

# Deep Learning-Based Multiscale Prediction of Composite Aerogels' Effective Thermal Conductivity

V.-T. Nguyen, V.-H. Nguyen, S. Naili\*

<sup>a</sup> Univ Paris Est Creteil, Univ Gustave Eiffel, CNRS, UMR 8208, MSME, F-94010 Créteil, France

---

## Abstract

In this work, we present a novel approach for predicting the effective properties of composite materials by integrating multiscale homogenization techniques with deep learning. High-fidelity datasets are generated using Milton's method and asymptotic homogenization through finite element simulations, capturing material behavior across scales from the nanoscale to the macroscale. A key contribution of this study is the development of an advanced multiscale homogenization model that accounts for heterogeneous property distributions within the representative volume element at the mesoscale, enabling more realistic and accurate material representation. The generated datasets are used to train multilayer perceptron models that achieve high predictive accuracy and strong generalization across both similar and dissimilar inclusion scenarios, while significantly reducing computational cost compared to conventional methods for predicting the effective properties of composite materials, specifically the effective thermal conductivity of hybrid aerogels. Additionally, the surrogate models show strong extrapolation capability, providing physically consistent predictions beyond the training domain. This scalable and robust framework offers a powerful tool for data-driven material design and can be readily extended to predict other effective properties.

**Keywords:** Multiscale modeling, asymptotic homogenization, deep learning, multilayer perceptron, effective thermal conductivity, hybrid aerogels.

---

## 1. Introduction

Composite materials play a critical role in various industries, including aerospace, marine, and construction, because of their properties derived from their complex and multiscale microstructures. In general, these remarkable characteristics arise from the interaction of different constituents and their arrangement within the microstructure. As a result, the behavior observed at the macroscopic scale, such as strength or conductivity, is directly influenced by the microstructure behavior of the material. Therefore, understanding the behavior of the microstructure is a key element in accurately predicting or modifying the macroscopic performance of composite materials.

Multiscale modeling involves phenomena occurring across various length and/or time scales, often spanning several orders of magnitude. It is a powerful method for gaining insight into the behavior of advanced materials, particularly composite materials. This approach produces results that are widely accepted in the literature and are crucial for accurately and efficiently predicting material properties to be used in structural simulations. This method has been developed and studied in a reference book written in the 2010s by Weinan [1] emphasizing the mathematical point of view. Since its introduction, multiscale modeling has received significant attention from researchers, particularly in the field of materials science. Its effectiveness and important role have been demonstrated in many recent applications [2, 3, 4, 5, 6].

Multiscale methods typically operate at a minimum of two scales: a coarse scale (often referred to as the macroscale) and a fine scale (commonly called the microscale). Each scale plays a critical role in the problem-solving process. The central idea is that the behavior occurring at a specific scale

---

\*Corresponding author

Email address: naili@u-pec.fr (S. Naili)

21 governs the system's behavior across larger scales; the scales are interconnected. At each scale, well-  
22 established methods have been developed over time to address the relevant phenomena. In general,  
23 finer scales rely on more detailed theories to provide in-depth information about the system.

24 Single-scale modeling, such as molecular dynamics or continuum theory, is now relatively well-  
25 developed and efficient. However, describing systems across multiple scales often requires different  
26 physical laws. For instance, consider solids: at the macroscale (*e.g.*, on the order of centimeters),  
27 solids can be accurately described using density and displacement fields (when the temperature field  
28 is decoupled) governed by the continuum Navier equations. Advances in computational power now  
29 make it feasible to tackle complex multiscale problems, which is increasingly important in fields like  
30 nanoscience, where multiscale modeling techniques are in high demand [7].

31 Multiscale methods can be broadly categorized into two types: sequential and concurrent ap-  
32 proaches: (i) sequential methods involve conducting fine-scale simulations as a pre-processing step to  
33 inform the coarse-scale model. In these approaches, the fine-scale model generates parameters that are  
34 used to describe the coarse scale, after which the fine-scale model is no longer needed. (ii) Concurrent  
35 methods, on the other hand, link models at multiple scales in an integrated framework, where the  
36 different scales operate simultaneously and communicate through a 'handshaking' procedure. This ap-  
37 proach is essential for inherently multiscale systems, where the behavior at one scale strongly depends  
38 on the dynamics at other scales. Unlike sequential methods, concurrent approaches are relatively new,  
39 and only a limited number of models have been developed to date [8, 9].

40 For many practical problems, fully solving the fine-scale problem is computationally infeasible  
41 due to the enormous costs involved. As a result, more efficient alternative approaches are being  
42 developed. Classical methods rely on analytical approaches under simplified conditions to derive  
43 effective models at the scale of interest. On the computational side, several important numerical  
44 techniques explicitly address the multiscale nature of the solutions. These include multigrid methods,  
45 domain decomposition methods, fast multipole methods, adaptive mesh refinement techniques, and  
46 multiresolution methods [7].

47 For composite materials, multiscale models establish a connection between microstructural param-  
48 eters and measurable macroscopic properties. However, developing these models is often a challenging  
49 and time-consuming process that requires significant expertise. Additionally, capturing certain com-  
50 plex phenomena, such as multiscale interactions, remains difficult.

51 To address those limitations, deep learning (DL) techniques, particularly artificial neural networks  
52 (ANNs), have emerged as highly effective solutions (*see* for example [10, 11, 12, 13]). Since multiscale  
53 analyses of composite structures typically require expertise knowledge of modeling and computational  
54 techniques, deep learning offers an efficient alternative by constructing pre-trained surrogate models,  
55 which are able to approximate the relationship between inputs (*e.g.* material properties, geometry)  
56 and outputs (*e.g.* effective properties). Surrogate models built using deep neural networks (DNNs)  
57 can accurately capture mapping between input data and output, resulting in highly precise predictive  
58 models for material behavior. Especially, after the training phase, these surrogate models can be used  
59 to predict the effective properties of various types of similar composites in a very short period of time.

60 Another significant advantage of such surrogate models is the ability to handle the complexity of  
61 the problem effectively, including the material structure complexity, the multiscale nature, the high-  
62 dimensional input space, the large and complex dataset and the uncertainty in data, providing accurate  
63 predictions of the effective properties based on the microstructural features of the materials. This is  
64 particularly important for the design and optimization of materials. Despite advances in the use of  
65 DNNs to predict the effective properties of composite materials, this approach remains a challenge  
66 that warrants further in-depth research.

67 In this paper, our goal is to combine the homogenization of a representative volume element (RVE)  
68 with deep neural networks, namely multilayer perceptron (MLP) models, to predict the effective  
69 properties of composite materials, specifically the effective thermal conductivity (ETC) of composite  
70 aerogels at mesoscopic and macroscopic scales. A deep learning network could provide much faster  
71 prediction compared to conventional multiscale modeling-based simulations. In other words, the use  
72 of DNNs can significantly reduce the complexity compared to traditional simulation methods while  
73 maintaining high prediction accuracy, making it convenient for a wide range of users, even those

74 without extensive expertise in multiscale modeling or numerical simulation. However, one of the  
75 main challenges in multiscale modeling is the difficulty in incorporating physical behavior into the  
76 models. To address this, we explore scenarios with dissimilar inclusions at the microscale, resulting in  
77 a heterogeneous representative volume element at the mesoscale. This approach offers a more accurate  
78 representation of the practical physical behavior of composite materials, accelerating the development  
79 of new aerogels with enhanced properties for various applications. Additionally, the development,  
80 training, and testing of an MLP model have become much more accessible thanks to libraries such as  
81 Keras [14], Tensorflow [15] and Scikit-learn [16].

82 As a consequence, the main contributions of this paper are as follows:

- 83 (i) We propose a novel technique for analyzing the effective properties of composite materials.  
84 Specifically, we develop deep neural network models trained on data generated from multiscale  
85 homogenization models to characterize the effective properties of composite materials, focusing  
86 on the effective thermal conductivity of composite aerogels. Using surrogate models enables  
87 the possibility of instantaneous prediction of effective properties without requiring finite element  
88 simulations, which are typically complex and require expertise in numerical modeling.
- 89 (ii) We first consider scenarios where inclusions in the polymer matrix are similar at the microscale.  
90 An artificial neural network model, specifically a multilayer perceptron, is developed using data  
91 generated from a three-scale homogenization model to predict the effective thermal conductivity  
92 of hybrid aerogels at the mesoscale. Subsequently, we construct another MLP model based  
93 on data from a two-scale asymptotic homogenization model to predict the effective thermal  
94 conductivity of hybrid aerogels at the macroscale.
- 95 (iii) Next, we address scenarios where inclusions in the polymer matrix are dissimilar at the mi-  
96 croscale. A new multiscale homogenization model is developed to handle dissimilar inclusion  
97 scenarios. Using data generated from this model, we constructed an MLP model to predict the  
98 effective thermal conductivity of composite aerogels at the mesoscale. Additionally, for dissimilar  
99 inclusions, we consider a heterogeneous representative volume element at the mesoscale, with  
100 randomly assigned properties in different regions. This approach enables us to construct an MLP  
101 model for predicting the effective thermal conductivity of composite aerogels at the macroscale,  
102 based on data generated from a two-scale asymptotic homogenization model, which provides a  
103 more realistic and accurate representation.

104 This article is structured as follows. The introduction and contributions are presented in this  
105 present section defined as Section 1. Section 2 discusses multiscale modeling, introduces our mul-  
106 tiscale homogenization approach, and formulates the problem. We also review recent literature and  
107 explain how to integrate deep neural networks with multiscale homogenization modeling, including the  
108 mathematical formulation, data generation process, and hyperparameters such as activation functions  
109 and loss functions. Furthermore, Section 2 provides an overview of the multilayer perceptron model,  
110 covering the basic elements of neural networks and the training procedure. In Section 3, we apply our  
111 approach to predict the effective thermal conductivity of composite aerogels. Specifically, Section 3.1  
112 demonstrates the high accuracy of our MLP models in predicting the effective thermal conductivity  
113 of hybrid aerogels for similar inclusion scenarios, while Section 3.2 addresses dissimilar inclusion sce-  
114 narios. Additionally, a comparison on computational time between our MLP models and traditional  
115 methods is also presented in this section. The applications of the developed surrogate models together  
116 with their performance on extrapolation are discussed in Section 4. Finally, Section 5 presents a con-  
117 clusion, perspectives, and potential directions for future research. The Appendix includes technical  
118 details concerning the homogenization methods used in this paper.

119 **2. Methodology**

120 In this section, we present the methodology adopted in this study, which combines multiscale mod-  
121 eling and homogenization techniques with deep learning to predict the effective thermal conductivity  
122 of composite materials. We begin by outlining the theoretical foundations of multiscale modeling,  
123 homogenization methods, and the formulation of the problem. We then describe how deep neural net-  
124 works are incorporated into the multiscale homogenization framework. To provide context, we include  
125 a brief review of recent literature followed by describing the important advantages of our approach  
126 compared to other neural network architectures or traditional regression models. Finally, we give an  
127 overview of the MLP architecture and training procedure, this includes the mathematical formula-  
128 tion, solving optimization problem, and key hyperparameters of the model such as activation functions  
129 and loss functions. The statistical parameters used to assess the training process are also presented.  
130 Finally, the section concludes by highlighting the significant influence of the training process on the  
131 MLP model's predictive performance, particularly in estimating effective properties.

132 *2.1. Overview of the approach*

133 This section outlines the theoretical foundations that support our proposed methodology. We be-  
134 gin with an introduction to multiscale modeling, which provides the general framework for capturing  
135 material behavior across different spatial scales. This is followed by a presentation of the homogeniza-  
136 tion methods used to bridge these scales and estimate effective properties. Finally, we describe the  
137 thermal problem formulation, which defines the physical context of our study and the target property  
138 effective thermal conductivity used throughout the multiscale analysis.

139 *2.1.1. Multiscale modeling*

140 In general, traditional approaches to modeling focus on one scale. If the interest is the macroscale  
141 behavior of a system, for instance, in an engineering application, the model must take into account  
142 the effect of the smaller scales by some constitutive relations. Typically, physical phenomena can be  
143 modeled at various levels of complexity and across different spatial and temporal scales. Multiscale  
144 modeling provides a framework, based on fundamental principles, for constructing mathematical mod-  
145 els of such phenomena, by examining the connection between models at different scales. The general  
146 principles of multiscale models consist of capturing the behavior at a given scale of complex systems  
147 with the help of models at a smaller scale. This characteristic makes multiscale modeling become a  
148 powerful approach used in the analysis and design of composite materials, which have complex struc-  
149 tures and behaviors due to their heterogeneous nature (see, for example [17, 18] and [19, 20, 21, 22, 23]).  
150 Such models are used in simulations at different spatial or temporal scales, enabling a comprehensive  
151 analysis of systems where interactions across scales significantly influence overall behavior. In this  
152 study, multiscale modeling provides insight into the prediction of effective properties of composite  
153 materials by bridging the nanoscopic to macroscopic scales. In particular, coupling between scales  
154 is achieved using homogenization techniques, where outputs from lower scales inform parameters at  
155 higher scales.

156 To illustrate the methodology developed in this paper, composite aerogels are considered. At the  
157 mesoscopic scale, these aerogels consist of two phases: a gas phase and a solid phase. The solid  
158 phase is composed of graphene oxide fillers dispersed within a polymer matrix, forming a network of  
159 thin alveolar structures with interconnected pores. These pores have diameters ranging from several  
160 dozen micrometers to 200 micrometers. The structure of composite aerogels is analyzed across four  
161 distinct scales: nanoscopic, microscopic, mesoscopic, and macroscopic scales. By leveraging multiscale  
162 modeling, this work enables a deeper understanding of the materials' effective properties. In addition,  
163 the integration of multiscale ensures both accuracy and computational efficiency, addressing the lack  
164 of predictive capacity or oversimplification in existing approaches.

165 *2.1.2. Homogenization method*

166 Composites are heterogeneous materials composed of multiple phases, with their macroscopic prop-  
167 erties depending on the distribution, orientation, and material characteristics of these phases. Beside  
168 experimental measurements, homogenization method is a process used in the physical description of

169 heterogeneous materials, where the material is replaced with an equivalent, in some sense, homogeneous  
 170 material. Such kind of method is often used to derive the effective properties of heterogeneous  
 171 materials (see, for example [24]). One of its advantages is the significantly lower computational cost  
 172 compared to other numerical formulations (as illustrated in [25]).

173 Given the advantage of determining the effective properties of composite materials at a single scale,  
 174 it is logical to consider extending homogenization methods to multiple scales. This leads to a coupling  
 175 approach, where multiscale modeling and homogenization method provide a systematic framework for  
 176 connecting different scales from nanoscopic to macroscopic scales, enabling accurate modeling of the  
 177 effective properties of composite materials. The core concept of this coupling method arises from the  
 178 complex interactions occurring at smaller scales, which have a direct impact on the behavior of larger  
 179 scales. As a consequence, the behavior of a material at a larger scale can be derived from its behavior  
 180 at smaller scales. This is achieved by homogenizing the properties of the material at smaller scales to  
 181 determine its effective properties at larger scales. By employing this method, we can capture critical  
 182 nanostructural details while avoiding the significant computational cost of simulating the entire system  
 183 at high resolution.

184 The accuracy and reliability of the results obtained through computational multiscale homoge-  
 185 nization depend entirely on the quality of the descriptions of materials, phases, and interfaces at the  
 186 smaller scales. This coupling approach has been successfully applied to efficiently and accurately char-  
 187 acterize fiber-reinforced composites (see [26]) and to model the mechanical response of heterogeneous  
 188 materials (see [27]).

189 Moreover, combining multiscale homogenization methods with machine learning techniques has  
 190 emerged as a promising direction to enhance their predictive capabilities (see [28]).

191 To illustrate the methodology developed in this paper, the asymptotic homogenization method (see,  
 192 for example [29, 30],[31, 32]) and Milton method (see, for example [33]) are used. For instance, the  
 193 asymptotic homogenization process involves solving boundary value problems based on a representative  
 194 volume element to extract effective properties such as stiffness, permeability, or conductivity. In the  
 195 framework of this approach, periodic boundary conditions are typically applied to ensure accurate  
 196 representation of the material's behavior. These methods are briefly described in the Appendix.

197 The structure of composite materials can be analyzed at several distinct scales. In this work, we  
 198 denote by  $\mathbb{K}_e^{(j+1)}$  for  $j = 1, 2, 3$  the second-order effective conductivity tensor corresponding to the  
 199 three scales (microscale, mesoscale, and macroscale).

200 For each scale, let  $\hat{\Theta}^{(i)} = (\hat{\Theta}_g^{(i)}, \hat{\Theta}_p^{(i)})$  represent the characteristic hyper-parameter of the heteroge-  
 201 neous structure, where  $\{\hat{\Theta}_g^{(i)}\}_{i=1}^4$  denotes the geometric hyper-parameter, and  $\{\hat{\Theta}_p^{(i)}\}_{i=1}^4$  denotes the  
 202 material properties hyper-parameter, in which the upper indice (1) refers to the nanoscale. Conse-  
 203 quently, the second-order effective conductivity tensor at each scale can be expressed as:

$$\mathbb{K}_e^{(j+1)} = \mathbb{K}_e^{(j+1)}(\hat{\Theta}_g^{(j)}, \hat{\Theta}_p^{(j)}), \quad \text{for } j = 1, 2, 3. \quad (1)$$

204 Let  $H_i$  for  $i = 1, 2, 3, 4$  be the vector spaces of  $\mathbb{R}^n$  which correspond to the four scales and  $\Psi_j$  for  
 205  $j = 1, 2, 3$  be the mapping defined as:

$$\begin{aligned} H_j &\xrightarrow{\Psi_j} H_{j+1} \\ \hat{\Theta}_p^{(j)} &\mapsto \hat{\Theta}_p^{(j+1)} \equiv \mathbb{K}_e^{(j+1)}. \end{aligned} \quad (2)$$

206 On the other hand, let  $\Psi = \Psi_2 \circ \Psi_1$  be the mapping from nano to mesocale, as it can be depicted as:

$$\begin{array}{c} H_1 \xrightarrow{\Psi_1} H_2 \xrightarrow{\Psi_2} H_3 \\ \curvearrowright \Psi = \Psi_2 \circ \Psi_1 \end{array} \quad (3)$$

207 One of our principal interests is to approximate the mapping  $\Psi$  without passing through  $\Psi_2 \circ \Psi_1$ .  
 208 To achieve this goal, deep learning techniques, such as multilayer perceptron models using deep neural  
 209 networks, become very potential candidates due to the ability of capturing mappings based on given

210 datasets with a very high accuracy. Such kind of data-driven methods were studied for a wide range  
 211 of applications, see [34, 35, 36, 37]. Thanks to the high predictive capability with low computational  
 212 cost, those surrogate models become more and more popular in more complex problems (multiscale  
 213 modeling, computational fluid dynamics, molecular dynamics, large-scale simulations) which require  
 214 very expensive cost in terms of calculation and time consuming. Among those problems, coupling  
 215 multiscale homogenization models with deep learning techniques is a very interesting direction and  
 216 deserves further in-depth research.

217 In this work, we aim to approximate the mapping  $\Psi$  by MLP surrogate models. Similarly to  
 218 what were described above: the effective conductivity tensor can be expressed as  $\tilde{\mathbb{K}}_e^{(j+1)}(\tilde{\Theta}_g^{(j)}, \tilde{\Theta}_p^{(j)})$   
 219 for  $j = 1, 2, 3$ , where  $\tilde{\Theta}^{(i)} = (\tilde{\Theta}_g^{(i)}, \tilde{\Theta}_p^{(i)})$  stands for the characteristic parameter consisting of  $\{\tilde{\Theta}_g^{(i)}\}_{i=1}^4$   
 220 the geometric hyper-parameter and  $\{\tilde{\Theta}_p^{(i)}\}_{i=1}^4$  the physical hyper-parameter obtained at each scale by  
 221 the MLP models, in which the upper indice (1) refers to the nanoscale.

222 Let  $\tilde{H}_i$  be the approximating spaces for  $H_i$ ,  $i = 1, 2, 3, 4$  and  $\tilde{\Psi}_j$  be the approximation for  $\Psi_j$ ,  
 223  $j = 1, 2, 3$  defined as:

$$\begin{aligned} \tilde{H}_j &\xrightarrow{\tilde{\Psi}_j} \tilde{H}_{j+1} \\ \hat{\Theta}^{(j)} &\mapsto \tilde{\Theta}_p^{(j+1)} \equiv \tilde{\mathbb{K}}_e^{(j+1)}. \end{aligned} \quad (4)$$

224 Furthermore, let  $\tilde{\Psi} = \tilde{\Psi}_2 \circ \tilde{\Psi}_1$  be the approximation for  $\Psi$ , as it can be expressed as:

$$\begin{aligned} \tilde{H}_1 &\xrightarrow{\tilde{\Psi}_1} \tilde{H}_2 \xrightarrow{\tilde{\Psi}_2} \tilde{H}_3 \\ &\swarrow \tilde{\Psi} = \tilde{\Psi}_2 \circ \tilde{\Psi}_1 \nearrow \end{aligned} \quad (5)$$

225 Hence,  $\tilde{\Psi}$  is an approximation of  $\Psi$  for Milton method, asymptotic homogenization method and  
 226 finite element method (FEM). Particularly,  $\tilde{\Psi}$  can be constructed by MLP surrogate models using  
 227 deep neural networks, which will be detailed in section 2.2.2.

