

Data driven Computational Mechanics at EXascale



DCoMEX

Data driven Computational Mechanics at EXascale

Work program topic: EuroHPC-01-2019

Type of action: Research and Innovation Action (RIA)

REPORT ON THE DCoMEX-MAT APPLICATION

DELIVERABLE D7.4 (theory manual)

Version No 1

<http://www.dcomex.eu/>

This project has received funding from the European High-Performance Computing Joint Undertaking Joint Undertaking ('the JU'), under Grant Agreement No 956201



DOCUMENT SUMMARY INFORMATION

Project Title	Data driven Computational Mechanics at EXascale
Project Acronym	DCoMEX
Project No:	956201
Call Identifier:	EuroHPC-01-2019
Project Start Date	01/04/2021
Related work package	WP 7
Related task(s)	Task 7.2
Lead Organisation	UCY
Submission date	18/02/2021
Re-submission date	
Dissemination Level	PU

Quality Control:

	Who	Affiliation	Date
Checked by internal reviewer	George Stavroulakis	NTUA	15/02/2024
Checked by WP Leader	T. Stylianopoulos	UCY	15/02/2024
Checked by Project	Vissarion Papadopoulos	NTUA	15/02/2024

Document Change History:

Version	Date	Author (s)	Affiliation	Comment

Contents

1. Introduction	4
2. Nonlinear computational homogenization on multiscale material models	4
3. Hierarchical Bayesian framework for multiscale material modeling.....	6
4. Sequence of feed forward neural networks as surrogate models	8
5. Numerical Application.....	11
5.1 Multiscale material models under investigation	11
5.1.1 CNT-reinforced cement paste experimental setup and multiscale model	11
5.1.2 CNT-reinforced mortar experimental setup and multiscale model	13
5.1.3 CNT-reinforced concrete experimental setup and multiscale model	14
5.2 Development of the FFNN sequence for each multiscale material model.....	16
5.3 Parameter identification of the CNT-cement paste interfacial properties	19

1 Introduction

In Deliverable 7.4 (theory manual) that is associated with **WP7“Applications”** of the DCoMEX project, a detailed report is provided regarding the implementation of the second application. In this application an efficient and robust framework is developed for the exploration of hyper-performance composite materials. The material models are studied by means of a hierarchical multiscale modeling strategy. The exceptional predictive capabilities of nonlinear computational homogenization are harnessed for the solution of the multiscale material systems. The theoretical backround of nonlinear computational homogenization is summarized in section 2. A hierarchical Bayesian framework that can efficiently integrate diverse data sources towards a joint inference problem is utilized for the parametric investigation of the studied materials. An overview of the proposed hierarchical Bayesian strategy is provided in section 3. To tackle the excessive computational requirements of such an elaborate algorithmic procedure feed forward neural networks (FFNNs) are employed as surrogate models. Their objective is to learn and accurately emulate the nonlinear homogenization equation that defines the material behavior across the multiple length scales of each material. Details regarding the development of the FFNN sequence are given in section 4. The validation of the aforementioned algorithmic machinery is done by means of a challenging application, which is the investigation of the carbon nanotubes (CNTs) interfacial mechanical properties in cementitious material configurations. The results of the application regarding both the offline stage of the FFNN training and the online stage of the hierarchical Bayesian analysis are provided in section 5.

2 Nonlinear computational homogenization on multiscale material models

For an elastic macroscopic body, the equilibrium equation without body forces is defined as:

$$\nabla_M \cdot \boldsymbol{\sigma}_M = \mathbf{0} \text{ in } \Omega_M \quad (2.1)$$

where $\boldsymbol{\sigma}_M$ and Ω_M are the macroscopic stresses and the macroscopic domain respectively.

The material behavior of the macroscopic system is described by the constitutive form:

$$\boldsymbol{\sigma}_M(t) = \mathcal{F}(\boldsymbol{\varepsilon}_M(\tau), \mathbf{h}_M(\tau), \tau \in [0, t]) \quad (2.2)$$

with $\boldsymbol{\varepsilon}_M$ and \mathbf{h}_M being the macroscopic strains and internal material variables. In computational homogenization the relation in eq. (3.2) is attained from the solution of a series of boundary value problems on the representative volume elements (RVE) of the fine scales, as shown in fig. 2.1.

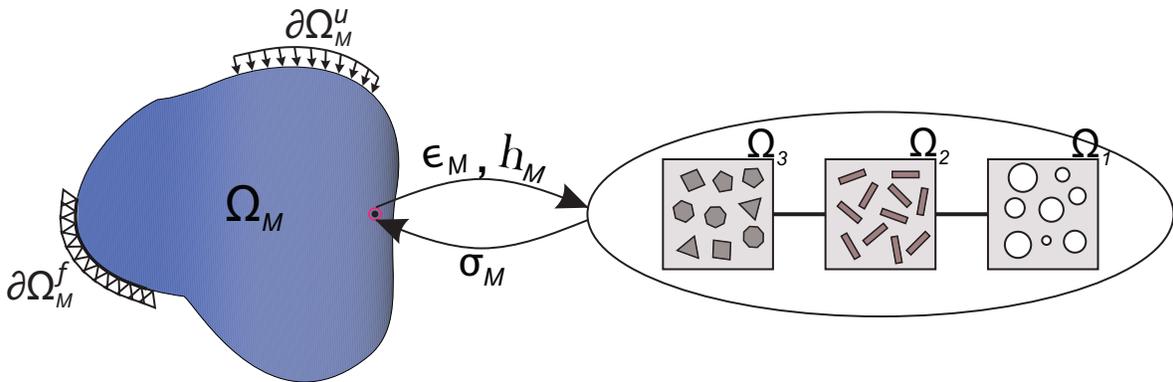


Figure 2.1: The constitutive behavior of a macroscopic system is defined through a series of finer scales

When there is a single distinct finer scale e.g. a microscale μ , then one RVE can be assigned to describe the “material genome” i.e. the properties of the material phases, their interaction and the morphological patterns.

The equilibrium equation in this scale is expressed as:

$$\nabla_{\mu} \boldsymbol{\sigma}_{\mu} = \mathbf{0} \quad \text{in } \Omega_{\mu} \quad (2.3)$$

where $\boldsymbol{\sigma}_{\mu}$ and Ω_{μ} are the stresses and the domain at the microscale.

In the microscale, the constitutive relation is postulated as:

$$\boldsymbol{\sigma}_{\mu}(t) = f(\boldsymbol{\varepsilon}_{\mu}(\tau), \mathbf{h}_{\mu}(\tau), \boldsymbol{\theta}_{\mu}, \tau \in [0, t]) \quad (2.4)$$

where $\boldsymbol{\varepsilon}_{\mu}(\tau)$ and $\mathbf{h}_{\mu}(\tau)$ are the strains and the internal variables of the microscale. In $\boldsymbol{\theta}_{\mu}$ are included parameters such as material or morphological descriptors in the microscale.

The macroscopic effective material properties are obtained after solving the equilibrium imposed by eq. (2.3). The boundary conditions applied on the RVE, such as the linear displacements on $\partial\Omega_{\mu}$ used in this application, have to be consistent with the criterion:

$$\boldsymbol{\sigma}_M : \delta \boldsymbol{\varepsilon}_M = \frac{1}{\mathcal{V}_{\mu}} \int_{\Omega_{\mu}} \boldsymbol{\sigma}_{\mu} : \delta \boldsymbol{\varepsilon}_{\mu} d\Omega_{\mu} \quad (2.5)$$

with \mathcal{V}_{μ} being the RVE volume.

The effective properties are acquired as the average value of the microscopic stress field:

$$\boldsymbol{\sigma}_M = \frac{1}{\mathcal{V}_{\mu}} \int_{\Omega_{\mu}} \boldsymbol{\sigma}_{\mu} d\Omega_{\mu} \quad (2.6)$$

The homogenized tangent modulus is computed through the differentiation of eq. (2.6) with the macroscopic strains:

$$\mathbf{C}_M = \frac{1}{\mathcal{V}_{\mu}} \partial_{\boldsymbol{\varepsilon}_M} \int_{\Omega_{\mu}} \boldsymbol{\sigma}_{\mu} d\Omega_{\mu} \quad (2.7)$$

It cases where the material has unique features in more distinct length scales, a further decomposition of the material model has to be done as shown in fig. 2.1. The macroscopic homogenization operation, in this case, is defined through the N -fold integral:

$$\boldsymbol{\sigma}_M = \frac{1}{\mathcal{V}_N} \int_{\Omega_N} \dots \underbrace{\frac{1}{\mathcal{V}_1} \int_{\Omega_1} \boldsymbol{\sigma}_1 d\Omega_1 \dots d\Omega_N}_{\boldsymbol{\sigma}_2} \dots \underbrace{\dots}_{\boldsymbol{\sigma}_N} \quad (2.8)$$

where the numbering of the scales starting from the finest one is declared by the indices $\{1, \dots, N\}$. The tangent modulus can be computed via a similar generalization as in eq. 2.8

For the solution of a macroscopic system the N -sequence of RVEs have to be solved for each integration point. In non-linear cases, the solution scheme is completed when all scales have converged numerically. This procedure is an extension of the well established FE^2 algorithm to the case of an arbitrary number N of fine scales as a FE^N analysis.

3 Hierarchical Bayesian framework for multiscale material modeling

Consider a set of data $\mathbf{D} = [\mathbf{D}^1, \dots, \mathbf{D}^K]$, with $\mathbf{D}^i = [d^{i,1}, \dots, d^{i,N_i}]$, that encompasses measurements (i.e. mechanical responses) obtained from a number of K independently performed experiments. These experimental scenarios are simulated by the respective multiscale computational models $\mathcal{M}^1, \dots, \mathcal{M}^K$ that are used to predict structural responses $\mathbf{m}^1, \dots, \mathbf{m}^K$ in line with the experimental datasets. The parameterization of these models is done through a distinct set $\boldsymbol{\theta}^i$, as $\boldsymbol{\theta}^i = [\boldsymbol{\theta}_1^i, \dots, \boldsymbol{\theta}_{S_i}^i]$, which includes a series of physical, topological or constitutive attributes for each length scale $s = 1, \dots, S_i$ of the respective system. A subset of these parameters ${}^c\boldsymbol{\theta}^i \subseteq \boldsymbol{\theta}^i$, with ${}^c\boldsymbol{\theta}^i = [{}^c\boldsymbol{\theta}_1^i, \dots, {}^c\boldsymbol{\theta}_{S_i}^i]$, is present in all the investigated multiscale material models.

In this application of the DCoMEX project, a novel hierarchical Bayesian computational framework is developed based on the module 4 of the project. With this framework a robust investigation of the common model parameters ${}^c\boldsymbol{\theta}$ is done by incorporating the knowledge from all the K experimental instances. As a graphical representation of the overall problem at hand, fig. 3.1 depicts a three-model scenario where the parameters of a specific scale are to be investigated. The data sources have the potential to encompass a wide range of experiments conducted across various length scales, for example, macroscale displacement measurements, mesoscale topology characterization through Scanning Electron Microscope (SEM) images and microscale strain field extraction using the Digital Image Correlation (DIG) technique. For the sake of brevity, the common parameters ${}^c\boldsymbol{\theta}$, will be plainly denoted as parameters $\boldsymbol{\theta}$ for the remainder of this report.

