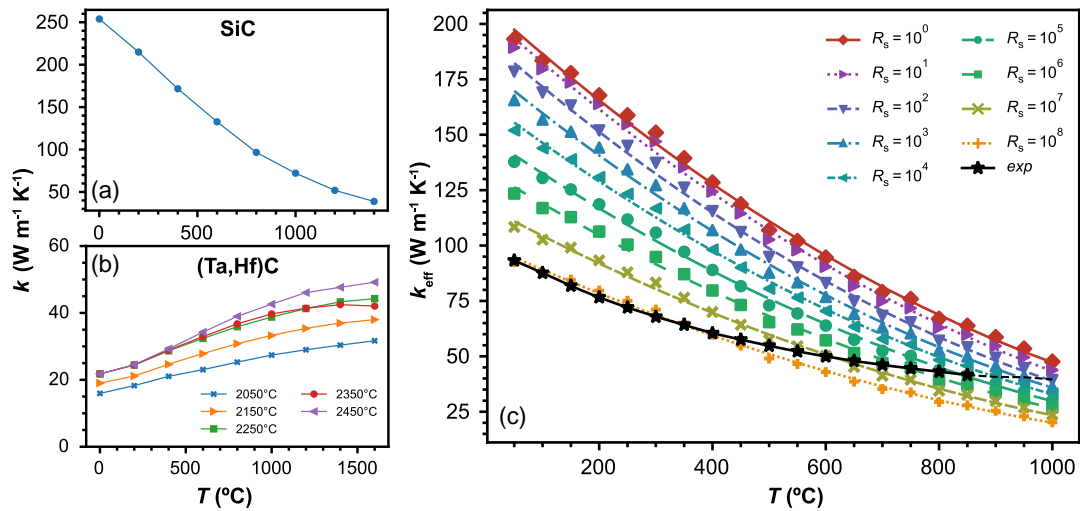
Figure 5. a) Thermal conductivity of SiC redrawn from ref. [54]. b) Thermal conductivity of (Hf,Ta)C redrawn from ref. [55]. c) Validation of the experimentally measured thermal conductivity with the computationally calculated thermal homogenization results obtained at various temperatures and interface thermal resistances.

As the exact value of R_s is unknown in the system, simulations are conducted for different values of R_s ranging from 10^{-1} to 10^8. The summarized results are presented in Figure 5c. The solid black line represents the experimental thermal conductivity results at various temperatures. The distinct markers in different colors depict the simulated results obtained at different values of R_s. Additionally, the various styles in matching colors illustrate the fitted lines corresponding to the simulation results of the same R_s.

Observations indicate that for temperatures below 400 °C, the experimental measurements agree with the simulation results, where the value of R_s is 10^8. As the temperature increases, the experimental results tend toward smaller R_s values. At the maximum temperature of 850 °C for which experimental data

are available, the experimental and simulation results coincide at R_s equal to 10^5. Extrapolation of the experimental line (represented by the dashed black line extending from the continuous black line) suggests that at higher temperatures, around 1000 °C, the R_s is expected to be about 10^2 for the agreement between experimental and simulation data.

This leads to the conclusion that the interface thermal resistance between phases has a significant and dominant influence on the thermal conductivity of the composite, especially at lower temperatures. In summary, R_s tends to be high at temperatures lower than 500 °C due to increased phonon scattering and thermal flux deviation at the interface. This behavior results from imperfect interfaces and phase mismatch on either side of the interface.[50]
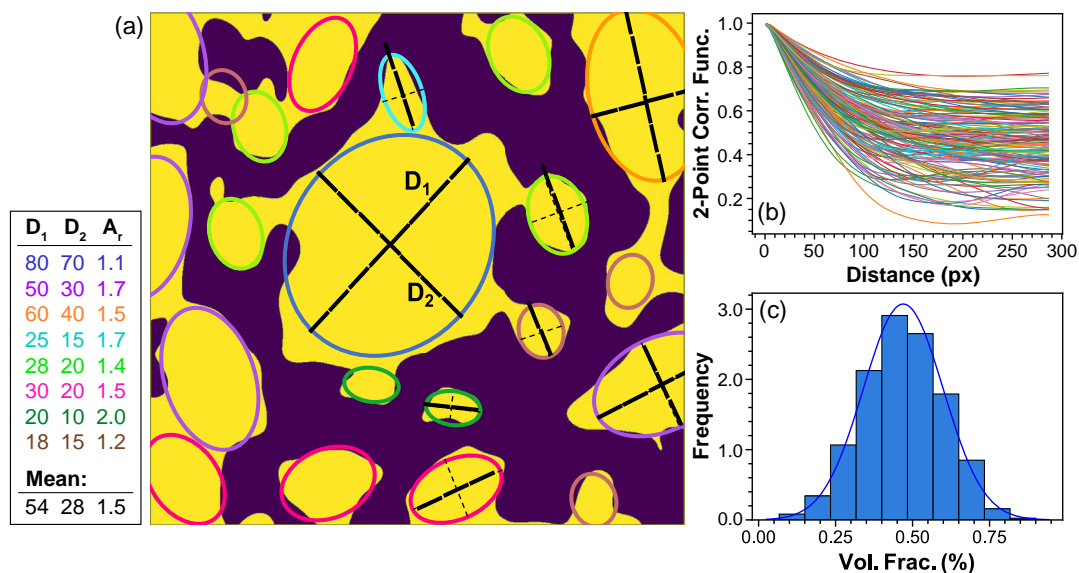