### 228 2.1.3. Thermal problem

229 Composite materials are widely used in engineering and industrial applications due to their superior  
 230 mechanical, thermal, and multifunctional properties. However, their heterogeneous microstructure,  
 231 consisting of distinct phases with varying thermal conductivities, presents significant challenges in  
 232 accurately predicting their effective thermal behavior. Therefore, gaining a clear understanding of the  
 233 thermal properties of composites is essential for applications such as thermal management systems,  
 234 aerospace components, or electronic packaging, where heat transfer plays a critical role.

235 The thermal problem in composite materials arises from the complex interaction of heat conduction  
 236 within and between their constituent phases. These interactions are influenced by various factors  
 237 such as geometry, distribution and thermal properties of the inclusions and matrix, as well as the  
 238 presence of interfacial resistance. Hence, precisely modeling these interactions is crucial for predicting  
 239 effective properties such as thermal conductivity, which determine the overall performance of composite  
 240 materials in practical applications.

241 To address this complexity, the multiscale homogenization method provides a systematic framework  
 242 for linking microscale thermal behavior with macroscale effective properties. By modeling heat transfer  
 243 at the microscale and upscaling the results, this method captures the influence of microstructural  
 244 features on the overall thermal performance of the composite material. This approach eliminates the  
 245 need for computationally expensive full-scale simulations, making it an efficient and accurate tool for  
 246 predicting effective thermal properties.

247 In this work, we are interested in characterizing the effective thermal conductivity in the steady-  
 248 state of composite materials. Particularly, for obtaining the ETC from scale  $(j)$  to scale  $(j+1)$   
 249 for  $j = 1, 2, 3$ , we consider a two-phase composite material which can be denoted as a domain  $\Omega$ ,  
 250 comprising two distinct phases  $\Omega_a$  and  $\Omega_b$  which are separated by the common boundary  $\Gamma$ . Under

251 the assumption of an absence of volumetric heat generation, the steady-state heat conduction equations  
 252 for each phase are outlined as follows:

$$\begin{cases} \operatorname{div}(\mathbb{K}^\alpha(\mathbf{x}) \operatorname{grad} T^\alpha(\mathbf{x})) = 0, & \text{in } \Omega_\alpha, \alpha = a, b, \\ T^a(\mathbf{x}) = T^b(\mathbf{x}), & \text{on } \Gamma, \\ \mathbb{K}^a(\mathbf{x}) \operatorname{grad} T^a(\mathbf{x}) \cdot \mathbf{n} = \mathbb{K}^b(\mathbf{x}) \operatorname{grad} T^b(\mathbf{x}) \cdot \mathbf{n}, & \text{on } \Gamma, \end{cases} \quad (6)$$

253 where  $T^\alpha$  denotes the temperature,  $\mathbb{K}^\alpha$  denotes the second-order thermal conductivity tensor of phase  
 254  $\alpha$  and  $\mathbf{n}$  stands for the normal vector to  $\Gamma$ . The operators  $\operatorname{div}$  and  $\operatorname{grad}$  represent the divergence  
 255 and gradient, respectively. For simplicity, we assume a perfect interface and apply Dirichlet boundary  
 256 conditions. These assumptions eliminate the need to introduce additional variables such as thermal  
 257 contact resistance, thereby reducing the number of unknowns in the system. This simplification leads  
 258 to faster numerical simulations and better convergence. More importantly, it allows us to focus on  
 259 the intrinsic thermal behavior of the constituent phases (solid and gas) without interference from  
 260 interfacial effects. Such assumptions are particularly useful in early-stage studies or when developing  
 261 and testing machine learning models like MLPs, as they streamline the modeling process and emphasize  
 262 the underlying structure-property relationships without unnecessary computational overhead.

263 This study applies the multiscale homogenization method to predict the effective thermal conduc-  
 264 tivity of composite materials, accounting for the heterogeneity of their microstructure. By incorporat-  
 265 ing deep neural networks and multiscale homogenization, we aim to provide insight into the thermal  
 266 behavior of composites and to offer a predictive tool for material design. The results will contribute to  
 267 the optimization of composite materials for thermal management applications and other heat-intensive  
 268 environments.

## 269 2.2. Deep learning method based on artificial neural networks and multiscale homogenization modeling

270 The increasing complexity of multiscale homogenization modeling, particularly for composite ma-  
 271 terials with intricate microstructures, necessitates efficient and accurate computational tools. Deep  
 272 learning methods based on artificial neural networks provide a powerful framework for approximat-  
 273 ing complex mappings and accelerating computations. By integrating deep learning with multiscale  
 274 homogenization, this study aims to enhance predictive capabilities and computational efficiency of  
 275 traditional multiscale homogenization approaches.

276 Artificial neural networks excel at capturing relationships and high-dimensional mappings, mak-  
 277 ing them ideal for approximating the effective properties of heterogeneous materials. In the context  
 278 of multiscale homogenization, ANNs can be employed to learn the relationship between input fea-  
 279 tures (material properties and geometrical configuration of phases) and effective properties (thermal  
 280 conductivity, stiffness) across distinct scales. This approach reduces the complexity associated with  
 281 solving multiscale problems from the nanoscopic to the macroscopic scales while maintaining accu-  
 282 racy. The idea of using deep neural networks to predict the effective thermal conductivity of two-phase  
 283 materials is an innovative approach that leverages the power of deep learning to model complex re-  
 284 lationships in materials science. Typically, two-phase materials, which consist of distinct phases such  
 285 as solid-liquid, solid-gas or liquid-gas, exhibit complex thermal behaviors due interactions between  
 286 the different phases. Traditional methods for predicting thermal conductivity, such as theoretical  
 287 models (Maxwell's equations or effective medium theory), often struggle to account for the nonlinear,  
 288 heterogeneous, and dynamic nature of these materials.

289 The proposed methodology uses a combination of numerical multiscale homogenization and deep  
 290 learning to predict effective material properties. First, numerical simulations using traditional mul-  
 291 tiscale homogenization methods are performed to generate a dataset of input-output pairs, where  
 292 inputs represent microstructural descriptors (volume fractions, material phases) at smaller scales and  
 293 outputs correspond to effective properties at larger scales. This dataset is then used to train an ANN,  
 294 which learns the underlying mapping between input features and output responses. Once trained,  
 295 the ANN serves as a surrogate model, providing rapid predictions of effective properties for new mi-  
 296 crostructural configurations. The artificial neural network employed in this study consists of multiple  
 297 fully connected layers with ReLU activation functions, designed to capture the mapping between mi-  
 298 crostructural features and effective properties. The network is trained using a supervised learning

approach with Adam optimizer to minimize the mean squared error between predicted and computed effective properties. The integration of ANNs into the multiscale homogenization framework significantly enhances its efficiency and scalability. The trained ANN acts as a surrogate model, enabling rapid predictions of effective properties without the need for extensive microscale numerical simulations. This capability is particularly beneficial for large-scale parametric studies or real-time applications, where computational cost is a limiting factor. Moreover, the DL approach captures complex relationships and inter-dependencies that may be challenging to model explicitly using traditional multiscale homogenization techniques.

By combining deep learning with multiscale homogenization, this study achieves several key advantages: computational efficiency, accuracy, scalability, and flexibility. The proposed methodology provides a robust framework for exploring the design space of advanced materials and optimizing their performance for a wide range of applications.

### 2.2.1. Recent advances

With the rapid and unstoppable growth of artificial intelligence, machine learning techniques are being applied in various fields, including numerical engineering and materials science. One such application involves the use of deep neural networks to predict the effective thermal conductivity of two-phase materials. By leveraging deep learning, these models can accurately capture complex relationships and predict the thermal conductivity of these materials. This approach is particularly beneficial due to the complexity and high computational cost associated with traditional methods of predicting thermal conductivity, see [38, 39, 40, 41, 42, 43]. In this method, a DNN is trained on a large dataset generated from preliminary analyses. The DNN model takes into account various factors such as grain size, temperature, and other material properties, see [43, 44] and [45, 46, 47]. The trained DNN model can then be used to validate and predict more complex analyses, providing predictions for the effective thermal conductivity of two-phase materials. This approach is not only more efficient but also offers comparable, if not superior, accuracy to conventional methods, as discussed in [43, 44, 38] and [48, 49]. One of the key advantages of this approach is its ability to handle large datasets and complex relationships between variables, making it a powerful tool for predicting the thermal conductivity of two-phase materials.

Deep Material Network (DMN), Deep Homogenization Network (DHN), and 3D Convolutional Neural Network (3D-CNN) are machine learning approaches for multiscale homogenization, each with unique strengths. DMN is a data-driven method that learns the topology of material on multiple scales to predict the effective properties of heterogeneous materials, see [50, 51]. DHN incorporates homogenization theory into neural networks, making it ideal for unidirectional composites with thermal conductivity, ensuring physically consistent predictions, see [52]. 3D-CNN uses high-fidelity datasets to predict the effective properties of heterogeneous materials, excelling in uncertainty quantification and model transferability, see [53, 54, 36]. These methods overcome the high computational costs of traditional techniques while enabling efficient, accurate multiscale material analysis and design.

MLPs are particularly well-suited for studying the effective thermal conductivity of aerogels due to the nature of the input data, which typically consists of structured numerical features. In the present work, these inputs include scalar material properties such as pore size, porosity, thickness, diameter, and phase conductivity. Compared to more complex architectures like 3D Convolutional Neural Networks, Graph Neural Networks (GNNs) [55, 56], Deep Material Networks or Deep Homogenization Networks, MLPs have a simpler structure, making them easier to train and interpret. They also offer a favorable balance between predictive performance and computational efficiency, especially for small to medium-sized datasets. This is particularly advantageous in material science, where data is often limited due to the high cost and effort to obtain. Moreover, MLPs require less memory and training time than more complex models like 3D-CNNs, GNNs, DMN or DHN. On the other hand, 3D-CNNs are best suited for image-based or spatially correlated data, such as microstructure images. GNNs are specifically designed for data with an explicit graph structure, such as atomic or molecular. DMN requires more prior knowledge and engineering effort to construct, while DHN involves higher architectural complexity and training costs. Overall, models like 3D-CNNs, GNNs, DMN, and DHN typically require larger datasets to achieve good generalization performance, due to their higher model complexity and greater number of parameters.

Apart from that, a very natural and important question to be posed, that is “What are the advantages of using MLPs over traditional regression models (such as linear or polynomial regression) for predicting the effective thermal conductivity of aerogels?” The first reason, that is, traditional regression models do not perform well for multiscale modeling of ETC in aerogels. This is due to the inherently complex relationships between various microstructural parameters such as porosity, pore size, phase-specific thermal conductivities and the resulting ETC. Traditional models, including linear and polynomial regression, typically require explicit feature engineering and the selection of fixed basis functions, which can limit their capacity to capture such complex dependencies. On the contrary, MLPs are universal function approximators that are well-suited for modeling high-dimensional relationships without manual intervention. They can naturally handle multiple interacting input features, even when these interactions are non-additive or strongly coupled, scenarios where classical regression methods often struggle unless highly customized. Secondly, When properly trained, MLPs generally outperform traditional regression models, especially in cases where the underlying physical behavior cannot be easily expressed through a predefined mathematical form. Thirdly, MLPs learn to extract and combine relevant features automatically during training, whereas traditional models typically require handcrafted features such as ratios or products. Furthermore, MLPs are also highly scalable and adaptable. Their performance can be enhanced through hyperparameter tuning or deeper network architectures as more training data becomes available. Additionally, they tend to generalize better to unseen data and integrate easily into end-to-end learning pipelines, including tasks like feature selection, uncertainty quantification, or multi-output prediction. In contrast, traditional regression methods are generally more limited in flexibility and may exhibit early performance saturation due to the constraints imposed by their fixed functional forms. For these reasons, MLPs represent a more robust and versatile approach for modeling the effective thermal behavior of complex materials like aerogels.

Therefore, in this work, our approach is to use MLP models, a type of supervised learning, to avoid the need for solving multiscale homogenization problems and to obtain results more efficiently. Specifically, we aim to construct 2 different MLP models to predict the effective properties of the composite materials at mesoscopic and macroscopic scales. For the first MLP model, the input data, or features, will include parameters describing geometrical configurations and physical properties from nanoscopic and microscopic scales to predict the effective properties of the composite materials at mesoscale. A second MLP model is then constructed to predict the effective properties of the composite materials at macroscale with the input including geometrical configurations at mesoscale and the physical properties resulted from the microscopic calculation.

### 2.2.2. Multilayer perceptron model

In general, an artificial neural network for regression tasks is trained to associate input features  $\mathbf{x}$  with the corresponding label  $\mathbf{y}$ . Specifically, a multilayer perceptron is an advanced feed forward ANN with multiple layers and nonlinear activation functions, capable of handling nonlinearly separable data. Unlike the basic perceptron from 1958, MLPs are trained using backpropagation, adjusting weights based on errors to improve accuracy. An MLP consists of an input layer, one or more hidden layers, and an output layer, with neurons connected between layers.

Two neurons are linked by a scalar “weight”. For instance in Fig. 1, for the neuron  $x_1^{(1)}$ , input features  $x_i^{(0)}$  are multiplied by weights  $w_{1i}$  (for  $i = 1, \dots, n$ ), summed with a bias term  $b_1$ , and then, passed through an activation function, introducing nonlinearity. The bias allows for the shift of the activation function, improving the accuracy of the model.

As a consequence, the output at the hidden layer, in what follows denoted by  $\mathbf{y}_l$ , is then expressed as:

$$\mathbf{y}_0 = \mathbf{x}, \quad (7)$$

$$\mathbf{y}_l = f(\mathbf{W}_l \mathbf{y}_{l-1} + \mathbf{b}_l), \quad l \in [1, L - 1], \quad (8)$$

$$\mathbf{y}_L = \mathbf{W}_L \mathbf{y}_{L-1} + \mathbf{b}_L, \quad l = L, \quad (9)$$

where  $L$  is the number of layers,  $f$  is the activation function,  $\mathbf{W}_l$  and  $\mathbf{b}_l$  are the weight matrix and bias vector, respectively, for layer  $l$ . The output of the last layer is denoted by  $\hat{\mathbf{y}}$ , the predicted values,

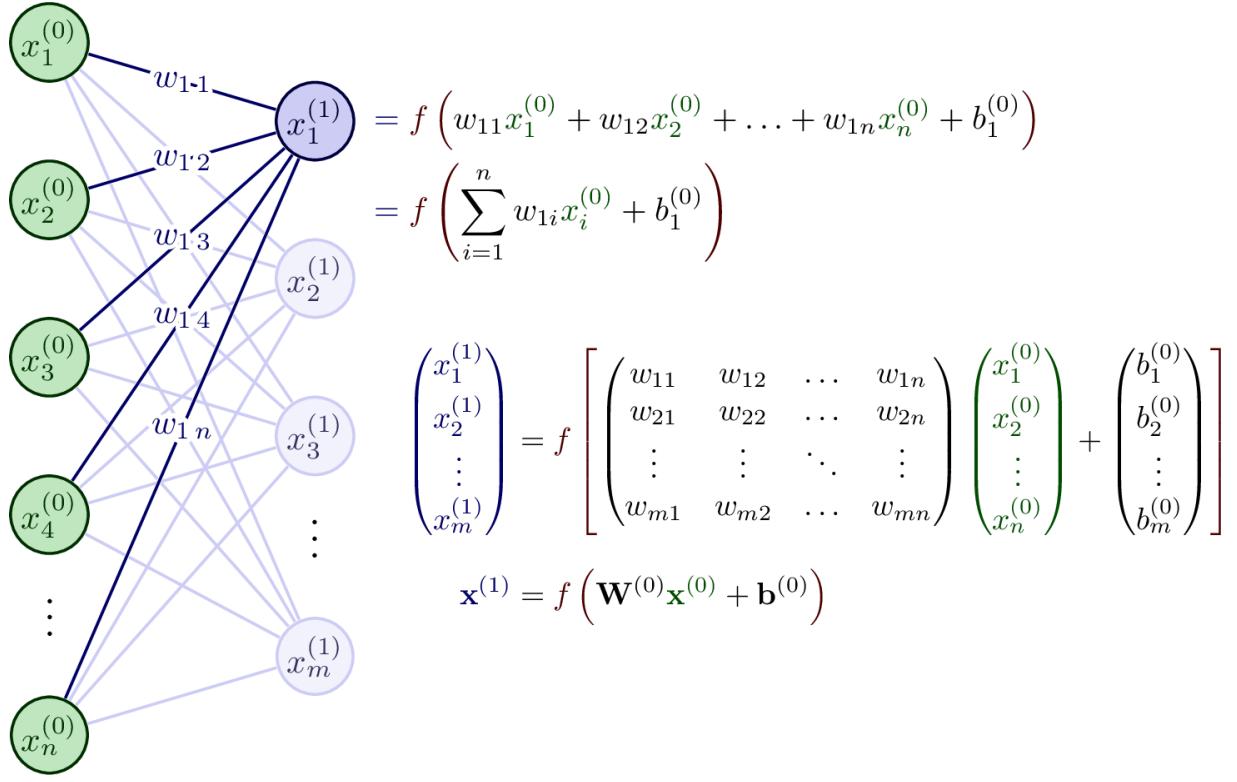


Figure 1: Schematic representation of the execution of the neuron  $x_1^{(1)}$ .

400 which are used with the ground truth  $\mathbf{y}$  to compute the loss function  $\mathcal{L}(\hat{\mathbf{y}}, \mathbf{y})$ .

401 The procedure is repeated through all hidden layers until the output of the last layer is computed.  
402 A loss function is then computed to assess the difference between predicted values  $\hat{\mathbf{y}}$  and ground truth  
403  $\mathbf{y}$ . If stopping criteria are not met or the maximum number of epochs is not reached, backpropagation  
404 iteratively adjusts the weights and biases, continuing until the network optimizes the mapping between  
405 inputs and labels, improving accuracy and performance. Specifically, the network's behavior can be  
406 represented by a mathematical function to approximate a mapping  $\tilde{\Psi}$  such that:

$$\hat{\mathbf{y}}_i = \tilde{\Psi}(\mathbf{x}_i, \mathbf{W}), \quad (10)$$

407 for any training input vector  $\mathbf{x}_i$  and predicted output vector  $\hat{\mathbf{y}}_i$ , where  $\mathbf{W}$  describes the weight matrix.  
408 The average loss function used for the minimizing process is defined over  $N$  samples  $\{\mathbf{x}_i, \mathbf{y}_i\}_{i=1}^N$  as:

$$\mathcal{L}(\mathbf{y}_i, \tilde{\Psi}(\mathbf{x}_i, \mathbf{W})) = \frac{1}{N} \sum_{i=1}^N e(\mathbf{y}_i, \tilde{\Psi}(\mathbf{x}_i, \mathbf{W})), \quad (11)$$

409 where  $e(\mathbf{y}_i, \tilde{\Psi}(\mathbf{x}_i, \mathbf{W}))$  describes the error function between  $\mathbf{y}_i$  and  $\tilde{\Psi}(\mathbf{x}_i, \mathbf{W})$  which can be defined as  
410 the squared error, absolute error or relative error.

411 Based on the available information, the network will attempt to determine the function  $\tilde{\Psi}$  that  
412 solves the optimization problem:

$$\mathbf{W}^* = \arg \min_{\mathbf{W}} \frac{1}{N} \sum_{i=1}^N e(\mathbf{y}_i, \tilde{\Psi}(\mathbf{x}_i, \mathbf{W})), \quad (12)$$

413 where the operator  $\arg \min_{\mathbf{W}} e(\mathbf{W})$  returns the set of  $\mathbf{W}$  for which  $e(\mathbf{W})$  attains the function's smallest  
414 value (if it exists). By using the weight matrix  $\mathbf{W}$ , problem (12) can be solved *via* gradient descent  
415 with backpropagation. Gradient descent updates function parameters iteratively by moving opposite  
416 to the gradient, which points to the steepest increase, thus guiding toward the minimum. At the  $t$ -th

417 iteration step, the components  $w_{kj}$  (for  $k = 1, \dots, m$  and  $j = 1, \dots, n$ ) of the weight matrix  $\mathbf{W}$   
 418 in Eq. (12) are expressed as:

$$w_{kj}^{(t)} = w_{kj}^{(t-1)} - r^{(t)} \frac{\partial \mathcal{L}(y_i, \tilde{\Psi}(x_i, \mathbf{W}))}{\partial w_{kj}^{(t-1)}}, \quad (13)$$

419 with  $r^{(t)}$  represents the learning rate.

420 One of the most often used activation function is the ReLU function, defined as  $f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto \max\{0, x\}$ , where  $\mathbb{R}$  is the set of real numbers. This function is favored for its simplicity and  
 421 efficiency. In this work, ReLU is employed as the activation function for the hidden layers, while the  
 422 linear activation function, defined by  $f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto x$ , which simply returns the input value, is  
 423 used for the output layer.

