AI-enhanced multiscale finite element methods for forward and inverse uncertainty quantification problems in structural mechanics

By
Stefanos Christos Pyrialakos

School of Civil Engineering
Institute of Structural Analysis and Antiseismic Research

National Technical University of Athens

Supervisor: Professor Vissarion Papadopoulos

A thesis submitted for the degree of
Doctor of Philosophy

June, 2024

APPROVAL

PhD THESIS EXAMINATION COMMITTEE:

Professor Vissarion Papadopoulos
(Supervisor and Principal Advisor of the Committee)
National Technical University of Athens
School of Civil Engineering

Professor Konstantinos Spiliopoulos
(Member Advisor of the Committee)
National Technical University of Athens
School of Civil Engineering

Professor Nikolaos Lagaros
(Member Advisor of the Committee)
National Technical University of Athens
School of Civil Engineering

Professor Christos Zeris
(Member of the Examination Committee)
National Technical University of Athens
School of Civil Engineering

Professor Costas Charitidis
(Member of the Examination Committee)
National Technical University of Athens
School of Chemical Engineering

Associate Professor Michalis Fragiadakis
(Member of the Examination Committee)
National Technical University of Athens
School of Civil Engineering

Associate Professor Savvas Triantafyllou
(Member of the Examination Committee)
National Technical University of Athens
School of Civil Engineering

TO MY PARENTS AND MY BROTHER.

Acknowledgments

Undertaking this PhD has profoundly transformed my life, and I owe immense gratitude to the many individuals who supported and guided me along the way.

First and foremost, I extend my heartfelt appreciation to my parents and my brother for their unwavering belief in me and relentless encouragement to pursue my aspirations. Their boundless support and unconditional love have been the cornerstone of my journey, for which I am profoundly thankful.

The completion of this thesis owes much to the invaluable guidance, support, and motivation provided by my supervisor, Professor Vissarion Papadopoulos. I am deeply grateful for his patience, unwavering faith in my abilities, and his generosity in sharing his time and wisdom. Above all, I am indebted to him for the exemplary leadership that he has demonstrated over the years.

Beyond my supervisor, I extend my gratitude to the members of my thesis committee for their insightful feedback and encouragement. I am particularly grateful to Emeritus Professor Konstantinos Spiliopoulos and Professor Nikolaos Lagaros for their contributions to my supervision committee and their readiness to provide guidance whenever needed.

My sincere thanks also go to my colleague, Dr. Ioannis Kalogeris, whose encouragement and mentorship were instrumental during the journey of pursuing my doctoral degree. Our conversations were always a wellspring of inspiration for me. I am also appreciative of my esteemed colleagues of the MGroup research group for all the enriching discussions and the fruitful collaborations, that provided me with valuable knowledge in a wide spectrum of computational mechanics topics and expanded my way of thinking.

AI-enhanced multiscale finite element methods for forward and inverse uncertainty quantification problems in structural mechanics

ABSTRACT

Over recent decades, there has been growing interest in high-performance materials tailored for complex engineering applications. By modifying material structures at fine scales, exceptional properties such as enhanced mechanical strength, improved thermal conductivity, and novel optical features can be achieved. To address the time-consuming and costly experimentation on these materials, several computational techniques have been developed. Among them, the multiscale computational homogenization method via the well-established FE^2 algorithm has gained significant attention. Despite its computational intensity, this algorithm is favored for its ability to reliably predict the complex macroscopic behavior of multiscale material systems due to non-linear phenomena at finer scales. However, identifying the parameters that characterize material behavior at fine scales still remains a nontrivial undertaking. This thesis presents a cost-efficient framework using machine learning strategies for implementing the computational homogenization modeling approach on multi-query analyses investgating fine-scale parameters. Novel computational methodologies are proposed for accurate and efficient forward and inverse uncertainty quantification analyses on multiscale material systems and are validated through real-world case studies.

First, this thesis presents a strategy for performing Bayesian inference on microscale material properties using experimental observations from the visible structure. To tackle the computational load of repeated FE^2 analyses, a feed-forward neural network (FFNN) is used to emulate material behavior affected by microstructural parameters. This is achieved by training the FFNN on a dataset from offline representative volume element (RVE) solutions. Next, the thesis generalizes the FE^2 algorithm by employing a sequence of FFNNs to represent different scales in the multiscale system, with each FFNN learning the constitutive law of its corresponding length scale. This results in a FFNN that emulates macroscopic behavior by incorporating mechanisms from each finer scale. Based on this scheme, the thesis, subsequently, proposes a methodology to identify optimal typologies of nanocomposite materials for desirable structural responses under uncertain conditions. Finally, a hierarchical Bayesian framework is introduced to utilize disjoint experimental measurements in multiscale material systems for joint parameter inference. This framework integrates experimental data from different scales and material compositions to yield informed parameters for future model predictions.

ΠΕΡΙΛΗΨΗ ΤΗΣ
ΔΙΔΑΚΤΟΡΙΚΗΣ ΔΙΑΤΡΙΒΗΣ

με τίτλο

'Μέθοδοι πεπερασμένων στοιχείων σε πολλαπλές κλίμακες ενισχυμένες με AI για ευθέα και αντίστροφα προβλήματα ποσοτικοποίησης αβεβαιότητας στη δομική μηχανική'

Ένα μεγάλο πλήθος υλικών που χρησιμοποιούνται στην επιστήμη του μηχανικού αλλά και γενικότερα στην καθημερινή ζωή είναι σύνθετα. Ως σύνθετα χαρακτηρίζονται τα υλικά που συντίθενται από δύο ή περισσότερες φάσεις συστατικών υλικών τα οποία λόγω της ισχυρής σύνδεσης μεταξύ τους σχηματίζουν ένα ενιαίο σώμα. Σύνθετα υλικά συναντώνται είτε αυτούσια στη φύση όπως το ξύλο και τα οστά είτε κατασκευάζονται από τον άνθρωπο όπως το σκυρόδεμα. Οι συνεχώς αυξανόμενες απαιτήσεις για υλικά υψηλών αποδόσεων με εξειδικευμένες μηχανικές, θερμικές, ηλεκτρικές κ.α. ιδιότητες έχει αποτελέσει ισχυρό κίνητρο για την εξερεύνηση και δημιουργία νέων σύνθετων υλικών τόσο στην έρευνα αλλά και στην βιομηχανία. Ως απόρροια αυτής της στοχευμένης προσπάθειας, ένα μεγάλο πλήθος υποσχόμενων υλικών έχει αναδειχθεί τις τελευταίες δεκαετίες, συμπεριλαμβανομένων των ενισχυμένων υλικών με παράγωγα γραφίτη όπως νανοσωλήνες άνθρακα, ανθρακόνημα και γραφένιο.

Ο χαρακτηρισμός των ιδιοτήτων των σύνθετων υλικών είναι συνήθως μια επίπονη διαδικασία. Αυτό αποδίδεται στο γεγονός του ότι θα πρέπει να ληφθούν υπόψη τόσο τα επιμέρους χαρακτηριστικά και οι ιδιότητες των συστατικών υλικών των σύνθετων καθώς και το αποτέλεσμα της αλληλεπίδρασης τους. Λόγω αυτού, ο σχηματισμός του μηχανικού αλλά και γενικότερα του μαθηματικού προβλήματος είναι ένα δύσκολο έργο το οποίο μελετάται ήδη περισσότερο από έναν αιώνα. Οι πρώτες προσπάθειες για την εύρεση της ενιαίας μηχανικής μακροσκοπικής συμπεριφοράς των συνθέτων υλικών έγιναν τον 18ο αιώνα με τις κομβικές εργασίες των Voigt [132] και Reuss [109]. Μέσα από τις εργασίες αυτές συντίθεται ο κανόνας των μιγμάτων (rule of mixtures), όπου είτε κάνοντας την υπόθεση του σταθερού πεδίου παραμόρφωσης στην πρώτη περίπτωση, είτε του σταθερού πεδίου τάσης στη δεύτερη περίπτωση, μπορούν να υπολογιστούν ενιαίες υλικές ιδιότητες. Οι παραδοχές της σταθερής παραμόρφωσης και αντίστοιχα της σταθερής τάσης, παρ΄ όλα αυτά, δεν ανταποκρίνονται σε ρεαλιστικά σενάρια και συνεπώς τα μοντέλα αυτά δε μπορούν να κάνουν καλή πρόβλεψη των ομογενοποιημένων χαρακτηριστικών σε περίπλοκα σύνθετα συστήματα. Στην προσπάθεια βελτίωσης των πρώτων αυτών αποπειρών, προτάθηκαν διάφορες μέθοδοι οι οποίες βασίστηκαν στην λογική του ελλειψοειδούς εγκλείσματος Eshelby [32]. Με βάση τις τεχνικές αυτές μπορούν να προκύψουν προσεγγιστικές λύσεις του ενιαίου προβλήματος, μέσω της θεώρησης ότι υπάρχει πλήρης απουσία ή ασθενής αλληλεπίδραση μεταξύ των εγκλεισμάτων. Βελτιώσεις αυτής της τεχνικής, με στόχο την αντιμετώπιση των ανακριβειών που δημιουργεί η παραδοχής της μη αλληλεπίδρασης, προτάθηκαν από τον Hill [56] και Budiansky [14] με την εισαγωγή των "Self Consistent" μεθόδων, αλλά και της "Generalized Self-Consistent" μεθόδου [20] από τον Christensen.

Όλες οι προαναφερθείσες μέθοδοι βασίζονται σε κάποιες προσεγγίσεις και σε φαινομενολογικές θεωρήσεις για το ετερογενές πρόβλημα. Παρόλο που μέσω τέτοιων παραδοχών επιτυγχάνεται η επίλυση του προβλήματος με μικρό υπολογιστικό κόστος μέσω αναλυτικών ή ημι-αναλυτικών εκφράσεων, οι προβλέψεις τους σε πολλές περιπτώσεις απέχουν αξιοσημείωτα από την πραγματικότητα. Μέσα από την ραγδαία αύξηση των υπολογιστικών δυνατοτήτων, έχει κατοχυρωθεί τα τελευταία χρόνια πως ο πιο αξιόπιστος τρόπος για την εύρεση των ομογενοποίημενων χαρακτηριστικών οποιουδήποτε σύνθετου υλικού συστήμα-

τος είναι μέσω της επίλυσης ενός προβλήματος οριακής τιμής (boundary value problem) στην μικροκλίμακα. Αυτή η μεθοδολογία έχει καθιερωθεί στην βιβλιογραφία ως υπολογιστική ομογενοποίηση (computational homogenization). Οι πρώτες απόπειρες για την επίλυση του μικροσκοπικού γραμμικού αριθμητικού προβλήματος με την μέθοδο των πεπερασμένων στοιχείων έγιναν το 1987 από τον Suquet [123]. Η επέκταση της υπολογιστικής ομογενοποίησης σε μη γραμμικά συστήματα πολλαπλών κλιμάκων έχει γίνει πιο πρόσφατα συμβαδίζοντας με την πολύ ισχυρή σύγχρονη υπολογιστική ισχύ. Οι μέθοδοι πολλαπλών κλιμάκων (multiscale methods) ή πολλαπλών επιπέδων (multilevel methods) όπως έχουν καθιερωθεί, έχουν την δυνατότητα να επιλύουν πολύπλοκα μη-γραμμικά προβλήματα που περιγράφουν εκλεπτυσμένα σύνθετα υλικά. Συγκεκριμένα, οι μέθοδοι μη γραμμικής υπολογιστικής ομογενοποίησης χρησιμοποιούν ένα εμφωλευμένο σχήμα σύμφωνα με το οποίο δημιουργείται μια σταθερή επικοινωνία μεταξύ των διαφορετικών κλιμάκων κατά τη διάρκεια της επίλυσης. Μια από τα πιο αξιοσημείωτες μεθόδους που εφαρμόζουν τη λογική της εμφώλευσης είναι ο αλγόριθμος FE^2, όπως έχει κατοχυρωθεί από τον Feyel [35], ο οποίος έχει ως βάση την ταυτόχρονη ανάλυση πεπερασμένων στοιχείων και στις δύο κλίμακες του συζευγμένου συστήματος. Η μεγάλη ακρίβεια που προσφέρουν αυτές οι μέθοδοι έχουν οδηγήσει στην επέκταση της χρήσης τους πέρα από μηχανικά προβλήματα και στην διερεύνηση άλλων φυσικών φαινομένων (π.χ. θερμότητας και ηλεκτρισμού) καθώς και πολυ-φυσικών προβλημάτων (π.χ. πιεζοηλεκτρικότητας). Οι εμφωλευμένες προσεγγίσεις θεωρούνται ως οι πιο ακριβείς για την πρόβλεψη της εξέλιξης ενός σύνθετου υλικού, ωστόσο, ένα σημαντικό μειονέκτημα που τους αποδίδεται είναι οι τεράστιες υπολογιστικές απαιτήσεις τους, οι οποίες σε πολλές περιπτώσεις μπορεί να είναι και απαγορευτικές. Με βάση τα παραπάνω, ένας βασικός προσανατολισμός της παρούσας διατριβής είναι η αξιοποίηση της ισχυρής προγνωστικής ικανότητας των εμφωλευμένων σχημάτων στις αναλύσεις πολλαπλών κλιμάκων. Στη συνέχεια περιγράφεται συνοπτικά το μηχανικό πρόβλημα πολλαπλών κλιμάκων:

Η ισχυρή διατύπωση της εξίσωσης στατικής ισορροπίας ενός ελαστικού σώματος M το οποίο τοποθετείται στη μακροσκοπική δομική κλίμακα και ορίζεται σε έναν πεδίο $\Omega^M \subset R^3$ έχει τη μορφή:

$$\nabla^M \cdot \boldsymbol{\sigma}^M = \boldsymbol{b}^M \quad \sigma\tau o \quad \Omega^M \tag{1}$$

Η καταστατική σχέση για το μακροσκοπικό πρόβλημα εκφράζεται μέσω της γενικής μορφής:

$$\boldsymbol{\sigma}^M(t) = \boldsymbol{\sigma}^M \{\boldsymbol{\epsilon}^M(t), \boldsymbol{\theta}^M(\bar{t}), \bar{t} \in [0,t]\} \tag{2}$$

όπου το t δηλώνει τον ψευδοχρόνο στην περίπτωση μιας ψευδοστατικής ανάλυσης και το $\bar{t}$ την ιστορική εξάρτηση της συμπεριφοράς του υλικού. Επιπλέον, το $\boldsymbol{\theta}^M$ χρησιμοποιείται για να δηλώσει ένα σύνολο εσωτερικών μεταβλητών που χαρακτηρίζουν την κατάσταση του υλικού (π.χ. πλαστικότητα, θραύση) και παραμέτρους του υλικού μοντέλου. Στο πλαίσιο της ανάλυσης πολλαπλών κλιμάκων όπως ορίζεται από την υπολογιστική ομογενοποίηση, η εξ. (2) δεν έχει ρητή συναρτησιακή μορφή. Αντίθετα, αυτή η σχέση προκύπτει μέσω

της επίλυσης ενός προβλήματος οριακής τιμής που επιβάλλεται από ένα αντιπροσωπευτικό στοιχείο όγκου (Representative Volume Element - RVE) , το οποίο περιγράφει τη δομή της λεπτής κλίμακας, όπως παρουσιάζεται στο σχήμα 1.

Η λεπτή κλίμακα μπορεί να αντιπροσωπεύει διαφορετικές κλίμακες μήκους, ωστόσο η διαμόρφωση της γενικεύεται συνήθως ως πρόβλημα μικροκλίμακας. Η ισχυρή διατύπωση της εξίσωσης στατικής ισορροπίας απουσία ογκομετρικών δυνάμεων εκφράζεται ως:

$$\nabla^\mu \boldsymbol{\sigma}^\mu = 0 \quad \sigma\tau o \quad \Omega^\mu \tag{3}$$

Η καταστατική σχέση για το πρόβλημα της μικροκλίμακας γράφεται ως:

$$\boldsymbol{\sigma}^\mu(t) = \boldsymbol{\sigma}^\mu \{\boldsymbol{\epsilon}^\mu(t), \boldsymbol{\theta}^\mu(\bar{t}), \bar{t} \in [0, t]\} \tag{4}$$

όπου το $\boldsymbol{\theta}^\mu$ συμπεριλαμβάνει υλικές παραμέτρους της μικροκλίμακας.
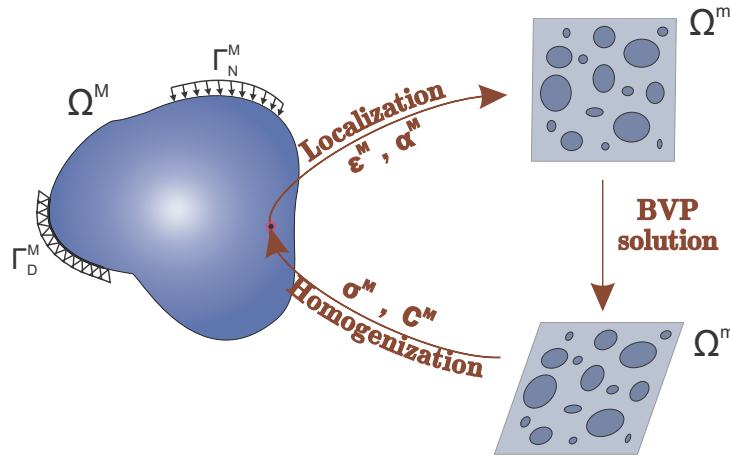
Σχήμα 1: Μετάβαση μεταξύ των κλιμάκων όπως ορίζεται από την υπολογιστική ομογενοποίηση

Σύμφωνα με την αρχή του διαχωρισμού των κλιμάκων, η θεωρία της ομογενοποίησης μπορεί να εφαρμοστεί αποτελεσματικά εάν ισχύει η ακόλουθη σχέση:

$$\lambda^\mu \ll \lambda^{RVE} \ll \lambda^M \tag{5}$$

όπου λ^μ, λ^{RVE}, λ^M είναι οι κλίμακες μήκους των διακυμάνσεων του μικροσκοπικού πεδίου, του μεγέθους RVE και των διακυμάνσεων του μακροσκοπικού πεδίου, αντίστοιχα.

Για την σύνδεση του συστήματος που ορίζεται από τις δύο κλίμακες, στη μικροκλίμακα πρέπει να δοθούν οριακές συνθήκες που είναι συνεπείς με την απόκριση της μακροδομής. Για το σκοπό αυτό, το πεδίο μετατόπισης $\boldsymbol{u}^\mu(\boldsymbol{x})$ συσχετίζεται με την παραμόρφωση $\boldsymbol{\epsilon}^M$ ενός σημείου μακροσκοπικού υλικού μέσω της σχέσης:

$$\boldsymbol{u}^\mu = \boldsymbol{\epsilon}^M \cdot \boldsymbol{x}^\mu + \tilde{\boldsymbol{u}}^\mu \qquad (6)$$

όπου $\tilde{\boldsymbol{u}}^\mu$ είναι το πεδίο διακύμανσης μετατόπισης που αποδίδεται στην ετερογένεια της μικροδομής. Μία επιλογή που πληροί αυτήν την απαίτηση και χρησιμοποιήθηκε στην παρούσα διατριβή είναι ο περιορισμός του $\tilde{\boldsymbol{u}}^\mu$ στο όριο του RVE, δίνοντας στην εξ. (6) την τελική μορφή:

$$\boldsymbol{u}^\mu = \boldsymbol{\epsilon}^M \boldsymbol{x}^\mu \ \ \sigma\tau\alpha \ \ \boldsymbol{x}^\mu \in \Gamma^\mu \qquad (7)$$

Στην υπολογιστική ομογενοποίηση, η τοπική μεταβολή του μακροσκοπικού έργου συνδέεται με τη μεταβολή του μικροσκοπικού έργου ανά μονάδα όγκου μέσω της σχέσης ενεργειακής συνέπειας Hill-Mandell.

$$\boldsymbol{\sigma}^M : \delta\boldsymbol{\epsilon}^M = \frac{1}{\|\Omega^\mu\|} \int_{\Omega^\mu} \boldsymbol{\sigma}^\mu : \delta\boldsymbol{\epsilon}^\mu d\Omega^\mu \qquad (8)$$

Με βάση την εξ. (8) και μετά από την εκτέλεση κάποιων αλγεβρικών πράξεων προκύπτει η σχέση που ορίζει τον κανόνα ομογενοποίησης, δηλαδή την μετάβαση από την μικροκλίμακα στην μακροκλίμακα, ως:

$$\boldsymbol{\sigma}^M = \frac{1}{\|\Omega^\mu\|} \int_{\Gamma^\mu} \hat{\boldsymbol{\tau}}^\mu \otimes \boldsymbol{x}^\mu d\Gamma^\mu = \frac{1}{\|\Omega^\mu\|} \int_{\Omega^\mu} \boldsymbol{\sigma}^\mu d\Omega^\mu \qquad (9)$$

όπου το πεδίο τάσεων $\boldsymbol{\sigma}^\mu$ μπορεί να αποκτηθεί μετά την επίλυση του προβλήματος της οριακής τιμής (BVP) που επιβάλλεται στο RVE σύμφωνα με τον κανόνα τοπικοποίησης (localization), που περιγράφεται από την εξ. (7) .

Επιπλέον, το εφαπτομενικό μητρώο που ορίζει τον καταστατικό νόμο στην εξ. (2) για το εκάστοτε μακροσκοπικό σημείο υλικού ορίζεται ως:

$$\boldsymbol{C}^M = \frac{1}{\|\Omega^\mu\|} \partial_{\boldsymbol{\epsilon}^M} \int_{\Omega^\mu} \boldsymbol{\sigma}^\mu d\Omega^\mu \qquad (10)$$

Με βάση τη σύνδεση μεταξύ της μικροκλίμακας και της μακροκλίμακας όπως ορίζεται από την υπολογιστική ομογενοποίηση, ο αλγόριθμος FE^2 [35] χρησιμοποιείται για τη συνεχή αλληλεπίδραση μεταξύ των δύο κλιμάκων κατά την επίλυση ενός μακροσκοπικού σύνθετου συστήματος. Σύμφωνα με αυτόν τον αλγόριθμο, για κάθε αύξηση του φορτίου της μακροδομής, η μακροσκοπική τάση $\boldsymbol{\epsilon}^M$ υπολογίζεται σε κάθε σημείο ολοκλήρωσης όλων των πεπερασμένων στοιχείων, όπου ορίζεται ένα μοναδικό RVE. Στην συνέχεια, εφαρμόζεται το σχήμα ομογενοποίησης που περιγράφηκε προηγουμένως, όπου επιλύεται το RVE και αποκτάται η μακροσκοπική τάση $\boldsymbol{\sigma}^M$ και το εφαπτομενικό μητρώο $\boldsymbol{C}^M$ στο εκάστοτε σημείο ολοκλήρωσης. Μέσω αυτών τον ποσοτήτων, μπορεί να υπολογιστεί το μακροσκοπικό διάνυσμα εσωτερικής δύναμης $\boldsymbol{F}_{int}^M(\boldsymbol{\sigma}^M)$ και ο πίνακας εφαπτομενικής ακαμψίας $\boldsymbol{K}^M(\boldsymbol{C}^M)$. Αυτή η διαδικασία επαναλαμβάνεται έως ότου η εσωτερική δύναμη γίνει ισοδύναμη με την εξωτερική δύναμη $\boldsymbol{F}_{ex}^M$ σε όλο τον μακροσκοπικό φορέα.

Οι μέθοδοι μη γραμμικής υπολογιστικής ομογενοποίησης παρότι είναι αδιαμφισβήτητα οι πλέον ιδανικές για την πρόβλεψη οποιασδήποτε υλικής συμπεριφοράς, δεν έχουν υιοθετηθεί ακόμη πλήρως από την επιστημονική κοινότητα. Αυτό οφείλεται στο γεγονός ότι το υπολογιστικό κόστος που απαιτεί μια ανάλυση κάνοντας χρήση τους είναι υπέρογκο, στην πλειονότητα των περιπτώσεων. Στην προσπάθεια εύρεσης απαντήσεων σε αυτήν την τροχοπέδη, έχουν γίνει εκτεταμένες διερευνήσεις σχετικά με την χρήση υποκατάστατης μοντελοποίησης (surrogate modeling) για την πρόβλεψη της μακροσκοπικής υλικής συμπεριφοράς. Η υιοθέτηση υποκατάστατων μοντέλων σε αυτήν την περιοχή της έρευνας, έχει ως στόχο την σημαντική μείωση ή ακόμη και αντικατάσταση της πολύ δαπανηρής επίλυσης της εξίσωσης της ομογενοποίησης όπως αυτή ορίζεται από την υπολογιστική ομογενοποίηση. Διαφορετικές μέθοδοι έχουν χρησιμοποιηθεί στην βιβλιογραφία προς επίτευξη αυτού του στόχου, όπως τεχνικές μείωσης τάξης μοντέλου (model order reduction) π.χ.proper orthogonal decomposition και autoencoders αλλά και εξομοιωτές εξισώσεων πχ. feed forward neural networks και Gaussian process regression. Στην παρούσα διατριβή γίνεται αξιοποίηση τεχνικών μηχανικής μάθησης για τη δημιουργία υποκατάστατης μοντελοποίησης με στόχο την αποδοτική επίλυση των εξαιρετικά απαιτητικών προβλημάτων που μελετώνται.

Εκτός από τον υπολογιστικό σχεδιασμό νανοσύνθετων υλικών, αυτά τα υλικά συχνά έχουν πολλές αβεβαιότητες που εντοπίζονται στη μικροδομή τους, όπως η μορφολογία, οι μηχανικές ιδιότητες των επιμέρους υλικών καθώς και οι ιδιότητες αλληλεπίδρασής τους. Ο πειραματικός μικροδομικός χαρακτηρισμός των υλικών αυτών είναι ένα πολύ δύσκολο και κοστοβόρο έργο καθώς απαιτείται εκλεπτυσμένος εργαστηριακός εξοπλισμός. Από την άλλη, υπάρχουν εξειδικευμένες υπολογιστικές τεχνικές για την επίλυση αντίστροφων προβλημάτων, δηλαδή την παραμετρική ταυτοποίηση μοντέλων μέσω πραγματικών μετρήσεων. Μια από τις πιο διαδεδομένες μεθόδους είναι η Μπεϋζιανή επικαιροποίηση (Bayesian Update) σύμφωνα με την οποία οι παράμετροι του μοντέλου μοντελοποιούνται ως τυχαίες μεταβλητές και μέσω της εφαρμογής του κανόνα Bayes οι εκ των υστέρων κατανομές τους μπορούν να προσδιοριστούν με βάση τα διαθέσιμα δεδομένα. Πέρα από τον χαρακτηρισμό τέτοιων υλικών, μια εξίσου σημαντική πρόκληση είναι και η εύρεση βέλτιστων τυπολογιών, ως προς την διάταξη των εγκλεισμάτων, που να μπορούν να προσδώσουν σε μία κατασκευή στοχευμένες μηχανικές ιδιότητες, ιδιαίτερα υπό το πρίσμα αβεβαιοτήτων π.χ. στο υλικό ή στην εξωτερική φόρτιση. Μια περαιτέρω κατεύθυνση της παρούσας διατριβής είναι η ανάπτυξη υπολογιστικών μεθοδολογιών που θα μπορούν να απαντήσουν στα ανωτέρω ερωτήματα με μεγάλη ακρίβεια αλλά και με μικρό υπολογιστικό φόρτο. Συγκεκριμένα, μέσα από την διερεύνηση των μηχανικών ιδιοτήτων ενός συγκεκριμένου τύπου νανοσύνθετου υλικού δηλαδή τα υλικά ενισχυμένα με νανοσωλήνες άνθρακα, φιλοδοξία της διατριβής είναι η ανάπτυξη γενικευμένων αριθμητικών σχημάτων που θα μπορούν να εφαρμοστούν σε αναλύσεις ποσοτικοποίησης αβεβαιότητας σε οποιαδήποτε υλικό, αλλά και στην μελέτη διαφορετικών φυσικών φαινομένων.

Στο πρώτο στάδιο της διατριβής προτείνεται μια νέα μέθοδος βασισμένη στην Μπεϋζιανή επικαιροποίηση για παραμετρική διερεύνηση νανοσύνθετων υλικών. Συγκεκριμένα το ενδιαφέρον εστιάζεται στην εκμάθηση των παραμέτρων που χαρακτηρίζουν τη διεπαφή

CNT/πολυμερούς στη μικροκλίμακα με δεδομένα διαθέσιμα από τις ανώτερες κλίμακες. Αυτές οι παράμετροι συνδέονται με μεγάλες αβεβαιότητες και ο χαρακτηρισμός τους είναι δύσκολο έργο, καθώς οι μετρήσεις στη μικροκλίμακα είναι δαπανηρές και δύσκολο να ληφθούν. Για να ξεπεραστεί αυτό, η παρούσα μελέτη εισάγει ένα υπολογιστικό πλαίσιο για την ενημέρωση των προηγούμενων πεποιθήσεων σχετικά με τις τιμές αυτών των παραμέτρων, χρησιμοποιώντας μετρήσεις μετατοπίσεων που λαμβάνουν χώρα στην μέση και/ή στην μακροκλίμακα που χαρακτηρίζουν το εξεταζόμενο σύνθετο υλικό. Όσον αφορά τη μοντελοποίηση, η διεπαφή CNT/πολυμερούς διαμορφώνεται χρησιμοποιώντας ένα μοντέλο συνεκτικής ζώνης και έναν διγραμμικό καταστατικό νόμο δεσμού-ολίσθησης. Στη συνέχεια, η μέθοδος FE^2 χρησιμοποιείται για την πρόβλεψη της απόκρισης των κατασκευών. Όπως έχει αναφερθεί, παρά την ακρίβειά της, αυτή η μέθοδος συνδέεται με τεράστιες υπολογιστικές απαιτήσεις όταν εφαρμόζεται σε μεγάλα προβλήματα. Επομένως, η εφαρμογή της στα πλαίσια της Μπεϋζιανής επικαιροποίησης η οποία απαιτεί πολλαπλές αξιολογήσεις μοντέλων είναι απαγορευτική. Για να μετριαστεί αυτό το τεράστιο κόστος, αναπτύχθηκε μια τεχνική υποκατάστατης μοντελοποίησης που χρησιμοποιεί τεχνητά νευρωνικά δίκτυα, τα οποία εκπαιδεύονται στο να προβλέπουν τη μη γραμμική σχέση τάσης-παραμόρφωσης όπως προκύπτει από την ανάλυση στοιχείων αντιπροσωπευτικού όγκου της μικροδομής. Το σύνολο δεδομένων στο οποίο εκπαιδεύεται το νευρωνικό δίκτυο, λαμβάνεται με την ανάλυση ενός περιορισμένου αριθμού διαφορετικών διαμορφώσεων RVE χρησιμοποιώντας λεπτομερείς αναλύσεις πεπερασμένων στοιχείων.

Συγκεκριμένα, οι τρεις παράμετροι που περιγράφουν τη διεπιφανειακή σχέση μεταξύ των CNT και της μήτρας πολυμερούς επιλέγονται ως τυχαίες παράμετροι προς ενημέρωση. Αυτές είναι η διεπιφανειακή διατμητική αντοχή $\tau_{1,s}$, η ελαστική κλίση D_{el} πριν από την ολίσθηση και η πλαστική κλίση D_{pl} εφόσον έχει ξεπεραστεί η $\tau_{1,s}$, δηλαδή, $\boldsymbol{\theta} = (\tau_{1,s}, D_{el}, D_{pl})$. Ο νόμος διεπαφής απεικονίζεται στο σχήμα 2. Επιπλέον, τα μηχανικά μοντέλα που διερευνώνται είναι σύνθετες κατασκευές ενισχυμένες με CNT και οι προβλέψεις μοντέλων $\mathcal{M}(\boldsymbol{\theta})$ για δεδομένες πραγματοποιήσεις του $\boldsymbol{\theta}$ λαμβάνονται με την επίλυση του προβλήματος πολλαπλών κλιμάκων με FEM. Επίσης, τα δεδομένα $\boldsymbol{D}$ που συλλέγονται περιλαμβάνουν μετρήσεις μετατοπίσεων σε καθορισμένες θέσεις της μακροδομής.

Σύμφωνα με την Μπεϋζιανή επικαιροποίηση το αντίστροφο πρόβλημα διατυπώνεται μέσω ενός πιθανοτικού πλαισίου. Ακολουθώντας την πιο κοινή προσέγγιση του προσθετικού σφάλματος, η σχέση μεταξύ των μετρούμενων δεδομένων και των προβλέψεων του μοντέλου ποσοτικοποιείται μέσω της έκφρασης:

$$\boldsymbol{y}^i(\boldsymbol{\theta}) = \boldsymbol{m}^i(\boldsymbol{\theta}) + \boldsymbol{\epsilon}^i \qquad i = 1, ..., K \tag{11}$$

όπου κάθε $\boldsymbol{\epsilon}^i$ υποδηλώνει τον όρο σφάλματος που εκφράζεται συνήθως ως τυχαία μεταβλητή με μια κανονική συνάρτηση πυκνότητας πιθανότητας $\sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}^i)$. Ο πίνακας διακύμανσης $\boldsymbol{\Sigma}^i$ ενσωματώνει το μέγεθος του σφάλματος πρόβλεψης του μοντέλου M^i και τα σφάλματα των πειραματικών παρατηρήσεων $\boldsymbol{D}^i$. Ο όρος $\boldsymbol{y}^i$ αντιπροσωπεύει την έξοδο του στοχαστικού ευθέου μοντέλου Y^i από όπου παράγονται τα δεδομένα $\boldsymbol{D}^i$.
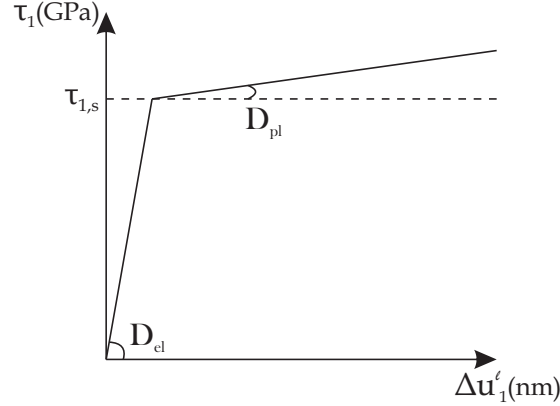
**Σχήμα 2:** Διγραμμικός καταστατικός νόμος της ολίσθησης μεταξύ CNT και μήτρας

Συγκεκριμένα το πρόβλημα της Μπεϋζιανής επικαιροποίησης ορίζεται ως η αναζήτηση μιας εκ των υστέρων κατανομής (posterior distribution) $P(\boldsymbol{\theta}|\boldsymbol{D})$ μέσω της σχέσης:

$$P(\boldsymbol{\theta}|\boldsymbol{D}) = \frac{P(\boldsymbol{D}|\boldsymbol{\theta})P(\boldsymbol{\theta})}{P(\boldsymbol{D})} = \frac{\pi(\boldsymbol{D}|\boldsymbol{\theta})P(\boldsymbol{\theta})}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty}\pi(\boldsymbol{D}|\boldsymbol{\theta})P(\boldsymbol{\theta})d\theta_1 d\theta_2...d\theta_n} \tag{12}$$

όπου το $P(\boldsymbol{\theta})$ είναι η εκ των προτέρων κατανομή (prior distribution), δηλαδή οι προηγούμενες πεποιθήσεις σχετικά με την πιθανολογική φύση των παραμέτρων. Στην περίπτωση που δεν μπορούν να γίνουν εκ των προτέρων υποθέσεις, συνήθως εφαρμόζονται μη ενημερωτικές κατανομές, όπως η ομοιόμορφη κατανομή. Ο όρος $P(\boldsymbol{D})$ ονομάζεται απόδειξη (evidence) και είναι ουσιαστικά ένα πολυδιάστατο ολοκλήρωμα, αφού περιλαμβάνει τον υπολογισμό της περιθωριοποιημένης εκ των υστέρων κατανομής ως προς τον παραμετρικό χώρο. Ο όρος $\pi(\boldsymbol{D}|\boldsymbol{\theta})$ υποδηλώνει την πιθανοφάνεια (likelihood) του $\boldsymbol{D}$ για δεδομένες τιμές $\boldsymbol{\theta}$ και ποσοτικοποιεί την ομοιότητα μεταξύ των δεδομένων και των προβλέψεων του μοντέλου με στοχαστικό τρόπο. Στην εξ. (12), ο αναλυτικός υπολογισμός του n-διάστατου ολοκληρώματος στον παρονομαστή δεν είναι εφικτός στη γενική περίπτωση.

Προκειμένου να αντληθούν δείγματα αποτελεσματικά από την εκ των υστέρων κατανομή, χρησιμοποιείται η τεχνική Markov Chain Monte Carlo (MCMC), σύμφωνα με την οποία παρακάμπτεται ο υπολογισμός του παρονομαστή της εξ. (12) και γίνεται δειγματοληψία μόνο από τον αριθμητή. Συγκεκριμένα, ο αλγόριθμος Metropolis-Hastings (MH) επιλέχθηκε σε αυτή τη μελέτη, ωστόσο, μπορεί στη θέση αυτού να εφαρμοστεί οποιοσδήποτε άλλος MCMC αλγόριθμος. Η σχηματική αναπαράσταση των αλγοριθμικών βημάτων του προτεινόμενου σχήματος Μπεϋζιανής επικαιροποίησης σε συστήματα πολλαπλών κλιμάκων απεικονίζεται στο σχήμα 3.

Ο υπολογισμός της συνάρτησης πιθανοφάνειας $\pi(\boldsymbol{D}|\boldsymbol{\theta}')$ απαιτεί την επίλυση του μοντέλου $\mathcal{M}$ για τις προτεινόμενες τιμές παραμέτρων $\boldsymbol{\theta}'$ σε κάθε επανάληψη του MH αλγορίθμου. Στα προβλήματα που μελετήθηκαν σε αυτή την εργασία, οι επιλύσεις μοντέλων
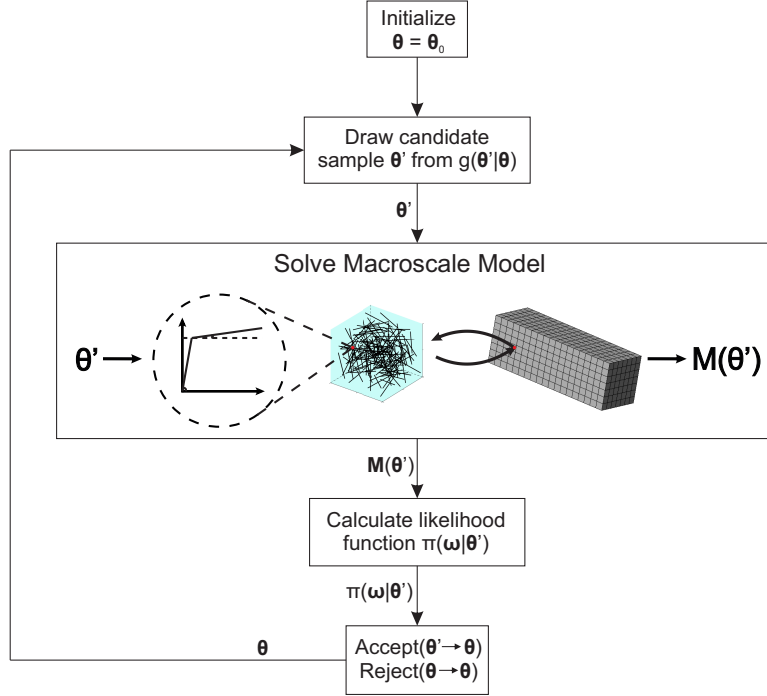
Σχήμα 3: Μπεϋζιανή επικαιροποίηση σε συστήματα πολλαπλών κλιμάκων χρησιμοποιώντας τον αλγόριθμο Metropolis-Hastings

είναι ιδιαίτερα χρονοβόρες και η άμεση εφαρμογή του αλγορίθμου MH θα ήταν ανέφικτη. Για την αντιμετώπιση αυτού του ζητήματος, προτείνεται να δημιουργηθεί ένα υποκατάστατο μοντέλο που θα αναπαράγει τη σχέση τάσης-παραμόρφωσης όπως ορίζεται από το RVE με χαμηλό κόστος. Με αυτό το τρόπο επιτυγχάνεται δραστική επιτάχυνση των επιλύσεων του μοντέλου για τις εκάστοτε τιμές παραμέτρων $\boldsymbol{\theta}'$.

Συγκεκριμένα, αναπτύσσεται ένα FFNN του οποίου ο στόχος είναι να μάθει τη μη γραμμική εξίσωση στο σχήμα ομογενοποίησης του RVE, για διαφορετικές παραμέτρους της διεπαφής CNT/μήτρας. Οι νευρώνες εισόδου του αποτελούνται από το μακροσκοπικό διάνυσμα παραμόρφωσης $\overline{\boldsymbol{\epsilon}}$ μαζί με τις τρεις παραμέτρους διεπαφής $\boldsymbol{\theta} = (\tau_{1,s}, D_{el}, D_{pl})$, ενώ οι νευρώνες εξόδου είναι οι τιμές του διανύσματος μακροσκοπικής τάσης $\overline{\boldsymbol{\sigma}}$. Λόγω της συμμετρίας των τανυστών $\overline{\boldsymbol{\epsilon}}$ και $\overline{\boldsymbol{\sigma}}$, ο κάθε ένας χαρακτηρίζεται από 3 άγνωστες μεταβλητές για 2D προβλήματα και 6 μεταβλητές για 3D προβλήματα.

Προκειμένου να συλλεχθούν τα απαραίτητα δείγματα για την εκπαίδευση του FFNN, επιλύεται ένα πλήθος από RVE για διάφορες τιμές εισόδου $\overline{\boldsymbol{\epsilon}}$ και $\boldsymbol{\theta}$. Αυτές οι τιμές επιλέγονται με ψευδοτυχαίο τρόπο, χρησιμοποιώντας την τεχνική Latin Hypercube Sampling (LHS). Τα δειγματικά εύρη επιλέγονται προσεκτικά, καθώς μικρά εύρη μπορεί να οδηγήσουν σε ανεπαρκείς πληροφορίες κατά την επίλυση του Μπεϋζιανού συστήματος πολλαπλών κλιμάκων,

ενώ μεγάλα εύρη απαιτούν την επίλυση ενός παράλογα μεγάλο αριθμού RVE για την ακριβή εκπαίδευση του FFNN. Δεδομένου ότι η υπολογιστική ομογενοποίηση απαιτεί χρήση αυξητικού-επαναληπτικού επιλύτη για μη γραμμικές εξισώσεις (π.χ. Newton-Raphson), για κάθε $\boldsymbol{\theta}$, οι συγκλίνουσες τιμές του $\overline{\boldsymbol{\sigma}}$ για όλες ενδιάμεσες προσαυξήσεις του $\overline{\boldsymbol{\epsilon}}$, συμπεριλαμβάνονται στα δεδομένα εκπαίδευσης.

Επιπλέον, το μακροσκοπικό εφαπτομενικό καταστατικό μητρώο $\overline{\boldsymbol{C}}$ μπορεί να εξαχθεί από το υποκατάστατο μοντέλο. Αυτό μπορεί να γίνει άμεσα μετά την εκπαίδευση του FFNN χρησιμοποιώντας την μέθοδο Αυτόματης Διαφοροποίησης (Automatic Differentiation - AD). Η AD επιτρέπει υπολογισμούς παραγώγων των στοιχείων εξόδου $\overline{\boldsymbol{\sigma}}$ σε σχέση με τα στοιχεία εισόδου $\overline{\boldsymbol{\epsilon}}$ εφαρμόζοντας τον κανόνα της αλυσίδας. Έτσι, οι συντελεστές $\bar{c}_{ij}$ του μακροσκοπικού εφαπτομενικού πίνακα $\overline{\boldsymbol{C}} = [\bar{c}_{ij}]$ μπορούν να υπολογιστούν ως:

$$\bar{c}_{ij} = \frac{\partial \overline{\sigma}_{ij}}{\partial h_k} \frac{\partial h_k}{\partial h_{k-1}} \cdots \frac{\partial h_1}{\partial \overline{\epsilon}_{ij}} \tag{13}$$

όπου h_k είναι η έξοδος στο αντίστοιχο κρυφό επίπεδο k_{th} του FFNN για είσοδο $\overline{\epsilon}_{ij}$.

Η όλη διαδικασία συνοψίζεται στα ακόλουθα βήματα:

1. Δημιουργία N τυχαίων δειγμάτων εισόδου $\hat{\boldsymbol{q}}_i = [\overline{\boldsymbol{\epsilon}}_i, \boldsymbol{\theta}_i]$, $i = 1, 2, .., N$ εντός ορισμένων δειγματικών περιοχών.

2. Λύση της μη γραμμικής εξίσωσης ομογενοποίησης του RVE για κάθε διάνυσμα εισόδου $\hat{\boldsymbol{q}}_i$ και λήψη του αντίστοιχου διανύσματος εξόδου $\hat{\boldsymbol{y}}_i = [\overline{\boldsymbol{\sigma}}_i]$ για όλες τις ενδιάμεσες επαυξητικές λύσεις n_{incr}

3. Επιλογή της αρχιτεκτονικής FFNN και εκπαίδευση του χρησιμοποιώντας τα ζεύγη εισόδου-εξόδου $N \times n_{incr}$

4. Υπολογισμός του εφαπτομενικού μητρώου $\overline{\boldsymbol{C}}$ μέσω της μεθόδου AD στο FFNN.
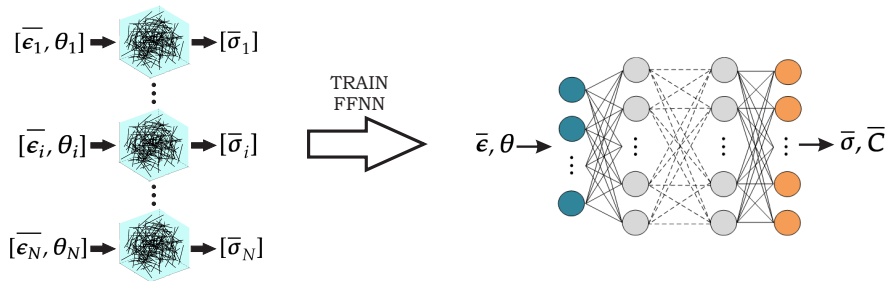
Σχήμα 4: Διαδικασία εκπαίδευσης του FFNN
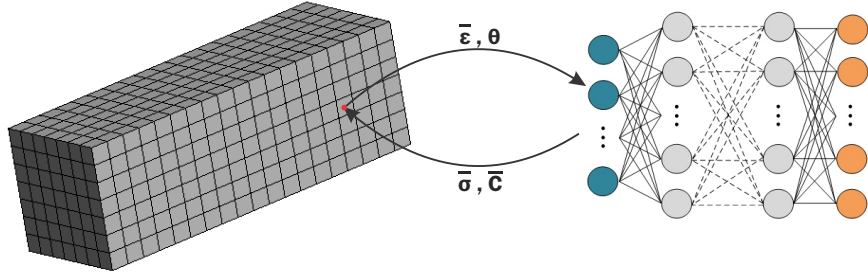
Σχήμα 5: Μπεϋζιανή επικαιροποίηση σε μοντέλα πολλαπλών κλιμάκων, επιταχυνόμενα με FFNN

Σε αρκετές περιπτώσεις δύο κλίμακες δεν επαρκούν για να περιγράψουν με ακρίβεια την συμπεριφορά κάποιων περίπλοκων υλικών. Η μοντελοποίηση των υλικών αυτών απαιτεί τον σχηματισμό περισσότερων από δύο κλιμάκων μήκους. Παρόμοια με την διαδικασία υπολογιστικής ομογενοποίησης που περιγράφηκε προηγουμένως, πρέπει να δημιουργηθεί μια σύνδεση μεταξύ αυτών των κλιμάκων κατά τη διάρκεια της λύσης. Αυτό μπορεί να επιτευχθεί διαδοχικά μεταξύ κάθε ζεύγους διαδοχικών κλιμάκων. Όπως στην εξ. (7) ο κανόνας τοπικοποίησης εφαρμόζεται αρχικά για κάθε ζεύγος ως:

$$\boldsymbol{\epsilon}^2(\boldsymbol{x}^2) = \frac{1}{\|\Omega^1\|}\int_{\Omega^1}\boldsymbol{\epsilon}^1(\boldsymbol{x}^1;\boldsymbol{x}^2)d\Omega^1 \ , \ \cdots \ , \ \boldsymbol{\epsilon}^M(\boldsymbol{x}^M) = \frac{1}{\|\Omega^N\|}\int_{\Omega^N}\boldsymbol{\epsilon}^N(\boldsymbol{x}^N;\boldsymbol{x}^M)d\Omega^N$$
$$(14)$$

Προς διευκρίνιση του παραπάνω συμβολισμού, η μεταβλητή μετά το ερωτηματικό, για παράδειγμα η $\boldsymbol{x}^2$ στο $\boldsymbol{\epsilon}^1(\boldsymbol{x}^1;\boldsymbol{x}^2)$, υποδηλώνει ότι το πεδίο παραμόρφωσης $\boldsymbol{\epsilon}^1$ μέσα στο ολοκλήρωμα είναι αυτό που αναπτύχθηκε από το RVE που σχετίζεται με το υλικό σημείο $\boldsymbol{x}^2$ της ανώτερης κλίμακας.

Στη συνέχεια, συνδυάζοντας όλα τα διαδοχικά βήματα τοπικοποίησης, μπορεί να αποκτηθεί μια σχέση που περιγράφει ολόκληρη την αλληλεπίδραση του συστήματος ως:

$$\boldsymbol{\epsilon}^M(\boldsymbol{x}^M) = \frac{1}{\|\Omega^N\|}\int_{\Omega^N}\cdots\underbrace{\underbrace{\frac{1}{\|\Omega^1\|}\int_{\Omega^1}\boldsymbol{\epsilon}^1(\boldsymbol{x}^1;\boldsymbol{x}^2)d\Omega^1}_{\boldsymbol{\epsilon}^2(\boldsymbol{x}^2;\boldsymbol{x}^3)}\ldots d\Omega^N}_{\boldsymbol{\epsilon}^N(\boldsymbol{x}^N;\boldsymbol{x}^M)} \quad (15)$$

Ξεκινώντας από την πιο λεπτή κλίμακα, η λύση των RVE μπορεί να επιτευχθεί εφαρμόζοντας το βήμα ομογενοποίησης της εξ. (9) για κάθε ζεύγος ως:

xvi

$$\boldsymbol{\sigma}^2(\boldsymbol{x}^2) = \frac{1}{\|\Omega^1\|} \int_{\Omega^1} \boldsymbol{\sigma}^1(\boldsymbol{x}^1, \boldsymbol{\alpha}^1; \boldsymbol{x}^2) d\Omega^1 \;,\; \cdots \;,\; \boldsymbol{\sigma}^M(\boldsymbol{x}^M) = \frac{1}{\|\Omega^N\|} \int_{\Omega^N} \boldsymbol{\sigma}^N(\boldsymbol{x}^N, \boldsymbol{\alpha}^N; \boldsymbol{x}^M) d\Omega^N$$

$$(16)$$

όπου $\boldsymbol{\alpha}^s$ είναι οι εσωτερικές μεταβλητές της s κλίμακας. Η σύνδεση των διαδοχικών σταδίων ομογενοποίησης δίνει στη συνέχεια:

$$\boldsymbol{\sigma}^M(\boldsymbol{x}^M) = \frac{1}{\|\Omega^N\|} \int_{\Omega^N} \cdots \underbrace{\frac{1}{\|\Omega^1\|} \int_{\Omega^1} \boldsymbol{\sigma}^1(\boldsymbol{x}^1, \boldsymbol{\alpha}^1; \boldsymbol{x}^2) d\Omega^1}_{\boldsymbol{\sigma}^2(\boldsymbol{x}^2, \boldsymbol{\alpha}^2; \boldsymbol{x}^3)} \ldots d\Omega^N \qquad (17)$$

Συνεπώς, ο εφαπτομενικός καταστατικός πίνακας διαμορφώνεται ως εξής:

$$\boldsymbol{C}^M(\boldsymbol{x}^M) = \partial_{\boldsymbol{\epsilon}^M} \frac{1}{\|\Omega^N\|} \int_{\Omega^N} \cdots \underbrace{\frac{1}{\|\Omega^1\|} \int_{\Omega^1} \boldsymbol{\sigma}^1(\boldsymbol{x}^1, \boldsymbol{\alpha}^1; \boldsymbol{x}^2) d\Omega^1}_{\boldsymbol{\sigma}^2(\boldsymbol{x}^2, \boldsymbol{\alpha}^2; \boldsymbol{x}^3)} \ldots d\Omega^N \qquad (18)$$

Εφαρμόζοντας την έννοια της μη γραμμικής υπολογιστικής ομογενοποίησης δύο κλιμάκων και επεκτείνοντάς την σε εμφωλευμένα συστήματα πολλαπλών κλιμάκων, μπορεί να επιτευχθεί μια εκτίμηση των συνολικών λύσεων RVE που απαιτούνται. Σε αυτήν την περίπτωση, η απαίτηση για επαναλαμβανόμενες λύσεις μεταξύ οποιουδήποτε τοπικού συστήματος δύο κλιμάκων πρέπει να προστεθεί πάνω στην την απαίτηση για επαναλαμβανόμενες λύσεις μεταξύ κάθε διαδοχικού ζεύγους κλιμάκων που βρίσκονται ψηλότερα από το τρέχον ζεύγος. Ο συνολικός αριθμός λύσεων προβλημάτων οριακής τιμής που απαιτούνται για μια πλήρη λύση ολόκληρου του συστήματος δίνεται από:

$$K_{RVE} = k^M + k^M \times k^N + \cdots + k^M \times k^N \times \cdots \times k^2 \qquad (19)$$

όπου για κάθε κλίμακα s, το k^s ορίζεται ως:

$$k^s = n_{int}^s \times n_{iter}^s \times n_{incr}^s \qquad (20)$$

με n_{int}^s να είναι ο αριθμός των σημείων ολοκλήρωσης στο διακριτοποιημένο σύστημα, n_{iter}^s ο αριθμός των επαναλήψεων της διαδικασίας Newton και n_{incr}^s ο αριθμός των επαυξήσεων της ανάλυσης.

Είναι εμφανές από την εξ. (19) ότι για να εκτελεστεί η πλήρης λύση της μακροδομής, η ποσότητα των προβλημάτων που πρέπει να λυθούν εξαρτάται άμεσα από τη συνολική ποσότητα των RVE που περιγράφουν το σύστημα πολλαπλών κλιμάκων και ξεφεύγει εκτός διαχειρίσιμων πλαισίων πολύ γρήγορα για αυξανόμενο αριθμό κλιμάκων. Κατά συνέπεια, αυτή η διαδικασία καθίσταται υπολογιστικά δυσεπίλυτη ακόμη και για μοντέλα μακροκλίμακας χαμηλής ακρίβειας και, ως εκ τούτου, η εύρεση ενός τρόπου παράκαμψης αυτού του τεράστιου υπολογιστικού κόστους είναι ζωτικής σημασίας.

Με βάση τα ανωτέρω, στη συνέχεια της διατριβής προτείνεται μια μη παρεμβατική στρατηγική υποκατάστασης μοντελοποίησης, αφιερωμένη σε προσεγγίσεις υπολογιστικής ομογενοποίησης FE^N σε συστήματα με πολλές κλίμακες ($N > 2$). Η ιδέα είναι να χρησιμοποιηθεί μια ακολουθία βαθέων νευρωνικών δικτύων (Deep Neural Networks - DNN) που αντιπροσωπεύουν την ιεραρχία των διαφορετικών κλιμάκων στο πρόβλημα πολλαπλών κλιμάκων. Κάθε νευρωνικό δίκτυο υποβάλλεται στην εκμάθηση του φυσικού νόμου στην αντίστοιχη κλίμακα μήκους του προβλήματος. Παρόμοια με το αρχικό πρόβλημα όπου κάθε λεπτότερη κλίμακα περιέχεται σε μια πιο αδρή κλίμακα, τα DNN που αντιπροσωπεύουν λεπτές κλίμακες περιέχονται στα DNN που αντιπροσωπεύουν πιο αδρές κλίμακες. Στο τέλος της διαδικασίας εκπαίδευσης, προκύπτει ένα ενιαίο DNN που μιμείται τη μακροσκοπική συμπεριφορά ενσωματώνοντας όλους τους φυσικούς μηχανισμούς που εντοπίζονται σε κάθε μία από τις λεπτότερες κλίμακες του προβλήματος. Αυτή η προσέγγιση εκμεταλλεύεται πλήρως την ακρίβεια και τις δυνατότητες μοντελοποίησης που παρέχουν τα σχήματα FE^N, ενώ ταυτόχρονα ξεπερνά τις τεράστιες υπολογιστικές τους απαιτήσεις. Συγκεκριμένα, τα DNN είναι επιφορτισμένα με την εκμάθηση παραμετροποιημένων εκδόσεων του καταστατικού νόμου σε κάθε κλίμακα, κάτι που επιτρέπει την μοντελοποίηση ενός ευρέος φάσματος πιθανών συμπεριφορών υλικού. Αυτό επιτυγχάνεται με την επαύξηση του επιπέδου εισόδου των DNN με το σύνολο των αβέβαιων παραμέτρων υλικού. Με αυτήν την προσέγγιση, κάθε DNN ενσωματώνει την αβέβαιη συμπεριφορά που προέρχεται από όλες τις προηγούμενες κλίμακες και τελικά μια μακροσκοπική καταστατική απόκριση που ενσωματώνει όλες αυτές τις πληροφορίες λαμβάνεται μέσω του τελικού DNN. Με τη σειρά του, αυτό το DNN μπορεί να εφαρμοστεί ως υποκατάστατο μοντέλο του υλικού σε οποιοδήποτε μακροδομικό σύστημα και για διάφορα προβλήματα πολλαπλών επιλύσεων (π.χ. ανάλυση ευαισθησίας, βελτιστοποίηση, Μπεϋζιανή επικαιροποίηση).

Συγκεκριμένα, τα βήματα για την κατασκευή του εμφωλευμένου σχήματος N κλιμάκων με τη βοήθεια των DNN είναι τα ακόλουθα:

- Εφαρμόζεται η διαδικασία διαδοχικής ομογενοποίησης σύμφωνα με την εξ. (7) και (9) σε κάθε ζεύγος διαδοχικών κλιμάκων ξεκινώντας από την λεπτότερη, δηλαδή την 1η κλίμακα

- Ορίζεται ένα σύνολο παραμέτρων/μεταβλητών που καθορίζουν την είσοδο και την έξοδο του αντίστοιχου DNN. Η είσοδος περιλαμβάνει τα στοιχεία του διανύσματος παραμόρφωσης στην ανώτερη κλίμακα $\boldsymbol{\epsilon}^{(2)}$, ενώ η έξοδος είναι το αντίστοιχο διάνυσμα τάσης $\boldsymbol{\sigma}^{(2)}$ που προκύπτει από την διαδικασία ομογενοποίησης. Εάν υποθέσουμε

περαιτέρω ότι η λύση του συστήματος επηρεάζεται από ένα διάνυσμα παραμέτρων $\boldsymbol{\alpha}^{(1)}$ που χαρακτηρίζουν τον καταστατικό νόμο του υλικού στην λεπτότερη κλίμακα, τότε αυτές οι παράμετροι θεωρούνται επίσης ως είσοδοι στο DNN.

- Το εφαπτομενικό καταστατικό μητρώο $\boldsymbol{C}^{(2)}$ για μια συγκεκριμένη κατάσταση παραμόρφωσης $\boldsymbol{\epsilon}^{(2)}$, που απαιτείται στις επαναλήψεις Newton-Raphson, υπολογίζεται άμεσα χρησιμοποιώντας την μέθοδο Automatic Differentiation (AD). Μέσω αυτής επιτυγχάνεται ο υπολογισμός των παραγώγων της εξόδου $\boldsymbol{\sigma}^{(2)}$ σε σχέση με την είσοδο $\boldsymbol{\epsilon}^{(2)}$ εφαρμόζοντας τον κανόνα της αλυσίδας στο DNN και, επομένως, τα στοιχεία $c_{ij}^{(2)}$ του μακροσκοπικού εφαπτομενικού πίνακα $\boldsymbol{C}^{(2)} = [c_{ij}^{(2)}]$ λαμβάνονται ως:

$$c_{ij}^{(2)} = \frac{\partial \sigma_{ij}^{(2)}}{\partial h_k} \frac{\partial h_k}{\partial h_{k-1}} \cdots \frac{\partial h_1}{\partial \epsilon_{ij}^{(2)}} \tag{21}$$

όπου h_k είναι η έξοδος του κρυφού επιπέδου k_{th} του DNN για την είσοδο $\epsilon_{ij}^{(2)}$.

- Μετά την επιτυχή εκπαίδευση και επαλήθευση, η λεπτότερη κλίμακα DNN (κλίμακα 1) μπορεί να εφαρμοστεί άμεσα για να αναπαραστήσει την καταστατική σχέση του υλικού μήτρας στην επόμενη κλίμακα (κλίμακα 2). Η διαδικασία επαναλαμβάνεται για αυτήν την κλίμακα όπου δημιουργείται ένα δεύτερο DNN χρησιμοποιώντας το διάνυσμα $\left(\boldsymbol{\epsilon}^{(3)}, \boldsymbol{\alpha}^{(2)}, \boldsymbol{\alpha}^{(1)}\right)$ ως είσοδο και το διάνυσμα $\boldsymbol{\sigma}^{(3)}$ ως έξοδο. Για απλοποίηση της σημειογραφίας, η είσοδος $\left(\boldsymbol{\epsilon}^{(3)}, \boldsymbol{\alpha}^{(2)}, \boldsymbol{\alpha}^{(1)}\right)$ γράφεται ως $\left(\boldsymbol{\epsilon}^{(3)}, \hat{\boldsymbol{\alpha}}^{(3)}\right)$, όπου, στη γενική περίπτωση, $\hat{\boldsymbol{\alpha}}^{(s+1)} = \left(\boldsymbol{\alpha}^{(s)}, \cdots, \boldsymbol{\alpha}^{(1)}\right)$ είναι το επαυξημένο παραμετρικό διάνυσμα. Όμοια, ο εφαπτομενικός πίνακας $\boldsymbol{C}^{(3)}$ είναι άμεσα διαθέσιμος μέσω της AD. Είναι σημαντικό να σημειωθεί εδώ ότι το DNN της δεύτερης κλίμακας περιλαμβάνει επίσης τις παραμέτρους υλικού $\boldsymbol{\alpha}^{(1)}$ της προηγούμενης κλίμακας ως είσοδο, καθώς αυτό θα του επιτρέψει να καταγράψει τη συνολική συμπεριφορά του υλικού, έως το σημείο αυτό, η οποία επηρεάζεται τόσο από την πρώτη όσο και από την δεύτερη κλίμακα.

- Αυτή η διαδικασία επαναλαμβάνεται για όλες τις κλίμακες μέχρι τη μακροκλίμακα, όπου τελικά καταλήγει σε ένα μόνο τελικό DNN που ενσωματώνει όλες τις πληροφορίες από τις χαμηλότερες κλίμακες και αποτελεί το υποκατάστατο μοντέλο της μακροσκοπικής συμπεριφοράς του σύνθετου υλικού.