Figure 6. a) Analysis of the mesoscale microstructure to obtain the parameterized descriptor, i.e., the aspect ratio and the minor and major diameters. The diversity and data distribution in terms of the two-point correlation function, i.e., $(S_2(r))$, and volume fraction are shown in (b,c), respectively.

5. Data-Driven Surrogate Model for Thermal Homogenization

The development of a ML-based surrogate model for predicting the properties of composite materials, trained on data obtained from physical models, provides a powerful and efficient approach to the design of new materials.[51,52] It accelerates the material design and optimization process, allowing researchers and engineers to explore a wider range of composite formulations and configurations. It facilitates the identification of key factors that influence material properties, contributing to a deeper understanding of the structure–property relationship in composites.

In this study, a surrogate model is developed to predict the thermal conductivity of composites with high accuracy using data-driven approaches, eliminating the need for complex and computationally expensive simulations or experiments. In the model, the input is the two-point correlation of the (Hf, Ta)C/SiC microstructures, and the output is the corresponding thermal conductivity. The following section details the data generation process used to train and develop the model.

5.1. Data Generation

To construct a surrogate model capable of predicting the thermal conductivity of diverse (Hf, Ta)C/SiC microstructures, a variety of different microstructures are essential. As it is impractical to obtain these different microstructures experimentally, synthetic
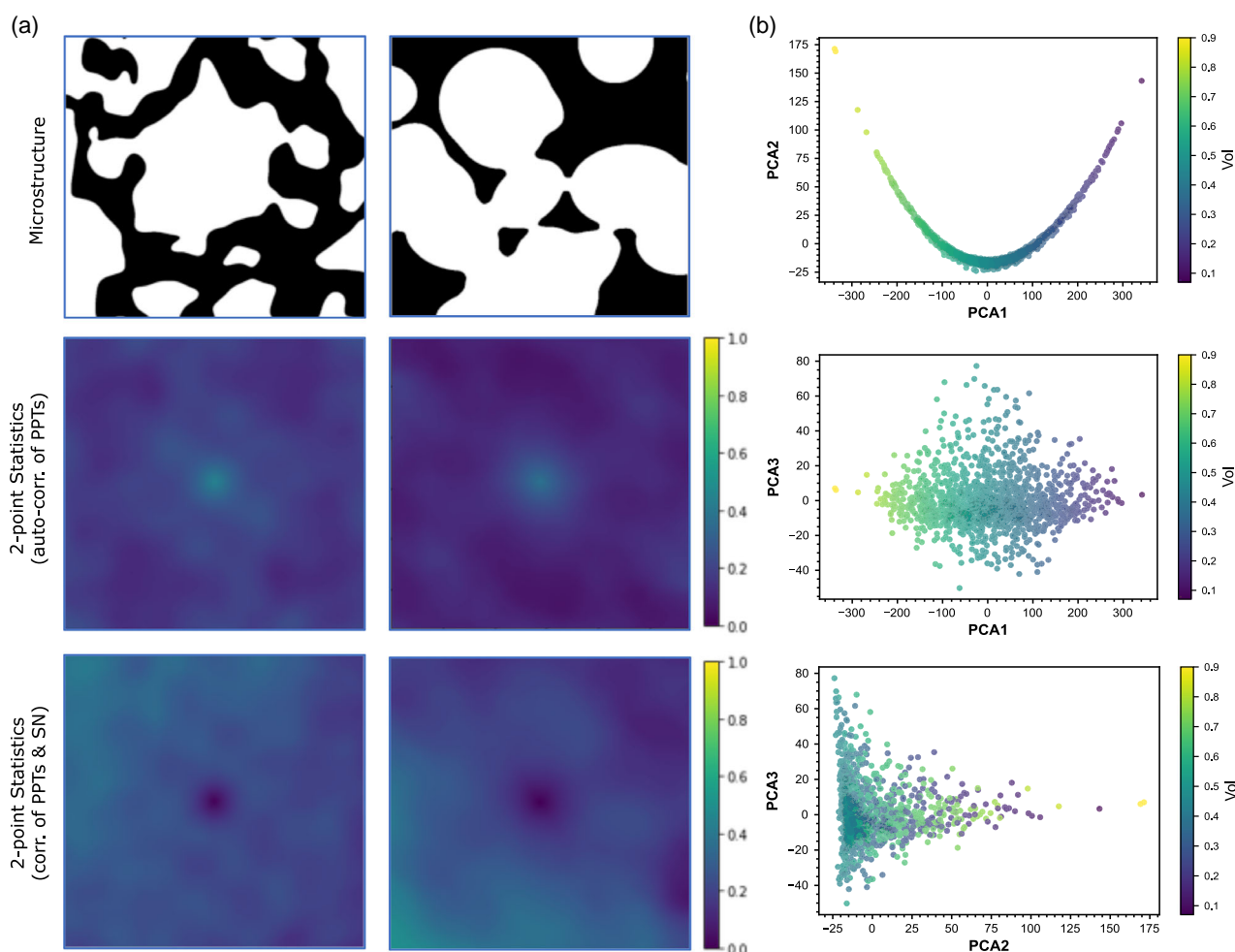
Table 1. Parameterized descriptors extracted from the mesoscale SEM image analysis.