424 Efficiently training a deep neural network model requires careful selection of the loss function. For  
 425 regression problems, mean square error (MSE) is commonly used as the loss function. Additionally,  
 426 mean absolute error (MAE) is often employed as a metric to evaluate the training process. In this  
 427 study, the MSE and MAE are computed using the following relations:

$$\mathcal{L}_{MSE}(\hat{\mathbf{y}}, \mathbf{y}) = \frac{1}{N} \sum_{i=1}^N |\hat{\mathbf{y}}_i - \mathbf{y}_i|^2, \quad (14)$$

$$\mathcal{L}_{MAE}(\hat{\mathbf{y}}, \mathbf{y}) = \frac{1}{N} \sum_{i=1}^N |\hat{\mathbf{y}}_i - \mathbf{y}_i|, \quad (15)$$

429 where  $\hat{\mathbf{y}}_i$  and  $\mathbf{y}_i$  are the predicted and actual  $i$ -th output values from the training, respectively. The  
 430 symbol  $N$  represents the total number of samples.

431 In order to verify the ability of the trained model in predicting unseen data, the test datasets are  
 432 used together with statistical error analysis parameters. The important statistical parameters used in  
 433 this work are the relative error ( $Re$ ) and coefficient of determination ( $R^2$ ), which can be calculated as:

$$Re_i = \frac{|\hat{\mathbf{y}}_i - \mathbf{y}_i|}{|\mathbf{y}_i|}, \quad (16)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N |\hat{\mathbf{y}}_i - \mathbf{y}_i|^2}{\sum_{i=1}^N |\hat{\mathbf{y}}_i - \frac{\sum_{i=1}^N \mathbf{y}_i}{N}|^2}. \quad (17)$$

434 Note that a  $R^2$  value approaching 1 suggests that the model fits the data very well, indicating a  
 435 high degree of precision in the prediction of outcomes.

436 The data will be obtained through a multiscale homogenization calculation by Milton's method  
 437 and an asymptotic homogenization method development with finite elements method.

438 The training phase in the first stage of applying a Multilayer Perceptron is crucial to the success  
 439 of the second stage, where the trained model is used to predict effective properties. The purpose of  
 440 training the MLPs is to learn a functional mapping between input parameters (e.g. porosity, pore  
 441 size, phase conductivity, thickness, etc.) and a target property (e.g. effective thermal conductivity).

442 During training, the MLP adjusts its internal weights to minimize the error between predicted and  
 443 true values. This process enables it to capture the relationships between material features and their  
 444 effective thermal property. A well-trained MLP can not only fit the training data but also generalize  
 445 to new, unseen data and this is essential for reliable predictions in the second stage. Moreover, it can  
 446 help preventing underfitting or overfitting, in other words, the model hasn't learned enough, it fails to  
 447 capture important relationships, leading to poor predictions or the model has memorized the training  
 448 data, it performs well on known inputs but poorly on new ones. If the model will be tested outside the  
 449 training data range (extrapolation), a properly trained and regularized MLP is more likely to yield  
 450 physically consistent results.

451 On the other hand, in the second stage, the trained MLPs are used to make fast, data-driven  
 452 predictions of effective thermal properties without the need for expensive simulations or experiments.

453 If training was successful, predictions will be accurate and physically plausible, obtained almost im-  
454 mediately which helps saving time, calculation costs and support material design and optimization.  
455 If training was poor, predictions may be unreliable or nonsensical leading to mislead further design  
456 decisions and fail to generalize to realistic material scenarios.

### 457 3. Application for predicting the ETC of composite aerogels

458 In this section, we present the application of the proposed methodology to predict the effective  
459 thermal conductivity of composite aerogels. We begin by briefly introducing aerogels and their im-  
460 portant role as thermal super-insulators. A multiscale homogenization framework is then established  
461 based on the analysis of the aerogels structure on four distinct scales. As a consequence, corresponding  
462 MLP models are developed to predict the ETC of aerogels from nano to mesoscales and from meso  
463 to macroscales, respectively. We then evaluate the performance of our MLP models across different  
464 scenarios. Specifically, Section 3.1 focuses on cases with similar inclusions and demonstrates the high  
465 predictive accuracy of the MLP approach. In Section 3.2, we extend the analysis to more challenging  
466 scenarios involving dissimilar inclusions. Furthermore, this section includes a comparison of com-  
467 putational time between the proposed MLP models and conventional multiscale modeling methods,  
468 highlighting the efficiency gains achieved through our approach.

469 Despite their name, aerogels are solid, elastic, and dry materials that do not resemble a gel in their  
470 physical properties. The name comes from the fact that they are made from gels. The highly porous  
471 structure of aerogels means that up to 99.99% of the air can be confined, giving thermal conductivity  
472 values at atmospheric pressure that are lower than those of air at rest. They are therefore described as  
473 thermal super-insulators that put a brake on heat loss. The use of this type of ultralight nanostructured  
474 solid means that the thickness of the insulating material can be reduced by up to three times, while  
475 maintaining the same thermal conductivity. It is noteworthy that more recently, transparent aerogels  
476 have been developed to replace window glass. However, in addition to their high cost, pure silica  
477 aerogels, which are of great interest in the field of super-insulation, suffer from poor resistance to  
478 mechanical stress and thermal shock, which is holding back their large-scale industrialization and  
479 commercial success.

480 Particularly, in [57, 58], the authors characterized the effective thermal conductivity of graphene  
481 oxide (GO)-reinforced polymer aerogels produced *via* an eco-friendly freeze-drying process. Images of  
482 those aerogels are shown in Fig. 2(b). The cross-sections of aerogels on these figures reveal their high  
483 porosity. Multiscale geometrical configurations were designed based on experimental data, followed by  
484 a homogenization process from nano- to macroscale. The Milton method was applied at the nanoscopic  
485 scale, while asymptotic homogenization and finite element methods addressed micro- and mesoscopic  
486 scales, using representative unit cells to model heterogeneous structures. The approach revealed how  
487 microstructural features affect the ETC at macroscale.

488 In this work, we aim to explore the capability of deep neural networks to efficiently and accurately  
489 predict the effective thermal conductivity of various two-phase materials, such as aerogels. By lever-  
490 aging nanoscopic and microscopic scale features, DNNs can address the complexity of these materials,  
491 providing reliable predictions for designing and optimizing insulation materials. Unlike traditional  
492 multiscale homogenization methods, the application of DNNs in predicting the ETC of aerogels re-  
493 mains underexplored, highlighting the need for further research.

494 As a result, we propose incorporating a multiscale homogenization model developed in [57, 58] with  
495 deep learning techniques, namely artificial neural networks. Homogenization methods, which rely on  
496 explicit multilayer laminated structure relationships at the nanoscale and asymptotic development  
497 approaches at larger scales (ranging from hundreds of nanometers to micrometers), will be used to  
498 determine effective properties. Specifically, our approach combines homogenization on a representative  
499 volume element with a multilayer perceptron model to predict the ETC at mesoscopic and macroscopic  
500 scales. Here, DNNs replace traditional multiscale homogenization calculations, effectively mapping  
501 input features to output labels with a very high accuracy. After training, these models can predict  
502 the ETC of various aerogels and related composite materials.

503 In addition to the ease of implementation facilitated by available libraries, deep learning provides a  
504 major advantage in simplifying complex models, making them more accessible to a broader audience,

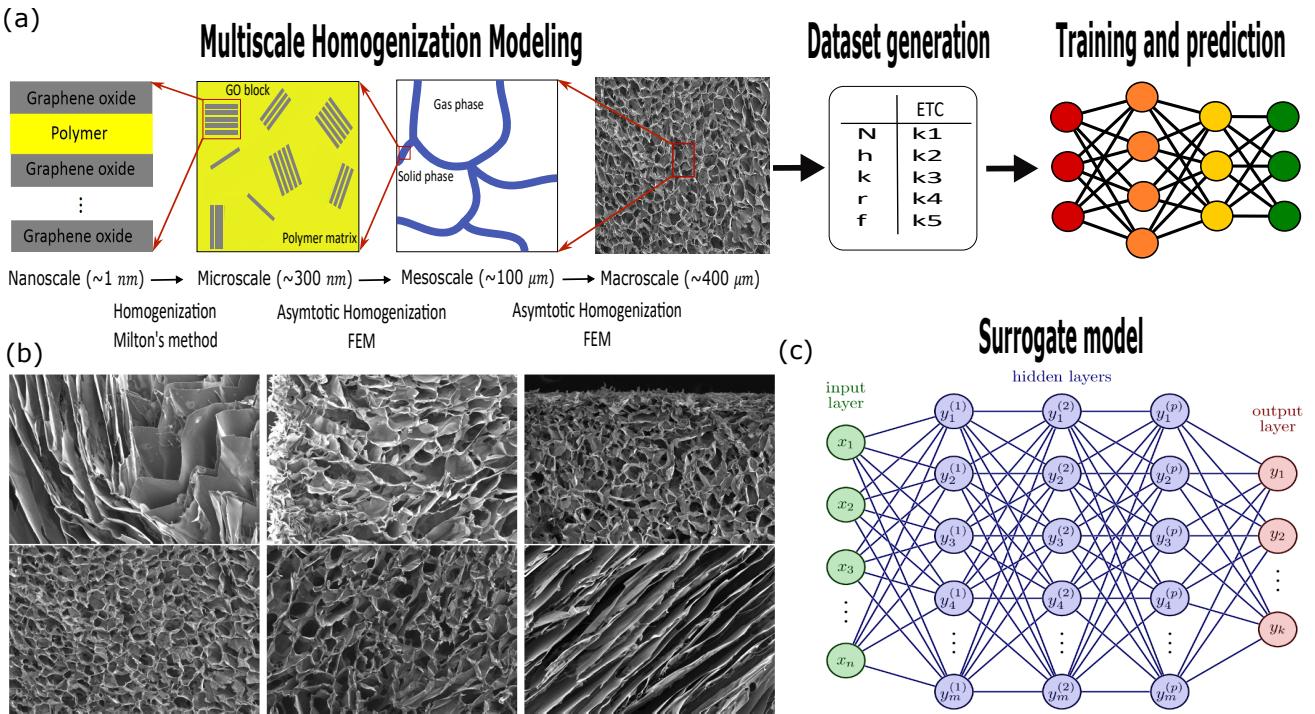


Figure 2: **Schematic illustration of the integration of multiscale homogenization modeling with deep learning for predicting the ETC of hybrid aerogels.** (a) Flowchart of the approach. (b) Scanning electron microscopy (SEM) images of the hybrid aerogels. (c) MLP models with  $n$  input parameters,  $p$  hidden layers,  $m$  neurons and  $k$  outputs. For instance, the nano-meso model is constructed with two hidden layers with the parameters defined as  $x_1 = N_C, x_2 = h_{GO}, x_3 = h_{CS}, x_4 = k_{GO}, x_5 = k_{CS}, x_6 = r_{GO}, x_7 = f_{GO}$  and  $y_1 = K_{11}^{(3)}, y_2 = K_{12}^{(3)}, y_3 = K_{13}^{(3)}, y_4 = K_{22}^{(3)}, y_5 = K_{23}^{(3)}, y_6 = K_{33}^{(3)}$ , associated with  $\hat{\Theta}_g^{(1)} = \{N_C, h_{GO}, h_{CS}\}$ ,  $\hat{\Theta}_p^{(1)} = \{k_{GO}, k_{CS}\}$ ,  $\hat{\Theta}_g^{(2)} = \{r_{GO}, f_{GO}\}$ .

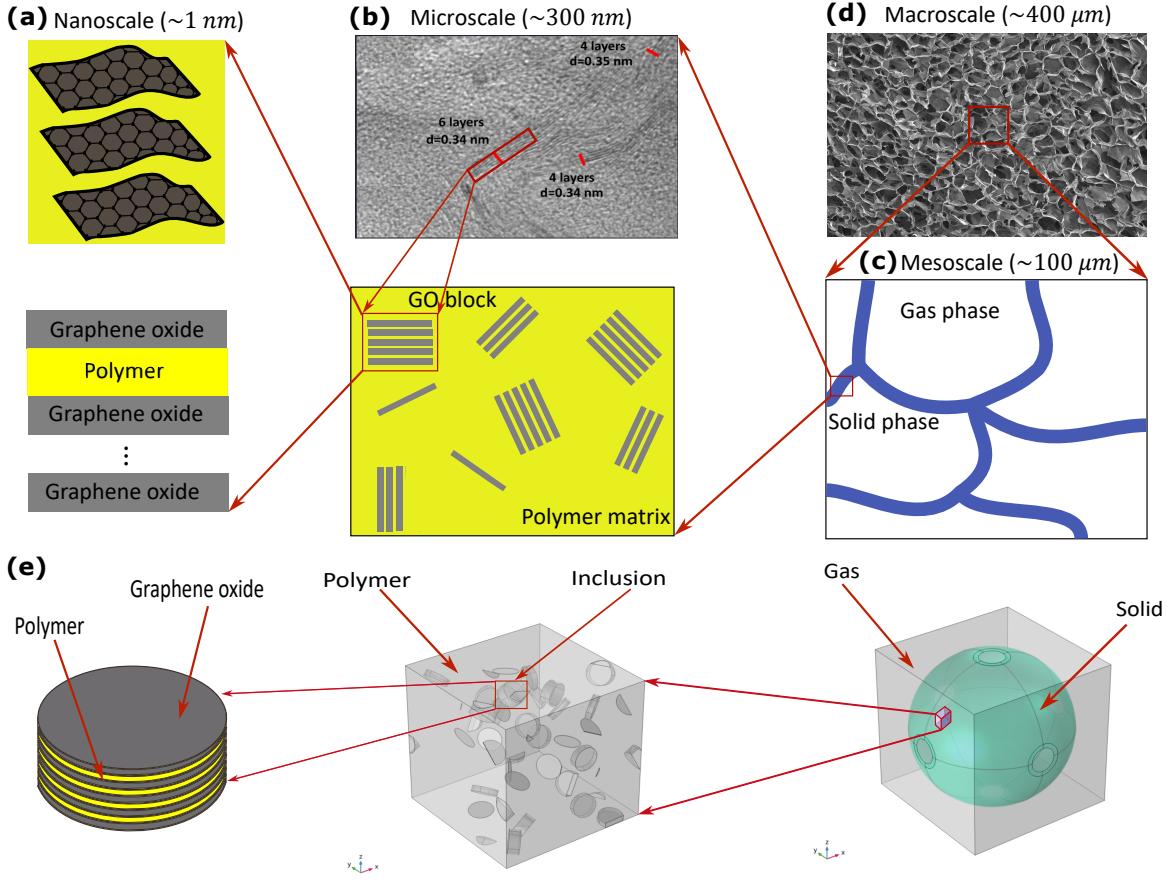
including those without expertise in numerical modeling. More specifically, by encapsulating intricate procedures in user-friendly interfaces, deep learning tools empower users to derive valuable insights simply by providing input parameters, without requiring expertise in advanced modeling techniques. This approach improves the understanding of the multiphysical behavior of composite materials and the effects of their microstructure on effective properties, ultimately improving efficiency in both the study and fabrication of these materials.

First, in section 3.1.1, we describe an MLP model designed to predict the ETC of elaborated hybrid aerogels at mesoscale. Similarly, in section 3.1.2, we introduce an MLP model that provides a prediction of the ETC of aerogels at macroscale. Those models enable the immediate determination of the ETC without requiring multiscale homogenization calculations. Furthermore, both MLP models demonstrate high prediction accuracy, highlighting the advantages of using deep neural networks to predict the ETC of elaborated hybrid aerogels.

As mentioned in Section 2.1.1, the structure of composite aerogels can be analyzed across four distinct scales. At the nanoscale, the structure consists of intercalated layers of graphene oxide and polymer. At the microscale, graphene oxide blocks are dispersed within the chitosan matrix to form the solid phase of the aerogel. Moving to the mesoscale, the material is composed of gas phase and solid phase. Finally, the macroscale consists of alveolar structures with interconnected pores. The flowchart from nano to macroscale is summarized in Fig. 3.

Before the training begins, a data generation process is carried out to produce the necessary dataset. More specifically, the data will be obtained through a multiscale homogenization calculation by Milton's method and an asymptotic homogenization method with finite elements method, as shown in Fig. 2(a).

In particular, an automated procedure is set up for the data generation process based on the multiscale homogenization models developed in [57, 58]. Those computations are performed across the four scales: nanoscopic, microscopic, mesoscopic, and macroscopic scales. To predict the ETC at



**Figure 3: Geometrical analysis of hybrid aerogels across nano-micro-meso-macroscales.** (a) Nanoscale with intercalated layers of graphene oxide and polymer. (b) Microscale where graphene oxide blocks are dispersed within the chitosan matrix. (c) Mesoscale includes gas phase and solid phase. (d) Macroscale formed by alveolar structures with interconnected pores. (e) 3D numerical simulation of nano-micro-mesocales with dissimilar inclusions in the polymer matrix at microscale (middle) and geometry of the representative unit cell at the mesoscale (right).

mesoscale, we generate the dataset  $(\hat{\Theta}_g^{(1)}, \hat{\Theta}_p^{(1)}, \hat{\Theta}_g^{(2)}, \hat{\Theta}_p^{(2)}, \hat{\Theta}_p^{(3)})$  where  $\hat{\Theta}_p^{(3)}$  is the ETC at mesoscale. Similarly, to predict the ETC at macroscale  $\hat{\Theta}_p^{(4)}$ , a dataset  $(\hat{\Theta}_g^{(3)}, \hat{\Theta}_p^{(3)}, \hat{\Theta}_p^{(4)})$  is generated. The number of samples are generated from 1000 to 1200, corresponding to the models developed in section 3.1 and section 3.2. The following discussion outlines the main steps of the procedure for calculating the ETC at mesoscale.

- **Step 1.** Generate random values for the parameters  $(\hat{\Theta}_g^{(1)}, \hat{\Theta}_p^{(1)}, \hat{\Theta}_g^{(2)})$  for nanoscopic and microscopic homogenization computations. The value ranges of the parameters are given in more details in section 3.1 and section 3.2.
- **Step 2.** Generate random inclusions in a cube. For the microscale, a method involving two processes in Matlab and Comsol Multiphysics environment was utilized to create random disc inclusions and assign distinct effective thermal conductivity to each inclusion. First, a Matlab script was developed to randomly generate ellipsoid inclusions within a cube of a given side length, ensuring a desired volume fraction. As detailed in [57], a specified number of ellipsoid inclusions with an initial null volume are randomly created with random velocities, angular velocities, and orientations. These inclusions are then set into motion, with their volumes progressively increasing from zero. During each iteration, two events can occur: binary collisions between inclusions or collisions between inclusions and the cube boundaries. In the case of binary collisions, the velocities and angular velocities of the involved inclusions are updated. For

548 collisions with the cube boundaries, the periodic part of the inclusion is created on the opposite  
 549 side, following periodic boundary conditions. The simulation ends once the prescribed volume  
 550 fraction is achieved. Subsequently, discs replace the ellipsoids without altering their position,  
 551 orientation, aspect ratio, and volume. Using the thickness of the graphene oxide sheet as a  
 552 reference, the model is rescaled to its actual size.

- 553 • **Step 3.** The effective thermal conductivity of each inclusion ( $\hat{\Theta}_p^{(2)}$ ) is calculated by solving the  
 554 Eq. (A.1) using explicit formulas from Milton method, as described in the Appendix.
- 555 • **Step 4.** The procedure is proceeded by solving the cell problems (A.6)-(A.9) by the finite  
 556 element method in Comsol Multiphysics with the input parameters ( $\hat{\Theta}_g^{(2)}, \hat{\Theta}_p^{(2)}$ ) to obtain  $\hat{\Theta}_p^{(3)}$ ,  
 557 which is the ETC at mesoscale.

558 For the calculation of ETC at macroscale, the procedure is much simpler as follow.

- 559 • **Step 1.** Generate random values for the parameters ( $\hat{\Theta}_g^{(3)}, \hat{\Theta}_p^{(3)}$ ) for mesoscopic asymptotic  
 560 homogenization. More details of value ranges of the parameters are given in section 3.1 and  
 561 section 3.2.
- 562 • **Step 2.** The effective thermal conductivity  $\hat{\Theta}_p^{(4)}$  is obtained by solving the cell problems (A.6)-  
 563 (A.9) by the finite element method in Comsol Multiphysics with the input parameters ( $\hat{\Theta}_g^{(3)}, \hat{\Theta}_p^{(3)}$ ).  
 564 Further details are provided in [57, 58].

565 The calculation procedure described above can be summarized in the Algorithms 1 and 2.

---

#### Algorithm 1: data generation for Nano-Meso problem

---

```

1 i = 0
2 while  $i < N_{sample}$  do
3   while Desired volume fraction not reached do
4     Randomly generate parameters ( $\hat{\Theta}_g^{(1)}, \hat{\Theta}_p^{(1)}, \hat{\Theta}_g^{(2)}$ )
5     Inclusion generating
6     Solving Eq. (6) by Milton method for  $\hat{\Theta}_p^{(2)}$ 
7     Solving Eqs. (A.6)-(A.9) by FEM for  $\hat{\Theta}_p^{(3)}$ 
8     Save ( $\hat{\Theta}_g^{(1)}, \hat{\Theta}_p^{(1)}, \hat{\Theta}_g^{(2)}, \hat{\Theta}_p^{(2)}, \hat{\Theta}_p^{(3)}$ )
9      $i = i + 1$ 
```

---



---

#### Algorithm 2: data generation for Meso-Macro problem

---

```

1 i = 0
2 while  $i < N_{sample}$  do
3   Randomly generate parameters ( $\hat{\Theta}_g^{(3)}, \hat{\Theta}_p^{(3)}$ ) (from mesoscale)
4   Solving Eqs. (A.6)-(A.9) by FEM for  $\hat{\Theta}_p^{(4)}$ 
5   Save ( $\hat{\Theta}_g^{(3)}, \hat{\Theta}_p^{(3)}, \hat{\Theta}_p^{(4)}$ )
6    $i = i + 1$ 
```

---

568 Concerning the randomness in generating inclusions, we set “rng shuffle” in Matlab, it is a com-  
 569 mand used to initialize the random number generator (RNG) using a seed based on the current time,  
 570 effectively making the random number sequence different in each program execution. Additionally,  
 571 all random parameters are generated using the command  $(b-a)*rand() + a$  in Matlab which generates  
 572 uniformly distributed random numbers on an interval  $(a, b)$ .