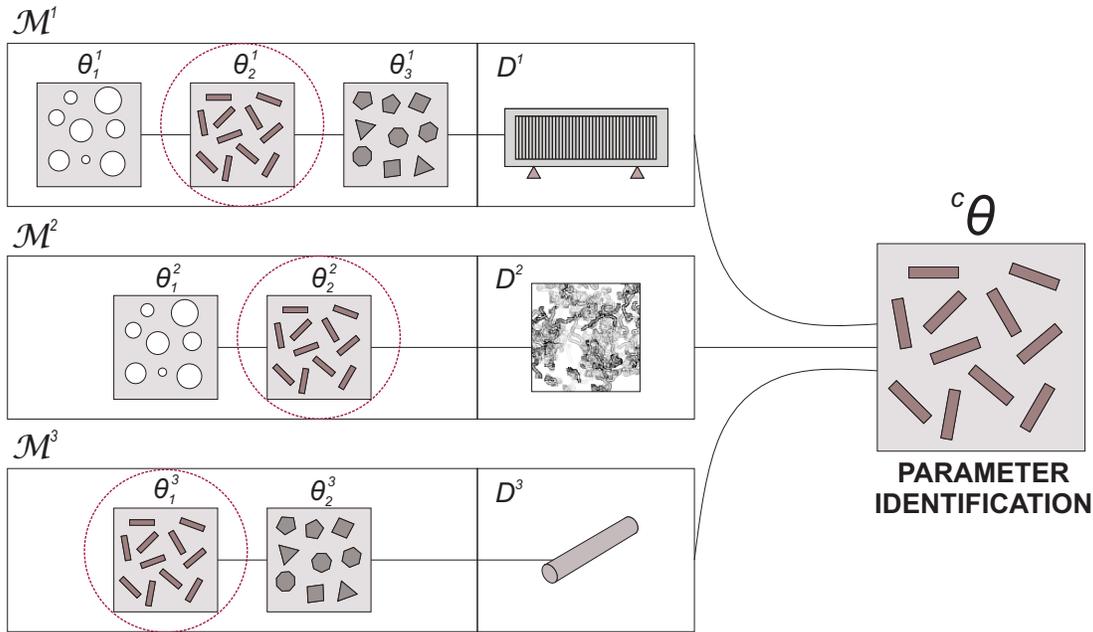


Figure 3.1: Parameter identification of the common properties ${}^c\boldsymbol{\theta}$ found in a series of diverse experimental cases such as a model \mathcal{M}^1 of a bending test on reinforced beam made of a composite material, a model \mathcal{M}^2 of a digital image of a composite material's microstructure and a model \mathcal{M}^3 of a rod specimen of the composite.

Based on the hierarchical Bayesian paradigm, each dataset \mathbf{D}^i obtained from a unique experimental setup is considered separately. For each one of them, the common parameters $\boldsymbol{\theta}$ have a distinct definition as $\boldsymbol{\theta}^i$. This distinction is necessary since these parameters, albeit having the same physical meaning amongst all datasets, they can ultimately be represented by different values due to the external variability. It is further assumed that each $\boldsymbol{\theta}^i$ is conditioned on a series of hyperparameters $\boldsymbol{\psi}$, denoted as $P(\boldsymbol{\theta}^i | \boldsymbol{\psi})$. These parameters are used to provide

in the mathematical framework the sense of affinity amongst θ^i for $i = 1, \dots, K$ and to explicitly account for the model variations across the respective datasets. The hyperparameters generally represent statistical parameters of a pre-specified family of distributions. The hierarchical Bayesian problem is formulated as:

$$y^i(\theta^i|\psi) = m^i(\theta^i|\psi) + \varepsilon^i \quad i = 1, \dots, K \quad (3.1)$$

The directed acyclic graph (DAG) of a standard hierarchical Bayesian scheme is depicted in fig. 3.2. The joint posterior distribution is expressed through Bayes theorem as:

$$P(\theta, \psi|\mathbf{D}) = \frac{P(\mathbf{D}|\theta)P(\theta|\psi)P(\psi)}{P(\mathbf{D})} = \prod_{i=1}^K \left[\prod_{j=1}^{N_i} [P(d^{i,j}|\theta^i)] P(\theta^i|\psi) \right] \frac{P(\psi)}{P(\mathbf{D})} \quad (3.2)$$

where $P(\psi)$ is the prior distribution of the hyperparameters.

An efficient way to solve the problem described in eq. (3.2) is to decouple the total solution process, by sequentially sampling from the marginalized posterior distributions of the physical parameters and the hierarchical hyperparameters respectively. The data are generally sparse and there are potentially strong nonlinear phenomena that describe the material behavior which could unpredictably alter the form of the target distributions. Therefore, to acquire the precise expression of the posterior PDFs the Transitional Markov Chain Monte Carlo (TMCMC) algorithm is utilized for drawing samples from both marginal distributions.

The first step is to sample from the marginal posterior distribution of the model hyperparameters, which is postulated as:

$$P(\psi|\mathbf{D}) = \int_{\Omega_\theta} P(\mathbf{D}|\theta)P(\theta|\psi)d\theta \frac{P(\psi)}{P(\mathbf{D})} = \prod_{i=1}^K \left[\int_{\Omega_{\theta^i}} \prod_{j=1}^{N_i} [P(d^{i,j}|\theta^i)] P(\theta^i|\psi) d\theta^i \right] \frac{P(\psi)}{P(\mathbf{D})} \quad (3.3)$$

where in the above equation we have used the fact that the likelihood function $P(\mathbf{D}|\psi)$ assumes the form:

$$P(\mathbf{D}|\psi) = \prod_{i=1}^K \left[\int_{\Omega_{\theta^i}} \prod_{j=1}^{N_i} [P(d^{i,j}|\theta^i)] P(\theta^i|\psi) d\theta^i \right] \quad (3.4)$$

The integral of eq. (3.3) can be approximated via Monte Carlo sampling:

$$P(\psi|\mathbf{D}) \simeq \prod_{i=1}^K \left[\frac{1}{N_{\theta^i}} \sum_{k=1}^{N_{\theta^i}} P(\theta_k^i|\psi) \right] \frac{P(\psi)}{P(\mathbf{D})} \quad (3.5)$$

To perform this integration, samples from each likelihood $P(\mathbf{D}^i|\theta^i)$ associated with the model M^i , have to first be collected. This is achieved efficiently by employing the TMCMC algorithm.

The marginal distribution of the updated multiscale model parameters θ^{new} that take into account all datasets and can be applied in future predictions is expressed as:

$$P(\theta^{new}|\mathbf{D}) = \int_{\Omega_\psi} P(\theta^{new}|\psi)P(\psi|\mathbf{D})d\psi \quad (3.6)$$

The approximate solution of eq. (3.6) is done by generating an amount of samples for ψ through the empirical distribution of $P(\psi|\mathbf{D})$ obtained from eq. (3.5), as follows:

$$P(\theta^{new}|\mathbf{D}) \simeq \sum_{k=1}^{N_\psi} P(\theta^{new}|\psi_k) \quad (3.7)$$

For the production of samples from eq. (3.6), the TCMCM algorithm is again used, where in this case the initial sampling is performed on the hyperparameter prior $P(\boldsymbol{\psi})$. On the contrary to sampling from $P(\mathbf{D}^i|\boldsymbol{\theta}^i)$, eq. (3.7) represents a cheap-to-evaluate procedure, since no multiscale model resolutions are needed.

After the posterior probabilistic form of $\boldsymbol{\theta}^{new}$ has been obtained following the hierarchical Bayesian scheme, these parameters can then be used towards any uncertainty propagation analysis on future simulations Y^{new} as shown in figure 3.2. These tests could be conducted on unseen material systems that are partially described by the inferred parameters. The quantity of interest (e.g. a structural response) \mathbf{y}^{new} is calculated as:

$$P(\mathbf{y}^{new}|\mathbf{D}) = \int_{\Omega_{\boldsymbol{\theta}^{new}}} P(\mathbf{y}^{new}|\boldsymbol{\theta}^{new})P(\boldsymbol{\theta}^{new}|\mathbf{D})d\boldsymbol{\theta}^{new} \quad (3.8)$$

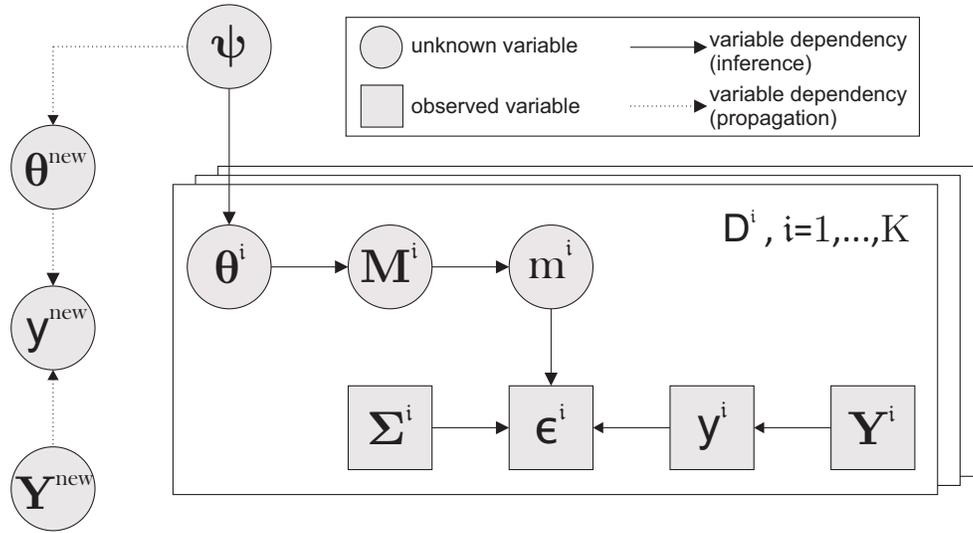


Figure 3.2: Directed Acyclic Graph (DAG) of the general hierarchical problem

4 Sequence of feed forward neural networks as surrogate models

To facilitate the proposed hierarchical Bayesian framework presented in section 3, a novel surrogate modeling strategy is utilized for the cost mitigation of the FE^N analyses in the context of computational homogenization. The key idea is to use a sequence of FFNNs, to learn the parametrized non-linear homogenized response of the RVEs on each scale in a hierarchical manner, starting from the finest scale and progressively substituting each RVE with a FFNN that encapsulates the material behavior at all previous scales.

The steps for constructing the FFNN-aided N -scale nested scheme are the following:

- A sequential homogenization procedure is implemented according to eqs. (2.6) on each pair of consecutive scales starting from the finest one, the 1st scale.
- A set of parameters/variables that define the input and output of the corresponding FFNN f_1^{NN} are defined and a number of total training samples is assigned. The samples include the number n of the different RVE solutions, while each solution further contains all the converged solution steps t of the nonlinear incremental analysis. The input samples involve the components of the strain vectors at the upper scale $\{\{\boldsymbol{\epsilon}_2\}_{1:t}\}_{1:n}$, while the outputs are the stress vectors $\{\{\boldsymbol{\sigma}_2\}_{1:t}\}_{1:n}$ at the end of the homogenization procedure for each RVE solution. If we further assume that each RVE solution is affected by a vector of parameters $\boldsymbol{\theta}_1$ that

characterize the material's constitutive law at the finest scale, then these parameters are also regarded as input samples to the FFNN f_1^{NN} as $\{\boldsymbol{\theta}_1\}_{1:n}$.