Ως επαλήθευση του προτεινόμενου σχήματος υποκατάστασης μοντελοποίησης, μελετάται ένα 4-κλιμάκων μοντέλο σκυροδέματος οπλισμένο με CNT. Αρχικά, το DNN^{micro} εκπαιδεύεται με σκοπό την αντικατάσταση του RVE της μικροκλίμακας, που αποτελείται από τσιμεντόπαστα και CNT, χρησιμοποιώντας τα $\left(\boldsymbol{\epsilon}^{fm}, \boldsymbol{\alpha}^{\mu}\right)$ ως είσοδο και τα $\boldsymbol{\sigma}^{fm}$ ως έξοδο. Στη συνέχεια, το $DNN^{meso,fine}$ που αντιπροσωπεύει το RVE της λεπτή μεσοκλίμακας, δηλαδή την τσιμεντοκονία, εκπαιδεύεται σε ζεύγη $\left(\boldsymbol{\epsilon}^{cm}, \boldsymbol{\alpha}^{fm}, \boldsymbol{\alpha}^{\mu}\right)$, όπου τώρα το DNN^{micro} θεωρείται ως το υλικό μήτρας μαζί με τα λεπτά αδρανή ως εγκλείσματα. Η διαδικασία επαναλαμβάνεται άλλη μια φορά για το $DNN^{meso,coarse}$ του RVE της αδρής μεσοκλίμακας,
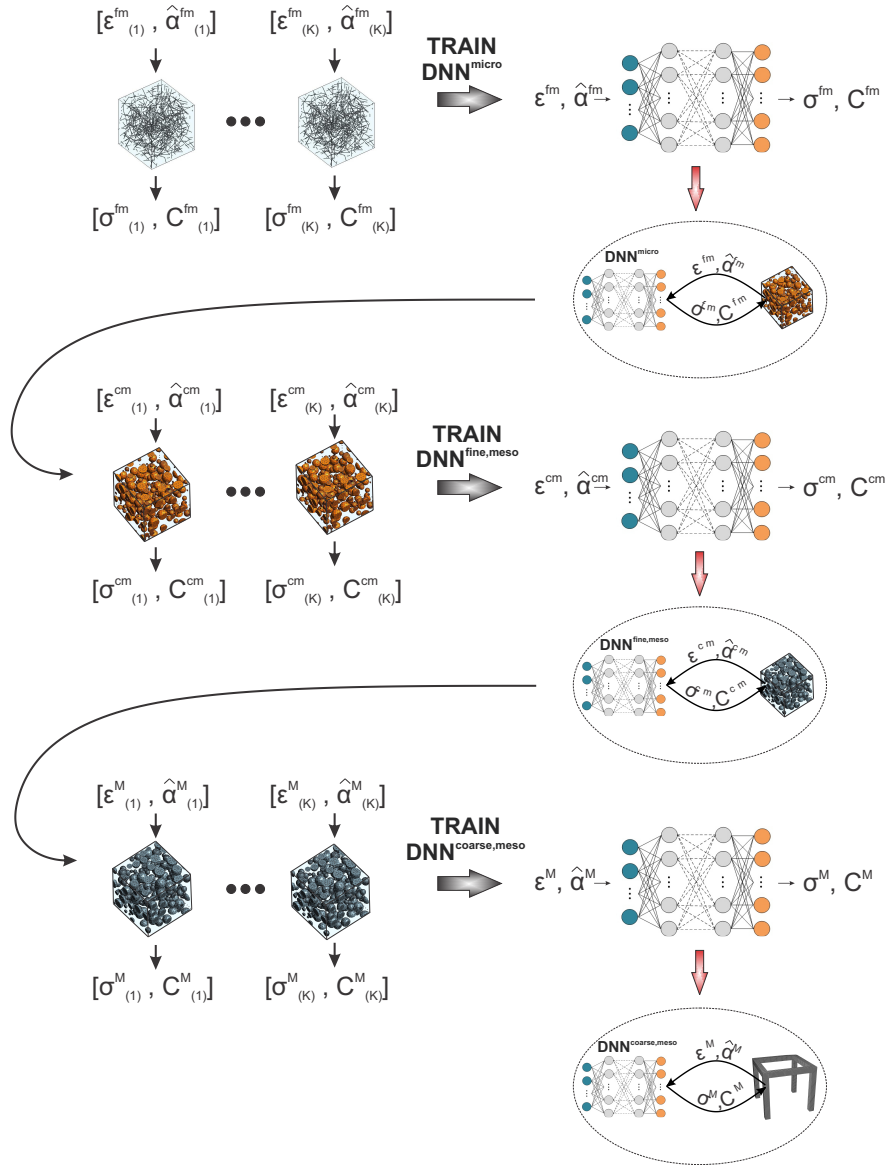
Σχήμα 6: Διαδικασία εκπαίδευσης σύμφωνα με την προτεινόμενη στρατηγική. Ξεκινώντας από τη μικροκλίμακα, ένα νευρωνικό δίκτυο, με όνομα DNN^{micro} εκπαιδεύεται να μιμείται τη συμπεριφορά τάσης-παραμόρφωσης του RVE της μικροκλίμακας. Το DNN^{micro} χρησιμοποιείται ως καταστατικός νόμος του υλικού της μήτρας στη λεπτή μεσοκλίμακα και, στη συνέχεια, το $DNN^{fine,meso}$ εκπαιδεύεται για να μιμείται τη συμπεριφορά τάσης-παραμόρφωσης του RVE της λεπτής μεσοκλίμακας. Η διαδικασία επαναλαμβάνεται άλλη μια φορά για την αδρή μεσοκλίμακα όπου διαμορφώνεται το τελικό δίκτυο $DNN^{coarse,meso}$ που ενσωματώνει με επιτυχία τη συνολική συμπεριφορά του σύνθετου υλικού.

(α΄)

(β΄)
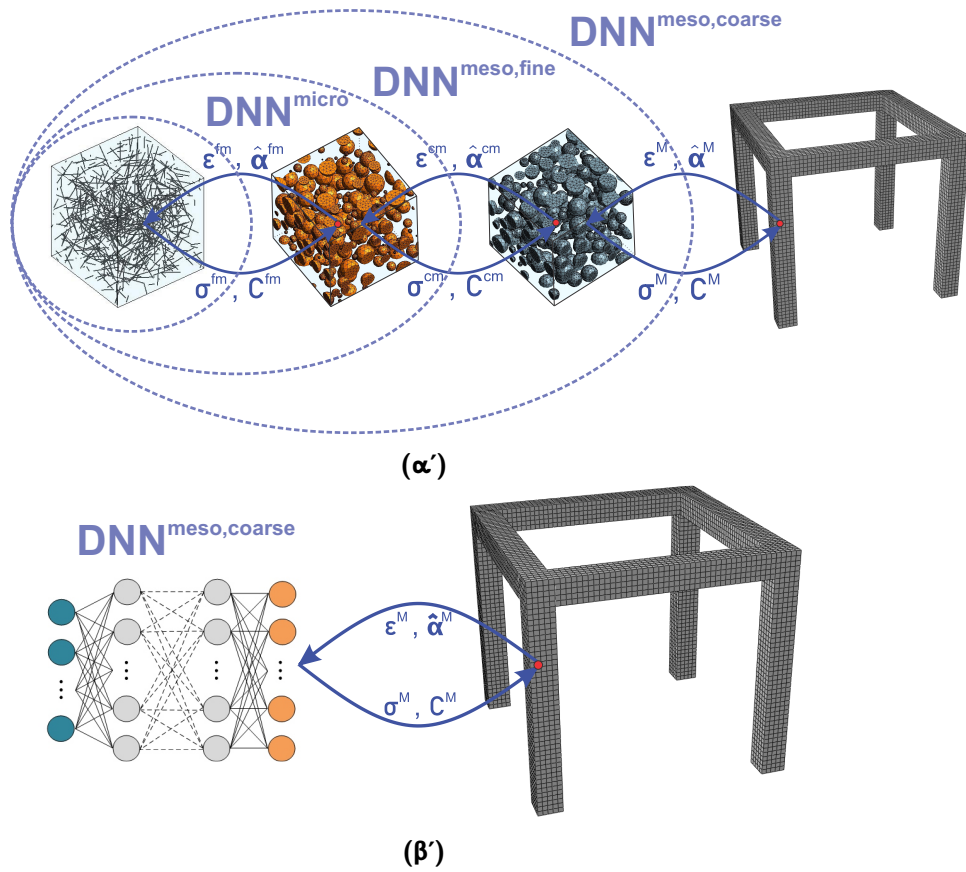
Σχήμα 7: (α) Ιεραρχική κατασκευή των υποκατάστατων μοντέλων DNN και (β) Επικοινωνία των κλιμάκων χρησιμοποιώντας το τελικό $DNN^{meso,coarse}$ ως τον καταστατικό νόμο της μακροκλίμακας

το οποίο είναι το τελικό DNN που ενσωματώνει όλα τα προηγούμενα DNN και αντιπροσωπεύει την καταστατική σχέση για το πρόβλημα μακροκλίμακας. Η διαδικασία εκπαίδευσης απεικονίζεται στο σχήμα 6 και το τελικό αποτέλεσμα της διαδικασίας στα σχήματα 7α΄ και 7β΄.

Για την περαιτέρω ανάδειξη των δυνατοτήτων της προτεινόμενης μεθόδου εκτελείται μια ανάλυση ευαισθησίας (Sentitivity Analysis - SA) και ειδικότερα μια καθολική ανάλυση ευαισθησίας με βάση την διακύμανση στο εξεταζόμενο μοντέλο. Συγκεκριμένα η SA εφαρμόζεται για εκτίμηση της ευαισθησίας των κατασκευών που αποτελούνται από σκυρόδεμα οπλισμένο με CNT όσων αφορά τις παραμέτρους υλικού στη μικροδομή. Εδώ, η SA εκτελείται στις παραμέτρους του μοντέλου πλαστικότητας Drucker-Prager, δηλαδή, τον συντελεστή τριβής ϕ, τον συντελεστή διαστολής ψ, την αρχική συνοχή c_0 και τον συντελεστή σκλήρυνσης H, σε σκυρόδεμα οπλισμένο με CNT. Οι παράμετροι του μοντέλου συμβολίζο-

νται συλλογικά ως $\hat{\boldsymbol{\alpha}}^M = (\phi, \psi, c_0, H)$. Στη συνέχεια, η SA αποδίδει μια εκτίμηση για το πώς η αβεβαιότητα διαδίδεται στις διάφορες κλίμακες του μοντέλου και επηρεάζει τη μακροσκοπική απόκριση. Σε αυτήν τη ρύθμιση, το μοντέλο $\mathcal{M} := f(\phi, \psi, c_0, H) := f(\hat{\boldsymbol{\alpha}}^M)$ μπορεί να αντιπροσωπεύει την ποσότητα ενδιαφέροντος, π.χ. μια μετατόπιση σε μια δεδομένη θέση της μακροσκοπικής κατασκευής, η οποία λαμβάνεται μετά την επίλυση του κατασκευαστικού προβλήματος με το σχήμα FE^4. Ο στόχος είναι να αξιολογηθούν οι δείκτες ευαισθησίας S_i και οι συνολικοί δείκτες S_{T_i} για $i = 1, 2, 3, 4$ που αντιστοιχούν στις τέσσερις παραμέτρους του μοντέλου ϕ, ψ, c_0, H, αντίστοιχα. Για να γίνει αυτό, πρέπει να παραχθούν $K \cdot (d+2)$ δειγματοληψίες από το παραμετρικό διάνυσμα $\{\hat{\boldsymbol{\alpha}}_i^M\}_{i=1}^{K(d+2)}$ σύμφωνα με μια προκαθορισμένη κατανομή πιθανότητας, όπου το K είναι συνήθως της τάξης των $10^3 \div 10^5$ ούτως ώστε να παραχθούν αξιόπιστες εκτιμήσεις των δεικτών ευαισθησίας και $d = dim(\hat{\boldsymbol{\alpha}}^M) = 4$. Για καθεμία από αυτές τις περιπτώσεις, το πρόβλημα μακροκλίμακας λύνεται προκειμένου να ληφθούν οι αποκρίσεις $\{\mathcal{M}_i\}_{i=1}^{K \cdot (d+2)}$ και να υπολογιστούν τα $f(\boldsymbol{R}_{(j)}), f(\boldsymbol{Q}_{(j)})$ και $f(\boldsymbol{Q}_{\boldsymbol{R},(j)}^{(i)})$, όπου $\boldsymbol{R}, \boldsymbol{Q}$ και $\boldsymbol{Q_R}$ είναι οι δειγματικοί πίνακες.

Λαμβάνοντας υπόψη την πολυπλοκότητα του μοντέλου, γίνεται προφανές ότι αυτός ο τύπος ανάλυσης θα ήταν υπολογιστικά μη πραγματοποιήσιμος με μια άμεση εφαρμογή του σχήματος FE^4. Ωστόσο, το προτεινόμενο υποκατάστατο σχήμα μοντελοποίησης μπορεί να χρησιμοποιηθεί για να επιταχύνει τις επαναλαμβανόμενες επιλύσεις μοντέλων για διάφορες περιπτώσεις των παραμέτρων ϕ, ψ, c_0, H. Μια σχηματική αναπαράσταση της διαδικασίας SA σε μια γενική ανάλυση πολλαπλών κλιμάκων υποβοηθούμενη από τα DNN απεικονίζεται στο σχήμα 8. Σύμφωνα με αυτό, $K \cdot (d+2)$ δειγματοληψίες του παραμετρικού διανύσματος $\{\hat{\boldsymbol{\alpha}}_i^M\}_{i=1}^{K \cdot (d+2)}$ δημιουργούνται μετά από μια καθορισμένη πιθανοτική κατανομή και για καθεμία από αυτές τις περιπτώσεις το πρόβλημα μακροκλίμακας επιλύεται για να ληφθούν οι αποκρίσεις $\{\mathcal{M}_i\}_{i=1}^{K(d+2)}$. Οι επιλύσεις μοντέλων $K \cdot (d+2)$ που είναι απαραίτητες για τη σύγκλιση της SA αναμένεται να εκτελεστούν σε προσιτό υπολογιστικό χρόνο, καθώς το τελικό DNN, δηλαδή το DNN^N, έχει αντικαταστήσει το ακριβό FE^N υπολογιστικό σχήμα. Αφού συλλεχθούν τα δείγματα και πραγματοποιηθεί η SA, μπορούν να εξαχθούν κρίσιμα συμπεράσματα για τις παραμέτρους του μοντέλου πολλαπλών κλιμάκων παρατηρώντας πώς επηρεάζουν την έξοδο του μοντέλου.
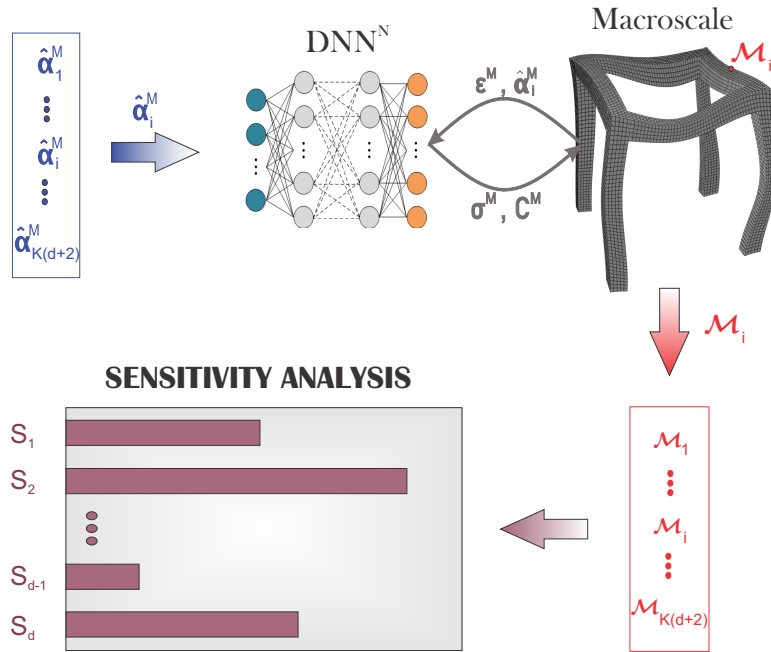
Σχήμα 8: Ανάλυση ευαισθησίας σε σύστημα πολλαπλών κλιμάκων επιταχυνόμενη από το προτεινόμενο υποκατάστατο μοντέλο. Το πρόβλημα μακροκλίμακας επιλύεται για περιπτώσεις $K \cdot (d+2)$ του διανύσματος παραμέτρων $\hat{\alpha}^M$. Η αντίστοιχη έξοδος μοντέλου $\{\mathcal{M}_i\}_{i=1}^{K(d+2)}$ μπορεί να υπολογιστεί αποτελεσματικά βάσει της προτεινόμενης εναλλακτικής στρατηγικής μοντελοποίησης. Στη συνέχεια, οι δείκτες ευαισθησίας των παραμέτρων του υλικού μπορούν να υπολογιστούν εύκολα μέσω στατιστικής μετα-επεξεργασίας.

Στην επόμενη φάση της διατριβής παρουσιάζεται ένα καινοτόμο πλαίσιο βελτιστοποίησης για τον εντοπισμό βέλτιστων τυπολογιών υλικού με στόχο τη βελτίωση της δομικής απόκρισης υπό την παρουσία αβεβαιοτήτων. Συγκεκριμένα, εξετάζεται ο βέλτιστος προσανατολισμός των CNT σε μοντέλα σκυροδέματος ενισχυμένα με CNT. Συνεπώς, οι χωρικές γωνίες των CNT έχουν τον ρόλο των μεταβλητών σχεδιασμού στο συγκεκριμένα πρόβλημα βελτιστοποίησης. Αυτή η συγκεκριμένη παράμετρος έχει αποδειχθεί ότι παίζει κρίσιμο ρόλο στην απόδοση του σύνθετου υλικού. Προς επίτευξης μιας τέτοια διερεύνησης, ως μέθοδος βελτιστοποίησης χρησιμοποιείται η μέθοδος Covariance Matrix Adaptation - Evolution Strategy (CMA-ES), η οποία καθώς δεν βασίζεται σε παραγώγους καθίσταται ικανή να χειριστεί μη κυρτά προβλήματα συνεχούς βελτιστοποίησης και είναι κατάλληλη για υπολογιστικά ακριβά αριθμητικά μοντέλα. Με αυτή τη μεθοδολογία γίνεται δυνατή η εξερεύνηση βέλτιστων προσανατολισμών CNT σε κάθε θέση της μακροκλίμακας των εξεταζόμενων μοντέλων με στόχο την ελαχιστοποίηση της δυνατής συνολικής μετατόπισης. Προχωρώντας ένα βήμα παραπέρα, στην προσπάθεια παροχής μιας πιο ορθολογικής και στιβαρής προσέγγισης, λαμβάνεται επίσης υπόψη η τυχαιότητα στην εξωτερική φόρτιση της υπό διερεύνηση

κατασκευής και επαναδιατυπώνεται το πρόβλημα στο πλαίσιο της στοχαστικής βελτιστοποίησης. Χάρη στον υπολογιστικό μηχανισμό της αλληλουχίας DNN που αναπτύχθηκε στην προηγούμενη μελέτη, γίνεται εφικτή η βελτιστοποίηση του υλικού με γνώμονα την απόδοση της κατασκευής που χαρακτηρίζεται από σκυρόδεμα οπλισμένο με CNT. Πιο συγκεκριμένα αναπτύσσεται ένα μοντέλο 3 κλιμάκων για να αναπαραστήσει το εξεταζόμενο υλικό και η επίλυση του επιτυγχάνεται μέσω της ανάπτυξης του κατάλληλου ιεραρχικού σχήματος από DNN με στόχο την παράκαμψη της εκτέλεσης της μη γραμμικής υπολογιστικής ομογενοποίησης.

Έστω $\mathcal{U}$ το μαθηματικό μοντέλο του υπό διερεύνηση δομικού συστήματος. Δεδομένου ότι στις περισσότερες περιπτώσεις πρακτικού ενδιαφέροντος το μαθηματικό μοντέλο δεν μπορεί να εξαχθεί αναλυτικά, η πιο κοινή προσέγγιση είναι η αντικατάστασή του με ένα αριθμητικό μοντέλο $\boldsymbol{U}$, που συνήθως λαμβάνεται με FEM. Από αυτή την άποψη, το $\boldsymbol{U} \in \mathbb{R}^d$ είναι ένα διάνυσμα διαστάσεων d που αντιστοιχεί στους d βαθμούς ελευθερίας της FEM διακριτοποίησης. Επιπλέον, εάν η κατασκευή υπόκειται σε συνθήκες τυχαίας φόρτωσης, που εκφράζονται μέσω του r-διάστατου τυχαίου διανύσματος $\boldsymbol{q}$, με το r να είναι ο αριθμός των τυχαίων μεταβλητών, τότε $\boldsymbol{U} := \boldsymbol{U}(\boldsymbol{q})$. Επιπλέον, η δομική απόκριση επηρεάζεται από το διαμόρφωση του υλικού στην μικροκλίμακα, του οποίου οι ιδιότητες παραμετροποιούνται σύμφωνα με τον προσανατολισμό των CNT, $\boldsymbol{\theta} := \boldsymbol{\theta}^\mu = (\theta_p, \theta_a)$ με $\theta_p \in [0\ pi]$ την πολική γωνία και $\theta_a \in [0\ 2\pi)$ το αζιμούθιο, όπως φαίνεται στο σχήμα 9. Επομένως, το τελικό μοντέλο μπορεί να εκφραστεί ως $\boldsymbol{U} := \boldsymbol{U}(\boldsymbol{q}, \boldsymbol{\theta})$.
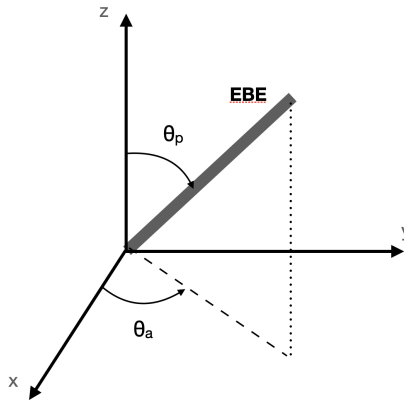
Σχήμα 9: Πολική γωνία θ_p και αζιμούθιο θ_a, που χαρακτηρίζουν τον προσανατολισμό ενός CNT στον τρισδιάστατο χώρο

Για να δημιουργηθεί μια βάση αναφοράς, εξετάζεται αρχικά η περίπτωση πλήρους απουσίας CNT στην τσιμεντοκονία. Αυτό συμβαίνει επειδή στόχος είναι η αξιολόγηση της συμπεριφοράς του συμβατικού σκυροδέματος σε σύγκριση με το οπλισμένο με CNT. Επομένως, πραγματοποιείται ένας μεγάλος αριθμός, N_{ref}, δειγματοληψιών $\boldsymbol{q}_1, ..., \boldsymbol{q}_{N_{ref}}$ και

εκτελούνται οι αντίστοιχες προσομοιώσεις μοντέλων $\boldsymbol{U}_{ref}^{i} = \boldsymbol{U}(\boldsymbol{q})$ για $i = 1, ..., N_{ref}$. Στη συνέχεια, υπολογίζεται ο μέσος όρος και η τυπική απόκλιση από αυτό το στατιστικό δείγμα ως:

$$\mathbb{E}[\boldsymbol{U}_{ref}] = \frac{1}{N_{ref}} \sum_{i=1}^{N_{ref}} \boldsymbol{U}_i \tag{22}$$

$$Std[\boldsymbol{U}_{ref}] = \sqrt{\sum_{i=1}^{N_{ref}} \frac{(\boldsymbol{U}_i - \mathbb{E}[\boldsymbol{U}_{ref}]) \odot (\boldsymbol{U}_i - \mathbb{E}[\boldsymbol{U}_{ref}])}{N_{ref} - 1}} \tag{23}$$

όπου το $\odot$ υποδηλώνει το γινόμενο Hadamard. Υπολογίζεται η μέση συνολική παραμόρφωση της κατασκευής, $\mathscr{M}$, και η τυπική απόκλιση, $\mathscr{S}$, με τον κανόνα L_2 των παραπάνω διανυσμάτων, δηλαδή,

$$\mathscr{M} = \|\mathbb{E}[\boldsymbol{U}_{ref}]\|_2 \tag{24}$$

$$\mathscr{S} = \|[Std[\boldsymbol{U}_{ref}]\|_2 \tag{25}$$

Το πρόβλημα στοχαστικής βελτιστοποίησης ορίζεται στη συνέχεια ως εξής: Για ένα δεδομένο κλάσμα βάρους CNTs στο κονίαμα, αναζητούνται οι τιμές του διανύσματος σχεδιασμού $\boldsymbol{\theta} = (\theta_p, \theta_a)$ που ελαχιστοποιούν το σταθμισμένο άθροισμα του μέσου όρου και της τυπικής απόκλισης της συνολικής παραμόρφωσης, κανονικοποιημένης από τις αντίστοιχες τιμές αναφοράς που δίνονται από τις εξ. (24) και (25). Αυτό μπορεί να εκφραστεί μαθηματικά ως εξής:

$$\boldsymbol{\theta}^{\star} = (\theta_p^{\star}, \theta_a^{\star}) = \arg\min 2_{\theta_p \in [0\ \pi], \theta_a \in [0\ 2\pi]} w_1 \frac{\mathbb{E}[\boldsymbol{U}]}{M} + w_2 \frac{Std[\boldsymbol{U}]}{S}$$

$$= \arg\min 2_{\theta_p \in [0\ \pi], \theta_a \in [0\ 2\pi]} \mathscr{L}(\boldsymbol{\theta}) \tag{26}$$

με $\mathscr{L}(\boldsymbol{\theta}) = w_1 \frac{\mathbb{E}[\boldsymbol{U}]}{M} + w_2 \frac{Std[\boldsymbol{U}]}{S}$ που δηλώνει τη συνάρτηση απώλειας (loss function).

Στο τελευταίο στάδιο της διατριβής προτείνεται μια νέα μέθοδος για τον προσδιορισμό των ιδιοτήτων των υλικών σε συστήματα πολλαπλών κλιμάκων μέσω δεδομένων που προέρχονται από ετερογενή πειραματικά σενάρια. Το παρουσιαζόμενο υπολογιστικό σχήμα έχει την δυνατότητα να συγχωνεύσει δεδομένα που αποκτήθηκαν από πειράματα τα οποία πραγματοποιήθηκαν σε υλικά διαφορετικής σύνθεσης και μπορεί να συμπεριλάβει μετρήσεις που έγιναν σε διάφορες κλίμακες μήκους, επιτρέποντας τη συστηματική ενσωμάτωση πολλαπλών πειραματικών πηγών δεδομένων σε ένα ενοποιημένο υπολογιστικό πλαίσιο. Για να επιτευχθεί αυτό, χρησιμοποιείται η μέθοδος Transitional Markov Chain Monte Carlo (TM-CMC) για δειγματοληψία από τις περιθωριοποιημένες εκ των υστέρων κατανομές τόσο των

παραμέτρων του μοντέλου πολλαπλών κλιμάκων όσο και των ιεραρχικών υπερπαραμέτρων. Αυτές οι υπερπαράμετροι χρησιμοποιούνται στη συνέχεια για την εξαγωγή ενημερωμένων φυσικών παραμέτρων, οι οποίες μπορούν να χρησιμοποιηθούν για μελλοντικές προβλέψεις μοντέλων.

Έστω ένα σύνολο δεδομένων $\boldsymbol{D} = [\boldsymbol{D}^1, ..., \boldsymbol{D}^K]$, με $\boldsymbol{D}^i = [d^{i,1}, ..., d^{i,N_i}]$, που περιλαμβάνει μετρήσεις (δηλαδή μηχανικές αποκρίσεις) που λαμβάνονται από έναν αριθμό K ανεξάρτητων πειραμάτων. Κάθε i πείραμα ορίζεται σε ένα πεδίο Ω^i και περιγράφεται από έναν συγκεκριμένο συνδυασμό συνοριακών συνθηκών $\partial\Omega^i$ και σύνθεσης υλικού. Στο πλαίσιο της ανάλυσης σύνθετων υλικών, αυτά τα πειραματικά σενάρια μπορούν να προσομοιωθούν από τα αντίστοιχα υπολογιστικά μοντέλα πολλαπλών κλιμάκων $\mathcal{M}^1, ..., \mathcal{M}^K$ που χρησιμοποιούνται για την πρόβλεψη δομικών αποκρίσεων $\boldsymbol{m}^1, ..., \boldsymbol{m}^K$ σύμφωνα με τα πειραματικά δεδομένα. Κάθε μοντέλο $\mathcal{M}^i$ χαρακτηρίζεται από μια μοναδική σύνθεση υλικού, για παράδειγμα με διαφορετικό αριθμό κλιμάκων μήκους, ανόμοια σύνθεση σε όλες ή ένα υποσύνολο των κλιμάκων μήκους, κ.λπ. Κατά συνέπεια, η παραμετροποίηση αυτών των μοντέλων γίνεται μέσω ενός διακριτού συνόλου $\boldsymbol{\theta}^i$, ως $\boldsymbol{\theta}^i = [\boldsymbol{\theta}^i_1, ..., \boldsymbol{\theta}^i_{S_i}]$, που περιλαμβάνει μια σειρά φυσικών, τοπολογικών ή καταστατικών χαρακτηριστικών για κάθε κλίμακα μήκους $s = 1, ..., S_i$ του αντίστοιχου συστήματος. Σε πολλές περιπτώσεις, ένα υποσύνολο αυτών των παραμέτρων $^c\boldsymbol{\theta}^i \subseteq \boldsymbol{\theta}^i$, με $^c\boldsymbol{\theta}^i = [^c\boldsymbol{\theta}^i_1, ..., ^c\boldsymbol{\theta}^i_{S_i}]$, υπάρχει σε όλα τα υλικά μοντέλα πολλαπλών κλιμάκων που διερευνώνται, πράγμα που σημαίνει ότι $\boldsymbol{\theta}^1 \cap ... \cap \boldsymbol{\theta}^K \equiv {}^c\boldsymbol{\theta}^1 \equiv ... \equiv {}^c\boldsymbol{\theta}^K \equiv {}^c\boldsymbol{\theta}$.

Ως γραφική αναπαράσταση του συνολικού προβλήματος το σχήμα 10 απεικονίζει ένα σενάριο τριών μοντέλων όπου πρέπει να διερευνηθούν οι παράμετροι μιας συγκεκριμένης κλίμακας. Η ιεραρχική στρατηγική Bayes, που συμπληρώνεται από τις υπερπαραμέτρους $\boldsymbol{\psi}$, επιτρέπει την αναγνώριση νέων φυσικών παραμέτρων $^c\boldsymbol{\theta}^{new}$ με αξιόπιστο τρόπο. Με τη σειρά του, τα $^c\boldsymbol{\theta}^{new}$ μπορούν να διευκολύνουν τεκμηριωμένες προβλέψεις $\boldsymbol{y}^{new}$ σε καινούργια συστήματα πολλαπλών κλιμάκων. Είναι σημαντικό να τονιστεί ότι οι πηγές δεδομένων έχουν τη δυνατότητα να περιλαμβάνουν ένα ευρύ φάσμα πειραμάτων που διεξάγονται σε διάφορες κλίμακες μήκους. Αυτά τα πειράματα, για παράδειγμα, μπορεί να περιλαμβάνουν μετρήσεις μετατόπισης στη μακροκλίμακα, χαρακτηρισμό τοπολογίας στη μεσοκλίμακα μέσω εικόνων που προέρχονται μέσω της μεθόδου Scanning Electron Microscopy (SEM) ή εξαγωγή πεδίου παραμόρφωσης στην μικροκλίμακα χρησιμοποιώντας την τεχνική Digital Image Correlation (DIG).
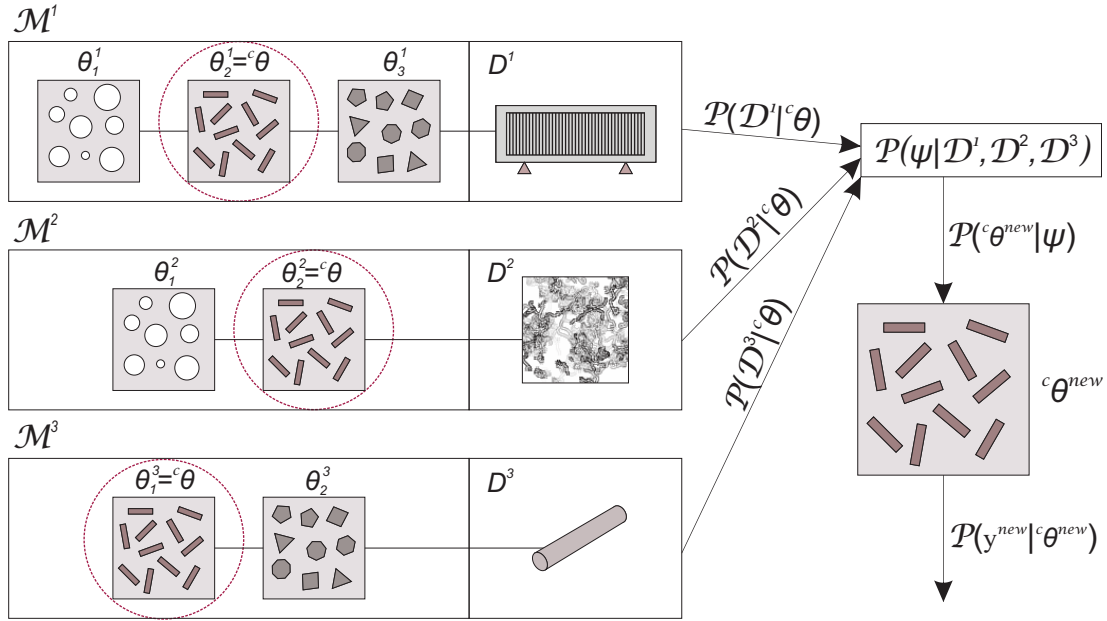
Σχήμα 10: Προτεινόμενο πλαίσιο για την επικαιροποίηση παραμέτρων και τη διάδοση της αβεβαιότητας των κοινών ιδιοτήτων $^c\boldsymbol{\theta}$ που εντοπίζονται σε μια σειρά από ποικίλες πειραματικές διαμορφώσεις, όπως ένα μοντέλο $\mathcal{M}^1$ δοκιμής κάμψης σε ενισχυμένη δοκό κατασκευασμένη από σύνθετο υλικό, μοντέλο $\mathcal{M}^2$ ψηφιακής εικόνας μικροδομής ενός άλλου σύνθετου υλικού και μοντέλο $\mathcal{M}^3$ δείγματος ράβδου ενός διαφορετικού σύνθετου.

Σε τέτοιες περιπτώσεις, η κλασσική Μπεϋζιανή επικαιροποίηση υποτιμά τη συνολική μεταβλητότητα. Αυτό οφείλεται στο ότι τείνει να αντιπροσωπεύει μόνο την «αβεβαιότητα εκτίμησης παραμέτρων» και δεν είναι σε θέση να υπολογίσει την εξωτερική μεταβλητότητα μεταξύ των συνόλων δεδομένων. Αυτή η μεταβλητότητα υποδηλώνει την αβεβαιότητα που δεν μπορεί να ποσοτικοποιηθεί ρητά στη διατύπωση του προβλήματος λόγω έλλειψης γνώσης σχετικά με τις συγκεκριμένες συνθήκες του πειράματος. Λαμβάνοντας υπόψη αυτά τα ευρήματα, γίνεται η επιλογή μιας ιεραρχικής Μπεϋζιανής προσέγγισης για τη διαμόρφωση της προτεινόμενης στρατηγικής.

Με βάση το ιεραρχικό παράδειγμα Bayes, κάθε σύνολο δεδομένων $\boldsymbol{D}^i$ που λαμβάνεται από μια μοναδική πειραματική ρύθμιση εξετάζεται ξεχωριστά. Για κάθε μία από αυτές, οι κοινές παράμετροι $\boldsymbol{\theta}$ έχουν έναν ξεχωριστό ορισμό ως $\boldsymbol{\theta}^i$. Αυτή η διάκριση είναι απαραίτητη αφού αυτές οι παράμετροι, αν και έχουν την ίδια φυσική σημασία μεταξύ όλων των συνόλων δεδομένων, μπορούν τελικά να αντιπροσωπεύονται από διαφορετικές τιμές λόγω της εξωτερικής μεταβλητότητας. Θεωρείται περαιτέρω ότι κάθε $\boldsymbol{\theta}^i$ εξαρτάται από μια σειρά υπερπαραμέτρων $\boldsymbol{\psi}$, που συμβολίζονται ως $P(\boldsymbol{\theta}^i|\boldsymbol{\psi})$. Αυτές οι παράμετροι χρησιμοποιούνται για να παρέχουν στο μαθηματικό πλαίσιο την αίσθηση συγγένειας μεταξύ των $\boldsymbol{\theta}^i$ για $i = 1, .., K$ και για να λάβουν ρητά υπόψη τις διαφοροποιήσεις του μοντέλου στα αντίστοιχα

σύνολα δεδομένων. Οι υπερπαράμετροι αντιπροσωπεύουν γενικά στατιστικές παραμέτρους μιας προκαθορισμένης οικογένειας κατανομών. Το κλασικό Μπεϋζιανό πρόβλημα που τίθεται στην εξ. (11) αναδιατυπώνεται στην ιεραρχική παραλλαγή ως:

$$\boldsymbol{y}^i(\boldsymbol{\theta}^i|\boldsymbol{\psi}) = \boldsymbol{m}^i(\boldsymbol{\theta}^i|\boldsymbol{\psi}) + \boldsymbol{\epsilon}^i \qquad i = 1, ..., K \tag{27}$$

Αντίστοιχα, η από κοινού εκ των υστέρων κατανομή εκφράζεται μέσω του θεωρήματος Bayes ως:

$$P(\boldsymbol{\theta}, \boldsymbol{\psi}|\boldsymbol{D}) = \frac{P(\boldsymbol{D}|\boldsymbol{\theta})P(\boldsymbol{\theta}|\boldsymbol{\psi})P(\boldsymbol{\psi})}{P(\boldsymbol{D})} = \prod_{i=1}^{K}\Big[\prod_{j=1}^{N_i}\big[P(d^{i,j}|\boldsymbol{\theta}^i)\big]P(\boldsymbol{\theta}^i|\boldsymbol{\psi})\Big]\frac{P(\boldsymbol{\psi})}{P(\boldsymbol{D})} \tag{28}$$

όπου $P(\boldsymbol{\psi})$ είναι η εκ των προτέρων κατανομή των υπερπαραμέτρων.

Για να αποφευχθεί το ασύμφορο υπολογιστικό κόστος της απευθείας δειγματοληψίας από την από κοινού κατανομή της εξ. (28), προτείνεται ένας εναλλακτικός τρόπος διαδοχικής δειγματοληψίας.

Το πρώτο βήμα είναι η δειγματοληψία από την περιθωριοποιημένη εκ των υστέρων κατανομή των υπερπαραμέτρων του μοντέλου, η οποία ορίζεται ως:

$$P(\boldsymbol{\psi}|\boldsymbol{D}) = \int_{\Omega_{\boldsymbol{\theta}}} P(\boldsymbol{D}|\boldsymbol{\theta})P(\boldsymbol{\theta}|\boldsymbol{\psi})d\boldsymbol{\theta}\frac{P(\boldsymbol{\psi})}{P(\boldsymbol{D})} = \prod_{i=1}^{K}\Big[\int_{\Omega_{\boldsymbol{\theta}^i}}\prod_{j=1}^{N_i}\big[P(d^{i,j}|\boldsymbol{\theta}^i)\big]P(\boldsymbol{\theta}^i|\boldsymbol{\psi})d\boldsymbol{\theta}^i\Big]\frac{P(\boldsymbol{\psi})}{P(\boldsymbol{D})} \tag{29}$$

όπου στην παραπάνω εξίσωση χρησιμοποιήθηκε το γεγονός ότι η συνάρτηση πιθανοφάνειας $P(\boldsymbol{D}|\boldsymbol{\psi})$ παίρνει τη μορφή:

$$P(\boldsymbol{D}|\boldsymbol{\psi}) = \prod_{i=1}^{K}\Big[\int_{\Omega_{\boldsymbol{\theta}^i}}\prod_{j=1}^{N_i}\big[P(d^{i,j}|\boldsymbol{\theta}^i)\big]P(\boldsymbol{\theta}^i|\boldsymbol{\psi})d\boldsymbol{\theta}^i\Big] \tag{30}$$

Το ολοκλήρωμα της εξ. (29) μπορεί να υπολογιστεί κατά προσέγγιση μέσω δειγματοληψίας Monte Carlo:

$$P(\boldsymbol{\psi}|\boldsymbol{D}) \simeq \prod_{i=1}^{K}\Big[\frac{1}{N_{\boldsymbol{\theta}^i}}\sum_{k=1}^{N_{\boldsymbol{\theta}^i}}P(\boldsymbol{\theta}_k^i|\boldsymbol{\psi})\Big]\frac{P(\boldsymbol{\psi})}{P(\boldsymbol{D})} \tag{31}$$

Για να πραγματοποιηθεί αυτή η ολοκλήρωση, πρέπει πρώτα να συλλεχθούν δείγματα από κάθε συνάρτηση πιθανοφάνειας $P(\boldsymbol{D}^i|\boldsymbol{\theta}^i)$ που σχετίζεται με το μοντέλο M^i. Όταν η TMCMC χρησιμοποιείται ως μέθοδος δειγματοληψίας για μία εκ των υστέρων κατανομή, το αρχικό βήμα είναι να συλλεχθεί μια ποσότητα δειγμάτων από την εκ των προτέρων κατανομή. Εφόσον στην παρούσα περίπτωσή η δειγματοληψία γίνεται απευθείας από την

$P(\boldsymbol{D}^i|\boldsymbol{\theta}^i)$, ένας εύκολος τρόπος για να αναπαραχθεί η τυπική αλγοριθμική διαδικασία της TMCMC είναι να χρησιμοποιηθούν βοηθητικές ομοιόμορφες εκ των προτέρων κατανομές $P(\boldsymbol{\theta}^i)$ για την αρχική δειγματοληψία από την κάθε κατανομή $P(\boldsymbol{D}^i|\boldsymbol{\theta}^i)$. Για να αποφευχθεί η εισαγωγή οποιασδήποτε μεροληψίας στη διαδικασία, θα πρέπει να επιλεγεί το κατώτερο όριο u_l και το ανώτερο όριο u_u των προηγούμενων $\mathcal{U}_{\boldsymbol{\theta}^i}(u_l, u_u)$ αρκετά ευρύ για να καλύψει το δείγμα του χώρου κάθε $\boldsymbol{\theta}^i$ που ορίζεται από την $P(\boldsymbol{D}^i|\boldsymbol{\theta}^i)$.

Η περιθωριοποιημένη κατανομή των ενημερωμένων παραμέτρων των μοντέλων πολλαπλών κλιμάκων $\boldsymbol{\theta}^{new}$ που λαμβάνουν υπόψη όλα τα σύνολα δεδομένων και μπορούν να εφαρμοστούν σε μελλοντικές προβλέψεις εκφράζεται ως:

$$P(\boldsymbol{\theta}^{new}|\boldsymbol{D}) = \int_{\Omega_{\boldsymbol{\psi}}} P(\boldsymbol{\theta}^{new}|\boldsymbol{\psi})P(\boldsymbol{\psi}|\boldsymbol{D})d\boldsymbol{\psi} \tag{32}$$

Η προσεγγιστική λύση της εξ. (32) γίνεται με τη δημιουργία μιας ποσότητας δειγμάτων $\boldsymbol{\psi}$ μέσω της εμπειρικής κατανομής $P(\boldsymbol{\psi}|\boldsymbol{D})$ που λαμβάνεται από την εξ. (31), ως εξής:

$$P(\boldsymbol{\theta}^{new}|\boldsymbol{D}) \simeq \sum_{k=1}^{N_{\boldsymbol{\psi}}} P(\boldsymbol{\theta}^{new}|\boldsymbol{\psi}_k) \tag{33}$$

Για την παραγωγή δειγμάτων από την εξ. (32), χρησιμοποιείται ξανά ο αλγόριθμος TM-CMC, όπου στην περίπτωση αυτή η αρχική δειγματοληψία εκτελείται στην εκ των προτέρων κατανομή των υπερπαραμέτρων $P(\boldsymbol{\psi})$. Αντίθετα με τη δειγματοληψία από την $P(\boldsymbol{D}^i|\boldsymbol{\theta}^i)$, η εξ. (33) αντιπροσωπεύει μια φθηνή υπολογιστική διαδικασία, καθώς δεν χρειάζονται αναλύσεις μοντέλων πολλαπλών κλιμάκων.

Αφού ληφθεί η εκ των υστέρων πιθανολογική μορφή του $\boldsymbol{\theta}^{new}$ σύμφωνα με το ιεραρχικό σχήμα Bayes , αυτές οι παράμετροι μπορούν στη συνέχεια να χρησιμοποιηθούν για οποιαδήποτε ανάλυση διάδοσης αβεβαιότητας σε μελλοντικές προσομοιώσεις Y^{new} όπως φαίνεται στο σχήμα 11. Αυτές οι δοκιμές θα μπορούσαν να πραγματοποιηθούν σε καινούργια υλικά συστήματα που περιγράφονται εν μέρει από τις επικαιροποιημένες παραμέτρους. Η ποσότητα ενδιαφέροντος (π.χ. μια δομική απόκριση) $\boldsymbol{y}^{new}$ υπολογίζεται ως:

$$P(\boldsymbol{y}^{new}|\boldsymbol{D}) = \int_{\Omega_{\boldsymbol{\theta}^{new}}} P(\boldsymbol{y}^{new}|\boldsymbol{\theta}^{new})P(\boldsymbol{\theta}^{new}|\boldsymbol{D})d\boldsymbol{\theta}^{new} \tag{34}$$
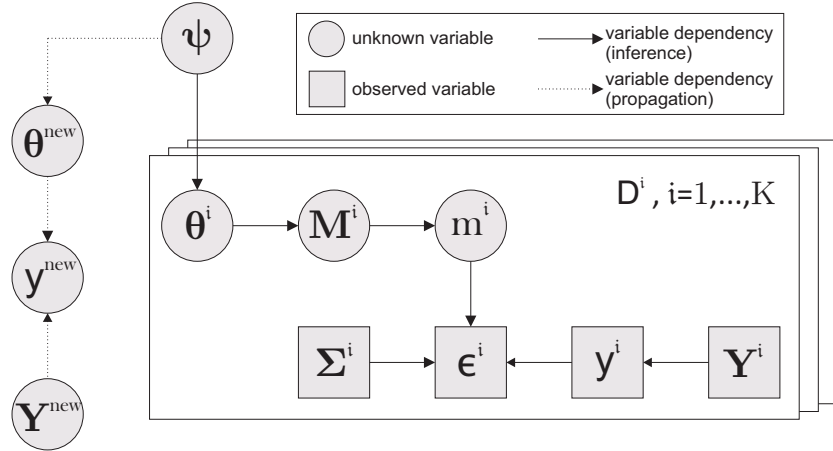
Σχήμα 11: DAG του ιεραρχικού προβλήματος Bayes

Το προτεινόμενο πλαισίο εφαρμόστηκε στην διερευνήση των μηχανικών ιδιοτήτων της διεπαφής των CNTs, σύμφωνα με τον καταστατικό νόμο που έχει οριστεί στο σχήμα 2, σε διαμορφώσεις τσιμεντοειδών υλικών. Για να πραγματοποιηθεί η ταυτοποίηση των παραμέτρων του υλικού που ερευνήθηκαν, ελήφθησαν τα αποτελέσματα από διαφορετικά πειράματα που ανακτήθηκαν από τη βιβλιογραφία. Αυτά περιλαμβάνουν ένα πείραμα κάμψης 3 σημείων ενός δείγματος τσιμεντόπαστας ενισχυμένης με CNT, ένα πείραμα εφελκυσμού που πραγματοποιήθηκε σε ράβδους τσιμεντοκονίας ενισχυμένων με ˜NT και ένα πείραμα κάμψης 4 σημείων σε μια δοκό σκυροδέματος ενισχυμένη με CNT. Για την προσομοίωση των πειραμάτων αναπτύχθηκαν τα κατάλληλα μοντέλα πολλαπλών κλιμάκων και για την επίλυση χρησιμοποιήθηκε η τεχνική της εμφώλευσης στα πλαίσια της υπολογιστικής ομογενοποίησης. Η τεχνική υποκατάστατης μοντελοποίησης μέσω της ανάπτυξης αλληλουχίας νευρωνικών που προτάθηκε σε προηγούμενη μελέτη αποδείχθηκε απαραίτητη για την δραστική μείωση του υπολογιστικού κόστους. Τα βήματα της αλγοριθμικής διαδικασίας της προτεινόμενης μεθόδου παρουσιάζονται παρακάτω.

Input	: Σύνολο δειγμάτων εκπαίδευσης n, συνολικές προσαυξήσεις t και υπερπαράμετροι FFNN. Παραμετροποιημένα υλικά μοντέλα πολλαπλών κλιμάκων $M^i(\boldsymbol{\theta}^i)$, προηγούμενες κατανομές $P(\boldsymbol{\psi})$ και $P(\boldsymbol{\theta}^i)$. Σύνολο δειγμάτων $N^{\boldsymbol{\theta}^i}$ και παραμέτρων TMCMC k^i και β^i. Σύνολο δειγμάτων $N^{\boldsymbol{\psi}}$ και παραμέτρων TMCMC k^{new} και β^{new}.
Output	: ενημερωμένη κατανομή των παραμέτρων $P(\boldsymbol{\theta}^{new}|\boldsymbol{D})$, ενημερωμένη κατανομή των ποσοτήτων ενδιαφέροντος $P(\boldsymbol{y}^{new}|\boldsymbol{D})$

Offline Stage:
for $i \leftarrow 1$ **to** K **do**
 for $j \leftarrow 1$ **to** S_i **do**
  Δημιουργία και αποθήκευση δειγμάτων εκπαίδευσης $\{\{\boldsymbol{\epsilon}_{j+1}\}_{1:t}, \hat{\boldsymbol{\theta}}_j^i\}_{1:n}$·
  Επίλυση του προβλήματος οριακής τιμής της κλίμακας j·
  Αποθήκευση των τάσεων από την επίλυση $\{\{\boldsymbol{\sigma}_{j+1}\}_{1:t}\}_{1:n}$·
  Εκπαίδευση του FFNN $f_j^{NN,i}$·
  Αποθήκευση του FFNN $f_j^{NN,i}$;
 end
end
Online Stage:
for $i \leftarrow 1$ **to** K **do**
 Δημιουργία δειγμάτων από την $P^{NN,i}(\boldsymbol{D}^i|\boldsymbol{\theta}^i)$ ως
  $\{\boldsymbol{\theta}^i\}_{1:N^{\boldsymbol{\theta}^i}} \leftarrow TMCMC(p(\boldsymbol{\theta}^i), P^{NN,i}(\boldsymbol{D}^i|\boldsymbol{\theta}^i), N^i, k^i, \beta^i)$·
 Αποθήκευση δειγμάτων $\{\boldsymbol{\theta}^i\}_{1:N^{\boldsymbol{\theta}^i}}$ και δεδομένα συνάρτησης πιθανοφάνειας
  $P^{NN,i}(\boldsymbol{D}^i|\{\boldsymbol{\theta}^i\}_{1:N^{\boldsymbol{\theta}^i}})$
end
Εκτίμηση $P(\boldsymbol{\psi}|\boldsymbol{D})$·
Δημιουργία δειγμάτων από την $P(\boldsymbol{\psi}|\boldsymbol{D})$ ως
 $\{\boldsymbol{\psi}\}_{1:N^{\boldsymbol{\psi}}} \leftarrow TMCMC(P(\boldsymbol{\psi}), P(\boldsymbol{D}|\boldsymbol{\psi}), N^{\boldsymbol{\psi}}, k^{\boldsymbol{\psi}}, \beta^{\boldsymbol{\psi}})$·
Εκτίμηση $P(\boldsymbol{\theta}^{new}|\boldsymbol{D})$·
Διάδωση της εκ των υστέρων αβεβαιότητας $P(\boldsymbol{y}^{new}|\boldsymbol{D})$;

Algorithm 1: Αλγόριθμος της προτεινόμενης μεθοδολογίας για παραμετρική διερεύνηση ετερογενών υλικών μοντέλων πολλαπλών κλίμακων κάνοντας χρήση υποκατάστατης μοντελοποίησης

Contents

List of Figures

xxxvii

0
Introduction

The need for advanced high-performance materials is undoubtedly one of the main drivers of today's innovation in industry and research. Over the past decades, several new technologies have emerged for the development of materials with enhanced properties (mechanical, thermal, electrical) based on appropriate modifications of their composition in finer scales. For instance, carbon nanotubes (CNTs) have been extensively used to improve the mechanical properties of polymers [127, 43] and cementitious materials [67, 68, 142], graphene oxide to enhance the electrical conductivity in polymers [78] and graphite nanoplatelets to enhance the thermal conductivity on epoxy-based composites [143]. Even though the addition of appropriate fillers provides the means to create high performance materials with targeted properties, yet, the resources required to study their behavior experimentally greatly slow down the progress in this field. In this regard, a promising approach towards replacing the experimental procedure is given by simulation-based material design that could drastically reduce the time and cost required for the characterization of these elaborate materials. On the downside, accurately modeling their behavior is a remarkably challenging computational mechanics problem due to the complexity of the physical phenomena spanning across multiple scales, such as the nanoscale, microscale and mesoscale, as well as the uncertainties associated with the individual material parameters and microstructural geometry.

To tackle the intrinsic limitations that characterize experimental procedures, several

computational techniques have been developed capable of accurately modeling complex composite materials. The most notable are arguably the multiscale methods [85], according to which the system is decomposed and defined in a number of distinct length scales. Each of these scales has its own material 'genome', i.e. individual topological, geometrical and physical properties, such as the different material phases, the interaction between them, their concentrations, spatial distributions, dimensions etc. The material 'genomes' can vary significantly between scales and, in this respect, dissimilar and scale dependent physical phenomena can emerge. The concept of multiscale mechanics consists of identifying appropriate relationships to bridge the various length scales and quantifying these dependencies. This is done either straightforwardly by defining all the scales on a single system and introducing 'handshake' regions between them or with the introduction of a representative volume element (RVE), which is a system that is decoupled from the main model and constitutes a statistically representative realization of a specific scale. Linking the scales in a hierarchical manner i.e from the finest scale and upwards, the so-called homogenization takes place, which is an averaging over the field variables of the RVE of a specific scale.

Early homogenization-based multiscale approaches provided approximations of the effective constitutive behavior of multi-phase materials via analytical or semi-analytical relations [33, 89, 21, 53]. Although these could provide a sufficient representation of the material in the linear regime, they tend to behave poorly on non-linear materials that construe complex physical phenomena. More recent research approaches were based on solving a boundary value problem posed at the RVE configuration. In this context, resulting methodologies can either correspond to algorithmically sequential or semi-concurrent procedures. In sequential procedures, the constitutive responses are predetermined from offline RVE solutions [114], while in the semi-concurrent procedures a nested scheme is used to establish a constant communication between the different scales during the solution. One of the most notable frameworks that implement a nested scheme is the FE^2 algorithm, which is based on performing finite element analysis on both scales of a two-scale system [118, 71, 35, 70]. The nested approaches are considered to be most accurate for predicting the evolution of a system through different stages, however, a major downside is their immense computational requirements, which in many cases can be prohibitive.

Despite the tremendous advancements in computational power nowadays, the application of nested computational homogenization schemes on multiscale material systems

remains a computationally challenging and, in many cases, intractable problem. An efficient way to combat this issue is via the development of surrogate models which are simpler mathematical constructs that aim to replicate the input-output relation of the original equations through a training procedure. In the context of computational material mechanics, surrogate models have been extensively used to learn non-linear constitutive relations such as hyperelastic laws [60] or more complex history dependent relations such as plastic or viscoplastic laws [79, 1, 81]. Beyond the purely data-driven surrogate models for material laws, attempts have been made recently on the development of mechanistically consistent and physics-informed models. In an effort to develop NN-based material models that can provide accurate results with less training data, while also preserving physical consistency, several works have proposed methodologies that take into account physical information either through specialized NN architectures [141, 84] or by augmenting the loss function with additional terms to ensure physical consistency[2]. In a multiscale setting, several techniques have been proposed to alleviate the enormous computational demands of nested schemes. For instance, reduced order models (ROMs) have been applied in the governing equations of the lower scale either with linear [146, 36] or manifold-based methods [10]. Additionally, non-intrusive approaches such as response surfaces [125], Gaussian Processes [110, 37], deep feedforward neural networks (DNNs) [60, 82, 9] and recurrent neural networks (RNNs) [136, 41, 139, 13] have also been successfully employed. In these data-driven schemes, the objective is to uncover the macroscopic stress-strain or strain-strain energy relations from experimental data and/or synthetic data.

Besides computational material design of nanocomposites, these materials often have many uncertainties located at their fine structure, such as the morphology, the volume fraction, the mechanical properties of the individual materials and their interaction properties. The challenging task of identifying phenomenological material parameters, which cannot be directly measured through experiments, has been extensively explored in numerous studies within the literature. The objective of inferring the parameter values that provide the maximum likelihood estimation, has led to a wide adoption of several optimization strategies. Representative instances of such approaches include employing the least squares method for the calibration of statistical constitutive parameters of rocks [25] or of strain rate sensitivity parameters of metals [15] and the utilization of a genetic algorithm towards microstructural parameter inference for a visco-plastic-damage model for hardened cement [48]. Nevertheless, the reliability of single-point estimations

3

is not consistent, mainly because inverse problems are often ill-posed, especially in the context of multiscale material systems, where highly nonlinear phenomena are at play. In response to this limitation, the Bayesian paradigm [6, 12] has more recently gained traction as a means of material parameter investigation. This paradigm re-contextualizes the inverse problem within a stochastic framework. Bayesian parameter inference has been applied to empirical constitutive laws in various scenarios, including heterogeneous mediums [69], a visco-elastic model [105], a crystal plasticity-damage model [96] and a thermo-visco-plastic model [144]. In the area of materials with multiscale properties, randomness in these parameters is introduced to the numerical model via appropriate probabilistic descriptions, either using random variables or random fields. For instance, in [116] the volume fraction of CNTs in representative volume elements (RVEs) of the composite material is modeled as a random variable. A similar approach can be found in [134] concerning the length of the CNTs, their waviness, the agglomeration parameters and the effective Young's modulus. Also, in [115] random fields were used to describe the waviness of CNTs in the polymer matrix. However, accurately selecting the parameters of these probabilistic models is not straightforward, as they depend on the manufacturing process. This fact suggests that their calibration would require microscale measurements, which are hard to obtain.

This is particularly true for the estimation of the interaction properties between the CNTs and the polymer matrix, because these are not directly observable quantities and require sophisticated experimental setups [121, 102, 26]. In this dissertation a Bayesian analysis framework is proposed in [104] to address this issue, with the aim of updating the prior beliefs on the mechanical properties of carbon based nanocomposites. The focus is placed on learning the parameters that characterize the CNT/polymer interface in the microscale with data available from upper scales. As mentioned, these parameters are associated with great uncertainties and their characterization is a difficult task, since microscale measurements are costly and hard to obtain. To overcome this, the present study introduces a computational framework for updating the prior beliefs on the values of these parameters, by using deformation measurements on meso- and/or macro-scale structures composed of the composite. In terms of modeling, the CNT/polymer interface is formulated using a cohesive zone model and a bilinear bond-slip constitutive law. The FE^2 method is then employed for predicting the response of the composite structures, but despite its accuracy, this method is associated with tremendous computational demands for large-scale problems. Therefore, its application to the Bayesian setting that requires

multiple model evaluations is prohibitive. To alleviate this enormous cost, a surrogate modeling technique is developed which utilizes artificial neural networks, trained to predict the nonlinear stress-strain relationship of representative volume elements of the microstructure. The data set over which the neural network is trained, is obtained by analyzing a limited number of different RVE configurations using a detailed finite element analysis. The elaborated methodology is first validated through a numerical example from 2D elasticity, which demonstrated its high accuracy and its significant cost reduction capabilities. It is then applied to a more challenging large-scale problem from 3D elasticity. Even though this research focuses on the characterization of the mechanical properties of composite materials, the proposed numerical procedure is generic and can be straightforwardly applied to other physically analogous phenomena related to nano-composite modeling, such as parameter identification in heat transfer or electrical conduction.

Undoubtedly, concrete is the most common construction material world-wide, with its uses ranging from house construction to bridges and dams. Due to the critical role it plays in people's lives and their general well-being, vast research has been devoted to better understanding and improving its properties. In the last decades, several new technologies have emerged, aiming to improve its strength [129, 108], durability [111, 98] and fatigue resistance [45, 73], just to mention a few. The focus in previous attempts has mostly been in enhancing the performance of cementitious materials by adding appropriate fillers to the mix such as glass fibers [106, 124] and steel fibers [122] or chemical reagents such as metakaolin [3]. A different approach, however, which has attracted a major interest in the scientific community, opening new research fields within concrete technology, is that of reinforcing concrete at the nanoscale level using nanomaterials as fillers. There are numerous applications of nanomaterials as concrete reinforcement, including nano silica [11, 64], nano titanium [77], nano alumina [80], nano clay[16, 62], and carbon based nanomaterials such as graphene sheets and carbon nanotubes (CNTs) [76, 97, 142], which are considered to be among the most promising ones. The huge interest of researchers from diverse fields towards carbon nanomaterials and especially CNTs is due to their extraordinary mechanical [130], thermal [8] and electrical properties [29]. As an example, CNTs, which may consist of a single or multiple rolled graphene sheets, have an estimated Young's modulus of 1TPa and their tensile strength is 100 times larger than that of the strongest steel [128]. For this reason, they have been extensively investigated in numerous experimental works and have been

shown to significantly improve the mechanical properties of any host material, including concrete [90, 94].

As mentioned previously, nonlinear computational homogenization analyses, albeit being very accurate, they are associated with an enormous computational cost. This is especially true for multiscale material systems, such as CNT-reinforced concrete, that are defined by more than two distinct length scales. In such cases, nested computational homogenization schemes are prohibited even for trivial applications. This dissertation proposes a surrogate modeling strategy, in [103], dedicated to FE^N computational homogenization approaches on systems with many scales ($N > 2$). The idea is to employ a sequence of neural networks that represent the hierarchy of the separate scales in the multiscale problem. Each neural network is being trained to learn the physical law at a corresponding length scale of the problem. In a similar manner to the original problem where each finer scale is contained in a coarser scale, neural networks representing fine scales are contained in the DNNs that represent coarser scales. At the end of the training process, a single deep network which emulates the macroscopic behavior by incorporating all physical mechanisms arising at each of the problem's finer scales is derived. This approach takes full advantage of the accuracy and modeling capabilities that FE^N schemes provide, while at the same time overcomes their immense computational requirements. Specifically, the DNNs are tasked with learning parameterized versions of the constitutive law in each scale, which allows the modeling of a wide range of possible material behaviors. This is accomplished by augmenting the input layer of the DNNs with the set of the uncertain material parameters. With this approach, each DNN incorporates the uncertain behavior that comes from all the previous scales and ultimately a macroscopic constitutive response that encapsulates all this information is obtained through the final DNN. In turn, this DNN can be applied as a surrogate of the material in any macrostructural system and for various multi-query problems (e.g. sensitivity analysis, optimization, Bayesian inference).

The elaborated methodology is demonstrated on the analysis of a large-scale building made of CNT-reinforced concrete. This particular structural system is modeled as a four-scale problem consisting of (i) carbon nanotube-reinforced cement paste at the microscale, (ii) reinforced cement mortar at a fine mesoscale level, (iii) reinforced concrete at a coarse mesoscale level and (iv) the macroscopic structural system. The composite material is characterized by different nonlinear constitutive laws at each scale. The solution of the full multiscale problem is attained by using a FE^4 scheme at a reasonable

computational time by virtue of the elaborated surrogate modeling setup. In turn, this enables laborious sensitivity analyses in order to assess the uncertainty in the microscopic material parameters and its propagation to the macroscopic structural response.

Next, in the present thesis, a novel computational framework is developed in [65], for identifying optimal material typologies to improve structural performance under the presence of uncertainties. Specifically, the focus in this work is on carbon nanotube(CNT)-reinforced concrete with the optimization problem consisting in finding the optimal CNT orientation in the host material so as to minimize the total deformation of structures made up from the composite. A computational model for CNT-reinforced concrete is proposed which utilizes a 3-level hierarchical approach for material characterization. In particular, cement mortar enhanced with carbon nanotubes is studied at a microscale level, while the reinforced cement paste along with concrete aggregates is studied at a mesoscale level using a continuum micromechanics model. This, in turn, enables the study of realistic structural problems made of the composite and the assessment of their performance. The material optimization analysis is facilitated by means of the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) [51, 49] optimization algorithm, which is derivative-free method for handling non-convex continuous optimization problems and is suited for computationally expensive numerical models. With this methodology the optimal CNT orientations can be obtained at every location of the macroscale structure that will lead to a reduced overall deformation. Going one step further, in the effort to provide a more rational and robust approach, the randomness in the external loading of the structure under investigation is considered and therefore the problem is reformulated in the context of stochastic optimization. To alleviate the vast computational burden associated with this endeavor, the aforementioned surrogate modeling technique for FE^N material systems is employed for the macroscopic material prediction the studied 3-scale system. With this approach, a drastic cost reduction is achieved per structural analysis which allows the efficient execution of the stochastic structural optimization problem.