573 *3.1. Predicting the ETC of composite aerogels with similar inclusions*

574 After generating the datasets using Algorithms 1 and 2, we proceed with training and predic-  
 575 tion. We first focus on scenarios involving similar inclusions within the polymer matrix. Section 3.1.1

576 presents the construction, training, and evaluation of the MLP model at the mesoscale, while Section 3.1.2 addresses the corresponding model at the macroscale.

### 578 3.1.1. MLP model at mesoscale

579 In this section, we construct an artificial neural network model, namely an MLP model, based  
580 on a three-scale (nano-micro-meso) homogenization model that can predict the ETC of elaborated  
581 hybrid aerogels at mesoscale. The inputs encompass characteristic nanoscopic parameters such as the  
582 number of graphene oxide sheets per block, the thickness of a graphene oxide sheet, the thickness of a  
583 polymer sheet, the thermal conductivity of a graphene oxide sheet, and the thermal conductivity of the  
584 polymer. Additionally, microscopic parameters include the ratio between the diameter and thickness  
585 of a graphene oxide sheet together with the volume fraction of graphene oxide. Homogenization  
586 calculations by Milton method at nanoscale and FEM at microscale are replaced by the MLP model.  
587 The results show the high prediction accuracy of the proposed MLP model with the output considered  
588 as a spherical tensor of order 2.

589 In order to train an artificial neural network model effectively, it is essential to start with a solid  
590 foundation: a comprehensive and reliable dataset. This dataset needs to encompass a diverse array  
591 of relevant input features. A crucial aspect of preparing this dataset is feature engineering, which  
592 entails making informed decisions about which input features are most pertinent and should thus be  
593 incorporated into the dataset.

Scale	Parameter	Notation	Value
Nanoscale	Number of GO sheets per block	$N_c$	{1,3,5}
	Thickness of a GO sheet (m)	$h_{GO}$	$10^{-9}$
	Thickness of a polymer sheet (m)	$h_{CS}$	$[1.5, 5.3] \times 10^{-9}$
	Thermal conductivity of an GO sheet, isotropic (W/(m·K))	$k_{GO}$	[16, 20]
	Thermal conductivity of polymer, isotropic (W/(m·K))	$k_{CS}$	[0.1, 0.5]
Microscale	Ratio between the diameter and the thickness of an GO sheet	$r_{GO}$	[10, 100]
	Volume fraction of GO	$f_{GO}$	[0.004, 0.03]

Table 1: Values and notations of the parameters at nano- and microscales for similar inclusions scenarios.

594 In this part, the input data, or features, will include all parameters from the nanoscale and mi-  
595 croscale covering geometrical configuration and thermal conductivity of GO and polymer sheets, as  
596 described in Tab. 1.

597 Specifically, the 5 parameters ( $N_c, h_{GO}, h_{CS}, k_{GO}, k_{CS}$ ) are used following the analytical formulas  
598 in Milton method, as described in the Appendix section 5, with the normal  $\mathbf{n} = \mathbf{e}_3 = (0, 0, 1)$  to obtain  
599 the effective thermal conductivity tensors ( $\mathbb{K}_{1M}^{(2)}, \mathbb{K}_{3M}^{(2)}, \mathbb{K}_{5M}^{(2)}$ ) corresponding to the cases  $N_c = 1, 3, 5$ ,  
600 respectively. In other words, the effective thermal conductivity tensor at microscale can be written as

$$\mathbb{K}_{N_c M}^{(2)} = \begin{bmatrix} K_{11} & 0 & 0 \\ 0 & K_{11} & 0 \\ 0 & 0 & K_{33} \end{bmatrix},$$

601 in which for  $N_c = 1$ , we obtain the isotropic ETC of the monolayer GO with  $K_{11} = K_{33} \in [16, 20]$ .  
602 For  $N_c = 3$ , we obtain the transverse isotropic ETC of the block GO with  $K_{11} \in [3.61, 10.25], K_{33} \in$   
603  $[0.13, 0.98]$ . And similarly for  $N_c = 5$ , we obtain the transverse isotropic ETC of the block GO with  
604  $K_{11} \in [3.13, 9.36], K_{33} \in [0.12, 0.90]$ . The effective thermal conductivity tensor at microscale  $\mathbb{K}_{N_c M}^{(2)}$   
605 is then used together with the 2 parameters  $r_{GO}$  and  $f_{GO}$  through an asymptotic homogenization  
606 procedure with FEM to obtain the effective thermal conductivity tensor of the solid phase at mesoscale:

$$\mathbb{K}^{(3)} = \begin{bmatrix} K_{11}^{(3)} & K_{12}^{(3)} & K_{13}^{(3)} \\ K_{21}^{(3)} & K_{22}^{(3)} & K_{23}^{(3)} \\ \text{sym} & & K_{33}^{(3)} \end{bmatrix}. \quad (18)$$

607 Since  $K_{ij}^{(3)} \ll K_{ii}^{(3)}, i, j = 1, 2, 3$ , the terms  $K_{ij}^{(3)}, i \neq j$  can be neglected, on the other hand,  $K_{11}^{(3)} \sim$   
608  $K_{22}^{(3)} \sim K_{33}^{(3)}$  so (18) can be expressed as a spherical tensor  $\mathbb{K}^{(3)} = K^{(3)}\mathbb{I}$  (where  $\mathbb{I}$  denotes the second-  
609 order identity tensor) and  $K^{(3)} = (K_{11}^{(3)} + K_{22}^{(3)} + K_{33}^{(3)})/3$ .

Without loss of generality, we consider the effective thermal conductivity of the solid phase at mesoscale to be a symmetric second-order tensor. Therefore, the outputs of the MLP model can be represented by the vector  $\mathbf{K}^{(3)}$ , which is expressed as

$$\mathbf{K}^{(3)} = \begin{bmatrix} K_{11}^{(3)} & K_{12}^{(3)} & K_{13}^{(3)} & K_{21}^{(3)} & K_{23}^{(3)} & K_{33}^{(3)} \end{bmatrix}.$$

Once an appropriate training dataset and network architecture are chosen, the artificial neural network undergoes training with continuous monitoring of the loss function. In this study, we utilize the Keras library implemented in Python to construct and train the MLP model.

In this paper, the best choices of hyperparameters were determined through empirical experimentation. Various numerical tests were performed by adjusting hyperparameters such as the number of hidden layers, number of neurons, batch size, maximum number of epochs, loss function, and optimizer. The optimal configuration was selected based on its ability to achieve good training performance, specifically, convergence of the loss function within a small number of epochs, without signs of overfitting or underfitting. Therefore, in this section, the MLP model is constructed with two hidden layers, each containing 20 neurons. The structure of the neural network is illustrated in Fig. 2(c). As discussed in section 2.2.2, the ReLU activation function is used thank to its simplicity and efficiency. The batch size is set to 1, and the training process is allowed to run for a maximum of 10 000 epochs. The input/output parameters and hyperparameters used for training are presented in Tab. 2. The MSE is employed as the loss function, and the Adam optimizer is used with the following parameters ( $\alpha = 0.001, \beta_1 = 0.9, \beta_2 = 0.999, \epsilon = 10^{-7}$ ). The training and validation processes are evaluated using the MAE. The dataset is divided into training, validation, and test sets with a ratio of 0.6:0.2:0.2.

A small dataset size can introduce several potential biases and limitations. One major concern is overfitting, where the model learns the training data too well, including noise or irrelevant patterns, which results in poor performance on new or unseen data. On the other hand, limited data may lead to biased predictions when applied to broader or more diverse inputs, as the model might not capture the full variability of the underlying problem. Additionally, small datasets can also lead to underfitting, where the model fails to learn complex relationships and true dependencies in the data. In this work, however, the dataset sizes were selected to ensure good training performance while minimizing these risks. As a consequence, a total of 1200 samples are randomly generated according to the procedure described in section 2.2.2. These samples are divided into three groups: 400 samples with  $N_c = 1$ , 400 samples with  $N_c = 3$ , and 400 samples with  $N_c = 5$ . For each group, the number of inclusions in the polymer matrix is set to 10. The thickness of a GO sheet ( $h_{GO}$ ) is fixed at  $10^{-9}$  meters. Other parameters, including the thickness of a polymer sheet ( $h_{CS}$ ), the thermal conductivity of a GO sheet ( $k_{GO}$ ), the thermal conductivity of the polymer ( $k_{CS}$ ), the aspect ratio of a GO sheet ( $r_{GO}$ ), and the volume fraction of GO ( $f_{GO}$ ), are randomly assigned using a uniform distribution within the intervals specified in Tab. 1.

Input layer	Hidden layer	Output layer	Loss function	Optimizer
$(N_c, h_{GO}, h_{CS}, \dots, k_{GO}, k_{CS}, r_{GO}, f_{GO})$	$2 \times 20 \times \text{ReLU}$	$(K_{11}^{(3)}, K_{12}^{(3)}, K_{13}^{(3)}, K_{21}^{(3)}, K_{23}^{(3)}, K_{33}^{(3)})$	MSE	Adam

Table 2: Architecture of the MLP model for prediction of ETC at mesoscale for similar inclusions scenarios.

Figure 4 illustrates the evolution of the loss function (see Fig. 4(a)) and the MAE (see Fig. 4(b)) as functions of the number of epochs, for both training and validation sets. The proposed neural network model demonstrates a rapid convergence, with the loss function decreasing significantly due to continuous learning and self-adjustment. Ultimately, the final MAE and loss function MSE values for the training set are 0.0059 and  $8.95 \times 10^{-5}$ , respectively, while for the validation set, they are 0.0066 and  $8.62 \times 10^{-5}$ .

To assess the precision of our proposed MLP model, we present in Fig. 4(c) the relative errors for the predicted components of effective thermal conductivity  $k_{11}^e, k_{22}^e, k_{33}^e$  and  $k_{mean}^e$  at 240 selected data points. In particular, the majority of relative errors are below 0.1 in all cases. Furthermore, Figure 4(d) illustrates the correlations between predicted and exact values, showing an exceptionally

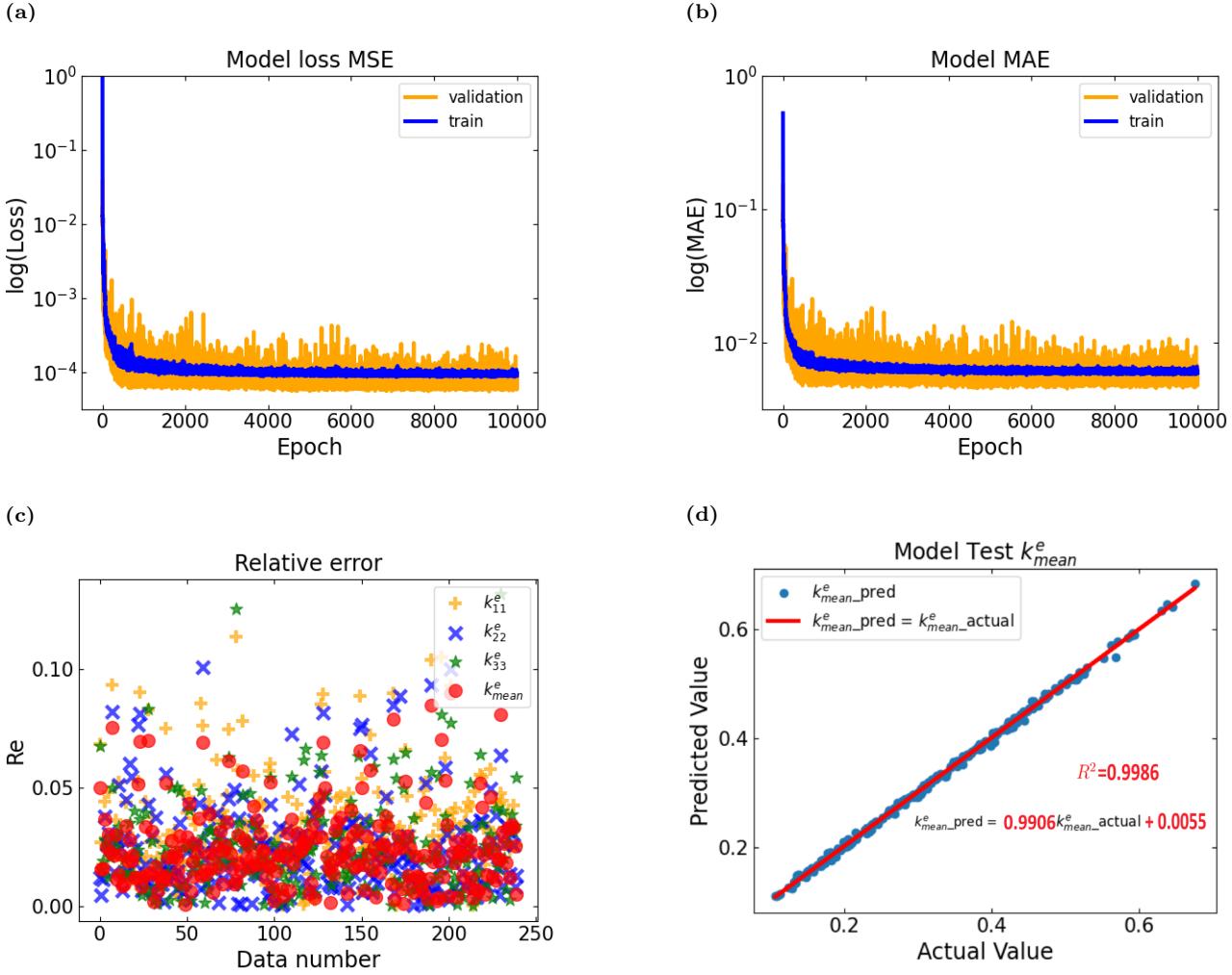


Figure 4: **Nano-Meso problem with similar inclusions.** (a) Evolution of loss MSE as a function of number of epochs between training set and validation set. (b) Evolution of MAE as a function of number of epochs between training set and validation set. (c) Relative error between predicted and exact values of predicted effective thermal conductivity coefficients  $k_{11}^e, k_{22}^e, k_{33}^e, k_{mean}^e = (k_{11}^e + k_{22}^e + k_{33}^e)/3$ . (d) Correlation between predicted and exact values.

high  $R^2$  value of 0.9986 for  $k_{mean}^e$ . These results show the robust predictive accuracy of our MLP model.

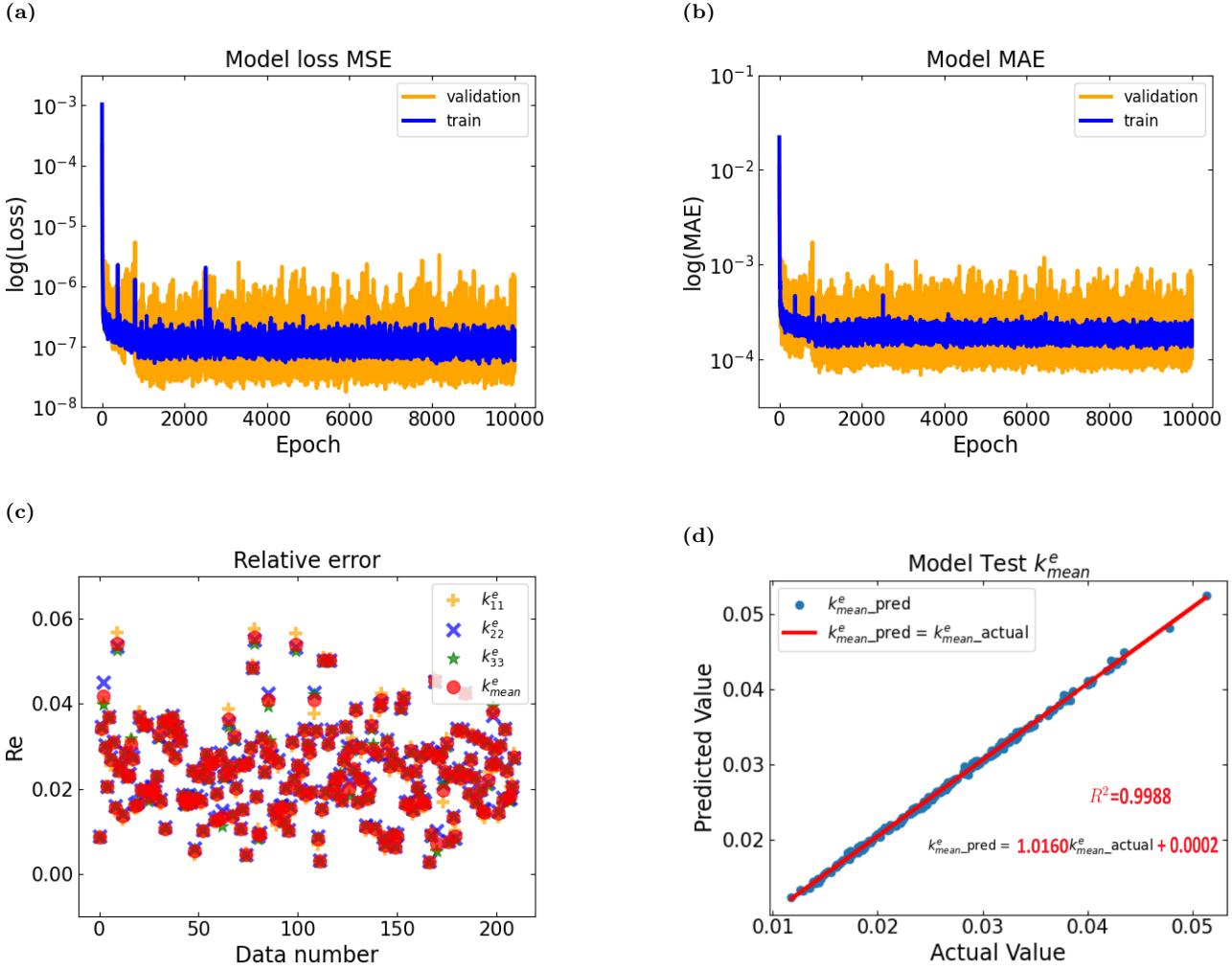


Figure 5: **Meso-Macro problem with similar inclusions.** (a) Evolution of loss MSE as a function of number of epochs between training set and validation set. (b) Evolution of MAE as a function of number of epochs between training set and validation set. (c) Relative error between predicted and exact values of predicted effective thermal conductivity coefficients  $k_{11}^e, k_{22}^e, k_{33}^e, k_{mean}^e = (k_{11}^e + k_{22}^e + k_{33}^e)/3$ . (d) Correlation between predicted and exact values.

### 656 3.1.2. MLP model at macroscale

657 In this section, we build another MLP model utilizing a two-scale (meso-macro) asymptotic ho-  
 658 mogenization model to predict the ETC of elaborated hybrid aerogels at the macroscale. In this study,  
 659 the representative volume element is modeled as a hollow sphere with 6 holes its surface to preserve  
 660 the interconnected nature of the pore network, as can be seen in Fig. 3(e), with the matrix (green)  
 661 and the pore (gray) phases. To simplify the complex morphology of the pores while maintaining a  
 662 representative geometric form, the pores are approximated as hollow spheres. This assumption aligns  
 663 closely with observed structural features and facilitates more efficient modeling and numerical im-  
 664 plementation. The mesoscopic inputs consist of the geometrical configuration of the sphere and the  
 665 thermal conductivity of the solid phase, which is derived from microscopic results. Computational  
 666 procedure by the asymptotic homogenization with FEM at mesoscale are substituted by the MLP  
 667 model. Similarly, the proposed model demonstrates a very high degree of accuracy in its predic-  
 668 tion with the output also considered as a spherical tensor of order 2. This shows the effectiveness  
 669 of coupling DNNs and multiscale homogenization models, opening up many potential interesting re-  
 670 search directions in which machine learning methods may show their superior advantages compared  
 671 to deterministic methods, particularly in handling and processing large volumes of data efficiently.

672 Specifically, at the mesoscale, we consider the RVE as a hollow sphere with six holes on its surface.

673 These holes ensure an interconnected pore network within the aerogels. The wall thickness of the sphere  
674 can be adjusted to achieve the desired porosity. Given that aerogels are highly porous structures, this  
675 work focuses on porosities ranging from 95% to 99.99%. Additionally, the gas phase considered here  
676 includes not only stationary air, with a thermal conductivity of 0.026 W/(m·K), but also various other  
677 types of gases with thermal conductivities ranging from 0.01 to 0.03 W/(m·K). Thus, the mesoscopic  
678 inputs include the sphere's diameter, the wall thickness of the sphere, the diameter of the six holes,  
679 the thermal conductivity of the gas phase, and the thermal conductivity of the solid phase, which is  
derived from microscopic results, as described in Tab. 3.