- During the online solution, the tangent constitutive matrix \mathbf{C}_2 for the tuple of the strain state and parametric vector $[\boldsymbol{\varepsilon}_2, \boldsymbol{\theta}_1]$, required for the Newton-Raphson iterations, is effortlessly computed using Automatic Differentiation (AD). By using differentiable activation functions in the FFNN, such as the logistic or hyperbolic tangent function, the FFNN becomes a differentiable function. This technique allows for the computation of the derivatives of the output $\boldsymbol{\sigma}_2$ with respect to the input $\boldsymbol{\varepsilon}_2$ by applying the chain rule on the FFNN and, thus, the elements $c_{2,ij}$ of the macroscopic tangent matrix $\mathbf{C}_2 = [c_{2,ij}]$ are obtained as:

$$c_{2,ij} = \frac{\partial \sigma_{2,ij}}{\partial h_k} \frac{\partial h_k}{\partial h_{k-1}} \dots \frac{\partial h_1}{\partial \varepsilon_{2,ij}} \quad (4.1)$$

where h_k is the output at the k_{th} hidden layer of the FFNN for input $\boldsymbol{\varepsilon}_{2,ij}$.

- After the successful training and validation, the finest scale FFNN f_1^{NN} can be straightforwardly applied to represent the constitutive relation of the matrix material at the next scale (scale 2). The process is iterated for this scale and a second FFNN f_2^{NN} is built using $[\{\{\boldsymbol{\varepsilon}_3\}_{1:t}, \boldsymbol{\theta}_2, \boldsymbol{\theta}_1\}_{1:n}]$ as input and $\{\{\boldsymbol{\sigma}_3\}_{1:t}\}_{1:n}$ as output. To simplify notation we write the input $[\{\{\boldsymbol{\varepsilon}_3\}_{1:t}, \boldsymbol{\theta}_2, \boldsymbol{\theta}_1\}_{1:n}]$ as $[\{\{\boldsymbol{\varepsilon}_3\}_{1:t}, \hat{\boldsymbol{\theta}}_2\}_{1:n}]$, where, in the general case, $\hat{\boldsymbol{\theta}}_{s+1} = [\boldsymbol{\theta}_s, \dots, \boldsymbol{\theta}_1]$ is the augmented parametric vector. Again, the tangent matrix \mathbf{C}_3 is readily available through AD during the online stage. It is important to note here that the FFNN f_2^{NN} of the second scale also involves the material parameters $\boldsymbol{\theta}_1$ of the previous scale as input, since this will allow it to capture the behavior of the material at both the first and the second scale.
- This procedure is repeated for all scales up to the macroscale M , where it ultimately results in a single final FFNN that incorporates all the information from the lower scales and constitutes the surrogate model of the composite material's behavior.

To better illustrate the proposed surrogate modelling strategy a 4-scale example of a CNT-reinforced concrete is presented herein. First, the FFNN f_1^{NN} is trained to substitute the RVE of the microscale, which consists of cement paste and CNTs, using $[\{\{\boldsymbol{\varepsilon}_2\}_{1:t}, \hat{\boldsymbol{\theta}}_1\}_{1:n}]$ as input and $\{\{\boldsymbol{\sigma}_2\}_{1:t}\}_{1:n}$ as output. Next, the FFNN f_2^{NN} which represents the fine mesoscale RVE (cement mortar) is trained on pairs $[\{\{\boldsymbol{\varepsilon}_3\}_{1:t}, \hat{\boldsymbol{\theta}}_2\}_{1:n}]$ and $\{\{\boldsymbol{\sigma}_3\}_{1:t}\}_{1:n}$, where now the FFNN f_1^{NN} is considered as the matrix material along with the fine aggregates as the inclusions. The process is repeated one more time for the FFNN f_3^{NN} of the coarse mesoscale RVE, which is the final FFNN that encapsulates all previous FFNNs and represents the constitutive relation for the macroscale problem. The algorithmic procedure of the training is illustrated in figure [4.1](#).

This surrogate modeling strategy can then be applied towards the cost reduction of the multi-query analysis imposed by the hierarchical Bayesian problem described in section [3](#). A unique hierarchy of FFNNs can be developed for each investigated multiscale model M^i . By that, each likelihood function $P(\mathbf{D}^i | \boldsymbol{\theta}^i)$ is replaced by an inexpensive to calculate counterpart which is denoted as $P^{NN,i}(\mathbf{D}^i | \boldsymbol{\theta}^i)$. The algorithm [1](#) encapsulates the complete algorithmic machinery, i.e. the offline stage of the FFNN development and the online stage of the hierarchical Bayesian analysis, for the solution of the studied problem.

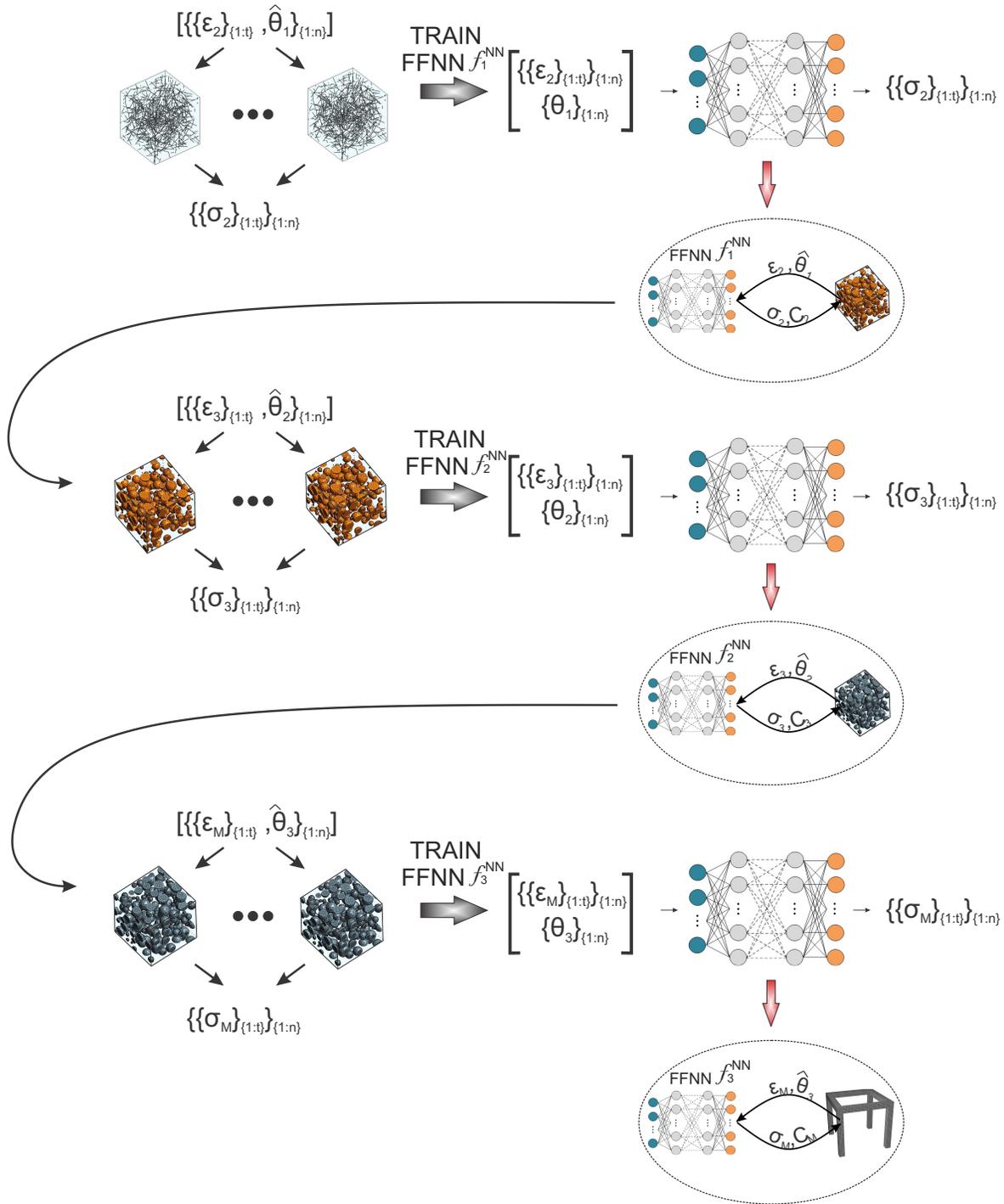


Figure 4.1: Training procedure according to the proposed strategy. Starting from the microscale, a neural network, f_1^{NN} is trained to emulate the stress-strain behavior of the microscale RVE. This f_1^{NN} is used in exchange of the host material in the fine mesoscale and, next, f_2^{NN} is trained to emulate the stress-strain behavior of the fine mesoscale RVE. The process is repeated one more time for the coarse mesoscale until the final network f_3^{NN} successfully encapsulates the overall composite material's behavior.

Input : Total training samples n , total increments t and FFNN hyperparameters. Parameterized multiscale material models $M^i(\boldsymbol{\theta}^i)$, prior distributions $P(\boldsymbol{\psi})$ and $P(\boldsymbol{\theta}^i)$. Total samples $N^{\boldsymbol{\theta}^i}$ and TMCMC parameters k^i and β^i . Total samples N^Ψ and TMCMC parameters k^{new} and β^{new} .

Output : parameter updated distribution $P(\boldsymbol{\theta}^{\text{new}}|\mathbf{D})$, quantity of interest updated distribution $P(\mathbf{y}^{\text{new}}|\mathbf{D})$

Offline Stage:

```

1 for  $i \leftarrow 1$  to  $K$  do
2   for  $j \leftarrow 1$  to  $S_i$  do
3     Generate and store training samples  $\{\{\boldsymbol{\varepsilon}_{j+1}\}_{1:t}, \hat{\boldsymbol{\theta}}_j^i\}_{1:n}$ ;
4     Solve the BVP of scale  $j$ ;
5     Store solution stresses  $\{\{\boldsymbol{\sigma}_{j+1}\}_{1:t}\}_{1:n}$ ;
6     Train the FFNN  $f_j^{NN,i}$ ;
7     Store FFNN  $f_j^{NN,i}$ ;
8   end
9 end
    
```

Online Stage:

```

10 for  $i \leftarrow 1$  to  $K$  do
11   Generate from  $P^{NN,i}(\mathbf{D}^i|\boldsymbol{\theta}^i)$  samples  $\{\boldsymbol{\theta}^i\}_{1:N^{\boldsymbol{\theta}^i}} \leftarrow \text{TMCMC}(p(\boldsymbol{\theta}^i), P^{NN,i}(\mathbf{D}^i|\boldsymbol{\theta}^i), N^i, k^i, \beta^i)$ ;
12   Store samples  $\{\boldsymbol{\theta}^i\}_{1:N^{\boldsymbol{\theta}^i}}$  and likelihood function data  $P^{NN,i}(\mathbf{D}^i|\{\boldsymbol{\theta}^i\}_{1:N^{\boldsymbol{\theta}^i}})$ 
13 end
14 Estimate  $P(\boldsymbol{\psi}|\mathbf{D})$ ;
15 Generate from  $P(\boldsymbol{\psi}|\mathbf{D})$  samples  $\{\boldsymbol{\psi}\}_{1:N^\Psi} \leftarrow \text{TMCMC}(P(\boldsymbol{\psi}), P(\mathbf{D}|\boldsymbol{\psi}), N^\Psi, k^\Psi, \beta^\Psi)$ ;
16 Estimate  $P(\boldsymbol{\theta}^{\text{new}}|\mathbf{D})$ ;
17 Propagate posterior uncertainty  $P(\mathbf{y}^{\text{new}}|\mathbf{D})$ ;
    
```

Algorithm 1: Surrogate model enhanced algorithm for multiscale material parameter inference and uncertainty propagation

5 Numerical Application

5.1 Multiscale material models under investigation

For the validation of the framework developed for the second application of the DCoMEX WP 7, a mechanical characterization is performed on the interfacial properties of CNTs in cementitious materials. Based on experimental measurements acquired from the literature, three different multiscale material models are investigated in the study. These include a 3-point bending test of a CNT-reinforced cement specimen, a tension test performed on CNT-reinforced mortar rods and a 4-point bending experiment on a CNT-reinforced concrete beam. Next, the details for each experimental case will be presented along with the multiscale model that reproduces each material behavior.