The ability to combine diverse datasets for parameter identification can provide significant benefits in computational material modeling. The motivation to do that is not exclusive to materials, given the current era of "big data". It has been recognized that in such situations, the conventional Bayesian method falls short in addressing the external variability due to the inability to explicitly quantify it within the model's structure. This variability stems from contradictory environmental, operational, and experimental

conditions across distinct datasets. With respect to material modeling, additional variability can be present due to dissimilar material composition among the experimental cases. All of these disparities lead to the underestimation of the overall uncertainty. To address the complex challenge, hierarchical Bayesian approaches have been proposed [22] in several scientific disciples such as structural dynamics [7], biomechanics [30] and machine learning [42]. This approach introduces an additional layer of parameters into the probabilistic model, referred to as hyperparameters, which act as statistical parameters, assuming the role of hyper-priors within the Bayesian framework. They effectively capture the external variability, leading to robust outcomes concerning the prediction of the posterior uncertainty. Regarding materials, hierarchical Bayesian strategies have been utilized to ascertain parameters that characterize macroscopic phenomenological material laws [92, 99], while in [126], a hierarchical Bayesian paradigm has been employed for microstructural parameter calibration via testing on tensile coupons with the same material layout.

This thesis, lastly, proposes a novel method for determining material properties within multiscale material systems through a range of experimental scenarios. The presented framework holds the promise of merging data acquired from experiments conducted on materials of different compositions and encompassing measurements taken at various length scales, allowing the systematic integration of multiple experimental data sources into a unified computational framework. To achieve this, the Transitional Markov Chain Monte Carlo (TMCMC) method is utilized for sampling from the marginalized posterior distributions of both multiscale model parameters and hierarchical hyperparameters. These hyperparameters are subsequently employed to derive informed physical parameters, which can be used for future model predictions. Crucially, feedforward neural networks (FFNNs) play a key role in reducing the computational complexity of implementing hierarchical Bayesian analysis on top of nonlinear computational homogenization. Their primary aim is to learn and accurately predict the nonlinear constitutive law across various scales. To evaluate the efficacy of the proposed approach, a study is carried out on the parameters that define the interfacial mechanical behavior of carbon nanotubes (CNTs) in CNT-reinforced cementitious material configurations. For this task, data have been collected from conventional experiments conducted on diverse material configurations defined at multiple length scales, each associated and characterized through a FE^2 based hierarchical multiscale computational model.

The first chapter familiarizes the reader with the concept of multiscale finite element

analysis in non-linear systems. The boundary value problems that define the macroscale and the microscale are posed followed by the scale transition process. Then the boundary valued coupled system is reformulated to systems with more than two scales. Focusing on the finite element method, the numerical solution of the microscale problem, the two-scale system and the N-scale system are detailed.

The second chapter introduces the modeling techniques that are applied towards the simulation of the tackled material systems. The common characteristic of all the composites studied in this dissertation is that they are reinforced with CNTs. The transition from molecular dynamics to structural mechanics for the CNT representation is first explained and then the fully bonded and cohesive formulations of the CNT/matrix composite systems are provided.

The third chapter is dedicated on the scientific fields of uncertainty quantification and optimization. Starting with a brief introduction in probability theory, several types of parametric analyses are then presented. The inverse problem by means of the single-level Bayesian Update method is formulated followed by the hierarchical Bayesian Update extension in cases of multiple datasets. Next, the global sensitivity concept is presented with a main focus on variance-based strategies. Lastly, the machinery that entails a stochastic optimization procedure is laid out.

The fourth chapter illustrates the concept of surrogate modeling in engineering application. Specifically, it focuses on machine learning based surrogate models, where concepts such as neural networks and automatic differentiation are reviewed. Then, ways to utilize machine learning techniques towards material modeling are outlined.

The fifth chapter presents a Bayesian framework for determining the mechanical properties of carbon based nanocomposites. In particular, Bayesian parameter inference is applied to learn the parameters that characterize the CNT/polymer interface in the microscale. The prior beliefs on the values of these parameters are updated, by using measurements on large-scale structures made of the composite. To alleviate the cost that multiple FE^2 analyses entail, a surrogate modeling technique is developed which utilizes DNNs. The DNN is trained to predict the nonlinear stress-strain relationship of representative volume elements of the microstructure. The elaborated methodology is validated through numerical examples on 2D and 3D elasticity.

The sixth chapter introduces a novel surrogate framework designed to accelerate the solution procedure of hierarchically formulated multiscale problems. The idea is to employ a sequence of deep neural networks (DNNs) that represent the hierarchy of the

separate scales in the multiscale problem. Each DNN is trained to learn the constitutive law of a corresponding length scale of the problem and ultimately, DNNs representing fine scales are contained in the DNNs that represent coarser scales. At the end of the training process, a single deep network is produced which emulates the parameterized macroscopic behavior. Based on this strategy, a global sensitivity analysis is performed on a four-scale CNT-reinforced concrete structural system described by a FE^4 model at reasonable computational times

The seventh chapter presents a numerical framework for the stochastic material optimization of structures made up of CNT-reinforced structures. Specifically, the focus lies in finding the optimal CNT orientation at different members of structural systems based on a minimization of the sum of the mean and standard deviation of the overall structural deformation. This methodology is demonstrated on a three-scale CNT-reinforced concrete structural system described by a FE^3 computational model.

The eighth chapter introduces a hierarchical Bayesian framework to infer the material properties of multiscale material systems through a variety of experimental data acquired from different length scales and/or different material compositions. An informed set of physical parameters that encapsulates the information from all the different experiments is produced and is used for making predictions in future material models. This framework is demonstrated on a case study of CNT-reinforced cementitious material configurations through the investigation of the CNT interfacial mechanical behavior, by utilizing experimental data on dissimilar material compositions.

The tenth chapter discusses the conclusions drawn from this research and presents a summary of the contributions.

1

Non-linear multiscale finite element analysis

1.1 Macroscale boundary value problem

Consider an elastic body M placed on the macroscopic structural scale and defined on a domain $\Omega^M \subset R^3$ with a Lipschitz-continuous boundary Γ^M. According to the infinitesimal strain theory, the Green-Lagrangian strain tensor $\boldsymbol{\epsilon}^M$ is expressed as:

$$\boldsymbol{\epsilon}^M = \frac{1}{2} \left(\nabla^M \boldsymbol{u}^M + \left(\nabla^M \boldsymbol{u}^M \right)^\mathsf{T} \right) \tag{1.1}$$

The strong formulation of the static equilibrium equation has the form:

$$\nabla^M \cdot \boldsymbol{\sigma}^M = \boldsymbol{b}^M \quad on \quad \Omega^M \tag{1.2}$$

and the boundary conditions are:

$$\boldsymbol{u}^M = \hat{\boldsymbol{u}}^M \quad on \quad \Gamma_D^M \tag{1.3a}$$

$$\boldsymbol{\sigma}^M \cdot \boldsymbol{n} = \hat{\boldsymbol{\tau}}^M \quad on \quad \Gamma_N^M \tag{1.3b}$$

$$\Gamma^M \equiv \Gamma_D^M \cup \Gamma_N^M \quad , \quad \Gamma_D^M \cap \Gamma_N^M \equiv \emptyset \tag{1.3c}$$

where $\boldsymbol{b}^M$ are the volumetric forces, $\boldsymbol{\sigma}^M$ is the Cauchy stress tensor, $\boldsymbol{n}$ is the outward unit normal vector, $\hat{\boldsymbol{u}}^M$ and $\hat{\boldsymbol{\tau}}^M$ are the prescribed displacements and forces on the boundaries, while Γ_D^M and Γ_N^M are the subsets of Γ^M with Dirichlet and Neumann conditions respectively.

Moreover, a constitutive relation for the macro-problem is established through the general form:

$$\boldsymbol{\sigma}^M(t) = \boldsymbol{\sigma}^M\{\boldsymbol{\epsilon}^M(t), \boldsymbol{\theta}^M(\bar{t}), \bar{t} \in [0, t]\} \tag{1.4}$$

with t denoting the pseudo-time in the case of a quasi-static analysis and $\bar{t}$ the history-dependence of the material's behavior. Additionally, $\boldsymbol{\theta}^M$ is used to declare a set of internal variables that characterize the state of the material (e.g. plasticity, damage) and physical or constitutive parameters of the material model. In the context of multiscale analysis as defined by computational homogenization, eq. (1.4) does not have an explicit functional form. Instead this relation is obtained through the solution of a boundary value problem imposed by a representative volume element (RVE) that characterizes the fine-scale structure.

1.2 Microscale boundary value problem

The fine scale can represent several possible length scales, however, without loss of generality, the fine scale formulation will be presented here in terms of a microscale problem. As shown in fig. 1.1, the elastic body μ, which embodies the RVE, is defined on a domain $\Omega^\mu \subset R^3$ with a Lipschitz-continuous boundary Γ^μ. The infinitesimal strain theory is also applied here:

$$\boldsymbol{\epsilon}^\mu = \frac{1}{2}(\nabla^\mu \boldsymbol{u}^\mu + (\nabla^\mu \boldsymbol{u}^\mu)^\mathsf{T}) \tag{1.5}$$

The strong formulation of the static equilibrium equation in the absense of volumetric forces is expressed as:

$$\nabla^\mu \boldsymbol{\sigma}^\mu = 0 \quad on \quad \Omega^\mu \tag{1.6}$$

while the boundary conditions are:

$$\boldsymbol{u}^\mu = \hat{\boldsymbol{u}}^\mu \quad on \quad \Gamma_D^m \equiv \Gamma^\mu \tag{1.7a}$$

12

where $\boldsymbol{\sigma}^\mu$ is the Cauchy stress tensor of the microscale and $\hat{\boldsymbol{u}}^\mu$ the prescribed displacements on the Dirichlet boundary Γ_D^m.

The constitutive relation for the microscale problem is written as:

$$\boldsymbol{\sigma}^\mu(t) = \boldsymbol{\sigma}^\mu\{\boldsymbol{\epsilon}^\mu(t), \boldsymbol{\theta}^\mu(\bar{t}), \bar{t} \in [0,t]\} \tag{1.8}$$

with $\boldsymbol{\theta}^\mu$ including the microscale internal variables and material parameters.
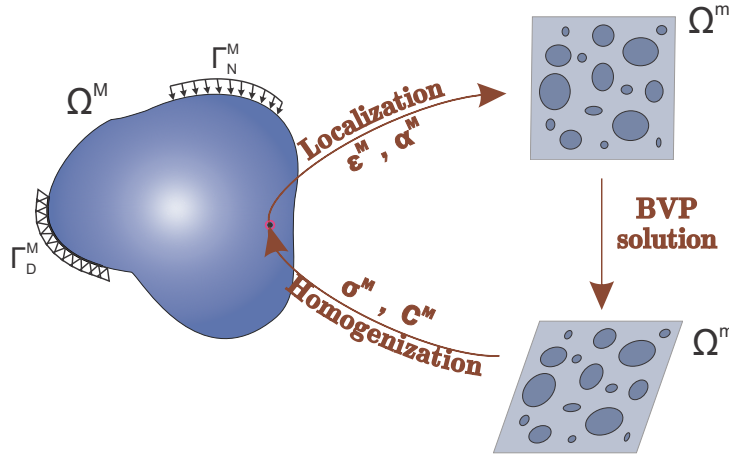
Figure 1.1: Scale transition during computational homogenization

1.3   Transition between the length scales

According to the principle of separation of scales, homogenization theory can be applied effectively if the following relation holds:

$$\lambda^\mu \ll \lambda^{RVE} \ll \lambda^M \tag{1.9}$$

where λ^μ, λ^{RVE}, λ^M are the length scales of the microscopic field fluctuations, of the RVE size and of the macroscopic field fluctuations, respectively.

To link the system defined by the two scales, boundary conditions that are consistent with the response of the macrostructure have to be assigned on the microscale. To this purpose, the displacement field $\boldsymbol{u}^\mu(\boldsymbol{x})$ is associated to the strain $\boldsymbol{\epsilon}^M$ of a macroscopic material point through the relation:

$$\boldsymbol{u}^{\mu} = \boldsymbol{\epsilon}^{M} \cdot \boldsymbol{x}^{\mu} + \tilde{\boldsymbol{u}}^{\mu} \tag{1.10}$$

where $\tilde{\boldsymbol{u}}^{\mu}$ is the displacement fluctuation field attributed to the heterogeneity of the microstructure.

Alternatively eq. (1.10) can be expressed as:

$$\boldsymbol{\epsilon}^{\mu} = \boldsymbol{\epsilon}^{M} + \nabla^{\mu} \tilde{\boldsymbol{u}}^{\mu} \tag{1.11}$$

Ultimately, the scale transition is established by utilizing averaging relations for the field variables of the two scale system. For the strain tensors this relation is written as:

$$\boldsymbol{\epsilon}^{M} = \frac{1}{\|\Omega^{\mu}\|} \int_{\Omega^{\mu}} \boldsymbol{\epsilon}^{\mu} d\Omega^{\mu} \tag{1.12}$$

Inserting eq. (1.11) into eq. (1.12) and after some algebraic manipulations yields:

$$\frac{1}{\|\Omega^{\mu}\|} \int_{\Gamma^{\mu}} \tilde{\boldsymbol{u}}^{\mu} \otimes \boldsymbol{n}^{\mu} d\Gamma^{\mu} = 0 \tag{1.13}$$

For the relation (1.13) to hold, appropriate boundary conditions have to be assigned on the fluctuation field $\tilde{\boldsymbol{u}}^{\mu}$. A choice that fulfills this requirement and was used in the present thesis is to restrain $\tilde{\boldsymbol{u}}^{\mu}$ on the boundary of the RVE as:

$$\tilde{\boldsymbol{u}}^{\mu} = \boldsymbol{0} \quad on \quad \boldsymbol{x}^{\mu} \in \Gamma^{\mu} \tag{1.14}$$

Applying eq. (1.14) into eq. (1.10) produces the final form of the localization i.e. the macro-to-micro scale transition:

$$\boldsymbol{u}^{\mu} = \boldsymbol{\epsilon}^{M} \boldsymbol{x}^{\mu} \quad on \quad \boldsymbol{x}^{\mu} \in \Gamma^{\mu} \tag{1.15}$$

It is evident that the above relation assigns a linear deformation on Γ^{μ} of the RVE. Other types of boundaries can be chosen, as well, such as uniform tractions or periodic boundary conditions [88].

In computational homogenization, the local variation of the macroscopic work is linked to the variation of the microscopic work per unit volume through the Hill-Mandell energy consistency relation [55]:

14

$$\boldsymbol{\sigma}^M : \delta \boldsymbol{\epsilon}^M = \frac{1}{\|\Omega^\mu\|} \int_{\Omega^\mu} \boldsymbol{\sigma}^\mu : \delta \boldsymbol{\epsilon}^\mu d\Omega^\mu \qquad (1.16)$$

Taking into account the microscopic equilibrium and then using the divergence theorem, eq. (1.16) can be restated as:

$$\boldsymbol{\sigma}^M : \delta \boldsymbol{\epsilon}^M = \frac{1}{\|\Omega^\mu\|} \int_{\Omega^\mu} \nabla^\mu \cdot (\boldsymbol{\sigma}^\mu \cdot \delta \boldsymbol{u}^\mu) d\Omega^\mu = \frac{1}{\|\Omega^\mu\|} \int_{\Gamma^\mu} \hat{\boldsymbol{\tau}}^\mu \cdot \delta \boldsymbol{u}^\mu d\Gamma^\mu \qquad (1.17)$$

where $\hat{\boldsymbol{\tau}}^\mu$ is the traction vector acting on the boundary Γ^μ.

Inserting eq. (1.10) into eq. (1.17) yields:

$$\boldsymbol{\sigma}^M : d\boldsymbol{\epsilon}^M = \frac{1}{\|\Omega^\mu\|} \left(\int_{\Gamma^\mu} \hat{\boldsymbol{\tau}}^\mu \otimes \boldsymbol{x}^\mu d\Gamma^\mu \right) : \delta \boldsymbol{\epsilon}^M + \frac{1}{\|\Omega^\mu\|} \int_{\Gamma^\mu} \hat{\boldsymbol{\tau}}^\mu \cdot \delta \tilde{\boldsymbol{u}}^\mu d\Gamma^\mu \qquad (1.18)$$

Based on eq. (1.18) and after applying the divergence theorem one more time leads to the micro-to-macro relation:

$$\boldsymbol{\sigma}^M = \frac{1}{\|\Omega^\mu\|} \int_{\Gamma^\mu} \hat{\boldsymbol{\tau}}^\mu \otimes \boldsymbol{x}^\mu d\Gamma^\mu = \frac{1}{\|\Omega^\mu\|} \int_{\Omega^\mu} \boldsymbol{\sigma}^\mu d\Omega^\mu \qquad (1.19)$$

where the stress field $\boldsymbol{\sigma}^\mu$ can be acquired after the solution of the boundary value problem (BVP) imposed by the localization rule of eq. (1.15) on the RVE.

Additionally, the tangent modulus that defines the constitutive law in eq. (1.4) at the particular macroscopic material point is:

$$\boldsymbol{C}^M = \frac{1}{\|\Omega^\mu\|} \partial_{\boldsymbol{\epsilon}^M} \int_{\Omega^\mu} \boldsymbol{\sigma}^\mu d\Omega^\mu \qquad (1.20)$$

The interaction of the macroscale and the microscale of the coupled two-scale system is illustrated in fig. 1.1.

1.4 N-scale boundary value problem

Consider a system defined by a macroscopic scale M as introduced in section 1.1 and a set of N successively finer length scales $s = 1, .., N$ (e.g. $s = 1 \rightarrow$ nanoscale, $s = 2 \rightarrow$ microscale, $s = 3 \rightarrow$ mesoscale). Similarly to section 1.2 each of the finer scales s is placed on a domain Ω^s where, by omitting the body forces, the balance equation is stated as:

15

$$\nabla \cdot \boldsymbol{\sigma}^s = 0 \quad on \quad \Omega^s \tag{1.21}$$

and the boundary conditions are as follows:

$$\boldsymbol{u}^s = \hat{\boldsymbol{u}}^s \quad on \quad \Gamma_D^s \equiv \Gamma^s \tag{1.22}$$

In a similar manner to section 1.3, a link between these scales has to be established during the solution. This can be achieved sequentially between each pair of consecutive scales. As in eq. (1.15) the localization rule is initially applied for each pair as:

$$\boldsymbol{\epsilon}^2(\boldsymbol{x}^2) = \frac{1}{\|\Omega^1\|} \int_{\Omega^1} \boldsymbol{\epsilon}^1(\boldsymbol{x}^1; \boldsymbol{x}^2) d\Omega^1 \ , \quad \cdots \quad , \ \boldsymbol{\epsilon}^M(\boldsymbol{x}^M) = \frac{1}{\|\Omega^N\|} \int_{\Omega^N} \boldsymbol{\epsilon}^N(\boldsymbol{x}^N; \boldsymbol{x}^M) d\Omega^N \tag{1.23}$$

To clarify the above notation, the variable after the semicolon, for instance $\boldsymbol{x}^2$ in $\boldsymbol{\epsilon}^1(\boldsymbol{x}^1; \boldsymbol{x}^2)$, denotes that the strain field $\boldsymbol{\epsilon}^1$ inside the integral is the one developed from the RVE that is associated to the material point $\boldsymbol{x}^2$ of the upper scale.

Then, combining all the successive localization steps, a relation that describes the whole system interaction can be acquired as:

$$\boldsymbol{\epsilon}^M(\boldsymbol{x}^M) = \frac{1}{\|\Omega^N\|} \int_{\Omega^N} \cdots \underbrace{\underbrace{\frac{1}{\|\Omega^1\|} \int_{\Omega^1} \boldsymbol{\epsilon}^1(\boldsymbol{x}^1; \boldsymbol{x}^2) d\Omega^1}_{\boldsymbol{\epsilon}^2(\boldsymbol{x}^2; \boldsymbol{x}^3)} \dots d\Omega^N}_{\underset{\boldsymbol{\epsilon}^N(\boldsymbol{x}^N; \boldsymbol{x}^M)}{\vdots}} \tag{1.24}$$

Next, starting from the finest scale, the solution of the RVEs can be obtained by applying the homogenization step of eq. (1.19) for each pair as:

$$\boldsymbol{\sigma}^2(\boldsymbol{x}^2) = \frac{1}{\|\Omega^1\|} \int_{\Omega^1} \boldsymbol{\sigma}^1(\boldsymbol{x}^1, \boldsymbol{\theta}^1; \boldsymbol{x}^2) d\Omega^1 \ , \quad \cdots \quad , \ \boldsymbol{\sigma}^M(\boldsymbol{x}^M) = \frac{1}{\|\Omega^N\|} \int_{\Omega^N} \boldsymbol{\sigma}^N(\boldsymbol{x}^N, \boldsymbol{\theta}^N; \boldsymbol{x}^M) d\Omega^N \tag{1.25}$$

where $\boldsymbol{\theta}^s$ are the sth scale internal variables and material parameters. The connection of the subsequent homogenization steps then gives:

16

$$\boldsymbol{\sigma}^M(\boldsymbol{x}^M) = \frac{1}{\|\Omega^N\|} \int_{\Omega^N} \cdots \underbrace{\underbrace{\frac{1}{\|\Omega^1\|} \int_{\Omega^1} \boldsymbol{\sigma}^1(\boldsymbol{x}^1, \boldsymbol{\theta}^1; \boldsymbol{x}^2) d\Omega^1}_{\boldsymbol{\sigma}^2(\boldsymbol{x}^2, \boldsymbol{\theta}^2; \boldsymbol{x}^3)} \cdots d\Omega^N}_{\boldsymbol{\sigma}^N(\boldsymbol{x}^N, \boldsymbol{\theta}^N; \boldsymbol{x}^M)} \tag{1.26}$$

Accordingly, the tangent constitutive matrix is formulated as:

$$\boldsymbol{C}^M(\boldsymbol{x}^M) = \partial_{\boldsymbol{\epsilon}^M} \frac{1}{\|\Omega^N\|} \int_{\Omega^N} \cdots \underbrace{\underbrace{\frac{1}{\|\Omega^1\|} \int_{\Omega^1} \boldsymbol{\sigma}^1(\boldsymbol{x}^1, \boldsymbol{\theta}^1; \boldsymbol{x}^2) d\Omega^1}_{\boldsymbol{\sigma}^2(\boldsymbol{x}^2, \boldsymbol{\theta}^2; \boldsymbol{x}^3)} \cdots d\Omega^N}_{\boldsymbol{\sigma}^N(\boldsymbol{x}^N, \boldsymbol{\theta}^N; \boldsymbol{x}^M)} \tag{1.27}$$

1.5 Finite element solution of the microscopic boundary value problem

The discrete version of the microscopic BVP problem can be obtained through the utilization of a numerical discretization technique such as the Finite Element Method (FEM), the Finite Difference Method (FDM), the Discrete Element Method (DEM) and others. In the context of FEM, which was applied in this thesis, the above equations can be recast in matrix form as follows. After discretizing the RVE, linear displacements are imposed on the boundary nodes according to the macroscopic strain $\bar{\boldsymbol{\epsilon}}$ (localization rule) with the following relation:

$$\boldsymbol{u}_n = \boldsymbol{D}_n \boldsymbol{\epsilon}^M \tag{1.28}$$

where $\boldsymbol{u}_n$ is the nodal displacement vector of node n and $\boldsymbol{D}_n$ is a nodal coordinate dependent matrix written in the 3D case as:

$$\boldsymbol{D}_n = \frac{1}{2} \begin{bmatrix} 2x_1 & 0 & 0 & x_2 & x_3 & 0 \\ 0 & 2x_2 & 0 & x_1 & 0 & x_3 \\ 0 & 0 & 2x_3 & 0 & x_1 & x_2 \end{bmatrix}_n \tag{1.29}$$

17

with x_1, x_2, x_3 being the nodal coordinates of node n. Then, the global set of equations for the problem becomes:

$$u_b = D\epsilon^M \qquad (1.30)$$

where the subscript b refers to the nodes on the boundary and $D = [D_1, D_2, ..., D_M]$ is the global coordinate matrix for all M nodes lying on the boundary.

The equilibrium of the RVE is imposed using Lagrange multipliers and the system of equations assumes the form:

$$u_b = D\epsilon^M \qquad (1.31\text{a})$$

$$f_b = \lambda \qquad (1.31\text{b})$$

$$f_\alpha = 0 \qquad (1.31\text{c})$$

where f denotes the internal force vector, while the indices α and b stand for the nodes at the interior of the RVE and the boundary, respectively. Additionally, the Lagrange multipliers λ express the external forces on the system's nodes. In incremental form and using matrix notation eqs. (1.31) can be restated as:

$$\begin{bmatrix} K_{\alpha\alpha} & K_{\alpha b} \\ K_{b\alpha} & K_{bb} \end{bmatrix} \begin{Bmatrix} \Delta u_\alpha \\ \Delta u_b \end{Bmatrix} + \begin{Bmatrix} f_\alpha \\ f_b \end{Bmatrix} = \begin{Bmatrix} 0 \\ \lambda + \Delta\lambda \end{Bmatrix} \qquad (1.32)$$

with K being the tangential stiffness matrix of the RVE.

The procedure to solve the nonlinear equation with a standard Newton-Raphson method is:

- For the first iteration of each increment where the microscopic equilibrium is satisfied, a displacement increment Δu_b is imposed as $\Delta u_b = D\Delta\epsilon^M$ resulting in a $\Delta u_\alpha = K_{\alpha\alpha}^{-1} K_{\alpha b} \Delta u_b$

- For the following iterations where $\Delta u_b = 0$ the internal deformations are updated as $\Delta u_\alpha = -K_{\alpha\alpha}^{-1} f_\alpha$ until convergence has been achieved in the sense that $\|f_\alpha\| = 0$

- If convergence has been reached the macroscopic stress and the macroscopic tangent modulus are calculated as:
  $\sigma^M = \frac{1}{\|\Omega^\mu\|} D f_b$, $C^M = \frac{1}{\|\Omega^\mu\|} D^T \tilde{K}_{bb} D$, where $\tilde{K}_{bb} = K_{bb} - K_{b\alpha} K_{\alpha\alpha}^{-1} K_{\alpha b}$

18

1.6 FINITE ELEMENT SOLUTION OF THE TWO-SCALE SYSTEM (FE^2 SCHEME)

So far the homogenization procedure has been presented in terms of a single macroscopic material point. The generalization of this procedure on a nested finite element scheme can be achieved by the well-established FE^2 algorithm. The main steps of this algorithm for a complete multiscale solution are summarized next:

1. Initialize macroscopic state variables ($\boldsymbol{F}_{ext}^M = 0, \boldsymbol{F}_{int}^M = 0, \boldsymbol{u}^M = 0, \Delta\boldsymbol{u}^M = 0$).

2. For each macroscopic incremental step

3. Apply macroscopic load increment $\Delta\boldsymbol{F}_{ext}^M$.

4. For each macroscopic integration point placed on x_M:

   - Define a unique discretized RVE on Ω^μ.

   - Assign boundary conditions on the RVE according to the localization rule of eq. (1.15)

   - Solve the microstructural problem and obtain $\boldsymbol{u}^\mu$ and $\boldsymbol{\sigma}^\mu$.

   - Apply the homogenization rule to acquire the homogenized stress $\boldsymbol{\sigma}^M$ from eq. (1.19) and homogenized tangent modulus C^M from eq. (1.20).

5. Calculate the macroscopic internal force vector $\boldsymbol{F}_{int}^M(\boldsymbol{\sigma}^M)$ and the tangent stiffness operator $\boldsymbol{K}^M(\boldsymbol{C}^M)$.

6. Solve the macrostructural problem and get $\Delta\boldsymbol{u}^M$.

7. Check the condition $\|\boldsymbol{F}_{ext}^M - \boldsymbol{F}_{int}^M\| \leq \epsilon$ (ϵ=prescribed tolerance):

   - If condition is met, then if it is the final macroscopic load increment $\Delta\boldsymbol{F}_{ext}^M$ exit, else return to 2.

   - If condition is not met return to 4.

The nonlinear computational homogenization solution procedure, as implemented by the FE^2 algorithm provided above has the ability to yield very reliable predictions even for complex composite materials with multiple phases and strong non-linearities. Nevertheless, the computational demands of such a framework render it prohibitive for non-trivial applications.

1.7 Finite element solution of the N-scale system (FE^N scheme)

Based on the concept of nonlinear computational homogenization in the context of a two-scale system as presented in section 1.6, for each integration point of the macroscale, a complete model solution of the micro scale has to be obtained and this has to be repeated for all the iterations of the nonlinear analysis [145]. In theory, this concept can be extended into a system of N length scales as a FE^N scheme, where the coarse scale of an initial FE^2 scheme represents the fine scale on the following FE^2. By applying this reasoning, an estimation of the total RVE solutions in the context of multiple nested scale systems can be acquired. In this case, the requirement for repeated solutions between any local two-scale system has to be added on top of the requirement for repeated solutions between each consecutive pair of scales that are located higher to the current pair. The total number of BVP solutions required for a complete solution of the whole system is then given by:

$$K_{RVE} = k^M + k^M \times k^N + \cdots + k^M \times k^N \times \cdots \times k^2 \qquad (1.33)$$

where for each scale s, k^s is defined as:

$$k^s = n_{int}^s \times n_{iter}^s \times n_{incr}^s \qquad (1.34)$$

with n_{int}^s being the number of integration points on the discretized system, n_{iter}^s the number of iterations of the Newton procedure and n_{incr}^s the number of increments of the analysis.

It is evident from eq. (1.33) that in order to perform the full solution of the macrostructure, the amount of BVPs that have to be solved is related to the total amount of RVEs that describe the multiscale system and it quickly blows up for increasing number of scales. As a consequence, this procedure becomes computationally intractable even for low-fidelity multiscale models.

<div style="text-align: right; color: orange; font-size: 3em;">**2**</div>

Modeling of CNT-reinforced composite systems

2.1 CARBON NANOTUBE FINITE ELEMENT REPRESENTATION

A single-walled carbon nanotube (SWCNT) can be conceptualized as a rolled-up graphene sheet, forming a hollow tube. On the other hand, a multi-walled carbon nanotube (MWCNT) consists of concentric cylinders made of graphitic layers with sealed ends, featuring a layer spacing of approximately 0.34 nm. Graphite takes the shape of a two-dimensional sheet with carbon atoms arranged in a hexagonal pattern, each connected to three nearest neighbors. Nanotubes possess an atomic arrangement characterized by their chirality, or helical nature, determined by parameters such as the chiral vector $\boldsymbol{C}_h$ and the chiral angle C_θ. The concept of slicing a graphite sheet along dotted lines and then rolling the resulting strip so that the chiral vector's tip meets its tail is illustrated in fig. 2.1. This chiral vector, alternatively termed the roll-up vector, is mathematically represented by the following equation:

$$\boldsymbol{C}_h = n\boldsymbol{a}_1 + m\boldsymbol{a}_2 \tag{2.1}$$

where the integers (n, m) represent the count of steps along the zigzag carbon bonds of

the hexagonal lattice, while $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ denote unit vectors. The chiral angle signifies the degree of twist in the tube. For the extremes known as zigzag and armchair, the chiral angles are 0 and 30 degrees, respectively. The zigzag nanotube is represented by the roll-up vector (n, 0), while the armchair nanotube is denoted by (n, n). Additionally, the roll-up vector of the nanotube determines its diameter.
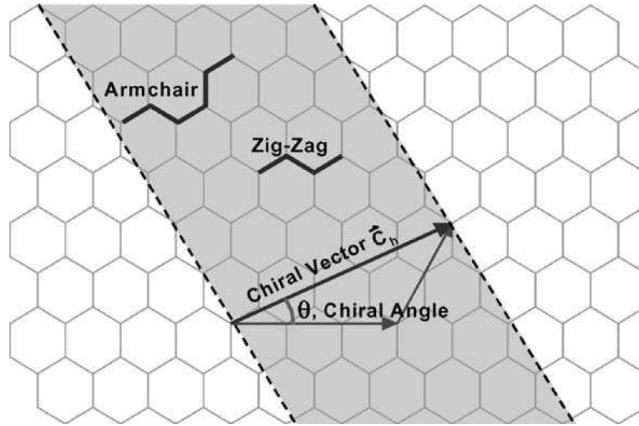
Figure 2.1: Derivation of a CNT from an initial graphene sheet

In light of the molecular structural mechanics approach (MSM) [75], CNTs can be viewed as space frames where the carbon atoms represent the nodes and the C-C bonds the structural elements (i.e. beams) that connect them. The force field developed in a CNT is attributed to the covalent bonds developed between the carbon atoms (C-C bond) and is expressed through a steric potential energy. The form of this energy, which is solely affected by the relative positions of the atoms, is expressed as a combination of energies arising due to the C-C bonding with the following equation:

$$U = \sum U_r + \sum U_\theta + \sum U_\phi \tag{2.2}$$

where U_r, U_θ, U_ϕ, are the bond stretch, the bond bending and torsional bending energies that are developed between each pair of carbon atoms. Each of these energy types is defined by the following forms:

$$U_r = \frac{1}{2}\kappa_r \Delta r^2 \quad , \quad with \quad \Delta r^2 = (r - r_0)^2 \tag{2.3a}$$

$$U_\theta = \frac{1}{2}\kappa_\omega \Delta \omega^2 \quad , \quad with \quad \Delta \omega^2 = (\omega - \omega_0)^2 \tag{2.3b}$$

$$U_\phi = \frac{1}{2}\kappa_\tau \Delta \tau^2 \quad , \quad with \quad \Delta \tau^2 = (\tau - \tau_0)^2 \tag{2.3c}$$

In this context, k_r, k_ω, and k_τ denote the force constants for bond stretching, bond angle bending, and torsional resistance, respectively. Additionally, Δr, $\Delta \omega$, and $\Delta \tau$, represent the increments in bond stretching, changes in bond angle, and alterations in bond twisting angle, respectively.

On the other hand, according to the theory of classical structural mechanics, the strain energies of a uniform beam of length L and double symmetric cross-section that is subjected to pure axial force, pure bending and pure torsion respectively, assume the following forms:

$$U_a = \frac{EA}{2L^B}\Delta L^2 \quad , \quad with \quad \Delta L^2 = (L^B - L_0^B)^2 \tag{2.4a}$$

$$U_m = \frac{2EI}{L^B}\Delta \alpha^2 \quad , \quad with \quad \Delta \alpha^2 = (\alpha - \alpha_0)^2 \tag{2.4b}$$

$$U_t = \frac{GJ}{2L^B}\Delta \beta^2 \quad , \quad with \quad \Delta \beta^2 = (\beta - \beta_0)^2 \tag{2.4c}$$

with E being the elastic modulus, L^B is the beam length, while A, I and J being the cross-section area, bending moment of inertia and torsional moment of inertia respectively. The terms ΔL, $\Delta \alpha$ and $\Delta \beta$ declare the axial stretching deformation, the bending angle shift and the torsional angle difference.

The relation between the potential energy of the force field of C-C bonds and the mechanical properties of the equivalent structural beam element can then be derived from the equations:

$$(EA)_{eq} = \kappa_r L^B \tag{2.5a}$$

$$(EI)_{eq} = \kappa_\omega L^B \tag{2.5b}$$

$$(GJ)_{eq} = \kappa_\tau L^B \tag{2.5c}$$

Although this structural representation is a straightforward and accurate transition from the molecular mechanics, the complexity of the system persists, since each space frame that simulates a CNT is inherently defined by an enormous amount of Degrees Of Freedom (DOFs). To be able to efficiently incorporate the CNTs in the microstructure modelling, a further simplification is necessitated. This is performed in a similar manner to the MSM approach, as shown in fig. 2.2, with the projection of each space frame structure onto an equivalent beam element(EBE) [100] with the relations:

$$(EA)_{eq}^{EBE} = \frac{F_x L^{SF}}{u_x} \tag{2.6a}$$

$$(EI)_{eq}^{EBE} = \frac{F_y (L^{SF})^3}{3u_y} \tag{2.6b}$$

$$(GJ)_{eq}^{EBE} = \frac{T L^{SF}}{\varphi} \tag{2.6c}$$

where the axial stiffness $(EA)_{eq}^{EBE}$, the bending rigidity $(EI)_{eq}^{EBE}$ and the torsional rigidity $(GJ)_{eq}^{EBE}$ of the EBE, are obtained by measuring the horizontal displacement u_x, the vertical displacement u_y and the angle of rotation φ that emerge from the imposed loads on the space frame of length L^{SF}.
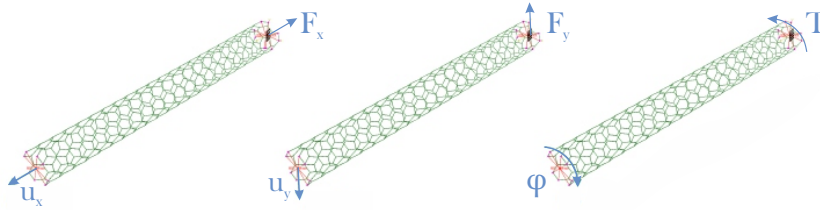
Figure 2.2: CNT as space frame subjected to three deformation cases

2.2 Fully-bonded formulation of the CNT/matrix interface

To simulate the interaction between the CNTs and the surrounding matrix, the embedding finite element technique is employed. With this approach at each of the EBEs (3-D beam finite elements) nodes k the respective DOFs $\boldsymbol{v}^k = [v_1^k, ..., v_6^k]$ are embedded in the DOFs $\boldsymbol{u}^s = [u_1^s, ..., u_{24}^s]$ of the surrounding solid element s (3-D hexahedral finite element) with the relation:

$$\boldsymbol{v}^k = \boldsymbol{T}^{k,s}\boldsymbol{u}^s \tag{2.7}$$

where the [6x24] transformation matrix $\boldsymbol{T}^{k,s}$ is formulated as:

$$\boldsymbol{T}^{k,s} = \begin{bmatrix} \boldsymbol{T}_1^{k,s} & ... & \boldsymbol{T}_8^{k,s} \end{bmatrix} \tag{2.8}$$

with:

$$\boldsymbol{T}_i^{k,s} = \begin{bmatrix} N_i^s(\boldsymbol{x}^k) & 0 & 0 \\ 0 & N_i^s(\boldsymbol{x}^k) & 0 \\ 0 & 0 & N_i^s(\boldsymbol{x}^k) \\ 0 & -\frac{1}{2}N_{i,z}^s(\boldsymbol{x}^k) & \frac{1}{2}N_{i,y}^s(\boldsymbol{x}^k) \\ \frac{1}{2}N_{i,z}^s(\boldsymbol{x}^k) & 0 & -\frac{1}{2}N_{i,x}^s(\boldsymbol{x}^k) \\ -\frac{1}{2}N_{i,y}^s(\boldsymbol{x}^k) & \frac{1}{2}N_{i,x}^s(\boldsymbol{x}^k) & 0 \end{bmatrix} \tag{2.9}$$

where N_i^s, $N_{i,x}^s$, $N_{i,y}^s$, $N_{i,z}^s$ are the shape functions and their derivatives on x, y, z of the node i of each solid element s, while x^k are the local coordinates of the EBE node k indicating its relevant position with respect to the host solid element. In the case of a two-node EBE ($k = 1, 2$), eq. (2.7) can be written as:

$$\boldsymbol{v} = \begin{Bmatrix} \boldsymbol{v}^1 \\ \boldsymbol{v}^2 \end{Bmatrix} = \begin{bmatrix} \boldsymbol{T}^{1,s_1} & 0 \\ 0 & \boldsymbol{T}^{2,s_2} \end{bmatrix} \begin{Bmatrix} \boldsymbol{u}^{s_1} \\ \boldsymbol{u}^{s_2} \end{Bmatrix} = \boldsymbol{T}\boldsymbol{u} \tag{2.10}$$

where s_1 and s_2 are the solid elements in which the respective EBE node k lies.

After applying the transformation $\boldsymbol{T}$, the beam's stiffness matrix is expressed as:

$$\boldsymbol{K}_b' = \boldsymbol{T}^T \boldsymbol{K}_b \boldsymbol{T} \tag{2.11}$$

The complete stiffness matrix of the whole system's interaction is:

$$\boldsymbol{K} = \boldsymbol{K}_s + \sum_{i=1}^{N} \boldsymbol{K}_{b_i}' \tag{2.12}$$

where $\boldsymbol{K}_s$ is the solid stiffness matrix, while $\boldsymbol{K}_{b_i}'$ is the stiffness matrix of the ith EBE for a system of N total EBEs.

2.3 Cohesive formulation of the CNT/matrix interface

For every EBE it is assumed that an auxiliary cohesive beam element (CBE) exists that is parallel and attached to the first, as shown in fig. 2.3. This element is characterized in terms of the host matrix DOFs with an embedding technique. The two elements are linked together with a cohesive zone formulation [4, 28]. The methodology with the FEM approach for the 3D case is described in detail below.
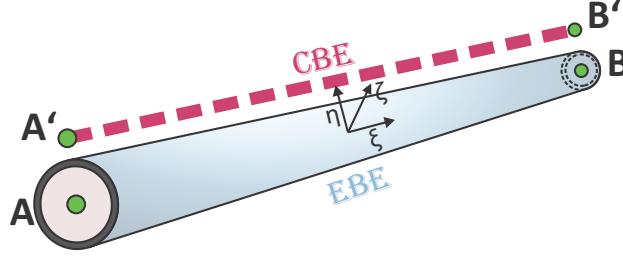
Figure 2.3: EBE with its respective CBE in the global coordinate system

The cohesive zone model relates the DOFs of the lower surface of a predetermined discontinuity with its upper surface. Here, the lower part is the EBE, while the upper one is the CBE. The cohesive stiffness matrix represents the stiffness of the interaction between these two parts and in isoparametric formulation can be expressed as:

$$\boldsymbol{K}_{coh} = \begin{bmatrix} \int_{-1}^{1} \boldsymbol{M} d\xi & -\int_{-1}^{1} \boldsymbol{M} d\xi \\ -\int_{-1}^{1} \boldsymbol{M} d\xi & \int_{-1}^{1} \boldsymbol{M} d\xi \end{bmatrix} \tag{2.13}$$

where ξ indicates the longitudinal direction of the EBE/CBE pair, while $\boldsymbol{M}$ is a [12x12] matrix given as:

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{M}_{aa} & 0 & \boldsymbol{M}_{ab} & 0 \\ 0 & \boldsymbol{M}_r & 0 & 0 \\ \boldsymbol{M}_{ba} & 0 & \boldsymbol{M}_{bb} & 0 \\ 0 & 0 & 0 & \boldsymbol{M}_r \end{bmatrix} \tag{2.14}$$

The [3x3] matrices $\boldsymbol{M}_{aa}$, $\boldsymbol{M}_{ab}$, $\boldsymbol{M}_{ba}$, $\boldsymbol{M}_{bb}$ describe the translational DOFs and are parts of a matrix $\boldsymbol{M}_t$ formulated as:

$$M_t = \begin{bmatrix} M_{aa} & M_{ab} \\ M_{ba} & M_{bb} \end{bmatrix} = N_b^T R_m D_{tan} R_m^T N_b \tag{2.15}$$

where N_b is a [3x9] matrix that contains the beam element shape functions:

$$N_b = \begin{bmatrix} N_1^b & 0 & 0 & N_2^b & 0 & 0 & N_3^b & 0 & 0 \\ 0 & N_1^b & 0 & 0 & N_2^b & 0 & 0 & N_3^b & 0 \\ 0 & 0 & N_1^b & 0 & 0 & N_2^b & 0 & 0 & N_3^b \end{bmatrix} \tag{2.16}$$

while R_m is the [3x3] rotation matrix from the local to the global coordinate system and D_{tan} is the constitutive matrix of the interfacial law. On the other hand, the [3x3] matrix M_r defines the rotational DOFs of the beam and its expressed as:

$$M_r = R_m^T C_r R_m \tag{2.17}$$

where C_r describes, in terms of a penalty stiffness factor, the relation of the rotational DOFs of the EBE with those of the CBE and is written as:

$$C_r = \begin{bmatrix} c_1 & 0 & 0 \\ 0 & c_2 & 0 \\ 0 & 0 & c_3 \end{bmatrix} \tag{2.18}$$

with each diagonal element taking values on the range $0 \leq c_i \leq 1$, in agreement with the rotational dependence of the two elements.

After applying numerical integration, eq. (2.13) can be rewritten as:

$$K_{coh} = \frac{L}{2} \sum_{i=1}^n w_i \begin{bmatrix} M(\xi_i) & -M(\xi_i) \\ -M(\xi_i) & M(\xi_i) \end{bmatrix} = \begin{bmatrix} K_{coh}^{11} & K_{coh}^{12} \\ K_{coh}^{21} & K_{coh}^{22} \end{bmatrix} \tag{2.19}$$

where L is the length of the EBE, while ξ_i and w_i are the points and weights according to the integration scheme.

Accordingly, the internal force vector of the cohesive zone is:

$$f_{coh} = \begin{bmatrix} \int_{-1}^1 f_{\text{CBE}} d\xi \\ -\int_{-1}^1 f_{\text{EBE}} d\xi \end{bmatrix} = \sum_{i=1}^n \begin{bmatrix} w_i N_b^T(\xi_i) R_m \tau \\ -w_i N_b^T(\xi_i) R_m \tau \end{bmatrix} = \begin{bmatrix} f_{coh}^1 \\ f_{coh}^2 \end{bmatrix} \tag{2.20}$$

where τ is the interfacial stress corresponding to the relative displacement between the EBE and CBE.

Having described the cohesive zone FEM equations, next the DOFs $\boldsymbol{v}^{k'} = [v_1^{k'}, ..., v_6^{k'}]$ of each node k' of the CBE, which correspond to the DOFs $\boldsymbol{v}^k = [v_1^k, ..., v_6^k]$ of their respective EBE, are embedded in the DOFs $\boldsymbol{u}^s = [u_1^s, ..., u_{24}^s]$ of the surrounding solid element s with the relation:

$$\boldsymbol{v}^{k'} = \boldsymbol{T}^{k',s} \boldsymbol{u}^s \tag{2.21}$$

where the [6x24] transformation matrix $\boldsymbol{T}^{k',s}$ is formulated as:

$$\boldsymbol{T}^{k',s} = \begin{bmatrix} \boldsymbol{T}_1^{k',s} & ... & \boldsymbol{T}_8^{k',s} \end{bmatrix} \tag{2.22}$$

with:

$$\boldsymbol{T}_i^{k',s} = \begin{bmatrix} N_i^s(\boldsymbol{x}^{k'}) & 0 & 0 \\ 0 & N_i^s(\boldsymbol{x}^{k'}) & 0 \\ 0 & 0 & N_i^s(\boldsymbol{x}^{k'}) \\ 0 & -\frac{1}{2}N_{i,z}^s(\boldsymbol{x}^{k'}) & \frac{1}{2}N_{i,y}^s(\boldsymbol{x}^{k'}) \\ \frac{1}{2}N_{i,z}^s(\boldsymbol{x}^{k'}) & 0 & -\frac{1}{2}N_{i,x}^s(\boldsymbol{x}^{k'}) \\ -\frac{1}{2}N_{i,y}^s(\boldsymbol{x}^{k'}) & \frac{1}{2}N_{i,x}^s(\boldsymbol{x}^{k'}) & 0 \end{bmatrix} \tag{2.23}$$

where N_i^s, $N_{i,x}^s$, $N_{i,y}^s$, $N_{i,z}^s$ are the shape functions and their derivatives on x, y, z of the node i of each solid element s, while $x^{k'}$ are the local coordinates of the CBE node k' indicating its relevant position with respect to the host solid element. In the case of a two-nodded CBE, eq. (2.21) can be written as:

$$\boldsymbol{v'} = \begin{Bmatrix} \boldsymbol{v}^{A'} \\ \boldsymbol{v}^{B'} \end{Bmatrix} = \begin{bmatrix} \boldsymbol{T}^{A',S_A} & 0 \\ 0 & \boldsymbol{T}^{B',S_B} \end{bmatrix} \begin{Bmatrix} \boldsymbol{u}^{S_A} \\ \boldsymbol{u}^{S_B} \end{Bmatrix} = \boldsymbol{Tu} \tag{2.24}$$

where $\boldsymbol{v}^{A'}$, $\boldsymbol{v}^{B'}$ and $\boldsymbol{T}^{A',S_A}$, $\boldsymbol{T}^{B',S_B}$ are the $\boldsymbol{v}^{k'}$ and $\boldsymbol{T}^{k',s}$ of the starting and ending node of the CBE, while $\boldsymbol{u}^{S_A}$ and $\boldsymbol{u}^{S_B}$ are the displacements $\boldsymbol{u}^s$ of the corresponding solid element in which each CBE node lies. The above formulation is illustrated in fig. 2.4 for the starting node A of an EBE/CBE pair, while for the ending node B the formulation is similar.

Applying the matrix $\boldsymbol{T}$ in eq. (2.19) the stiffness matrix takes the final form:

$$\boldsymbol{K}_{coh}^* = \begin{bmatrix} \boldsymbol{T}^T & 0 \\ 0 & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{K}_{coh}^{11} & \boldsymbol{K}_{coh}^{12} \\ \boldsymbol{K}_{coh}^{21} & \boldsymbol{K}_{coh}^{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{T} & 0 \\ 0 & \boldsymbol{I} \end{bmatrix} = \begin{bmatrix} \boldsymbol{K}_{coh}^{*11} & \boldsymbol{K}_{coh}^{*12} \\ \boldsymbol{K}_{coh}^{*21} & \boldsymbol{K}_{coh}^{*22} \end{bmatrix} \tag{2.25}$$

28

The complete stiffness matrix of the whole system's interaction is:

$$K = \begin{bmatrix} K_s + K_{coh}^{*11} & K_{coh}^{*12} \\ K_{coh}^{*21} & K_b + K_{coh}^{*22} \end{bmatrix} \tag{2.26}$$

where K_s and K_b are the solid and EBE stiffness matrices respectively.

In a similar way, the internal force vector in eq. (2.20) can be expressed as:

$$f = \begin{bmatrix} f_s + T^T f_{coh}^1 \\ f_b + f_{coh}^2 \end{bmatrix} \tag{2.27}$$
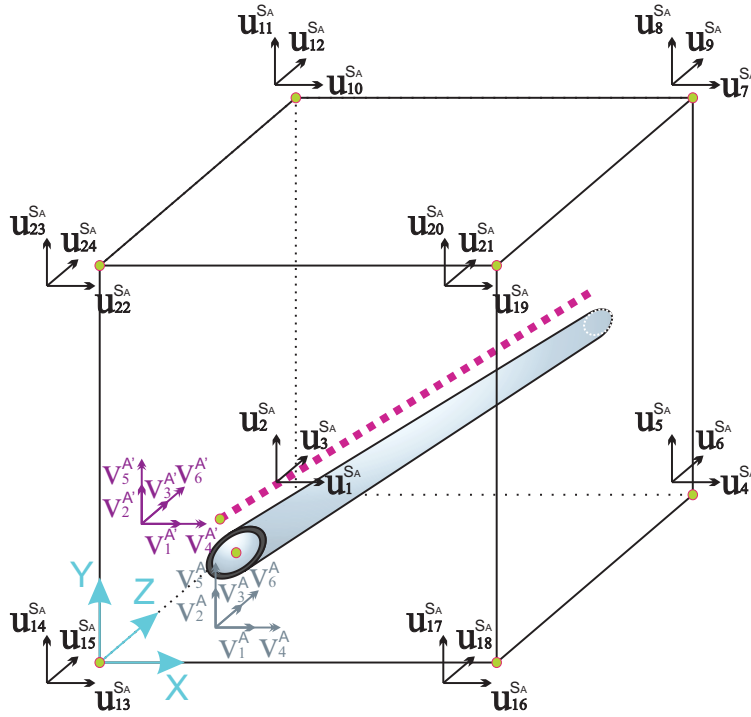
Figure 2.4: Embedding of a beam element into the host solids

The definition of the interfacial constitutive equations between each EBE and the surrounding matrix is usually done through the choice of specialized constitutive laws such as bond-slip or traction-separation relations.

3

Uncertainty quantification and optimization

3.1 Probability theory

Consider an experiment where the outcome is uncertain. The set of possible outcomes that can emerge each time the experiment is executed is called sample space Ω. As a structural mechanics example consider a concrete specimen where the compressive yield stress is not constant each time a uniaxial compression test is performed due to randomness in the material properties. In this case Ω contains all the possible values of the yield stress that can emerge after the completion of the test. The set of all measurable subsets included in Ω constitutes the σ-algebra $\mathscr{F}$. For each subset in $\mathscr{F}$ we can assign a measure according to a function $\mu : \mathscr{F} \rightarrow \mathbb{R}$. In probability theory this measure declares the probability P of each event happening and is defined as $P : \mathscr{F} \rightarrow [0, 1]$. The probability of the union of all possible outcomes, namely the sample space Ω, equals always unity i.e. $P_\Omega = 1$. Going back to the concrete example, for each possible yield stress value a probability can be assigned based on experimental results or computer simulations. The triplet of $(\Omega, \mathscr{F}, P)$ is called a probability space and can be used to model non-deterministic features in real-world applications.

A random variable is a mapping from a sample space Ω, which defined over a probability

space $(\Omega, \mathscr{F}, P)$, to a measurable space, most commonly the $\mathbb{R}$ space. It is expressed as:

$$X : \Omega \to \mathbb{R} \tag{3.1}$$

Depending on whether the image of a random variable is countable or continuous it is called discrete or continuous random variable respectively. For a continuous random variable, the probability that it will take values between a lower bound x_l and an upper bound x_u is expressed though the probability density function (PDF). This is defined as:

$$P(x_l \leq X \leq x_u) = \int_{x_l}^{x_u} f(x)dx \ , \ \ where \ \ \int_{-\infty}^{\infty} f(x)dx = 1 \tag{3.2}$$

The probability that a random variable will be equal or less to an upper value x_u is described by the cumulative distribution function (CDF), postulated as:

$$F(x) = \int_{-\infty}^{x_u} f(x)dx \tag{3.3}$$

A collection of random variables $X_1, ..., X_d$ is called a random vector $\boldsymbol{X} = [X_1, ..., X_d] \in \mathbb{R}^d$. The expressions for the CDF and PDF are straightforwardly generalized in the random vector case. The joint PDF for a lower bound vector $\boldsymbol{x}_l = [x_{l,1}, ..., x_{l,d}]$ and an upper bound vector $\boldsymbol{x}_u = [x_{u,1}, ..., x_{u,d}]$ has the form:

$$P(\boldsymbol{x}_l \leq \boldsymbol{x} \leq \boldsymbol{x}_u) = \int_{\boldsymbol{x}_l}^{\boldsymbol{x}_u} f(\boldsymbol{x})d\boldsymbol{x} \ , \ \ where \ \ \int_{-\infty}^{\infty} f(\boldsymbol{x})d\boldsymbol{x} = 1 \tag{3.4}$$

with $\boldsymbol{x} = [x_1, ..., x_d]$. The respective joint CDF for an upper vector $\boldsymbol{x}_u = [x_{u,1}, ..., x_{u,d}]$ is:

$$F(\boldsymbol{x}) = \int_{-\infty}^{\boldsymbol{x}_u} f(\boldsymbol{x})d\boldsymbol{x} \tag{3.5}$$

When the interest lies on a subset of the random vector $\boldsymbol{X}_s \subseteq \boldsymbol{X}$, the rest random variables $\boldsymbol{X}_{\sim s}$ can be integrated out. This is called the marginalization operation and the marginal PDF is obtained as:

$$f_s(\boldsymbol{x}_s) = \int_{\boldsymbol{x}_{\sim s}} f(\boldsymbol{x})d\boldsymbol{x} \tag{3.6}$$

The conditional probability of a random variable $X_i, i \in [1, ..., d]$ given the other random variables $\boldsymbol{X}_{\sim i}$ is defined as:

31

$$P(X_i|\boldsymbol{X}_{\sim i}) = \frac{P(X_i, \boldsymbol{X}_{\sim i})}{P(\boldsymbol{X}_{\sim i})} \tag{3.7}$$

When the random variables of a random vector are conditionally independent, meaning that the knowledge of one's value does not provide any knowledge about the values of the other random variables, the joint PDF can be rewritten as:

$$f(\boldsymbol{x}) = f(x_1, ..., x_d) = f_{X_1}(x_1)...fX_d((x_d) \tag{3.8}$$

For each random variable we can define several metrics that quantify different aspects of its probabilistic form. These metrics are expressed as linear operators applied on the PDF of the random variable X. The most common metrics are the expectation or mean value μ which is the first order moment and the variance σ^2 which is the second order central moment. These are expressed as:

$$E[X] = \mu = \int_{-\infty}^{\infty} x f(x) dx \ \ and \ \ E[(X - \mu)^2] = \sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx \tag{3.9}$$

Moments and central moments of different order can be defined accordingly as:

$$E[X^n] = \int_{-\infty}^{\infty} x^n f(x) dx \ \ and \ \ E[(X - \mu)^n] = \int_{-\infty}^{\infty} (x - \mu)^n f(x) dx \ \ , \ \ n = 1, ... \tag{3.10}$$

An additional useful metric is the coefficient of variation which is postulated as:

$$CoV = \frac{\sigma}{\mu} \tag{3.11}$$

where the term σ, namely the squared root of the variance, is called standard deviation.

The joint variability between two random variables X_1 and X_2 is quantified through the covariance as:

$$E[X_1, X_2] = E[X_1 X_2] - E[X_1]E[X_2] = \int_{-\infty}^{\infty} (x - \mu_1)(x - \mu_2) f(x_1, x_2) dx_1 dx_2 \tag{3.12}$$

3.2 Single-level Bayesian Inference

An effective technique for combining newly emerged observations with preexisting models is based on the Bayes theorem and it is called Bayesian Update or Bayesian Inference. With this technique, prior probabilistic information about the uncertain parameters are updated according to previously unseen data of the mechanical system's response. In view of the Bayesian paradigm the problem is formulated in a probabilistic framework. The relation between the measured data and the model predictions can be quantified through different approaches.

Consider a model $\mathcal{M}(\boldsymbol{\theta})$ that is described by a series of parameters $\boldsymbol{\theta} = [\theta_1, ..., \theta_N]$ which are assumed to be random variables. For instance the model $\mathcal{M}$ could be a computational model of a structural assembly, while the parameters $\boldsymbol{\theta}$ could define a series of material or geometric properties of $\mathcal{M}$. Additionally, consider a series of experimental measurements or observations $\boldsymbol{D}$, with $\boldsymbol{D} = [d_1, ..., d_m]$, that have been acquired from experiments on the physical replica of the computational model $\mathcal{M}$. The most used approach to relate the model predictions with the data is formulated with the introduction of an additive error:

$$\boldsymbol{y}(\boldsymbol{\theta}) = \boldsymbol{m}(\boldsymbol{\theta}) + \boldsymbol{\epsilon} \tag{3.13}$$

where $\boldsymbol{\epsilon}$ denotes the error term most commonly expressed as a random variable with a Gaussian probability density function (PDF) $\sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$. The autocovariance tensor $\boldsymbol{\Sigma}$ encapsulates the magnitude of the $\mathcal{M}$ model prediction error and errors of the experimental observations $\boldsymbol{D}$. The term $\boldsymbol{y}^i$ represents the output of the stochastic forward model $\mathcal{Y}$ according to which the data $\boldsymbol{D}$ are realized.

After relating the data with the model predictions through $P(\boldsymbol{D}|\boldsymbol{\theta})$, the Bayes' theorem can be readily applied to update the probabilistic form of the investigated parameters, as:

$$P(\boldsymbol{\theta}|\boldsymbol{D}) = \frac{P(\boldsymbol{D}|\boldsymbol{\theta})P(\boldsymbol{\theta})}{P(\boldsymbol{D})} = \frac{\pi(\boldsymbol{D}|\boldsymbol{\theta})P(\boldsymbol{\theta})}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty}\pi(\boldsymbol{D}|\boldsymbol{\theta})P(\boldsymbol{\theta})d\theta_1 d\theta_2...d\theta_n} \tag{3.14}$$

where $P(\boldsymbol{\theta})$ is the prior distribution i.e. the previous beliefs regarding the probabilistic nature of the parameters. In the case that no prior assumptions can be made, non-informative priors, such as the uniform distribution are most commonly applied. The

term $P(\boldsymbol{D})$ is called evidence and it is essentially a multidimensional integral, since it involves the marginalization operation over the parametric space, resulting in an analytically intractable solution in most cases. The term $\pi(\boldsymbol{D}|\boldsymbol{\theta})$ denotes the likelihood of $\boldsymbol{D}$ for given values of $\boldsymbol{\theta}$, which essentially quantifies the similarity between the data and the model predictions in a stochastic manner. In eq. (3.14), direct evaluation of the n-fold integral in the denominator is not feasible in the general case.

The typical assumption that the data $\boldsymbol{D}$ are uncorrelated is made in most practical applications. Under this condition, eq. (3.14) is restated as:

$$P(\boldsymbol{\theta}|\boldsymbol{D}) = \frac{\prod_{i=1}^{K} \prod_{j=1}^{N_i} [P(d^{i,j}|\boldsymbol{\theta})] P(\boldsymbol{\theta})}{P(\boldsymbol{D})} \tag{3.15}$$

where $\pi(\boldsymbol{D}|\boldsymbol{\theta})$ denotes the likelihood of $\boldsymbol{D}$ for given values of $\boldsymbol{\theta}$. In this equation, direct evaluation of the n-fold integral in the denominator is not feasible in the general case. This fact has motivated the use of Markov Chain Monte Carlo (MCMC) sampling techniques, which allow direct sampling from the posterior distribution without the need to solve the high-dimensional integral in (3.14), based on the fact that:

$$P(\boldsymbol{\theta}|\boldsymbol{D}) \propto P^{uns}(\boldsymbol{\theta}|\boldsymbol{D}) \tag{3.16}$$

with $P^{uns}(\boldsymbol{\theta}|\boldsymbol{D}) = \pi(\boldsymbol{D}|\boldsymbol{\theta}) P(\boldsymbol{\theta})$ being the unscaled probabilistic model after taking into consideration the new data. Representative algorithms based on the MCMC paradigm are the Metropolis-Hastings algorithm and the Transitional Markov Chain Monte Carlo (TMCMC) algorithm.

The Metropolis-Hastings (MH) algorithm [54], while being one of the oldest, it is the most well-established algorithm in the realm of MCMC methods. The implementation steps of the MH algorithm are summarized below:

1. Start with an initial value of parameters $\boldsymbol{\theta}_0$

2. Select an arbitrary candidate density function $g(\cdot|\cdot)$. The purpose of this function is to propose the next candidate sample at each MCMC step. In this notation, the first placeholder refers to the variables of the density function and the second to its parameters. A typical choice is a Gaussian distribution with the previously accepted sample as mean value and a pre-selected standard deviation.

3. Select a burn-in period N_{burn}. During this period the samples acquired from the

MH algorithm are discarded. This is important in order to ensure that the Markov Chain has converged to its stationary distribution, meaning that the samples are drawn from the posterior distribution.