Scale	Parameter	Notation	Value
Mesoscale	Diameter of the sphere (m)	$D_s$	$[60, 200] \times 10^{-6}$
	Wall thickness of the sphere (m)	$t_s$	$[0.34, 1.88] \times 10^{-6}$
	Diameter of the 6 holes on the wall of the sphere (m)	$D_h$	$[6.51, 40] \times 10^{-6}$
	Thermal conductivity of the gas phase, isotropic (W/(m·K))	$k_g$	$[0.01, 0.03]$
	Thermal conductivity of the solid phase, isotropic (W/(m·K))	$k_s$	$[0.107, 0.712]$

Table 3: Values and notations of the parameters at mesoscale for similar inclusions scenarios.

680 Similar to section 3.1.1, the effective thermal conductivity tensor at macroscale obtained through  
681 an asymptotic homogenization procedure with FEM can be expressed as:

$$\mathbb{K}^{(4)} = \begin{bmatrix} K_{11}^{(4)} & K_{12}^{(4)} & K_{13}^{(4)} \\ & K_{22}^{(4)} & K_{23}^{(4)} \\ \text{sym} & & K_{33}^{(4)} \end{bmatrix}. \quad (19)$$

683 Here,  $K_{ij}^{(4)} \ll K_{ii}^{(4)}$ ,  $i, j = 1, 2, 3$ , the terms  $K_{ij}^{(4)}, i \neq j$  can be neglected, additionally,  $K_{11}^{(4)} \sim K_{22}^{(4)} \sim$   
684  $K_{33}^{(4)}$  so (19) can be expressed as a spherical tensor  $\mathbb{K}^{(4)} = K^{(4)}\mathbb{I}$  (where  $\mathbb{I}$  represents the second-order  
685 identity tensor) and  $K^{(4)} = (K_{11}^{(4)} + K_{22}^{(4)} + K_{33}^{(4)})/3$ .

686 Similarly, without loss of generality, we consider the effective thermal conductivity at the macroscale  
687 to be a symmetric second-order tensor. Thus, the outputs of the MLP model can be characterized by  
688 the vector  $\mathbf{K}^{(4)}$ , which is written as

$$\mathbf{K}^{(4)} = [K_{11}^{(4)} \quad K_{12}^{(4)} \quad K_{13}^{(4)} \quad K_{22}^{(4)} \quad K_{23}^{(4)} \quad K_{33}^{(4)}].$$

689 The MLP model is specifically designed to predict the components of the effective thermal con-  
690 ductivity tensor, which is a symmetric positive definite second-order tensor. To enforce symmetry, the  
691 network outputs only the six independent components corresponding to the upper triangular part of  
692 the tensor, thereby ensuring a symmetric structure by design. Although explicit physical constraints  
693 are not embedded within the MLP architecture, positive definiteness is consistently observed in the  
694 predictions. This is supported by the empirical observation that the diagonal components are signif-  
695 icantly larger than the off-diagonal ones, aligning with the known property that symmetric, strictly  
696 diagonally dominant matrices with positive diagonal entries are positive definite. Importantly, the  
697 MLP is trained on high-fidelity data generated from asymptotic homogenization, a rigorous multi-  
698 scale method that inherently captures key physical principles such as heterogeneity, anisotropy, and  
699 conservation of heat flux. As a result, the MLP learns a mapping grounded in physically consistent  
700 behavior.

701 In this section, we perform a normalization process to scale all features to a common range of  
702  $[0, 1]$ . This step is necessary due to the different ranges and units of the input features. Features  
703 with larger scales can dominate the objective function, leading to sub-optimal model performance.  
704 Without normalization, features with larger numerical ranges may unduly influence the model's pre-  
705 dictions compared to those with smaller ranges. Therefore, scaling ensures that all features contribute  
706 proportionately. Normalization is also crucial because machine learning algorithms using gradient  
707 descent optimization converge much faster when the input features are normalized or standardized.  
708 The optimization process works better when features have similar ranges. As a result, in this work, we

709 use the MinMaxScaler tool from the scikit-learn library to normalize the input features. Specifically,  
 710 the input features are normalized using the formula:

$$\hat{\mathbf{x}}_i = \frac{\mathbf{x}_i - \mathbf{x}_{\min}}{\mathbf{x}_{\max} - \mathbf{x}_{\min}}, \quad (20)$$

711 where  $i$  is the serial number of the dataset,  $\mathbf{x}_i$  is the actual value of the data,  $\hat{\mathbf{x}}_i$  is the normalized value,  
 712 and  $\mathbf{x}_{\min}$  and  $\mathbf{x}_{\max}$  are the minimum and maximum values of  $\mathbf{x}_i$  in the data samples, respectively.

713 A total of 1054 samples are randomly generated with a uniform distribution following the intervals  
 714 specified in Tab. 3. The neural network model, as illustrated in Fig. 2(c), has 2 hidden layers with  
 715 16 neurons each, a batch size of 1, and is trained for 10 000 epochs using the MSE loss function and  
 716 the Adam optimizer. The training and validation processes are evaluated in terms of the MAE. The  
 717 dataset is divided into training, validation, and test sets with a ratio of 0.6:0.2:0.2.

Input layer	Hidden layer	Output layer	Loss function	Optimizer
$(D_s, t_s, D_h, k_g, k_s)$	$2 \times 16 \times \text{ReLU}$	$(K_{11}^{(4)}, K_{12}^{(4)}, K_{13}^{(4)}, K_{22}^{(4)}, K_{23}^{(4)}, K_{33}^{(4)})$	MSE	Adam

Table 4: Architecture of the MLP model for prediction of ETC at macroscale for similar inclusions scenarios.

718 The evolution of the loss function (see Fig. 5(a)) and the MAE (see Fig. 5(b)) for the training and  
 719 validation sets are illustrated as functions of the number of epochs in Fig. 5. The proposed neural  
 720 network model converges quickly, with the loss function dropping rapidly after continuous learning  
 721 and self-adjustment. The final MAE and loss function MSE values are 0.00018 and  $1.03 \times 10^{-7}$  for  
 722 the training dataset, and 0.00029 and  $2.37 \times 10^{-7}$  for the validation dataset.

723 To evaluate the accuracy of the proposed MLP model, Figure 5(c) shows the relative errors of  
 724 the predicted effective thermal conductivity components  $k_{11}^e, k_{22}^e, k_{33}^e$  and  $k_{mean}^e$  for 210 selected data  
 725 points. The relative errors are less than 0.06 for all cases. The correlation between the predicted  
 726 and exact values is plotted in Fig. 5(d), with a very high  $R^2$  value of 0.9988 for  $k_{mean}^e$ . These results  
 727 demonstrate the high prediction accuracy of the proposed MLP model.

728 In this section and the previous, we focused on predicting the ETC of elaborated hybrid aerogels  
 729 at mesoscopic and macroscopic scales with identical inclusions, using the multiscale homogenization  
 730 model developed in [57, 58]. This means that we considered scenarios where the inclusions within  
 731 the polymer matrix at the microscale are similar. Specifically, there are three types of inclusions:  
 732 exfoliated structures (monolayer graphene oxide) and intercalated structures (block-3 and block-5  
 733 graphene oxide). These inclusions differ in their geometrical configurations, defined by the number of  
 734 graphene oxide sheets in each block, as well as in their effective properties. In section 3.2, we examine  
 735 a multiscale homogenization model with dissimilar inclusions, which represents a different approach  
 736 from the one in [57, 58]. This model allows for more general scenarios where the inclusions are varied  
 737 in type and quantity within the polymer matrix, leading to diverse and interesting results.

### 738 3.2. Predicting the ETC of elaborated hybrid aerogels with dissimilar inclusions

739 In this section, we first briefly introduce in section 3.2.1 the idea of developing a new multiscale  
 740 homogenization approach with dissimilar inclusions. Then, we present MLP models for predicting the  
 741 effective thermal conductivity of hybrid aerogels with dissimilar inclusion scenarios at both mesoscale  
 742 and macroscale in section 3.2.2 and section 3.2.3, respectively.

#### 743 3.2.1. A new multiscale homogenization model with dissimilar inclusions

744 In the preceding section, the multiscale homogenization model presented in [57, 58] was limited to  
 745 scenarios at the microscale where all inclusions within the polymer matrix were identical. Specifically,  
 746 this included cases featuring solely monolayer graphene oxide, block-3 graphene oxide, or block-5  
 747 graphene oxide inclusions. In such scenarios, where similar inclusions uniformly dispersed at the  
 748 microscale, the RVE was also assumed to have a homogeneous property distribution.

749 In this section, we introduce a new multiscale homogenization model that considers dissimilar  
 750 inclusions at the microscale. This more comprehensive model allows for the simultaneous presence  
 751 of all three types of inclusions within the polymer matrix, creating broader scenarios where varying

types and quantities of inclusions coexist. Additionally, this approach also introduces heterogeneity in the property distribution of the RVE at mesoscale, which provides a more realistic representation of aerogel materials.

By adopting an RVE with heterogeneous property distribution at the mesoscale, we enhance the accuracy of predicting the effective thermal conductivity of aerogels. This methodology aligns computational models more closely with real-world material behaviors observed in experimental studies. Ultimately, it advances both the understanding and the application of aerogels in thermal management technologies by capturing the complexities of their internal structures and thermal properties more effectively.

The new multiscale homogenization model involves several key steps for randomly generating dissimilar inclusions within a cube, summarized as follows:

- **Step 1.** Random Generation of Inclusions: Initially,  $N$  inclusions of type block-5 graphene oxide are randomly generated within the cube, similar to scenarios involving similar inclusions. This step determines various properties of the inclusions such as their positions and orientations.
- **Step 2.** Assignment of Inclusion Thickness: Each inclusion's thickness is randomly adjusted corresponding to the type of inclusions (monolayer, block-3 or block-5 GO) and assigned within the cube. This step adds variability to the structural composition of the material.
- **Step 3.** Microscale Thermal Conductivity Calculation: Milton method is employed to calculate the effective thermal conductivity of each inclusion at the microscale.
- **Step 4.** Mesoscale Homogenization and FEM Analysis: The RVE's geometry with randomly generated inclusions, is imported into Comsol Multiphysics. An asymptotic homogenization procedure is subsequently applied. Finite element method analysis using Comsol Multiphysics is employed to determine the effective thermal conductivity of the solid phase at mesoscale. This phase integrates the contributions of all inclusions to predict the overall thermal properties of the material.

By considering dissimililar inclusions, the effective thermal conductivity is observed to be a second-order spherical tensor, which can be expressed as  $\mathbb{K}^{(3)} = K^{(3)}\mathbb{I}$ .

In the next section, we construct an MLP model at mesoscale to predict the effective thermal conductivity of hybrid aerogels with dissimilar inclusions.

### 3.2.2. MLP model at mesoscale with dissimilar inclusions

In this section, we develop an artificial neural network model, specifically an MLP model, using the new three-scale (nano-micro-meso) homogenization model described in section 3.2.1. This model is designed to predict the effective thermal conductivity of hybrid aerogels at the mesoscale for scenarios involving dissimilar inclusions. By employing this more comprehensive model, we can examine a wide range of cases, and the MLP model is expected to provide highly accurate and realistic predictions compared to the model presented in section 3.1.1.

Furthermore, this approach allows us to consider playing with random parameters, namely the number of each type of inclusions, to generate different effective properties which can be exploited at the mesoscopic scale. In such scenarios, where dissimilar inclusions lead to increased complexity and randomness, deep learning methods may be advantageous. These methods can effectively handle the inherent randomness and intricacies present in the model, enhancing the fidelity of outputs and the prediction ability of the models. This is a promising research direction which will enable opportunities to potentially obtain results which are new and interesting.

In this part, we study scenarios where the three types of inclusions are combined within the polymer matrix, as shown in Fig. 3(e), here for instance with 11 inclusions of monolayer type, 11 inclusions of block-3 graphene oxide type and 8 inclusions of block-5 graphene oxide type. A new multiscale homogenization model was developed in section 3.2.1. This new model could produce interesting results since dissimilar inclusions can disrupt the uniformity of the thermal conduction pathways, leading to variations in thermal conductivity. Similar to section 3.1, an MLP model was constructed using the new multiscale homogenization model to predict the ETC of hybrid aerogels at mesoscale.

Scale	Parameter	Notation	Value
Nanoscale	Number of monolayer GO	$N_m$	[0,30]
	Number of block-3 GO	$N_{b3}$	[0,30]
	Number of block-5 GO	$N_{b5}$	[0,30]
	Thickness of a GO sheet (m)	$h_{GO}$	$10^{-9}$
	Thickness of a polymer sheet (m)	$h_{CS}$	$[1.5, 5.3] \times 10^{-9}$
	Thermal conductivity of an GO sheet, isotropic (W/(m·K))	$k_{GO}$	[16, 20]
Microscale	Thermal conductivity of polymer, isotropic (W/(m·K))	$k_{CS}$	[0.1, 0.5]
	Ratio between the diameter and the thickness of an GO sheet	$r_{GO}$	[10, 100]
Microscale	Volume fraction of GO	$f_{GO}$	[0.004, 0.03]

Table 5: Values and notations of the parameters at nanoscale and microscale with dissimilar inclusions.

802 Specifically, the number of GO sheets per block ( $N_c$ ) is replaced with three parameters  $N_m$ ,  $N_{b3}$   
803 and  $N_{b5}$ , representing the number of monolayer GO inclusions, block-3 GO inclusions, and block-5  
804 GO inclusions, respectively. The total number of inclusions ( $N_{inc}$ ) is randomly assigned within the  
805 interval [10, 30]. The number of each type of inclusion is also randomly generated within the interval  
806 [0, 30], ensuring that  $N_m + N_{b3} + N_{b5} = N_{inc}$ . All other parameters remain the same with the similar  
807 inclusion scenarios, as described in Tab. 5.

Input layer	Hidden layer	Output layer	Loss function	Optimizer
$(N_m, N_{b3}, N_{b5}, h_{GO}, \dots, h_{CS}, k_{GO}, k_{CS}, r_{GO}, f_{GO})$	$4 \times 60 \times \text{ReLU}$	$(K_{11}^{(3)}, K_{12}^{(3)}, K_{13}^{(3)}, K_{22}^{(3)}, K_{23}^{(3)}, K_{33}^{(3)})$	MSE	Adam

Table 6: Architecture of the MLP model for prediction of ETC at mesoscale with dissimilar inclusions.

808 A total of 1000 samples are randomly generated with a uniform distribution following the intervals  
809 specified in Tab. 5. The MLP model is configured with 4 hidden layers of 60 neurons each as shown in  
810 Fig. 2(c). For dissimilar inclusions scenarios, the nano-meso model is constructed with 4 hidden layers  
811 with the parameters defined as  $x_1 = N_m, x_2 = N_{b3}, x_3 = N_{b5}, x_4 = h_{GO}, x_5 = h_{CS}, x_6 = k_{GO}, x_7 =$   
812  $k_{GO}, x_8 = r_{GO}, x_9 = f_{GO}$  and  $y_1 = K_{11}^{(3)}, y_2 = K_{12}^{(3)}, y_3 = K_{13}^{(3)}, y_4 = K_{22}^{(3)}, y_5 = K_{23}^{(3)}, y_6 = K_{33}^{(3)}$ ,  
813 associated with  $\hat{\Theta}_g^{(1)} = \{N_m, N_{b3}, N_{b5}, h_{GO}, h_{CS}\}$ ,  $\hat{\Theta}_p^{(1)} = \{k_{GO}, k_{CS}\}$ ,  $\hat{\Theta}_g^{(2)} = \{r_{GO}, f_{GO}\}$ . The  
814 training is conducted with a batch size of 1, over 10 000 epochs, using MSE loss and the Adam  
815 optimizer. The training and validation processes are evaluated in terms of MAE. The dataset is  
816 divided into training, validation, and test sets with a ratio of 0.6:0.2:0.2.

817 Figure 6 depicts how the loss function (see Fig. 6(a)) and MAE (see Fig. 6(b)) change over the  
818 epochs for both the training and validation sets during the training process. The proposed neural  
819 network model shows rapid convergence. The loss function decreases quickly after continuous learning  
820 and self-adjustment. The final MAE and loss function MSE values are  $0.00214$  and  $1.004 \times 10^{-5}$  for  
821 the training dataset, and  $0.0017$  and  $7.106 \times 10^{-6}$  for the validation dataset.

822 To evaluate the performance of the MLP model, we show in Fig. 6(c) the relative errors of predicted  
823 effective thermal conductivity components  $k_{11}^e, k_{22}^e, k_{33}^e$  and  $k_{mean}^e$  for 200 selected data points. It can  
824 be seen that the relative errors are less than 0.04 for most cases. The correlations of predicted and  
825 exact values are plotted in Fig. 6(d) with a very high value of  $R^2$  (0.9994 for  $k_{mean}^e$ ). These results  
826 show the high prediction accuracy of the proposed MLP model.

### 827 3.2.3. MLP model at macroscale with dissimilar inclusions

828 We have developed a new multiscale homogenization model that incorporates dissimilar inclusions  
829 within the polymer matrix at the microscale. This model now allows us to explore broader scenarios  
830 in which the thermal conductivity distribution within the RVE is heterogeneous at the mesoscale. By  
831 introducing randomly assigned properties across different regions of the RVE, we can more accurately  
832 capture the inherent variability and randomness typical of real heterogeneous materials. Typically,  
833 the effective properties of materials, such as thermal conductivity, result from intricate interactions  
834 at the microscale. As a result, utilizing an RVE with a heterogeneous distribution of thermal conduc-  
835 tivity enables a deeper understanding of how these microscale variations influence the overall thermal

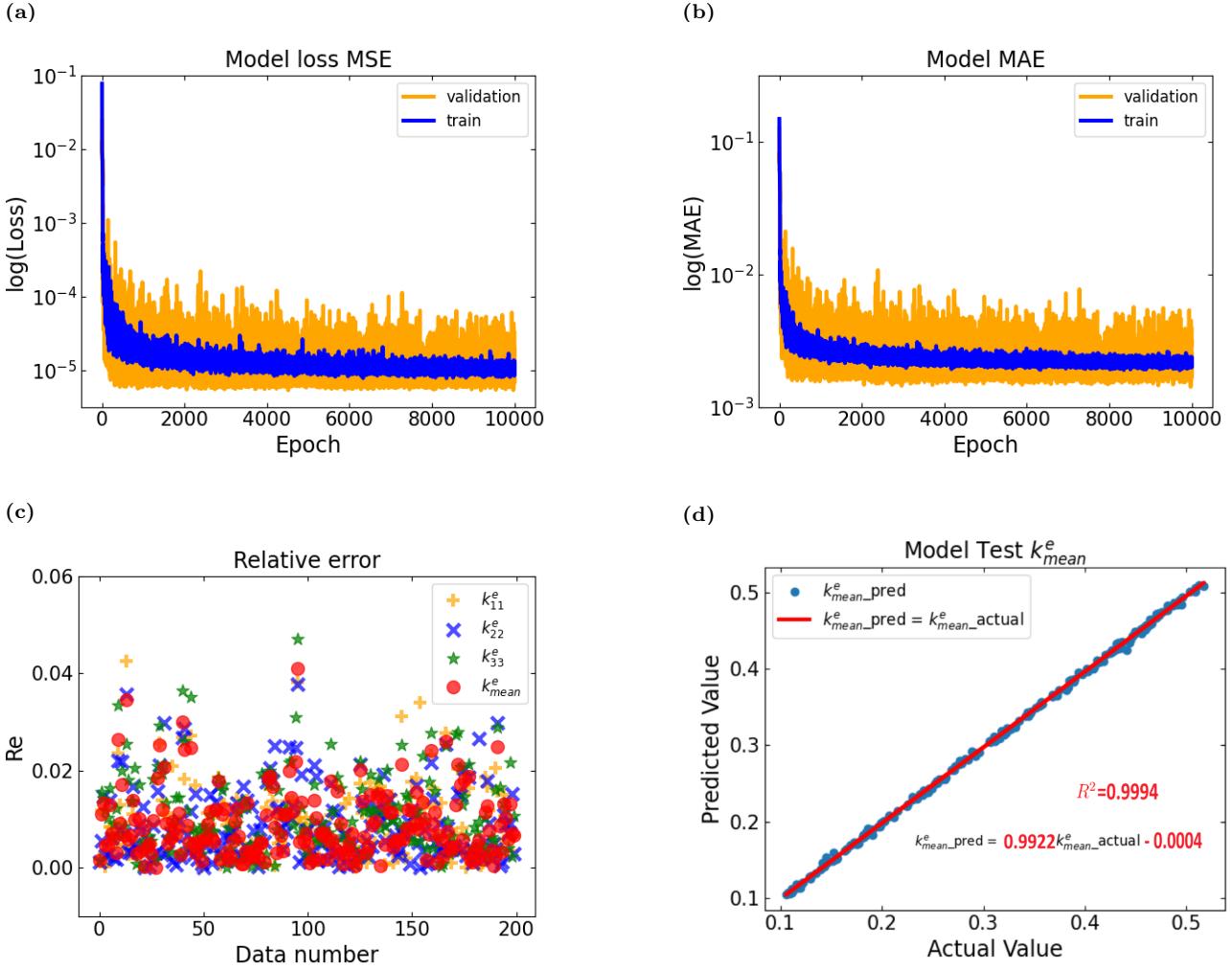


Figure 6: **Nano-Meso problem with dissimilar inclusions.** (a) Evolution of loss MSE as a function of number of epochs between training set and validation set. (b) Evolution of MAE as a function of number of epochs between training set and validation set. (c) Relative error between predicted and exact values of predicted effective thermal conductivity coefficients  $k_{11}^e, k_{22}^e, k_{33}^e, k_{mean}^e = (k_{11}^e + k_{22}^e + k_{33}^e)/3$ . (d) Correlation between predicted and exact values.

behavior of aerogels. This understanding is pivotal for optimizing material design and performance. Furthermore, employing an RVE approach at the mesoscale with a heterogeneous thermal conductivity distribution provides a more realistic and precise method to predict the effective thermal conductivity of aerogels. It facilitates a detailed examination of how microstructural characteristics affect thermal transport, which is crucial for advancing aerogel technologies in applications such as thermal insulation. This research is particularly promising for predicting the effective properties of hybrid aerogels. Notably, to our knowledge, there are currently not many studies in the literature that have explored this specific approach, highlighting its novelty and potential impact in materials science and engineering.

