NATIONAL TECHNICAL UNIVERSITY OF ATHENS

School of Chemical Engineering

Postgraduate Program: ''Computational Mechanics''

Bayesian Update on Nonlinear Multiscale Systems

Master Thesis

Stefanos-Christos Pyrialakos

Supervisor: Associate Professor Vissarion Papadopoulos

Athens 2020

Contents

List of Figures

1.

2.

3.

Acknowledgements

I would like to express my sincere gratitude to Associate Professor Vissarion Papadopoulos for giving me the opportunity to work on a dissertation that provided me with a lot of knowledge on a wide spectrum of topics and made me realize my high interest in computational engineering, while at the same time guiding me on every stage of this project. Also, I would like to thank Ioannis Kalogeris and Gerasimos Sotiropoulos for helping me willingly whenever I needed it while completing this dissertation.

Introduction

Mechanicals systems are often uncertain in their response under different conditions. Examples of such uncertainties include experimental errors, lack of data, uncertain model parameters, and systematic model inadequacy. This uncertainty is most of the times a hard task to define properly. The overall uncertainty can be decomposed on a number of different parameters that affect the system with several ways. These parameters are not always straight forward and the connection of their probabilistic structure with that of the whole system's may not be possible to be found from beforehand. A characteristic example of this issue is the investigation of the uncertainty of the microstructure for heterogeneous materials. The direct investigation of the microstructure is a very challenging and costly task. This issue can be countered efficiently by taking advantage of the high amount of information and data that can be collected and stored while observing the system's response macroscopically and with a multiscale approach update the random microscale parameters.

The Bayesian framework, which combines new information with preexisting models, can be used in order to update the microscopic scales with acquired data from the investigation of either those scales or related macroscopic scales. Several studies have been done for the stochastic formulation of microstructures that ultimately define the macrostructure (i.e. random wrinkling of graphene sheets [16], stochastic geometric formulation of carbonate rocks [17]), without however any consideration of real data that can reformulate the microstructure. On the other hand, multiscale Bayesian updating has been implemented mainly on image processing studies (i.e. speckle removal in ultrasound images [18], image segmentation [19,20]). However, Bayesian updating on microscopic stochastic parameters of heterogeneous materials based on multiscale procedures have yet to be investigated thoroughly, in spite of the useful information that can be gathered from it.

Here the Bayesian updating of the probabilistic structure of the uncertain parameters describing the interaction between carbon nanotubes (CNTs) and the surrounding polymer matrix is considered. A MCMC technique with the use of the Metropolis Hastings is implemented on a semi-concurrent multiscale approach (FE^2) in order to update the beliefs of the microscopic parameters according to a macroscale deformation observation. The posterior distributions of these parameters are then compared to deterministic solutions of the multiscale model in the direction of verifying the efficiency of the procedure.

1. Multiscale analysis

The modeling of heterogeneous materials can be accomplished with several ways. The most straightforward way it to create a single scale macroscopic model. This model relies on brute computational force, since it has to be densely discretized in order to include all the heterogeneities. It is obvious that such a model can be used in some special occasions only, such as structures with linear material behavior and highly distinguishable heterogeneities. In an effort to overcome this issue multiscale approaches have been developed. These approaches can be classified in two main categories [2]:

2. Sequential methods.

A thorough microscopic analysis gives all those mechanisms responsible for the macroscopic behavior of the structure. A macroscopic model is produced by taking into consideration all of the above mechanisms. This model can then be solved by discretizing it with some method. Although this method is easy to implement and with low computational effort, much of the information is lost and the progression of the model in different phases cannot be accurately predicted.

2. Integrated methods.

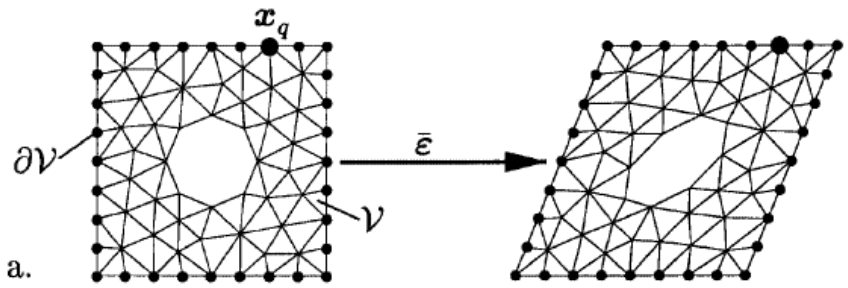
The microscopic scale can be linked with the macroscopic by using some homogenization rule. This can be achieved by discretizing the microstructure and with the use of some analytical equations describe the model progression. Then all the information is transferred to the macromodel through the established link. This method is computationally heavier than the previous approach, but it can accurately predict the model progression through different phases.