4. Set i=1

5. While $i \leq N_{burn} + N_{samples}$ ($N_{samples}$ being the required number of samples)

   (a) Generate a candidate sample $\boldsymbol{\theta}'$ from $g(\boldsymbol{\theta}'|\boldsymbol{\theta}_{i-1})$

   (b) Evaluate the model response $\boldsymbol{M}(\boldsymbol{\theta}')$ and calculate the likelihood $\pi(\boldsymbol{\omega}|\boldsymbol{\theta}')$

   (c) Calculate the acceptance probability $\alpha(\boldsymbol{\theta}_{i-1}, \boldsymbol{\theta}') = min \left[1, \frac{\pi_{post}^{uns}(\boldsymbol{\theta}'|\boldsymbol{\omega})g(\boldsymbol{\theta}'|\boldsymbol{\theta}_{i-1})}{\pi_{post}^{uns}(\boldsymbol{\theta}_{i-1}|\boldsymbol{\omega})g(\boldsymbol{\theta}_{i-1}|\boldsymbol{\theta}')} \right]$

   (d) Draw v from the uniform distribution $\mathcal{U}(0, 1)$

   (e) If $v < \alpha(\boldsymbol{\theta}_{i-1}, \boldsymbol{\theta}')$ then let $\boldsymbol{\theta}_i = \boldsymbol{\theta}'$ and set $i = i + 1$, else let $\boldsymbol{\theta}_i = \boldsymbol{\theta}_{i-1}$

6. Discard the first N_{burn} samples.

An alternative sampling method is the TMCMC algorithm, which aims at tackling the difficulties that arise during sampling with the Metropolis-Hastings algorithm from more complicated posterior distributions $P(\boldsymbol{\theta}|\boldsymbol{D})$ (e.g. multimodal, very peaked, very flat). The TMCMC algorithm introduced in [18] proposes the construction and the sequential sampling from a number of intermediate PDFs of the form:

$$P_j(\boldsymbol{\theta}) \simeq P(\boldsymbol{D}|\boldsymbol{\theta})^{q_j} P(\boldsymbol{\theta}) \ , \quad j = 0, ..., m \ \ with \ \ 0 = q_0 < q_1 < ... < q_m = 1 \qquad (3.17)$$

The main steps of the algorithmic procedure are:

TMCMC parameters: prior distribution $P(\boldsymbol{\theta})$, likelihood function $P(\boldsymbol{D}|\boldsymbol{\theta})$, $N_s = $ number of samples at each level, $k = $ coefficient of variation for $w_j = P(\boldsymbol{D}|\boldsymbol{\theta}_{j-1})^{q_j - q_{j-1}}$, $\beta = $ scaling factor of the proposal distribution

For $j = 0$, N_s samples $\boldsymbol{\theta}_{0,1}, ..., \boldsymbol{\theta}_{0,N_s}$ are drawn from the prior distribution $P(\boldsymbol{\theta})$. Subsequently, for all $j > 0$:

1. Find q_j by solving $q_j = argmin_q(|CoV_j(q) - k|)$, where $CoV_j(q)$ is the sample coefficient of variation of the set $\{P(\boldsymbol{D}|\boldsymbol{\theta}_{j-1,k})^{q_j - q_{j-1}}\}_{k=1}^{N_s}$.

2. Calculate the weighting coefficients $w_{j,k} = P(\boldsymbol{D}|\boldsymbol{\theta}_{j-1,k})^{q_j - q_{j-1}}$.

3. Compute the covariance matrix of the proposal distribution $\boldsymbol{\Sigma}_j = \beta^2 \sum_{k=1}^{N_s} \left[w_{j,k} (\boldsymbol{\theta}_{j-1,k} - \frac{\sum_{l=1}^{N_s} w_{j,l}\boldsymbol{\theta}_{j-1,l}}{\sum_{l=1}^{N_s} w_{j,l}}) (\boldsymbol{\theta}_{j-1,k} - \frac{\sum_{l=1}^{N_s} w_{j,l}\boldsymbol{\theta}_{j-1,l}}{\sum_{l=1}^{N_s} w_{j,l}})^T \right]$.

4. For each l in $1, ..., N_s$ set $\boldsymbol{\theta}_{j,l}^c = \boldsymbol{\theta}_{j-1,l}$, where the superscript c denotes the candidate sample. For $k = 1, ..., N_s$, do the following:

   - Select index l from the set $1, ..., N_s$ with probability $\frac{w_{j,l}}{\sum_{n=1}^{N_s} w_{j,n}}$.

   - Propose sample $\boldsymbol{\theta}^c$ from the normal distribution $\mathcal{N}(\boldsymbol{\theta}_{j,l}^c, \boldsymbol{\Sigma}_j)$.

   - Generate sample u from the uniform distribution $\mathcal{U}(0,1)$.

   - If $u \leq \frac{P_j(\boldsymbol{\theta}^c)}{P_j(\boldsymbol{\theta}_{j,l}^c)}$, set $\boldsymbol{\theta}_{j,k} = \boldsymbol{\theta}^c$ and $\boldsymbol{\theta}_{j,l}^c = \boldsymbol{\theta}^c$; else set $\boldsymbol{\theta}_{j,k} = \boldsymbol{\theta}_{j,l}^c$.

5. If $q_j = 1$ terminate the iterations, else continue with $j = j + 1$.

3.3 Hierarchical (Multi-level) Bayesian Inference

Let a series of models $\mathcal{M}^i(\boldsymbol{\theta}^i)$, with $i = 1, ..., K$, where each one is characterized by a series of parameters $\boldsymbol{\theta}^i = [\theta_1^i, ..., \theta_N^i]$ considered as random variables. For instance the model $\mathcal{M}$ could be a computational model of a structural assembly, while the parameters $\boldsymbol{\theta}$ could define a series of material or geometric properties of $\mathcal{M}$. Additionally, consider a series of experimental measurements or observations $\boldsymbol{D}$, with $\boldsymbol{D} = [d_1, ..., d_m]$, that have been acquired from experiments on the physical asset simulated by the computational model $\mathcal{M}$.

Based on the hierarchical Bayesian paradigm, each dataset $\boldsymbol{D}^i$ 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 $\boldsymbol{\theta}^i$ for $i = 1, .., K$ 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 classical Bayesian problem posed in eq. (3.14) is reformulated in the hierarchical variation as:

$$\boldsymbol{y}^i(\boldsymbol{\theta}^i|\boldsymbol{\psi}) = \boldsymbol{m}^i(\boldsymbol{\theta}^i|\boldsymbol{\psi}) + \boldsymbol{\epsilon}^i \qquad i = 1, ..., K \tag{3.18}$$

Each error term $\boldsymbol{\epsilon}^i$ in eq. (3.18) can be considered as a random variable with a predetermined PDF. Alternatively, the statistical parameters of this PDF could also be considered as unknown parameters to be inferred. The joint posterior distribution is expressed through Bayes theorem as:

$$P(\boldsymbol{\theta}, \boldsymbol{\psi}|\boldsymbol{D}) = \frac{P(\boldsymbol{D}|\boldsymbol{\theta})P(\boldsymbol{\theta}|\boldsymbol{\psi})P(\boldsymbol{\psi})}{P(\boldsymbol{D})} = \prod_{i=1}^{K}\Big[\prod_{j=1}^{N_i}\big[P(d^{i,j}|\boldsymbol{\theta}^i)\big]P(\boldsymbol{\theta}^i|\boldsymbol{\psi})\Big]\frac{P(\boldsymbol{\psi})}{P(\boldsymbol{D})} \tag{3.19}$$

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

The form of both the conditional prior PDFs of $\boldsymbol{\theta}^i$ and the prior PDF of $\boldsymbol{\psi}$ have to be determined. As in eq. (3.14), in situations where there is absence of prior information regarding the investigated parameters, uniform distributions are generally preferred for both PDFs. Alternatively, conjugate priors has been used for $\boldsymbol{\psi}$ in [19], however that was done mainly for alleviating the computational demands, since a closed form solution can be retrieved for the posterior.

3.4 GLOBAL SENSITIVITY ANALYSIS

Sensitivity Analysis (SA) is the study of how different sources of uncertainty in the model's input space can impact the uncertainty of the model output. SA is in close relation with uncertainty analysis, with SA going a step further and instead of simply aiming at the investigation of the uncertainty in the system output, it can also quantify how this uncertainty is affected by each of the input factors. SA can serve a number of valuable objectives in the economy of modelling, as it can uncover technical modelling errors, identify critical regions in the input space, establish priorities for research and simplify models. To overcome the limitations of derivative-driven SA methods, the so-called global methods have gained significant traction. These methods consider the input factors as uncertain and investigate simultaneously a handful of data points in the input space. The global methods can outperform significantly the local ones, especially in uncertain and nonlinear models. Here the variance-based (global) sensitivity analysis (VBSA) will be reviewed, since it has been applied in the present thesis, but for a more

thorough study of SA the reader is referred to [113].

Let $\mathcal{M}$ be a (scalar) model of the form $\mathcal{M} = f(\boldsymbol{X})$, where $\boldsymbol{X} = (X_1, ..., X_d) \in [0,1]^d \subset \mathbb{R}^d$ is the input vector of random variables, with X_i being the i-factor and $f(\boldsymbol{X}) \in \mathcal{L}(\mathbb{R}^d)$, a square-integrable function. By applying a functional decomposition, known as Hoeffding decomposition [57], f can be written in terms of elementary functions as:

$$f(\boldsymbol{X}) = f_0 + \sum_{i=1}^{d} f_i(X_i) + \sum_{i}^{d} \sum_{j>i}^{d} f_{ij}(X_i, X_j) + ... + f_{1...d}(X_1, ..., X_d) \tag{3.20}$$

The functions $f_{i_1, i_2, ..., i_k}$ are uniquely given by the equations:

$$f_0 = \mathbb{E}(\mathcal{M}) \tag{3.21}$$

$$f_i = \mathbb{E}_{\boldsymbol{X}_{\sim i}}(\mathcal{M}|X_i) - \mathbb{E}(\mathcal{M}) \tag{3.22}$$

$$f_{ij} = \mathbb{E}_{\boldsymbol{X}_{\sim ij}}(\mathcal{M}|X_i, X_j) - f_i - f_j - \mathbb{E}(\mathcal{M}) \tag{3.23}$$

(similarly for higher order terms)

which satisfy the condition [120]

$$\int_0^1 \cdots \int_0^1 f_{i_r \cdots k}(x_{i_1}, \ldots, x_{i_k}) dx_{i_1} \cdots dx_{i_k} = 0 \quad , \quad \{i_i, ..., i_k\} \subseteq \{i_1, .., i_d\} \tag{3.24}$$

In the above, the notation $\mathbb{E}_{\boldsymbol{X}_{\sim i}}(\mathcal{M}|X_i)$ refers to the mean of $\mathcal{M}$, taken over all possible values of $\boldsymbol{X}_{\sim i}$, while keeping X_i fixed.

In a similar fashion to eq. (3.20), a functional decomposition for the model variance termed as ANOVA [31] is defined as:

$$V(\mathcal{M}) = \sum_{i=1}^{d} V(f_i) + \sum_{i=1}^{d} \sum_{j>i}^{d} V(f_{ij}) + ... + V(f_{12..d}) \tag{3.25}$$

where $V(\cdot)$ is the variance of the respective term.

The variance of each function of the ANOVA series can be written in terms of partial variances as:

$$V(f_i) = V\left(\mathbb{E}_{\boldsymbol{X}_{\sim i}}(\mathcal{M}|X_i)\right) \ , \ \ V(f_{ij}) = V\left(\mathbb{E}_{\boldsymbol{X}_{\sim ij}}(\mathcal{M}|X_i, X_j)\right) - V(f_i) - V(f_j) \ , \ \ ... \ (3.26)$$

and so on for higher order interactions.

If the expansion in eq. (3.25) is normalized i.e. divided by the total variance $V(\mathcal{M})$ of the model then it takes the form:

$$\sum_{i=1}^{d} S(f_i) + \sum_{i=1}^{d}\sum_{j>i}^{d} S(f_{ij}) + ... + S(f_{12..d}) := \sum_{i=1}^{d} S_i + \sum_{i=1}^{d}\sum_{j>i}^{d} S_{ij} + ... + S_{1..d} = 1 \ (3.27)$$

with $S_i, S_{ij}, \ldots, S_{1..d}$ denoting the sensitivity indices or Sobol indices of the respective order. These indices, whose values reside in the domain $[0, 1]$, essentially provide percentages of how the information of the model output is affected by each input factor or by the synergy of them. For practical purposes, the most useful indices are that of the first and second order, since generally these are the most dominant terms of the expansion.

An additional metric was introduced in [58] with the so-called total effect indices. In this case, for an input factor X_i, in the normalized expansion of eq. (3.27) all the terms that do not contain X_i are left out. Thus, the expansion for a certain X_i is expressed as:

$$S_{Ti} = S_i + \sum_{j\neq i}^{d} S_{ij} + \sum_{j\neq i, k\neq i, k>j}^{d} S_{ijk} + ... + S_{1..d} = 1 - S_{Ci} \tag{3.28}$$

where S_{Ti} is the total effect index of X_i and S_{Ci} is the sum of the additional terms of the expansion that do not involve X_i. Essentially by employing these indices, the higher order terms that include X_i are also taken into consideration.

An efficient way to compute S_i and S_{Ti} was presented in [112] by using Monte Carlo sampling or for better exploration of the input space via a pseudo-random procedure. The general concept is that from two independent sampling matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$ one can create a series of additional matrices $\boldsymbol{Q}_{\boldsymbol{R}}^{(i)}$, where all the columns are taken from $\boldsymbol{Q}$ except for the column i which is taken from $\boldsymbol{R}$ and the same procedure holds for the generation of $\boldsymbol{R}_{\boldsymbol{Q}}^{(i)}$. The rows and columns of these matrices represent the samples and input factors accordingly. Next, the estimators for the various indices can be computed. A commonly applied estimator for the calculation of S_i [112] is:

$$S_i \simeq \frac{1}{K} \frac{1}{V^*(\mathcal{M})} \sum_{j=1}^{K} f(\boldsymbol{Q}_{(j)})(f(\boldsymbol{R}_{\boldsymbol{Q},(j)}^{(i)}) - f(\boldsymbol{R}_{(j)})) \qquad (3.29)$$

and accordingly the estimator for S_{Ti} is:

$$S_{Ti} \simeq \frac{1}{2K} \frac{1}{V^*(\mathcal{M})} \sum_{j=1}^{K} (f(\boldsymbol{Q}_{(j)}) - f(\boldsymbol{Q}_{\boldsymbol{R},(j)}^{(i)}))^2 \qquad (3.30)$$

where the subscript j denotes the row of the respective matrix, while K is the sample size and $V^*(\mathcal{M})$ is the unbiased estimator of the unconditional model variance.

An effective way to reduce the required evaluations is to restate eq. (3.29) as:

$$S_i \simeq \frac{1}{K} \frac{1}{V^*(\mathcal{M})} \sum_{j=1}^{K} f(\boldsymbol{R}_{(j)})(f(\boldsymbol{Q}_{\boldsymbol{R},(j)}^{(i)}) - f(\boldsymbol{Q}_{(j)})) \qquad (3.31)$$

leading to $K(d+2)$ model solutions.

The estimator of the variance is then given by:

$$V^*(\mathcal{M}) = \frac{1}{K(d+2)-1} \sum_{j=1}^{K} \sum_{\mu=1}^{d+2} \left| F_{j,\mu} - \frac{1}{K(d+2)} \sum_{j=1}^{K} \sum_{\mu=1}^{d+2} F_{j,\mu} \right|^2 \qquad (3.32)$$

where $F_{j,\mu}$ are the components of a matrix $\boldsymbol{F}$ that includes the model output for each row j of the respective sampling matrix μ and is expressed as:

$$\boldsymbol{F} = \begin{bmatrix} f(\boldsymbol{Q}_{(1)}) & f(\boldsymbol{R}_{(1)}) & f(\boldsymbol{Q}_{\boldsymbol{R},(1)}^{(1)}) & \cdots & f(\boldsymbol{Q}_{\boldsymbol{R},(1)}^{(d)}) \\ \vdots & \vdots & \vdots & & \vdots \\ f(\boldsymbol{Q}_{(j)}) & f(\boldsymbol{R}_{(j)}) & f(\boldsymbol{Q}_{\boldsymbol{R},(j)}^{(1)}) & \cdots & f(\boldsymbol{Q}_{\boldsymbol{R},(j)}^{(d)}) \\ \vdots & \vdots & \vdots & & \vdots \\ f(\boldsymbol{Q}_{(K)}) & f(\boldsymbol{R}_{(K)}) & f(\boldsymbol{Q}_{\boldsymbol{R},(K)}^{(1)}) & \cdots & f(\boldsymbol{Q}_{\boldsymbol{R},(K)}^{(d)}) \end{bmatrix} \qquad (3.33)$$

3.5 STOCHASTIC OPTIMIZATION

One fundamental aspect of optimization problems encountered in engineering lies in the uncertainty surrounding the parameters involved. For instance, material parameters such as yield stresses, along with external loading, manufacturing errors, numerical

approximation errors, are not known a priori. Instead, they must be regarded as random variables with specific probability distributions. Therefore, the optimization aims at finding the optimal solution in a stochastic problem, which translates to finding the optimal probability density functions that best describe the parameter values that maximize or minimize a stochastic objective function.

Let's think about a model, denoted as $\mathcal{M}(\boldsymbol{\theta}, \mathfrak{X})$, characterized by a set of parameters $\boldsymbol{\theta} = [\theta_1, ..., \theta_N]$, termed as design variables. For example, $\mathcal{M}$ might represent a computational model of a structural assembly, with $\boldsymbol{\theta}$ specifying various loading, material or geometric attributes of $\mathcal{M}$. Also, let a set of input variables $\mathfrak{X} = [\mathfrak{X}_1, ..., \mathfrak{X}_K]$ that directly affect the outcome of $\mathcal{M}$ and are assumed to be random variables. Additionally, let an objective (cost) function f, dependent on $\boldsymbol{\theta}$ and on statistics of the outcome of $\mathcal{M}$ (e.g $E[m]$, $E[m^2]$). The minimization of f provides optimal targeted actions via selecting the most appropriate $\boldsymbol{\theta}$. For example, this could be the choice of the optimal material or geometry of a structure towards minimization of the expected displacement under random loading conditions that constitute $\mathfrak{X}$.

A well-established algorithm to solve stochastic optimization problems is called Co-variance Matrix Adaptation - Evolution Strategy (CMA-ES) [49]. This algorithm seeks to find the minimum of the objective function f by sampling λ points $\{\boldsymbol{\theta}_i\}_{i=1}^{\lambda}$ from a Gaussian distribution $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with the mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$ being updated at each iteration (generation) of the algorithm. At each iteration the samples are evaluated and sorted based on their corresponding function values. Then, $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are adapted based on the k best samples in order to increase the probability of sampling future individuals in the direction of favorable samples. The implementation steps of the algorithm are outlined in algorithm 1.

- *Initialize model parameters*: $\sigma^0, \boldsymbol{\mu}^0, \boldsymbol{C}^0, \boldsymbol{p}_c^0, \boldsymbol{p}_\sigma^0, k_{eff}, c_c, c_\sigma, c_1, c_k, d_\sigma$

- While *termination criteria not met*

  1. Draw new candidate samples

$$\boldsymbol{\theta}_i^{g+1} = \boldsymbol{\mu}^g + \sigma^g \boldsymbol{L}^g \boldsymbol{z}_i, \text{ with } \boldsymbol{z}_i \sim \mathcal{N}(0, \boldsymbol{I}) \tag{3.34}$$

  where $\sigma^g > 0$ is the step size at the g-th generation and $\boldsymbol{L}^g$ is the lower

triangular matrix obtained from the Cholesky decomposition of

$$\mathbf{\Sigma} = \sigma^2 \boldsymbol{C} = \sigma^2 \boldsymbol{L}\boldsymbol{L}^T \tag{3.35}$$

2. Evaluate objective function f at the obtained samples $\boldsymbol{\theta}_i^{g+1}$ and sort them according to their fitness

$$f(\boldsymbol{\theta}_{1:\lambda}) \leq f(\boldsymbol{\theta}_{2:\lambda}) \leq \cdots \leq f(\boldsymbol{\theta}_{\lambda:\lambda}) \tag{3.36}$$

with $\boldsymbol{\theta}_{i:\lambda}$ denoting the i-th fittest sample.

3. Choose the k first samples that produce the smallest function values and update the mean as:

$$\boldsymbol{\mu}^{g+1} = \frac{1}{k}\sum_{i=1}^{k}\boldsymbol{\theta}_{i:\lambda}^{g+1} \tag{3.37}$$

4. Update the covariance matrix in order to increase the probability of selecting favorable samples

$$\boldsymbol{p}_c^{g+1} = (1 - c_c)\boldsymbol{p}_c^g + \sqrt{c_c(2 - c_c)k_{eff}}\frac{\boldsymbol{\mu}^{g+1} - \boldsymbol{\mu}^g}{\sigma^g} \tag{3.38}$$

$$\boldsymbol{p}_\sigma^{g+1} = (1 - c_\sigma)\boldsymbol{p}_\sigma^g + \sqrt{c_\sigma(2 - c_\sigma)k_{eff}}(\boldsymbol{C}^g)^{-\frac{1}{2}}\frac{\boldsymbol{\mu}^{g+1} - \boldsymbol{\mu}^g}{\sigma^g} \tag{3.39}$$

$$\sigma^{g+1} = \sigma^g exp\left(\frac{c_\sigma}{d_\sigma}\left(\frac{\|\boldsymbol{p}_\sigma^{g+1}\|}{\mathbb{E}\left[\mathcal{N}(0, \boldsymbol{I}\right]} - 1\right)\right) \tag{3.40}$$

$$\boldsymbol{C}^{g+1} = (1 - c_1 - c_k)\boldsymbol{C}^g + c_1\boldsymbol{p}_c^{g+1}\left(\boldsymbol{p}_c^{g+1}\right)^T + c_k\sum_{i=1}^{k} w_i\bar{\boldsymbol{\theta}}_{i:\lambda}^{g+1}\left(\bar{\boldsymbol{\theta}}_{i:\lambda}^{g+1}\right)^T,$$
$$\tag{3.41}$$

with $\bar{\boldsymbol{\theta}}_{i:\lambda}^{g+1} = \frac{\boldsymbol{\theta}_{i:\lambda}^{g+1} - \boldsymbol{\mu}^g}{\sigma^g}$

5. Proceed to the next generation: $g \leftarrow g + 1$

• Return solution

4

Machine learning and surrogate modeling

4.1 MACHINE LEARNING-BASED SURROGATE MODELS

In the past two decades, significant strides in computational technology have facilitated the utilization of sophisticated, physics-based mathematical models to characterize complex physical phenomena encountered across various domains of engineering and science. These intricate models offer enhanced accuracy in portraying the fundamental physics of a given problem, thereby yielding precise depictions of system behavior. However, despite their efficacy, computational expenses remain a considerable challenge, particularly in scenarios involving nonlinear or dynamic systems with high-dimensionalities. This challenge becomes especially pronounced in parametrized systems, common in fields such as stochastic analysis, sensitivity analysis, or optimization, where conducting numerous simulations becomes imperative. To address this issue, surrogate models emerge as a viable solution, enabling the substitution of the original model with a function that replicates the complex system's behavior at significantly reduced computational costs per evaluation.

To describe the problem in a more concrete setting, let as consider a computational model $\mathbb{M}$, which takes $\mathbb{M}$-dimensional vectors $\mathbf{x}$ as inputs and maps them to $\mathbb{R}^N$:

$$\mathbb{M} : \mathbf{x} \in \mathscr{D} \subset \mathbb{R}^M \mapsto \mathbf{y} \in \mathbb{R}^N \qquad (4.1)$$

A surrogate model $\tilde{\mathbb{M}}$ can then be defined as:

$$\tilde{\mathbb{M}} : \mathbf{x} \in \mathscr{D} \subset \mathbb{R}^M \mapsto \mathbf{y} \in \mathbb{R}^N \tag{4.2}$$

such that

$$\tilde{\mathbb{M}} \approx \mathbb{M} \tag{4.3}$$

The surrogate model's construction may rely on an assumed functional form of $\mathbb{M}$ and/or partial data gathered from restricted executions of the original model. Additionally, $\mathbb{M}$ is treated as a black box, where the internal workings remain unknown, and only the output $\mathbf{y} = \mathbb{M}(\mathbf{x})$ is observable. For instance, a finite element model serves as a surrogate, approximating the solution to the system's governing equations that lack analytical solutions.

In recent decades, various methodologies have emerged for the development of surrogate models. Noteworthy among these are polynomial chaos expansions, Gaussian process regression, support vector machines, radial basis functions, principal component analysis, and artificial neural networks. Next, the focus is on neural networks (NNs) and especially on feed-forward neural networks (FFNNs) which have been extensively employed in this thesis.

NNs are information-processing mathematical models inspired by the biological neural networks that constitute the human brain. As its original counterpart, they are able to learn from observational data, that is, by considering examples without being programmed with any task-specific rules. The basic component of an NN is the artificial neuron. An artificial neuron, denoted with j is a processing unit which performs the following operations:

1. It receives an input signal x_i from the synapse i

2. It multiplies the signal by the synaptic weight w_{ji}

3. It sums all input signals x_i with their respective weights w_{ji}, for all the synapses $i = 1, ..., n$ and adds a bias term b_j.

4. It processes the sum of the input signals through an activation function $\varphi(\cdot)$, for example the sigmoid or the hyperbolic tangent function, and outputs the result y_j.

In mathematical terms, the neuron j can be described by the equation:

44

$$y_k = \varphi(\sum_{i=1}^{n} w_{ji} + b_j) \tag{4.4}$$
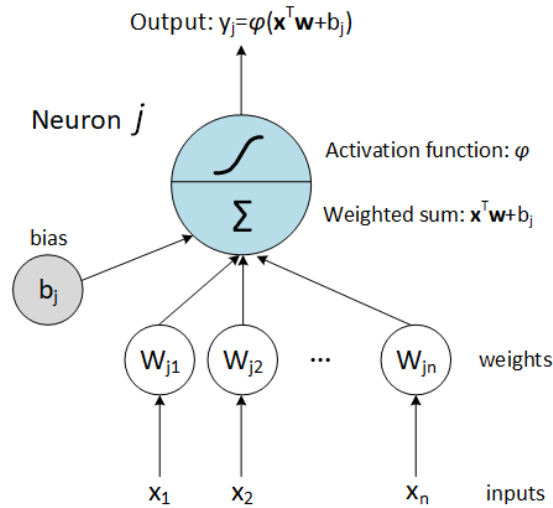
Figure 4.1: Nonlinear model of a neuron, labeled j

A schematic representation of the computational model of a nonlinear neuron is depicted in fig. 4.1.

In this context, an NN is an oriented graph with neurons being the nodes of the graph and the synapses being the oriented edges. The synaptic weights are calibrated through a training process based on observational data. Depending on the interconnection of neurons, different types of neural networks arise. Amongst them, the most popular and widely applied type is the feed-forward neural network (FFNN), also known as perceptron. In terms of the architecture, an FFNN consists of the input layer, the hidden layer(s) and the output layer. NNs with more than one hidden layer are referred to as deep neural networks or multi-layer perceptrons. In terms of connectivity, in FFNN neurons from a layer can only be connected with neurons from the next layer towards the output layer. This means that the information moves in only one direction, forward, from the input nodes, through the hidden nodes (if they exist) and to the output nodes. An example of an FFNN network architecture is given in figure 4.2.
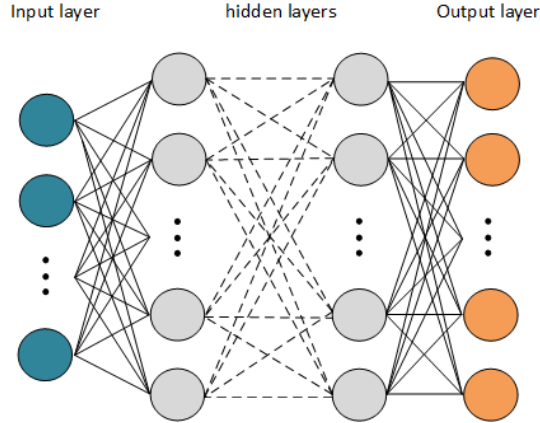
Figure 4.2: Example of a generic feedforward neural network

For a specific choice of network architecture, to train the network a set of N_{tr} labeled data $\{\overline{\boldsymbol{x}}, \overline{\boldsymbol{y}})\}_{1 \leq k \leq N_{tr}}$ is first provided. Next, a loss function is specified, such as the mean square error, which is given by the expression:

$$MSE = \sum_{k=1}^{N_{tr}} \|\overline{\boldsymbol{y}}_k - \hat{\boldsymbol{y}}_k\|_2^2 \in \mathbb{R} \tag{4.5}$$

where, $\overline{\boldsymbol{y}}_k$ is the target output for the input $\overline{\boldsymbol{x}}_k$ and $\hat{\boldsymbol{y}}_k$ is the respective network's output. Then, the training of the network consists in finding the optimal weights $\boldsymbol{w} = (w_{ij})$ that minimize $MSE = MSE(\boldsymbol{w})$. Some commonly preferred algorithms to solve this (non-convex) optimization problem are the stochastic gradient descent algorithm, the Levenberg-Marquardt algorithm [47] and Adam [66].

Based on the above, FFNNs essentially establish a non-linear map from the space of the input data to the space of the output data. Their powerful approximation properties are well-established from numerous applications, but, also, from a theoretical standpoint, from the universal function approximation theorem which states that:*A FFNN with one hidden layer, that contains a finite number of neurons and has non-constant, bounded and continuous activation functions, can approximate any continuous function defined on a compact subset of $\mathbb{R}^m$.* Or, equivalently, the Cybenko's theorems with respect to one-hidden layer [24] and two-hidden layer networks [23].

Training an NN involves optimizing its set of weights, which can be approached using various methods ranging from evolutionary algorithms to gradient-based techniques like BFGS, with stochastic gradient descent and its many variants being the most common.

In gradient-based methods, which are the most popular for NN training, evaluating the gradient of a function f can be done using numerical differentiation, symbolic differentiation, or automatic differentiation. Numerical differentiation approximates derivatives using finite differences from the function's values at sample points, but it has several drawbacks. It requires $O(n)$ evaluations of f for a gradient in nn dimensions and is prone to being ill-conditioned and unstable. On the other hand, automatic differentiation involves automatically manipulating expressions to obtain derivative expressions by applying differentiation rules. While symbolic derivatives can produce exact analytical expressions, they are not suitable for efficient runtime calculation as they can become exponentially larger than the original expression.

Automatic Differentiation (AD) performs a unique interpretation of a computer program by adjusting the domain of variables to include derivative values and redefining operator semantics to propagate derivatives according to the chain rule of differential calculus. As a technical term, AD refers to a set of techniques that compute derivatives by accumulating values during code execution, resulting in numerical derivative evaluations instead of symbolic derivative expressions. This approach allows for precise derivative evaluation at machine precision, with minimal overhead and optimal asymptotic efficiency. Unlike the complexity of reformatting code into closed-form expressions required by symbolic differentiation, AD can be applied to standard code with minimal modifications, accommodating branching, loops, and recursion.

In machine learning, the backpropagation algorithm, a specialized form of Automatic Differentiation (AD), is essential for training neural networks. Essentially, backpropagation represents learning as gradient descent within the neural network's weight space, aiming to find the minima of an objective function. The necessary gradient is determined by propagating the sensitivity of the objective value backward from the output, using the chain rule to calculate the partial derivatives of the objective function with respect to each weight.

A function $f : R^n \to R^m$ is constructed using intermediate variables v_i such that:

- variables $v_{i-n} = x_i$, $i = 1, ..., n$ are the input variables,

- variables v_i , $i = 1, ..., l$ are the working (intermediate) variables, and

- variables $y_{m-i} = v_{l-i}$, $i = m - 1, ..., 0$ are the output variables.

AD in the reverse accumulation mode corresponds to a generalized backpropagation

47

algorithm, in that it propagates derivatives backward from a given output. This is done by complementing each intermediate variable v_i with an adjoint:

$$\bar{v}_i = \frac{\partial y_j}{\partial v_i} \tag{4.6}$$

which represents the sensitivity of a given output y_j with respect to changes in v_i. In the context of backpropagation i.e. the training procedure, y corresponds to the scalar error E. Additionally, AD has the ability to calculate the Jacobian J of the NN model, i.e. the derivative matrix of the m-dimensional vector valued NN output with respect to the n-dimensional vector valued NN input, expressed as:

$$J = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_1}{\partial x_n} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix} \tag{4.7}$$

Reverse mode AD involves a two-phase process for computing derivatives. Initially, during the first phase, the original function code is executed forward, generating intermediate variables v_i and logging dependencies in a computational graph through a bookkeeping procedure. Subsequently, in the second phase, derivatives are determined by backpropagating adjoints $\bar{v}_i$ from the outputs to the inputs.

4.2 Machine learning-based material modeling

The idea to harness the exceptional predictive capabilities of machine learning algorithms and especially NNs for the development of material models, based on experiments, has already been explored more than 30 years ago [40]. The motivation for that is that if the experimental results contain the relevant information about the material behavior, then the trained NN would contain sufficient information about the material behavior to qualify as a material model. Such a trained NN not only would be able to reproduce the experimental results it was trained on, but through its generalization capability it should be able to approximate the results of other experiments.

There are several ways to create an appropriate functional form for the material behavior, e.g. based on the choice of a strain-based or stress based formulation or whether the material is path-dependent or not. A NN can be trained on the results of several proportional and non-proportional strain or stress paths and then can simulate

the test results for other proportional and non-proportional strain or stress paths. The degree of accuracy in this generalization depends on how comprehensive the training set is. For instance, for a strain-controlled problem with no path-dependency, training means to present the network with the experimental data and have it self-organize, or modify its weights, such that it correctly reproduces the total stress state when presented with the total strain state as shown in fig. 4.3 for a 2D material. Alternatively, when there is path dependency, the current stress and strain states along with the strain increment can be used as the input of the NN towards the prediction of the stress increment output as illustrated in fig. 4.4 for a 2D material. By using the appropriate strain and stress components, the extension to 3D materials is straightforward.
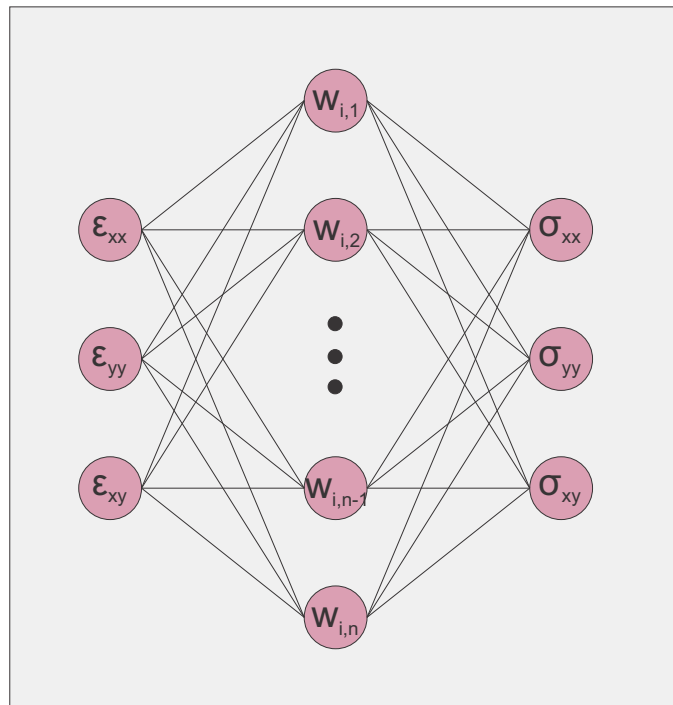
Figure 4.3: Strain-based NN for 2D non-path dependent material behavior
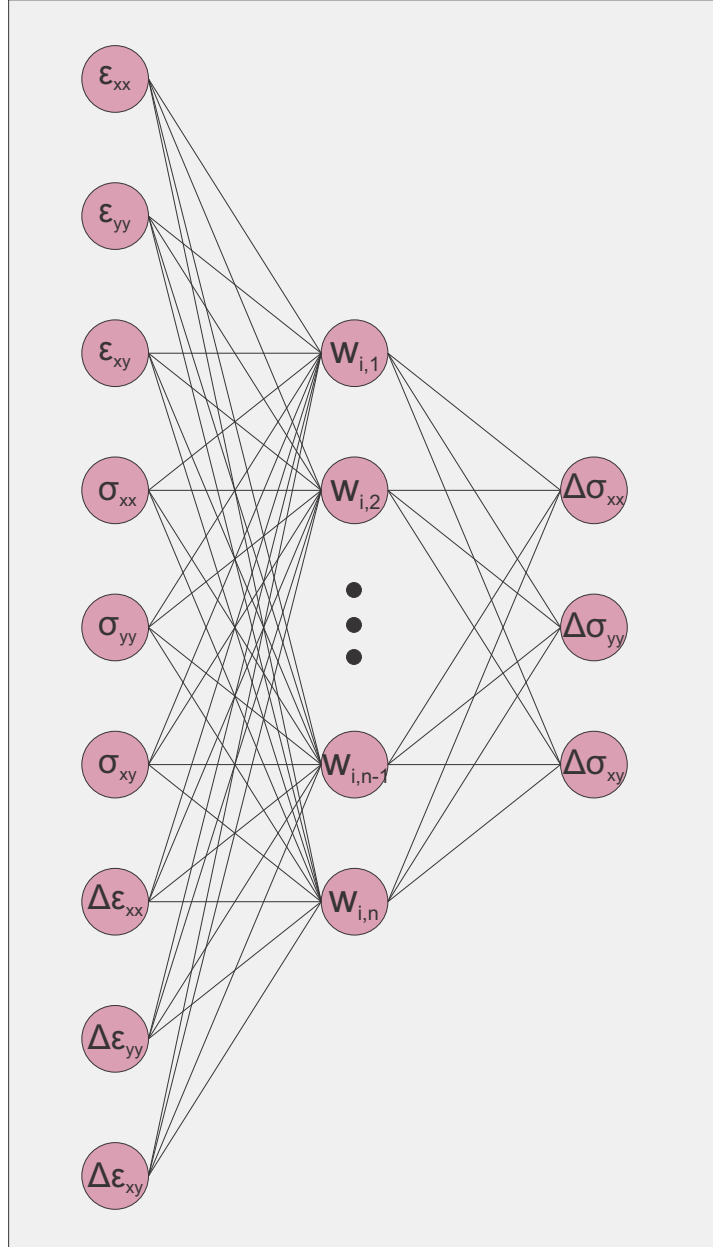
Figure 4.4: Strain-based NN for 2D path dependent material behavior

A significant benefit of NN material models is that the tangent material stiffness matrix can be straightforwardly be extracted through differentiation on the trained NN [52]. The most efficient way to do this is through Automatic Differentiation as explained

in section 4.1. The tangent matrix of a material model is expressed as:

$$C = \frac{\partial \boldsymbol{\sigma}}{\partial \boldsymbol{\epsilon}} = \begin{bmatrix} \frac{\partial \sigma_1}{\partial \epsilon_1} & \cdots & \frac{\partial \sigma_m}{\partial \epsilon_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial \sigma_1}{\partial \epsilon_n} & \cdots & \frac{\partial \sigma_m}{\partial \epsilon_n} \end{bmatrix} \tag{4.8}$$

Observing eq. (4.8), it can be realized that it corresponds to a standard case of a Jacobian computation. Therefore, by applying the chain rule on the output layer, i.e. the stress vector with respect to the input layer, i.e. the strain vector, one can directly extract the components of the tangent matrix as:

$$C_{ij} = \frac{\partial \sigma_{ij}}{\partial h_k} \frac{\partial h_k}{\partial h_{k-1}} \cdots \frac{\partial h_1}{\partial \epsilon_{ij}} \tag{4.9}$$

where h are the hidden layers of the NN.

When experimental observations are absent, instead microstructural simulations can be utilized to create the necessary dataset for the NN training. This idea has been proposed more recently (e.g. [74] for 2D RVE simulations and [72] for 3D RVE simulations) where computational homogenization is applied towards the extraction of effective strain-stress or strain-strain energy data pairs. In that sense the NN that is tasked with emulating the behavior of a composite material system through RVE simulations can be considered as a surrogate model. That is because when a multiscale computational homogenization analysis takes place (section 1.6), the NN completely bypasses the need to perform elaborate FE analyses each time the algorithmic procedure enters an integration point. This surrogate modeling technique is an integral part of the present thesis and the concept is expanded in more detail during the presentation of the original works (see chapters 5-8).

5

A neural network-aided Bayesian identification framework for multiscale modeling of nanocomposites

5.1 Introduction

This chapter proposes a framework for learning the parameters that characterize the CNT/polymer interface in the microscale with data available from upper scales. These parameters are associated with great uncertainties and their characterization is a difficult task, since microscale measurements are costly and hard to obtain. To overcome this, the present study introduces a computational framework for updating the prior beliefs on the values of these parameters, by using deformation measurements on meso- and/or macro-scale composite structures. In terms of modeling, the CNT/polymer interface is formulated using a cohesive zone model and a bilinear bond-slip constitutive law. The FE^2 method is then employed for predicting the response of the composite structures, but despite its accuracy, this method is associated with immense computational demands for large-scale problems. Therefore, its application to the Bayesian setting that requires multiple model evaluations is prohibitive. To alleviate this enormous cost, a surrogate modeling technique is developed which utilizes artificial neural networks, trained to

predict the nonlinear stress-strain relationship of representative volume elements of the microstructure. The data set over which the neural network is trained, is obtained by analyzing a limited number of different RVE configurations using a detailed finite element analysis. The elaborated methodology is first validated through a numerical example from 2D elasticity, which demonstrated its high accuracy and its significant cost reduction capabilities. It is then applied to a more challenging large-scale problem from 3D elasticity. Even though this research focuses on the characterization of the mechanical properties of composite materials, the proposed numerical procedure is generic and can be straightforwardly applied to other physically analogous phenomena related to nano-composite modeling, such as parameter identification in heat transfer or electrical conduction.

The remaining of this chapter is organized as follows. In section 5.2 the formulation of the multiscale problem and the modeling of the CNT/polymer interface with the cohesive zone model are presented. Section 5.3 introduces the proposed Bayesian framework for multiscale systems. In section 5.4 the usage of the neural network as a surrogate model is illustrated. Lastly, section 5.5 provides numerical applications to test the efficiency of the surrogate model compared to the full-scale system solution, in terms of precision and computational cost reduction.

5.2 MULTISCALE MODEL AND SOLUTION PROCEDURE

For the accurate description of the composite material, a multiscale model with three stages is implemented in this work. The atomic scale is linked with the microscopic scale with a hierarchical approach, while the microscopic is linked with the macroscopic scale with a semi-concurrent approach.

At the atomic scale each pair of carbon's C-C covalent bonds are modeled with the molecular structural mechanics (MSM) method [17]. According to this approach, the lattice is simulated as a space frame structure, where every bond between adjacent carbon atoms is modeled as a continuous circular beam. A more thorough description of MSM is given in section 2.1. Despite this being a fairly accurate representation of the atomic structure it leads to excessive computational cost for each CNT. Hence, before proceeding to the next scale each space frame is projected to an equivalent beam element (EBE), which reduces the total degrees of freedom of the structure to a reasonable amount. As detailed in section 2.1, the structural properties of the EBEs are calculated by subjecting

53

a cantilever beam to an axial F_x, a transverse F_y and a torsional T load. Subsequently, the axial stiffness $(EA)_{eq}$, the bending rigidity $(EI)_{eq}$ and the torsional rigidity $(GI)_{eq}$ are obtained, via eq. (2.6), by measuring the horizontal displacement u_x, the vertical displacement u_y and the angle of rotation φ corresponding to the aforementioned loads .

At the microscopic scale a representative volume element (RVE) is chosen for the formulation of the surrounding polymer matrix. A number of EBEs are then added as inclusions of the matrix in order to achieve a specific volume fraction of CNT/polymer. The addition of these elements is usually done with a random positioning generator. In terms of the finite element modeling of the CNT/polymer composite system, the cohesive formulation, provided in section 2.3, is used for simulating the interaction between the two phases.

The non-linearity in the interface is given in the form of a bond-slip law expressed as:

$$\boldsymbol{\tau} = f(\Delta \boldsymbol{u}) \tag{5.1}$$

A linearization of eq. (5.1) yields the tangential relation between increments of the traction $d\boldsymbol{\tau}$ and the relative displacement $d\Delta \boldsymbol{u}$ in terms of a tangent constitutive matrix $\boldsymbol{D}_{tan}$. Thus, eq. (5.1) is restated in incremental form as:

$$\boldsymbol{\tau} = \boldsymbol{D}_{tan} \Delta \boldsymbol{v}^l \tag{5.2}$$

where $\Delta \boldsymbol{v}^l$ is the relative displacement of the CBE with its corresponding EBE in local coordinates and can be calculated as:

$$\Delta \boldsymbol{v}^l = \boldsymbol{R}_m^T \boldsymbol{N}_b (\boldsymbol{v}' - \boldsymbol{v}) \tag{5.3}$$

In the case that there is no coupling between the displacements in the three local directions, $\boldsymbol{D}_{tan}$ is written as:

$$\boldsymbol{D}_{tan} = \begin{bmatrix} D_{11} & 0 & 0 \\ 0 & D_{22} & 0 \\ 0 & 0 & D_{33} \end{bmatrix} \tag{5.4}$$

where D_{11} is the slip component, while D_{22} and D_{33} are the normal components. It is assumed that perfect bonding exists between the solid and the EBE for the two normal components, thus having linear behavior. On the contrary, for the slip component a

54

bi-linear approach is selected, as depicted in fig. 5.1 and is formulated as:

$$D_{11} = \begin{cases} D_{el}, & \tau_1 \leq \tau_{1,s} \\ D_{pl} = pD_{el}, & \tau_1 > \tau_{1,s} \end{cases} \tag{5.5}$$

where D_{el} is the elastic slope and D_{pl} is the plastic slope of the bi-linear diagram, while $\tau_{1,s}$ is the interfacial shear strength i.e the shear stress at which the interface is leaving the elastic region. Also, $p \in [0,1)$ represents the percentage of the hardening in the plastic region.
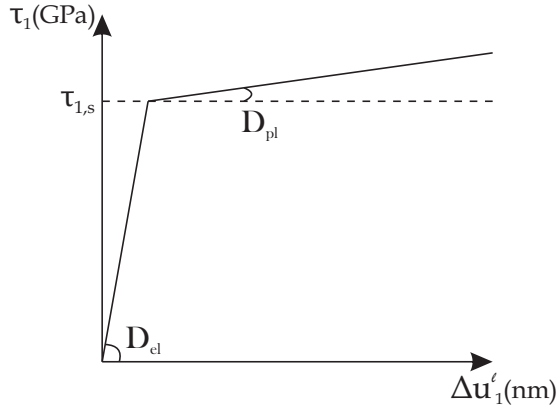
Figure 5.1: Bi-linear constitutive law of the slippage

This particular bi-linear law is straightforwardly parametrizable by the parameters $\tau_{1,s}$, D_{el}, D_{pl} and will allow us to demonstrate the application of the proposed Bayesian methodology for updating our initial beliefs on these parameters based on available data. However, other more involved constitutive laws could be used as well, such as bi-linear relations with softening, tri-linear [131], or exponential [95].

Before advancing to the macroscopic scale, a homogenization scheme [88] must be implemented towards a connection with the microstructure. This scheme is applied to the microscopic model of an RVE of polymer matrix and linear EBEs representing the CNT reinforced inside the matrix. According to it, for a given macroscopic strain $\bar{\epsilon}$ a linear displacement function is applied to the boundary of the RVE as:

$$\boldsymbol{u}(\boldsymbol{x}) = \bar{\boldsymbol{\epsilon}}\boldsymbol{x} \quad \text{at } \boldsymbol{x} \in \partial\mathcal{V} \tag{5.6}$$

where $\boldsymbol{x}$ is the position vector of a point on the boundary $\partial\mathcal{V}$ of the RVE. This relation

provides the boundary conditions that are applied to the RVE with respect to the macroscopic variable and is referred to as the localization rule.

The solution of the aforementioned boundary value problem will produce the macroscopic stress $\overline{\boldsymbol{\sigma}} = [\overline{\sigma}_{11}, \overline{\sigma}_{22}, \overline{\tau}_{12}]^T$, calculated by the volume average of the microscopic stress field $\boldsymbol{\sigma}(\boldsymbol{x})$ via the relation:

$$\overline{\boldsymbol{\sigma}} = \frac{1}{\|\mathscr{V}\|} \int_{\mathscr{V}} \boldsymbol{\sigma} d\boldsymbol{x} \tag{5.7}$$

with $\|\mathscr{V}\|$ denoting the volume of the RVE. This equation expresses the homogenization rule, which yields the macroscopic state variables as a function of the microscopic stress state. The transition from the microscopic level to macroscopic and vice versa is schematically presented in fig. 5.2.
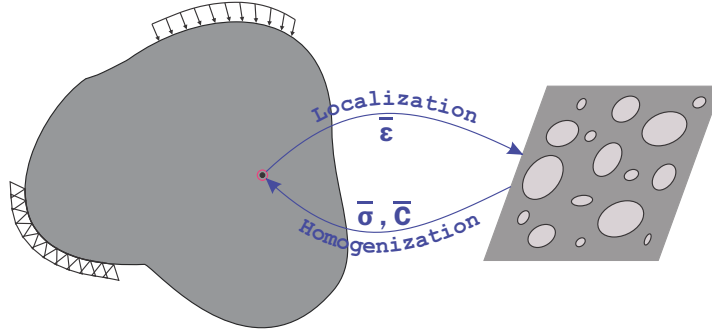
Figure 5.2: Transition between scales during FE^2

Lastly, the macroscopic tangent (effective) modulus $\overline{\boldsymbol{C}}$ is calculated as the derivative of the macroscopic stress $\overline{\boldsymbol{\sigma}}$ with respect to the macroscopic strain $\overline{\boldsymbol{\epsilon}}$, that is:

$$\overline{\boldsymbol{C}} = \partial_{\overline{\boldsymbol{\epsilon}}} \overline{\boldsymbol{\sigma}} \tag{5.8}$$

A more detailed description regarding the transition between two-scale material systems is given in section 1.3.

The semi-concurrent FE^2 algorithm [35] is used for the online transition between the two scales during the solution of the macroscopic composite system. According to this algorithm, for each load increment of the macrostructure, the macroscopic strain $\overline{\boldsymbol{\epsilon}}$ is calculated on every integration point of each finite element, where it is assumed that an RVE exists. After that, the previously described homogenization scheme is

56

implemented, from which the macroscopic stress $\overline{\boldsymbol{\sigma}}$ and tangent modulus $\overline{\boldsymbol{C}}$ are acquired for each Gauss point. With these, the macroscopic internal force vector $\overline{\boldsymbol{F}}_{int}(\overline{\boldsymbol{\sigma}})$ and tangential stiffness matrix $\overline{\boldsymbol{K}}(\overline{\boldsymbol{C}})$ can be calculated. This procedure is repeated until the internal force vector is equivalent to the external one, namely $\overline{\boldsymbol{F}}_{ex}$. The interaction of the two scales during the solution of the algorithm was illustrated in fig. 5.2. The algorithmic steps of FE^2 are presented in section 1.6.

5.3   BAYESIAN UPDATE OF THE INTERFACE PROPERTIES

The three previously introduced parameters that describe the interfacial relation between the CNTs and the polymer matrix are selected as random parameters to be updated. These are the interfacial shear strength $\tau_{1,s}$, the elastic slope D_{el} before the slippage and the plastic slope D_{pl} after $\tau_{1,s}$ has been surpassed, namely, $\boldsymbol{\theta} = (\tau_{1,s}, D_{el}, D_{pl})$. In addition, the mechanical models to be investigated are CNT reinforced composite structures and the model predictions $M_i(\boldsymbol{\theta})$ for given realizations of $\boldsymbol{\theta}$ are obtained by solving the multiscale problem with FEM. Also, the collected data d_i involve measurements of deformations at specific locations of the macrostructure.

The inverse problem as postulated by the Bayesian paradigm is provided in eq. (3.14). The interfacial parameters are assumed to be independent to each other, therefore the Bayesian update formulation of eq. (3.15) is adopted here. The discrepancy between the data D and the model predictions is quantified by means of the additive error approach of eq. (3.13). In order to efficiently draw samples from the posterior distribution the Markov Chain Monte Carlo (MCMC) technique is employed based on equation (3.16). Specifically, the Metropolis-Hastings (MH) algorithm [54] is selected in this study, however, more sophisticated versions can also be applied such as the adaptive MH [46] or the transitional MH [18], when challenging posteriors are to be anticipated. The algorithmic procedure of MH is provided in section 3.2. A schematic representation of the algorithmic steps of the MH algorithm for Bayesian update on multiscale systems is depicted in fig. 5.3.

Evidently, the evaluation of the likelihood function $\pi(\boldsymbol{D}|\boldsymbol{\theta}')$ in the expression of $\pi_{post}^{uns}(\boldsymbol{\theta}'|\boldsymbol{D})$ requires the model M to be evaluated at the new parameter values $\boldsymbol{\theta}'$ each time. In the problems studied in this work, model evaluations are particularly time consuming and a direct application of the MH algorithm would be unfeasible. To address this issue, it is proposed in this work to build a surrogate model that will output the

57

relation between the strain/interfacial parameters and the stresses as defined by RVE simulations, which will drastically accelerate the model evaluations at new parameter values $\boldsymbol{\theta}'$.

As a general remark, selecting a candidate density $g(\boldsymbol{\theta}'|\boldsymbol{\theta})$ close to the target density $\pi_{post}(\boldsymbol{\theta}|\boldsymbol{D})$ will lead to more candidates being accepted in the MH algorithm. However, this choice is problem dependent and cannot be determined a priori, since we do not know $\pi_{post}(\boldsymbol{\theta}|\boldsymbol{D})$. Again, the utilization of the aforementioned surrogate will minimize the negative impact a poor choice of candidate density will have on the MH efficiency. The procedure to build the proposed surrogate will be illustrated in the following section.
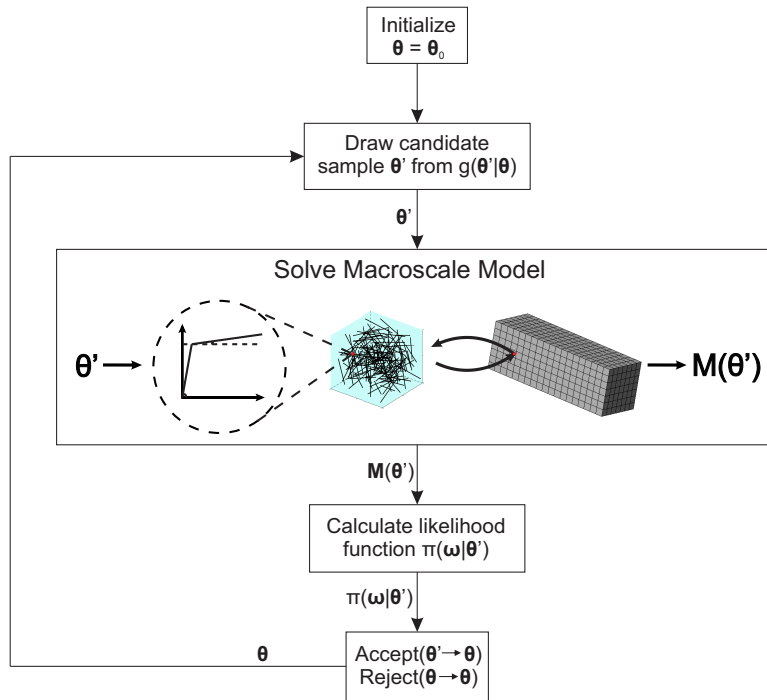
Figure 5.3: Bayesian update on multiscale systems using the Metropolis-Hastings algorithm

5.4 SURROGATE MODELING OF THE RVE USING DEEP NEURAL NETWORKS AND AUTOMATIC DIFFERENTIATION

Needless to say, the previously described procedure to update the parameters that lie in the microscale during a FE^2 algorithm, requires an immense computational effort. To reduce this cost and make the problem computationally tractable, a surrogate modeling

scheme is developed in this section. Specifically, a feed-forward neural network, or FFNN for short (see 4.1), is deployed that will emulate the nonlinear equation in the RVE's homogenization scheme, for different parameters of the CNT/matrix interface. Its input neurons consist of the macroscopic strain vector $\bar{\epsilon}$ along with the three interfacial parameters $\boldsymbol{\theta} = (\tau_{1,s}, D_{el}, D_{pl})$, while the output neurons are the values of macroscopic stress vector $\bar{\sigma}$. Due to the symmetry of the $\bar{\epsilon}$ and $\bar{\sigma}$ tensors, they involve 3 unknown variables for 2D problems and 6 variables for 3D problems.

In order to collect the necessary samples for the FFNN training, a number of RVEs have to be solved for various input values $\bar{\epsilon}$ and $\boldsymbol{\theta}$. Usually these values are chosen in a pseudo-random manner, by using sampling techniques such as Latin Hypercube Sampling (LHS) [87], within some specified ranges. These have to be carefully selected, as small ranges may lead to insufficient information during the solution of the bayesian multiscale system, while large ones may require the evaluation of an unreasonable amount of RVEs in order to accurately train the FFNN. Therefore, some intuition on the mechanical system's behavior from beforehand is required. Also, since the homogenization requires an incremental-iterative algorithm, for each $\boldsymbol{\theta}$, the converged values of $\bar{\epsilon}$ at each increment are considered as input samples.

Additionally, the macroscopic tangent modulus $\overline{C}$ needs to be derived from the surrogate. This can be straightforwardly implemented after the training of the FFNN using Automatic Differentiation (AD) [5]. In essence, AD is a set of techniques developed to numerically evaluate the gradient of a function specified by a computer program. It exploits the fact that every operation performed by the program, no matter how complicated, executes a sequence of elementary arithmetic operations (addition, subtraction, multiplication, division, etc.) and elementary functions (*exp*, *log*, *sin*, *cos*, etc.). By applying the chain rule to these operations, derivatives of arbitrary order can be computed to working precision. In the frame of this work, AD allows us to compute the derivatives of the output $\bar{\sigma}$ with respect to the input $\bar{\epsilon}$. Thus, the elements $\bar{c}_{ij}$ of the macroscopic tangent matrix $\overline{C} = [\bar{c}_{ij}]$ can be obtained as

$$\bar{c}_{ij} = \frac{\partial \bar{\sigma}_{ij}}{\partial h_k} \frac{\partial h_k}{\partial h_{k-1}} \cdots \frac{\partial h_1}{\partial \bar{\epsilon}_{ij}} \tag{5.9}$$

where h_k is the output at the k_{th} hidden layer of the FFNN for input $\bar{\epsilon}_{ij}$.

The whole procedure is summarized in the following steps:

1. Generate N random input samples $\hat{\boldsymbol{q}}_i = [\overline{\boldsymbol{\epsilon}}_i, \boldsymbol{\theta}_i]$, $i = 1, 2, .., N$ within some specified ranges.

2. Solve the nonlinear equation of the RVE for each input vector $\hat{\boldsymbol{q}}_i$ and get the respective output $\hat{\boldsymbol{y}}_i = [\overline{\boldsymbol{\sigma}}_i]$, while keeping all the intermediate increment solutions n_{incr}

3. Choose the FFNN architecture and train it using the $N \times n_{incr}$ pairs of input-output

4. Calculate $\overline{\boldsymbol{C}}$ using AD on the FFNN.
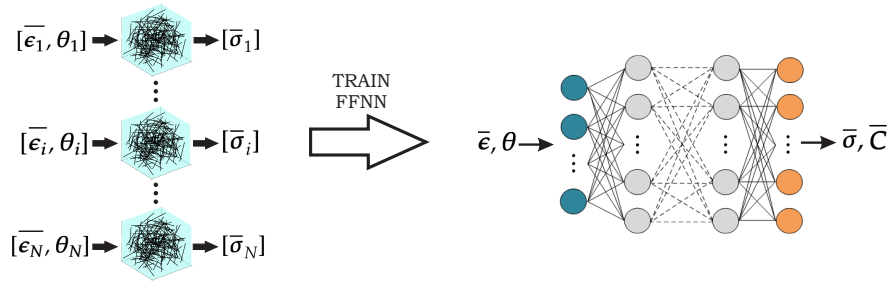
Figure 5.4: Offline (training) procedure
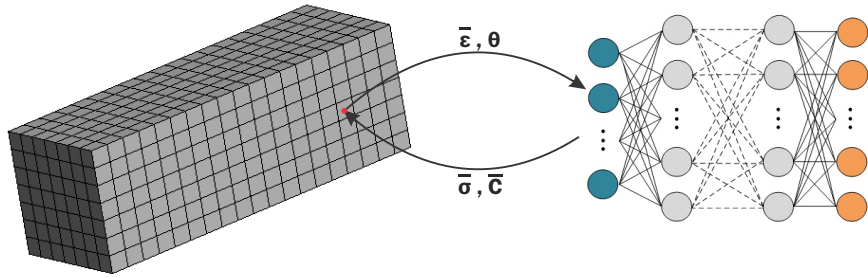
Figure 5.5: Multiscale Bayesian Update, accelerated with surrogate modeling

5.5 NUMERICAL APPLICATIONS

5.5.1 EXAMPLE 1.

For the first application, the fixed composite panel depicted in fig. 5.6 is considered, which is subjected to a bending test. It is made of a polyether-ether-ketone (PEEK)

polymer reinforced with straight CNTs(8,8) as inclusions. The polymer is assumed to be linear elastic with Young's modulus $E = 4GPa$ and Poisson's ratio $v = 0.4$. Each CNT of length $L_0 = 50nm$ is projected to an EBE as described in section 2.1 and after applying eq. (2.6) the axial, bending and torsional stiffness are found to be $EA_{eq} = 694.77nN$ $EI_{eq} = 100.18nN \cdot nm^2$ and $GJ_{eq} = 68.77GPa \cdot nm/rad$ respectively. The total number of CNTs is chosen so as to achieve a volume fraction of 3%.
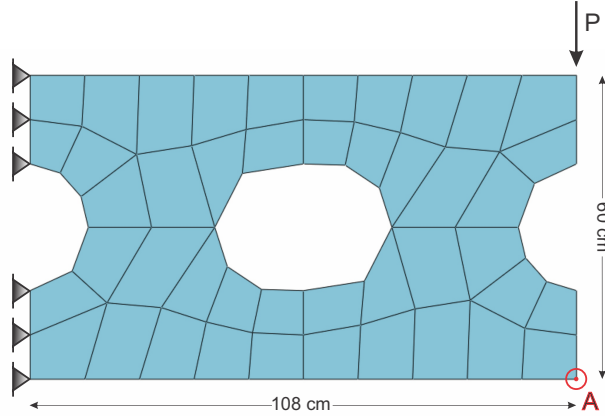
Figure 5.6: Fixed composite panel

The microstructure is characterized by an $100 \times 100 \times 20$ (nm) RVE as the one shown in fig. 5.7, in which the CNTs were placed with a random position and orientation generator. The dimensions of the RVE were selected in order to eliminate the stochastic effects of the geometry, thus, making a single realization representative of the response of the microstructure. Following a similar procedure to the one described in [9] for verifying this condition, 50 realizations of volume elements with the same interfacial parameters and 3%vf, but with different dispersions of CNTs, were generated and solved for the same loading conditions in order to evaluate the corresponding stress vectors. Next, the coefficients of variation (CoV) for each component of the stress vectors were computed. The max value of the CoV of all stress components was found to be 5.78% which was considered low enough to justify the use of a single RVE instead of multiple Stochastic Volume Elements (SVEs).

Subsequently, the RVE is discretized with 100 quadrilateral plane stress (2D) finite elements. In the macroscopic scale, the model is discretized with the same finite element type and for each integration point of the 44 elements, an RVE is assigned.

61
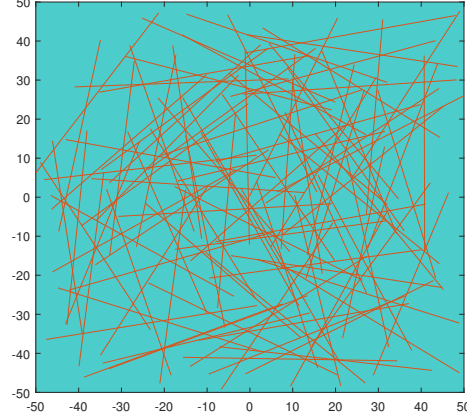
Figure 5.7: Representation of a 2D RVE for volume fraction 3%

The stochastic parameters D_{el}, D_{pl} and $\tau_{1,s}$ in this example are modeled as independent random variables, each having a normal probability density function (PDF) as its prior distribution. It is initially assumed that the parameters follow Gaussian distributions, $\mathcal{N}(\mu, \sigma)$ with μ being the mean and σ the standard deviation. Specifically, the assigned prior distributions are $D_{el} \sim \mathcal{N}(10, 2) \ (GPa/nm)$, $D_{pl} \sim \mathcal{N}(1, 0.2) \ (GPa/nm)$ and $\tau_{1,s} \sim \mathcal{N}(0.1, 0.02) \ (GPa)$. In the macrostructure, a measurement of the vertical displacement in point A was recorded as $u_A = 4.2cm$ with an approximate error of $\epsilon_A = 0.1cm$ for the vertical load $P = 100kN$. Thus the likelihood function is defined as a Gaussian distribution with mean value equal to the measurement u_A and with standard deviation equal to ϵ_A. To ensure that samples with negative values will not be proposed, the selected candidate distribution for each parameter is a trivariate truncated Gaussian on the domain $[0, +\infty) \times [0, +\infty) \times [0, +\infty)$. It is given by the following equation:

$$g(\boldsymbol{\theta}'|\boldsymbol{\theta}_{i-1}) = \frac{g_{\mathcal{N}}(\boldsymbol{\theta}'|\boldsymbol{\theta}_{i-1})}{\Phi(\boldsymbol{\theta}_{i-1})} \tag{5.10}$$

where $g_{\mathcal{N}}(\boldsymbol{\theta}'|\boldsymbol{\theta}_{i-1})$ is a trivariate Gaussian distribution with mean vector $\boldsymbol{\theta}_{i-1}$ and a diagonal covariance matrix $\boldsymbol{\Sigma}$ with entries:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma^2_{\tau_{1,s}} = 0.004^2 & 0 & 0 \\ 0 & \sigma^2_{D_{el}} = 0.4^2 & 0 \\ 0 & 0 & \sigma^2_{D_{pl}} = 0.04^2 \end{bmatrix} \tag{5.11}$$

62

while Φ is the cumulative distribution function of a Gaussian distribution centered at zero with the same covariance matrix $\mathbf{\Sigma}$.