Several studies have explored the use of heterogeneous RVEs to predict effective thermal conductivity across diverse applications, including composite materials, porous media, and multi-phase materials. Noteworthy examples discussing the effectiveness of this approach can be found in [59], [60], [61]. These studies illustrated the application and advantages of employing heterogeneous RVEs in accurately predicting thermal properties across various material systems.

More specifically, we consider cases in which the RVE is divided into 8 quarters, as it can be seen in Fig. 3(e), with randomly attributed conductivities produced from the 3 types of inclusions: exfoliated structure (monolayer GO), intercalated structures (block-3 and block-5 GO) from microscale. Considering an heterogeneous RVE with randomly assigned properties in different parts better represents the

854 inherent randomness and heterogeneity present in many materials, enabling more accurate predictions  
 855 of effective properties.

Scale	Parameter	Notation	Value
Mesoscale	Diameter of the sphere (m)	$D_s$	$[60, 200] \times 10^{-6}$
	Wall thickness of the sphere (m)	$t_s$	$[0.34, 1.3] \times 10^{-6}$
	Diameter of the 6 holes on the wall of the sphere (m)	$D_h$	$[6.51, 40] \times 10^{-6}$
	Thermal conductivity of gas phase, isotropic (W/(m·K))	$k_g$	$[0.01, 0.03]$
	Thermal conductivity of solid phase, quarter $i$ , $i \in [1, 8]$ (W/(m·K))	$k_{si}$	$[0.0987, 0.8167]$

Table 7: Values and notations of the parameters at mesoscale for dissimilar inclusions.

856 Similar to the approach discussed in section 3.1.2, the RVE used in this study is a hollow sphere  
 857 with six holes on its surface, characterized by consistent geometrical parameters such as sphere's  
 858 diameter, sphere's wall thickness, and hole's diameter. The thermal conductivity of the gas phase  
 859 remains unchanged. However, the solid phase's thermal conductivity is segmented into eight distinct  
 860 parts, each corresponding to different thermal conductivity values resulting from various combinations  
 861 of three types of inclusions (monolayer, block-3, and block-5 graphene oxide) at the microscale. A  
 862 study was conducted to determine the range of effective thermal conductivity at the mesoscale using  
 863 10 inclusions in equal ratios, with other parameters specified in Tab. 5. Various numerical tests shown  
 864 that increasing the number of inclusions or changing the ratios do not significantly affect the ranges.  
 865 Thus, in this work, these thermal conductivity values range within the interval  $[0.0987, 0.8167]$ , as  
 866 detailed in Tab. 7.

867 Similar to the approach described in section 3.1.2, input features are normalized using formula (20).  
 868 A total of 1200 samples are randomly generated with a uniform distribution based on the intervals  
 869 specified in Tab. 7. The neural network model is designed with two hidden layers, each contains 18  
 870 neurons, and is trained with a batch size of 1 for 10 000 epochs. It utilizes MSE as the loss function  
 871 and the Adam optimizer. The training and validation processes are evaluated using MAE. The dataset  
 872 is divided into training, validation, and test sets with a ratio of 0.6:0.2:0.2.

Input layer	Hidden layer	Output layer	Loss function	Optimizer
$(D_s, t_s, D_h, k_g)$ $k_{si}, i \in [1, 8]$	$2 \times 18 \times \text{ReLU}$	$(K_{11}^{(4)}, K_{12}^{(4)}, K_{13}^{(4)}, K_{22}^{(4)}, K_{23}^{(4)}, K_{33}^{(4)})$	MSE	Adam

Table 8: Architecture of the MLP model for prediction of ETC at macroscale for dissimilar inclusions.

873 The training process of the proposed neural network model is illustrated in Fig. 7, showing the  
 874 evolution of the loss function (see Fig. 7(a)) and MAE (see Fig. 7(b)) for both the training and  
 875 validation sets across epochs. The model demonstrates rapid convergence, with the loss function  
 876 decreasing steadily as the network adjusts. Ultimately, the MAE reaches final values of 0.0004 for the  
 877 training dataset and 0.0003 for the validation dataset, accompanied by very low final loss function  
 878 MSE values ( $2.93 \times 10^{-7}$  for training and  $1.82 \times 10^{-7}$  for validation).

879 To assess the accuracy of the proposed MLP model, Figure 7(c) displays the relative errors of  
 880 predicted effective thermal conductivity components  $k_{11}^e, k_{22}^e, k_{33}^e$  and  $k_{mean}^e$  for 240 selected data.  
 881 The majority of relative errors are below 0.06, indicating the model's high accuracy. Additionally,  
 882 Figure 7(d) shows the correlation between predicted and exact values, with a very high coefficient of  
 883 determination ( $R^2$  value of 0.9989 for  $k_{mean}^e$ ), demonstrating strong predictive performance across all  
 884 cases. These results show the high prediction accuracy of the proposed MLP model.

885 In the following discussion, we compare the total computational time between the deep learning  
 886 approach and traditional methods. The comparison considers scenarios with dissimilar inclusions,  
 887 specifically addressing both Nano-Meso and Meso-Macro problems, as outlined in Algorithms 1 and 2,  
 888 respectively. Detailed computational times, measured in seconds (s), are presented in Tables 9 and 10.  
 889 All computations and model training were performed on a system with a 12th Gen Intel(R) Core(TM)  
 890 i7-12700 CPU @ 2.10 GHz and 32 GB of RAM. For the deep learning approach, 1000 samples were  
 891 generated and used for training, with the maximum number of training epochs set to 10000. It can be  
 892 observed that, after the training phase, the deep learning approach is capable of generating predictions

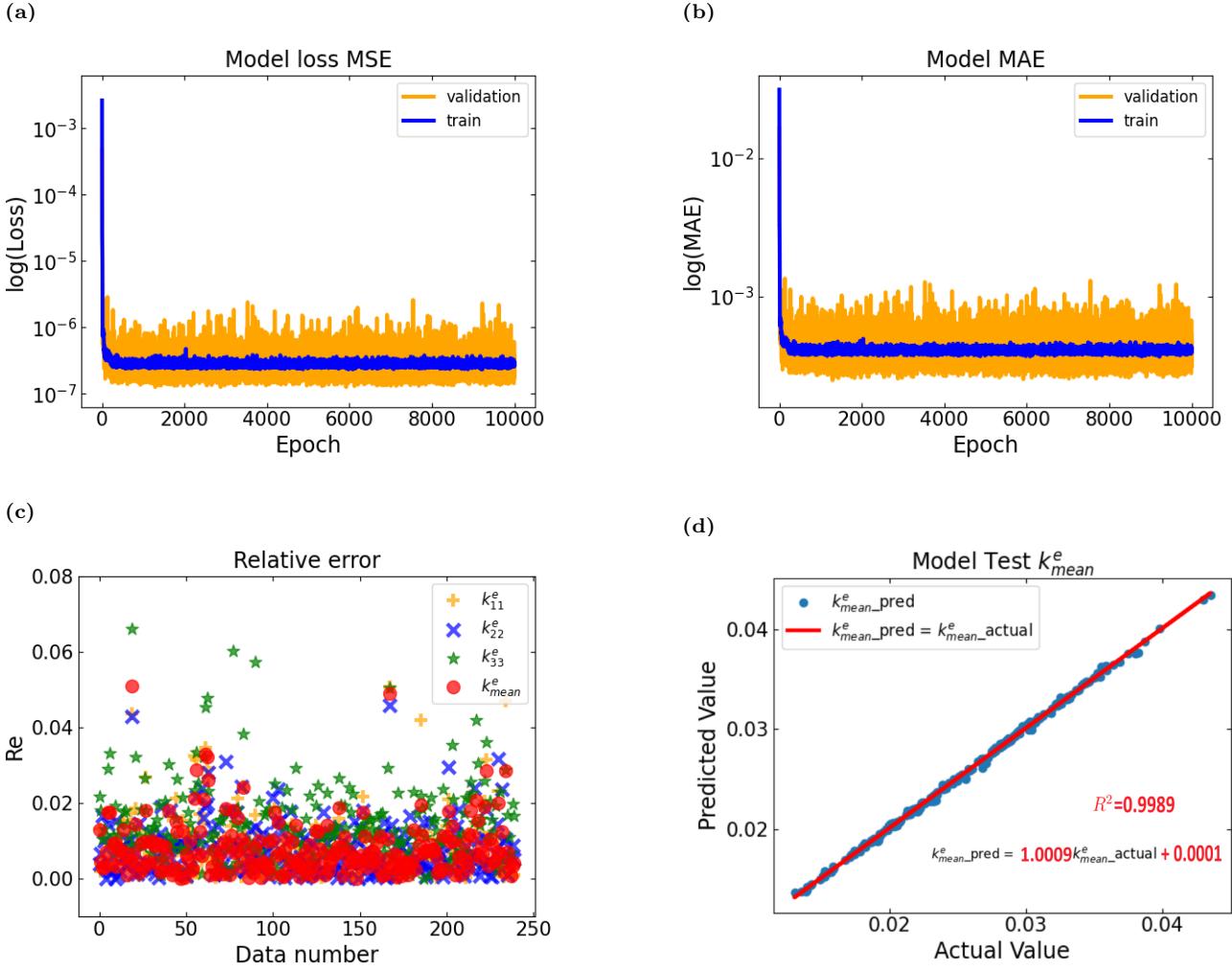


Figure 7: **Meso-Macro problem with dissimilar inclusions.** (a) Evolution of loss MSE as a function of number of epochs between training set and validation set. (b) Evolution of MAE as a function of number of epochs between training set and validation set. (c) Relative error between predicted and exact values of predicted effective thermal conductivity coefficients  $k_{11}^e, k_{22}^e, k_{33}^e, k_{mean}^e = (k_{11}^e + k_{22}^e + k_{33}^e)/3$ . (d) Correlation between predicted and exact values.

almost instantaneously.

Scale	Step	Traditional methods
Nano-Meso	Generate random parameters, inclusion generating	10
	Milton's method	180
	FEM	360
Meso-Macro	Generate random parameters, FEM	1584

Table 9: Computational time (s) of traditional methods in dissimilar inclusions scenarios.

Scale	Step	Deep learning approach
Nano-Meso	Data generation	$(10+180+360)\times 1000$
	Training process	1800
	Prediction	0.5
Meso-Macro	Data generation	$1584\times 1000$
	Training process	1800
	Prediction	0.5

Table 10: Computational time (s) of deep learning approach in dissimilar inclusions scenarios for training on 1000 samples and 10000 epochs.

894 **4. Numerical results**

895 In this section, we explore the application of the developed MLP models, with a particular emphasis  
896 on their predictive performance. After establishing the high accuracy of the trained models, we  
897 investigate how different microstructural features influence the ETC at both mesoscopic and macro-  
898 scopic scales. Specifically, we analyze the effect of various inclusion types and mixing ratios on thermal  
899 behavior, providing insight into the structure–property relationships learned by the surrogate models.  
900 The influence of the monolayer GO content in the polymer matrix is also examined across scales.  
901 Additionally, we assess the extrapolation capability of the MLP models, highlighting their ability to  
902 produce physically consistent and reliable predictions even beyond the training domain.

903 *4.1. Prediction of mesoscale ETC of hybrid aerogels using Nano-Meso MLP model*

904 In the work of Le *et al.* [57, 58], the developed homogenization models were limited at producing  
905 result for only one type of inclusions (monolayer GO, block-3 GO or block-5 GO). However, our new  
906 homogenization models presented in section 3.2.1 allow simulating cases with dissimilar inclusions in  
907 the polymer matrix. This helps us to explain their interaction and their contribution to the effective  
908 thermal conductivity at mesoscale. With the use of our new surrogate deep learning models, we can  
909 now study the insight of their mechanism without passing through homogenization and FE calcula-  
910 tions. More specifically, the ETC is expressed as a function of the volume fraction of GO, as shown  
911 in Fig. 8(a). It can be observed that the thermal conductivity of monolayer GO is the largest with  
912 0.224 W/(m·K) at 3% volume fraction of GO, block-3 GO is smaller with 0.222 W/(m·K) and block-5  
913 GO is the smallest with 0.220 W/(m·K). When there is more than one type of inclusions, the larger  
914 the number of inclusions of monolayer type, the larger the thermal conductivity.

915 For a clearer observation, Fig. 8(b) shows the thermal conductivity as a function of the ratio of  
916 monolayer GO inclusions at 2% volume fraction of GO. It can be observed that the thermal conductiv-  
917 ity increases as the number of inclusions of type monolayer GO increases. Namely, the combination of  
918 monolayer and block-3 GO is the largest and increases from 0.2143 to 0.2163 W/(m·K), the combina-  
919 tion of the 3 types of inclusions is smaller and increases from 0.2134 to 0.2163 W/(m·K) and the com-  
920 bination of monolayer and block-5 GO is the smallest and increases from 0.2123 to 0.2163 W/(m·K).  
921 In general, the increase of thermal conductivity is not significant.

922 *4.2. Prediction of macroscale ETC of hybrid aerogels using Meso-Macro MLP model*

923 We now use the developed models to predict the behavior of ETC at macroscale. As discussed  
924 in section 3.2.3, the RVE in this case is divided into 8 quarters with different thermal conductivities.  
925 From various numerical tests, we take the conductivity of quarters corresponding to monolayer GO  
926 as 0.6, block-3 GO as 0.4 and block-5 GO as 0.3, conventionally. Figure 8(c) illustrates the variation  
927 in the ETC as a function of the ratio of the number of quarters corresponding to monolayer GO at a  
928 2.32% volume fraction of GO. Similarly, we obtain the same behavior as the ETC at mesoscale, the  
929 thermal conductivity increases as the number of quarters corresponding to monolayer GO increases.  
930 Namely, the combination of monolayer and block-3 GO is the largest and increases from 0.01496 to  
931 0.01646 W/(m·K), the combination of the 3 types of inclusions is smaller and increases from 0.01469  
932 to 0.01646 W/(m·K), the combination of monolayer and block-5 GO is the smallest and increases from  
933 0.01455 to 0.01646 W/(m·K). Overall, the increase in thermal conductivity is relatively minor. The  
934 fact that adding fillers such as GO into the polymer matrix is to enhance the mechanical properties  
935 while maintain the proper thermal conductivity of aerogels. This is coincided with the numerical  
936 results obtained from the developed surrogate models in which the mechanical properties are expected  
937 to be much improved.

938 *4.3. Extrapolation capability of developed models*

939 In this section, we examine the extrapolation capability of the developed models, specifically their  
940 ability to make reliable predictions outside the range of training parameters. We assess the model’s  
941 generalization performance beyond the training domain by testing scenarios that include similar in-  
942 clusions. To evaluate extrapolation, input parameters are intentionally assigned values outside their  
943 original training ranges. The numerical results indicate that the models exhibit strong extrapolation  
944 capability, with predictions aligning well with established physical relationships.

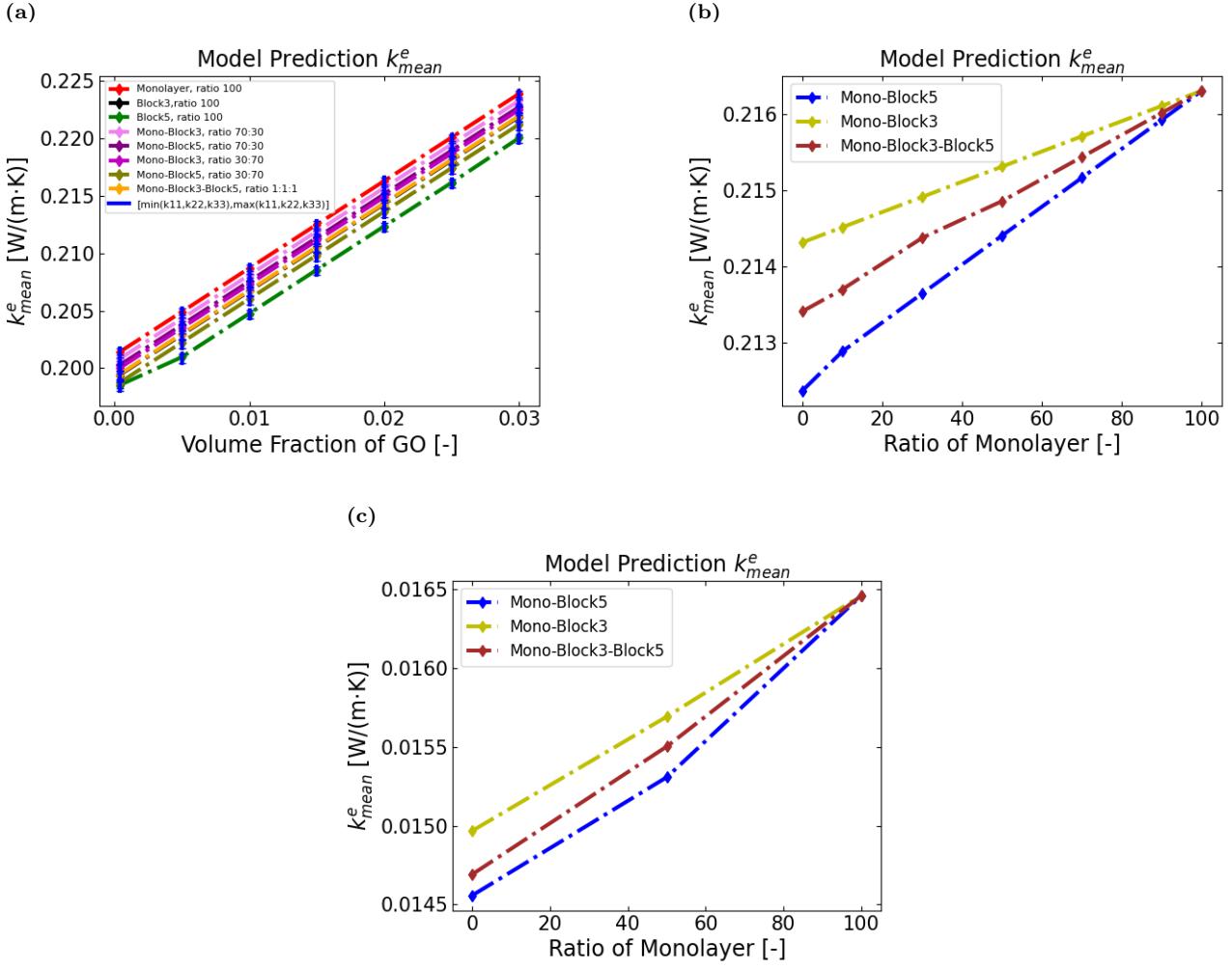


Figure 8: **Applications of developed MLP models.** (a)  $K^{(3)}$  as a function of the volume fraction of GO, with 30 inclusions,  $h_{GO} = 10^{-9}$  m,  $h_{CS} = 2 \times 10^{-9}$  m,  $k_{GO} = 18$  W/(m·K),  $k_{CS} = 0.2$  W/(m·K),  $r_{GO} = 67$ ,  $f_{GO} \in [0.0004, 0.03]$ . (b)  $K^{(3)}$  as a function of the ratio of monolayer GO, with 30 inclusions,  $h_{GO} = 10^{-9}$  m,  $h_{CS} = 2 \times 10^{-9}$  m,  $k_{GO} = 18$  W/(m·K),  $k_{CS} = 0.2$  W/(m·K),  $r_{GO} = 67$ ,  $f_{GO} = 0.02$ . (c)  $K^{(4)}$  as a function of the ratio of monolayer GO at macroscale with  $D_s = 85 \times 10^{-6}$  m,  $t_s = 0.63 \times 10^{-6}$  m,  $D_h = 26 \times 10^{-6}$  m.

#### 945 4.3.1. Nano-Meso MLP model

946 We begin by introducing the test values for extrapolation capability of the developed Nano-Meso  
947 MLP models, which are given in Tab. 11. The numerical results are shown in Fig. 9(a), which are  
948 consistent with the numerical results from Fig. 8 in section 4.1. In particular, increasing input pa-  
949 rameters such as the number of GO inclusions, their geometric characteristics, thermal conductivities,  
950 and volume fraction leads to a higher effective thermal conductivity of the material.