Param.	Mean[a] [unit]	Std.
D_1	40, 50, 60 [μm]	20%
D_2	20, 25, 26, 30, 33, 40 [μm]	30%
A_r	1.5, 2	–
V_f	35, 40, 45, 55, 60 [%]	–

[a]1440 images in total.

Figure 7. a) Visualization of the two-point statistics on two microstructure examples shown in the top row. The second and third rows show the auto-correlation for the bright phase (FPPs) and the correlation for the dark and bright phases, respectively. b) The correlation of the first three principal components (PC1, PC2, and PC3) obtained from the PCA on the entire dataset.

data are generated. The homogenization scheme involved a hierarchical model with three different length scales in the material system. In this study, the surrogate model focuses specifically on the mesoscale microstructure. Consequently, the data generation process only considers the mesoscale microstructure of the material system as well, which includes two distinct regions of SNs and FPPs.

In the mesoscale microstructure, we consider the SNs as the matrix and the FPPs as the second phase. Key features such as volume fraction and size of the second phase are essential for synthetic image generation. These parameters are obtained by manually modeling the FPPs phase using ellipsoids of different sizes and shapes, as illustrated in **Figure 6**a.

In this binary image, the bright and dark regions correspond to the FPPs and SNs, respectively. The FPPs are represented by ellipsoids of different sizes, with a distinct color. The major and minor axes of these ellipsoids are labeled as D_1 and D_2, respectively. The aspect ratios of the ellipsoids are defined as $A_r = D_1/D_2$. On the left side of the figure, the color-coded values of D_1, D_2, and A_r, corresponding to the ellipsoid colors, along with their mean value are also reported. The final microstructure parameters are determined based on the ellipsoid model analysis on the SEM image and are presented in **Table 1**. By using these parameters as microstructure features, it is possible to generate synthetic microstructures that closely resemble the SEM images. For this task, the commercial software GeoDict 2023 (Math2Market, GmbH, Kaiserslautern, Germany) is employed.

The combination of parameters, as described in Table 1, is used for microstructure generation. As GeoDict generates 3D

morphology, cross-sections are extracted for analysis. A total of 180 3D microstructures were generated, from which 1440 cross-sectional images were extracted for the input dataset. To assess the diversity of the input microstructures, the distribution of the dataset is visualized through the two-point correlation functions, i.e., $S_2(r)$ and volume fraction histograms, as illustrated in Figure 6b,c, respectively. The calculation of the two-point correlation function is performed on all images by poreSpy,[41] a Python package designed for the analysis of microstructure images.