To construct the FFNN surrogate, a set of macroscopic stress-strain pairs are used as training data. These are obtained from solving $N = 300$ RVEs using the computational homogenization scheme. Also, $n_{incr} = 10$ increments were chosen for each nonlinear solution, resulting in a total amount of 3000 pairs of input-output. The adopted strategy in order to select the appropriate ranges for the training samples is described next. For the interfacial parameters $\boldsymbol{\theta}$ the selection is straightforward by establishing a sufficiently broad range around the mean value of each prior distribution. On the other hand for the strain vector $\bar{\boldsymbol{\epsilon}}$ the intention is to gain first some knowledge of the macroscopic model's behavior. To achieve this, a plain parent material i.e. without any CNTs, is solved for this specific model configuration. The computational cost for this analysis is negligible in comparison to the multiscale FE^2 setting. From this solution, the extreme macroscopic strains are obtained and placed as the sampling minimum and maximum limits for $\bar{\boldsymbol{\epsilon}}$. This is due to the fact that any analogous model with the addition of CNTs would only mitigate the extreme strain values, that is, they would be inside of the proposed sample limits. The ranges over which the FFNN was ultimately trained are depicted in table 5.1. The first component of the stress-strain relation for several values of the interfacial parameters $\boldsymbol{\theta}$ is displayed in fig. 5.8, where the RVE was subjected to loading in $\bar{\epsilon}_{11}$ and studied for the cases of a bulk matrix, a composite with fully bonded interfacial behavior and composites with intermediate values of $\boldsymbol{\theta}$.
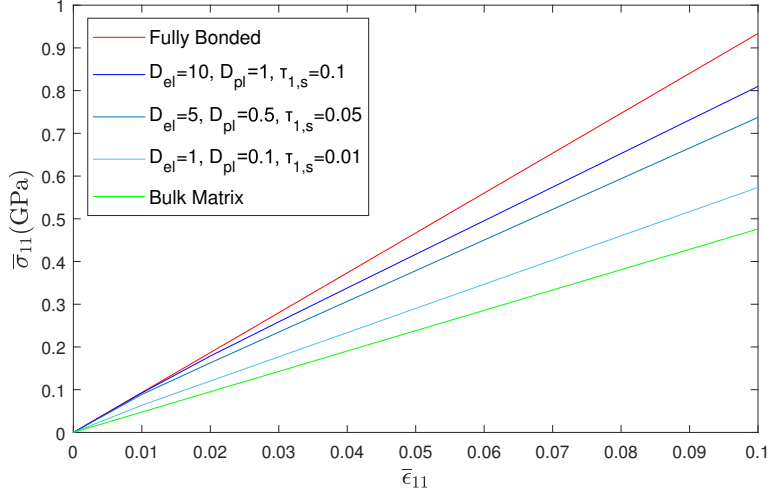
Figure 5.8: Stress-Strain relation for several material cases

Regarding the FFNN architecture, 3 hidden layers of 20 neurons each were considered with the hyperbolic tangent as the activation function. The optimization algorithm used to train the FFNN was the Levenberg-Marquardt algorithm. For the training process the data were split in three subsets, namely the train, test and validation subset with ratios 0.7, 0.15 and 0.15 respectively. The progression of the training is shown in fig. 5.9, which required approximately 300 epochs to reach acceptable levels of accuracy.

Due to the relatively low complexity of the RVE model and the macroscale structure in this particular application, it is feasible to perform BU on both the full and the surrogate model. The comparison between the results obtained from the two models will serve as validation for the accuracy of the surrogate. To this end, a number of 15000 samples are drawn from both model solutions using the MH algorithm and the posterior distribution of each stochastic parameter along with the respective prior is presented in fig. 5.10. As evidenced by these results, the two models are in close agreement. It should be mentioned that in order to obtain 15000 acceptable samples, the samples rejected from the MH algorithm were approximately 8000. Additionally, the initial 500 samples were considered to be burn-in steps and so they were discarded.

To quantify the speed-up achieved by the surrogate modeling approach, table 5.2 presents the computational costs required by both models. BU on the the full scale model took up $189h$, which was attributed to the computationally demanding nature of the FE^2 algorithm. Specifically, for the calculation of the likelihood function of a

candidate sample $\boldsymbol{\theta}'$, a complete multiscale FE^2 model evaluation had to be performed, resulting in as many RVE solutions as the total integration points of the macroscopic system multiplied by the number of iterations for the nonlinear procedure. Hence, this framework inevitably requires millions of RVE solutions even for simple problems such as the one studied here. On the other hand, the surrogate model bypasses the RVE homogenization process and thus requires only $0.8h$ in total. These include the offline cost of collecting the 3000 initial training samples (300 realizations of $\boldsymbol{\theta} \times 10$ increments each) and training the network, as well as the online cost of running the MH algorithm. These results indicate a remarkable speedup, as the surrogate can reduce the cost by two orders of magnitude, while maintaining high accuracy levels.
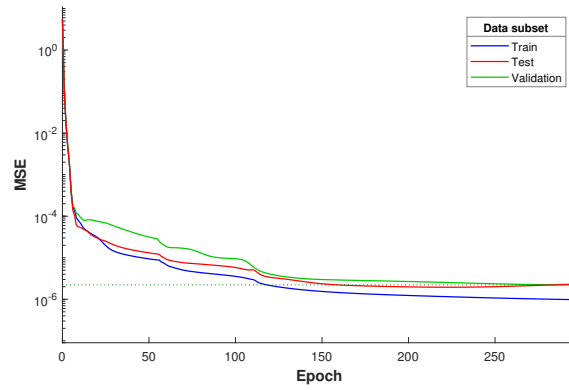
Figure 5.9: Progression of the FFNN training process

	ϵ_{11} $[-]$	ϵ_{12} $[-]$	ϵ_{22} $[-]$	D_{el} [GPa/nm]	D_{pl} [GPa/nm]	$\tau_{1,s}$ [GPa]
min	-0.1	-0.1	-0.1	1	0.1	0.01
max	0.1	0.1	0.1	20	2	0.2

Table 5.1: Input sample ranges for the FFNN training
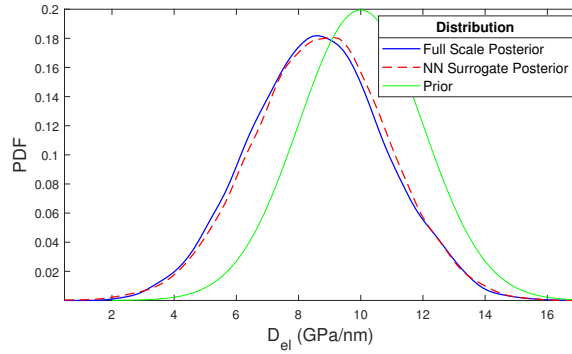
(a) Comparison of the posterior PDFs of D_{el}
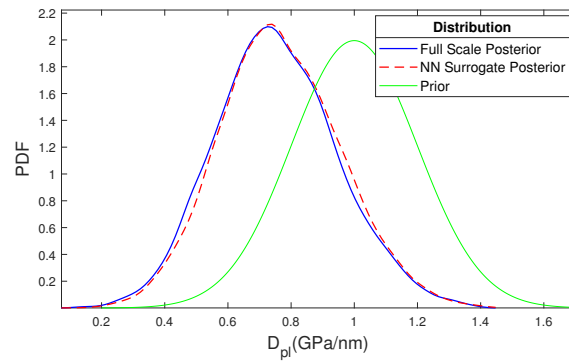
(b) Comparison of the posterior PDFs of D_{pl}
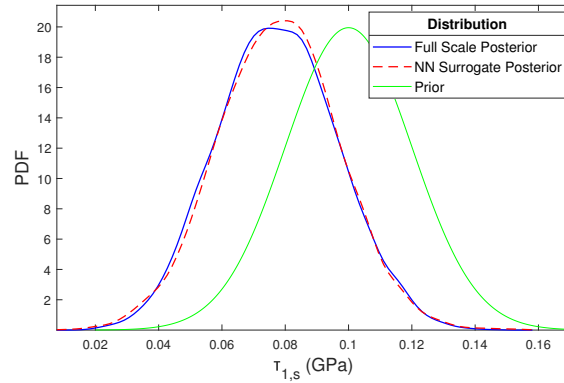
(c) Comparison of the posterior PDFs of $\tau_{1,s}$

Figure 5.10: Comparison of the posterior PDFs of each parameter obtained by performing BU on the full scale model and the surrogate model

66

	Computational time (sec)						
	Offline		Online				Total
Model	FFNN sampling	FFNN training	RVE solution	FFNN call/$\overline{C}$	FE^2	MCMC	
Full scale	-	-	0.022	-	30.5	680850	680850 (189hours)
Surrogate	174	31	-	8e-5	0.115	2588	2793 (0.8hours)

Table 5.2: Computational time of each stage of the algorithm for both models

5.5.2 EXAMPLE 2.

For the second application, a wrench as the one displayed in fig. 5.11 is studied. The structure is fixed at the inner part of the head (blue area) and it is subjected to a uniform pressure at the top side of the tail (purple area). The material properties of both the polymer and the CNTs are the same as in example 1. The number of CNTs is selected so as to attain a weight fraction of 4.5%. The representation of the microscopic scale is done with a $100 \times 100 \times 100$ (nm) RVE as depicted in fig. 5.12, where the CNTs are randomly scattered inside the parent material. The discretization of the macro-structure consists of approximately 50000 tetrahedral elements, while that of the micro-structure of 1000 hexahedral elements. It becomes obvious that the complexity of this model does not permit a direct application of the BU framework and a surrogate modeling approach is the only viable option.

Initially an investigation had to be performed to verify that the selected RVE is in fact representative of the response of the microstructure. The same strategy as in the first example was employed and 50 different solutions from volume elements with varying morphology in the microstructure were acquired. The maximum CoV of the components of the stress vectors in this case was found to be 3.13%, which suggests that a single realization of the selected RVE is sufficient for continuing to the rest of the process.
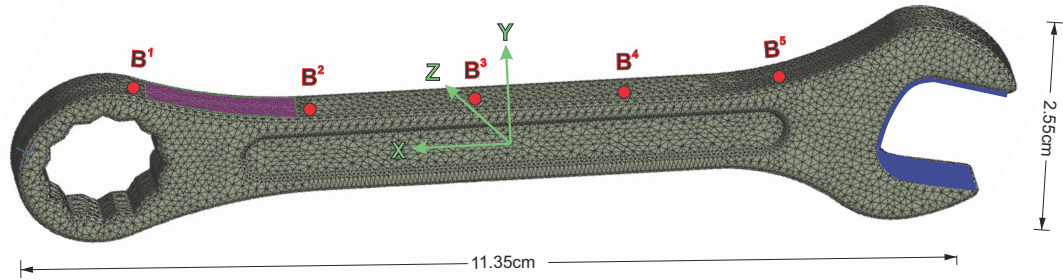
Figure 5.11: Composite Wrench
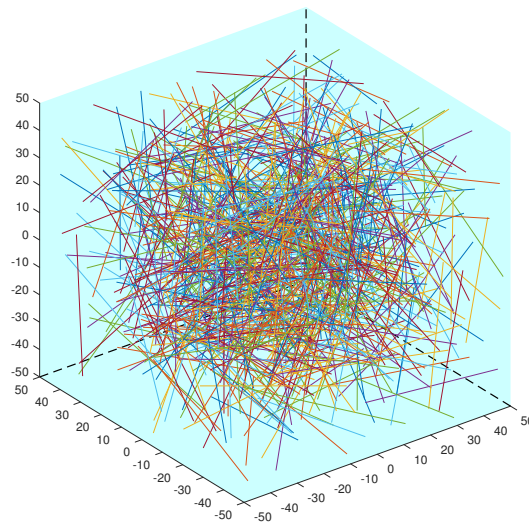
Figure 5.12: Representation of a 3D RVE for volume fraction 4.5%

Similarly to example 1, the normal PDFs $D_{el} \sim Normal(10, 2)$ (GPa/nm), $D_{pl} \sim Normal(1, 0.2)$ (GPa/nm) and $\tau_{1,s} \sim Normal(0.1, 0.02)$ (GPa) were chosen as priors for the (independent) uncertain parameters. The vertical displacements of the structure were measured at the points $\boldsymbol{B} = [B^1, ..., B^5]$ along its torso for two pressure loads, namely, $L_1 = 150KPa$ and $L_2 = 300KPa$, which are uniformly distributed over the purple area

68

shown in fig. 5.11, resulting in 10 observations in total. Thus, the vector of the measured responses is written as $\boldsymbol{D} = [\boldsymbol{d}_1, \boldsymbol{d}_2]$, where $\boldsymbol{D}_1 = [d_1^1, ..., d_1^5]$ and $\boldsymbol{d}_2 = [d_2^1, ..., d_2^5]$ are the responses for L_1 and L_2, as shown in table 5.3.

Although the measurements $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ are assumed to be independent, a statistical dependence is taken into account between the points in $\boldsymbol{B}$ per loading. The auto-covariance matrix that determines this dependency is the 5×5 matrix $\boldsymbol{K} = (k_{ij})$ with k_{ij} given by the autocovariance function $K(s) = \epsilon^2 exp\frac{(-|s|)}{\alpha}$, where s is the distance between the locations B^i and B^j, $\epsilon = 0.02$ cm is the approximate error for each measurement and $\alpha = 1$ cm is the correlation length. Obviously, $\boldsymbol{K}$ is the same for both measurements $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$. Therefore, the likelihood function can be written as:

$$\pi(\boldsymbol{D}|\boldsymbol{\theta}) = \prod_{i=1}^{2} \left((2\pi)^{-\frac{5}{2}} det(\boldsymbol{K})^{-\frac{1}{2}} exp \left(-\frac{1}{2}(\boldsymbol{M}_i(\boldsymbol{\theta}) - \boldsymbol{d}_i)^T \boldsymbol{K}^{-1}(\boldsymbol{M}_i(\boldsymbol{\theta}) - \boldsymbol{d}_i) \right) \right) \quad (5.12)$$

with $\boldsymbol{M}_i(\boldsymbol{\theta})$ being the model predictions for the respective loading case. The candidate density selected in this example is the same trivariate truncated Gaussian used in the previous example.

The initial training data set of the FFNN consists of the macroscopic stress-strain pairs obtained after solving $N = 300$ RVEs for different $\boldsymbol{\theta}$ and considering $n_{incr} = 10$ increments for each nonlinear solution. In this $3D$ case, the input vector consists of the strain $\bar{\boldsymbol{\epsilon}} = [\epsilon_{11}, \epsilon_{12}, \epsilon_{13}, \epsilon_{22}, \epsilon_{23}, \epsilon_{33}]$ and the parameters $\boldsymbol{\theta}$, while the output vector of the stress $\bar{\boldsymbol{\sigma}} = [\sigma_{11}, \sigma_{12}, \sigma_{13}, \sigma_{22}, \sigma_{23}, \sigma_{33}]$. For the FFNN architecture, 3 hidden layers with 20 neurons each and a hyperbolic tangent activation function were chosen. The FFNN was trained with the Levenberg-Marquardt algorithm and for the training process the data were split in ratios 0.7 for the train, 0.15 for the test and 0.15 for the validation subset. The best performance was obtained in epoch 470 with 10^{-6} mean squared error. The ranges over which the FFNN was trained and the progression of the training are shown in table 5.4 and fig. 5.13 respectively. Additionally, fig. 5.14 illustrates how the first component of the stress-strain relation is varied for composite materials with different properties when the RVE is subjected to $\bar{\epsilon}_{11}$ increments.

Loading case	Measurement(mm)				
	B^1	B^2	B^3	B^4	B^5
L_1	-0.15	-6.50	-5.31	-4.35	-5.23
L_2	-0.31	-13.51	-11.03	-9.04	-10.88

Table 5.3: Displacement measurements
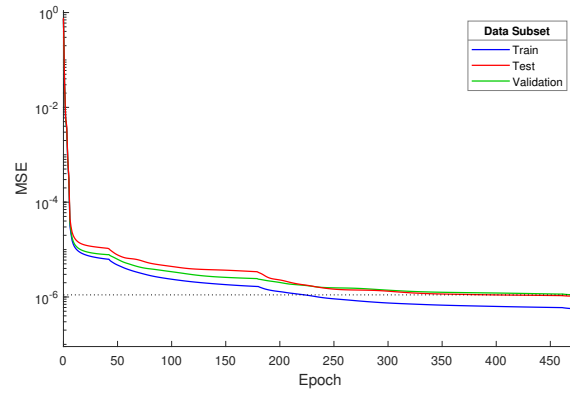
Figure 5.13: Progression of the FFNN training process

	ϵ_{11} [−]	ϵ_{12} [−]	ϵ_{13} [−]	ϵ_{22} [−]	ϵ_{23} [−]	ϵ_{33} [−]	D_{el} [GPa /nm]	D_{pl} [GPa /nm]	$\tau_{1,s}$ [GPa]
min	-0.08	-0.04	-0.04	-0.08	-0.04	-0.04	0.1	0.01	0.001
max	0.08	0.04	0.04	0.08	0.04	0.04	20	2	0.2

Table 5.4: Input sample ranges for the FFNN training
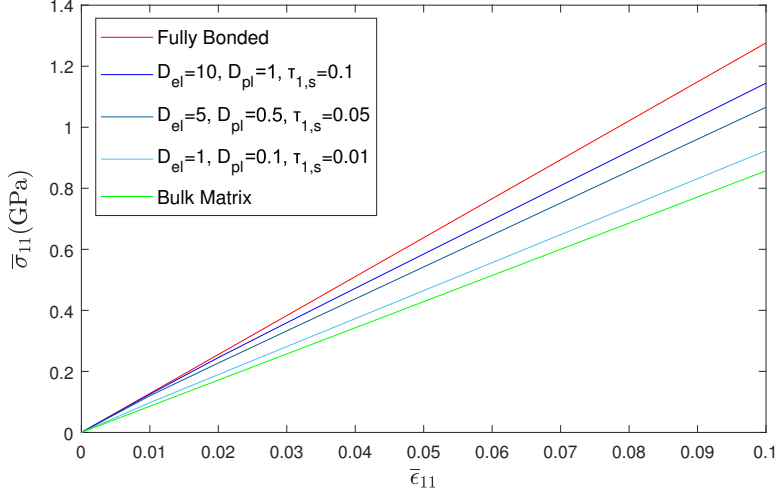
Figure 5.14: Stress-Strain relation for several material cases

Due to the complexity of the model, BU was carried out only on the surrogate model, since the implementation on the full scale model was computationally infeasible. Nevertheless, to confirm the accuracy of the surrogate in this example before proceeding with the BU, the following strategy was adopted. A set of 1000 new input samples $\{(\bar{\boldsymbol{\epsilon}}_j, \boldsymbol{\theta}_j)\}_{j=1}^{1000}$ were randomly generated and the surrogate's accuracy was measured by the errors:

$$err_{\overline{\boldsymbol{C}}} = \frac{\sum_{j=1}^{1000} \left\| \overline{\boldsymbol{C}}^{sur} - \overline{\boldsymbol{C}}^{FEM} \right\|_2}{\sum_{j=1}^{1000} \left\| \overline{\boldsymbol{C}}^{FEM} \right\|_2} \tag{5.13}$$

and

$$err_{\overline{\boldsymbol{\sigma}}} = \frac{\sum_{j=1}^{1000} \left\| \overline{\boldsymbol{\sigma}}^{sur} - \overline{\boldsymbol{\sigma}}^{FEM} \right\|_2}{\sum_{j=1}^{1000} \left\| \overline{\boldsymbol{\sigma}}^{FEM} \right\|_2} \tag{5.14}$$

where $(\cdot)^{sur}$ refers to the macroscopic quantities predicted by the surrogate, $(\cdot)^{FEM}$ to the FEM solutions of the homogenization and $\|\cdot\|_2$ denotes the standard 2-norm. These errors were found to be $err_{\overline{\boldsymbol{C}}} = 1.76\%$ and $err_{\overline{\boldsymbol{\sigma}}} = 0.80\%$, which demonstrated the high accuracy achieved by the surrogate.

Having verified the surrogate's accuracy, the posterior distribution of each component of $\boldsymbol{\theta}$ is obtained using 15000 samples. The acceptance rate was around 60%, resulting in

71

the calculation of approximately 25000 likelihood functions, while the burn-in period was 500 samples. The results are presented in fig. 5.15. These figs. indicate that the initial assumptions on the interface parameters change significantly due to the new measurements in the macrostructure. This is particularly noticeable in the pdf of the shear stress $\tau_{1,s}$, which is found to have smaller mean and less variance than those initially assumed. This outcome is of great importance to material design applications, since $\tau_{1,s}$ is one of the most dominant parameters in the microstructure with regards to its impact on the macroscopic behavior of the structure.

Moreover, in table 5.5 the computational time is illustrated for the surrogate model, while for the full scale model the respective time is estimated with an approximate prediction of the total RVE evaluations. This estimate is based on the fact that a single FE^2 solution on the full scale model required almost 10^6 RVE solutions. Therefore, performing BU on the full scale model would have taken up 105812 years, while the surrogate required 359.9 hrs, or seven orders of magnitude lower cost. Comparing this finding with that of example 1 it leads to the conclusion that the computational acceleration achieved by the surrogate scales accordingly to the RVE's and the macroscopic model's complexity.

	Computational time (hours)						
	Offline		Online				Total
	FFNN sampling	FFNN training	RVE solution	FFNN call/ $\overline{C}$	FE^2	MCMC	
Full scale (Prediction)	-	-	0.033	-	37076	926909667	926909667 (105812 years)
Surrogate	9.83	0.07	-	2.2e-8	0.014	350	359.9 (15days)

Table 5.5: Computational time of each stage of the algorithm for both models
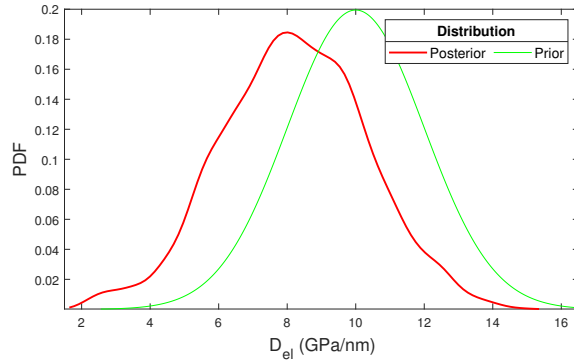
(a) Prior and posterior PDFs of D_{el}
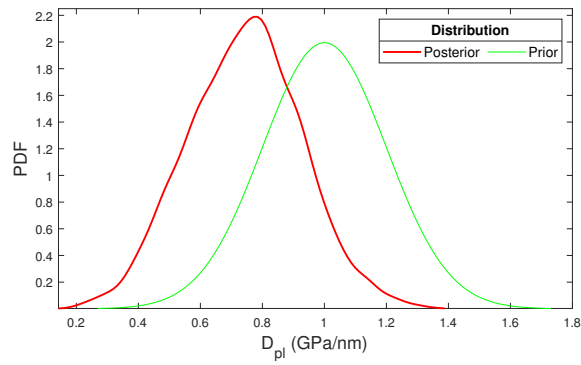
(b) Prior and posterior PDFs of D_{pl}
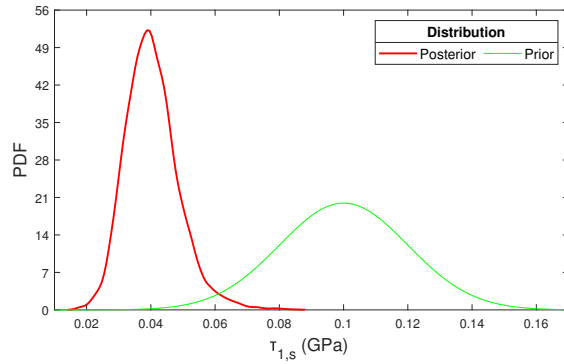
(c) Prior and posterior PDFs of $\tau_{1,s}$

Figure 5.15: Prior and posterior PDFs of each parameter

5.6   CONCLUSIONS

In this work, a methodology for updating the beliefs of the uncertain parameters that lie in the microscale of nanocomposite systems has been proposed. In particular, the bayesian framework with the MCMC technique has been employed on the FE^2 algorithm to learn the parameters of the CNT/polymer interface. The elaborated methodology utilized measurements from the macroscale structure to update the prior beliefs on the nonlinear parameters, rather than expensive and hard-to-obtain microscale measurements. In addition, to tackle the immense computational effort of performing bayesian update on this type of problem, a neural network surrogate was developed in order to replace the nonlinear relation of the homogenization scheme. This surrogate model displayed a high level of accuracy compared to the full scale system solution as well as a remarkable cost reduction. This allowed us to perform BU on complex large-scale problems, which would otherwise be infeasible. The proposed methodology was demonstrated on mechanical problems, where the uncertain parameters were successfully updated in a reasonable computational time. However, this framework can be extended to a diversity of nanocomposite-based applications, such as electrical conductivity or heat transfer.

6

Multiscale analysis of nonlinear systems using a hierarchy of deep neural networks

6.1 INTRODUCTION

This chapter proposes a non-intrusive surrogate modelling strategy, dedicated to FE^N computational homogenization approaches on systems with many scales ($N > 2$). The idea is to employ a sequence of neural networks that represent the hierarchy of the separate scales in the multiscale problem. Each neural network is being trained to learn the physical law at a corresponding length scale of the problem. In a similar manner to the original problem where each finer scale is contained in a coarser scale, neural networks representing fine scales are contained in the DNNs that represent coarser scales. At the end of the training process, a single deep network which emulates the macroscopic behavior by incorporating all physical mechanisms arising at each of the problem's finer scales is derived. This approach takes full advantage of the accuracy and modelling capabilities that FE^N schemes provide, while at the same time overcomes their immense computational requirements. Specifically, the DNNs are tasked with learning parameterized versions of the constitutive law in each scale, which allows us to model a wide range of possible material behaviors. This is accomplished by augmenting the input layer of the DNNs with the set of the uncertain material parameters. With

this approach, each DNN incorporates the uncertain behavior that comes from all the previous scales and ultimately a macroscopic constitutive response that encapsulates all this information is obtained through the final DNN. In turn, this DNN can be applied as a surrogate of the material in any macrostructural system and for various multi-query problems (e.g. sensitivity analysis, optimization, Bayesian inference).

The elaborated methodology is demonstrated on the analysis of a large-scale building made of CNT-reinforced concrete. This particular structural system is modeled as a four-scale problem consisting of (i) carbon nanotube-reinforced cement paste at the microscale, (ii) reinforced cement mortar at a fine mesoscale level, (iii) reinforced concrete at a coarse mesoscale level and (iv) the macroscopic structural system. The composite material is characterized by different nonlinear constitutive laws at each scale. The solution of the full multiscale problem is attained by using a FE^4 scheme at a reasonable computational time by virtue of the elaborated surrogate modelling setup. In turn, this allows us to perform laborious sensitivity analyses in order to assess the uncertainty in the material parameters and its propagation to the macroscopic structural response.

The rest of this chapter is organized as follows. Section 6.2 provides the constituents of the CNT-reinforced concrete multiscale model developed in this work and the solution framework based on a FE^4 algorithm. In Section 6.3 the reinstatement of the nested solution scheme with the utilization of neural networks as surrogate modelling technique is illustrated. Section 6.4 presents the application of global sensitivity analysis in the context of nested multiscale analyses. Section 6.5 demonstrates a numerical example to test the efficiency of the aforementioned methodology.

6.2 Multiscale model and solution procedure

This section illustrates a composite material modelling paradigm, which requires performing computational homogenization on multiple scales. In particular, a multiscale model of CNT-reinforced concrete is presented that consists of cement paste, CNTs and aggregate particles. Each of these constituents and their interactions are defined on appropriate length scales, as depicted in fig. 6.1. This approach will enable us to accurately model the interaction of each constituent with the cement paste and eventually study the macroscopic behavior of real-life structures composed of the composite. The details on the implementation aspects of this approach are presented next.

6.2.1 Cement paste material model

The cement paste is assumed to be fully hydrated and it is represented by a homogeneous material. For the cement material a phenomenological relation is adopted. Specifically, the transition into the inelastic regime is defined by the Drucker-Prager plasticity model [27]. The yield function that describes this transition is of the following form:

$$\Phi = J_2(\boldsymbol{\sigma}) + \eta p(\boldsymbol{\sigma}) - \xi c \tag{6.1}$$

NANOSCALE MICROSCALE FINE MESOSCALE COARSE MESOSCALE

(a) (b) (c) (d)
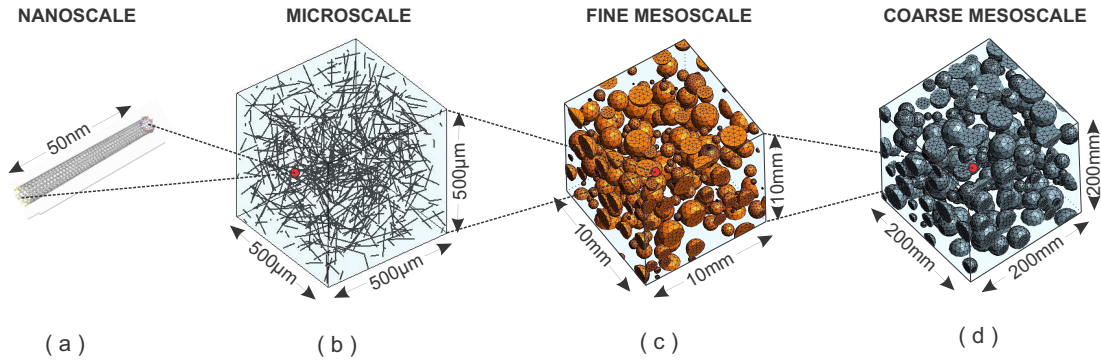
Figure 6.1: Multiscale model of the CNT-reinforced concrete material: (a) the nanoscale represents each individual CNT, (b) the microscale represents the cement paste reinforced with CNTs, (c) the fine mesoscale with the addition of sand particles represents the CNT-reinforced mortar, (d) the coarse mesoscale with the addition of high volume aggregates e.g. gravel represents the CNT-reinforced concrete.

where J_2 is the second deviatoric invariant of the stress tensor σ, p is the hydrostatic pressure, c is the cohesion, while η and ξ are constants that are chosen in a way to approximate an equivalent Mohr-Coulomb yield surface. Here these constants are used for the approximation of the outer edges of the Mohr-Coulomb criterion by the DP surface and have the form:

$$\eta = \frac{6\sin\phi}{\sqrt{3}(3 - \sin\phi)} \quad \& \quad \xi = \frac{6\cos\phi}{\sqrt{3}(3 - \sin\phi)} \tag{6.2}$$

with ϕ being the friction angle.

The classical strain rate decomposition is applied:

$$\dot{\boldsymbol{\epsilon}} = \dot{\boldsymbol{\epsilon}}_e + \dot{\boldsymbol{\epsilon}}_p \tag{6.3}$$

with initial conditions:

$$\boldsymbol{\epsilon}(t_0) = \boldsymbol{\epsilon}_e(t_0) + \boldsymbol{\epsilon}_p(t_0), \tag{6.4}$$

where the decomposition of the strain tensor into the elastic component ϵ_e and the plastic component ϵ_p is performed.

According to the plastic flow rule, the rate of the plastic strain is postulated as:

$$\dot{\boldsymbol{\epsilon}}_p = \dot{\lambda}\frac{\partial \Psi}{\partial \boldsymbol{\sigma}} \tag{6.5}$$

where λ is the plastic multiplier and Ψ is the plastic flow potential. Here, non-associative plastic flow is assumed, hence the plastic potential function takes a different form than the yield function ($\Phi \neq \Psi$). The potential function that was chosen here is:

$$\Psi = J_2(\boldsymbol{\sigma}) + \bar{\eta}p(\boldsymbol{\sigma}) \tag{6.6}$$

where $\bar{\eta}$ has a similar meaning with the approximation constants of Φ and is given by:

$$\bar{\eta} = \frac{6\sin\psi}{\sqrt{3}(3 - \sin\psi)} \tag{6.7}$$

with ψ being the dilatancy angle.

Furthermore, an isotropic linear hardening model is considered, hence $c(\bar{\boldsymbol{\epsilon}}_p)$ is expressed as:

$$c(\bar{\boldsymbol{\epsilon}}_p) = c_0 + H\bar{\boldsymbol{\epsilon}}_p \tag{6.8}$$

where c_0 is the initial cohesion, i.e. the initial yield stress and H is the hardening modulus.

Additionally, the rates of the plastic strain and the equivalent plastic strain are given by:

$$\dot{\boldsymbol{\epsilon}}_p = \left(\frac{1}{2\sqrt{J_2(\boldsymbol{\sigma})}}(\boldsymbol{\sigma} - \frac{1}{3}trace(\boldsymbol{\sigma})I) + \bar{\eta}\right)\dot{\lambda} \quad \& \quad \dot{\bar{\boldsymbol{\epsilon}}}_p = \xi\dot{\lambda} \tag{6.9}$$

The preceding equations form an optimization problem that is completed with the addition of the Kuhn-Tucker loading/unloading conditions:

78

$$\Phi \leq 0, \quad \dot{\lambda} \geq 0, \quad \Phi\dot{\lambda} = 0 \tag{6.10}$$

6.2.2 MICROSTRUCTURAL REPRESENTATION OF CNT-REINFORCED CEMENT PASTE

Every CNT is described as an assembly of covalent bonds developed between the carbon atoms (C-C bond) in hexagonal shapes. The molecular structural mechanics approach (MSM), which is detailed in section 2.1, transforms the molecular structure of CNTs to continuum structural models in the form of space frames. This transition is achieved with the relations in eq. (2.5). While the structural representation offers a clear and precise transition from molecular mechanics, the system's complexity remains a challenge due to the multitude of Degrees Of Freedom (DOFs) inherent in each space frame simulating a CNT. To streamline the integration of CNTs into microstructure modeling, further simplification is necessary. This simplification mirrors the approach of the MSM method, whereby each space frame structure is projected onto an equivalent beam element (EBE) using the relationships outlined in eq. (2.6).

To simulate the interaction between the CNTs and the cement paste, the embedding finite element technique via the fully bonded formulation of section 2.2 is implemented. A number of EBEs are scattered inside the microstructure RVE with random position and orientation as depicted in fig. 6.1 until a prescribed weight fraction for the CNTs is achieved. Since EBEs are 2-D elements there is no issue of overlapping with each other and additional checks for their valid placement are not required. The dimensions of the cement paste RVE are $(500\mu\mu)^3$.

6.2.3 MESOSTRUCTURAL REPRESENTATIONS OF REINFORCED MORTAR AND CONCRETE

A Fuller grading curve is applied in order to determine the size distribution of the aggregate particles in the mesoscale, as follows:

$$P(d) = (\frac{d}{d_{max}})^n \quad , \quad 0 \leqslant P(d) \leqslant 1 \tag{6.11}$$

where d is the aperture diameter of a specific sieve, d_{max} is the maximum diameter of the coarse aggregates and $P(d)$ is the cumulative percentage of the aggregates that pass

through d, while n is a constant that defines the shape of the curve.

The volume fraction of the aggregates within each of the segments $[d_s, d_{s+1}]$ is then estimated as:

$$V_a^{[d_s, d_{s+1}]} = \frac{P(d_s) - P(d_{s-1})}{P(d_{max}) - P(d_{min})} V_a^{tot} \tag{6.12}$$

where V_a^{tot} is the total volume fraction of the aggregates inside the specimen.

An algorithm is implemented for the placing of the aggregates inside the RVEs by utilizing eq. 6.12. For simplicity, all the inclusions are modeled as spheres. Other options, however, such as polygonal shapes are also applicable in a straightforward manner. Each newly added inclusion has to satisfy the condition of non-overlapping with the already existing particles. Nevertheless, the inclusions are allowed to intersect with the boundaries of the RVE since periodic topology is adopted here. For more details on the placing algorithm the reader is referred to [138]. Since the mesoscale is described by two separate scales, the diameter bounds d_{max} and d_{min} are assigned independently for each one of them.

In the fine mesoscale, the aggregates which are essentially sand particles have a minimum diameter $d_{min}^{fm} = 0.1mm$ and a maximum diameter $d_{max}^{fm} = 2mm$. The particles are assumed to behave linearly elastic with a Young's modulus $E = 30GPa$ and Poisson ratio $\nu = 0.22$ and their volume fraction inside the mortar is taken as 20%. The interaction between the two phases of the material is direct i.e. without an Interfacial Transition Zone (ITZ). The dimensions of the mortar RVE are $(10mm)^3$. A sample of the fine mesoscale representation of the mortar is included in fig. 6.1.

The coarse mesoscale model represents the reinforced concrete material. In this representation, the range of the diameters of the coarse aggregates are $d_{min}^{cm} = 2mm$ and $d_{max}^{cm} = 20mm$. Similarly to the sand particles, the coarse aggregates are taken as linearly elastic and no ITZ between the different phases exist. For these aggregates the Young's Modulus is $E = 45GPa$ and the Poisson ratio is $\nu = 0.22$. The volume fraction of the aggregates in this scale is chosen as 40%. The dimensions of the concrete RVE are $(200mm)^3$. A coarse mesoscale specimen is depicted as part of the multiscale model in fig. 6.1.

6.2.4 Computational homogenization scheme for the CNT-reinforced concrete

The macroscopic system is described by a homogeneous material with a constitutive relation formulated by the coupling of all the finer scales as these are described in sections 6.2.2 and 6.2.3. This coupling is achieved with a standard computational homogenization scheme as the one described in section 1.3. The information is initially propagated through the localization rule. Starting from a macroscopic integration point, strain driven boundary conditions are assigned on the RVE of the following scale. The resulted strains on the integration points of the matrix material are then used to apply the localization step onto the RVE that describes the next scale. This procedure continues step-wise between each pair of consecutive scales until the finest scale RVE has been reached, which is the microscale in the present model. Then, beginning from the finest scale, the microscale RVE is solved for the boundary conditions that resulted from the previous procedure and the homogenization step is applied to propagate the homogenized stresses and tangent moduli onto the coarser scale. The homogenization step is repeated between successive scale RVEs until the process has returned to the macroscale. It is important to note here that this procedure is performed in the frame of a nonlinear analysis, requiring internal iterations until the whole multiscale system is in a converged state.

The nested multiscale scheme of CNT-reinforced concrete is visualized in fig. 6.2. The connection of the four scales and the solution of the system is done according to the procedure described in sections 1.6 and 1.7.
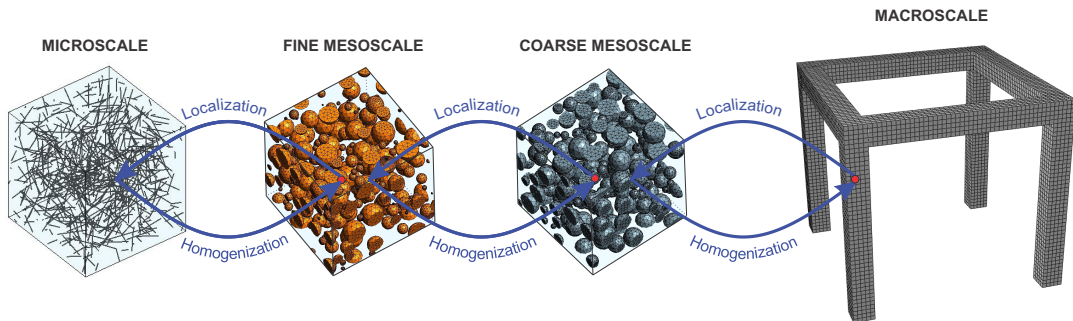
Figure 6.2: Hierarchical scale coupling scheme. The material behavior at an integration point in the macroscale is obtained from the analysis of a coarse mesoscale RVE. However, this requires the analysis of a fine mesoscale RVE at each integration point in the coarse mesoscale, which, in turn, requires the analysis of a microscale RVE at each of the corresponding integration points.

As pointed out in the previous sections, the task of performing this FE^4 scheme is directly associated to a prohibitive computational cost. This problem is tackled with the use of hierarchical surrogate models, as described in the following section.

6.3 SURROGATE MODELLING OF THE FE^N SCHEME UTILIZING A HIERARCHY OF NEURAL NETWORKS

6.3.1 MAIN CONCEPT

The powerful approximation capabilities of DNNs have been validated from numerous applications, but also, from the mathematical viewpoint, though the universal approximation theorem [23, 24, 59]. Therefore, DNNs are considered capable of capturing complex material behavior. In this framework, a dedicated surrogate modelling strategy is presented herein for mitigating the cost of FE^N solution schemes in the context of computational homogenization. The key idea is to use DNNs [44, 39], 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 DNN that encapsulates the material behavior at all previous scales.

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

- A sequential homogenization procedure is implemented according to eqs. (1.23) and (1.25) 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 DNN are defined. The input involves the components of the strain vector at the upper scale $\epsilon^{(2)}$, while the the output is the stress vector $\sigma^{(2)}$ at the end of the homogenization procedure. If we further assume that the solution of the system is affected by a vector of parameters $\alpha^{(1)}$ that characterize the material's constitutive law at the finest scale, then these parameters are also regarded as input to the DNN.

- The tangent constitutive matrix $C^{(2)}$ at a certain strain state $\epsilon^{(2)}$, required for the Newton-Raphson iterations, is effortlessly computed using Automatic Differentiation (AD) [5]. By using differentiable activation functions in the DNN, such as the logistic or hyperbolic tangent function, the DNN becomes a differentiable function.

82

This technique allows for the computation of the derivatives of the output $\boldsymbol{\sigma}^{(2)}$ with respect to the input $\boldsymbol{\epsilon}^{(2)}$ by applying the chain rule on the DNN and, thus, the elements $c_{ij}^{(2)}$ of the macroscopic tangent matrix $\boldsymbol{C}^{(2)} = [c_{ij}^{(2)}]$ are obtained as:

$$c_{ij}^{(2)} = \frac{\partial \sigma_{ij}^{(2)}}{\partial h_k} \frac{\partial h_k}{\partial h_{k-1}} \cdots \frac{\partial h_1}{\partial \epsilon_{ij}^{(2)}} \tag{6.13}$$

where h_k is the output at the k_{th} hidden layer of the DNN for input $\epsilon_{ij}^{(2)}$.

- After the successful training and validation, the finest scale DNN (scale 1) 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 DNN is built using $(\boldsymbol{\epsilon}^{(3)}, \boldsymbol{\alpha}^{(2)}, \boldsymbol{\alpha}^{(1)})$ as input and $\boldsymbol{\sigma}^{(3)}$ as output. To simplify notation we write the input $(\boldsymbol{\epsilon}^{(3)}, \boldsymbol{\alpha}^{(2)}, \boldsymbol{\alpha}^{(1)})$ as $(\boldsymbol{\epsilon}^{(3)}, \hat{\boldsymbol{\alpha}}^{(3)})$, where, in the general case, $\hat{\boldsymbol{\alpha}}^{(s+1)} = (\boldsymbol{\alpha}^{(s)}, \cdots, \boldsymbol{\alpha}^{(1)})$ is the augmented parametric vector. Again, the tangent matrix $\boldsymbol{C}^{(3)}$ is readily available through AD. It is important to note here that the DNN of the second scale also involves the material parameters $\boldsymbol{\alpha}^{(1)}$ 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, where it ultimately results in a single final DNN that incorporates all the information from the lower scales and constitutes the surrogate model of the composite material's behavior.

All aforementioned steps are summarized in the following algorithm:

Output: The trained DNN^N that emulates the composite material's behavior
$\boldsymbol{\sigma}^{(M)}(\boldsymbol{\epsilon}^{(M)}, \hat{\boldsymbol{\alpha}}^{(M)})$

for *scale* $s = 1$ **to** N **do**

$\quad$ **if** $s = 1$ **then**

$\quad\quad$ Assign the matrix material of RVE^s as the original matrix material of the composite;

$\quad$ **else**

$\quad\quad$ Assign the matrix material of RVE^s as the previously trained DNN^{s-1};

$\quad$ **end**

$\quad$ Choose the samples ranges in the input space $[\boldsymbol{\epsilon}^{(s+1)}, \hat{\boldsymbol{\alpha}}^{(s+1)}]$, where $\hat{\boldsymbol{\alpha}}^{(s+1)}$ is the augmented parametric vector $\hat{\boldsymbol{\alpha}}^{(s+1)} = [\boldsymbol{\alpha}^{(s)}, ..., \boldsymbol{\alpha}^{(1)}]$;

end

for *sample* $j = 1$ **to** K **do**

$\quad$ Generate a sample $[\boldsymbol{\epsilon}_j^{(s+1)}, \hat{\boldsymbol{\alpha}}_j^{(s+1)}]$ using the preferred sampling technique;

$\quad$ Solve the RVE^s and obtain the relation $[\boldsymbol{\epsilon}_j^{(s+1)}, \hat{\boldsymbol{\alpha}}_j^{(s+1)}] - [\boldsymbol{\sigma}_j^{(s+1)}]$;

$\quad$ Select the architecture and train DNN^s with the acquired RVE solutions;

end

Algorithm 2: Offline DNN training

To better illustrate the proposed surrogate modelling strategy let us focus on the 4-scale example of CNT-reinforced concrete presented in section 6.2. First, the DNN^{micro} is trained to substitute the RVE of the microscale, which consists of cement paste and CNTs, using $\left(\boldsymbol{\epsilon}^{fm}, \boldsymbol{\alpha}^{\mu}\right)$ as input and $\boldsymbol{\sigma}^{fm}$ as output. Next, the $DNN^{meso,fine}$ which represents the fine mesoscale RVE (cement mortar) is trained on pairs $\left(\boldsymbol{\epsilon}^{cm}, \boldsymbol{\alpha}^{fm}, \boldsymbol{\alpha}^{\mu}\right)$, where now DNN^{micro} is considered as the matrix material along with the fine aggregates as the inclusions. The process is repeated one more time for the $DNN^{meso,coarse}$ of the coarse mesoscale RVE, which is the final DNN that encapsulates all previous DNNs and represents the constitutive relation for the macroscale problem. The training procedure is illustrated in fig. 6.3 and the final outcome of the process in figs. 6.4a-6.4b.
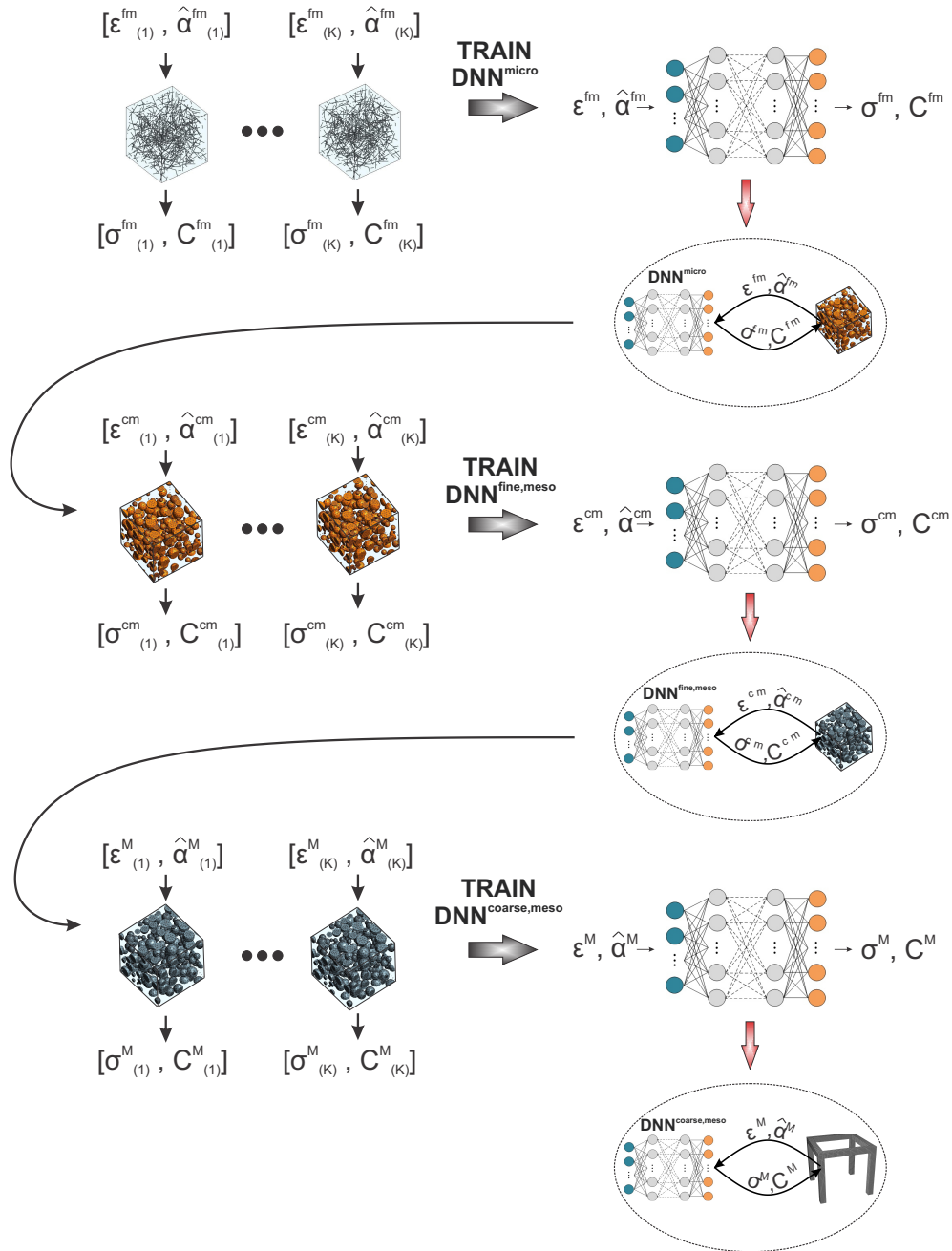
Figure 6.3: Training procedure according to the proposed strategy. Staring from the microscale, a neural network, DNN^{micro} is trained to emulate the stress-strain behavior of the microscale RVE. This DNN^{micro} is used in exchange of the host material in the fine mesoscale and, next, $DNN^{fine,meso}$ 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 $DNN^{coarse,meso}$ successfully encapsulates the overall composite material's behavior.
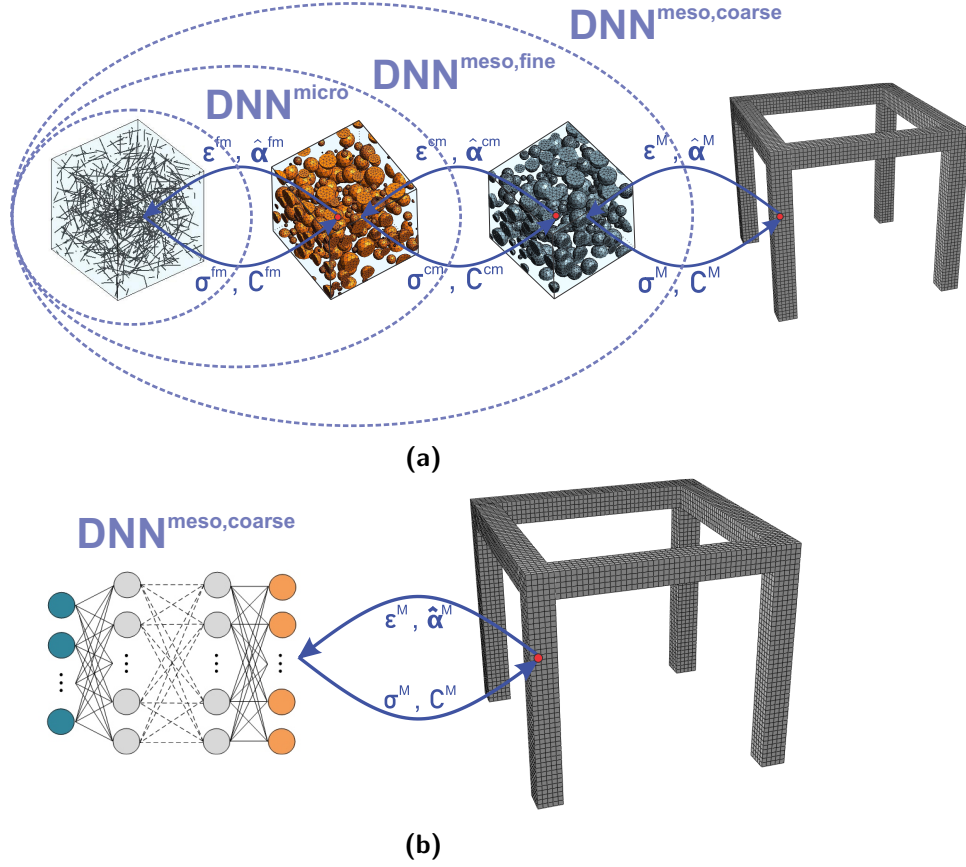
Figure 6.4: (a) Hierarchical construction of the surrogate DNN model and (b) Scale coupling using the final $DNN^{meso,coarse}$ as constitutive model

6.3.2 IMPLEMENTATION ASPECTS

Regarding the implementation aspects of the surrogate modelling strategy, there are a few key points that require further elaboration: To obtain the data for each DNN training, a series of RVEs need to be solved for various combinations of strains and parameter values. Starting from the finest (first) scale, a number of detailed FE simulations are required for the preparation of the training data set. These calculations do not involve any direct FE^2 simulations, since they are performed at the lowest scale. However, moving a step up to the 2nd scale, the training of the $DNN^{meso,fine}$ requires a FE^2 analysis between the first and second scale. Based on the premise that the first DNN^{micro} is well-trained and proven capable of accurately capturing the material behavior at the finest scale,

86

the DNN^{micro} is employed in lieu of the costly FE analyses. Then, this procedure iterates one more time in order to establish $DNN^{meso,coarse}$, which will represent the constitutive behavior of the composite material. With this approach, the hierarchical training of the successive DNNs at each scale is performed in a computationally efficient and tractable manner.

Furthermore, the problems studied here all assume small deformations and consider materials with isotropic behavior on every scale. Thus, the strain and stress tensors are symmetric and can be described with six independent variables, while the size of $\boldsymbol{\alpha}$ may vary on each scale. The accumulation of parameters $\boldsymbol{\alpha}^{cm}, \boldsymbol{\alpha}^{fm}, \boldsymbol{\alpha}^{\mu}$ as the training procedure progresses might lead to high dimensional input spaces for the DNNs that will impede an accurate training. Therefore, the sampling technique used for generating the input data plays a critical role in the procedure and sampling with simple Monte Carlo methods tend to not give representative results in the high dimensional space. More efficient choices include space filling techniques such as Latin Hypercube Sampling (LHS) [87] or quasi-Monte Carlo methods [119]. Additionally, attention must be paid on the range of the sampling in order for the DNN to be able to predict the full response of a system during every possible solution, while also avoid over-training the DNN for extreme values with infinitesimal chance to be reproduced in the online solution and increase unnecessarily the offline computational time.

As it is generally the case for all the data-driven material models, it should be highlighted that with the proposed data-driven material modelling strategy, the DNNs that learn to imitate the constitutive law of each scale are ultimately limited by the level of detail that has been applied to describe the original materials' behavior. These may include the material formulation for each separate phase of the composite or the interaction mechanisms between them. Nevertheless, this is merely a choice that has to be made based on the desired material accuracy in exchange of computational complexity and time. The core of the proposed scheme is invariant to different preferences for the material formulations.

6.4 Variance-based sensitivity analysis on composite structures

In this section, the variance-based SA framework, presented in section 3.4, is applied in order to assess how sensitive are the structures made up of CNT-reinforced concrete with respect to material parameters in the microstructure. Herein, SA is performed

on the Drucker-Prager plasticity model parameters, namely, the friction ϕ, the dilation ψ, the initial cohesion c_0 and the hardening modulus H, in a CNT-reinforced concrete building. The model parameters are collectively denoted as $\hat{\boldsymbol{\alpha}}^M = (\phi, \psi, c_0, H)$. Then, SA will give us a measure of how uncertainty propagates across the various scales of the model and affects the macroscopic response. In this setting, we can consider our model $\mathcal{M} := f(\phi, \psi, c_0, H) := f(\hat{\boldsymbol{\alpha}}^M)$ to be a quantity of interest of the structure in the macroscale, such as a displacement at a given location, which is obtained after solving the structural problem with the FE^4 scheme. The aim is to evaluate the sensitivity indices S_i and total indices S_{T_i} for $i = 1, 2, 3, 4$ corresponding to the four model parameters ϕ, ψ, c_0, H, respectively. To do this, we need to generate $K \cdot (d + 2)$ instances of the parameter vector $\{\hat{\boldsymbol{\alpha}}_i^M\}_{i=1}^{K(d+2)}$ according to a specified probability distribution, with K typically being of the order of $10^3 \div 10^5$ to produce reliable estimates of the sensitivity indices and $d = dim(\hat{\boldsymbol{\alpha}}^M) = 4$. For each of these instances, the macroscale problem is solved to obtain the responses $\{\mathcal{M}_i\}_{i=1}^{K \cdot (d+2)}$ and compute $f(\boldsymbol{R}_{(j)})$, $f(\boldsymbol{Q}_{(j)})$ and $f(\boldsymbol{Q}_{\boldsymbol{R},(j)}^{(i)})$, according to equations (3.30) and (3.31).

Taking into account the complexity of the model, it becomes apparent that this type of analysis would be computationally unrealizable with a direct application of the FE^4 solution scheme. However, the surrogate modelling scheme developed in the previous section can be employed to accelerate the repeated model evaluations for various instances of the microscopic model parameters ϕ, ψ, c_0, H. A schematic representation of the SA procedure in a general DNN enhanced multiscale analysis is depicted in fig. 6.5. According to this, $K \cdot (d + 2)$ instances of the parameter vector $\{\hat{\boldsymbol{\alpha}}_i^M\}_{i=1}^{K \cdot (d+2)}$ are generated following a specified probability distribution and for each of these instances the macroscale problem is solved to obtain the responses $\{\mathcal{M}_i\}_{i=1}^{K(d+2)}$. The $K \cdot (d + 2)$ model evaluations necessary for the convergence of the SA are expected to require affordable computational time since the final DNN, namely the DNN^N, has replaced the costly direct FE^N scheme. After the samples have been collected and the SA has been performed, critical deductions of the multiscale model parameters can be done by observing how they affect the model output.
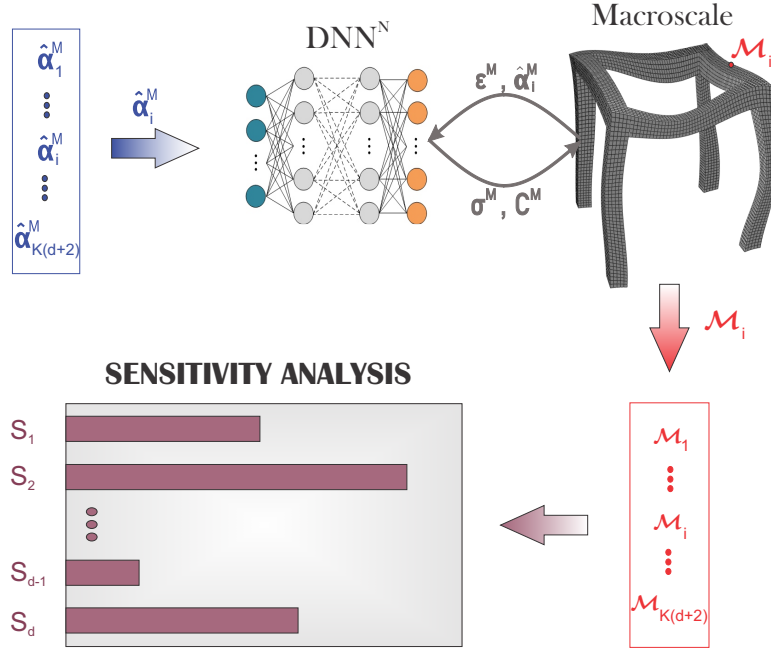
Figure 6.5: Sensitivity analysis on a multiscale system accelerated by the proposed surrogate model. The macroscale problem is solved for $K \cdot (d+2)$ instances of the parameter vector $\hat{\boldsymbol{\alpha}}^M$. The corresponding model output $\{\mathcal{M}_i\}_{i=1}^{K(d+2)}$ can be efficiently computed by virtue of the proposed surrogate modelling strategy. Then, the sensitivity indices of the material parameters can be straightforwardly computed via statistical post-processing.

6.5 NUMERICAL APPLICATION

In this section we showcase the potential of the proposed framework in a CNT-reinforced concrete structural model as the one presented in section 6.2. The structure to be analyzed in the macroscale is a two-storey CNT-reinforced concrete building as depicted in fig. 6.6. Each story has a height of $2.80m$, while both slabs have dimensions of $2.50m \times 2.50m \times 0.30m$. Both the columns and the beams have a rectangular cross-section with dimensions $0.30m \times 0.30m$. The four columns are fixed on the ground. At the edge of each storey, a lateral line load with magnitude $P_L = -5.28kN/cm$ is applied on the z-direction while a dead pressure load of magnitude $P_D = -0.16kN/cm^2$ is assigned on each slab. The macroscale model is discretized with 2308 hexahedral elements, which in turn results in 12276 DOFs.
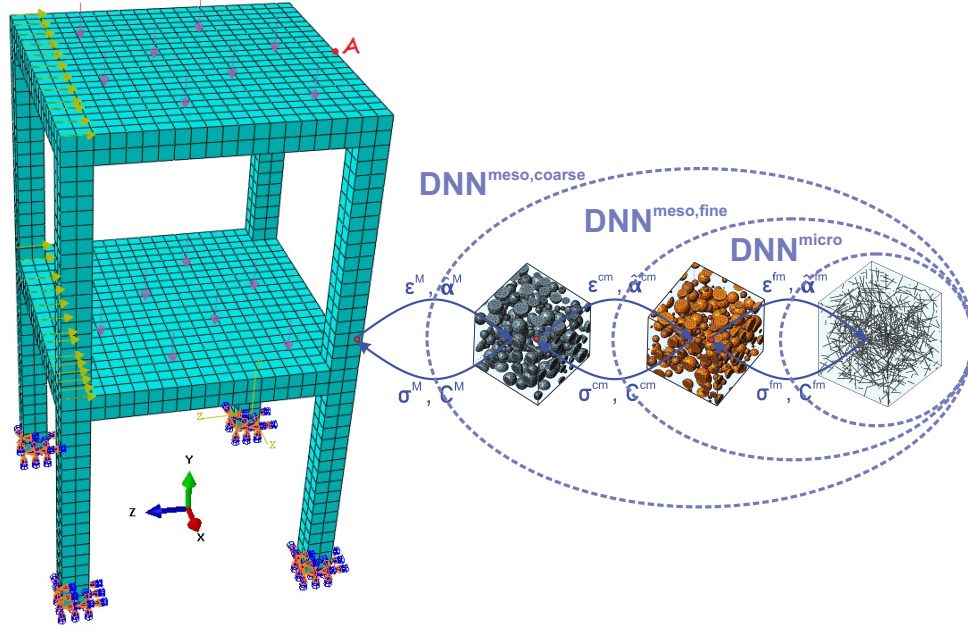
Figure 6.6: Multiscale model of a CNT-reinforced structure. Each Gauss point of the macroscale is associated with a constitutive law delivered by the DNN hierarchy

6.5.1 OFFLINE PROCEDURE

The initial cement paste matrix material is formed according to the Drucker-Prager(DP) plasticity model presented in section 6.2.1. A series of parameters are selected, based on which the SA of the macrostuctural response will be performed. As mentioned previously, these parameters are the ones that characterize the cement paste material, namely the friction angle ϕ, the dilation angle ψ, the initial cohesion c_0 and the hardening modulus H. To propagate the influence of these parameters on the macroscale behavior, a ten parameter input vector for the DNN^i of each scale i defined in section 6.3.1 is used, containing the strains $\boldsymbol{\epsilon}^i = [\epsilon_{xx}^i, \epsilon_{yy}^i, \epsilon_{zz}^i, \epsilon_{xy}^i, \epsilon_{xz}^i, \epsilon_{yz}^i]$ and the parametric vector $\boldsymbol{\alpha} = [\phi, \psi, c_0, H]$. The output vectors are the six stress components $\boldsymbol{\sigma}^i = [\sigma_{xx}^i, \sigma_{yy}^i, \sigma_{zz}^i, \sigma_{xy}^i, \sigma_{xz}^i, \sigma_{yz}^i]$ resulting from the homogenization equations given by eq. 1.25.