Scale	Notation	Training range	Test range
Nanoscale	$N_m$	[0,30]	[35,50]
	$N_{b3}$	[0,30]	[35,50]
	$N_{b5}$	[0,30]	[35,50]
	$h_{GO}$	$10^{-9}$	$[1.5, 3] \times 10^{-9}$
	$h_{CS}$	$[1.5, 5.3] \times 10^{-9}$	$[3, 6] \times 10^{-9}$
	$k_{GO}$	[16, 20]	[22, 25]
	$k_{CS}$	[0.1, 0.5]	[0.52, 0.88]
Microscale	$r_{GO}$	[10, 100]	[105, 120]
	$f_{GO}$	[0.004, 0.03]	[0.035, 0.05]

Table 11: Values and notations of the parameters at nanoscale and microscale for extrapolation tests.

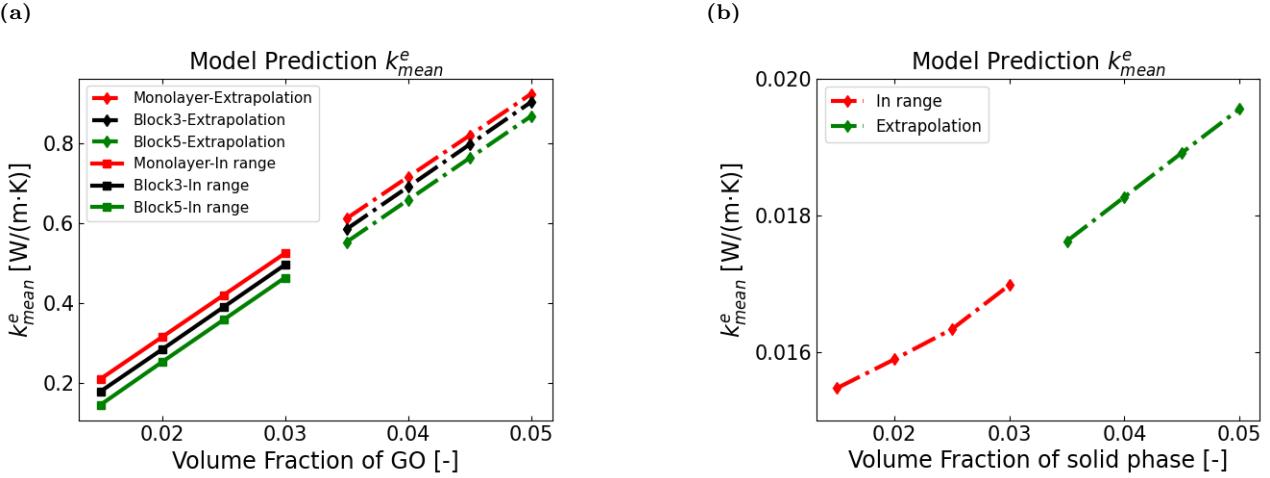


Figure 9: **Extrapolation tests.** (a)  $K^{(3)}$  as a function of the volume fraction of GO, parameters are detailed in Tab. 11. (b)  $K^{(4)}$  as a function of the volume fraction of solid phase, details of parameters are given in Tab. 12.

#### 951 4.3.2. Meso-Macro MLP model

952 Similarly to the previous section, we define the test values for extrapolation capability of the  
953 developed Meso-Macro MLP models, which are detailed in Tab. 12. The volume fraction of solid phase  
954 is considered on the interval [0.015, 0.03] for the training tests and [0.035, 0.05] for the extrapolation  
955 tests. The numerical results in Fig. 9(b) demonstrate the consistency of the model. Specifically, an  
956 increase in the solid phase volume fraction, along with higher thermal conductivities of both the solid  
957 and gas phases, results in an overall enhancement of the material's effective thermal conductivity.

Scale	Notation	Training range	Test range
Mesoscale	$D_s$	$[60, 200] \times 10^{-6}$	$[246, 296] \times 10^{-6}$
	$t_s$	$[0.34, 1.3] \times 10^{-6}$	$[1.8, 2.3] \times 10^{-6}$
	$D_h$	$[6.51, 40] \times 10^{-6}$	$[49, 56] \times 10^{-6}$
	$k_g$	$[0.01, 0.03]$	$[0.035, 0.05]$
	$k_{si}$	$[0.0987, 0.8167]$	$[0.85, 0.99]$

Table 12: Values and notations of the parameters at mesoscale for extrapolation tests.

## 958 5. Conclusion and perspectives

959 In this study, we presented a novel framework that integrates deep neural networks with mul-  
960 tiscale homogenization techniques to predict the ETC of composite materials, focusing on hybrid  
961 aerogels. Specifically, MLP models were trained on high-fidelity datasets generated from validated  
962 multiscale homogenization simulations. These models were applied to two scenarios, composites with  
963 similar and dissimilar inclusions, demonstrating excellent predictive accuracy across both mesoscale  
964 and macroscale regimes.

965 A key advancement of this work lies in the development of a multiscale homogenization model  
966 capable of incorporating three types of randomly distributed inclusions within a polymer matrix.  
967 This model marks a significant improvement over earlier approaches, as proposed in [57, 58], by  
968 incorporating heterogeneous mesoscale property distributions within the RVE, enabling more realistic  
969 characterization of aerogel structures. As a result, our approach more accurately reflects the complex  
970 microstructures found in real materials and offers a powerful tool for thermal material design.

971 In addition to accuracy, the proposed MLP models demonstrate substantial efficiency gains. This  
972 study includes a direct comparison of computational time between the MLP-based surrogate models  
973 and conventional multiscale modeling methods. The results highlight the superior computational

efficiency of our approach, making it well-suited for rapid material screening and design optimization, particularly when large-scale simulations are computationally prohibitive.

Furthermore, we evaluated the extrapolation capabilities of the developed MLP models. The results show that the models are able to produce physically consistent and reliable predictions even beyond the domain of the training data. This strong extrapolation performance is critical for practical engineering applications where unseen microstructural configurations may arise. It also emphasizes the robustness and generalization capacity of the proposed deep learning framework.

While this study is primarily focused on developing and validating the proposed methodology through comprehensive numerical simulations, the multiscale homogenization models employed to generate the dataset for training have already been benchmarked against experimental results, as stated in [58]. This prior validation ensures that the data used to train the MLP models is physically grounded and reliably represents thermal behavior, as evidenced by the strong agreement of the MLP predictions with established physical laws and their demonstrated high accuracy. Moving forward, we aim to strengthen the methodology by engaging in direct collaboration with experimentalists. Future efforts will focus on validating the predicted effective thermal conductivities through comparison with existing or newly acquired experimental data. This will further enhance the credibility and practical applicability of the proposed framework.

Looking ahead, we plan to extend the current framework in several directions. One key avenue involves developing MLP models capable of predicting macroscopic effective properties directly from nanoscopic input parameters, thus eliminating the need for sequential homogenization steps. Additionally, we propose enhancing the mesoscale geometric representation by adopting more realistic structures such as polyhedral cells based on Kelvin or Schwarz P minimal surfaces. These modifications could better capture the pore morphology of aerogels and improve prediction reliability.

Beyond thermal conductivity, the framework is also applicable to predicting mechanical properties of hybrid aerogels. Integrating CNNs with MLPs represents another promising direction, where CNNs can extract features from geometrical or tomographic images to directly predict material properties. Such hybrid models could offer a powerful image-to-property pipeline, facilitating automated and accurate property estimation for a broad class of advanced materials.

In summary, this work establishes a physically grounded, scalable, and computationally efficient approach for predicting effective properties of complex composite materials. It not only achieves high accuracy and generalization performance but also sets the stage for future advancements in data-driven multiscale modeling and intelligent material design.

## Appendix A.

In this appendix, the multiscale and homogenization methods are presented in order to determine the effective thermal properties. The Cartesian reference system was used where the orthonormal basis was specified by  $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$  and the coordinates of a point were denoted by  $\mathbf{x} = (x_1, x_2, x_3)$ .

### Appendix A.1. Milton's method

In the context of a laminate region denoted as domain  $\Omega$ , comprising two distinct phases, and under the assumption of an absence of volumetric heat generation, the conductivity equations for each phase are outlined below:

$$\operatorname{div}(\mathbb{K}(\mathbf{x}) \operatorname{grad} T(\mathbf{x})) = 0, \quad (\text{A.1})$$

where  $T$  denotes the temperature and  $\mathbb{K}$  stands for the second-order thermal conductivity tensor. The operators  $\operatorname{div}$  and  $\operatorname{grad}$  represent the divergence and gradient, respectively.

By applying Milton method, the effective thermal conductivity tensor  $\mathbb{K}^e$  can be computed by the following steps:

Step 1: Compute  $\mathbb{L} = \langle [\mathbb{M}(\mathbf{x}) - \mathbb{F}_1(\mathbf{n})]^{-1} \rangle_{\Omega}$ ,

Step 2: Compute  $\mathbb{M}^e = \mathbb{L}^{-1} + \mathbb{F}_1(\mathbf{n})$ ,

Step 3: Compute  $\mathbb{K}^e = c_0(\mathbb{I} - (\mathbb{M}^e)^{-1})$ ,

1021 where  $c_0$  is an arbitrary constant,  $\mathbb{I}$  denotes the second-order identity tensor and the symbol  $\langle \cdot \rangle_\Omega$   
1022 represents the volume average of  $(\cdot)$  over domain  $\Omega$ ,  $\Gamma_1(\mathbf{n}) = \mathbf{n} \otimes \mathbf{n}$  is the projection on the out of  
1023 plane of the normal  $\mathbf{n}$ , the symbol  $\otimes$  represents the tensor product. Finally,  $\mathbb{M}(\mathbf{x})$  is a second-order  
1024 tensor which can be expressed as

$$\mathbb{M}(\mathbf{x}) = c_0(c_0 \mathbb{I} - \mathbb{K}(\mathbf{x}))^{-1}.$$

1025 More details of the calculation procedure can be found in [58].

## 1026 Appendix A.2. Asymptotic homogenization method

1027 We examine a two-phase composite material structure within a periodic domain. The mesoscopic  
1028 domain of a unit cell is denoted by  $Y \subset \mathbb{R}^3$ , where  $l_c$  represents the mesoscopic characteristic length.  
1029 At the macroscale, the characteristic length is denoted by  $L_c$ , with the relationship  $\frac{l_c}{L_c} = \epsilon \ll 1$ . The  
1030 overall domain  $Y$  is composed of two phases,  $a$  and  $b$ , which occupy the subdomains  $Y^a$  and  $Y^b$ ,  
1031 respectively. The interface between these two phases is denoted by  $\Gamma$ :

$$Y = Y^a \cup Y^b, \quad Y^a \cap Y^b = \emptyset, \quad \partial Y^a \cup \partial Y^b = \Gamma.$$

1032 Let  $T^\alpha$  represent the temperature in the domain  $Y^\alpha$ . In the absence of a volumetric heat source,  
1033 the steady-state heat conduction problem is described by the following equations:

$$\operatorname{div}(\mathbb{K}^\alpha \operatorname{grad} T^\alpha) = 0 \quad \text{in } Y^\alpha, \quad \alpha = a, b, \quad (\text{A.2})$$

1034 with the assumption of perfect thermal contact between phase  $a$  and phase  $b$ , and the continuity  
1035 conditions for the heat transfer problem are given by:

$$T^a = T^b \quad \text{on } \Gamma, \quad (\text{A.3})$$

$$(\mathbb{K}^a \operatorname{grad} T^a - \mathbb{K}^b \operatorname{grad} T^b) \cdot \mathbf{n} = 0 \quad \text{on } \Gamma, \quad (\text{A.4})$$

1036 where  $\mathbf{n}$  is the unit vector that points outward and is locally normal to the boundary  $\Gamma$ .

1037 The thermal conductivity tensor is a second-order symmetric positive-definite tensor, which can  
1038 be written as:

$$\mathbb{K}^\alpha(\mathbf{x}) = \begin{bmatrix} K_{11}^\alpha & K_{12}^\alpha & K_{13}^\alpha \\ K_{12}^\alpha & K_{22}^\alpha & K_{23}^\alpha \\ K_{13}^\alpha & K_{23}^\alpha & K_{33}^\alpha \end{bmatrix},$$

1039 where  $K_{ij}^\alpha, 1 \leq i, j \leq 3$  are functions of vector  $\mathbf{x}$ .

1040 We introduce the fast coordinates  $\mathbf{x} = (x_1, x_2, x_3)$  at the mesoscale and the slow coordinates  
1041  $\mathbf{y} = (y_1, y_2, y_3)$  at the macroscale, such that  $y_i = \epsilon x_i$ .

1042 The asymptotic expansion of temperature is expressed as a power series in  $\epsilon$  as follows:

$$T^\alpha = T^{\alpha(0)}(\mathbf{x}, \mathbf{y}) + \epsilon T^{\alpha(1)}(\mathbf{x}, \mathbf{y}) + \epsilon^2 T^{\alpha(2)}(\mathbf{x}, \mathbf{y}) + \dots, \quad (\text{A.5})$$

1043 where  $T^{\alpha(0)}, T^{\alpha(1)}, T^{\alpha(2)}, \dots$  are  $Y$ -periodic in the variable  $\mathbf{y}$ .

1044 We denote terms of order  $\epsilon^n$  with the superscript  $(n)$  where macroscale variables are expressed at  
1045  $\epsilon^0$  order.

1046 After a homogenization calculations which were described in [58], the thermal problem 6 is solved  
1047 by introducing the  $Y$ -periodic characteristic function  $\mathbf{w}^\alpha = (w_j^\alpha(\mathbf{x}))$  which is governed by the following  
1048 boundary-value problem:

$$\frac{\partial}{\partial x_i} \left( K_{ij}^\alpha \left( \delta_{jl} + \frac{\partial w_l^\alpha}{\partial x_j} \right) \right) = 0, \quad \text{in } Y^\alpha, \quad \alpha = a, b, \quad (\text{A.6})$$

$$w_l^a = w_l^b, \quad \text{on } \Gamma, \quad (\text{A.7})$$

$$K_{ij}^a \left( \delta_{jl} + \frac{\partial w_l^a}{\partial x_j} \right) n_i = K_{ij}^b \left( \delta_{jl} + \frac{\partial w_l^b}{\partial x_j} \right) n_i, \quad \text{on } \Gamma, \quad (\text{A.8})$$

$$\langle \mathbf{w}^\alpha \rangle_Y = 0, \quad (\text{A.9})$$

1049 where  $\delta_{jl}$  represents the Kronecker symbol and  $\langle *\rangle_Y$  denotes the volume average of  $(*)$  over the domain  
 1050  $Y$ :

$$\langle *\rangle_Y = \frac{1}{|Y|} \left( \int_{Y^a} (*) dV + \int_{Y^b} (*) dV \right),$$

1051 where  $dV$  is differential volume element.

1052 The condition (A.9) ensures that the solution  $\mathbf{w}^\alpha$  is uniquely determined.

1053 The macroscopic equation can be written as:

$$\operatorname{div}(\mathbb{K}^e \operatorname{grad} T^{\alpha(0)}(\mathbf{x})) = 0 \quad \text{in } Y, \quad (\text{A.10})$$

1054 where the components of the effective thermal conductivity tensor  $\mathbb{K}^e$  are defined as:

$$K_{ij}^e = \left\langle K_{il}^\alpha \left( \frac{\partial w_j^\alpha}{\partial x_l} + \delta_{jl} \right) \right\rangle_Y. \quad (\text{A.11})$$

1055 The effective thermal conductivity tensor is obtained from Eq. (A.11), which derives from solving  
 1056 the cell problem defined by Eqs. (A.6)-(A.9).

1057 Multiplying (A.6) by a test function  $\mathbf{v} \in V = \{\mathbf{v} | \mathbf{v} \text{ satisfies Eqs. (A.7)-(A.9)}\}$ , and from integration  
 1058 by parts over each constituent, we have:

$$\begin{aligned} \int_{Y^\alpha} v_i \frac{\partial}{\partial x_i} \left( K_{ij}^\alpha \left( \delta_{jl} + \frac{\partial w_l^\alpha}{\partial x_j} \right) \right) dV &= 0, \quad x_i \in Y^\alpha, \alpha = a, b, \\ \int_{\partial Y^\alpha} v_i \left( K_{ij}^\alpha \left( \delta_{jl} + \frac{\partial w_l^\alpha}{\partial x_j} \right) \right) n_{i\alpha} dS - \int_{Y^\alpha} \left( K_{ij}^\alpha \left( \delta_{jl} + \frac{\partial w_l^\alpha}{\partial x_j} \right) \right) \frac{\partial v_i}{\partial x_i} dV &= 0, \end{aligned}$$

1059 where  $dV$  and  $dS$  are differential elements of volume and area, respectively. Due to  $\mathbf{n}_a = -\mathbf{n}_b$  and the  
 1060 condition of flux continuity (A.8), we obtain the weak form of (A.6):

$$\int_{Y^a} \left( K_{ij}^a \left( \delta_{jl} + \frac{\partial w_l^a}{\partial x_j} \right) \right) \frac{\partial v_i}{\partial x_i} dV + \int_{Y^b} \left( K_{ij}^b \left( \delta_{jl} + \frac{\partial w_l^b}{\partial x_j} \right) \right) \frac{\partial v_i}{\partial x_i} dV = 0. \quad (\text{A.12})$$

1061 The variational formulation (A.12) can be rewritten as:

$$a(\mathbf{w}, \mathbf{v}) = \varphi(\mathbf{v}), \quad \forall \mathbf{v} \in V, \quad (\text{A.13})$$

1062 where the bilinear and linear forms are defined by:

$$a(\mathbf{w}, \mathbf{v}) = \int_{Y^a} \left( K_{ij}^a \left( \delta_{jl} + \frac{\partial w_l^a}{\partial x_j} \right) \right) \frac{\partial v_i}{\partial x_i} dV + \int_{Y^b} \left( K_{ij}^b \left( \delta_{jl} + \frac{\partial w_l^b}{\partial x_j} \right) \right) \frac{\partial v_i}{\partial x_i} dV, \quad (\text{A.14})$$

$$\varphi(\mathbf{v}) = 0. \quad (\text{A.15})$$

1063 The form  $a(\mathbf{w}, \mathbf{v})$  is bilinear, symmetric positive-definite and coercive, as a consequence, the variational  
 1064 formulation defined by (A.13) admits a unique solution  $\mathbf{w}$  following Lax-Milgram Theorem.

1065 The weak formulation (A.13) of the cell problem of (A.6) can be solved by FEM using Comsol  
 1066 Multiphysics software.

1067 The main steps of the homogenization procedure can be summarized by the following Algorithm 3:

---

### Algorithm 3: key homogenization steps

---

- 1 Define Multiscale Problem: Composite with phases a and b, periodic domain Y
  - 2 Define Governing Equations: Eqs. (A.2)-(A.4)
  - 3 Apply Asymptotic Expansion: Eq. (A.5)
  - 4 Derive Cell Problem: Solve Eqs. (A.6)-(A.9) for characteristic functions  $w_l^\alpha$
  - 5 Compute Effective Conductivity Tensor: ETC tensor is obtained from Eq. (A.11)
- 

1069 Acknowledgements

1070 This work has benefited from a French government grant managed by ANR within the frame of  
 1071 the national program of Investments for the Future ANR-11-LABX-0022-01 (LabEx MMCD project).