5.1.1 CNT-reinforced cement paste experimental setup and multiscale model

The first dataset was obtained from a 3-point bending test on a fully hydrated (28 days) cement paste coupon enhanced with a 0.3% weight fraction of CNTs. The testing beam specimen had dimensions $160\text{mm} \times 40\text{mm} \times 40\text{mm}$, while for its support two rollers, 100mm apart, were used. A single gradual point load was applied on the center of the upper part via a third roller. The diameter of the CNTs varied between 10nm and 20nm and their length between $10\mu\text{m}$ and $20\mu\text{m}$. The experimental setting and the measurements that relate the flexural strain

with the respective stress based on the experimental findings are depicted in fig. 5.1. To integrate the CNTs into the FE analysis the Molecular Structural Mechanics (MSM) technique was used for their simulation. Following the MSM method, the covalent bonds that are developed between the carbon atoms are reproduced by structural space frame elements with tailored mechanical properties to replicate the effect of the force field constants of the carbon-carbon bonds. To reach the desired weight fraction, a significant amount of CNTs need to be inserted into the RVE. To this purpose the high degree-of-freedom (DOF) space frame CNT molecular models were projected into Equivalent Beam Elements (EBE) by mapping several structural responses of the space frame to equivalent mechanical properties of the EBE. Subsequently, a series of EBEs were positioned randomly inside the volume of the RVE until the weight fraction requirement is achieved. A visual representation of the CNT/cement paste RVE and the macroscopic model are given in fig. 5.2. The Drucker-Prager (DP) plasticity material law was used for the modelling of the constitutive behavior of the cement paste matrix. The CNTs were assumed to have an elastic behavior, while their interaction with the surrounding cement paste was modeled through a bond-slip bi-linear constitutive law. As shown in fig. 5.3 this interfacial law is constituted of three parameters, namely the interfacial shear strength $\tau_{int}^{y,1}$, the elastic stiffness $k_{int}^{el,1}$ before the slippage and the inelastic stiffness $k_{int}^{pl,1}$ after the slippage. Therefore, the parametric vector θ^1 is comprised by these three microscale parameters $\theta^1 = [\tau_{int}^{y,1}, k_{int}^{el,1}, k_{int}^{pl,1}]$. In the context of the finite element analysis, to integrate the CNTs in the composite material system the cohesive zone method was used in combination with an embedding technique.

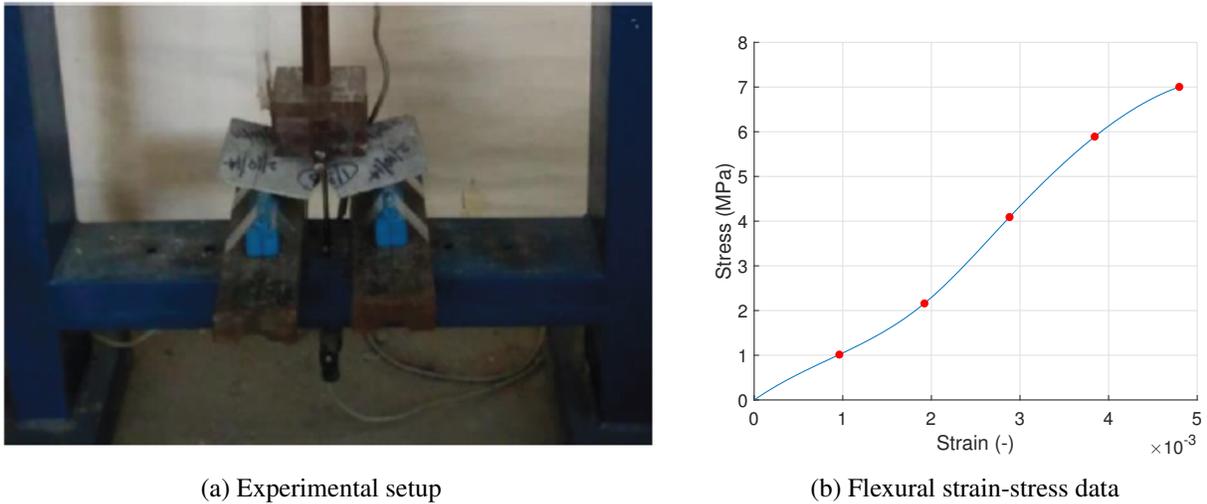


Figure 5.1: CNT-reinforced cement specimen

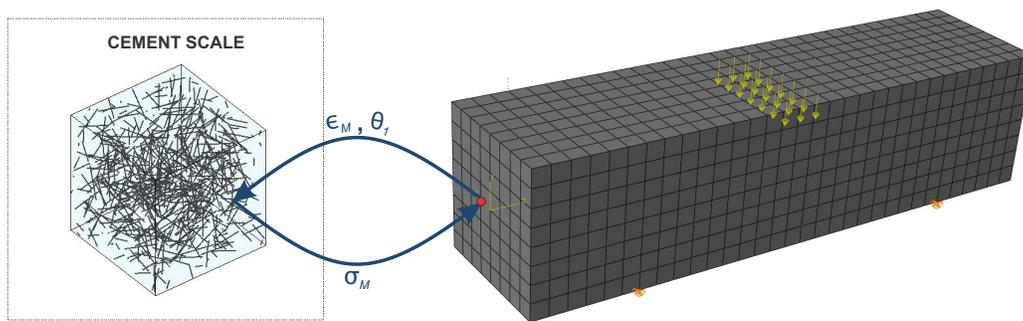


Figure 5.2: CNT-reinforced cement multiscale model

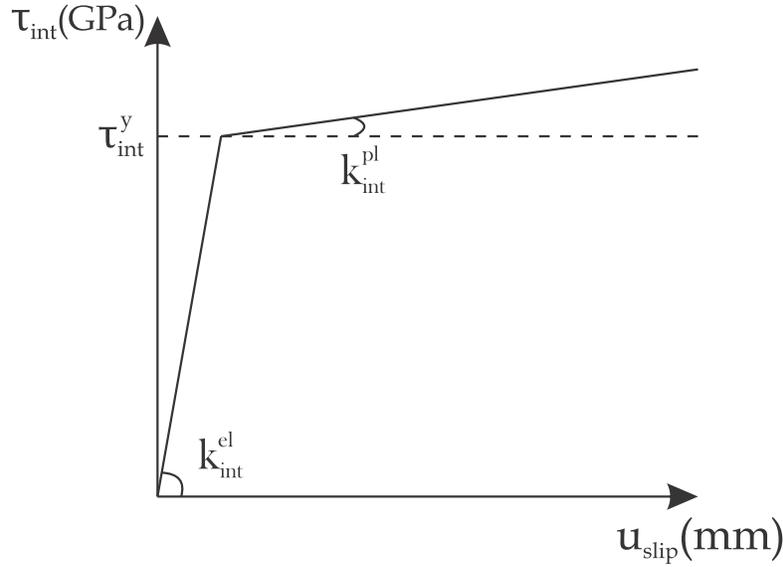


Figure 5.3: Constitutive law that defines the interaction between the CNTs and the matrix material

5.1.2 CNT-reinforced mortar experimental setup and multiscale model

The second experimental data source is a tension test conducted on a cylindrical rod made of mortar with a 0.5% weight fraction of CNTs inserted as fillers. The specifications of the rod coupon were 500mm for the length and 30mm for the diameter. The rod was fixed at the end, while at the other one a gradually increasing tension load was applied. In this laboratory study the CNTs had diameters between 10nm and 30nm, while the length fluctuated between 1 μ m and 2 μ m. Characteristic specimens and the tensile strain-stress dataset for the multi-walled CNT enhanced mortar bar are provided in fig. 5.4. The finest length scale of the material, which is the cement paste reinforced with the CNTs was formed according to the previously presented model of the cement scale as shown in fig. 5.5. For constructing the model of the mortar scale a mesoscale RVE was developed, which includes the sand particles inclusions that construe the mortar, an additional length scale was added. These aggregates were modeled as spherical inclusions with varying diameters. To replicate as realistically as possible the diameter distribution of the inclusions a Fuller grading curve was enforced. The minimum diameter of the candidate spheres were taken as $d_{min} = 0.1\text{mm}$, while the maximum diameter as $d_{max} = 2\text{mm}$. Their positional placement was again done randomly with a special caution that non-overlapping conditions between the inclusions are satisfied. Perfectly elastic conditions were assigned on the inclusions. The two-scale material configuration that characterizes the macroscopic behavior and the FE macromodel used to replicate the test are illustrated in fig. 5.5. Likewise to the previous CNT-reinforced cement material model, the material constitutive behavior is affected by the parameters that specify the CNT/matrix interfacial behavior in fig. 5.3. According to the hierarchical Bayesian scheme presented in section 3, these parameters assume a separate formulation for model \mathcal{M}^2 as $\theta^2 = [\tau_{int}^{y,2}, k_{int}^{el,2}, k_{int}^{pl,2}]$.