To prepare the images as input for ML, two crucial steps are performed. First, the two-point statistics (based on Equation (4)) of all the images are extracted and stored. These statistics provide a detailed representation of spatial relationships, but they remain as 2D data, posing computational challenges for ML training. To address this, PCA is used to reduce the dimensionality of the data by transforming it into uncorrelated variables known as principal components (see Section 3.3). This simplification streamlines the learning process and improves the efficiency of the ML model. In this study, only the first three principal components of PCA are used, as they showed promising performance in our ML training. Examples to illustrate the two-point statistics of two microstructures are shown in **Figure 7**a. There, the two-point statistics are represented in both the autocorrelation between the FPPs and also the correlation between PPTs and SNs. In addition, the relationship between the principal components in our datasets is also shown in Figure 7b. The colors of the points correspond to the volume fraction of the second phase in the artificially generated microstructures. In the first two plots,
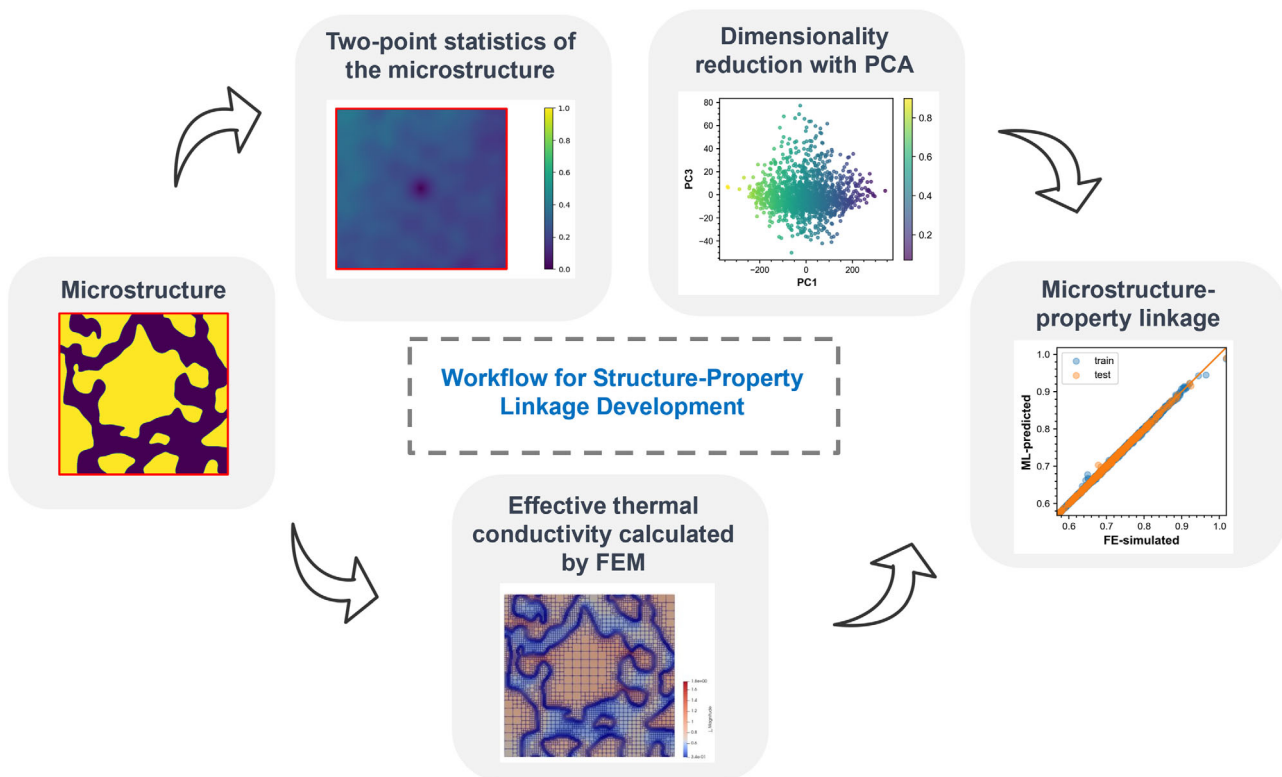

Figure 8. ML-based structure–property relation development workflow, adopted by ref. [25].

which show the correlation of the first principal component with the second and third ones, the points representing microstructures are arranged according to the volume fraction of the second phase. As a result, the ML input is transformed from a large, pixel-wise image to just three principal components obtained from the two-point statistics of the microstructures. The two-point statistics are derived from PyMKS,[43] a Python implementation of methods developed within the Materials Knowledge System framework.