As stated in the implementation aspects section, the choice of the input sample ranges play a crucial role on the efficiency of the proposed approach. For the material parameters $\boldsymbol{\alpha}$ this choice is straightforward, as their statistical properties are known

from their specifications. However, in order to set up reasonable and functional upper and lower bounds for the strain vectors $\boldsymbol{\epsilon}^i$, even though the offline procedure is decoupled from the macrostructure, a proper choice requires some prior information about the macromodel's behavior under certain conditions e.g. loading, boundaries.

A strategy is proposed here in order to pre-select the strain input ranges efficiently as follows: An inverse hierarchical procedure is implemented starting from the macroscale and then moving to the finer scales. A DP model is assigned as the material of the macro-model, where the parametric vector $\boldsymbol{\alpha}$ of the material is chosen in a way that would yield the most critical response, i.e. the largest displacements on any point of the structure under the loading conditions specified by the model configuration. To achieve that, in the examined problem all the components of $\boldsymbol{\alpha}$ are given their lowest values based on the material specifications as imposed by the problem. After the solution of the macroscale problem is completed, a pair of strain vectors is formed, namely the $\boldsymbol{\epsilon}_{max}^M$ and $\boldsymbol{\epsilon}_{min}^M$. Each of these strains contain respectively the maximum and minimum values of each component of the strain vector that resulted from the analysis. Since the loading of the following scale, i.e. the coarse mesoscale, is defined by the macroscopic strains via the localization rule, the $\boldsymbol{\epsilon}_{max}^M$ and $\boldsymbol{\epsilon}_{min}^M$ vectors represent the extreme candidate loading conditions. Therefore these vectors are used to define the maximum and minimum strain sampling ranges for the $DNN^{coarse,meso}$ training. In a similar manner, by solving the coarse mesoscale model with two loading cases imposed by the $\boldsymbol{\epsilon}_{max}^M$ and $\boldsymbol{\epsilon}_{min}^M$ we can track the maximum and minimum strains that resulted from the corresponding analysis. These extreme strain vectors are used to form the $\boldsymbol{\epsilon}_{max}^{cm}$ and $\boldsymbol{\epsilon}_{min}^{cm}$ respectively and can be readily applied as the strain sampling ranges for the $DNN^{fine,meso}$ training. This procedure is repeated one last time in order to determine the $\boldsymbol{\epsilon}_{max}^{fm}$ and $\boldsymbol{\epsilon}_{max}^{fm}$ that will represent the strain bounds for the DNN^{micro} training. As there is no involvement of any multiscale analysis during this process, the computational cost for this pre-selection step involves minimal FE solutions at each scale, hence the additional computational time is insignificant.

It has to be highlighted that the produced DNN from each step of the offline procedure i.e. the training of the respective scale, represents a sub-multiscale system. The number of scales that define this sub-multiscale system depends on all the preceding scales which have already been used to train the previous DNNs. The optimal way to validate the training quality of each of these intermediate DNNs is to compare them with direct solutions of detailed sub-multiscale systems. However, for subsystems with more than 2

scales, this is computationally intractable. Hence, a different strategy has to be employed for the cases where a $\geq FE^3$ analysis is required for the direct simulation. Since the validation of the first and the second scale require FE and FE^2 analyses respectively, these can be performed straightforwardly and without any further adjustments. For validating each higher scale s (with $s > 2$) instead of solving FE^s systems, the already trained DNN^{s-2} can be used as the matrix material of the $s-1$ scale. This enables the verification of all the higher DNNs by performing FE^2 analyses within a reasonable time. Obviously this approach is based on the premise that the preceding DNNs are accurately trained and therefore extensive testing is performed to verify the accuracy of each DNN starting from the finest scale.

The number of RVE simulations in order to collect the training data for each scale was determined by a concise trial and error procedure. For the number of samples that were ultimately selected, we took into account both the factors of the total computation burden for the RVE solutions and the achieved level of accuracy of the NNs.

Microscale

The first scale to be analyzed is the microscale described by the CNT-reinforced cement paste as presented in section 6.2.2. The bounds of the DNN^{micro} input parameters are given in table 6.1, where for the strain vector these bounds were chosen based on the aforementioned sampling strategy . For the DNN^{micro} training, 500 RVEs were solved and the non-linear solution for each of them included 20 increments, thus ultimately collecting a total of 10000 solutions. The DNN^{micro} architecture was selected after a trial-and-error process and consisted of 3 hidden layers with 20 neurons each. The weights and biases were initialized with the Nguyen-Widrow algorithm, while the optimization of the DNN^{micro} during the back-propagation was done by means of the Levenberg–Marquardt algorithm. For the training process, the data were split in three subsets, namely the train, test and validation subset with ratios 0.7, 0.15 and 0.15 respectively. A maximum limit of 3000 epochs were appointed. The total CPU time required for the collection of the data and the DNN^{micro} training where 28.4 hours and 5.8 hours, respectively. The progression of the training and the training error histogram of the fully trained DNN^{micro} are shown in fig. 6.7.

	ϵ^{cm}_{xx} [$-$]	ϵ^{cm}_{yy} [$-$]	ϵ^{cm}_{zz} [$-$]	ϵ^{cm}_{xy} [$-$]	ϵ^{cm}_{xz} [$-$]	ϵ^{cm}_{yz} [$-$]	ϕ [deg]	ψ [deg]	c_0 [GPa]	H [GPa]
min	-0.0152	-0.0201	-0.0138	-0.0303	-0.0280	-0.0242	10	10	0.01	1
max	0.0396	0.0352	0.0382	0.0540	0.0463	0.0389	40	40	0.05	5

Table 6.1: Input sample ranges for the DNN of the microscale

| (a) | (b) |
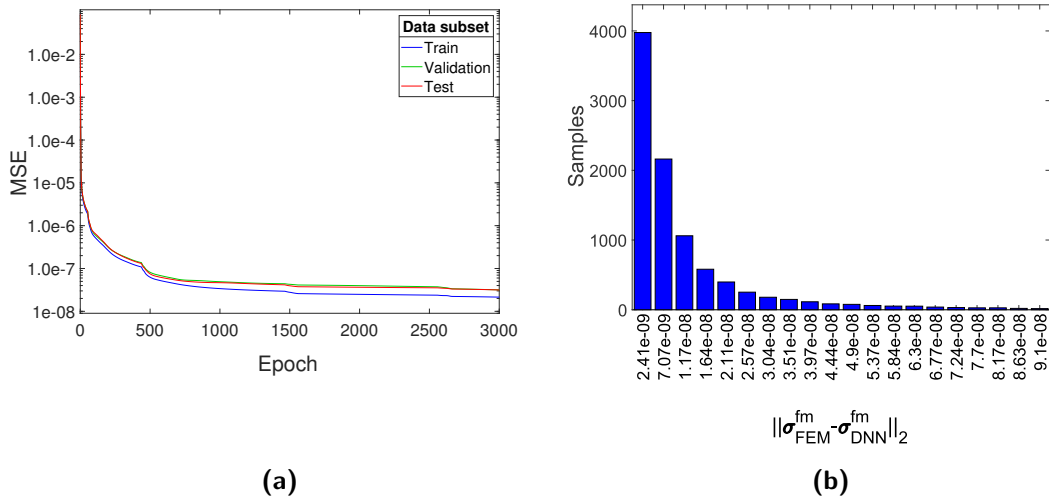
Figure 6.7: (a). Progression of mean squared error during the DNN^{micro} training (b). Training error histogram of DNN^{micro} on fully trained stage

To further assess the quality of the trained DNNs the normalized $L2$ norm of the discrepancy between 50 direct numerical solutions and surrogate solutions is employed. Specifically, the norm is computed for the differences of both the homogenized stress vectors and the tangent modulus matrices as:

$$err_{\boldsymbol{C}^{(i)}} = \frac{\sum_{j=1}^{50} \left\| \boldsymbol{C}^{(i)}_{FEM,j} - \boldsymbol{C}^{(i)}_{DNN,j} \right\|_2}{\sum_{j=1}^{50} \left\| \boldsymbol{C}^{(i)}_{FEM,j} \right\|_2} \tag{6.14}$$

and

93

$$err_{\boldsymbol{\sigma}^{(i)}} = \frac{\sum_{j=1}^{50} \left\| \boldsymbol{\sigma}_{FEM,j}^{(i)} - \boldsymbol{\sigma}_{DNN,j}^{(i)} \right\|_2}{\sum_{j=1}^{50} \left\| \boldsymbol{\sigma}_{FEM,j}^{(i)} \right\|_2} \tag{6.15}$$

where $(\cdot)_{DNN}$ refers to the macroscopic quantities predicted by the surrogate, $(\cdot)_{FEM}$ to the FEM solutions of the homogenization and $\|\cdot\|_2$ denotes the standard L_2 norm. The superscript (i) denotes the scale, hence $(i) = fm$ in the present case. These errors were found to be $err_{\boldsymbol{C}^{fm}} = 3.28\%$ and $err_{\boldsymbol{\sigma}^{fm}} = 2.16\%$, which demonstrated the high accuracy achieved by the surrogate.

A comparison between the stress-strain relations for the two analyses was performed next. Specifically two randomly selected loading cases were produced from the strain space of $\boldsymbol{\epsilon}^{fm}$ that was used for the DNN^{micro} training. These cases are used as tests for validating the DNN^{micro} predictions against the corresponding detailed FEM analyses. Figs. 6.8a and 6.8b present the comparison of the stress-strain relation obtained with FEM and the DNN^{micro} surrogate for the strain vectors $\boldsymbol{\epsilon}_1^{fm} = [2.21, -4.99, -1.98, -3.53, -4.08, -3.14] \cdot 10^{-3}$ and $\boldsymbol{\epsilon}_2^{fm} = [0.47, 4.73, 2.15, 1.98, -2.84, -4.76] \cdot 10^{-3}$ respectively, while also these cases were tested for three different sets of parametric vectors $\boldsymbol{\alpha}$, namely $\boldsymbol{\alpha}_1 = [\phi = 10°, \ \psi = 10°, \ c_0 = 0.01 GPa, \ H = 1 GPa]$, $\boldsymbol{\alpha}_2 = [\phi = 25°, \ \psi = 25°, \ c_0 = 0.03 GPa, \ H = 3 GPa]$, $\boldsymbol{\alpha}_3 = [\phi = 40°, \ \psi = 40°, \ c_0 = 0.05 GPa, \ H = 5 GPa]$. As can be seen from figs. 6.8a and 6.8b the curves are in good agreement for all cases.

PARAMETRIC CASE	$\alpha_1=[\phi=10,\ \psi=10,\ c_0=0.01,\ H=1]$	$\alpha_2=[\phi=25,\ \psi=25,\ c_0=0.03,\ H=3]$	$\alpha_3=[\phi=40,\ \psi=40,\ c_0=0.05,\ H=5]$
DIRECT FEM ANALYSIS	▭	◯	▷
DNN PREDICTION	▬▬	▬·▬·	▬ ▬
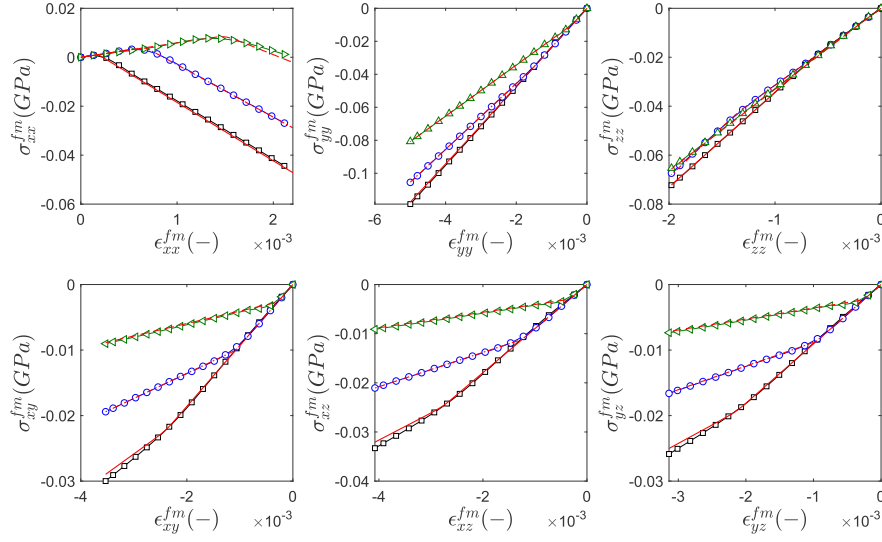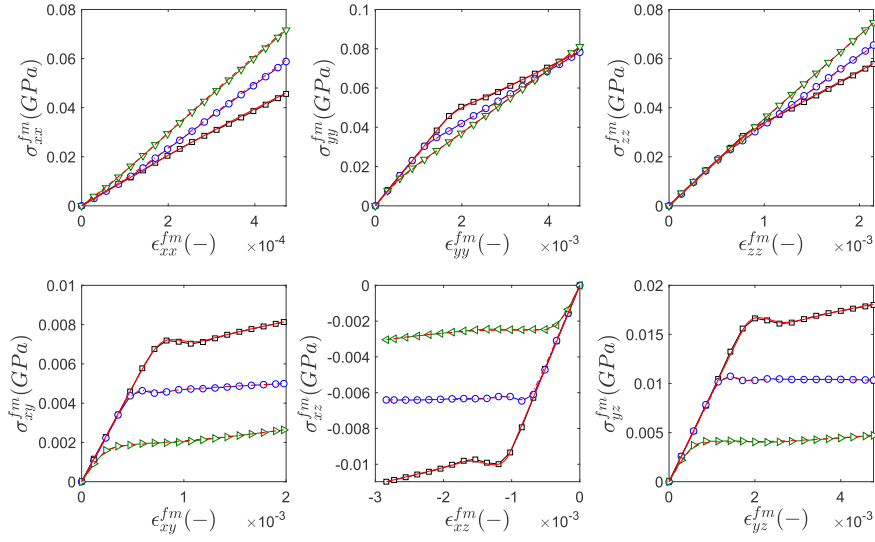
Figure 6.8: Comparison of DNN^{micro} and FEM produced strain-stress curves for the two randomly sampled loading cases (a) $\boldsymbol{\epsilon}_1^{fm} = [2.21, -4.99, -1.98, -3.53, -4.08, -3.14] \cdot 10^{-3}$ and (b) $\boldsymbol{\epsilon}_2^{fm} = [0.47, 4.73, 2.15, 1.98, -2.84, -4.76] \cdot 10^{-3}$.

Once the homogenized response of the microscale has been learned accurately by the DNN^{micro}, the transition on the following scale can be done. This scale describes the fine mesoscale of the reinforced mortar as mentioned in section 6.2.3. Similarly to the preceding scale, the data obtained from 500 RVE solutions, with 20 time steps each, were used for the training of the second DNN, namely the $DNN^{fine,meso}$. The choice of the hyperparameters and the general architecture of the $DNN^{fine,meso}$ is identical to the previous one. For the collection of the training data the computational cost was 25.2 hours, while for the DNN training 4.2 hours. The performance of the training and the error histogram are illustrated in fig. 6.9. Also the $L2$ norm of the discrepancy of the stresses and tangent moduli between direct FE^2 fine mesoscale analyses and $DNN^{fine,meso}$ predictions, for 50 RVE resolutions were found to be $err_{\boldsymbol{C}^{cm}} = 2.64\%$ and $err_{\boldsymbol{\sigma}^{cm}} = 1.89\%$.

| | ϵ_{xx}^{cm} | ϵ_{yy}^{cm} | ϵ_{zz}^{cm} | ϵ_{xy}^{cm} | ϵ_{xz}^{cm} | ϵ_{yz}^{cm} | ϕ | ψ | c_0 | H |
	$[-]$	$[-]$	$[-]$	$[-]$	$[-]$	$[-]$	[deg]	[deg]	[GPa]	[GPa]
min	-0.0087	-0.0192	-0.0081	-0.0200	-0.0167	-0.0183	10	10	0.01	1
max	0.0197	0.0200	0.0181	0.0316	0.0241	0.0245	40	40	0.05	5

Table 6.2: Input sample ranges for the DNN of the microscale
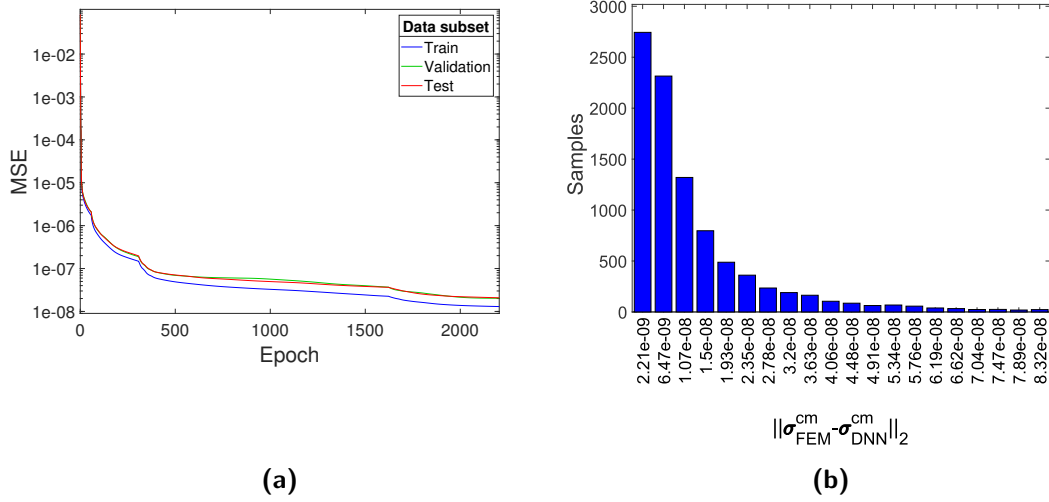
Figure 6.9: (a). Progression of mean squared error during the $DNN^{meso,fine}$ training (b). Training error histogram of $DNN^{meso,fine}$ on fully trained stage

Next, the stress-strain curves of direct FE^2 solutions in the fine mesoscale were compared with the predictions of the $DNN^{fine,meso}$. These tests were performed for the same strain vectors $\boldsymbol{\epsilon}_1^{cm} = \boldsymbol{\epsilon}_1^{fm} = [2.21, -4.99, -1.98, -3.53, -4.08, -3.14] \cdot 10^{-3}$ and $\boldsymbol{\epsilon}_2^{cm} = \boldsymbol{\epsilon}_2^{fm} = [0.47, 4.73, 2.15, 1.98, -2.84, -4.76] \cdot 10^{-3}$ as in the microscale and are shown in figs. 6.10a and 6.10b. Given the closeness of the results, we can confidently deduce that the quality of the $DNN^{fine,meso}$ is highly satisfactory.

PARAMETRIC CASE	α_1=[ϕ=10, ψ=10, c_0=0.01, H=1]	α_2=[ϕ=25, ψ=25, c_0=0.03, H=3]	α_3=[ϕ=40, ψ=40, c_0=0.05, H=5]
DIRECT FEM ANALYSIS	—□—	—○—	—▷—
DNN PREDICTION	▬▬	▬·▬·▬	▬ ▬ ▬
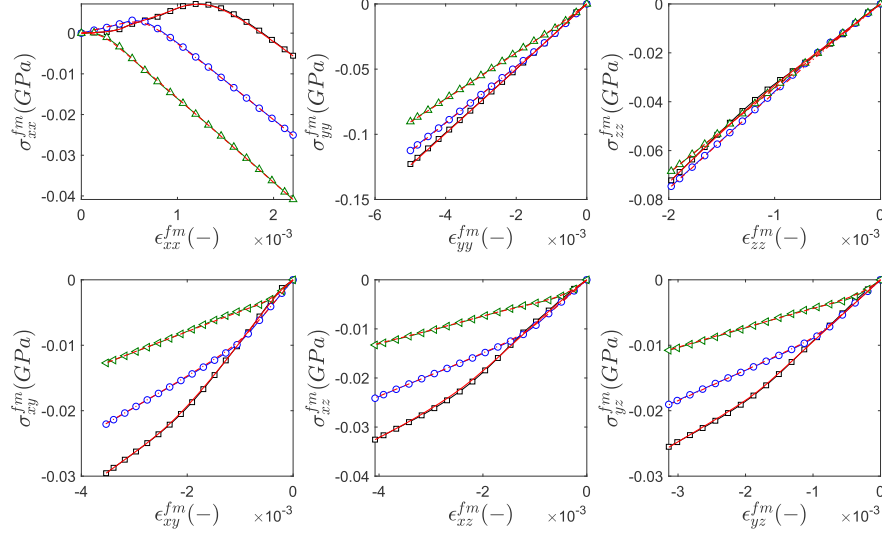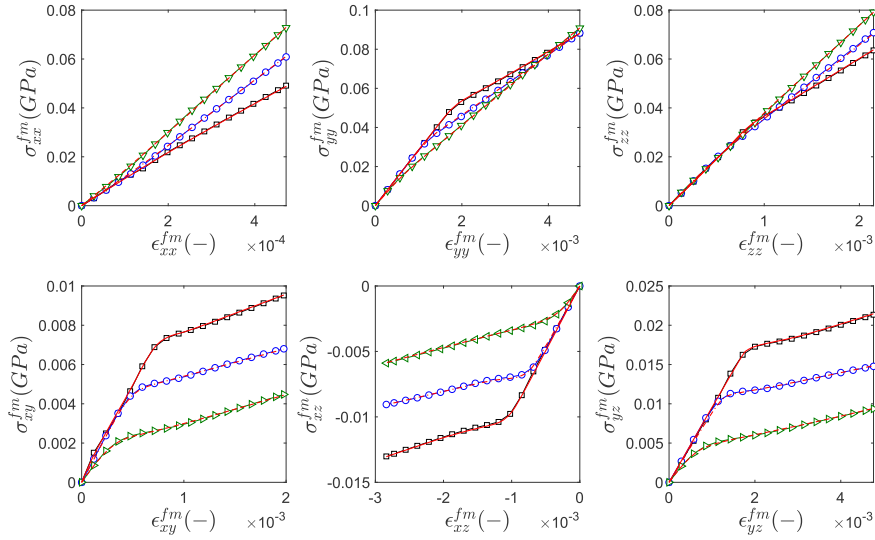
Figure 6.10: Comparison of DNN^{micro} and FEM produced strain-stress curves for the two randomly sampled loading cases (a) $\epsilon_1^{fm} = [2.21, -4.99, -1.98, -3.53, -4.08, -3.14] \cdot 10^{-3}$ and (b) $\epsilon_2^{fm} = [0.47, 4.73, 2.15, 1.98, -2.84, -4.76] \cdot 10^{-3}$.

Lastly, the procedure is repeated for the coarse mesoscale of reinforced concrete which is described in section 6.2.3. The $DNN^{coarse,meso}$, which has the same formulation with the previously constructed DNNs, is trained with the 10000 samples acquired from the RVE solutions, by following the same sampling strategy. Approximately 23.4 hours were employed for the collection of the samples and 1.5 hours for the training of the DNN. The progression of the training and the error histogram are reported in fig. 6.11. Since the direct FEM analyses in this case require FE^3 system solutions, these are computationally intractable and the strategy described in section 6.2.4 is employed. To be able to perform comparisons between direct simulations and the respective $DNN^{coarse,meso}$ predictions we replace the microscale with the accurately trained DNN^{micro}. With this, direct FEM solutions are enabled by solving FE^2 systems on the coarse mesoscale. By means of that, the $L2$ norm of the homogenized variables between FE^2 analyses and $DNN^{coarse,meso}$ predictions were estimated as $err_{\boldsymbol{C}M} = 2.23\%$ and $err_{\boldsymbol{\sigma}M} = 1.58\%$.

	ϵ_{xx}^{cm}	ϵ_{yy}^{cm}	ϵ_{zz}^{cm}	ϵ_{xy}^{cm}	ϵ_{xz}^{cm}	ϵ_{yz}^{cm}	ϕ	ψ	c_0	H
	[−]	[−]	[−]	[−]	[−]	[−]	[deg]	[deg]	[GPa]	[GPa]
min	-0.0056	-0.0171	-0.0052	-0.0163	-0.0089	-0.0153	10	10	0.01	1
max	0.0113	0.0135	0.0077	0.0163	0.0089	0.0135	40	40	0.05	5

**Table 6.3:** Input sample ranges for the DNN of the microscale
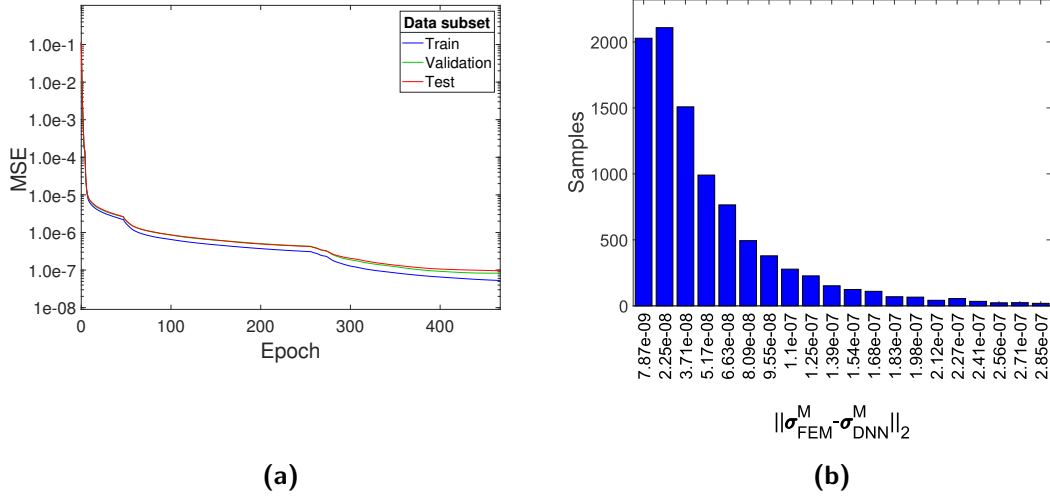
Figure 6.11: (a). Progression of mean squared error during the $DNN^{meso,coarse}$ training (b). Training error histogram of $DNN^{meso,coarse}$ on fully trained stage

The loading modes defined by the strain vectors $\boldsymbol{\epsilon}_1^M = \boldsymbol{\epsilon}_1^{fm} = [2.21, -4.99, -1.98, -3.53, -4.08, -3.14] \cdot 10^{-3}$ and $\boldsymbol{\epsilon}_2^M = \boldsymbol{\epsilon}_2^{fm} = [0.47, 4.73, 2.15, 1.98, -2.84, -4.76] \cdot 10^{-3}$ are again utilized in the direction of validating the performance of the $DNN^{coarse,meso}$. The comparison of the strain-stress curves between the detailed FE^2 resolutions and the $DNN^{coarse,meso}$ predictions is depicted in figs. 6.12a and 6.12b. Based on these figs., the excellent predictive capabilities of the $DNN^{coarse,meso}$ are evident.

PARAMETRIC CASE	$\alpha_1=[\phi=10,\ \psi=10,\ c_0=0.01,\ H=1]$	$\alpha_2=[\phi=25,\ \psi=25,\ c_0=0.03,\ H=3]$	$\alpha_3=[\phi=40,\ \psi=40,\ c_0=0.05,\ H=5]$
DIRECT FEM ANALYSIS	□	○	▷
DNN PREDICTION	▬	▪▪▪	▬ ▬
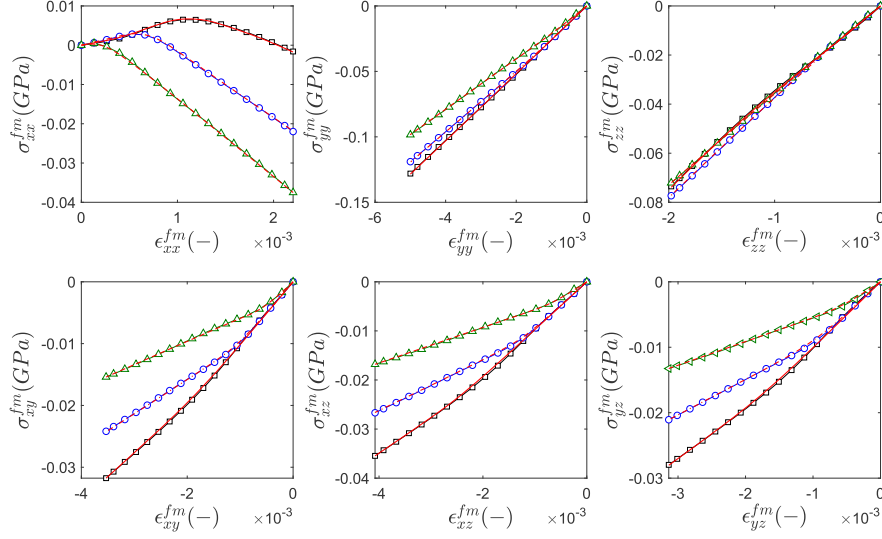
(a)
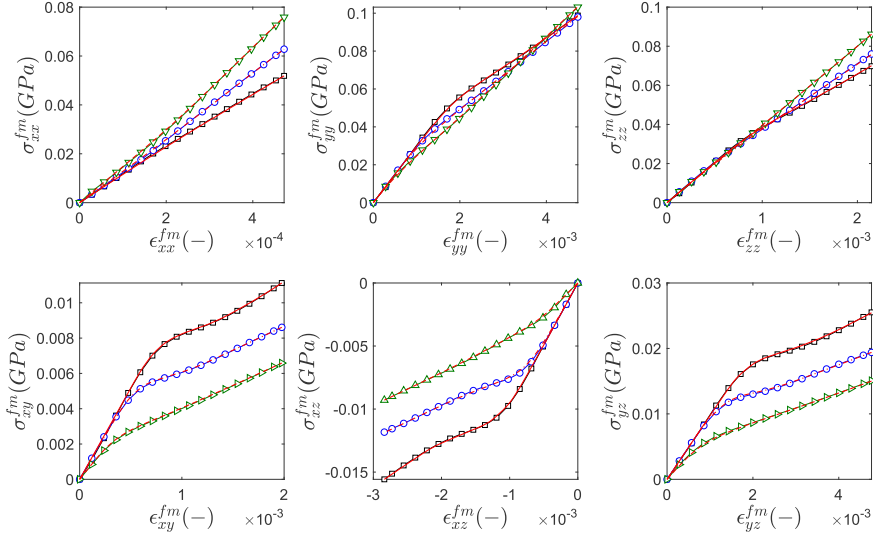
(b)

Figure 6.12: Comparison of DNN^{micro} and FEM produced strain-stress curves for the two randomly sampled loading cases (a) $\boldsymbol{\epsilon}_1^{fm} = [2.21, -4.99, -1.98, -3.53, -4.08, -3.14] \cdot 10^{-3}$ and (b) $\boldsymbol{\epsilon}_2^{fm} = [0.47, 4.73, 2.15, 1.98, -2.84, -4.76] \cdot 10^{-3}$.

101

6.5.2 SENSITIVITY ANALYSIS ON THE MACROSTRUCTURE

A sensitivity analysis, as described in section 3.4, is performed, based on the lateral displacement at the z-direction of node $\mathcal{A}$, i.e. the monitored node, whose location can be seen in fig. 6.6. For the purpose of this analysis, a stochastic parametric variability is introduced to the model parameters. Thus, the friction angle ϕ, the dilation angle ψ, the initial cohesion c_0 and the hardening modulus H are assumed to follow independent normal Gaussian distributions with properties listed in table 6.4. The propagation of the uncertainty from the parameter space to the solution is investigated by extracting the probability density function of $|u_z^A|$ from 48000 simulations and the results are shown in fig. 6.13.

Subsequently, SA is performed for the calculation of the sensitivity indices. Figs. 6.14a and 6.14b present the convergence behavior of the first order indices S and the total effect indices S_t, respectively, while fig. 6.15 illustrates both the first-order and total indices. As can be seen in these figs., approximately 40000 simulations were necessary for the first order sensitivity indices to converge. It is evident that for this type of analysis a direct computational homogenization procedure would be impossible. Instead, by utilizing the proposed surrogate model, the total analyses required for the collection of all the samples were performed at reasonable and affordable computational times. Also, from fig. 6.15 it is apparent that the initial cohesion c_0 is the most dominant parameter, in the sense that it has the largest impact on the investigated model. Moreover, from the S_t indices it can be concluded that the interactions between the parameters also have a noteworthy effect on the model's response.

$\mathcal{N}(\mu,\sigma)$	friction ϕ (degrees)	dilation ψ (degrees)	initial cohesion c_0 (GPa)	hardening modulus H (GPa)
μ	25	25	0.03	3
σ	5	5	0.004	0.6

Table 6.4: Probability distributions of the Drucker-Prager model uncertain parameters
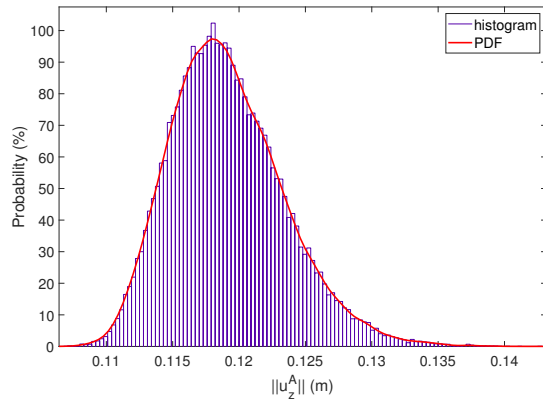
Figure 6.13: Probability Density Function of the horizontal displacement on the monitored node A.

(a)

(b)
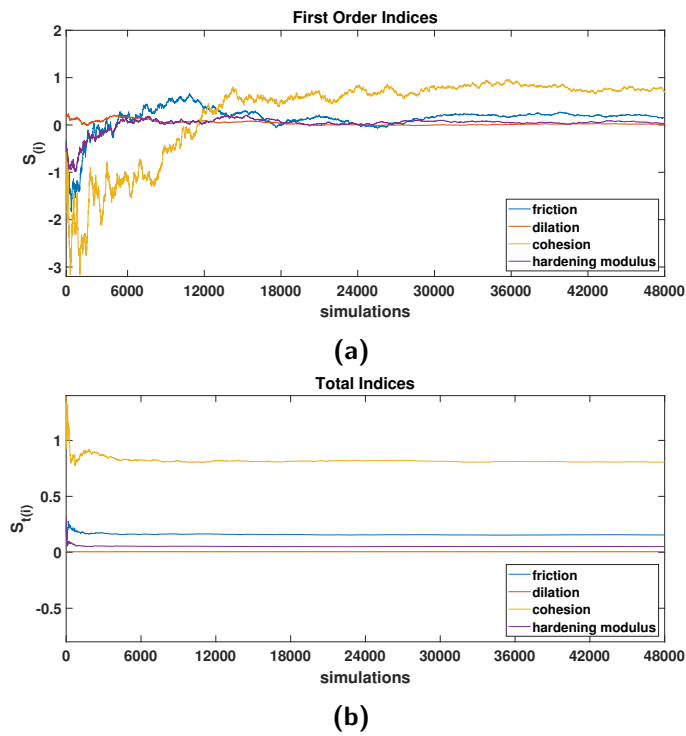
Figure 6.14: Convergence investigation of (a): first order indices $S_{(i)}$ and (b): total effect indices $S_{t(i)}$
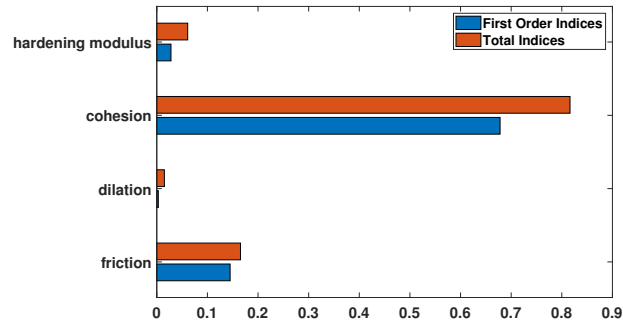
Figure 6.15: Sensitivity indices of the investigated parameters

To further assess the way each parameter affects the response of the model, scatter plots were produced in fig. 6.16. These plots depict the response u_z^A for various parameter instances. Based on these results we can confirm that the initial cohesion has indeed a strong correlation with the monitored response of the macrostructure. On the contrary, for the rest of the investigated parameters, the macroscopic response is more or less unaffected by their variability.
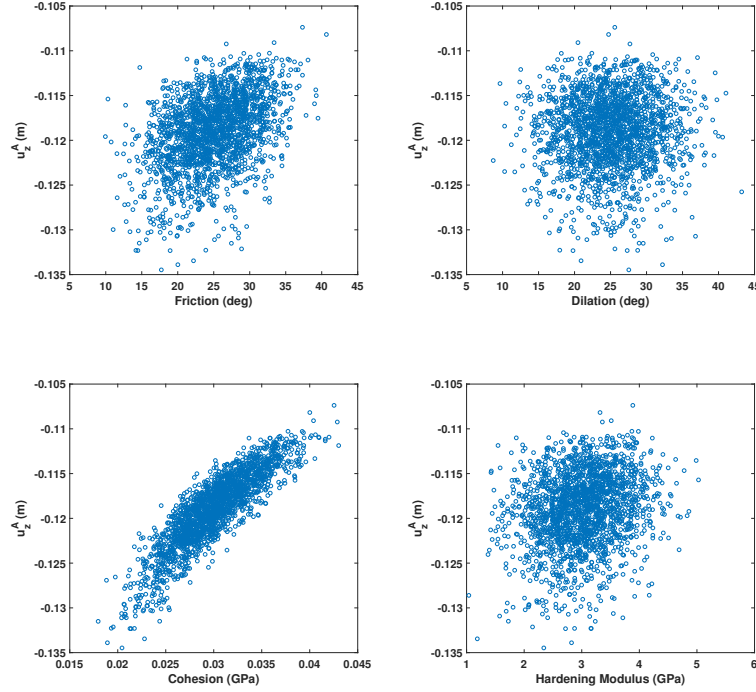
Figure 6.16: Scatter Plots of the investigated parameters against the response. The cohesion has the most notable influence in the macroscopic z-direction displacement of node A, while the friction, dilation and hardening modulus do not impact considerably this response.

In addition to the scatter plots, a series of force-displacement curves are plotted in fig. 6.17. Fig. 6.17a presents the evolution of the monitored response, if we assume that the constitutive law that defines the material macroscopically is described by each of the previously trained DNNs, namely the DNN^{micro}, the $DNN^{fine,meso}$ and the $DNN^{coarse,meso}$, corresponding to the reinforced cement paste, reinforced mortar and reinforced concrete materials, respectively. The aim of this fig. is to provide us with the information on how the material law progresses with the addition of each scale and how the macromodel's behavior is ultimately determined, for a fixed parametric vector $\boldsymbol{\alpha}$. In this case this investigation was carried out for the parametric vector $\boldsymbol{\alpha}_2 = [\phi = 25°, \psi = 25°, c_0 = 0.03GPa]$ and it can be observed that each added scale increases structural stiffness. Moreover, fig. 6.17b illustrates the evolution of the displacement, with regards to different combinations of the studied parameters. Here, this evolution was examined for the cases of the parametric vectors $\boldsymbol{\alpha}_1 = [\phi = 10°, \psi = 10°,$

$c_0 = 0.01 GPa$ (case A), $H = 1GPa$], $\boldsymbol{\alpha}_2 = [\phi = 25°, \psi = 25°, c_0 = 0.03GPa$ (case B), $H = 3GPa]$ and $\boldsymbol{\alpha}_3 = [\phi = 40°, \psi = 40°, c_0 = 0.05GPa, H = 5GPa]$ (case C). These analyses were performed on the basis that the macroscopic material is the reinforced concrete, described by the $DNN^{coarse,meso}$.

(a)   **(b)**
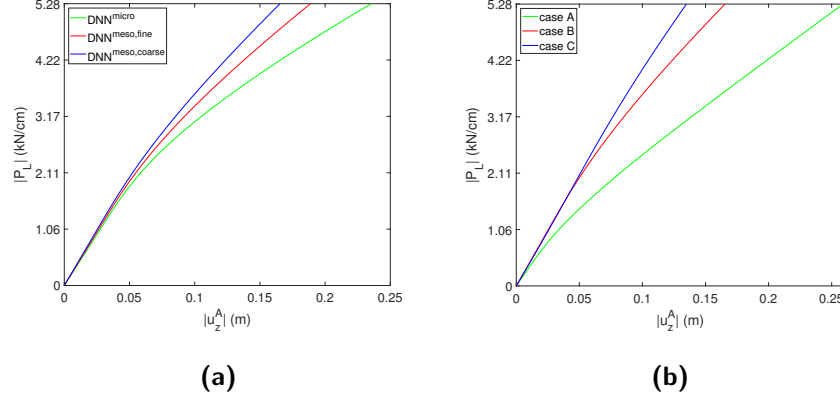
Figure 6.17: (a) Monitored response on the macrostructure by assigning each trained DNN as the macroscopic material. (b) Monitored response on the macrostructure for various combinations of the investigated parameters.

6.5.3 COMPUTATIONAL COST ASPECTS

A rough estimation of the computational time of the direct solution can be performed with the help of eq. (1.33). Taking into account that the number of integration points are 16356, 15932 and 12276 for the fine mesoscale, the coarse mesoscale and the macroscale respectively, while assuming that 10 incremental steps are assigned for the solution of each scale with each of them requiring 3 newton iterations to converge, then the total RVE evaluations for each macroscopic material point can be estimated as:

$$
\begin{aligned}
K_{RVE} &\simeq k^M + k^M \times k^{cm} + k^M \times k^{cm} \times k^{fm} \\
&= (12276 \times 3 \times 10) + (12276 \times 3 \times 10) \times (15932 \times 3 \times 10) + \\
&\quad + (12276 \times 3 \times 10) \times (15932 \times 3 \times 10) \times (16356 \times 3 \times 10) \\
&= 8.64 * 10^{16}
\end{aligned}
\tag{6.16}
$$

The enormous amount of RVE solutions required for performing a complete analysis of the macrostructure is evident from eq. (6.16). This amount is of several orders of

106

magnitude higher if we take into account the repetitive model evaluations enforced by the SA. In table 6.5 a comparison of the computational effort required for the direct numerical simulation and the surrogate model is carried out. For the surrogate model, the offline computations include the collection of the training data and the network training for the DNNs of all scales. During the system solution, instead of solving BVPs dictated by all the RVEs of the finer scales, minimal cost DNN computations take place. These computations include the evaluation of the $DNN^{coarse,meso}$ output and its jacobian, in order to acquire the macroscopic stresses and tangent moduli respectively. For the full scale model, since the solution is computationally intractable, only a prediction of the computational time is attempted here. This was done based on the total RVE resolutions estimated by eq. (6.16) and by taking into account the 40000 repetitive model solutions needed for the SA convergence. Note that all the computational effort has been documented in terms of serial programming procedures.

	Computational time (hours)						
	Offline		Online				Total
	DNN sampling	DNN training	RVE solution	DNN calculations	FE^4	SA	
Full scale (Prediction)	-	-	0.0513	-	4.435e15	1.774e20	1.774e20
Surrogate	77	11.5	-	7.34e-9	0.0027	108	196.5

Table 6.5: Computational time of each stage of the solution for the full scale and the surrogate model

As can be seen from table 6.5, carrying out the SA on the full scale model would have taken up to $1.77e20\ hrs$, while the surrogate required only 196.5 hrs, which is 18 orders of magnitude lower. As a result, the otherwise unapproachable task of performing SA on the four scale system, is implemented in a reasonable computational time.

6.6 Conclusions

In this work a novel accurate and computationally efficient surrogate modelling strategy is proposed for performing computational homogenization on nonlinear and parametrized multiscale systems that involve more than two scales. A DNN hierarchy has been employed as a surrogate model of the material behavior, where each DNN has been assigned to learn the constitutive law at the respective length scale of the problem. In contrast to the direct computational homogenization, which is inherently coupled, the proposed approach enables the decoupling of the whole procedure and it results in a dramatic decrease in computational cost. The proposed material modelling framework is demonstrated on the analysis of a large-scale structure made of CNT-reinforced concrete, where a sensitivity analysis has been performed to assess the influence of the constitutive parameters on the macroscopic response. This otherwise unreachable problem, in terms of computational effort, was solved in a reasonable time, by virtue of the elaborated surrogate modelling strategy. Future extensions of this work include the application of the proposed methodology to more complex analyses, such as post-fracture investigation and crack propagation of concrete as well as enhancing the presented surrogate strategy to work efficiently under non-monotonic loading paths. For the latter, several works already exist in the literature in the direction of capturing arbitrary loading conditions of complex path dependent materials. For that, different techniques can be utilized, such as mechanistically informed neural networks [81] recurrent neural networks [139, 91] or the enrichment of the neural network's input layer with internal variables that characterize the material's nonlinear progression [137].

7

Stochastic optimization of carbon nanotube reinforced concrete for enhanced structural performance

7.1 Introduction

In this chapter, a novel computational method is proposed for performing material optimization towards improved structural performance. In this regard, we consider the orientation of CNTs as the design variables to our optimization problem. This particular parameter is shown to play a critical role in the composite's performance [135]. Next, we utilize an optimization method called Covariance Matrix Adaptation Evolution Strategy (CMA-ES) [51, 49], which is derivative-free method for handling non-convex continuous optimization problems and is suited for computationally expensive numerical models. With this methodology we can obtain the optimal CNT orientations at every location of the macroscale structure that will lead to a reduced overall deformation. Going one step further, in the effort to provide a more rational and robust approach, we also take into account the randomness in the external loading of the structure under investigation and reformulate the problem in the context of stochastic optimization.

Stochastic structural optimization, as a procedure, entails a cumbersome computational

effort and the task of performing it on the FE^3 multiscale model that is developed to simulate the CNT-concrete structural system is very challenging. To alleviate the immense computational burden associated with this endeavour, we employ a surrogate modeling technique to simulate the composite material's behavior. Specifically, the methodology presented in the previous chapter is utilized and an appropriate sequence of FFNNs is developed to represent the material of the FE^3 model. With this approach, we achieve a drastic cost reduction per structural analysis that allows us to combat efficiently the stochastic structural optimization problem.

The remaining of this chapter is organized as follows: In section 7.2, we present the multiscale model for the analysis of structures made up of the reinforced concrete. Next, in section 7.3 the surrogate modeling strategy is proposed, which will enable us to train a highly efficient neural network to emulate the complex material's behavior. In section 7.4 we describe the problem of stochastic material optimization and elaborate on the numerical tools used in this work to tackle it. In section 7.5 numerical examples are provided that demonstrate the application of the proposed methodology to structural engineering applications.

7.2 Multiscale model and solution procedure

CNT-reinforced concrete is regarded as a 3-phase material in this work with its constituents being a) CNTs, b) cement mortar, that is cement paste and fine aggregates (e.g. sand) and c) large aggregates (e.g. gravel). This section presents the numerical models used to describe each of the constituents, as well as the final model of the composite.

7.2.1 Carbon nanotubes

CNTs are initially modeled as space frame structural models according to the Molecular Structural Mechanics (MSM) approach. Subsequently, these space frames are projected to Equivalent Beam Elements, aiming at systems with manageable DOFs. A more detailed description about the complete transition process from the molecular mechanics to the EBE representation is given in section 2.1.

7.2.2 CEMENT MORTAR

The cement mortar, which consists of fully hydrated cement paste and fine aggregates, is modeled as an isotropic homogeneous material, which obeys a non-linear constitutive relation and exhibits elastoplastic behavior. We employed a Drucker-Prager type yield function [34] to indicate the transition from the elastic to the inelastic regime, which has the form:

$$\Phi(\boldsymbol{\sigma}) = \sqrt{(1-\alpha){c_0}^2 + J_2(\boldsymbol{\sigma})^2} + \eta p(\boldsymbol{\sigma}) - \xi\sqrt{(1-\alpha){c_0}^2 + c^2} \tag{7.1}$$

with J_2 being the second deviatoric invariant of the stress tensor $\boldsymbol{\sigma}$, p the hydrostatic pressure, c_0 the cohesion, η and ξ are constants that are usually chosen in a way to approximate an equivalent Mohr-Coulomb yield surface, while α is a material dependent constant that indicates the ratio between its tension and compression strength. The approximation constants are chosen as:

$$\eta = \tan\phi/\sqrt{3} \quad \& \quad \xi = (3 + \tan\phi)/3\sqrt{3} \tag{7.2}$$

with ϕ being the friction angle.

The classical strain rate decomposition is applied:

$$\dot{\boldsymbol{\epsilon}} = \dot{\boldsymbol{\epsilon}}_e + \dot{\boldsymbol{\epsilon}}_p \tag{7.3}$$

with initial conditions:

$$\boldsymbol{\epsilon}(t_0) = \boldsymbol{\epsilon}_e(t_0) + \boldsymbol{\epsilon}_p(t_0), \tag{7.4}$$

at a pseudotime t_0, with ϵ_e and ϵ_p the elastic and plastic part of the strain tensor.

According to the plastic flow rule, the rate of the plastic strain is:

$$\dot{\boldsymbol{\epsilon}}_p = \dot{\lambda}\frac{\partial\Psi}{\partial\boldsymbol{\sigma}} \tag{7.5}$$

where λ is the plastic multiplier and Ψ is the flow potential function. Here, non-associative plastic flow is assumed, hence the flow potential function takes a different form than the yield function ($\Phi \neq \Psi$). The potential function that was chosen here is:

$$\Psi(\boldsymbol{\sigma}) = \sqrt{(1-\alpha){c_0}^2 + J_2(\boldsymbol{\sigma})^2} + \bar{\eta}p(\boldsymbol{\sigma}) \tag{7.6}$$

111

where $\bar{\eta}$ has a similar meaning with the approximation constants of Φ and is given by:

$$\bar{\eta} = \tan \psi / \sqrt{3} \tag{7.7}$$

with ψ being the dilatancy angle.

The evolution of the hardening function $c(\bar{\epsilon}_p)$ is defined by an isotropic linear softening model and the rate of the equivalent plastic strain is given by:

$$\dot{\bar{\epsilon}}_p = \dot{\lambda} \tag{7.8}$$

In the above equations the Kuhn-Tucker conditions have to hold:

$$\Phi \leq 0, \quad \dot{\lambda} \geq 0, \quad \Phi\dot{\lambda} = 0 \tag{7.9}$$

The mechanical properties and Druger-Prager model parameters considered in this work are given in table 7.1.

	Mortar
Young's modulus	$20 GPa$
Poisson ratio	0.2
α (eq. (7.1))	0.1
c_0 (eq. (7.1))	0.05

Table 7.1: Mechanical properties of mortar

7.2.3   CNT-REINFORCED MORTAR

For practical applications, we are interested in generating a representative volume element (RVE) of the reinforced mortar that will allow us to study the material's behavior. This RVE will consist of the mortar matrix and a specified number of EBEs, added as inclusions to the matrix, so as to achieve a specific volume or weight fraction. The contribution of each EBE to the overall stiffness is taken into account using the embedded FE technique of equation presented in section 2.2. The addition of these elements is usually done with a random positioning generator and figure 7.1 illustrates a RVE of the microstructure for 0.5% weight fraction. The geometrical and mechanical properties of the CNTs considered in this work are given in table 7.2.

	CNT
Young's modulus	$1.051 TPa$
Poisson ratio	0.35
length	$100 nm$
diameter	$1.063 nm$

Table 7.2: Geometrical and mechanical properties of CNTs
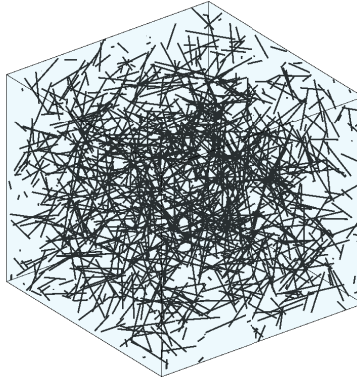
Figure 7.1: A RVE of the CNT reinforced mortar at the microscale. Its dimensions are $0.5 \ \mu m \times 0.5 \ \mu m \times 0.5 \ \mu m$

7.2.4 CNT-REINFORCED CONCRETE

CNT-reinforced concrete is studied at the mesoscale, represented by a computational model of a two-phase composite consisting of the reinforced mortar matrix along with coarse aggregates. The aggregates are considered to be linear elastic, while the nonlinear behavior of the composite originates from its microstructure. For the coarse aggregates, the following distribution is considered according to standard AB16 [133].

	Cement Matrix	Fine Aggregates						Coarse Aggregates					
Size [mm]	-	0.063	0.125	0.25	0.5	1	2	2.8	4	5.6	8	11.2	16
Vol. fraction [%]	29.26	1.50	1.62	1.76	1.76	3.63	12.17	5.063	5.15	6.74	16.61	2.90	11.83
Total [%]	29.26	22.44						48.30					
Total [%]	29.26	70.74											

Table 7.3: Concrete compositions according to standard AB16

Knowing the composition of the coarse aggregates in the matrix, we can generate a mesoscale RVE using a random geometry generator, as shown in figure 7.2. The Young's Modulus and Poisson ratio of the aggregates are taken to be 60 GPa and 0.22, respectively.
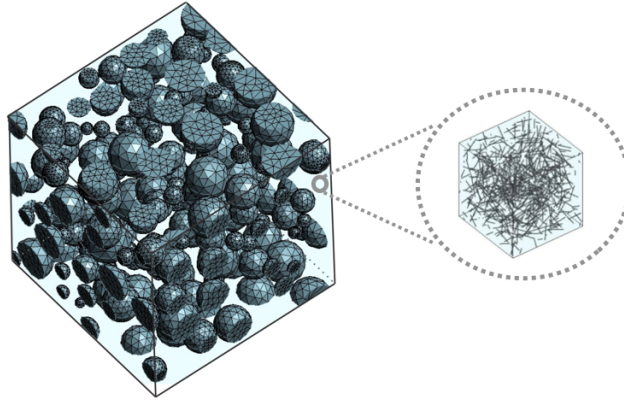
Figure 7.2: A RVE of the CNT reinforced concrete at the mesoscale. Its dimensions are $150 \ mm \times 150 \ mm \times 150 \ mm$

This section presents the numerical framework employed in order to investigate the behavior of structures at the macroscale, which are made up of the composite. Since the proposed material modeling approach involves separate model descriptions at the microscale and mesoscale, an interaction between them is required in order to pass information from one scale to the other. The interaction among scales is realised through (successive) homogenisation and localisation procedures [88, 38].

7.2.5 FE^3 SOLUTION SCHEME

Having fully established the connections between the micro-, the meso- and the macroscale, a variation of the semi-concurrent FE^2 algorithm [35], termed FE^3 herein, is used for the

online transition between the three scales during the solution of the macroscopic composite system. According to this algorithm, for each load increment of the macrostructure, the macroscopic strain $\overline{\boldsymbol{\epsilon}}^M$ is calculated on every integration point of each finite element, where it is assumed that a mesoscale RVE exists. The next step it to convert the mesoscopic strain at each integration point of the mesoscale RVE to a displacement field on a microscale RVE through a further localisation. Then, performing homogenisation on the microscale RVE will return the mesoscale stresses $\overline{\boldsymbol{\sigma}}^m$ and tangent moduli $\overline{\boldsymbol{C}}^m$ at all integration points of the mesoscale RVE. Lastly, the homogenization scheme is applied again at the mesoscale RVE, from which the macroscopic stress $\overline{\boldsymbol{\sigma}}^M$ and tangent modulus $\overline{\boldsymbol{C}}^M$ are acquired for each integration point. With these, the macroscopic internal force vector $\overline{\boldsymbol{F}}_{int}^M(\overline{\boldsymbol{\sigma}}^M)$ and tangential stiffness matrix $\overline{\boldsymbol{K}}^M(\overline{\boldsymbol{C}}^M)$ can be calculated. This procedure is repeated until the internal force vector is equivalent to the external one, namely $\overline{\boldsymbol{F}}_{ex}^M$. However, the nested nature of this scheme further implies that for each macroscopic iteration an additional set of mesoscopic iterations are required so that $\overline{\boldsymbol{F}}_{int}^m(\overline{\boldsymbol{\sigma}}^m) = \overline{\boldsymbol{F}}_{ex}^m$. The interaction of the three scales during the solution of the algorithm is illustrated in fig. 7.3.
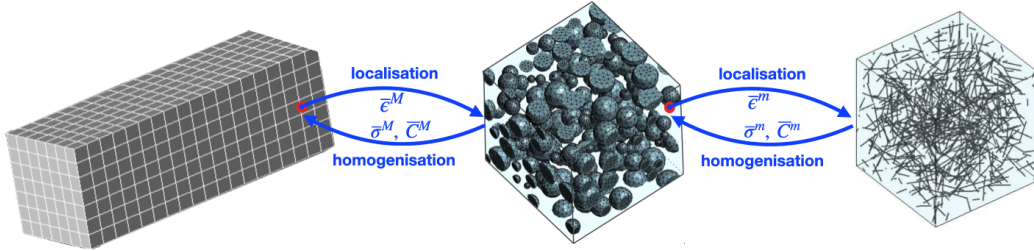
Figure 7.3: Transition between scales during the FE^3 solution algorithm

The advantage of this approach is the fact that we can model the composite material at multiple resolutions and take into account the physical mechanisms at each scale in a decoupled manner. This gives tremendous expressive capabilities to the proposed material modeling paradigm, however, the number of RVE evaluations needed, even for small problems in the macroscale, renders this approach computationally intractable. This issue will be remedied using a dedicated surrogate modeling technique, which is the focus of the next section.

7.3 Surrogate model of the FE^3 scheme

A neural network-based surrogate model is implemented in order to mitigate the computational cost of FE^3 analysis and allow the macroscale problem to be solved at a reasonable time. The aim is to use deep neural networks (DNN) to discover the parametrized non-linear homogenized response of the RVEs on each scale. Ultimately, this will enable us to effectively substitute both the micro- and the mesoscale RVE, with a single DNN that captures the mesoscale's behavior. The steps to construct this surrogate are as follows:

Step 1: Surrogate of the microscale RVE

Starting from the microscale, a set of input and output variables need to be defined, upon which the DNN will be trained in order to learn the nonlinear constitutive law of the microscale RVE. The equations that describe the micro-to-meso homogenization procedure as presented in section 1.3 relate the strain vector $\overline{\epsilon}^m$ of the mesoscale with the respective stress vector $\overline{\sigma}^m$ after a solution of a boundary value problem on the microscale. We can further assume that the solution of that system is also affected by a vector $\boldsymbol{\theta}^\mu$, containing microscale parameters. In this regard, the input neurons of the DNN can be represented by a vector that contains both the strain vector $\overline{\epsilon}^m$ and the parameters $\boldsymbol{\theta}^\mu$, thus assuming the form $[\overline{\epsilon}^m, \boldsymbol{\theta}^\mu]$. Additionally, the output neurons produce the stress vector $\overline{\sigma}^m$ for a given input vector.

To incorporate DNNs in the context of newton type iterative solutions, the output of the tangent modulus is also required by the DNN. In problems where the purpose of the surrogate is the simulation of a constitutive law, a great flexibility is provided by DNNs though the Automatic Differentiation (AD) technique [5]. AD allows for a straightforward computation of the tangent constitutive matrix $\overline{C}^m$ at a certain strain state $\overline{\epsilon}^m$, using the chain rule to effortlessly compute the derivatives of the output $\overline{\sigma}^m$ with respect to the input $\overline{\epsilon}^m$. Thus, the elements $\overline{c}_{ij}^m$ of the mesoscopic tangent matrix $\overline{C}^m = [\overline{c}_{ij}^m]$ can be obtained as

$$\overline{c}_{ij}^m = \frac{\partial \overline{\sigma}_{ij}^m}{\partial h_k} \frac{\partial h_k}{\partial h_{k-1}} \cdots \frac{\partial h_1}{\partial \overline{\epsilon}_{ij}^m} \tag{7.10}$$

where h_k is the output at the k_{th} hidden layer of the DNN for input $\overline{\epsilon}_{ij}$.

For the training of the DNN, a series of microscale RVEs are solved for various strain

combinations and parameter $\boldsymbol{\theta}^{\mu}$ realisations, leading to a training data set consisting of pairs $\{[\overline{\boldsymbol{\epsilon}}_i^m, \boldsymbol{\theta}_i^{\mu}] - \overline{\boldsymbol{\sigma}}_i^m\}_{i=1}^{N_{tr}^m}$, with N_{tr}^m being the number of training samples used for training the microscale's surrogate. Figures 7.4 and 7.5 illustrate the training procedure of the elaborated surrogate and its function, respectively.
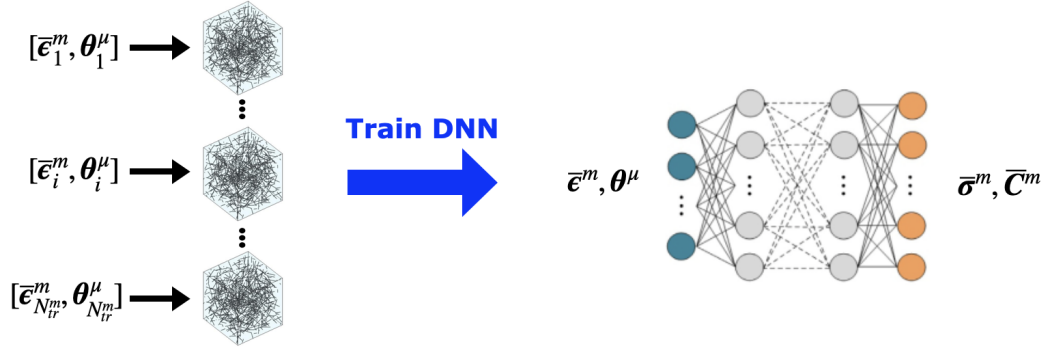
Figure 7.4: Training of the microscale DNN surrogate
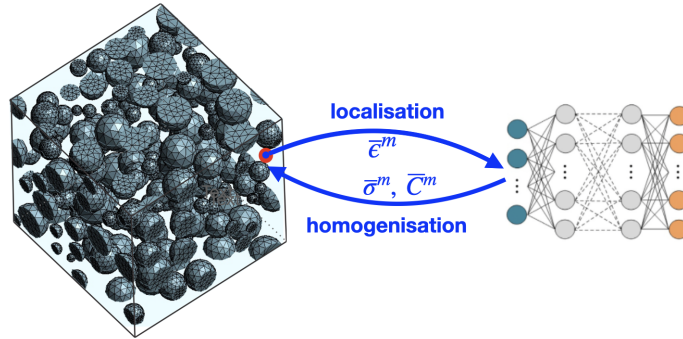
Figure 7.5: Replacement of the microscale RVE with its DNN surrogate

Step 2: Surrogate of the mesoscale RVE

From the previous step, we have established a 'cheap-to-evaluate' function that gives us the nonlinear constitutive law of the material at the microscale. Based on this, it is now possible to solve multiple mesoscale RVEs at a reasonable time, that will allow us to train the mesoscale DNN surrogate. Similarly to the procedure in Step 1, various strain combinations and parameter $\boldsymbol{\theta}^m$ realisations are generated, where in this case

the parameters $\boldsymbol{\theta}^m$ refer the mesoscale RVE. Solving the mesoscale RVE for each of these combinations, we can construct the training data set $\{[\bar{\boldsymbol{\epsilon}}_i^M, \boldsymbol{\theta}_i^m, \boldsymbol{\theta}_i^\mu] - \bar{\boldsymbol{\sigma}}_i^M\}_{i=1}^{N_{tr}^M}$, with N_{tr}^M being the number of training samples used. It is instrumental to include in the training process of the mesoscale surrogate, the microscale parameters $\boldsymbol{\theta}^\mu$, defined in the previous step, since with this approach we can indeed take into account the composite material's behavior at multiple scales. The training process of the surrogate is schematically depicted in fig. 7.6. Furthermore, figure 7.7 demonstrates how the surrogate can be used to replace the both the meso- and microscale RVE and return the homogenised stress and tangent modulus at each integration point of the macroscale problem at negligible cost.