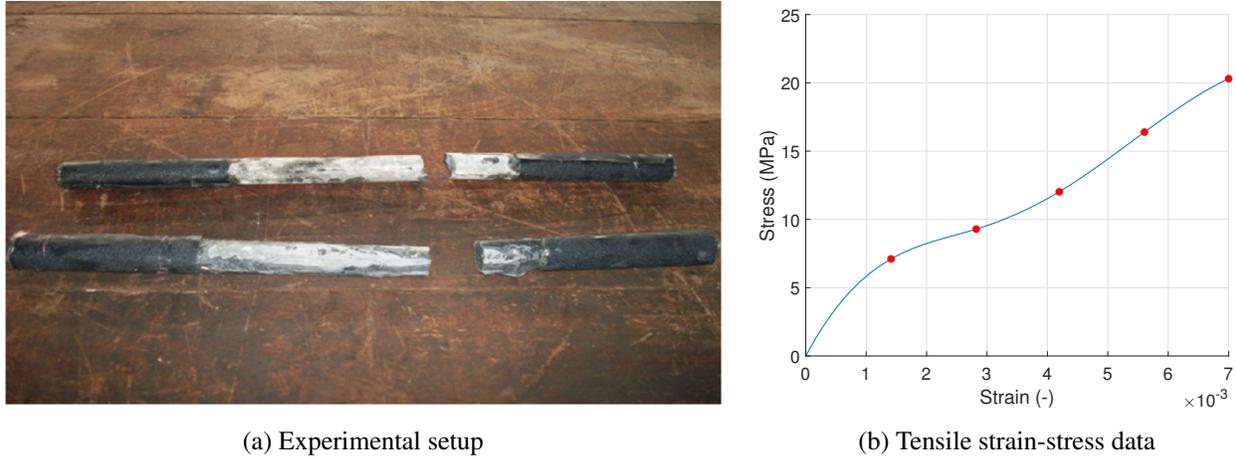


Figure 5.4: CNT-reinforced mortar specimen

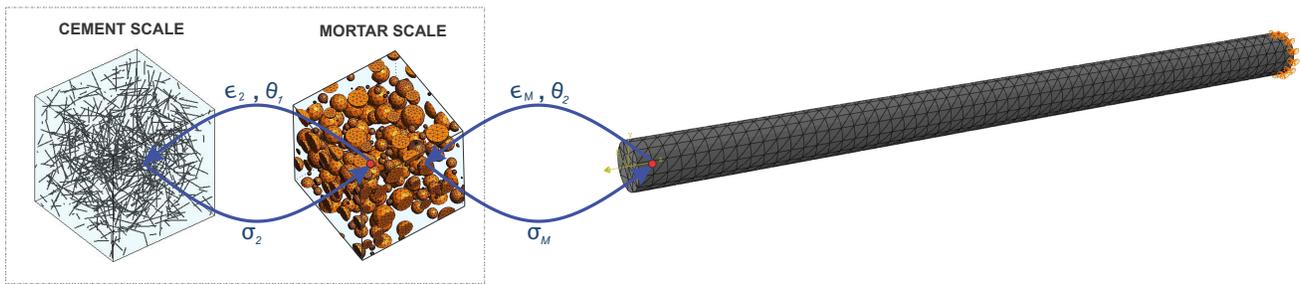
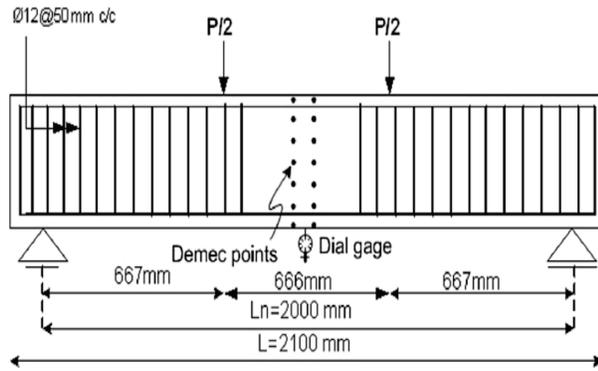


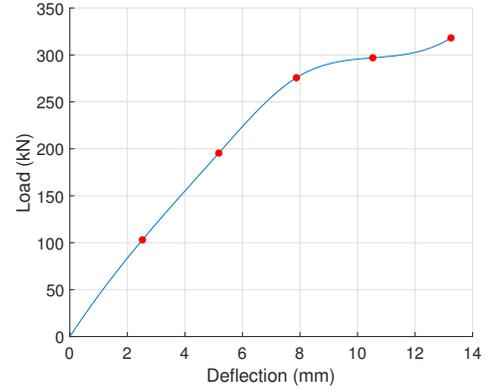
Figure 5.5: CNT-reinforced mortar multiscale model

5.1.3 CNT-reinforced concrete experimental setup and multiscale model

The third dataset comes from a 4-point bending test on a concrete beam reinforced with steel rebar and further strengthened with CNTs of 1% weight fraction. The length of the beam was 2100mm and the cross-section $150\text{mm} \times 250\text{mm}$, while supports were used, 2000m apart, at its lower part. The CNTs had diameters of $3 - 15\text{nm}$ and lengths of $15 - 330\mu\text{m}$. For the flexural test, two gradual point loads were applied on the upper part of the beam. The specifications of this experiment and the corresponding flexural displacement-load observations are presented in fig. 5.6. In this scenario the material is represented by a cement, mortar and concrete three-scale model as shown in fig. 5.7 linked hierarchically to the macroscopic FE model of the beam. The first two scales were formulated in accordance with the procedure described in subsections 5.1.1 and 5.1.2, while the final scale models the coarse aggregates at the mesoscale of the concrete specimen. For this simulation a Fuller grading curve was once more applied for the generation of inclusions of various sizes inside the RVE. The minimum and maximum diameters in this case are $d_{min} = 2\text{mm}$ and $d_{max} = 20\text{mm}$ respectively. The coarse aggregates were assumed to behave linearly elastic. Again, following the hierarchical Bayesian concept, the parametric vector of the constitutive law in fig. 5.3 that describes the interaction between the CNTs and the matrix in this model \mathcal{M}^3 is explicitly defined as $\theta^3 = [\tau_{int}^{v,3}, k_{int}^{el,3}, k_{int}^{pl,3}]$.

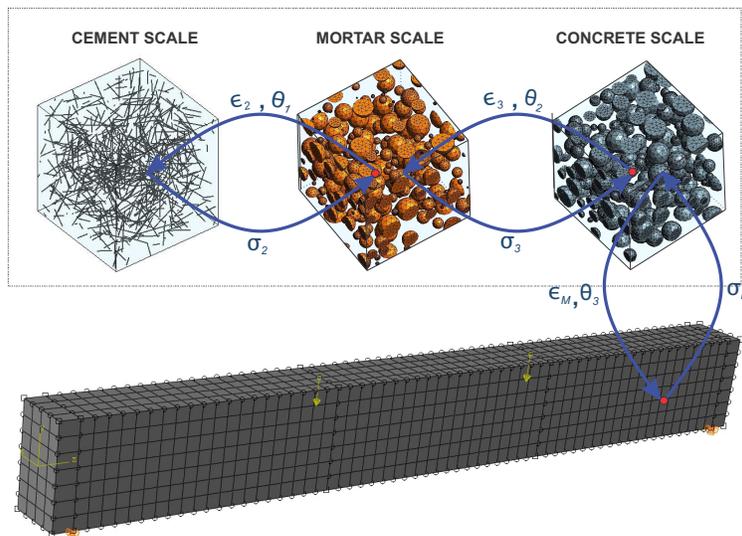


(a) Experimental setup

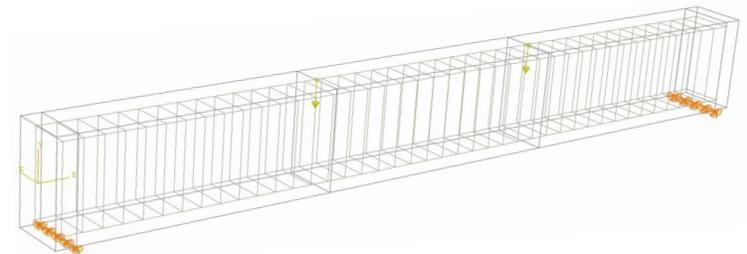


(b) Flexural Displacement-Load data

Figure 5.6: CNT-reinforced concrete specimen



(a)



(b)

Figure 5.7: (a) CNT-reinforced concrete multiscale model (b) geometric configuration of concrete beam

5.2 Development of the FFNN sequence for each multiscale material model

According to the strategy summarized in section 4 only one FFNN - $f_1^{NN,1}$ is sufficient for the reproduction of the macroscopic constitutive response of the first model presented in fig. 5.2. To acquire the dataset needed for the training of $f_1^{NN,1}$, 1000 unique parametrized strain sequences with 20 incremental steps each were generated. The selection of the lower and upper bounds for the sampling of input training space is presented in table 5.1. After the solution of the 1000 microstructure BVPs of the RVE model of the microscale, the data pairs $\{\{\boldsymbol{\varepsilon}_M\}_{1:20}, \hat{\boldsymbol{\theta}}_1^1\}_{1:1000}$ and $\{\{\boldsymbol{\sigma}_M\}_{1:20}\}_{1:1000}$ were used as input and output pairs for the training and testing of $f_1^{NN,1}$. The Adam optimizer with a learning rate of $\eta = 0.001$ and a batch size of 128 was chosen for the calibration of the $f_1^{NN,1}$ parameters. The Mean Squared Error (MSE) among the directly simulated stresses $\{\{\boldsymbol{\sigma}_M\}_{1:20}\}_{1:1000}$ and the respective predicted stresses from $f_1^{NN,1}$ was used as the loss metric. Regarding the $f_1^{NN,1}$ architecture, 3 hidden dense layers with 30 neurons and a hyperbolic tangent activation function for each one were selected. To prevent excessive computational times, a limit of 2000 epochs was set for the training. For the training process, the data were split in three subsets, namely the train, test and validation subset with ratios 70%, 15% and 15% respectively. In fig. 5.8, the training progress tracks the MSE loss for each optimization iteration (epoch), while the prediction accuracy was calculated by employing the L2 norm of each stress component in the test dataset and comparing it to the prediction of the $f_1^{NN,1}$.

Bounds	ε_{11} [-]	ε_{12} [-]	ε_{22} [-]	k_{int}^{el} [GPa/nm]	k_{int}^{pl} [GPa/nm]	τ_{int}^y [GPa]
Lower	-0.03	-0.03	-0.03	0	0	0
Upper	0.03	0.03	0.03	30	3	0.3

Table 5.1: Input sample ranges for the FFNN training

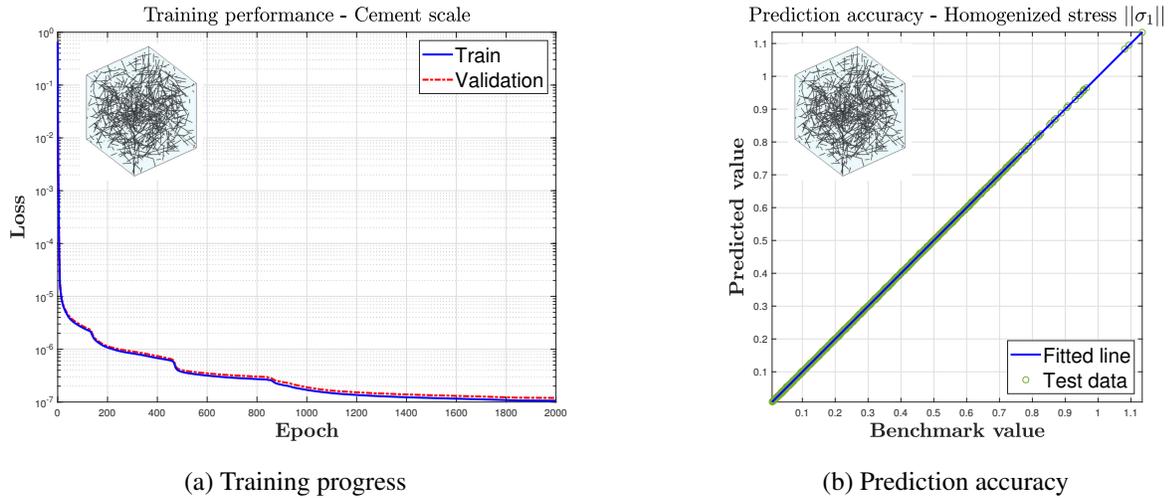


Figure 5.8: FFNN training results for the CNT-reinforced cement specimen - Cement scale

The training process was repeated again for the second material model shown in fig. 5.5 by starting from the microscale. In this scenario, the differences in the material properties, compared to the first model described in sec. 5.1.1, necessitate the training to start from the formulation of the FFNN - $f_1^{NN,2}$ which learns the CNT/cement homogenized behavior. The next step is the development of the FFNN - $f_2^{NN,2}$ which gives the total

macroscopic behavior. A set of data pairs $\{\{\boldsymbol{\epsilon}_2\}_{1:20}, \hat{\boldsymbol{\theta}}_1^2\}_{1:1000}$ and $\{\{\boldsymbol{\sigma}_2\}_{1:20}\}_{1:1000}$ were initially obtained by 1000 CNT/cement RVE solutions. These were then used for the training and testing of $f_1^{NN,2}$, which was subsequently used as the matrix material of the CNT/mortar scale. The next step was to gather data $\{\{\boldsymbol{\epsilon}_M\}_{1:20}, \hat{\boldsymbol{\theta}}_2^2\}_{1:1000}$ and $\{\{\boldsymbol{\sigma}_M\}_{1:20}\}_{1:1000}$ via 1000 solutions of the CNT/mortar RVE and then used them to train and test $f_2^{NN,2}$. For both $f_1^{NN,2}$ and $f_2^{NN,2}$ the same choices as the previous model were made regarding the FFNN architecture, the preprocessing and the training hyperparameters. The accuracy of both FFNNs is captured in figs. 5.9 and 5.10, where the progression of the MSE metric during the training and the quality of the predictions are visualised, respectively.