The output of our ML model is the predicted effective thermal conductivity. To generate the output data for training the ML model, cross-sectional images of 3D microstructures generated by GeoDict are meshed and transformed into the diffuse-interface microstructures (see Section 4.1). Subsequently, the k_{eff} is calculated using the computational thermal homogenization diffuse-interface model described in Section 3.1. In our material system, the calculated k_{eff} using the computational homogenization model at each temperature for both phases, namely, FPPs and SNs, resulted in nearly identical values. As part of the methodological approach in the ML model, we adopt a thermal conductivity ratio of approximately 0.5 between the two phases. This assumption enables us to explore differences in k_{eff} for different microstructures, i.e., developing the structure–property relation.

5.2. ML Model Training and Validation

The process of building an ML-based surrogate model is outlined in **Figure 8**. It starts with the binary image of the microstructure and illustrates the steps of ML input preparation in the form of principal components at the top. At the bottom, the generation of the output dataset for the ML model through computational homogenization is shown. Training is conducted on 80% of the dataset, with the remaining 20% used for testing. The training steps, which take a few seconds on a PC with a 2.5 GHz Intel Core i5-7200 processor and 8 GB of RAM, are due to the use of a simple data structure of the two-point statistics rather than using large images, sufficient amount of training data, and well-defined features. After training ML models, it is crucial to evaluate their predictive performance. The microstructure–property plot in Figure 8 visualizes FE simulated and ML predicted results and shows the complete agreement between the two values. Regarding the performance evaluation based on the well-known metrics such as the mean square error (MSE) and the coefficient of determination or R^2,[53] the model shows an excellent performance with MSE$= 10^{-5}$ and $R^2 = 0.99$. By obtaining the promising $R^2 = 0.99$ from the initial model results, hyperparameter tuning is unnecessary in this work.

6. Conclusion

This study focused on the development of a hierarchical diffuse-interface thermal homogenization scheme for the calculation of the effective thermal conductivity of $(Hf_{0.7}Ta_{0.3})C/SiC$ UHTC-NC. It was achieved by combining experimental insights such as SEM images and the measured thermal conductivity of the NC. In addition, a ML surrogate model is developed to predict the effective thermal conductivity of new microstructures in order

to establish a useful structure–property relationship that can be useful in the design of new materials. The point-wise conclusions derived from the study are listed as follows: 1) An approach was introduced to define representative images for computational homogenization from both nanoscale FPPs and microscale SNs regions. In addition, several image processing steps, such as Gaussian smoothing and thresholding, were employed to streamline the generation of the meshed image for finite element simulations. 2) Computational thermal homogenization results were demonstrated over nanoscale, microscale, and mesoscale regions at the working temperature of 1600 °C. The thermal conductivity values obtained from both FPPs and SNs were utilized as the reference thermal conductivity for the mesoscale homogenization. 3) The numerically obtained effective thermal conductivity values at different temperatures were validated by experimental measurements. They showed a good agreement at temperatures below 500 °C, considering a normalized interface thermal resistance value of 10^8. This highlights the temperature-dependent nature of the thermal resistance of the material system. 4) A ML-based surrogate model was developed to predict the effective thermal conductivity of composites with different microstructure designs. The model was trained on synthetically generated microstructures to overcome the challenge of obtaining expensive experimental samples and images. The reduced form of the two-point statistics of the microstructure images is fed to the model for training, resulting in a strong performance with $R^2 = 0.99$.

Further investigation of the dependence of the effective thermal conductivity on the porosity is planned for future work. Regarding the ML aspect, it is proposed to extend the dataset to different material systems with different thermal conductivity contrast ratios between the phases. Furthermore, a systematic sensitivity study will be performed to determine the number of principal components that should be used as input to achieve a good performance of the ML model.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

[1] W. G. Fahrenholtz, G. E. Hilmas, *Scr. Mater.* **2017**, *129* 94.