Figure 7.6: Training of the mesoscale DNN surrogate

Figure 7.7: Replacement of the mesoscale RVE with its DNN surrogate

General remarks

118

- The problems studied herein assume small deformations and consider materials with isotropic behavior on every scale, thus the tensors $\overline{\boldsymbol{\epsilon}}^m, \overline{\boldsymbol{\epsilon}}^M$ and $\overline{\boldsymbol{\sigma}}^m, \overline{\boldsymbol{\sigma}}^M$ are symmetric and can be outlined with six independent variables, while the sizes of $\boldsymbol{\theta}^\mu, \boldsymbol{\theta}^m$ may vary on each scale, according to the individual choices.

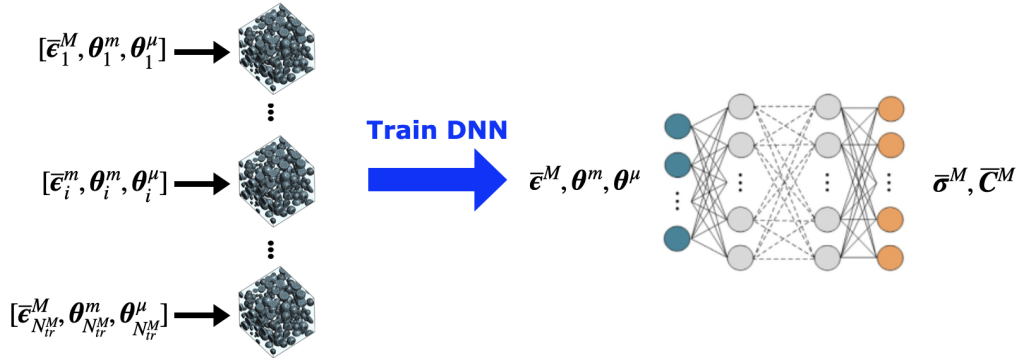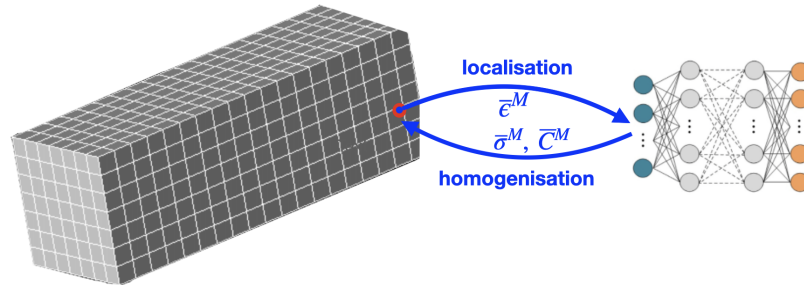- An important factor to the accuracy of the surrogate is the sampling technique used for generating the set of input training vectors. Sampling with simple Monte Carlo methods can be inefficient as it might fail to give representative results in high dimensional spaces with low samples. A more efficient alternative would be to draw samples from the space spanned by the DNN input vector using space-filling algorithms or variance reduction techniques. In this work, the input samples are obtained using the Latin Hypercube Sampling (LHS) algorithm [87].

- Lastly, attention must be paid on the range of the sampling in order for the DNN to be able to predict the full response of a system during every possible solution and at same time avoid over-training for extreme values with infinitesimal chance to be reproduced in the online solution, thus unnecessarily increasing the offline computational time.

7.4 Stochastic material optimization

7.4.1 Problem setting

Owing to the computational machinery developed in the previous sections, we are now able to optimize our material based on the performance of structures constituted by CNT-reinforced concrete. Let $\mathcal{U}$ be the mathematical model of the structural system under investigation. Since in most cases of practical interest the mathematical model cannot be derived analytically, the most common approach is to substitute it with a numerical model $\boldsymbol{U}$, typically obtained by the FE method. In this regard, $\boldsymbol{U} \in \mathbb{R}^d$ is a d-dimensional vector corresponding to the d dofs of the FE discretization. Further, if the structure is subjected to random loading conditions, expressed through the r-dimensional random vector $\boldsymbol{q}$, with r being the number of random variables, then $\boldsymbol{U} := \boldsymbol{U}(\boldsymbol{q})$. In addition, the structural response is affected by the material at the microscale, whose properties we chose to parametrize in this work according to the CNT orientation, $\boldsymbol{\theta} := \boldsymbol{\theta}^\mu = (\theta_p, \theta_a)$ with $\theta_p \in [0\ \pi]$ the polar angle and $\theta_a \in [0\ 2\pi)$ the azimuthal, as shown in figure 7.8. Therefore, we can express our final numerical model as $\boldsymbol{U} := \boldsymbol{U}(\boldsymbol{q}, \boldsymbol{\theta})$.
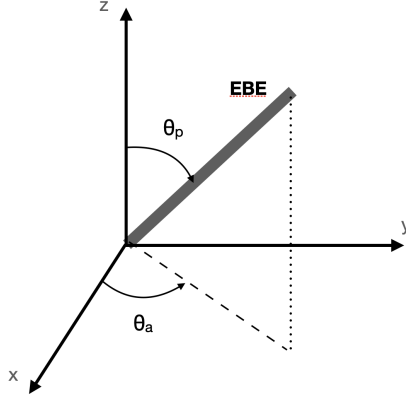
Figure 7.8: Polar angle θ_p and azimuthal θ_a, characterising the CNT's orientation in 3d space

To establish a basis of reference, we initially consider the absence of CNTs in the cement mortar. This is because we want to assess how conventional concrete behaves compared to the CNT-reinforced one. Therefore, we generate a large number, N_{ref}, of realizations $\boldsymbol{q}_1, ..., \boldsymbol{q}_{N_{ref}}$ and perform the corresponding model simulations $\boldsymbol{U}_{ref}^i = \boldsymbol{U}(\boldsymbol{q})$ for $i = 1, ..., N_{ref}$. Subsequently, we calculate the mean and standard deviation from this statistical sample as:

$$\mathbb{E}[\boldsymbol{U}_{ref}] = \frac{1}{N_{ref}} \sum_{i=1}^{N_{ref}} \boldsymbol{U}_i \tag{7.11}$$

$$Std[\boldsymbol{U}_{ref}] = \sqrt{\sum_{i=1}^{N_{ref}} \frac{(\boldsymbol{U}_i - \mathbb{E}[\boldsymbol{U}_{ref}]) \odot (\boldsymbol{U}_i - \mathbb{E}[\boldsymbol{U}_{ref}])}{N_{ref} - 1}} \tag{7.12}$$

where $\odot$ denotes the Hadamard product (elementwise multiplication). We compute the structure's mean total deformation, $\mathcal{M}$, and standard deviation, $\mathcal{S}$, by the L_2-norm of the above vectors, that is,

$$\mathcal{M} = \|\mathbb{E}[\boldsymbol{U}_{ref}]\|_2 \tag{7.13}$$

$$\mathcal{S} = \|[Std[\boldsymbol{U}_{ref}]\|_2 \tag{7.14}$$

The stochastic optimization problem is then defined as: For a given a weight fraction

of CNTs in the mortar, find the values of the design vector $\boldsymbol{\theta} = (\theta_p, \theta_a)$ that minimize the weighted sum of the mean and standard deviation of the total deformation, normalized by the respective reference values given by eqs. (7.13) and (7.14). This statement can be mathematically expressed as:

$$\boldsymbol{\theta}^\star = (\theta_p^\star, \theta_a^\star) = \underset{\theta_p \in [0\ \pi], \theta_a \in [0\ 2\pi]}{\arg\min} w_1 \frac{\mathbb{E}[\boldsymbol{U}]}{M} + w_2 \frac{Std[\boldsymbol{U}]}{S}$$

$$= \underset{\theta_p \in [0\ \pi], \theta_a \in [0\ 2\pi]}{\arg\min} \mathscr{L}(\boldsymbol{\theta}) \tag{7.15}$$

with $\mathscr{L}(\boldsymbol{\theta}) = w_1 \frac{\mathbb{E}[\boldsymbol{U}]}{M} + w_2 \frac{Std[\boldsymbol{U}]}{S}$ denoting the loss function. Also, the weight factors w_1, w_2 in eq. (7.15) are chosen as 0.6 and 0.4, respectively, so as to put more emphasis on the minimization of the mean value.

It is straightforward to extend the presented formulation to include separate CNT orientation for each structural member or specific regions of the structure. Evidently, the resulting structural designs after the optimization process will not be realisable with current manufacturing capabilities. Nevertheless, performing stochastic material optimisation based on structural performance can be regarded as one of the most challenging problems in computational mechanics and, hopefully, the outcomes of this investigation will indicate if material optimization is a research direction worth pursuing in the effort to design stronger and safer structures.

To solve the minimization problem posed in eq. (7.15) we will employ the Covariance Matrix Adaptation Evolution Strategy (CMA-ES)[50]. The algorithmic procedure of CMA-ES is given in section 3.5. In the problems studied in this work, the sample points $\{\boldsymbol{x}_i\}_{i=1}^{\lambda}$ in the CMA-ES algorithm refer to the design variables $\{\boldsymbol{\theta}_i\}_{i=1}^{\lambda}$ and the objective function f is replaced by $\mathscr{L}$ of eq. (7.15). However, the application of the CMA-ES algorithm for minimizing $\mathscr{L}$ requires some modifications, since each evaluation of $\mathscr{L}$ involves the estimation of the mean and standard deviation of the structural responses due to the random loading conditions. This is achieved by performing N_{MC} Monte Carlo simulations for various loading conditions $\{\boldsymbol{q}_j\}_{j=1}^{N_{MC}}$ at each evaluation $\mathscr{L}(\boldsymbol{\theta}_i)$. Therefore, the number of FE analyses required to solve the stochastic optimization problem becomes $\lambda \times N_{MC} \times N_{gen}$, with N_{MC} and λ chosen to be 1000 and 300, respectively, while N_{gen} is the number of generations required for the algorithm to converge. Evidently, this problem would be computationally unattainable without the surrogate modeling strategy

121

put forth in section 7.3.

7.5   NUMERICAL APPLICATIONS

In this section, we investigate two numerical examples to demonstrate the applicability of the proposed methodology to structural problems and assess the improvement achieved in structural performance.

7.5.1   TEST CASE 1: 3D BEAM FIXED AT BOTH ENDS

The first test case involves an illustrative example of a 3d beam, fixed at both ends, which is subdivided into two smaller beams, A and B, that meet in the middle. The beam is made up of CNT-reinforced concrete with 0.5% weight fraction of CNTs. The structure is subjected to a uniform load q as shown in fig. 7.9, which follows a truncated Gaussian distribution with mean $10MN/m^2$, standard deviation $2MN/m^2$ and its values are restricted within the interval $[2, 18]$ (MN/m^2). Further, we assume that the CNTs on each of the two beams can have different orientations, therefore, the vector of design variables for this problems becomes $\boldsymbol{\theta} = \left(\theta_a^A, \theta_p^A, \theta_a^B, \theta_p^B\right)$. The beam is discretized into 48 hexahedral elements, leading to 297 (free) dofs and 384 integration points.
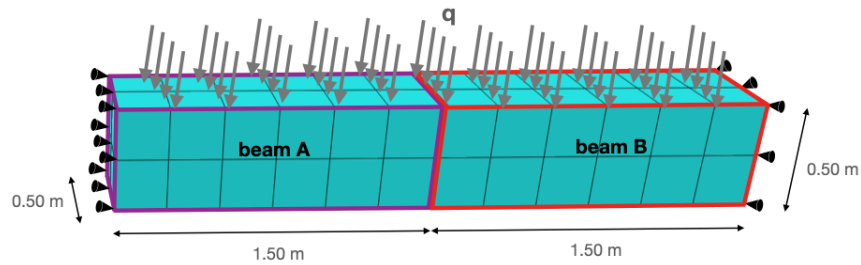
Figure 7.9: 3d beam, fixed at both ends

The first step to apply the algorithmic framework developed in the previous sections, is to establish the DNN surrogate for the microscale RVE. The input variables and their ranges are shown in table 7.4. The DNN's architecture consists of 2 hidden layers with 40 nodes each using the hyperbolic tangent activation function and the Levenberg-Marquardt optimization algorithm [83]. For the training, 5000 initial training

samples (500 realizations of $[(\overline{\boldsymbol{\epsilon}}^m, \theta_p, \theta_a)] \times 10$ increments each for the nonlinear solution algorithm) were used. The data were split in three subsets, namely the train,test and validation subset with ratios 0.7,0.15 and 0.15, respectively. The progression of the training is shown in fig. 7.10a, which required approximately 800 epochs to reach acceptable levels of accuracy. The training of the mesoscale DNN is performed following the same procedure as the microscale DNN, with its training progression given in fig. 7.10b.

	ϵ_{11}	ϵ_{22}	ϵ_{33}	ϵ_{12}	ϵ_{13}	ϵ_{23}	θ_p	θ_a
min	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	0	0
max	0.04	0.04	0.04	0.04	0.04	0.04	π	2π

Table 7.4: Input sample ranges for training the microscale DNN surrogate
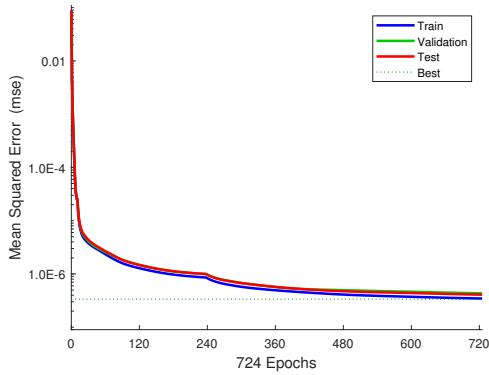
(a) Progression of the microscale DNN training process **(b)** Progression of the mesoscale DNN training process
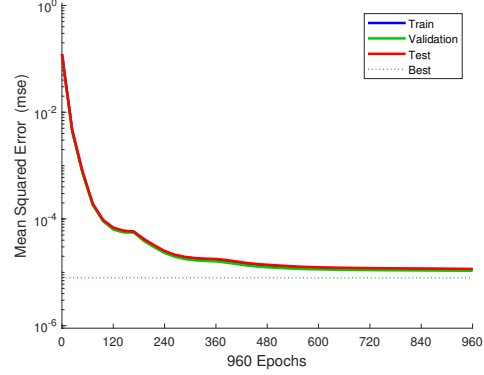
The optimal values of the design variables are obtained after applying the (100,300)-CMA-ES algorithm. These are presented in table 7.5, while figure 7.11 provides a schematic illustration of them. The reduction of the loss function $\mathcal{L}$ as a function of the objective function evaluations is shown in fig. 7.12,from which we observe that the loss function converges after $341 \times D$ objective function evaluations, where $D = 4$ is the number of design variables in this example, to the value of $\mathcal{L} = 0.81$. Lastly, we repeat this procedure from the beginning five more times in order to investigate the effect that different CNT contents in the mortar have on the structural performance. The results for

$wt\% = 0.25 - 1.50$ are displayed in figure 7.13. Additionally, the corresponding results that we could obtain by adding randomly oriented CNTs (without any preferred orientation, or alignment) are also included in this figure for the purposes of comparison. Upon inspection of this figure it becomes evident that CNTs are highly promising candidates for the development of high-performance concrete. Moreover, perfectly aligned CNTs that are optimally oriented with respect to the structure's loading conditions can give a significant additional improvement in mitigating structural deformations.

	Beam A	Beam B
$\theta_p \ (deg)$	90.47°	89.51°
$\theta_a \ (deg)$	0.02°	0.01°

Table 7.5: Optimal angles of CNTs in each structural member
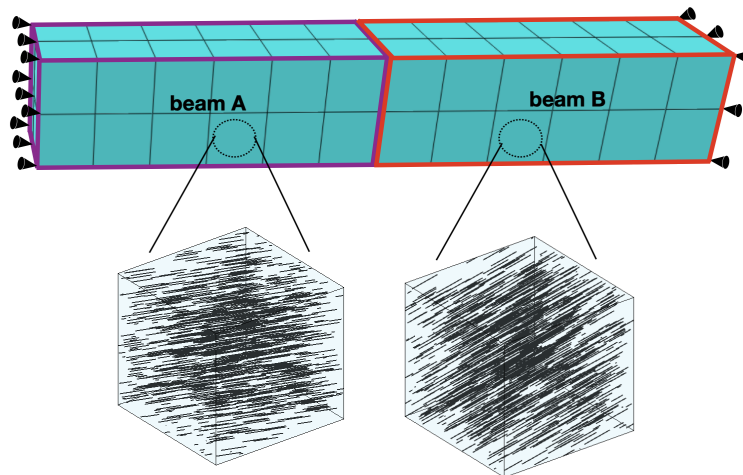
Figure 7.11: Schematic depiction of optimal CNT angles
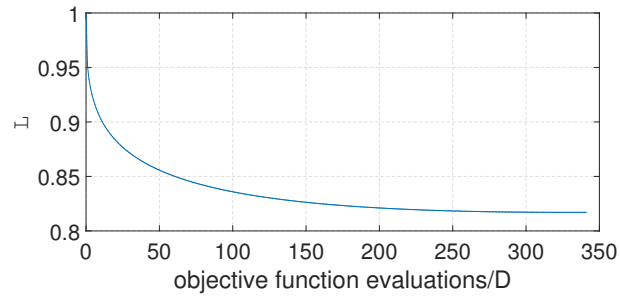
Figure 7.12: Optimization algorithm convergence of the loss function $\mathcal{L}$ vs objective function evaluations for CNT weight fraction equal to 0.5%
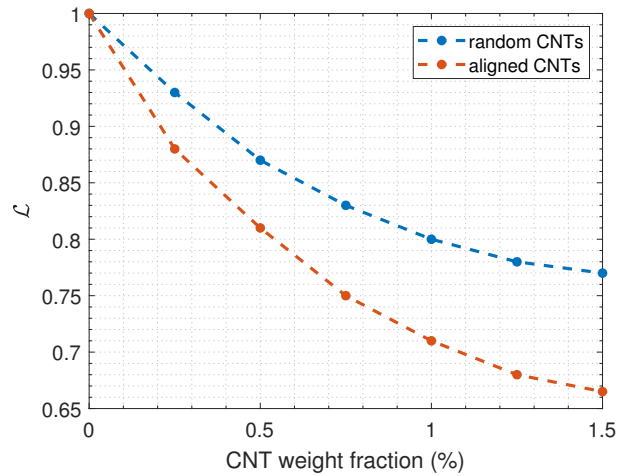
Figure 7.13: Comparison in the minimization of the loss $\mathcal{L}$ for randomly oriented CNTs and CNTs aligned with the chosen orientation, as a function of the CNT weight fraction in the mortar

7.5.2 TEST CASE: 3D FRAME

For our our second test case we examine the 3d frame of fig. 7.14. This structure is subjected to a random lateral load q_1, which follows a truncated Gaussian distribution with mean $2MN/m^2$, standard deviation $0.15MN/m^2$, restricted within the interval $[1,\ 3]$ (MN/m^2), and a vertical load q_2 that also follows a truncated Gaussian with mean $5MN/m^2$, standard deviation $0.25MN/m^2$, restricted within the interval $[2,\ 8]$ (MN/m^2). The frame is subdivided into the three structural members, columns A,C and beam B, with each member having its separate CNT orientation. As a result,

125

the stochastic optimization problem in this example involves $D = 6$ design variables $\boldsymbol{\theta} = (\theta_p^A, \theta_a^A, \theta_p^B, \theta_a^B, \theta_p^C, \theta_a^C)$. Regarding the computational mesh, the structure is discretised into 260 hexahedral finite elements, resulting in 2080 integration points.
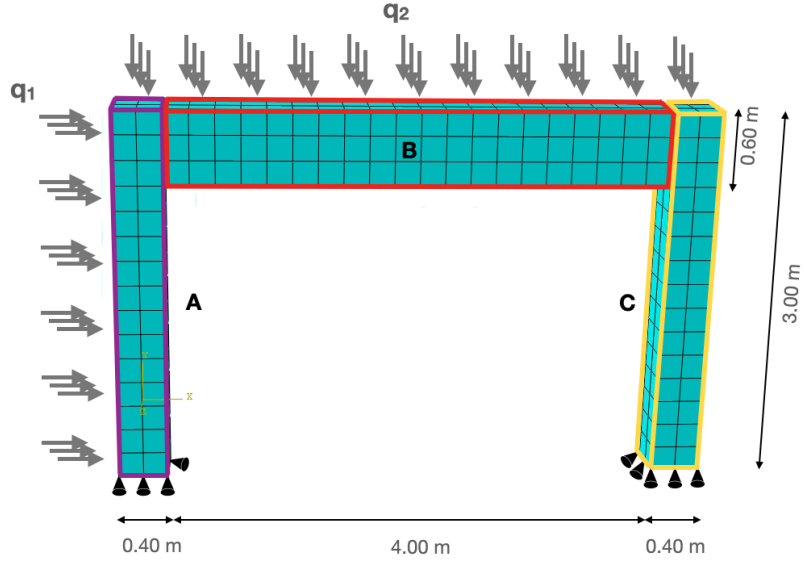
Figure 7.14: 3d frame

In this example, we also assume that the CNT content in the mortar is 0.5%. This enables us to use the already trained mesoscale DNN of the previous example. With this tool available, we proceed with finding the optimal CNT orientation angles of each structural member. These results are collectively presented in table 7.6 and a schematic depiction is provided in fig. 7.15. The minimization of the loss function with respect to the number of function evaluations is displayed in fig. 7.16, which converges ultimately to the value of 0.755 after $503 \times D$ evaluations. Finally, we repeat the process for various CNT contents in the RVE of the microstructure to assess the structural performance improvement. The results of this investigation are depicted in fig. 7.17, where also a comparison with respect to the case of randomly aligned CNT is included. An interesting conclusion is drawn after juxtaposing figs. 7.15 and 7.17 with the corresponding figs. 7.11 and 7.13 of example 1. Specifically, it can be observed that in the second example we attained better improvement in structural performance by selecting the appropriate CNT orientation angles. This leads us to speculate that in more complex problems, with

126

more structural members involved, even better results can be expected.

	Column A	Beam B	Column C
θ_p (deg)	1.03°	89.20°	0.72°
θ_a (deg)	0.01°	0.01°	0.02°

Table 7.6: Optimal angles of CNTs in each structural member
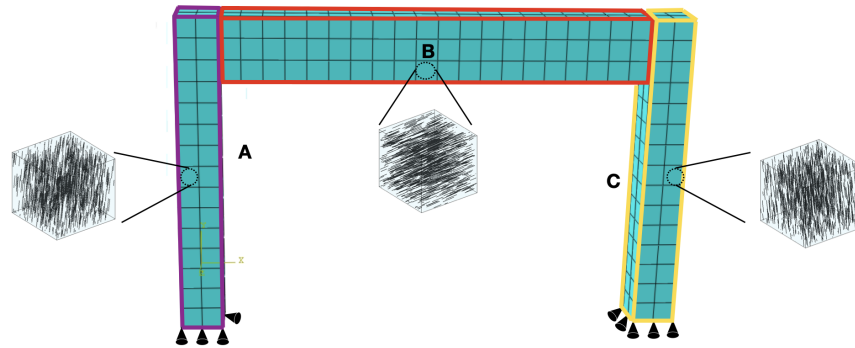
Figure 7.15: Optimal angles of CNTs in each structural member
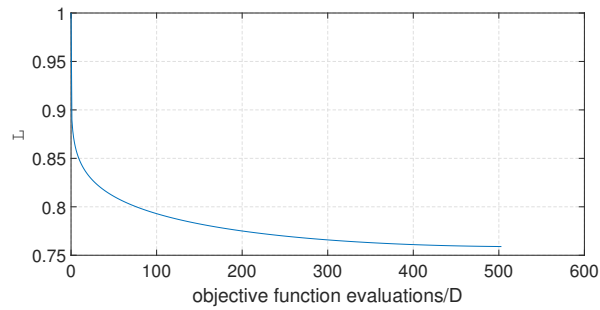
Figure 7.16: Optimization algorithm convergence of the loss function $\mathscr{L}$ vs objective function evaluations for CNT weight fraction equal to 0.5%
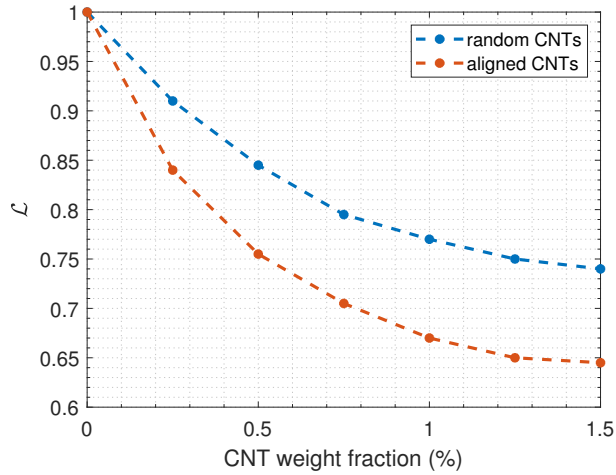
Figure 7.17: Comparison in the minimization of the loss $\mathscr{L}$ for randomly oriented CNTs and CNTs aligned with the chosen orientation, as a function of the CNT weight fraction in the mortar

7.6 Conclusions

In this work a novel numerical framework has been developed for the analysis of structures made up of CNT-reinforced structures. In particular, a two-level hierarchical material model has been proposed based on continuum micromechanics to characterize CNT-reinforced concrete. An extension of the FE^2 method, called FE^3 has been employed to assess structural performance, where the excessive computational demands of the method are effectively bypassed by using a neural network based surrogate as a substitute for the composite material. With this implementation, we perform stochastic material optimization, aiming to find the optimal CNT orientation at different members of the macroscale structural systems that minimize the sum of the mean and standard deviation of the overall structural deformation. The results of this investigation indicate that the reinforcement of concrete with CNTs can lead to significant enhancement in structural performance. In addition, the extension of the current manufacturing capabilities to the design of structural members with appropriately aligned and oriented CNTs can revolutionise structural design, leading to stronger, safer and more elegant structures. The elaborated methodology is demonstrated on structural problems, nevertheless it can be applied in a straightforward manner to other problems of engineering interest such as heat transfer and electrical conduction.

128

<div align="right">

8

</div>

An efficient hierarchical Bayesian framework for multiscale material modeling

8.1 Introduction

The present chapter proposes an innovative method for determining material properties within multiscale material systems through a range of experimental scenarios. The presented framework holds the promise of merging data acquired from experiments conducted on materials of different compositions and encompassing measurements taken at various length scales, allowing the systematic integration of multiple experimental data sources into a unified computational framework. To achieve this, we utilize the Transitional Markov Chain Monte Carlo (TMCMC) method to sample from the marginalized posterior distributions of both multiscale model parameters and hierarchical hyperparameters. These hyperparameters are subsequently employed to derive informed physical parameters, which can be used for future model predictions. Crucially, feedforward neural networks (FFNNs) play a key role in reducing the computational complexity of implementing hierarchical Bayesian analysis on top of nonlinear computational homogenization. Their primary aim is to learn and accurately predict the nonlinear constitutive

law across various scales. To evaluate the efficacy of the proposed approach, a study is carried out on the parameters that define the interfacial mechanical behavior of carbon nanotubes (CNTs) in CNT-reinforced cementitious material configurations. For this task, we collected data from conventional experiments conducted on diverse material configurations defined at multiple length scales, each associated and characterized through a FE^2 based hierarchical multiscale computational model.

The rest of this chapter is structured as follows: Section 8.2 presents the general problem of multiscale material parameter identification tackled in this study and details the proposed hierarchical Bayesian framework. Section 8.3 revisits the theoretical background of computational homogenization and the application of neural networks toward the mitigation of its computational cost. Section 8.4 demonstrates an illustrative example for the calibration of the interfacial mechanical properties in CNT-reinforced cementitious materials. Lastly, Section 8.5 summarizes the key points and results of this work and discusses possible extensions to the proposed strategy.

8.2 Bayesian inference on multiscale material systems.

8.2.1 Problem definition

Consider a set of data $\boldsymbol{D} = [\boldsymbol{D}^1, ..., \boldsymbol{D}^K]$, with $\boldsymbol{D}^i = [d^{i,1}, ..., d^{i,N_i}]$, that encompasses measurements (i.e. mechanical responses) obtained from a number of K independently performed experiments. Each i-th experiment is defined on a domain Ω^i and described by a particular combination of boundary conditions $\partial\Omega^i$ and material composition. In the context of material analysis, these experimental scenarios can be simulated by the respective multiscale computational models $\mathscr{M}^1, ..., \mathscr{M}^K$ that are used to predict structural responses $\boldsymbol{m}^1, ..., \boldsymbol{m}^K$ in line with the experimental datasets. Each model $\mathscr{M}^i$ is characterized by a unique material formulation, for instance with a different number of length scales, a dissimilar composition in all or a subset of the length scales, etc. Consequently, the parameterization of these models is done through a distinct set $\boldsymbol{\theta}^i$, as $\boldsymbol{\theta}^i = [\boldsymbol{\theta}^i_1, ..., \boldsymbol{\theta}^i_{S_i}]$, which includes a series of physical, topological or constitutive attributes for each length scale $s = 1, ..., S_i$ of the respective system. In many cases, a subset of these parameters ${}^c\boldsymbol{\theta}^i \subseteq \boldsymbol{\theta}^i$, with ${}^c\boldsymbol{\theta}^i = [{}^c\boldsymbol{\theta}^i_1, ..., {}^c\boldsymbol{\theta}^i_{S_i}]$, is present in all the investigated multiscale material models, meaning that $\boldsymbol{\theta}^1 \cap ... \cap \boldsymbol{\theta}^K \equiv {}^c\boldsymbol{\theta}^1 \equiv ... \equiv {}^c\boldsymbol{\theta}^K \equiv {}^c\boldsymbol{\theta}$.

In the present study, the focus is on the investigation of these common characteristics

130

$^{c}\boldsymbol{\theta}$ by properly incorporating the knowledge from all the K experimental instances. Towards this goal, we propose a hierarchical Bayesian framework tailored to tackle this intricate task. As a graphical representation of the overall problem at hand and without loss of generality, fig. 8.1 depicts a three-model scenario where common material parameters are to be investigated. The hierarchical Bayesian strategy, complemented by the hyperparameters $\boldsymbol{\psi}$, enables the identification of new physical parameters $^{c}\boldsymbol{\theta}^{new}$ in a reliable manner. In turn, $^{c}\boldsymbol{\theta}^{new}$ can facilitate informed predictions $\boldsymbol{y^{new}}$ on unobserved multiscale material systems. It is important to emphasize that the data sources have the potential to encompass a wide range of experiments conducted across various length scales. These experiments, for example, may include macroscale displacement measurements, mesoscale topology characterization through Scanning Electron Microscope (SEM) images [107], or microscale strain field extraction using the Digital Image Correlation (DIG) technique [86]. Since we are only concerned with the inference of the common parameters $^{c}\boldsymbol{\theta}$, for the sake of brevity, these will be plainly denoted as parameters $\boldsymbol{\theta}$ for the remainder of the chapter.
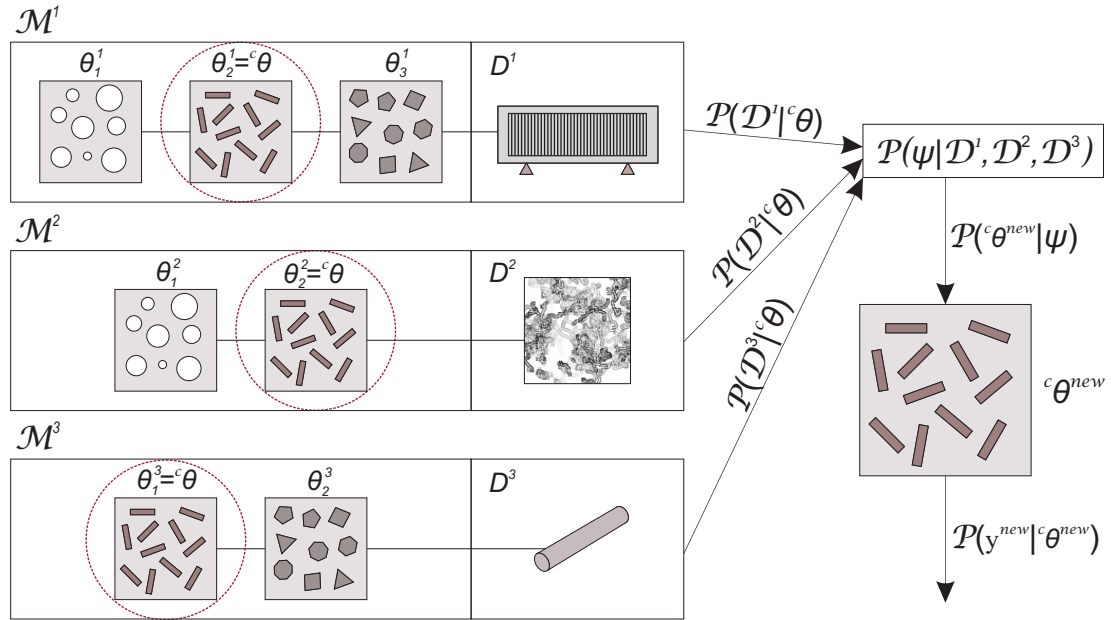

Figure 8.1: Parameter identification of the common properties $^{c}\boldsymbol{\theta}$ found in a series of diverse experimental cases such as a model $\mathcal{M}^1$ of a bending test on reinforced beam made of a composite material, a model $\mathcal{M}^2$ of a digital image of a composite material's microstructure and a model $\mathcal{M}^3$ of a rod specimen of the composite.

8.2.2 OVERVIEW OF THE HIERARCHICAL BAYESIAN FRAMEWORK

In this study we explore the integration of data from diverse independent sources with focus on parameter identification in multiscale material model analysis. Such cases can occur, for example, from experiments that took place in non-communicating laboratories, experiments from a single laboratory but for varying material compositions, environmental conditions etc. It has been showcased from a series of previous works, mainly in the field of structural dynamics, that in such cases the single-level Bayes underestimates the total variability [117, 63]. That is because it tends to only represent the "parameter estimation uncertainty" and is unable to account for the external variability amongst the datasets. This variability denotes the aleatoric uncertainty that cannot be explicitly quantified within the problem formulation due to lack of knowledge regarding the experiment specific conditions [7]. Taking these findings into consideration we opt for a hierarchical Bayesian approach for the formulation of the proposed strategy.

Several techniques have been proposed in the literature for the solution of eq. (3.19). The most straightforward way is to directly sample from the joint distribution of eq. (3.19) [93]. Despite this being a reliable approach, the high dimensionality of the parameter space leads to enormous computational costs and cannot be effectively applied in realistic scenarios. An alternative is to decouple the process for the solution of eq. (3.19), by sequentially sampling from the marginalized posterior distributions of the physical parameters and the hierarchical hyperparameters respectively [140].

Here we will implement the latter approach of the sequential sampling. In cases where many data points are available, to alleviate the computational demands, in [117], it was assumed that the form of the posterior is close to that of a Gaussian and analytical expressions through Laplace's approximation where derived. In this work we deal with sparse data and potentially strong nonlinear phenomena that could unpredictably alter the form of the target distributions. Therefore, to acquire the precise expression of the posterior PDFs we opt for the utilization of the Transitional Markov Chain Monte Carlo (TMCMC) [18] for drawing samples from both marginal distributions [140, 101]. A different technique will be used for the reduction of the computational cost and this will be presented in the following section.

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

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

$$(8.1)$$

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

$$P(\boldsymbol{D}|\boldsymbol{\psi}) = \prod_{i=1}^{K}\left[\int_{\Omega_{\boldsymbol{\theta}^i}}\prod_{j=1}^{N_i}\left[P(d^{i,j}|\boldsymbol{\theta}^i)\right]P(\boldsymbol{\theta}^i|\boldsymbol{\psi})d\boldsymbol{\theta}^i\right] \qquad (8.2)$$

The integral of eq. (8.1) can be approximated via Monte Carlo sampling:

$$P(\boldsymbol{\psi}|\boldsymbol{D}) \simeq \prod_{i=1}^{K}\left[\frac{1}{N_{\boldsymbol{\theta}^i}}\sum_{k=1}^{N_{\boldsymbol{\theta}^i}}P(\boldsymbol{\theta}_k^i|\boldsymbol{\psi})\right]\frac{P(\boldsymbol{\psi})}{P(\boldsymbol{D})} \qquad (8.3)$$

To perform this integration, samples from each likelihood $P(\boldsymbol{D}^i|\boldsymbol{\theta}^i)$ associated with the model M^i, have to first be collected. This is achieved efficiently by employing the TMCMC algorithm [18] (see section 3.2). When TMCMC is used as a sampler for a posterior PDF, the initial step is to gather an amount of samples from the prior PDF. Since in our case we want to sample directly from $P(\boldsymbol{D}^i|\boldsymbol{\theta}^i)$ an easy way to replicate the standard algorithmic procedure of TMCMC is to assume auxiliary uniform prior PDFs $P(\boldsymbol{\theta}^i)$ for the initial sampling from each likelihood. To avoid introducing any bias in the process, the lower limit u_l and upper limit u_u of the priors $\mathcal{U}_{\boldsymbol{\theta}^i}(u_l, u_u)$ should be selected broad enough to cover the sample space of each $\boldsymbol{\theta}^i$ postulated by $P(\boldsymbol{D}^i|\boldsymbol{\theta}^i)$.

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

$$P(\boldsymbol{\theta}^{new}|\boldsymbol{D}) = \int_{\Omega_{\boldsymbol{\psi}}} P(\boldsymbol{\theta}^{new}|\boldsymbol{\psi})P(\boldsymbol{\psi}|\boldsymbol{D})d\boldsymbol{\psi} \qquad (8.4)$$

The approximate solution of eq. (8.4) is done by generating an amount of samples for $\boldsymbol{\psi}$ through the empirical distribution of $P(\boldsymbol{\psi}|\boldsymbol{D})$ obtained from eq. (8.3), as follows:

$$P(\boldsymbol{\theta}^{new}|\boldsymbol{D}) \simeq \sum_{k=1}^{N_{\boldsymbol{\psi}}} P(\boldsymbol{\theta}^{new}|\boldsymbol{\psi}_k) \qquad (8.5)$$

For the production of samples from eq. (8.4), the TMCMC 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 $P(\boldsymbol{D}^i|\boldsymbol{\theta}^i)$, eq. (8.5) represents a cheap-to-evaluate procedure, since no multiscale model resolutions are needed.

The assumption of globally identifiable parameters in the form of $\boldsymbol{\theta}^{new}$ is valid when the hyperparameters $\boldsymbol{\psi}$ can account for the model variations across different datasets, thus rendering the principle of total probability valid [117]. Since the material models that involve common characteristics are similar by nature, this assumption can be confidently made in the present context. 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 fig. 8.2. 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) $\boldsymbol{y}^{new}$ is calculated as:

$$P(\boldsymbol{y}^{new}|\boldsymbol{D}) = \int_{\Omega_{\boldsymbol{\theta}new}} P(\boldsymbol{y}^{new}|\boldsymbol{\theta}^{new})P(\boldsymbol{\theta}^{new}|\boldsymbol{D})d\boldsymbol{\theta}^{new} \tag{8.6}$$
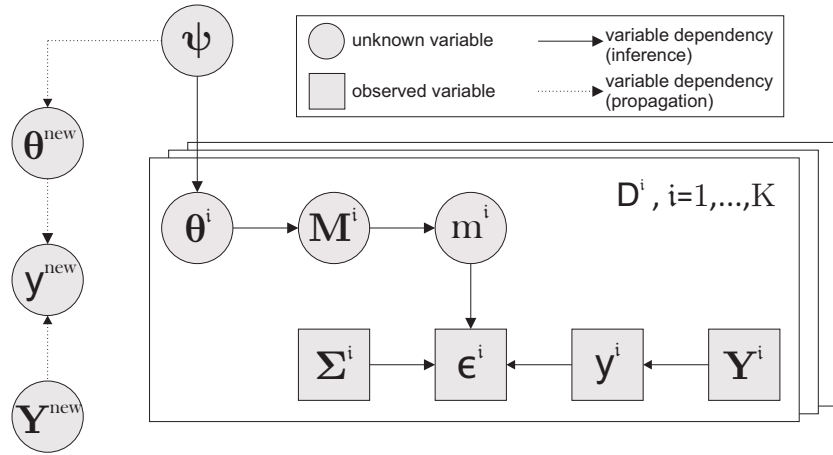
Figure 8.2: Directed Acyclic Graph (DAG) of the general hierarchical problem

Algorithm 3: Algorithm for hierarchical Bayesian inference and uncertainty propagation of heterogeneous multiscale material models

The proposed strategy is invariant to the technique that will be used for performing the multiscale material analysis. For instance, approaches such as rule of mixtures [132] or mean field homogenization [89] could be applied. Here we opt for a computational homogenization procedure to bridge the scales of the system, since it can provide the most accurate results in material cases with a notable non-linear behavior [145].

8.3 Efficient computational homogenization at multiple scales through neural network approximations

The key concept is to employ feed forward neural networks (FFNNs) towards the replacement of the costly RVE solutions that are repeatedly required during the FE^N analysis. This is accomplished by assigning the FFNNs to learn the stain-stress relation imposed by the homogenization equation for each scale. For each multiscale model M^i this relation is expressed per scale j as:

$$\boldsymbol{\sigma}_{j+1}(t) = f_j^{NN,i}\big(\boldsymbol{\epsilon}_{j+1}(t), \hat{\boldsymbol{\theta}}_j^i\big) \qquad (8.7)$$

where $\hat{\boldsymbol{\theta}}_j^i$ contains all the static parameters from the finest scale up to the j scale and the superscript i corresponds to the model $\mathcal{M}^i$. The expression in eq. (8.7) describes a plain strain-stress relation without the need to take into account the strain history or other internal variables. In this study, the knowledge of eq. (8.7) by the FFNNs is sufficient, since the database of experimental results includes solely monotonic loading scenarios. As a result, there will be a one-to-one correspondence for the strain-stress pairs and for that reason no further information in the form of internal variables is required.

Starting from the finest scale, namely the scale 1, a number n of RVE solutions for different cases of strain vector $\boldsymbol{\epsilon}_2$ and parameter vector $\hat{\boldsymbol{\theta}}_1^i$ tuples are performed. By retaining the converged states of all the solutions, a total amount of $\{\{\boldsymbol{\epsilon}_2\}_{1:t}, \{\boldsymbol{\sigma}_2\}_{1:t}, \hat{\boldsymbol{\theta}}_1^i\}_{1:n}$ of parametrized strain/stress pairs is ultimately collected, where t is the number of the increments of each solution. After choosing the architecture of the candidate neural network f_1^{NN}, the components $\{\{\boldsymbol{\epsilon}_2\}_{1:t}, \hat{\boldsymbol{\theta}}_1^i\}_{1:n}$ and $\{\{\boldsymbol{\sigma}_2\}_{1:t}\}_{1:n}$ are then used for the training as the input and output data respectively. The newly developed f_1^{NN} can be considered as a computationally cheap phenomenological relation determined by a series of non-physical parameters i.e. the optimized weights and biases of the network. It can be readily applied as the material law of the scale 2 towards the development of the next FFNN, namely the f_2^{NN}. Following the same concept as in the first scale, a dataset $\{\{\boldsymbol{\epsilon}_3\}_{1:t}, \{\boldsymbol{\sigma}_3\}_{1:t}, \hat{\boldsymbol{\theta}}_2^i\}_{1:n}$ is obtained through a series of solutions of the RVE representing the scale 2. The consistent constitutive matrix of the material represented by $f_1^{NN,i}$ can be effortlessly retrieved through Automatic Differentiation [5]. By sequentially applying the aforementioned procedure for each pair of scales until the macroscale, FFNNs that emulate the material law of each scale are ultimately retrieved. Additionally, all produced FFNNs inherit the knowledge of how the constitutive relation is affected by a number of parameters that characterize the finer structure. Therefore, they can be effectively utilized in the hierarchical Bayesian framework of multiscale material systems presented in section 8.2.2. By that, each likelihood function $P(\boldsymbol{D}^i|\boldsymbol{\theta}^i)$ is replaced by an inexpensive to calculate counterpart which we will denote as $P^{NN,i}(\boldsymbol{D}^i|\boldsymbol{\theta}^i)$.

Further details and implementation aspects for the development of the FFNN sequence can be found in chapter 6. The algorithmic procedure presented in alg. 4 constitutes the specialized version of the general algorithm 3, when surrogate models are employed

for the replacement of the finer scales of each multiscale model M^i.

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 $N^{\boldsymbol{\psi}}$ and TMCMC parameters k^{new} and β^{new}.

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

Offline Stage:

for $i \leftarrow 1$ **to** K **do**

    for $j \leftarrow 1$ **to** S_i **do**

        Generate and store training samples $\{\{\boldsymbol{\epsilon}_{j+1}\}_{1:t}, \hat{\boldsymbol{\theta}}_j^i\}_{1:n}$;

        Solve the BVP of scale j;

        Store solution stresses $\{\{\boldsymbol{\sigma}_{j+1}\}_{1:t}\}_{1:n}$;

        Train the FFNN $f_j^{NN,i}$;

        Store FFNN $f_j^{NN,i}$;

    end

end

Online Stage:

for $i \leftarrow 1$ **to** K **do**

    Generate from $P^{NN,i}(\boldsymbol{D}^i|\boldsymbol{\theta}^i)$ samples

    $\{\boldsymbol{\theta}^i\}_{1:N^{\boldsymbol{\theta}^i}} \leftarrow TMCMC(p(\boldsymbol{\theta}^i), P^{NN,i}(\boldsymbol{D}^i|\boldsymbol{\theta}^i), N^i, k^i, \beta^i)$;

    Store samples $\{\boldsymbol{\theta}^i\}_{1:N^{\boldsymbol{\theta}^i}}$ and likelihood function data $P^{NN,i}(\boldsymbol{D}^i|\{\boldsymbol{\theta}^i\}_{1:N^{\boldsymbol{\theta}^i}})$

end

Estimate $P(\boldsymbol{\psi}|\boldsymbol{D})$;

Generate from $P(\boldsymbol{\psi}|\boldsymbol{D})$ samples

  $\{\boldsymbol{\psi}\}_{1:N^{\boldsymbol{\psi}}} \leftarrow TMCMC(P(\boldsymbol{\psi}), P(\boldsymbol{D}|\boldsymbol{\psi}), N^{\boldsymbol{\psi}}, k^{\boldsymbol{\psi}}, \beta^{\boldsymbol{\psi}})$;

Estimate $P(\boldsymbol{\theta}^{new}|\boldsymbol{D})$;

Propagate posterior uncertainty $P(\boldsymbol{y}^{new}|\boldsymbol{D})$;

Algorithm 4: Algorithm of the proposed methodology for hierarchical Bayesian inference and uncertainty propagation of heterogeneous multiscale material models using surrogate modeling

8.4 Numerical application

8.4.1 Overview of experimental layouts and the respective multi-scale material models

As a means to illustrate the presented framework, the interfacial mechanical properties of CNTs in cementitious material configurations has been explored. To perform the calibration of the investigated material parameters, the results from thee separate tests were obtained from the literature. These include a 3-point bending test of a CNT-reinforced cement specimen [147], a tension test performed on CNT-reinforced mortar rods [61] and a 4-point bending experiment on a CNT-reinforced concrete beam [94]. Next, the details for each experimental case will be presented along with the multiscale model that reproduces the material's scales as well as the overall structural and material behavior.

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 $160mm \times 40mm \times 40mm$, 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 $10\mu m$ and $20\mu m$. The experimental setting and the measurements that relate the flexural strain with the respective stress based on the experimental findings are depicted in fig. 8.3. To integrate the CNTs into the FE analysis the Molecular Structural Mechanics (MSM) technique [75] 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 [75]. 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) [100] 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. 8.4. The Drucker-Prager (DP) plasticity material law [27] 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. 8.5 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 $\boldsymbol{\theta}^1$ comprises these three microscale parameters $\boldsymbol{\theta}^1 = \left[\tau_{int}^{y,1}, k_{int}^{el,1}, k_{int}^{pl,1}\right]$. In the context of the finite element analysis, to integrate the CNTs in the composite material system the cohesive zone method [4] was used in combination with an embedding technique. A more detailed description of the finite element formulation of the CNT/matrix interaction is given in 2.3.

(a) Experimental setup [147]
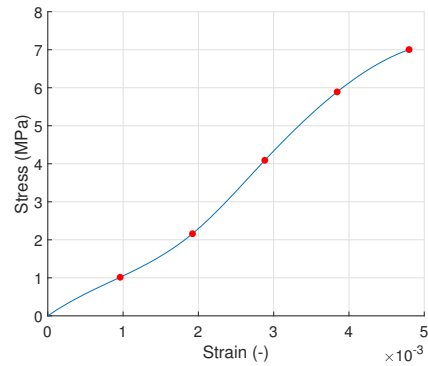
(b) Flexural strain-stress data [147]

Figure 8.3: CNT-reinforced cement specimen

139
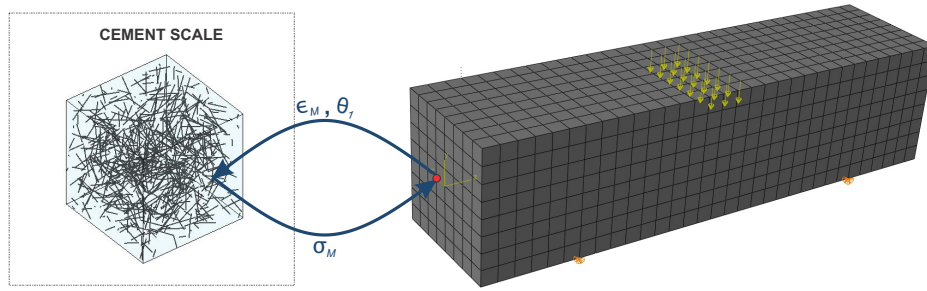
Figure 8.4: CNT-reinforced cement multiscale model
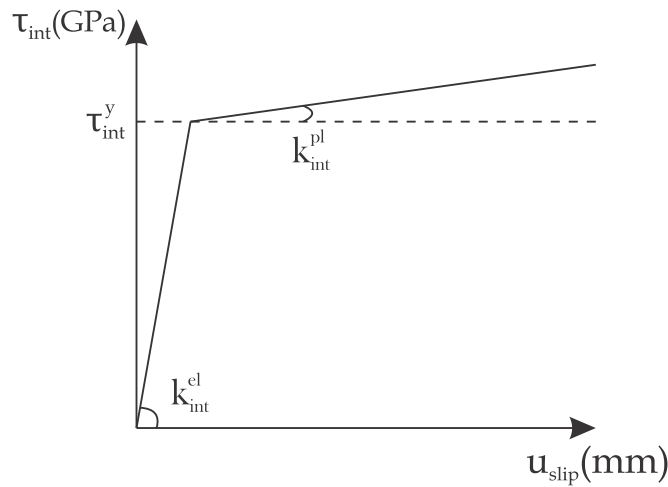
Figure 8.5: Constitutive law that defines the interaction between the CNTs and the matrix material

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 $1\mu m$ and $2\mu m$ Characteristic specimens and the tensile strain-stress dataset for the multi-walled CNT enhanced mortar bar are provided in fig. 8.6. 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. 8.7. For constructing the model of the mortal 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, as explained in [138], was enforced. The minimum diameter of the candidate spheres were taken as $d_{min} = 0.1mm$, while the maximum diameter as $d_{max} = 2mm$. 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 characterize the macroscopic behavior and the FE macromodel used to replicate the test are illustrated in fig. 8.7. 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. 8.5. According to the hierarchical Bayesian scheme presented in section 3.3, these parameters assume a separate formulation for model $\mathcal{M}^2$ as $\boldsymbol{\theta}^2 = \left[\tau_{int}^{y,2}, k_{int}^{el,2}, k_{int}^{pl,2}\right]$.

(a) Experimental setup [61]   **(b)** Tensile strain-stress data [61]
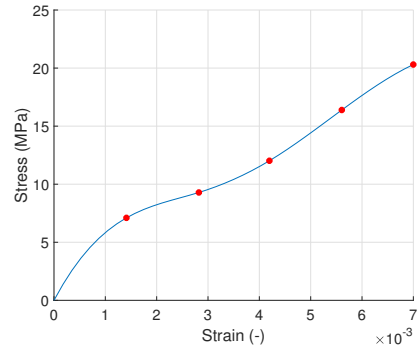
Figure 8.6: CNT-reinforced mortar specimen
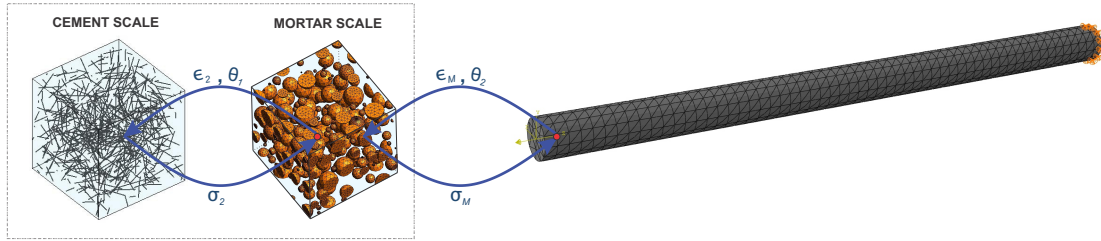
Figure 8.7: CNT-reinforced mortar multiscale model

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 $150mm \times 250mm$, while supports were used, $2000m$ apart, at its lower part. The CNTs had diameters of $3 - 15nm$ and lengths of $15 - 330\mu m$. 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. 8.8. In this scenario the material is represented by a cement, mortal and concrete three-scale model as shown in fig. 8.9 linked hierarchically to the macroscopic FE model of the beam. The first two scales were formulated in accordance with the procedure described in sections 8.4.1 and 8.4.1, 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 $d_{min} = 2mm$ and $d_{max} = 20mm$ respectively. The coarse aggregates were assumed to behave linearly elastic. Again, following the hierarchical Bayesian concept, the parametric vector of the constitutive law 8.5 that describes the interaction between the CNTs and the matrix in this model $\mathscr{M}^3$ is explicitly defined as $\boldsymbol{\theta}^3 = \left[\tau_{int}^{y,3}, k_{int}^{el,3}, k_{int}^{pl,3}\right]$.

(a) Experimental setup[94]

(b) Flexural Displacement-Load data [94]
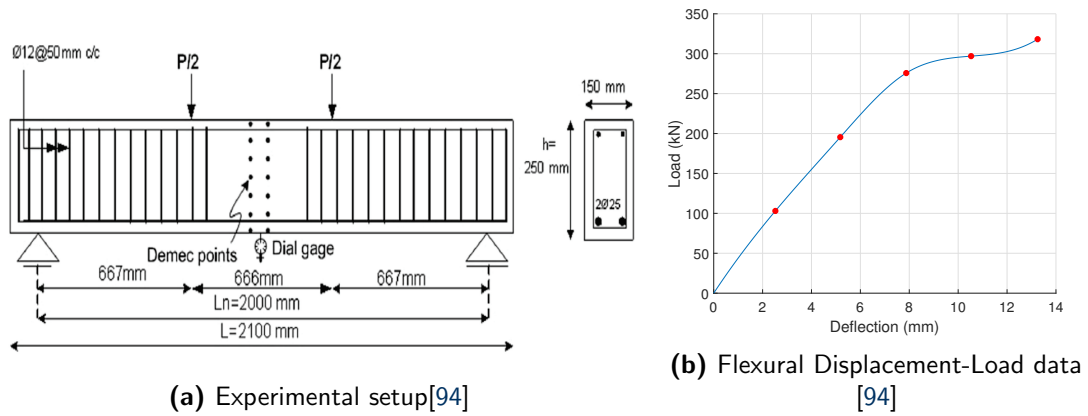
Figure 8.8: CNT-reinforced concrete specimen

(a)

(b)
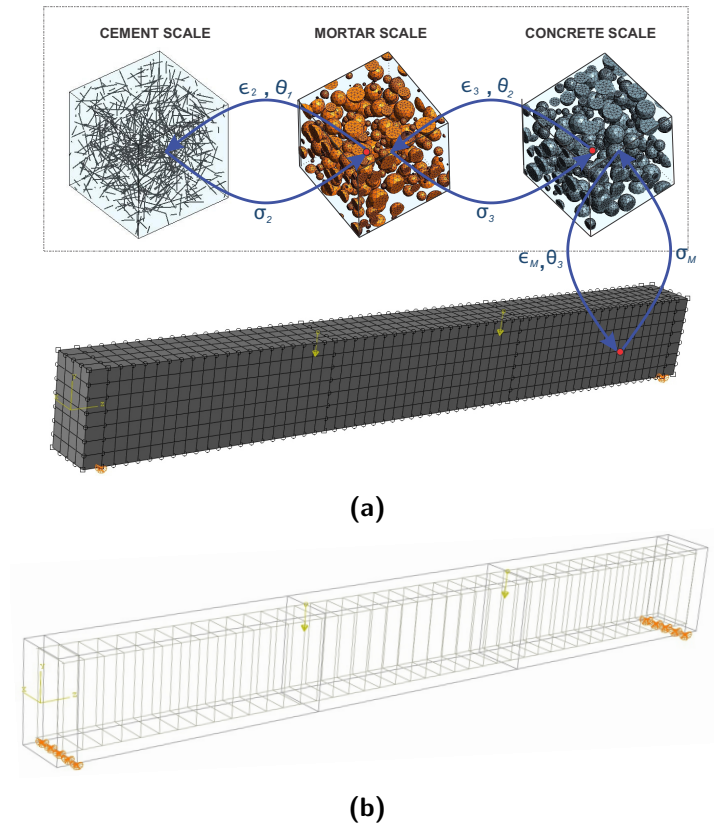
Figure 8.9: (a) CNT-reinforced concrete multiscale model (b) geometric configuration of concrete beam with rebar and stirrup reinforcement

143

8.4.2 Numerical Results

In what follows, the results from the procedure outlined in algorithm 4 will be presented. These include the offline stage of the FFNN sequence development for each multiscale model described in section 8.3 and the online stage of performing the hierarchical Bayesian analysis.

Surrogate model training

The material of the first model presented in fig. 8.4 is defined through the single scale of the CNT/cement paste mixture. Therefore, according to the strategy summarized in section 8.3 only one FFNN - $f_1^{NN,1}$ is sufficient for the reproduction of the macroscopic constitutive response. 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. To avoid unrealistic strain cases and parameter values, lower and upper bounds were set for the sampling as presented in table 8.1. After the solution of the 1000 microstructure BVPs of the RVE model of the microscale, the data pairs $\{\{\epsilon_M\}_{1:20}, \hat{\theta}_1^1\}_{1:1000}$ and $\{\{\sigma_M\}_{1:20}\}_{1:1000}$ were used as input and output pairs for the training and testing of $f_1^{NN,1}$. The Adam optimizer [66] 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 $\{\{\sigma_M\}_{1:20}\}_{1:1000}$ and the respective predicted stresses from $f_1^{NN,1}$ was used as the loss metric. To ensure a more reliable training, a Min-Max normalization was applied to both the input and output data. 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. The training curves and the predictions provided in fig. 8.10 demonstrate great performance. 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 8.1: Input sample ranges for the FFNN training

(a) Training progress

(b) Prediction accuracy
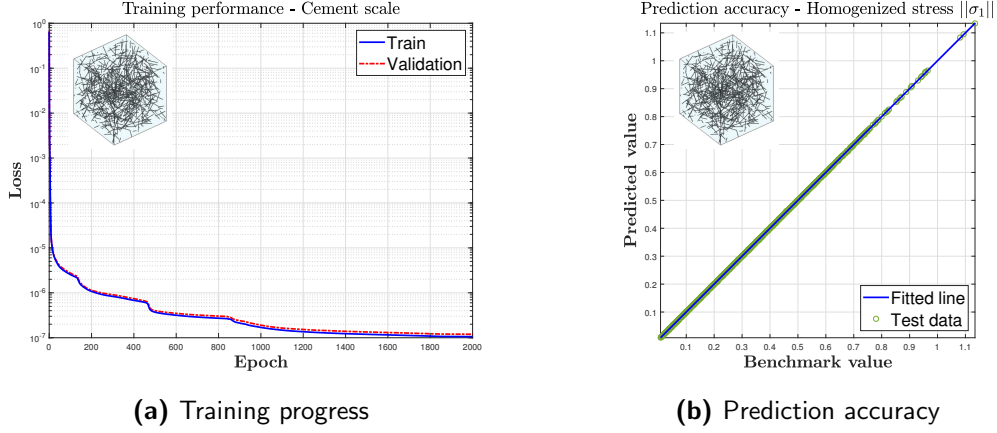
Figure 8.10: FFNN training results for the CNT-reinforced cement specimen - Cement scale

The second model of the CNT/mortar rod specimen is represented by the two-scale material model of fig. 8.7. The training process was repeated again for this material model starting from the microscale. In this scenario, the differences in the material properties, compared to the first model described in sec. 8.4.1, 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 high accuracy of both FFNNs is captured in figs. 8.11 and 8.12, where the progression of the MSE metric during the training and the quality of the

145

predictions are visualised, respectively.

(a) Training progress

(b) Prediction accuracy
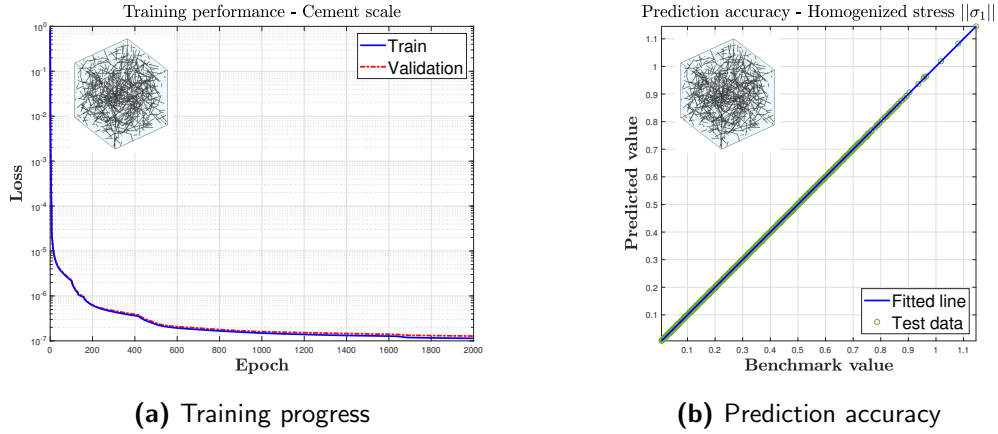
Figure 8.11: FFNN training results for the CNT-reinforced mortar specimen - Cement scale

(a) Training progress

(b) Prediction accuracy
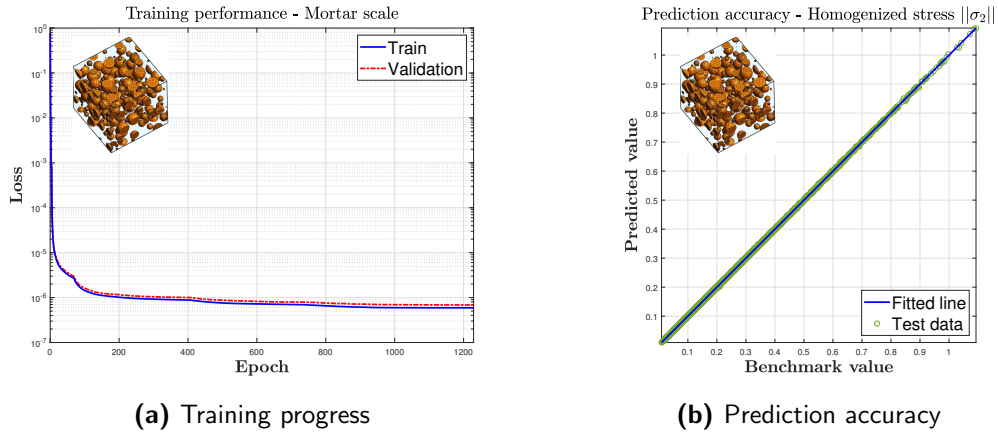
Figure 8.12: FFNN training results for the CNT-reinforced mortar specimen - Mortar scale

For the final model of the CNT/concrete beam specimen in fig. 8.9, 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, we started 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 which are given in figs. 8.13, 8.14 and 8.15 for the three scales respectively, demonstrate once more the high quality of the FFNN predictions.