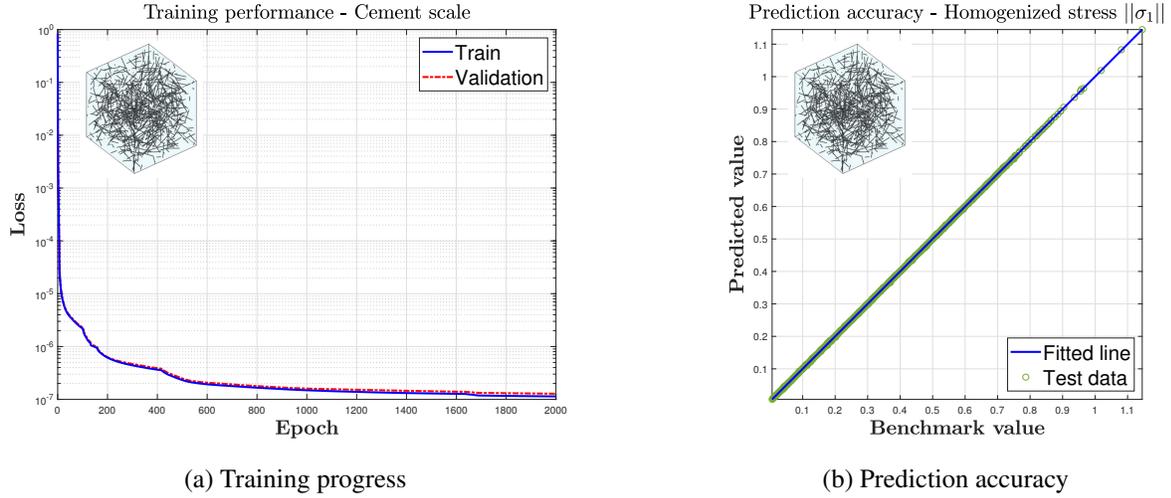


Figure 5.9: FFNN training results for the CNT-reinforced mortar specimen - Cement scale

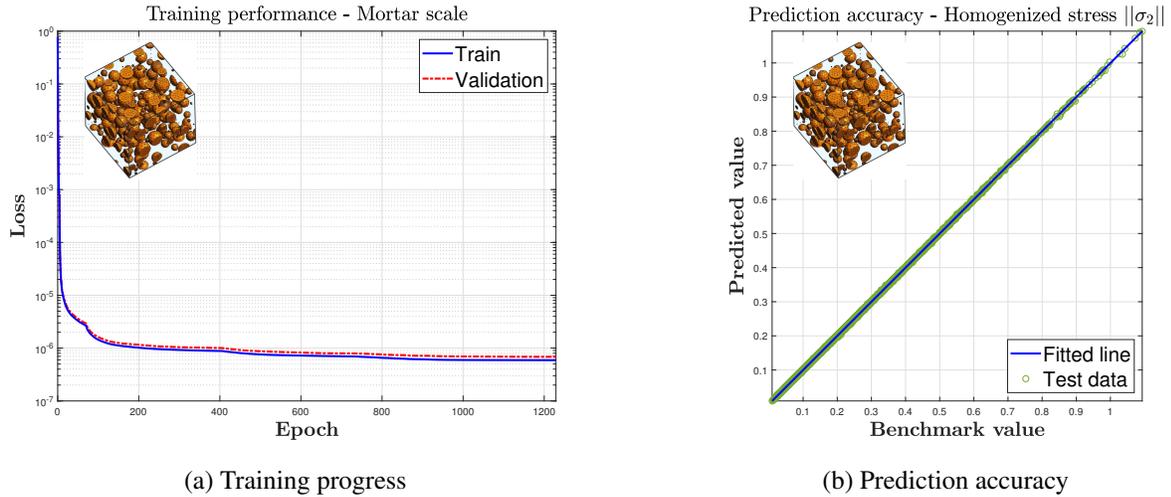


Figure 5.10: FFNN training results for the CNT-reinforced mortar specimen - Mortar scale

For the final model of the CNT/concrete beam specimen in fig. 5.7, a sequence of three FFNNs, namely the $f_1^{NN,3}$, the $f_2^{NN,3}$ and the $f_3^{NN,3}$ were constructed. Following the same concept as in the previous models, the procedure was initialized from the finest scale, which is the CNT/cement scale, by solving the BVP imposed by the

homogenization equation for 1000 different strain sequence/parameter cases $\{\{\boldsymbol{\epsilon}_2\}_{1:20}, \hat{\boldsymbol{\theta}}_1^3\}_{1:1000}$ and retrieving the respective stress outputs $\{\{\boldsymbol{\sigma}_2\}_{1:20}\}_{1:1000}$. With these data we were able to train and test the first FFNN - $f_1^{NN,3}$ which was then used towards the realization of the CNT/mortar data i.e. $\{\{\boldsymbol{\epsilon}_3\}_{1:20}, \hat{\boldsymbol{\theta}}_2^3\}_{1:1000}$ and $\{\{\boldsymbol{\sigma}_3\}_{1:20}\}_{1:1000}$. After training and testing the second FFNN - $f_2^{NN,3}$ through the utilization of the CNT/mortar dataset, the last step was to generate the CNT/concrete data pairs $\{\{\boldsymbol{\epsilon}_M\}_{1:20}, \hat{\boldsymbol{\theta}}_3^3\}_{1:1000}$ and $\{\{\boldsymbol{\sigma}_M\}_{1:20}\}_{1:1000}$ and use them to train and test the third FFNN - $f_3^{NN,3}$. All the choices regarding the formulation and training aspects of the FFNNs were made likewise to the two previous models. The results concerning the training process and the prediction accuracy based on the stress L2 norms are given in figs. 5.11, 5.12 and 5.13 for the three scales respectively.

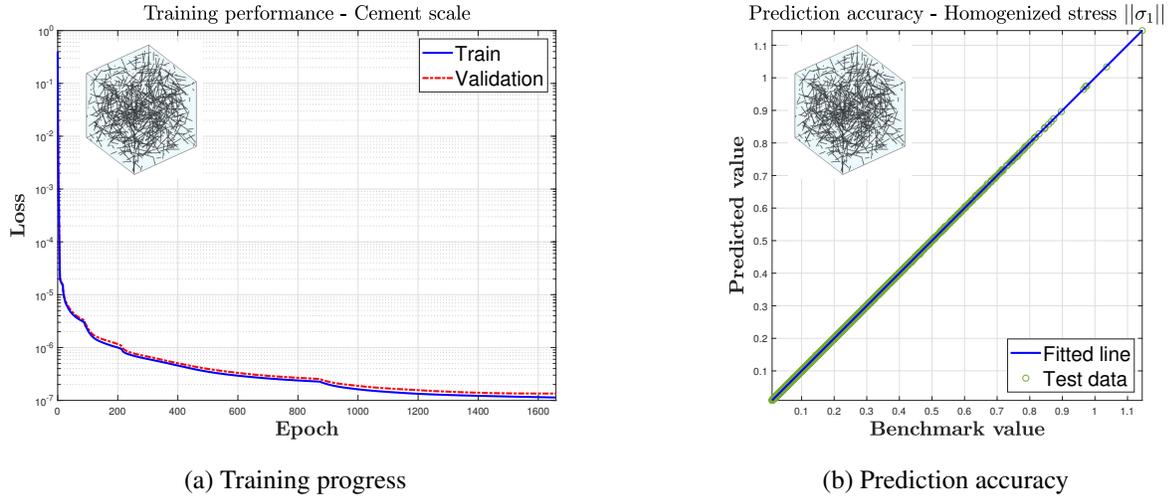


Figure 5.11: FFNN training results for the CNT-reinforced concrete specimen - Cement scale

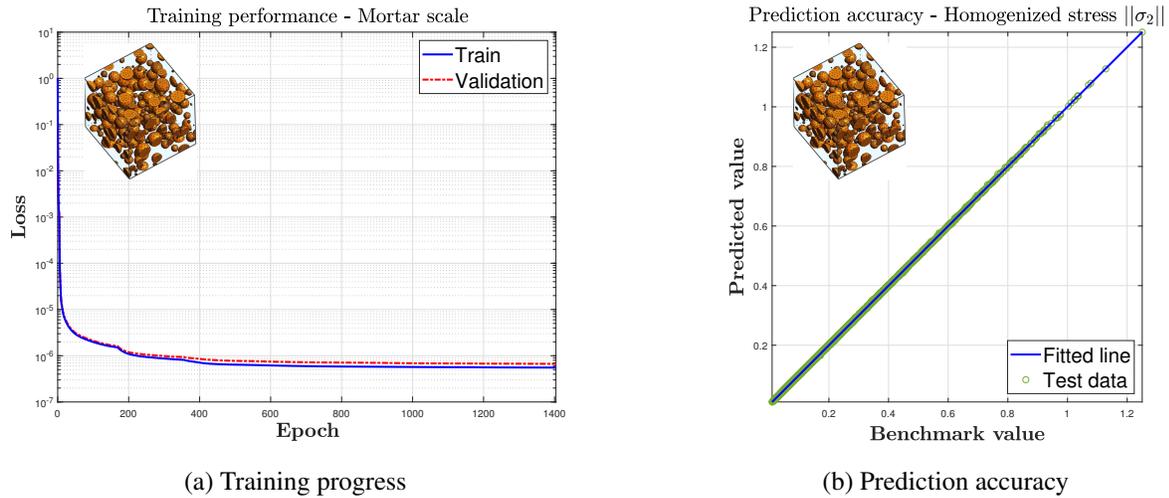


Figure 5.12: FFNN training results for the CNT-reinforced concrete specimen - Mortar scale

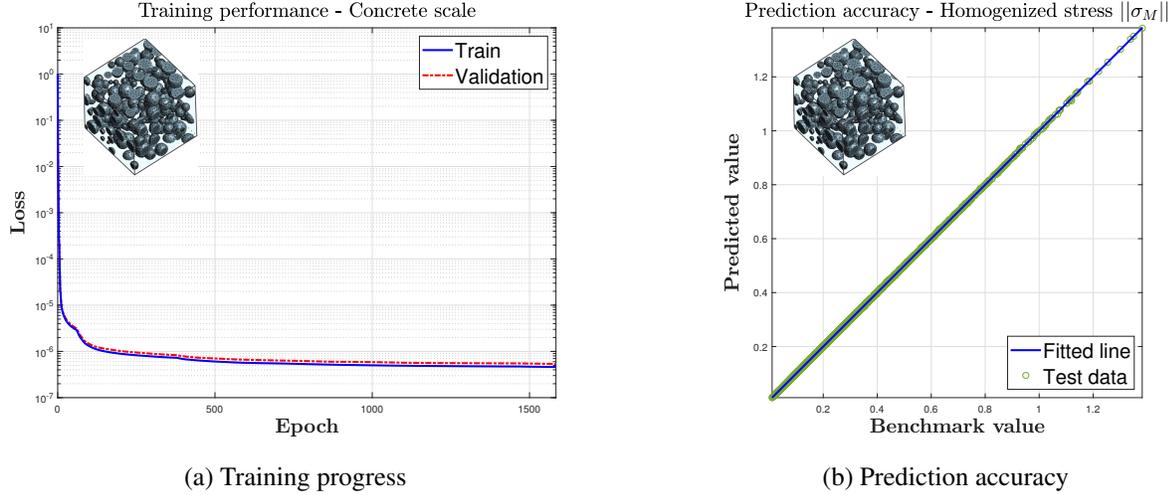


Figure 5.13: FFNN training results for the CNT-reinforced concrete specimen - Concrete scale