[2] N. Thor, G. Winkens, J. Bernauer, N.-C. Petry, K. Beck, J. Wang, R. Schwaiger, R. Riedel, U. Kolb, M. Leppe, A. Pundt, Microstructure Characterization and Mechanical Properties of Polymer- Derived (HfxTa₁₋ₓ)C/SiC Ceramic Prepared Upon Field-Assisted Sintering Technique (FAST/SPS), In Revision, Special Issue Special Issue MatCom – ComMat to AEM, **2023**, https://doi.org/10.1002/adem.202301841.

[3] E. Ionescu, S. Bernard, R. Lucas, P. Kroll, S. Ushakov, A. Navrotsky, R. Riedel, *Adv. Eng. Mater.* **2019**, *21*, 1900269.

[4] E. Wuchina, E. Opila, M. Opeka, B. Fahrenholtz, I. Talmy, *Electrochem. Soc. Interface* **2007**, *16*, 30.

[5] M. M. Opeka, I. G. Talmy, E. J. Wuchina, J. A. Zaykoski, S. J. Causey, *J. Eur. Ceram. Soc.* **1999**, *19*, 2405.

[6] S. Ghaffari, M. Faghihi-Sani, F. Golestani-Fard, H. Mandal, *J. Eur. Ceram. Soc.* **2013**, *33*, 1479.

[7] A. Bellosi, F. Monteverde, D. Sciti, *Int. J. Appl. Ceram. Technol.* **2006**, *3*, 32.

[8] S. Fu, M. Zhu, Y. Zhu, *J. Adv. Ceram.* **2019**, *8*, 457.

[9] E. Kroke, Y.-L. Li, C. Konetschny, E. Lecomte, C. Fasel, R. Riedel, *Mater. Sci. Eng.: R: Rep.* **2000**, *26*, 97.

[10] R. P. Chaudhary, C. Parameswaran, M. Idrees, A. S. Rasaki, C. Liu, Z. Chen, P. Colombo, *Prog. Mater. Sci.* **2022**, *128*, 100969.

[11] P. Colombo, G. Mera, R. Riedel, G. D. Sorarù, *J. Am. Ceram. Soc.* **2010**, *93*, 1805.

[12] S. Zhai, P. Zhang, Y. Xian, J. Zeng, B. Shi, *Int. J. Heat Mass Transfer.* **2018**, *117* 358.

[13] J.-P. Cao, X. Zhao, J. Zhao, J.-W. Zha, G.-H. Hu, Z.-M. Dang, *ACS Appl. Mater. Interfaces* **2013**, *5*, 6915.

[14] Z. Han, A. Fina, *Prog. Polym. Sci.* **2011**, *36*, 914.

[15] M. R. P. Elenchezhian, V. Vadlamudi, R. Raihan, K. Reifsnider, *Proceedings of CAMX 2018 Conf.*, Dallas, TX, **2018**.

[16] M. Sahimi, *Heterogeneous Materials I: Linear Transport and Optical Properties*, Vol. *1*, Springer **2003**.

[17] M. Sahimi, *Heterogeneous Materials: Nonlinear And Breakdown Properties And Atomistic Modeling*, Vol. *2*, Springer, **2003**.

[18] C. F. Matt, M. E. Cruz, *Numer. Heat Transf. Part A Appl.* **2008**, *53*, 577.

[19] A. M. Thiele, A. Kumar, G. Sant, L. Pilon, *Int. J. Heat Mass Transf.* **2014**, *73* 177.

[20] Y. Yang, M. Fathidoost, T. D. Oyedeji, P. Bondi, X. Zhou, H. Egger, B.-X. Xu, *Scr. Mater.* **2022**, *212* 114537.

[21] A. Aalilija, C.-A. Gandin, E. Hachem, *Int. J. Therm. Sci.* **2021**, *166* 106817.

[22] J. Yvonnet, Q.-C. He, Q.-Z. Zhu, J.-F. Shao, *Comput. Mater. Sci.* **2011**, *50*, 1220.

[23] R. Bostanabad, Y. Zhang, X. Li, T. Kearney, L. C. Brinson, D. W. Apley, W. K. Liu, W. Chen, *Prog. Mater. Sci.* **2018**, *95*, 1.

[24] S. Cheng, Y. Jiao, Y. Ren, *Acta Mater.* **2022**, *229* 117800.

[25] M. Li, H. Zhang, S. Li, W. Zhu, Y. Ke, *Mater. Des.* **2022**, *224* 111340.

[26] S. Torquato, *Random Heterogeneous Materials: Microstructure And Macroscopic Properties*, Vol. *16*, Springer Science & Business Media, **2005**.