1072 **References**

- [1] E. Weinan, Principles of multiscale modeling, Cambridge University Press, 2011.
- [2] F. Meng, Z. Cui, Y. Liang, Z. Wang, T. Yu, Z. Ma, J. Zhao, Multiscale model of material removal for ultrasonic assisted polishing of cylindrical surfaces, *Tribology International* 202 (2025) 110383. doi:<https://doi.org/10.1016/j.triboint.2024.110383>. URL <https://www.sciencedirect.com/science/article/pii/S0301679X24011356>
- [3] S. Pyrialakos, I. Kalogeris, V. Papadopoulos, An efficient hierarchical bayesian framework for multiscale material modeling, *Composite Structures* 351 (2025) 118570. doi:<https://doi.org/10.1016/j.compstruct.2024.118570>. URL <https://www.sciencedirect.com/science/article/pii/S0263822324006986>
- [4] L.-Y. Zhao, L. Ren, L.-H. Liu, Y.-M. Lai, F.-J. Niu, T. You, Multiscale modeling of diffuse damage and localized cracking in quasi-brittle materials under compression with a quadratic friction law, *International Journal of Solids and Structures* 304 (2024) 113038. doi:<https://doi.org/10.1016/j.ijsolstr.2024.113038>. URL <https://www.sciencedirect.com/science/article/pii/S0020768324003974>
- [5] J. Ruuskanen, A. Marteau, I. Niyonzima, A. Halbach, J. Vesa, G. Meunier, T. Tarhasaari, P. Rasilo, Multiharmonic multiscale modelling in 3-d nonlinear magnetoquasistatics: Composite material made of insulated particles, *Computer Methods in Applied Mechanics and Engineering* 425 (2024) 116945. doi:<https://doi.org/10.1016/j.cma.2024.116945>. URL <https://www.sciencedirect.com/science/article/pii/S0045782524002019>
- [6] T. Kaiser, N. von der Höh, A. Menzel, Computational multiscale modelling of material interfaces in electrical conductors, *Journal of the Mechanics and Physics of Solids* 186 (2024) 105601. doi:<https://doi.org/10.1016/j.jmps.2024.105601>. URL <https://www.sciencedirect.com/science/article/pii/S002250962400067X>
- [7] E. Weinan, B. Engquist, Multiscale modeling and computation, *Notices of the AMS* 50 (9) (2003) 1962–1070.
- [8] G. Lu, E. Kaxiras, Handbook of theoretical and computational nanotechnology, Vol. 10, American Scientific Publishers, 2005, Ch. 22: Overview of multiscale simulations of materials, pp. 1–33.
- [9] E. Tadmor, R. Miller, Modeling materials: continuum, atomistic and multiscale techniques, Cambridge University Press, 2011.
- [10] Y. Wang, C. Soutis, D. Ando, Y. Sutou, F. Narita, Application of deep neural network learning in composites design, *European Journal of Materials* 2 (1) (2022) 117–170. arXiv:<https://doi.org/10.1080/26889277.2022.2053302>, doi:[10.1080/26889277.2022.2053302](https://doi.org/10.1080/26889277.2022.2053302). URL <https://doi.org/10.1080/26889277.2022.2053302>
- [11] M. Mirkhalaf, I. Rocha, Micromechanics-based deep-learning for composites: Challenges and future perspectives, *European Journal of Mechanics - A/Solids* 105 (2024) 105242. doi:<https://doi.org/10.1016/j.euromechsol.2024.105242>. URL <https://www.sciencedirect.com/science/article/pii/S0997753824000226>
- [12] C. Qiu, Y. Han, L. Shanmugam, Y. Zhao, S. Dong, S. Du, J. Yang, A deep learning-based composite design strategy for efficient selection of material and layup sequences from a given database, *Composites Science and Technology* 230 (2022) 109154, recent Advances in Composites Science & Technology - A special issue in memoriam Prof. Dr.-Ing. Dr. h.c. mult. Klaus Friedrich. doi:<https://doi.org/10.1016/j.compscitech.2021.109154>. URL <https://www.sciencedirect.com/science/article/pii/S0266353821005108>

- 1116 [13] K. Gholami, F. Ege, R. Barzegar, Prediction of composite mechanical properties: Integration of  
 1117 deep neural network methods and finite element analysis, *Journal of Composites Science* 7 (2)  
 1118 (2023). doi:10.3390/jcs7020054.  
 1119 URL <https://www.mdpi.com/2504-477X/7/2/54>
- 1120 [14] G. Tanner, Introduction to deep learning with keras, <https://gilberttanner.com/blog/introduction-to-deep-learning-withkeras/> (2019).
- 1122 [15] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving,  
 1123 M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D. G. Murray, B. Steiner, P. Tucker,  
 1124 V. Vasudevan, P. Warden, M. Wicke, Y. Yu, X. Zheng, Tensorflow: A system for large-scale  
 1125 machine learning, in: Proceedings of the 12th USENIX Conference on Operating Systems Design  
 1126 and Implementation, OSDI'16, USENIX Association, USA, 2016, p. 265–283.
- 1127 [16] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Pret-  
 1128 tenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot,  
 1129 E. Duchesnay, Scikit-learn: Machine learning in Python, *Journal of Machine Learning Research*  
 1130 12 (2011) 2825–2830.
- 1131 [17] P. Kanouté, D. P. Boso, J. L. Chaboche, B. A. Schrefler, Multiscale methods for composites: A  
 1132 review, *Archives of Computational Methods in Engineering* 16 (1) (2009) 31–75. doi:10.1007/s11831-008-9028-8.  
 1133 URL <https://doi.org/10.1007/s11831-008-9028-8>
- 1135 [18] G. Chatzigeorgiou, F. Meraghni, N. Charalambakis, Multiscale Modeling Approaches for  
 1136 Composites, Elsevier, 2022. doi:<https://doi.org/10.1016/C2019-0-05214-4>.  
 1137 URL <https://www.sciencedirect.com/book/9780128231432/multiscale-modeling-approaches-for-co>  
 1138 via=ihub=
- 1139 [19] S. Seitl, J. Man, M. Jambor, Preface to the special issue of theoretical and applied fracture  
 1140 mechanics - fatigue crack growth - multiscale modelling and crack path analysis in engineering  
 1141 materials, *Theoretical and Applied Fracture Mechanics* 136 (2025) 104860. doi:<https://doi.org/10.1016/j.tafmec.2025.104860>.  
 1142 URL <https://www.sciencedirect.com/science/article/pii/S0167844225000187>
- 1144 [20] Z. Li, Z. Mao, R. Corcolle, A new multiscale model for ferroelectric materials considering hys-  
 1145 teresis, *Mechanics of Materials* 187 (2023) 104842. doi:<https://doi.org/10.1016/j.mechmat.2023.104842>.  
 1146 URL <https://www.sciencedirect.com/science/article/pii/S0167663623002880>
- 1148 [21] K. Liu, L. Meng, A. Zhao, Z. Wang, L. Chen, P. Li, A hybrid direct fe2 method for modeling  
 1149 of multiscale materials and structures with strain localization, *Computer Methods in Applied  
 1150 Mechanics and Engineering* 412 (2023) 116080. doi:<https://doi.org/10.1016/j.cma.2023.116080>.  
 1151 URL <https://www.sciencedirect.com/science/article/pii/S0045782523002049>
- 1153 [22] J. Selvaraj, B. El Said, Multiscale modelling of strongly heterogeneous materials using geometry  
 1154 informed clustering, *International Journal of Solids and Structures* 280 (2023) 112369. doi:  
 1155 <https://doi.org/10.1016/j.ijsolstr.2023.112369>.  
 1156 URL <https://www.sciencedirect.com/science/article/pii/S0020768323002664>
- 1157 [23] B. Elguezabal, J. Martínez-Esnaola, R. Soler, E. Paños, J. Alkorta, A multiscale material model  
 1158 for metallic powder compaction during hot isostatic pressing, *Powder Technology* 425 (2023)  
 1159 118599. doi:<https://doi.org/10.1016/j.powtec.2023.118599>.  
 1160 URL <https://www.sciencedirect.com/science/article/pii/S0032591023003832>

- 1161 [24] E. Andreassen, C. S. Andreasen, How to determine composite material properties using numerical  
1162 homogenization, Computational Materials Science 83 (2014) 488–495. doi:<https://doi.org/10.1016/j.commatsci.2013.09.006>.  
1164 URL <https://www.sciencedirect.com/science/article/pii/S0927025613005302>
- 1165 [25] F. Otero, S. Oller, X. Martinez, O. Salomón, Numerical homogenization for composite materials  
1166 analysis. comparison with other micro mechanical formulations, Composite Structures 122 (2015)  
1167 405–416. doi:<https://doi.org/10.1016/j.compstruct.2014.11.041>.  
1168 URL <https://www.sciencedirect.com/science/article/pii/S0263822314006102>
- 1169 [26] W. Shu, I. Stanciulesscu, Multiscale homogenization method for the prediction of elastic properties  
1170 of fiber-reinforced composites, International Journal of Solids and Structures 203 (2020) 249–263.  
1171 doi:<https://doi.org/10.1016/j.ijsolstr.2020.08.009>.  
1172 URL <https://www.sciencedirect.com/science/article/pii/S0020768320303085>
- 1173 [27] M. G. D. Geers, V. G. Kouznetsova, K. Matouš, J. Yvonnet, Homogenization Methods and  
1174 Multiscale Modeling: Nonlinear Problems, John Wiley and Sons, Ltd, 2017, pp. 1–34. arXiv:  
1175 <https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781119176817.ecm2107>, doi:<https://doi.org/10.1002/9781119176817.ecm2107>.  
1177 URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119176817.ecm2107>
- 1178 [28] D. Bishara, Y. Xie, W. Liu, S. Li, A state-of-the-art review on machine learning-based multi-  
1179 scale modeling, simulation, homogenization and design of materials, Archives of Computational  
1180 Methods in Engineering 30 (08 2022). doi:[10.1007/s11831-022-09795-8](https://doi.org/10.1007/s11831-022-09795-8).
- 1181 [29] J. Castillero, J. Otero, R. Ramos, A. Bourgeat, Asymptotic homogenization of laminated piezo-  
1182 composite materials, International Journal of Solids and Structures 35 (5) (1998) 527–541.  
1183 doi:[https://doi.org/10.1016/S0020-7683\(97\)00028-0](https://doi.org/10.1016/S0020-7683(97)00028-0).  
1184 URL <https://www.sciencedirect.com/science/article/pii/S0020768397000280>
- 1185 [30] A. L. Kalamkarov, I. V. Andrianov, V. V. Danishevskyy, Asymptotic homogenization of  
1186 composite materials and structures, Applied Mechanics Reviews 62 (3), 030802 (Mar 2009).  
1187 doi:[10.1115/1.3090830](https://doi.org/10.1115/1.3090830).  
1188 URL <https://doi.org/10.1115/1.3090830>
- 1189 [31] J. Fish, Z. Yang, Z. Yuan, A second-order reduced asymptotic homogenization approach for non-  
1190 linear periodic heterogeneous materials, International Journal for Numerical Methods in Engi-  
1191 neering 119 (6) (2019) 469–489. arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1002/nme.6058>, doi:<https://doi.org/10.1002/nme.6058>.  
1192 URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/nme.6058>
- 1193 [32] H. You, M. Xiang, Y. Jing, L. Guo, Z. Yang, A strain gradient phase field model for heterogeneous  
1194 materials based on two-scale asymptotic homogenization, Journal of the Mechanics and Physics  
1195 of Solids 200 (2025) 106104. doi:<https://doi.org/10.1016/j.jmps.2025.106104>.  
1196 URL <https://www.sciencedirect.com/science/article/pii/S0022509625000808>
- 1197 [33] G. Milton, The theory of composites, Cambridge University Press, 2002. doi:[10.1017/CBO9780511613357](https://doi.org/10.1017/CBO9780511613357).
- 1198 [34] J. Xue, Y. Cao, J. Shao, N. Burlion, Study of effective elastic properties of heterogeneous materials  
1199 with an artificial neural network model, Mechanics of Materials 179 (2023) 104597. doi:<https://doi.org/10.1016/j.mechmat.2023.104597>.  
1200 URL <https://www.sciencedirect.com/science/article/pii/S0167663623000431>
- 1201 [35] X. Lu, D. G. Giovanis, J. Yvonnet, V. Papadopoulos, F. Detrez, J. Bai, A data-driven computa-  
1202 tional homogenization method based on neural networks for the nonlinear anisotropic electrical  
1203 response of graphene/polymer nanocomposites, Computational Mechanics 64 (2) (2019) 307–321.
- 1204

- 1207 doi:10.1007/s00466-018-1643-0.  
1208 URL <https://doi.org/10.1007/s00466-018-1643-0>
- 1209 [36] F. Aldakheel, E. S. Elsayed, T. I. Zohdi, P. Wriggers, Efficient multiscale modeling of heterogeneous materials using deep neural networks, Computational Mechanics 72 (1) (2023) 155–171.  
1210 doi:10.1007/s00466-023-02324-9.  
1211 URL <https://doi.org/10.1007/s00466-023-02324-9>
- 1212  
1213 [37] J. Xue, J. Shao, N. Burlion, Estimation of constituent properties of concrete materials with an artificial neural network based method, Cement and Concrete Research 150 (2021) 106614.  
1214 doi:<https://doi.org/10.1016/j.cemconres.2021.106614>.  
1215 URL <https://www.sciencedirect.com/science/article/pii/S0008884621002635>
- 1216  
1217 [38] Q. Huang, D. Hong, B. Niu, D. Long, Y. Zhang, An interpretable deep learning strategy for effective thermal conductivity prediction of porous materials, International Journal of Heat and Mass Transfer 221 (2024) 125064. doi:<https://doi.org/10.1016/j.ijheatmasstransfer.2023.125064>.  
1218 URL <https://www.sciencedirect.com/science/article/pii/S0017931023012097>
- 1219  
1220 [39] Z. Qu, Y. Fu, Y. Liu, L. Zhou, Approach for predicting effective thermal conductivity of aerogel materials through a modified lattice boltzmann method, Applied Thermal Engineering 132 (03 2018). doi:10.1016/j.aplthermaleng.2018.01.013.
- 1221  
1222 [40] C. Bi, G. Tang, Effective thermal conductivity of the solid backbone of aerogel, International Journal of Heat and Mass Transfer 64 (2013) 452–456. doi:<https://doi.org/10.1016/j.ijheatmasstransfer.2013.04.053>.  
1223 URL <https://www.sciencedirect.com/science/article/pii/S0017931013003670>
- 1224  
1225 [41] R. Li, E. Lee, T. Luo, A unified deep neural network potential capable of predicting thermal conductivity of silicon in different phases, Materials Today Physics 12 (2020) 100181. doi:<https://doi.org/10.1016/j.mtphys.2020.100181>.  
1226 URL <https://www.sciencedirect.com/science/article/pii/S2542529320300055>
- 1227  
1228 [42] Y. Luo, Predicting lattice thermal conductivity via machine learning: a mini review, npj Computational Materials 9 (01 2023). doi:10.1038/s41524-023-00964-2.  
1229 URL <https://doi.org/10.1038/s41524-023-00964-2>
- 1230  
1231 [43] S. Wen, X. Li, B. Wang, J. Tan, Y. Liu, J. Lv, Z. Tan, L. Yin, Y. Du, Deep neural network-evaluated thermal conductivity for two-phase wc-m ( $m = ag, co$ ) cemented carbides, Materials 15 (18) (2022). doi:10.3390/ma15186269.  
1232 URL <https://www.mdpi.com/1996-1944/15/18/6269>
- 1233  
1234 [44] M. Kim, J. H. Moon, Deep neural network prediction for effective thermal conductivity and spreading thermal resistance for flat heat pipe, International Journal of Numerical Methods for Heat and Fluid Flow 33 (01 2023). doi:10.1108/HFF-10-2021-0685.  
1235 URL <https://doi.org/10.1108/HFF-10-2021-0685>
- 1236  
1237 [45] Y. Dong, X. Cao, X. Peng, J. Lv, Q. Li, In-pile heat conduction model of the dispersion nuclear fuel plate with particle agglomeration. part ii: Predicting the effective thermal conductivity under the in-pile thermal transfer pattern based on a deep neural network, Journal of Nuclear Materials 607 (2025) 155692. doi:<https://doi.org/10.1016/j.jnucmat.2025.155692>.  
1238 URL <https://www.sciencedirect.com/science/article/pii/S002231152500087X>
- 1239  
1240 [46] C. Shen, Q. Sheng, H. Zhao, Predicting effective thermal conductivity of fibrous and particulate composite materials using convolutional neural network, Mechanics of Materials 186 (2023) 104804. doi:<https://doi.org/10.1016/j.mechmat.2023.104804>.  
1241 URL <https://www.sciencedirect.com/science/article/pii/S0167663623002508>
- 1242  
1243  
1244  
1245  
1246  
1247  
1248  
1249  
1250  
1251  
1252

- 1253 [47] Z. Yang, X. Wu, X. He, X. Guan, A multiscale analysis-assisted two-stage reduced-order deep  
1254 learning approach for effective thermal conductivity of arbitrary contrast heterogeneous materials,  
1255 *Engineering Applications of Artificial Intelligence* 136 (2024) 108916. doi:<https://doi.org/10.1016/j.engappai.2024.108916>.  
1256 URL <https://www.sciencedirect.com/science/article/pii/S0952197624010741>
- 1258 [48] A. Adam, H. Fang, X. Li, Effective thermal conductivity estimation using a convolutional neural  
1259 network and its application in topology optimization, *Energy and AI* 15 (2024) 100310. doi:  
1260 <https://doi.org/10.1016/j.egyai.2023.100310>.  
1261 URL <https://www.sciencedirect.com/science/article/pii/S2666546823000824>
- 1262 [49] P. Liu, Z. Han, W. Wu, Y. Zhao, Y. Song, M. Chai, A developed convolutional neural net-  
1263 work model for accurately and stably predicting effective thermal conductivity of gradient  
1264 porous ceramic materials, *International Journal of Heat and Mass Transfer* 225 (2024) 125428.  
1265 doi:<https://doi.org/10.1016/j.ijheatmasstransfer.2024.125428>.  
1266 URL <https://www.sciencedirect.com/science/article/pii/S001793102400259X>
- 1267 [50] Z. Liu, C. Wu, M. Koishi, A deep material network for multiscale topology learning and accel-  
1268 erated nonlinear modeling of heterogeneous materials, *Computer Methods in Applied Mechanics  
1269 and Engineering* 345 (2019) 1138–1168. doi:<https://doi.org/10.1016/j.cma.2018.09.020>.  
1270 URL <https://www.sciencedirect.com/science/article/pii/S0045782518304729>
- 1271 [51] Z. Liu, C. Wu, Exploring the 3d architectures of deep material network in data-driven multiscale  
1272 mechanics, *Journal of the Mechanics and Physics of Solids* 127 (2019) 20–46. doi:<https://doi.org/10.1016/j.jmps.2019.03.004>.  
1273 URL <https://www.sciencedirect.com/science/article/pii/S0022509618310688>
- 1275 [52] J. Jiang, J. Wu, Q. Chen, G. Chatzigeorgiou, F. Meraghni, Physically informed deep homoge-  
1276 nization neural network for unidirectional multiphase/multi-inclusion thermoconductive compos-  
1277 ites, *Computer Methods in Applied Mechanics and Engineering* 409 (2023) 115972. doi:<https://doi.org/10.1016/j.cma.2023.115972>.  
1278 URL <https://www.sciencedirect.com/science/article/pii/S0045782523000956>
- 1280 [53] R. Chengping, L. Yang, Three-dimensional convolutional neural network (3D-CNN) for hetero-  
1281 geneous material homogenization, *CoRR* abs/2002.07600 (2020). arXiv:2002.07600.  
1282 URL <https://arxiv.org/abs/2002.07600>
- 1283 [54] H. Béji, T. Kanit, T. Messager, Prediction of effective elastic and thermal properties of hetero-  
1284 geneous materials using convolutional neural networks, *Applied Mechanics* 4 (1) (2023) 287–303.  
1285 doi:10.3390/app1mech4010016.  
1286 URL <https://www.mdpi.com/2673-3161/4/1/16>
- 1287 [55] C. Sarada Devi, S. Sahaaya Arul Mary, N. Karthikeyan, S. Varalakshmi, V. Talasila, G. Rama  
1288 Naidu, Graph neural network-based multiscale thermal modeling for heterogeneous materials with  
1289 complex structures, *Thermal Science and Engineering Progress* 55 (2024) 102983. doi:<https://doi.org/10.1016/j.tsep.2024.102983>.  
1290 URL <https://www.sciencedirect.com/science/article/pii/S2451904924006012>
- 1292 [56] J. Hou, Z. Gong, X. Ding, J. Sun, R. Tang, H. Xiao, J. Deng, The accurate prediction and  
1293 further optimization of thermal conductivity for 3d fully ceramic microencapsulated fuel via  
1294 graph convolutional neural network, *Materials Today Communications* 42 (2025) 111557. doi:  
1295 <https://doi.org/10.1016/j.mtcomm.2025.111557>.  
1296 URL <https://www.sciencedirect.com/science/article/pii/S2352492825000698>
- 1297 [57] D.-T. Le, V.-H. Nguyen, S. Mahouche-Chergui, B. Carbonnier, D. Grande, S. Naili, Multiscale  
1298 characterization of effective mechanical properties of graphene-chitosan composite aerogels, *Ma-  
1299 terials Today Communications* 40 (2024) 109491. doi:<https://doi.org/10.1016/j.mtcomm>.

- 1300        2024.109491.  
1301        URL <https://www.sciencedirect.com/science/article/pii/S2352492824014727>
- 1302 [58] D.-T. Le, V.-H. Nguyen, S. Mahouche-Chergui, B. Carbonnier, D. Grande, S. Naili, Multiscale  
1303        characterization of effective thermal properties of graphene/polymer composite aerogels, Com-  
1304        posites Part B 293 (2025) 112106.
- 1305 [59] S. Torquato, Random Heterogeneous Materials: Microstructure and Macroscopic Properties,  
1306        Vol. 55, 2002. doi:[10.1115/1.1483342](https://doi.org/10.1115/1.1483342).
- 1307 [60] H. Sankara, D. Baillis, O. Coulibaly, R. Coquard, N. Naouar, Z. Saghrouni, Computational model  
1308        of effective thermal conductivity of green insulating fibrous media, Materials (Basel, Switzerland)  
1309        17 1 (2024) 252.  
1310        URL <https://doi.org/10.3390/ma17010252>
- 1311 [61] S. Zhao, J. Zhao, Y. Lai, Multiscale modeling of thermo-mechanical responses of granular ma-  
1312        terials: A hierarchical continuum-discrete coupling approach, Computer Methods in Applied  
1313        Mechanics and Engineering 367 (2020) 113100. doi:<https://doi.org/10.1016/j.cma.2020.113100>.  
1315        URL <https://www.sciencedirect.com/science/article/pii/S004578252030284X>