5.3 Parameter identification of the CNT-cement paste interfacial properties

After the parametrized constitutive response of the three investigated models has been learned by the FFNN surrogates, we are ready to move to the online procedure of the proposed hierarchical Bayesian strategy as presented in algorithm 1. The data via which the investigated parameters will be updated were obtained from the figures 5.1b, 5.4b and 5.6b for the three models respectively. From each of the three curves, five equally spaced points were retrieved and used towards the uncertainty quantification as shown in the aforementioned figures. The initial step was to perform an independent TMCMC sampling from the likelihood function $P(\mathbf{D}^i | \boldsymbol{\theta}^i)$ of each model $\mathcal{M}^i, i = 1, 2, 3$. Since the solution process at this stage is decoupled, the posterior sampling for all models was enforced in a computationally parallel manner. This is an important feature as it partially counters the drawback of the MCMC algorithms which are serial by default and enables the option to employ a high number of models in the hierarchical Bayesian framework. The prior distributions $P(\boldsymbol{\theta}^i)$ were chosen as uniform distributions with their upper and lower bounds selected based on the parameter training bounds of the FFNNs. Therefore, for each model the priors were defined as $P(k_{int}^{el,i}) \sim \mathcal{U}(0, 30)$, $P(k_{int}^{pl,i}) \sim \mathcal{U}(0, 3)$ and $P(t_{int}^{y,i}) \sim \mathcal{U}(0, 0.3)$. The standard deviation of the error was calculated based on a coefficient of variation of 0.02. For the hyperparameters of the TMCMC algorithm, the values $k^i = 1.0$ and $\beta^i = 0.2$ were selected, while the number of the samples was chosen as $N^i = 10000$. The posterior form of each $\boldsymbol{\theta}^i$ is illustrated in figs. 5.14, 5.15 and 5.16.

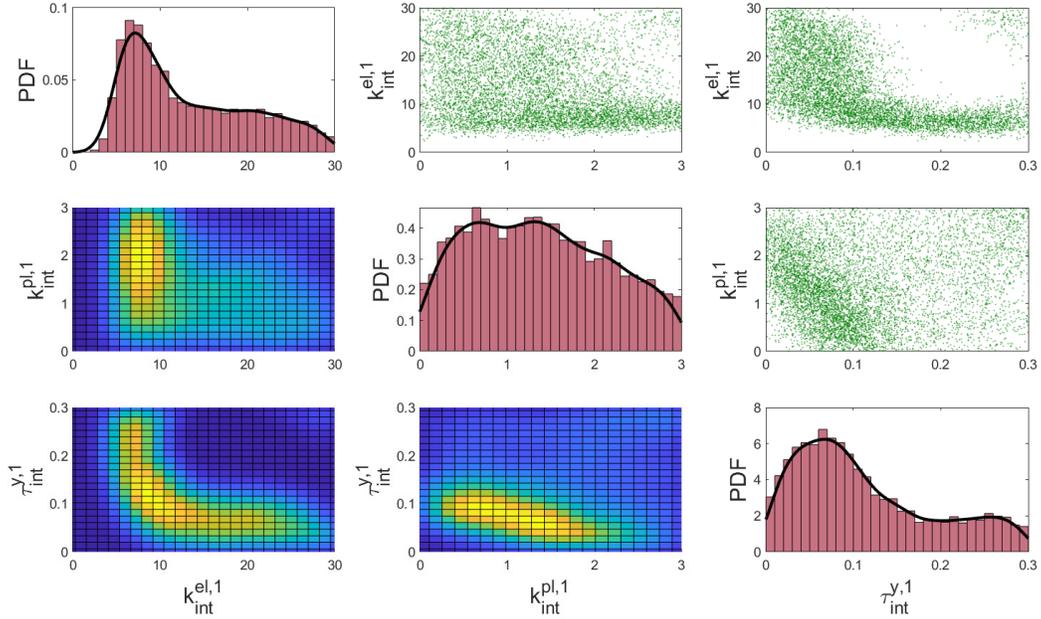


Figure 5.14: Results of the Bayesian analysis on the CNT/cement interfacial parameters of model \mathcal{M}^1 . Diagonal - Marginal probability density functions of the investigated parameters. Upper triangle - Scatter plots for each parameter pair. Lower triangle - Joint probability density functions for each parameter pair

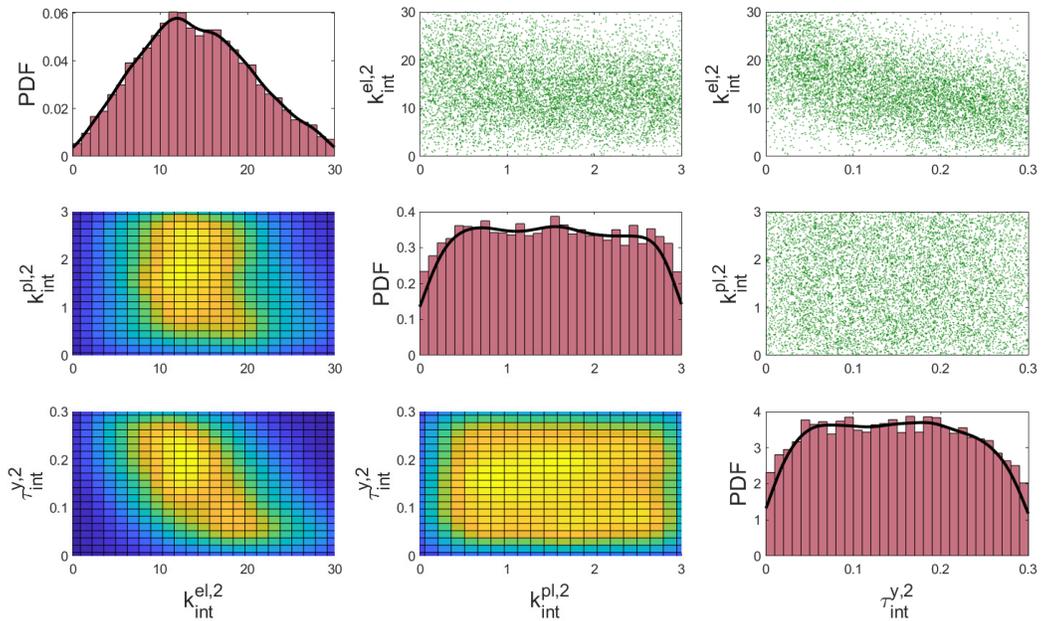


Figure 5.15: Results of the Bayesian analysis on the CNT/mortar interfacial parameters of model \mathcal{M}^2 . Diagonal - Marginal probability density functions of the investigated parameters. Upper triangle - Scatter plots for each parameter pair. Lower triangle - Joint probability density functions for each parameter pair

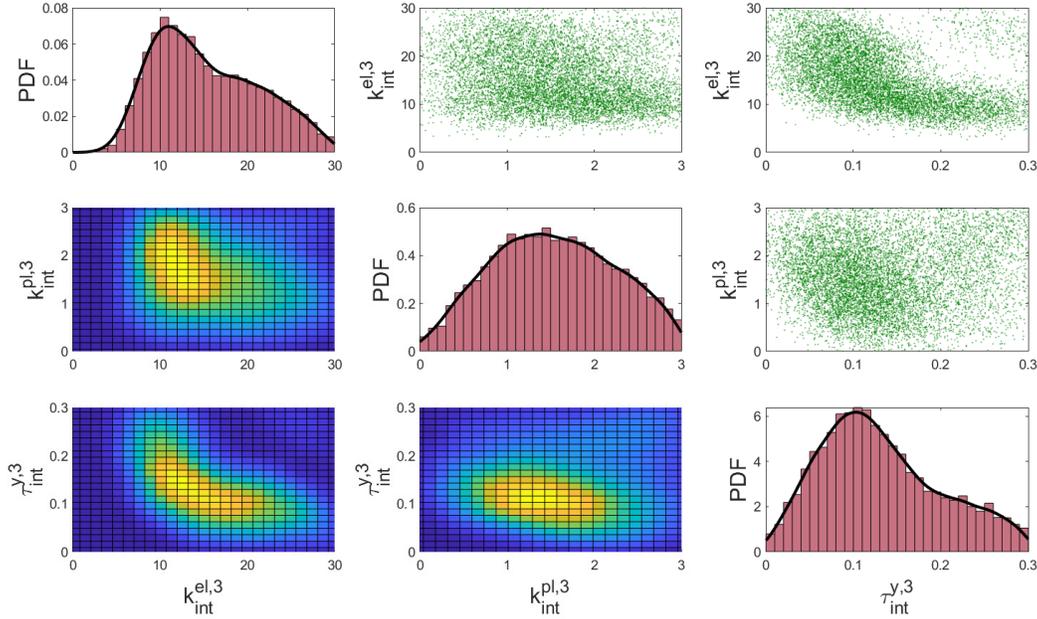


Figure 5.16: Results of the Bayesian analysis on the CNT/concrete interfacial parameters of model \mathcal{M}^3 . Diagonal - Marginal probability density functions of the investigated parameters. Upper triangle - Scatter plots for each parameter pair. Lower triangle - Joint probability density functions for each parameter pair

After all the necessary samples $\{\boldsymbol{\theta}^i\}_{1:N^{\theta^i}}$ have been collected, the next step is to formulate the hyperparameter posterior PDF as postulated in eq. (2.7). Again, following the process of algorithm ?? we used the TMCMC to sample from the empirical distribution $P(\boldsymbol{\psi}|\mathbf{D})$. The selection of the hyperparameter priors $P(\boldsymbol{\psi})$ and the parameter priors $P(\boldsymbol{\theta}^i|\boldsymbol{\psi})$ is given in table 5.2. At this stage, the TMCMC ran for $N^\Psi=50000$ samples and the hyperparameters were again appointed as $k^{new} = 1.0$ and $\beta^{new} = 0.2$. The hyperparameter posterior PDFs are given in fig. 5.17. The last step of algorithm 1 is to construct the probabilistic form of the new parameters $\boldsymbol{\theta}^{new}$ by utilizing the posterior samples $\{\boldsymbol{\psi}\}_{1:N^\Psi}$. Since each sample in $\{\boldsymbol{\psi}\}_{1:N^\Psi}$ defines a uniform distribution with certain bounds, a discrete mixture distribution could be formed by considering all these uniform PDFs $P(\boldsymbol{\theta}^{new}|\{\boldsymbol{\psi}\}_{1:N^\Psi})$. By doing so, eq. (??) is expressed as:

$$P(\boldsymbol{\theta}^{new}|\mathbf{D}) = \sum_{i=1}^{N^\Psi} w_i \mathcal{U}(\{\boldsymbol{\psi}^1\}_i, \{\boldsymbol{\psi}^1\}_i + \{\boldsymbol{\psi}^2\}_i) , \text{ with } w_i = \frac{1}{N^\Psi} , i = 1, \dots, N^\Psi \quad (5.1)$$

Sampling from such an empirical mixture PDF was straightforward. The resulting distributions $P(\boldsymbol{\theta}^{new}|\mathbf{D})$ are presented in fig. 5.18.