[27] J. C. Maxwell, *A Treatise On Electricity And Magnetism*, Vol. *1*, Clarendon Press, **1873**.

[28] V. D. Bruggeman, *Ann. Phys.* **1935**, *416*, 636.

[29] D. C. Pham, S. Torquato, *J. Appl. Phys.* **2003** *94*, 6591.

[30] M. Safdari, M. Baniassadi, H. Garmestani, M. S. Al-Haik, *J. Appl. Phys.* **2012** *112*, 114318.

[31] N. Thor, *Ph.D. Thesis*, Technische Universität Darmstadt, **2023**.

[32] S. G. Kim, W. T. Kim, T. Suzuki, *60*, 7186.

[33] M. Nicoli, M. Plapp, H. Henry, *Phys. Rev. E* **2011**, *84*, 046707.

[34] M. Fathidoost, Y. Yang, M. Oechsner, B.-X. Xu, *Mater. Des.* **2023**, *227*, 111746.

[35] M. R. Tonks, D. Gaston, P. C. Millett, D. Andrs, P. Talbot, *Comput. Mater. Sci.* **2012**, *51*, 20.

[36] C. J. Permann, D. R. Gaston, D. Andrš, R. W. Carlsen, F. Kong, A. D. Lindsay, J. M. Miller, J. W. Peterson, A. E. Slaughter, R. H. Stogner, R. C. Martineau, *SoftwareX* **2020**, *11*, 100430.

[37] B. Lin, *Ph.D. Thesis*, Technische Universität Darmstadt, **2023**.

[38] M. Skolnick, S. Torquato, *Phys. Rev. E* **2021**, *104*, 045306.

[39] K. Storey-Fisher, D. W. Hogg, *Astrophys. J.* **2021**, *909*, 220.

[40] P. Debye, H. Anderson Jr, H. Brumberger, *J. Appl. Phys.* **1957**, *28*, 679.

[41] J. T. Gostick, Z. A. Khan, T. G. Tranter, M. D. Kok, M. Agnaou, M. Sadeghi, R. Jervis, *J. Open Source Softw.* **2019**, *4*, 1296.

[42] D. T. Fullwood, S. R. Niezgoda, S. R. Kalidindi, *Acta Mater.* **2008**, *56*, 942.

[43] D. B. Brough, D. Wheeler, S. R. Kalidindi, *Integr. Mater. Manuf. Innov.* **2017**, *6*, 36.

[44] A. Cecen, T. Fast, S. R. Kalidindi, *Integr. Mater. Manuf. Innov.* **2016**, *5* 1.

[45] H. Hotelling, *J. Educ. Psychol.* **1933**, *24*, 417.

[46] M. Greenacre, P. J. Groenen, T. Hastie, A. I. d'Enza, A. Markos, E. Tuzhilina, *Nat. Rev. Methods Primers* **2022**, *2*, 100.

[47] X. Hu, J. Zhao, Y. Chen, Y. Wang, J. Li, Q. Wu, Z. Wang, J. Wang, *J. Mater. Inform.* **2022**, *2*, 5.

[48] S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, T. Yu, the Scikit-Image Contributors, *PeerJ* **2014**, *2*, e453.

[49] J. Bernauer, N.-C. Petry, N. Thor, S. A. Kredel, D. Teja Teppala, M. Galetz, M. Lepple, A. Pundt, E. Ionescu, R. Riedel, *Adv. Eng. Mater.*, https://doi.org/10.1002/adem.20230186.

[50] W. Tian, R. Yang, *Tech Sci. Press CMES* **2008**, *24* 123.

[51] B. Lin, Y. Bai, B.-X. Xu, *Mater. Des.* **2021**, *197* 109193.

[52] B. Lin, S. Medghalchi, S. Korte-Kerzel, B.-X. Xu, *PAMM* **2023**, *22*, e202200110.

[53] A. V. Tatachar, *Int. J. Innov. Technol. Explor. Eng.* **2021**, *8*, 853.

[54] S. Perevislov, M. Markov, Y. A. Kuznetsov, I. Kravchenko, A. Krasikov, *Russ. Metall.* **2020**, *2020*, 1477.

[55] O. Cedillos-Barraza, S. Grasso, N. Al Nasiri, D. D. Jayaseelan, M. J. Reece, W. E. Lee, *J. Eur. Ceram. Soc.* **2016**, *36*, 1539.