(a) Training progress

(b) Prediction accuracy
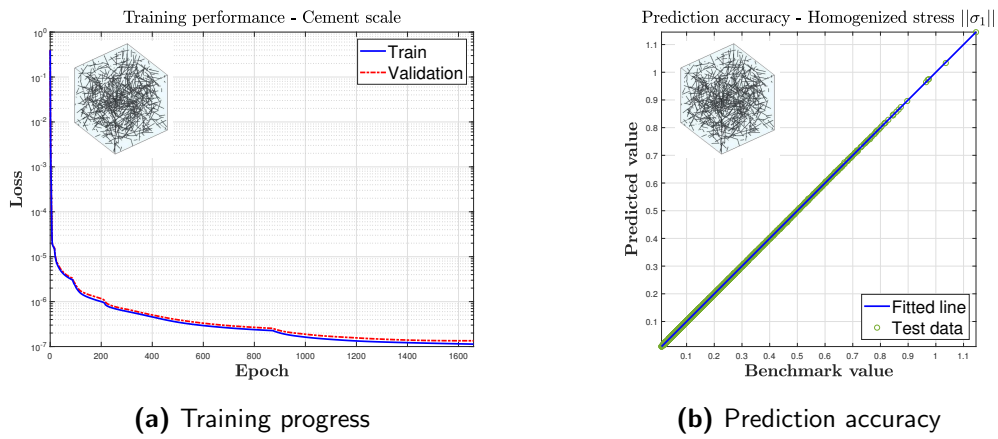
Figure 8.13: FFNN training results for the CNT-reinforced concrete specimen - Cement scale

(a) Training progress

(b) Prediction accuracy
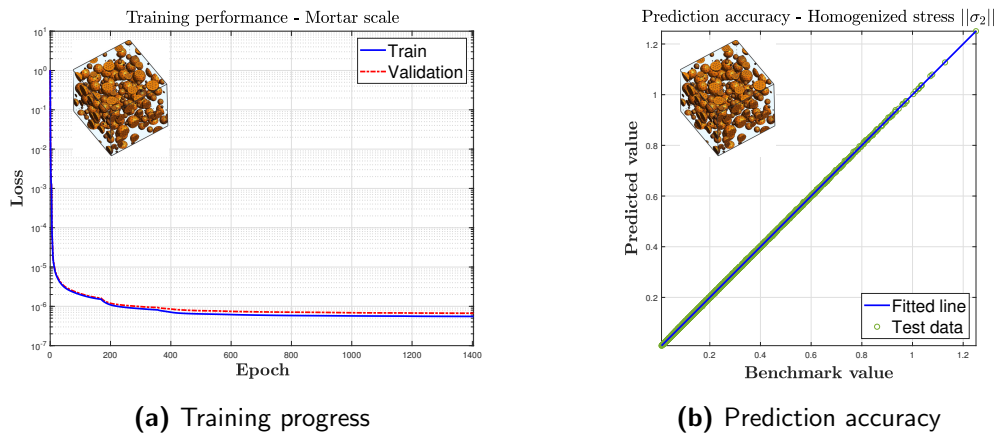
Figure 8.14: FFNN training results for the CNT-reinforced concrete specimen - Mortar scale

147

(a) Training progress   **(b)** Prediction accuracy
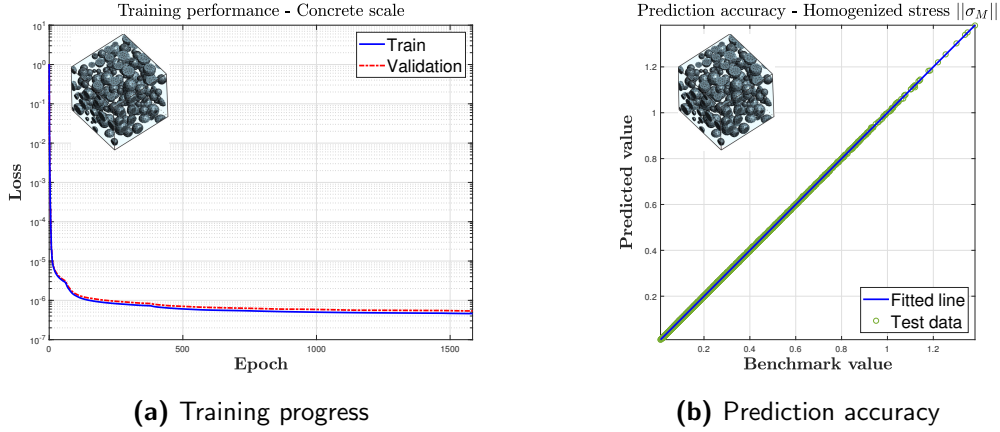
Figure 8.15: FFNN training results for the CNT-reinforced concrete specimen - Concrete scale

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 4. The data via which the investigated parameters will be updated were obtained from the figs. 8.3b, 8.6b and 8.8b for the three models respectively. From each of the three curves, five equally spaced points were retrieved and used towards the uncertainty quantification. The initial step was to perform an independent TMCMC sampling from the likelihood function $P(\boldsymbol{D}^i|\boldsymbol{\theta}^i)$ 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. As explained in section 3.3, to sample from each $P(\boldsymbol{D}^i|\boldsymbol{\theta}^i)$, first the prior distributions $P(\boldsymbol{\theta}^i)$ need to be defined. These 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 $P(t_{int}^{y,i}) \sim \mathcal{U}(0, 0.3)$. For the error terms in eq. 3.18, the standard deviation was calculated based on a coefficient of variation of 0.02. For the hyperparameters of the TMCMC algorithm, we followed the suggested values by the authors [18] i.e. $k^i = 1.0$ and $\beta^i = 0.2$, while the number of the

148

samples was chosen as $N^i = 10000$. The posterior form of each $\boldsymbol{\theta}^i$ is illustrated in figs. 8.16, 8.17 and 8.18.
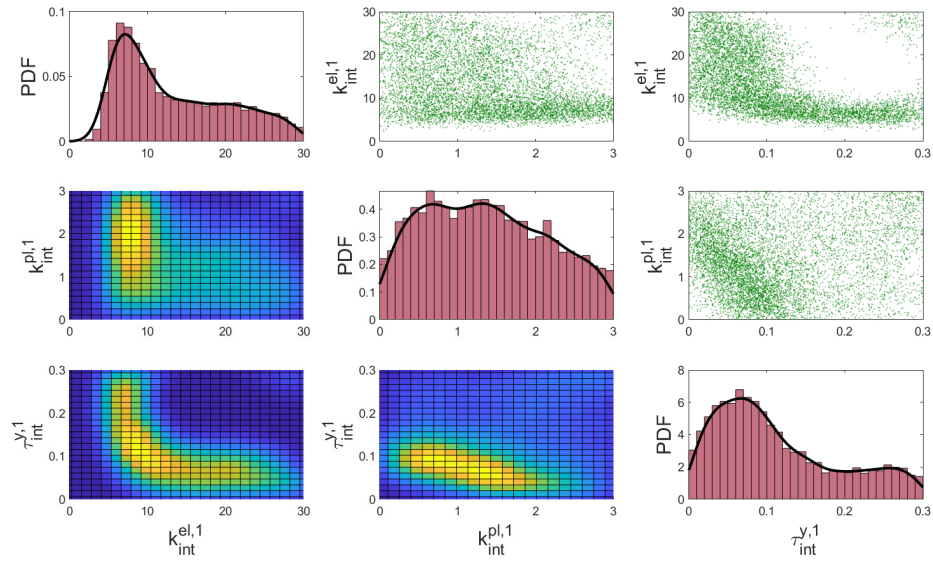
Figure 8.16: Results of the Bayesian analysis on the CNT/cement interfacial parameters of model $\mathscr{M}^1$. Diagonal - Marginal probability density functions of the investigated parameters. Upper triangle - Scatter plots for each parameter pair. Lower triangle - Joint probability density functions for each parameter pair

149
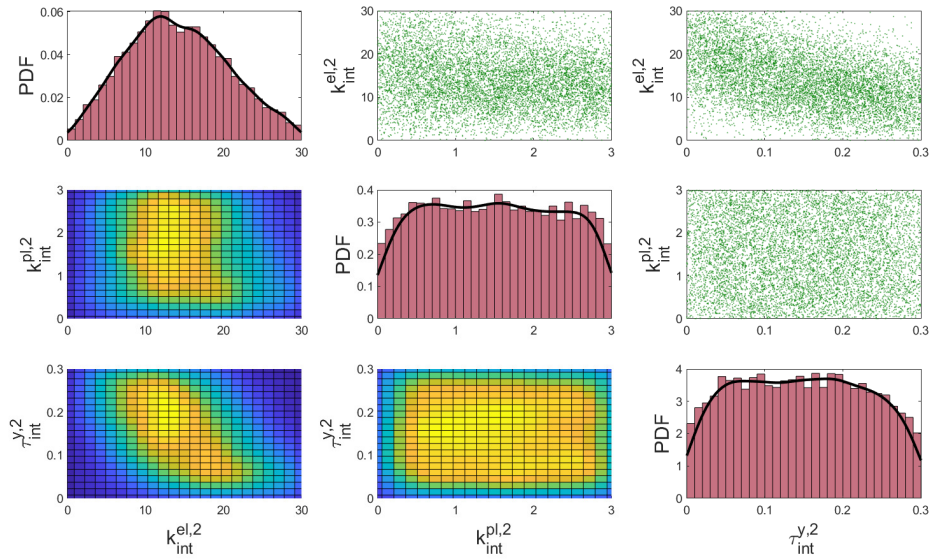
Figure 8.17: Results of the Bayesian analysis on the CNT/mortar interfacial parameters of model $\mathcal{M}^2$. Diagonal - Marginal probability density functions of the investigated parameters. Upper triangle - Scatter plots for each parameter pair. Lower triangle - Joint probability density functions for each parameter pair
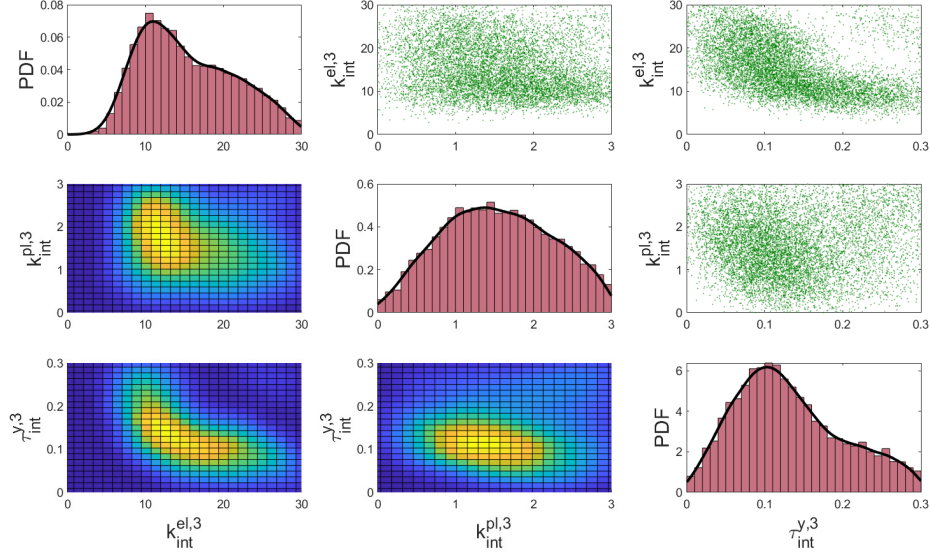
Figure 8.18: Results of the Bayesian analysis on the CNT/concrete interfacial parameters of model $\mathcal{M}^3$. Diagonal - Marginal probability density functions of the investigated parameters. Upper triangle - Scatter plots for each parameter pair. Lower triangle - Joint probability density functions 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. (8.3). Again, following the process of algorithm 4 we used the TMCMC to sample from the empirical distribution $P(\boldsymbol{\psi}|\boldsymbol{D})$. The selection of the hyperparameter priors $P(\boldsymbol{\psi})$ and the parameter priors $P(\boldsymbol{\theta}^i|\boldsymbol{\psi})$ is given in table 8.2. At this stage, the TMCMC ran for N^{ψ}=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. 8.19. The last step of algorithm 4 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. (8.5) is expressed as:

$$P(\boldsymbol{\theta}^{new}|\boldsymbol{D}) = \sum_{i=1}^{N^{\psi}} w_i \mathcal{U}(\{\boldsymbol{\psi}^1\}_i, \{\boldsymbol{\psi}^1\}_i + \{\boldsymbol{\psi}^2\}_i) \quad , \quad with \quad w_i = \frac{1}{N^{\psi}} \ , \ i = 1, ..., N^{\psi} \quad (8.8)$$

Sampling from such an empirical mixture PDF was straightforward. The resulting distributions $P(\boldsymbol{\theta}^{new}|\boldsymbol{D})$ are presented in fig. 8.20.

Hyperparameter	$\psi^1_{k^{el}_{int}}$	$\psi^2_{k^{el}_{int}}$	$\psi^1_{k^{pl}_{int}}$	$\psi^2_{k^{pl}_{int}}$	$\psi^1_{\tau^y_{int}}$	$\psi^2_{\tau^y_{int}}$
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)	$k^{el,i}_{int}$		$k^{pl,i}_{int}$		$\tau^{y,i}_{int}$	
Prior PDF	$\mathcal{U}(\psi^1_{k^{el}_{int}}, \psi^1_{k^{el}_{int}} + \psi^2_{k^{el}_{int}})$		$\mathcal{U}(\psi^1_{k^{pl}_{int}}, \psi^1_{k^{pl}_{int}} + \psi^2_{k^{pl}_{int}})$		$\mathcal{U}(\psi^1_{\tau^y_{int}}, \psi^1_{\tau^y_{int}} + \psi^2_{\tau^y_{int}})$	

Table 8.2: Prior distributions of the hyperparameters and the parameters of the tackled hierarchical Bayesian problem
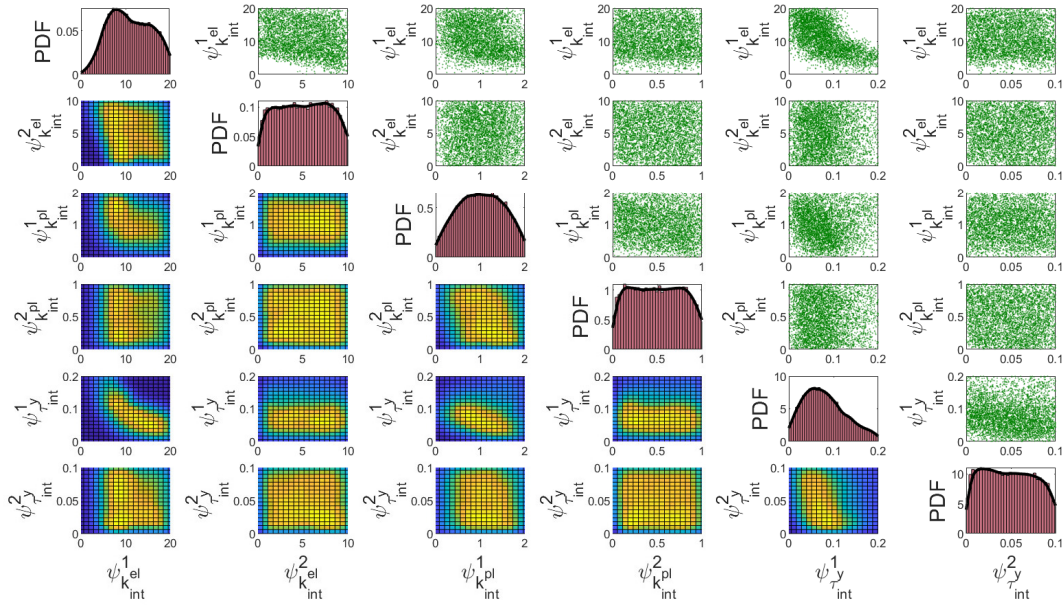
Figure 8.19: Results of the hierarchical Bayesian analysis on the hyperparameters. Diagonal - Marginal probability density functions of the hyperparameters. Upper triangle - Scatter plots for each hyperparameter pair. Lower triangle - Joint probability density functions for each hyperparameter pair
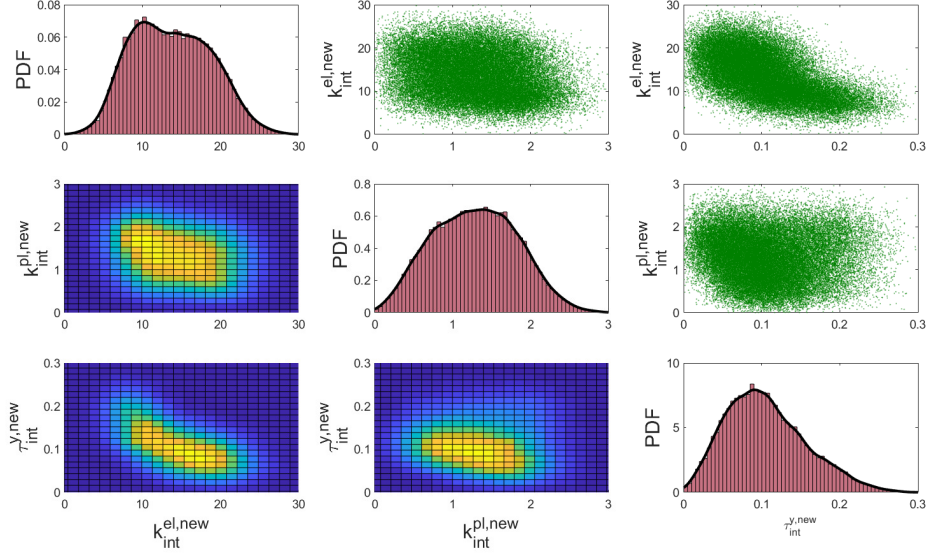
Figure 8.20: Results the of hierarchical Bayesian analysis on the newly formulated CNT/matrix interfacial parameters. Diagonal - Marginal probability density functions of the new parameters. Upper triangle - Scatter plots for each new parameter pair. Lower triangle - Joint probability density functions for each new parameter pair

Finally, we performed an uncertainty propagation analysis on how the informed parameters $\boldsymbol{\theta}^{new} = [k_{int}^{el,new}, k_{int}^{pl,new}, t_{int}^{y,new}]$ impact the mechanical behavior of the respective material for each model $\mathcal{M}^i, w \ i = 1, 2, 3$. For that, we investigated the influence of the CNT/matrix cohesive properties $\boldsymbol{\theta}^{new}$ on the macroscopic elastic stiffness matrix $\boldsymbol{C}_M^i$. Essentially, we aim for the solution of the problem postulated by eq. (8.6), where the quantity of interest $\boldsymbol{y}^{new}$ is the stiffness increase due the CNT reinforcement by considering the parameters $\boldsymbol{\theta}^{new}$. 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. 8.4 for the CNT/cement specimen, the two-scale model of fig. 8.7 for the CNT/mortar specimen and the three-scale model of fig. 8.9 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

153

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. 8.21. 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 $\mathcal{M}^2$ 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 8.21: 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 $P(\boldsymbol{\theta}^{new}|\boldsymbol{D})$ of the interfacial parameters.
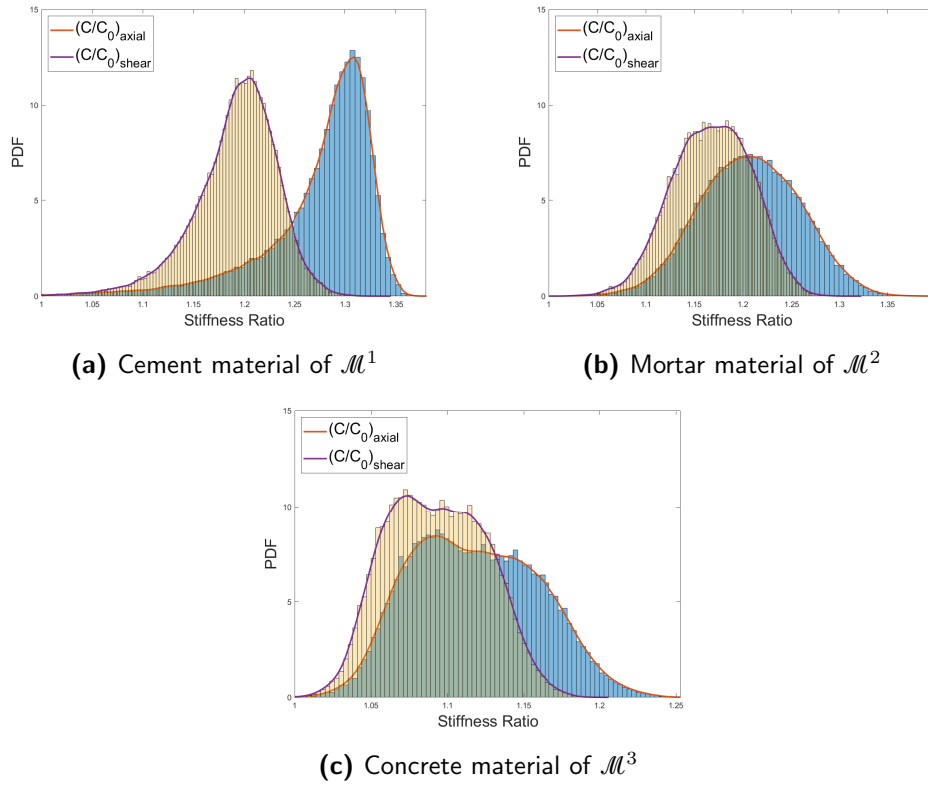
8.5 Conclusions

A hierarchical Bayesian strategy has been put forth in this work towards parameter identification of multiscale material systems. The suggested framework has the potential to integrate data that have been obtained from experiments performed on specimens with dissimilar material composition and incorporate measurements that take place in various length scales. To account for inherent model uncertainties due to environmental, operational and material discrepancies across the tests, the physical parameters were assigned to have a distinct definition for each dataset. This enabled a joint inference, where we utilized the posterior form of the hyperparameters to construct a new probabilistic representation of the physical parameters for future model predictions. To acquire precise forms of the posterior distributions from both the hyperparameters and the physical parameters, we employed the TMCMC method. To facilitate multiscale analyses, we opted for a computational homogenization approach due to its reliability, particularly in complex materials. The immerse computational demands were tackled through the employment of FFNNs that were tasked with predicting the constitutive response across multiple scales.

This study delved into the interfacial mechanical properties of CNTs within cementitious material configurations. Our approach involves calibrating the parameters that define a bond-slip cohesive law, utilizing experimental measurements from the literature. These experiments were conducted on different material compositions, specifically a CNT-reinforced cement paste specimen, a CNT-reinforced mortar coupon, and a CNT-reinforced concrete large-scale test. The initial step of our computational procedure focused on training surrogate models for each scale. These models were trained using a series of strain-stress responses derived from offline RVE solutions. The exceptional predictive accuracy achieved during the training phase enabled us to conduct the hierarchical Bayesian analysis in a cost-effective manner without compromising the precision of the results. The analysis yielded the posterior distributions of the studied interfacial parameters, which were then used for the investigation of the mechanical behavior of the three CNT-reinforced materials.

The computational framework put forth herein can be straightforwardly applied towards parameter inference to other multiscale material models and enables the efficient study of multiphysics problems such as the joint investigation of mechanical, thermal and electrical material attributes. The proposed strategy can also be generalized to

problems where data are derived from measurements across various length scales such as, for example, measurements recorded at the microscale and/or the mesoscale using the Digital Image Correlation technique. These directions remain in the scope of a future work.

9
Summary

This thesis presented cost-efficient and accurate computational frameworks for contemporary composite material design, characterization and discovery. The exceptional predictive capabilities of nested computational homogenization schemes were harnessed herein. Surrogate modeling strategies based on machine learning were utilized to dramatically reduce the prohibitive computational requirements of nested schemes on multi-query analyses. Based on this, a series of methodologies were proposed for forward and inverse uncertainty quantification problems in material engineering. Although, all the introduced methodologies were presented through the lens of a specific nanocomposite material, namely CNT-reinforced composites, they can be straightforwardly generalized to other composite materials.

9.1 INNOVATION OF THESIS

First, a methodology for updating the beliefs of the uncertain parameters that lie in the microscale of nanocomposite systems was proposed. In particular, the Bayesian framework by means of the MCMC technique was employed on top of FE^2 analyses to learn the parameters of the CNT/polymer interface. The elaborated methodology utilized measurements from the macroscale structure to update the prior beliefs on the nonlinear parameters, rather than expensive and hard-to-obtain microscale measurements. In

addition, to tackle the immense computational effort of performing Bayesian update on this type of problem, a neural network surrogate was developed in order to replace the nonlinear relation of the homogenization scheme. This surrogate model displayed a high level of accuracy compared to the full scale system solution as well as a remarkable cost reduction. This allowed to perform BU on complex large-scale problems, which would otherwise be unfeasible. The proposed methodology was demonstrated on mechanical problems, where the uncertain parameters were successfully updated in a reasonable computational time.

Furthermore, a novel accurate and computationally efficient surrogate modeling strategy was proposed, in this thesis, for performing computational homogenization on nonlinear and parametrized multiscale systems that involve more than two length scales. A DNN hierarchy was employed as a surrogate model of the material behavior, where each DNN was assigned to learn the constitutive law at the respective scale of the problem. In contrast to the direct computational homogenization, which is inherently coupled, the proposed approach enables the decoupling of the whole procedure and it results in a dramatic decrease in computational cost. The proposed material modeling framework was demonstrated on the analysis of a large-scale structure comprised of CNT-reinforced concrete, modeled as a 4-scale system, where a sensitivity analysis was performed to assess the influence of the constitutive parameters on the macroscopic response. This otherwise unreachable problem, in terms of computational effort, was solved in a reasonable time, by virtue of the elaborated surrogate modelling strategy.

In addition, a novel numerical framework was developed for the analysis of structures made up of CNT-reinforced structures. In particular, a two-level hierarchical material model was proposed based on continuum micromechanics to characterize CNT-reinforced concrete. An extension of the FE^2 method, called FE^3 was employed to assess structural performance, where the excessive computational demands of the method were effectively bypassed by using a DNN sequence based surrogate as a substitute for the composite material. With this implementation, stochastic material optimization was performed, aiming at finding the optimal CNT orientation at different members of the macroscale structural systems that minimize the sum of the mean and standard deviation of the overall structural deformation. The results of this investigation indicated that the reinforcement of concrete with CNTs can lead to significant enhancement in structural performance. In addition, the extension of the current manufacturing capabilities to the design of structural members with appropriately aligned and oriented CNTs can

158

revolutionize structural design, leading to stronger, safer and more elegant structures.

Lastly, a hierarchical Bayesian strategy was put forth in this thesis towards parameter identification of multiscale material systems. The suggested framework has the potential to integrate data that have been obtained from experiments performed on specimens with dissimilar material composition and incorporate measurements that take place in various length scales. To account for inherent model uncertainties due to environmental, operational and material discrepancies across the tests, the physical parameters were assigned to have a distinct definition for each dataset. This enabled a joint inference, where the posterior form of the hyperparameters was utilized to construct a new probabilistic representation of the physical parameters for future model predictions. To acquire precise forms of the posterior distributions from both the hyperparameters and the physical parameters the TMCMC method was employed. To facilitate multiscale analyses, a computational homogenization approach was preferred due to its reliability, particularly in complex materials. The intense computational demands were tackled through the employment of DNNs that were tasked with predicting the constitutive response across multiple scales. This study delved into the interfacial mechanical properties of CNTs within cementitious material configurations. This approach involves calibrating the parameters that define a bond-slip cohesive law, utilizing experimental measurements from the literature. These experiments were conducted on different material compositions, specifically a CNT-reinforced cement paste specimen, a CNT-reinforced mortar coupon, and a CNT-reinforced concrete large-scale test.

9.2 Limitations and assumptions

In any PhD thesis, acknowledging the limitations and assumptions is crucial for providing a transparent and comprehensive understanding of the research conducted. This section elucidates the key limitations of the research, such as constraints in data availability, methodological constraints, or external factors that may have influenced the outcomes. Simultaneously, it outlines the assumptions made during the study, including the theoretical, methodological, or empirical premises accepted as true for the purpose of analysis.

Throughout this thesis, various surrogate modeling strategies were presented, all utilizing feed-forward neural networks (FFNNs). FFNNs are the most popular type of neural networks, favored for their simplicity and exceptional predictive capabilities.

They are particularly well-suited for predicting one-to-one strain-stress relationships, which were a central focus of the analyses conducted in this research. However, when addressing non-monotonic analyses, such as cyclic loading cases, it becomes essential to account for the path dependency inherent in complex materials like those exhibiting plasticity or damage. Predicting the behavior of these materials poses a significantly more intricate challenge for machine learning models. This complexity arises due to the extreme dimensionality of the data, which introduces numerous bottlenecks, including increased computational demands and the potential for overfitting.

In this thesis, a simplification was made in the finite element modeling of the studied applications. Specifically, low-fidelity models were formulated across all length scales of the multiscale models, resulting in a coarse finite element discretization. This choice was deliberate to maintain focus on the core objectives of each proposed numerical framework without unnecessary complexity. However, to achieve exceptional accuracy in finite element simulations, higher-fidelity models should be considered. These models, while providing greater precision, come with the inevitable drawback of significantly increased computational costs.

Several simplifications were made concerning the material behavior in the studied multiscale material models. For instance, simple constitutive responses were selected for the matrix materials, with linear elastic behavior assumed for polymers and plastic behavior for concrete. Additionally, in formulating the Representative Volume Elements (RVEs), several assumptions were made. These included ignoring carbon nanotube (CNT) functionalization and curvature, and assuming perfect bonding between aggregates and cement. These choices were deliberate, reflecting the specific objectives and directions of this thesis. While more elaborate constitutive models for the matrix materials, along with the inclusion of additional micro- and meso-structural phenomena, could provide a more detailed understanding, they would also introduce an additional layer of complexity.

The applications studied in this thesis, which illustrate the proposed inverse uncertainty quantification frameworks, rely on experimental data. It is important to note that all measurements used in this thesis are either artificially generated or sourced from external literature. This decision was necessitated by the lack of both the necessary expertise and access to laboratory equipment to conduct in-situ experiments. Consequently, the effective implementation of the proposed model updating computational frameworks depended heavily on a key assumption: the reliability of the sourced data. This assumption is based on the expectation that standard procedures and guidelines were

followed during the experiments referenced in the external sources. While the frameworks account for moderate uncertainties through their probabilistic approach to parameter identification, they are not equipped to address cases involving fundamentally erroneous data.

9.3 FUTURE RESEARCH DIRECTIONS

Building on the limitations and assumptions discussed in the previous section, several promising avenues for future research emerge. These directions are proposed to address the constraints identified in this thesis and to further explore the unresolved aspects of the study. By investigating these areas, future research can build upon the current findings, refine the theoretical framework, and potentially offer new insights into the topic.

The application of more sophisticated neural network architectures beyond feed-forward neural networks (FFNNs), such as recurrent neural networks (RNNs), long short-term memory networks (LSTMs), or transformers, should be investigated. These advanced architectures may be better suited to address the path dependency and sequential nature of non-monotonic analyses, such as cyclic loading cases, and might enhance predictive accuracy for complex material behaviors. Additionally, hybrid models that combine FFNNs with other machine learning techniques or traditional computational methods should be developed and tested. For instance, integrating FFNNs with convolutional neural networks (CNNs) could capture both spatial and temporal dependencies in material behavior, potentially alleviating challenges associated with high-dimensional data. Methods for augmenting training datasets or generating synthetic data to improve the robustness and generalizability of machine learning models should also be explored. Techniques such as Generative Adversarial Networks (GANs) could be employed to create diverse and comprehensive datasets that better represent the complexities of material behavior under non-monotonic loading conditions. Furthermore, research should focus on enhancing the scalability and computational efficiency of surrogate modeling strategies, which could involve optimizing neural network architectures for faster training and inference e.g. by using Bayesian optimization for optimal neural network hyperparameter selection.

Future research could focus on the development and application of high-fidelity finite element models that offer greater accuracy compared to the low-fidelity models utilized

161

in this thesis. Future works could investigate advanced techniques and methodologies to balance the trade-off between accuracy and computational cost in high-fidelity models. This may include the use of adaptive mesh refinement techniques, or reduced-order modeling approaches to improve computational efficiency while maintaining accuracy. Linear techniques such as Principal Component Analysis (PCA), or nonlinear methods that uncover the data manifold such as kernel PCA or autoencoders may be employed to reduce computational demands and mitigate the risk of overfitting while preserving essential features of material behavior. Additionally, exploring innovative computational strategies and technologies to manage the increased computational demands of high-fidelity models would be valuable. Research could encompass parallel computing, high-performance computing (HPC) frameworks, or cloud-based solutions to improve scalability and reduce simulation time.

Future research could focus on developing and incorporating more sophisticated constitutive models for matrix materials. Investigating thermoviscoelastic, or thermoviscoplastic behaviors in polymers and employing phase field damage modeling in concrete will lead to more accurate simulations and predictions of material performance under various conditions. Additionally, further studies could explore the effects of carbon nanotube (CNT) functionalization and curvature on the mechanical properties of composite materials. Such research could offer insights into how these factors influence material behavior and performance at different scales. The development of more refined Representative Volume Elements (RVEs) that account for real-world complexities, such as imperfect bonding between aggregates and cement, is also warranted. This could involve creating models that incorporate interface properties and potential degradation effects to better simulate actual material behavior. Furthermore, future investigations might integrate additional stochastic micro-, meso- and macro- structural phenomena into multiscale models, including the effects of random topological features of the microstructure, or uncertain environmental and operational factors, to provide a more comprehensive understanding of material performance.

Vital further research could arise from conducting in-situ experiments to generate empirical data. This is essential for validating and refining the proposed inverse uncertainty quantification frameworks. This approach will help to confidently assess the effectiveness of the frameworks when applied to fully accounted experimental measurements, and counter the high uncertainty that comes with literature-sourced data. Additionally, expanding the range of datasets from various sources and experimental conditions will

facilitate the generalizability and adaptability across different scenarios. This could achieved through the utilization of data acquired from different material length scales, such as via Digital Image Correlation measurements on the micro- or meso- structure. Lastly, developing hybrid frameworks that integrate experimental data with simulated data can further enhance model accuracy and reliability. The combination of real data with high-quality simulated data, can mitigate the impact of data deficiencies and improve the overall performance of the proposed models.

References

[1] Abueidda, D. W., Koric, S., Sobh, N. A., & Sehitoglu, H. (2021). Deep learning for plasticity and thermo-viscoplasticity. *International Journal of Plasticity*, 136, 102852.

[2] As'ad, F., Avery, P., & Farhat, C. (2022). A mechanics-informed artificial neural network approach in data-driven constitutive modeling. *International Journal for Numerical Methods in Engineering*, 123(12), 2738–2759.

[3] Badogiannis, E. G., Sfikas, I. P., Voukia, D. V., Trezos, K. G., & Tsivilis, S. G. (2015). Durability of metakaolin self-compacting concrete. *Construction and Building Materials*, 82, 133–141.

[4] Barenblatt, G. (1962). The mathematical theory of equilibrium cracks in brittle fracture. *Advances in Applied Mechanics*, 7, 55 – 129.

[5] Baydin, A. G., Pearlmutter, B. A., Radul, A. A., & Siskind, J. M. (2017). Automatic differentiation in machine learning: A survey. *J. Mach. Learn. Res.*, 18(1), 5595–5637.

[6] Beck, J. L. & Katafygiotis, L. S. (1998). Updating models and their uncertainties. i: Bayesian statistical framework. *Journal of Engineering Mechanics*, 124(4), 455–461.

[7] Behmanesh, I., Moaveni, B., Lombaert, G., & Papadimitriou, C. (2015). Hierarchical bayesian model updating for structural identification. *Mechanical Systems and Signal Processing*, 64, 360–376.

[8] Berber, Kwon, & Tománek (2000). Unusually high thermal conductivity of carbon nanotubes. *Physical review letters*, 84 20, 4613–6.

[9] Bessa, M. A., Bostanabad, R., Liu, Z., Hu, A., Apley, D. W., Brinson, C., Chen, W., & Liu, W. K. (2017). A framework for data-driven analysis of materials under

uncertainty: Countering the curse of dimensionality. *Computer Methods in Applied Mechanics and Engineering*, 320, 633–667.

[10] Bhattacharjee, S. & Matouš, K. (2016). A nonlinear manifold-based reduced order model for multiscale analysis of heterogeneous hyperelastic materials. *Journal of Computational Physics*, 313, 635–653.

[11] Bjürnström, J., Martinelli, A., Matic, A., Bürjesson, L., & Panas, I. (2004). Accelerating effects of colloidal nano-silica for beneficial calcium–silicate–hydrate formation in cement. *Chemical Physics Letters*, 392(1), 242–248.

[12] Bolstad, W. M. (2009). *Understanding computational Bayesian statistics*, volume 644. John Wiley & Sons.

[13] Bonatti, C., Berisha, B., & Mohr, D. (2022). From cp-fft to cp-rnn: Recurrent neural network surrogate model of crystal plasticity. *International Journal of Plasticity*, 158, 103430.

[14] Budiansky, B. (1965). On the elastic moduli of some heterogeneous materials: Journal of the mechanics physics of solids.

[15] Burley, M., Campbell, J., Dean, J., & Clyne, T. (2018). Johnson-cook parameter evaluation from ballistic impact data via iterative fem modelling. *International Journal of Impact Engineering*, 112, 180–192.

[16] Chang, T.-P., Shih, J.-Y., Yang, K.-M., & Hsiao, T.-C. (2007). Material properties of portland cement paste with nano-montmorillonite. *Journal of Materials Science*, 42(17), 7478–7487.

[17] Chen, W.-H., Cheng, H.-C., & Liu, Y.-L. (2010). Radial mechanical properties of single-walled carbon nanotubes using modified molecular structure mechanics. *Computational Materials Science*, 47(4), 985–993.

[18] Ching, J. & Chen, Y.-C. (2007). Transitional markov chain monte carlo method for bayesian model updating, model class selection, and model averaging. *Journal of engineering mechanics*, 133(7), 816–832.

[19] Ching, J., Wu, S., & Phoon, K.-K. (2021). Constructing quasi-site-specific multivariate probability distribution using hierarchical bayesian model. *Journal of Engineering Mechanics*, 147(10), 04021069.

[20] Christensen, R. & Lo, K. (1979a). Solutions for effective shear properties in three phase sphere and cylinder models. *Journal of the Mechanics and Physics of Solids*, 27(4), 315–330.

[21] Christensen, R. & Lo, K. (1979b). Solutions for effective shear properties in three phase sphere and cylinder models. *Journal of the Mechanics and Physics of Solids*, 27(4), 315–330.

[22] Congdon, P. D. (2010). *Applied Bayesian hierarchical methods.* CRC Press.

[23] Cybenko, G. (1988). *Continuous Valued Neural Networks with Two Hidden Layers are Sufficient.* Technical report, Department of Computer Science, Tufts University.

[24] Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Math. Control Signal Systems*, 2, 303–314.

[25] Deng, J. & Gu, D. (2011). On a statistical damage constitutive model for rock materials. *Computers & Geosciences*, 37(2), 122–128.

[26] Deng, W., Qiu, W., Li, Q., Kang, Y., Guo, J., Li, Y., & Han, S. (2014). Multi-scale experiments and interfacial mechanical modeling of carbon nanotube fiber. *Experimental Mechanics*, 54, 3–10.

[27] Drucker, D. C. & Prager, W. (1952). Soil mechanics and plastic analysis or limit design. *Quarterly of Applied Mathematics*, 10, 157–165.

[28] Dugdale, D. (1960). Yielding of steel sheets containing slits. *Journal of the Mechanics and Physics of Solids*, 8(2), 100 – 104.

[29] Ebbesen, T. W., Lezec, H. J., Hiura, H., Bennett, J. W., Ghaemi, H. F., & Thio, T. (1996). Electrical conductivity of individual carbon nanotubes. *Nature*, 382, 54–56.

[30] Economides, A., Arampatzis, G., Alexeev, D., Litvinov, S., Amoudruz, L., Kulakova, L., Papadimitriou, C., & Koumoutsakos, P. (2021). Hierarchical bayesian uncertainty quantification for a model of the red blood cell. *Physical Review Applied*, 15(3), 034062.

[31] Efron, B. & Stein, C. (1981). The Jackknife Estimate of Variance. *The Annals of Statistics*, 9(3), 586 – 596.

166

[32] Eshelby, J. D. (1957). The determination of the elastic field of an ellipsoidal inclusion, and related problems. *Proceedings of the royal society of London. Series A. Mathematical and physical sciences*, 241(1226), 376–396.

[33] Eshelby, J. D. & Peierls, R. E. (1957). The determination of the elastic field of an ellipsoidal inclusion, and related problems. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 241(1226), 376–396.

[34] Feenstra, P. H. & De Borst, R. (1996). A composite plasticity model for concrete. *International Journal of Solids and Structures*, 33(5), 707–730.

[35] Feyel, F. (2003). A multilevel finite element method (fe2) to describe the response of highly non-linear structures using generalized continua. *Computer Methods in applied Mechanics and engineering*, 192(28-30), 3233–3244.

[36] Fritzen, F. & Kunc, O. (2018). Two-stage data-driven homogenization for nonlinear solids using a reduced order model. *European Journal of Mechanics - A/Solids*, 69, 201 – 220.

[37] Fuhg, J. N., Böhm, C., Bouklas, N., Fau, A., Wriggers, P., & Marino, M. (2021). Model-data-driven constitutive responses: Application to a multiscale computational framework. *International Journal of Engineering Science*, 167, 103522.

[38] Geers, M., Kouznetsova, V., & Brekelmans, W. (2010). Multi-scale computational homogenization: Trends and challenges. *Journal of Computational and Applied Mathematics*, 234(7), 2175 – 2182.

[39] Géron, A. (2017). *Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media, Inc., 1st edition.

[40] Ghaboussi, J., Garrett, J. H., & Wu, X. (1991). Knowledge-based modeling of material behavior with neural networks. *Journal of Engineering Mechanics*, 117(1), 132–153.

[41] Ghavamian, F. & Simone, A. (2019). Accelerating multiscale finite element simulations of history-dependent materials using a recurrent neural network. *Computer Methods in Applied Mechanics and Engineering*, 357, 112594.

[42] Girolami, M. & Rogers, S. (2005). Hierarchic bayesian models for kernel learning. In *Proceedings of the 22nd international conference on Machine learning* (pp. 241–248).

[43] Gojny, F., Wichmann, M., Köpke, U., Fiedler, B., & Schulte, K. (2004). Carbon nanotube-reinforced epoxy-composites: enhanced stiffness and fracture toughness at low nanotube content. *Composites Science and Technology*, 64(15), 2363–2371.

[44] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning.* The MIT Press.

[45] Graeff, A. G., Pilakoutas, K., Neocleous, K., & Peres, M. V. N. (2012). Fatigue resistance and cracking mechanism of concrete pavements reinforced with recycled steel fibres recovered from post-consumer tyres. *Engineering Structures*, 45, 385–395.

[46] Haario, H., Saksman, E., & Tamminen, J. (2001). An adaptive metropolis algorithm. *Bernoulli*, 7.

[47] Hagan, M. T. & Menhaj, M. B. (1994). Training feedforward networks with the marquardt algorithm. *IEEE Transactions on Neural Networks*, 5(6), 989–993.

[48] Hain, M. & Wriggers, P. (2008). Numerical homogenization of hardened cement paste. *Computational Mechanics*, 42, 197–212.

[49] Hansen, N. (2006). *The CMA Evolution Strategy: A Comparing Review*, (pp. 75–102). Springer Berlin Heidelberg: Berlin, Heidelberg.

[50] Hansen, N. (2016). The CMA evolution strategy: A tutorial. *CoRR*, abs/1604.00772.

[51] Hansen, N. & Kern, S. (2004). Evaluating the cma evolution strategy on multimodal test functions. In *Parallel Problem Solving from Nature - PPSN VIII* (pp. 282–291). Berlin, Heidelberg: Springer Berlin Heidelberg.

[52] Hashash, Y., Jung, S., & Ghaboussi, J. (2004). Numerical implementation of a neural network based material model in finite element analysis. *International Journal for numerical methods in engineering*, 59(7), 989–1005.

[53] Hashin, Z. & Rosen, B. W. (1964). The Elastic Moduli of Fiber-Reinforced Materials. *Journal of Applied Mechanics*, 31(2), 223–232.

[54] Hastings, W. K. (1970). Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1), 97–109.

[55] Hill, R. (1963). Elastic properties of reinforced solids: Some theoretical principles. *Journal of the Mechanics and Physics of Solids*, 11(5), 357–372.

[56] Hill, R. (1965). A self-consistent mechanics of composite materials. *Journal of the Mechanics and Physics of Solids*, 13(4), 213–222.

[57] Hoeffding, W. (1948). A Class of Statistics with Asymptotically Normal Distribution. *The Annals of Mathematical Statistics*, 19(3), 293 – 325.

[58] Homma, T. & Saltelli, A. (1996). Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering and System Safety*, 52(1), 1–17.

[59] Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5), 359–366.

[60] Huang, D., Fuhg, J. N., Weißenfels, C., & Wriggers, P. (2020). A machine learning based plasticity model using proper orthogonal decomposition. *Computer Methods in Applied Mechanics and Engineering*, 365, 113008.

[61] Hunashyal, A., Tippa, S. V., Quadri, S., & Banapurmath, N. (2011). Experimental investigation on effect of carbon nanotubes and carbon fibres on the behavior of plain cement mortar composite round bars under direct tension. *International Scholarly Research Notices*, 2011.

[62] Irshidat, M. R., Al-Saleh, M. H., & Sanad, S. H. (2015). Effect of nanoclay on expansive potential of cement mortar due to alkali-silica reaction. *Aci Materials Journal*, 112, 801–808.

[63] Jia, X., Sedehi, O., Papadimitriou, C., Katafygiotis, L. S., & Moaveni, B. (2022). Nonlinear model updating through a hierarchical bayesian modeling framework. *Computer Methods in Applied Mechanics and Engineering*, 392, 114646.

[64] Jo, B.-W., Kim, C.-H., ho Tae, G., & Park, J.-B. (2007). Characteristics of cement mortar with nano-sio2 particles. *Construction and Building Materials*, 21(6), 1351–1355.

169

[65] Kalogeris, I., Pyrialakos, S., Kokkinos, O., & Papadopoulos, V. (2022). Stochastic optimization of carbon nanotube reinforced concrete for enhanced structural performance. *Engineering with Computers*.

[66] Kingma, D. P. & Ba, J. (2015). Adam: A method for stochastic optimization. *CoRR*, abs/1412.6980.

[67] Konsta-Gdoutos, M. S., Metaxa, Z. S., & Shah, S. P. (2010a). Highly dispersed carbon nanotube reinforced cement based materials. *Cement and Concrete Research*, 40(7), 1052–1059.

[68] Konsta-Gdoutos, M. S., Metaxa, Z. S., & Shah, S. P. (2010b). Multi-scale mechanical and fracture characteristics and early-age strain capacity of high performance carbon nanotube/cement nanocomposites. *Cement and Concrete Composites*, 32(2), 110–115.

[69] Koutsourelakis, P.-S. (2012). A novel bayesian strategy for the identification of spatially varying material properties and model validation: an application to static elastography. *International Journal for Numerical Methods in Engineering*, 91(3), 249–268.

[70] Kouznetsova, V., Geers, M., & Brekelmans, W. (2004). Multi-scale second-order computational homogenization of multi-phase materials: a nested finite element solution strategy. *Computer Methods in Applied Mechanics and Engineering*, 193(48), 5525 – 5550.

[71] Kouznetsova, V., Geers, M. G. D., & Brekelmans, W. A. M. (2002). Multi-scale constitutive modelling of heterogeneous materials with a gradient-enhanced computational homogenization scheme. *International Journal for Numerical Methods in Engineering*, 54(8), 1235–1260.

[72] Le, B., Yvonnet, J., & He, Q.-C. (2015). Computational homogenization of nonlinear elastic materials using neural networks. *International Journal for Numerical Methods in Engineering*, 104(12), 1061–1084.

[73] Lee, S. J., Rust, J. P., Hamouda, H., Kim, Y. R., & Borden, R. H. (2005). Fatigue cracking resistance of fiber-reinforced asphalt concrete. *Textile Research Journal*, 75(2), 123–128.

[74] Lefik, M., Boso, D., & Schrefler, B. (2009). Artificial neural networks in numerical modelling of composites. *Computer Methods in Applied Mechanics and Engineering*, 198(21-26), 1785–1804.

[75] Li, C. & Chou, T.-W. (2003). A structural mechanics approach for the analysis of carbon nanotubes. *International Journal of Solids and Structures*, 40(10), 2487–2499.

[76] Li, G. Y., Wang, P. M., & Zhao, X. (2005). Mechanical behavior and microstructure of cement composites incorporating surface-treated multi-walled carbon nanotubes. *Carbon*, 43(6), 1239–1245.

[77] Li, H., hua Zhang, M., & ping Ou, J. (2007). Flexural fatigue performance of concrete containing nano-particles for pavement. *International Journal of Fatigue*, 29(7), 1292–1301.

[78] Li, M., Gao, C., Hu, H., & Zhao, Z. (2013). Electrical conductivity of thermally reduced graphene oxide/polymer composites with a segregated structure. *Carbon*, 65, 371–373.

[79] Li, X., Roth, C. C., & Mohr, D. (2019). Machine-learning based temperature- and rate-dependent plasticity model: Application to analysis of fracture experiments on dp steel. *International Journal of Plasticity*, 118, 320–344.

[80] Li, Z., Wang, H., He, S., Lu, Y., & Wang, M. (2006). Investigations on the preparation and mechanical properties of the nano-alumina reinforced cement composite. *Materials Letters*, 60(3), 356–359.

[81] Liu, D., Yang, H., Elkhodary, K., Tang, S., Liu, W. K., & Guo, X. (2022). Mechanistically informed data-driven modeling of cyclic plasticity via artificial neural networks. *Computer Methods in Applied Mechanics and Engineering*, 393, 114766.

[82] Lu, X., Giovanis, D., Yvonnet, J., Papadopoulos, V., Detrez, F., & Bai, J. (2018). A data-driven computational homogenization method based on neural networks for the nonlinear anisotropic electrical response of graphene/polymer nanocomposites. *Computational Mechanics*.

[83] Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2), 431–441.

[84] Masi, F., Stefanou, I., Vannucci, P., & Maffi-Berthier, V. (2021). Thermodynamics-based artificial neural networks for constitutive modeling. *Journal of the Mechanics and Physics of Solids*, 147, 104277.

[85] Matouš, K., Geers, M. G., Kouznetsova, V. G., & Gillman, A. (2017). A review of predictive nonlinear theories for multiscale modeling of heterogeneous materials. *Journal of Computational Physics*, 330, 192–220.

[86] McCormick, N. & Lord, J. (2010). Digital image correlation. *Materials today*, 13(12), 52–54.

[87] Mckay, M. D., Beckman, R. J., & Conover, W. J. (2000). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 42(1), 55–61.

[88] Miehe, C. & Koch, A. (2002). Computational micro-to-macro transitions of discretized microstructures undergoing small strains. *Archive of Applied Mechanics*, 72, 300–317.

[89] Mori, T. & Tanaka, K. (1973). Average stress in matrix and average elastic energy of materials with misfitting inclusions. *Acta Metallurgica*, 21(5), 571–574.

[90] Mousavi, M. A., Sadeghi-Nik, A., Bahari, A., Jin, C., Ahmed, R., Ozbakkaloglu, T., & de Brito, J. (2021). Strength optimization of cementitious composites reinforced by carbon nanotubes and titania nanoparticles. *Construction and Building Materials*, 303, 124510.

[91] Mozaffar, M., Bostanabad, R., Chen, W., Ehmann, K., Cao, J., & Bessa, M. A. (2019). Deep learning predicts path-dependent plasticity. *Proceedings of the National Academy of Sciences*, 116(52), 26414–26420.

[92] Nagel, J. B. & Sudret, B. (2014). A bayesian multilevel approach to optimally estimate material properties. In *Vulnerability, uncertainty, and risk: Quantification, mitigation, and management* (pp. 1504–1513). ASCE.

[93] Nagel, J. B. & Sudret, B. (2016). A unified framework for multilevel uncertainty quantification in bayesian inverse problems. *Probabilistic Engineering Mechanics*, 43, 68–84.

[94] Naji, H. F., Khalid, N. N., Alsaraj, W. K., Habouh, M. I., & Marchetty, S. (2021). Experimental investigation of flexural enhancement of rc beams with multi-walled carbon nanotubes. *Case Studies in Construction Materials*, 14, e00480.

[95] Needleman, A. & RI, P. (1990). An analysis of decohesion along an imperfect interface. *International Journal of Fracture - INT J FRACTURE*, 42, 21–40.

[96] Nguyen, T., Francom, D. C., Luscher, D. J., & Wilkerson, J. (2021). Bayesian calibration of a physics-based crystal plasticity and damage model. *Journal of the Mechanics and Physics of Solids*, 149, 104284.

[97] Nochaiya, T. & Chaipanich, A. (2011). Behavior of multi-walled carbon nanotubes on the porosity and microstructure of cement-based materials. *Applied Surface Science*, 257(6), 1941–1945.

[98] Pacheco-Torgal, F. & Labrincha, J. (2013). Biotech cementitious materials: Some aspects of an innovative approach for concrete with enhanced durability. *Construction and Building Materials*, 40, 1136–1141. Special Section on Recycling Wastes for Use as Construction Materials.

[99] Papadimas, N. & Dodwell, T. (2021). A hierarchical bayesian approach for calibration of stochastic material models. *Data-Centric Engineering*, 2, e20.

[100] Papanikos, P., Nikolopoulos, D., & Tserpes, K. (2008). Equivalent beams for carbon nanotubes. *Computational Materials Science*, 43(2), 345–352.

[101] Patsialis, D., Kyprioti, A. P., & Taflanidis, A. A. (2020). Bayesian calibration of hysteretic reduced order structural models for earthquake engineering applications. *Engineering Structures*, 224, 111204.

[102] Puck, A. & Schürmann, H. (2002). Failure analysis of frp laminates by means of physically based phenomenological models. *Composites Science and Technology*, 62(12), 1633 – 1662.

[103] Pyrialakos, S., Kalogeris, I., & Papadopoulos, V. (2023). Multiscale analysis of nonlinear systems using a hierarchy of deep neural networks. *International Journal of Solids and Structures*, 271, 112261.

[104] Pyrialakos, S., Kalogeris, I., Sotiropoulos, G., & Papadopoulos, V. (2021). A neural network-aided bayesian identification framework for multiscale modeling of nanocomposites. *Computer Methods in Applied Mechanics and Engineering*, 384, 113937.

[105] Rappel, H., Beex, L. A., & Bordas, S. P. (2018). Bayesian inference to identify parameters in viscoelasticity. *Mechanics of Time-Dependent Materials*, 22, 221–258.

[106] Raza, S. S., Qureshi, L. A., Ali, B., Raza, A., & Khan, M. M. (2021). Effect of different fibers (steel fibers, glass fibers, and carbon fibers) on mechanical properties of reactive powder concrete. *Structural Concrete*, 22(1), 334–346.

[107] Reimer, L. (2000). Scanning electron microscopy: physics of image formation and microanalysis. *Measurement Science and Technology*, 11(12), 1826–1826.

[108] Rejeb, S. K. (1996). Improving compressive strength of concrete by a two-step mixing method. *Cement and Concrete Research*, 26(4), 585–592.

[109] Reuß, A. (1929). Berechnung der fließgrenze von mischkristallen auf grund der plastizitätsbedingung für einkristalle. *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik*, 9(1), 49–58.

[110] Rocha, I., Kerfriden, P., & van der Meer, F. (2021). On-the-fly construction of surrogate constitutive models for concurrent multiscale mechanical analysis through probabilistic machine learning. *Journal of Computational Physics: X*, 9, 100083.

[111] Rodriguez-Camacho, R. & Uribe-Afif, R. (2002). Importance of using the natural pozzolans on concrete durability. *Cement and Concrete Research*, 32(12), 1851–1858.

[112] Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., & Tarantola, S. (2010). Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index. *Computer Physics Communications*, 181(2), 259–270.

174

[113] Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., & Tarantola, S. (2008). *Global Sensitivity Analysis. The Primer.* Wiley.

[114] Savvas, D., Papadopoulos, V., & Papadrakakis, M. (2012). The effect of interfacial shear strength on damping behavior of carbon nanotube reinforced composites. *International Journal of Solids and Structures*, 49(26), 3823–3837.

[115] Savvas, D., Stefanou, G., Papadopoulos, V., & Papadrakakis, M. (2016a). Effect of waviness and orientation of carbon nanotubes on random apparent material properties and rve size of cnt reinforced composites. *Composite Structures*, 152, 870 – 882.

[116] Savvas, D., Stefanou, G., & Papadrakakis, M. (2016b). Determination of rve size for random composites with local volume fraction variation. *Computer Methods in Applied Mechanics and Engineering*, 305, 340 – 358.

[117] Sedehi, O., Papadimitriou, C., Teymouri, D., & Katafygiotis, L. S. (2019). Sequential bayesian estimation of state and input in dynamical systems using output-only measurements. *Mechanical Systems and Signal Processing*, 131, 659 – 688.

[118] Smit, R., Brekelmans, W., & Meijer, H. (1998). Prediction of the mechanical behavior of nonlinear heterogeneous systems by multi-level finite element modeling. *Computer Methods in Applied Mechanics and Engineering*, 155(1), 181–192.

[119] Sobol', I. (1967). On the distribution of points in a cube and the approximate evaluation of integrals. *USSR Computational Mathematics and Mathematical Physics*, 7(4), 86–112.

[120] Sobol, I. (2001). Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. *Mathematics and Computers in Simulation*, 55(1), 271–280.

[121] Soden, P., Hinton, M., & Kaddour, A. (2002). Biaxial test results for strength and deformation of a range of e-glass and carbon fibre reinforced composite laminates: failure exercise benchmark data. *Composites Science and Technology*, 62(12), 1489 – 1514.

175

[122] Song, P. & Hwang, S. (2004). Mechanical properties of high-strength steel fiber-reinforced concrete. *Construction and Building Materials*, 18(9), 669–673.

[123] Suquet, P. (1987). Elements of homogenization for inelastic solid mechanics. *Homogenization techniques for composite media*, 272, 193–278.

[124] Tassew, S. & Lubell, A. (2014). Mechanical properties of glass fiber reinforced ceramic concrete. *Construction and Building Materials*, 51, 215–224.

[125] Temizer, I. & Zohdi, T. (2007). A numerical method for homogenization in non-linear elasticity. *Computational Mechanics*, 40(2), 281–298.

[126] Thomas, A. J., Barocio, E., Bilionis, I., & Pipes, R. B. (2022). Bayesian inference of fiber orientation and polymer properties in short fiber-reinforced polymer composites. *Composites Science and Technology*, 228, 109630.

[127] Thostenson, E. T. & Chou, T.-W. (2003). On the elastic properties of carbon nanotube-based composites: modelling and characterization. *Journal of Physics D: Applied Physics*, 36(5), 573–582.

[128] Thostenson, E. T., Ren, Z., & Chou, T.-W. (2001). Advances in the science and technology of carbon nanotubes and their composites: a review. *Composites Science and Technology*, 61(13), 1899–1912.

[129] Toutanji, H., Delatte, N., Aggoun, S., Duval, R., & Danson, A. (2004). Effect of supplementary cementitious materials on the compressive strength and durability of short-term cured concrete. *Cement and Concrete Research*, 34(2), 311–319.

[130] Treacy, M., Ebbesen, T., & Gibson, J. (1996). Exceptionally high young's modulus observed for individual carbon nanotubes. *Nature*, 381(6584), 678–680.

[131] Tvergaard, V. & Hutchinson, J. W. (1992). The relation between crack growth resistance and fracture process parameters in elastic-plastic solids. *Journal of the Mechanics and Physics of Solids*, 40(6), 1377–1397.

[132] Voigt, W. (1889). Ueber die beziehung zwischen den beiden elasticitätsconstanten isotroper körper. *Annalen der physik*, 274(12), 573–587.

[133] Vu, G., Diewald, F., Timothy, J. J., Gehlen, C., & Meschke, G. (2021). Reduced order multiscale simulation of diffuse damage in concrete. *Materials*, 14(14).

[134] Vu-Bac, N., Rafiee, R., Zhuang, X., Lahmer, T., & Rabczuk, T. (2015). Uncertainty quantification for multiscale modeling of polymer nanocomposites with correlated parameters. *Composites Part B: Engineering*, 68, 446 – 464.

[135] Wang, J., Zhang, L., & Liew, K. (2017). Multiscale simulation of mechanical properties and microstructure of cnt-reinforced cement-based composites. *Computer Methods in Applied Mechanics and Engineering*, 319, 393–413.

[136] Wang, K. & Sun, W. (2018). A multiscale multi-permeability poroplasticity model linked by recursive homogenizations and deep learning. *Computer Methods in Applied Mechanics and Engineering*.

[137] Wang, Y., Cheung, S. W., Chung, E. T., Efendiev, Y., & Wang, M. (2020). Deep multiscale model learning. *Journal of Computational Physics*, 406, 109071.

[138] Wriggers, P. & Moftah, S. (2006). Mesoscale models for concrete: Homogenisation and damage behaviour. *Finite Elements in Analysis and Design*, 42(7), 623–636. The Seventeenth Annual Robert J. Melosh Competition.

[139] Wu, L., Zulueta, K., Major, Z., Arriaga, A., & Noels, L. (2020). Bayesian inference of non-linear multiscale model parameters accelerated by a deep neural network. *Computer Methods in Applied Mechanics and Engineering*, 360, 112693.

[140] Wu, S., Angelikopoulos, P., Beck, J. L., & Koumoutsakos, P. (2019). Hierarchical stochastic model in bayesian inference for engineering applications: Theoretical implications and efficient approximation. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, 5(1), 011006.

[141] Xu, K., Huang, D. Z., & Darve, E. (2021). Learning constitutive relations using symmetric positive definite neural networks. *Journal of Computational Physics*, 428, 110072.

[142] Xu, S., Liu, J., & Li, Q. (2015). Mechanical properties and microstructure of multi-walled carbon nanotube-reinforced cement paste. *Construction and Building Materials*, 76, 16–23.

[143] Yu, A., Ramesh, P., Itkis, M. E., Bekyarova, E., & Haddon, R. C. (2007). Graphite nanoplatelet-epoxy composite thermal interface materials. *The Journal of Physical Chemistry C*, 111(21), 7565–7569.

177

[144] Yue, L., Heuzey, M.-C., Jalbert, J., & Lévesque, M. (2022). On the parameters identification of three-dimensional aging-temperature-dependent viscoelastic solids through a bayesian approach. *Mechanics of Time-Dependent Materials*, (pp. 1–23).

[145] Yvonnet, J. (2019). *Computational Homogenization of Heterogeneous Materials with Finite Elements*. Springer.

[146] Yvonnet, J. & He, Q.-C. (2007). The reduced model multiscale method (r3m) for the non-linear homogenization of hyperelastic media at finite strains. *Journal of Computational Physics*, 223(1), 341–368.

[147] Zaheer, M. M., Jafri, M. S., & Sharma, R. (2019). Effect of diameter of mwcnt reinforcements on the mechanical properties of cement composites. *Adv. Concr. Constr*, 8(3), 207–215.