Hyperparameter	$\Psi_{k_{int}^{el}}^1$	$\Psi_{k_{int}^{el}}^2$	$\Psi_{k_{int}^{pl}}^1$	$\Psi_{k_{int}^{pl}}^2$	$\Psi_{\tau_{int}^y}^1$	$\Psi_{\tau_{int}^y}^2$
Prior PDF	$\mathcal{U}(0, 20)$	$\mathcal{U}(0, 10)$	$\mathcal{U}(0, 2)$	$\mathcal{U}(0, 1)$	$\mathcal{U}(0, 0.2)$	$\mathcal{U}(0, 0.1)$
Parameter (i=1,2,3)	$k_{int}^{el,i}$		$k_{int}^{pl,i}$		$\tau_{int}^{y,i}$	
Prior PDF	$\mathcal{U}(\Psi_{k_{int}^{el}}^1, \Psi_{k_{int}^{el}}^1 + \Psi_{k_{int}^{el}}^2)$		$\mathcal{U}(\Psi_{k_{int}^{pl}}^1, \Psi_{k_{int}^{pl}}^1 + \Psi_{k_{int}^{pl}}^2)$		$\mathcal{U}(\Psi_{\tau_{int}^y}^1, \Psi_{\tau_{int}^y}^1 + \Psi_{\tau_{int}^y}^2)$	

Table 5.2: Prior distributions of the hyperparameters and the parameters of the tackled hierarchical Bayesian problem

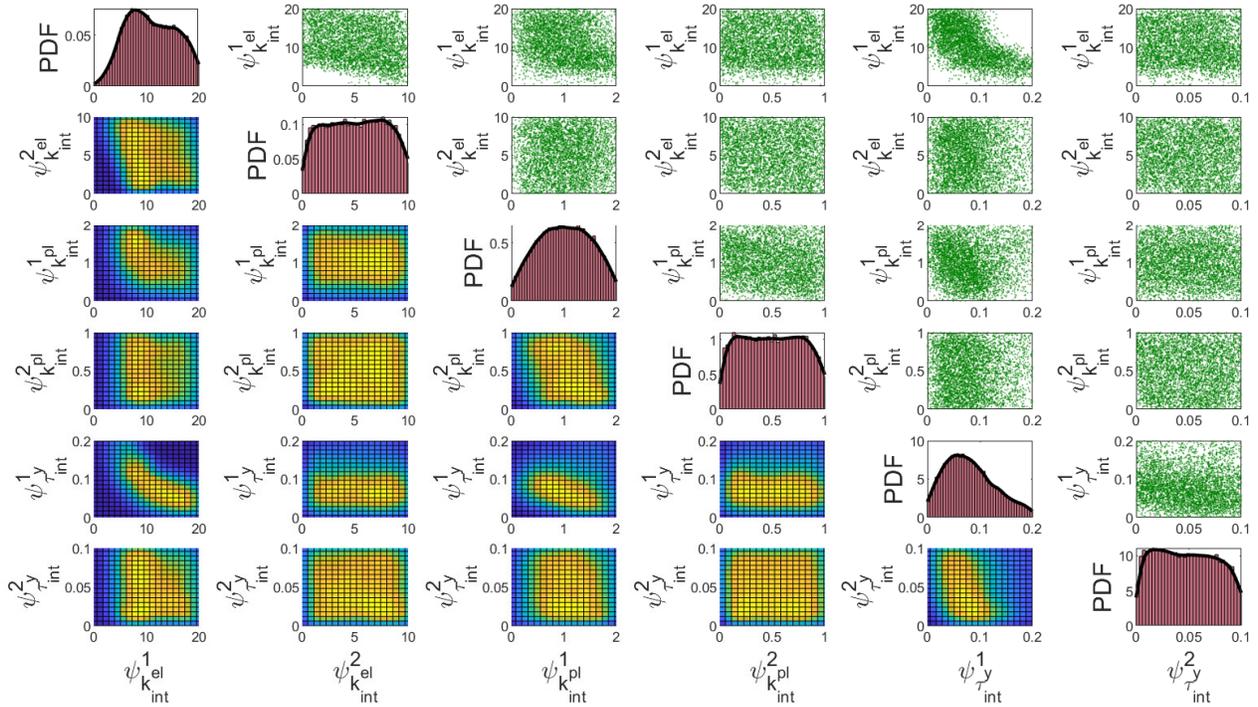


Figure 5.17: Results of the hierarchical Bayesian analysis on the hyperparameters. Diagonal - Marginal probability density functions of the hyperparameters. Upper triangle - Scatter plots for each hyperparameter pair. Lower triangle - Joint probability density functions for each hyperparameter pair

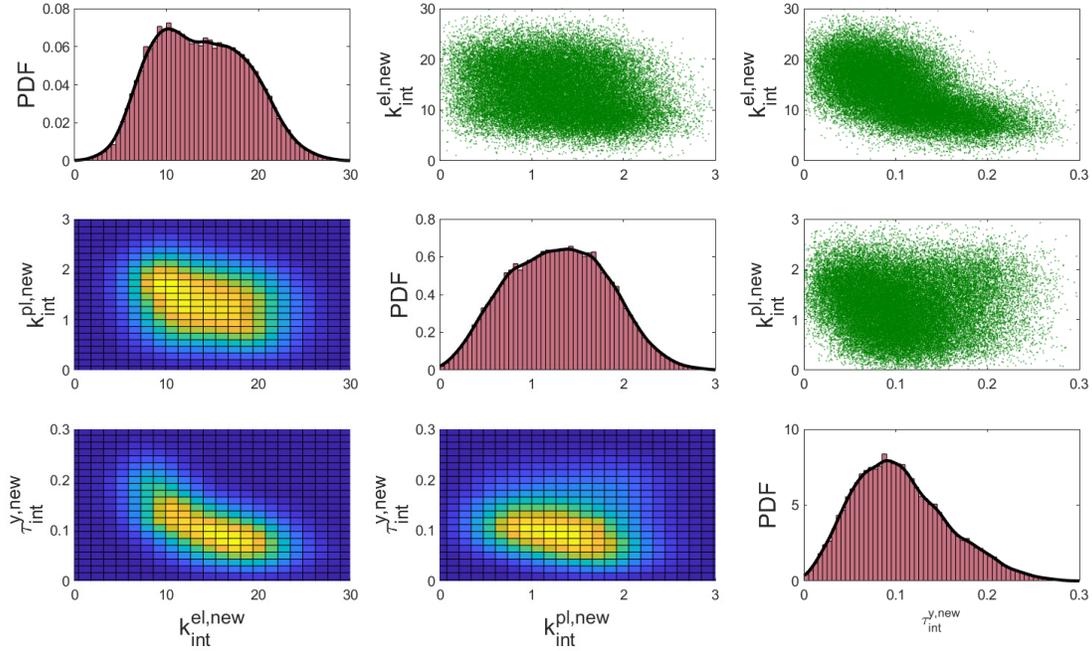
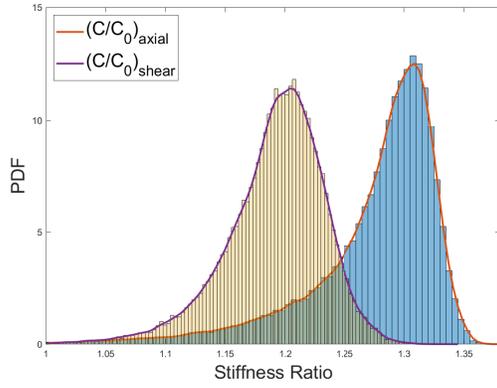
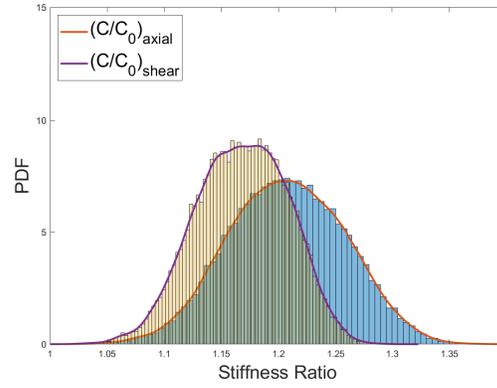


Figure 5.18: Results the of hierarchical Bayesian analysis on the newly formulated CNT/matrix interfacial parameters. Diagonal - Marginal probability density functions of the new parameters. Upper triangle - Scatter plots for each new parameter pair. Lower triangle - Joint probability density functions for each new parameter pair

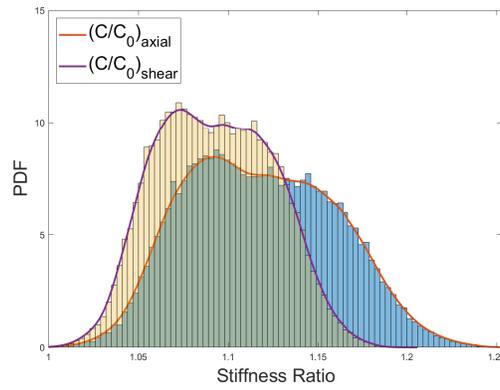
Finally, we performed an uncertainty propagation analysis on how the informed parameters $\theta^{new} = [k_{int}^{el,new}, k_{int}^{pl,new}, \gamma_{int}^{y,new}]$ impact the mechanical behavior of the respective material for each model $\mathcal{M}^i, w = 1, 2, 3$. For that, we investigated the influence of the CNT/matrix cohesive properties θ^{new} on the macroscopic elastic stiffness matrix \mathbf{C}_M^i . A standard Monte Carlo analysis was conducted on the full material composition of each model, which is described by the single scale model of fig. 5.2 for the CNT/cement specimen, the two-scale model of fig. 5.5 for the CNT/mortar specimen and the three-scale model of fig. 5.7 for the CNT/concrete specimen. The objective is to find the distributions for the axial and shear components of the stiffness tensor. To ensure that the plasticity conditions are not met, we applied a relatively small strain value and subsequently solved the homogenization problem. The inexpensive emulators in the form of the FFNNs enabled us to perform these analyses in negligible computational time since the homogenized strain-stress relation can be instantly extracted from each FFNN, while the constitutive matrix can then be easily obtained using Automatic Differentiation. The material in all cases is isotropic due the random orientation of the CNTs, therefore both the axial and shear stiffness is identical in all directions. The components $(C/C^0)_{axial}$ and $(C/C^0)_{shear}$ of the elasticity tensor, where C^0 is the respective stiffness value of the plain material, are presented in fig. 5.19. On average, the stiffness ratio for both axial and shear components is higher in model \mathcal{M}^1 , while the lowest ratio is observed in model \mathcal{M}^3 . The results are consistent with the expected outcome, since compared to the cement model \mathcal{M}^1 the addition of fine aggregates in model \mathcal{M}^2 and fine and coarse aggregates in model \mathcal{M}^3 leads to stiffer materials and reduces the impact of the CNT reinforcement on that aspect.



(a) Cement material of \mathcal{M}^1



(b) Mortar material of \mathcal{M}^2



(c) Concrete material of \mathcal{M}^3

Figure 5.19: Posterior distribution of the stiffness improvement due to the CNT reinforcement in the axial and shear components for each material model by considering the informed PDF $P(\boldsymbol{\theta}^{new}|\mathbf{D})$ of the interfacial parameters.