The discretized microstructure is expressed in the form of the so-called Representative Volume Element (RVE), in which numerical computations take place and then according to a homogenization scheme the results are transferred to the macrostructure.

The second method is used in this thesis with a two-scale first-order homogenization analysis.

Homogenization scheme

The homogenization procedure was developed according to [1] by imposing linear displacement as boundaries on the RVE.



1.1 Linear displacements as boundary conditions

We assume an equilibrium state:

$$div\sigma + \rho\ddot{u} = 0 \text{ in } B$$

where B the solid part of the RVE

By integrating and applying the Gauss theorem in the above relation:

$$\int_{\partial B} \sigma n dA - \int_{B} \rho_0 \ddot{u} dV = 0$$

or

$$\int_{\partial B} x \times \sigma n dA - \int_{B} x \times \rho_0 \ddot{u} dV = 0$$

The averaging theorem [8] equates the average microscopic stress with the macroscopic stress:

$$\bar{\sigma} = \frac{1}{|V|} sym \int_{V} \sigma dV$$

or

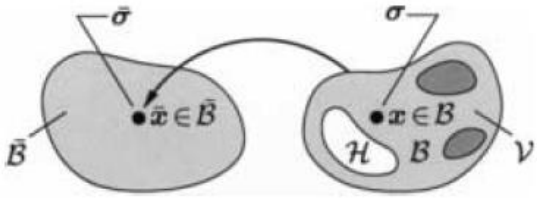
$$\bar{\sigma} = \frac{1}{|V|} \int_{dV} sym[t \otimes x] dA$$

In a similar manner:

$$\bar{\varepsilon} = \frac{1}{|V|}\int_{V} \varepsilon dV$$

or

$$\bar{\varepsilon} = \frac{1}{|V|}\int_{dV} sym[\mathrm{u} \otimes n] dA$$



1.2 Association of a point in the macrostructure with a heterogenous microstructure

The boundary conditions on the microstructure are chosen in order for the condition

$$\bar{\boldsymbol{\sigma}} : \dot{\bar{\boldsymbol{\varepsilon}}} = \frac{1}{|\mathcal{V}|}\int_{\partial \mathcal{V}} \mathbf{t} \cdot \dot{\mathbf{u}}\, \mathrm{d}A$$

to be satisfied a priori.

The assumption that the tractions are zero in the interior part of the RVE is made.

On the other hand, the boundary conditions on the dV must be chosen in a relative manner to the macroscopic point behavior. These conditions may take the form of periodic/aperiodic displacements or tractions.

By formulating the boundary conditions in terms of displacements that are dependent on the macroscopic strain $\bar{\varepsilon}$ we get the form:

$$u(x,t) = \bar{\varepsilon}(t)x \text{ at } x \in dV$$

which is a linear deformation of the boundary.

A discretization technique has to be implemented on the above analytical equations of the RVE for the purpose of computing the macroscopic properties. On top of that in the nonlinear case a convergent iterative procedure must be carried out. A FEM with a standard Newton-Raphson method will be used here.

If n are the nodes on the boundary dV after the discretization the imposed boundary equation takes the form:

$u_n = \bar{\varepsilon} x_n$ for $n = 1,2,3...$

By writing $\bar{\varepsilon}$ and u in the form:

$\bar{\varepsilon} = [\bar{\varepsilon}_{11} \;\; \bar{\varepsilon}_{22} \;\; \bar{\varepsilon}_{12}]^{\mathrm{T}}$

$uq = [u_1 \;\; u_2]_n^T$

The previous equation can be written as

$u_n = D_n^T \bar{\varepsilon}$, where

$$D_n^T = \frac{1}{2}\begin{bmatrix} 2x_1 & 0 & x_2 \\ 0 & 2x_2 & x_1 \end{bmatrix}_n$$

The complete set of the system's nonlinear equations is:

$$f_a = 0$$
$$f_b = \lambda$$
$$u_b = D^T \bar{\varepsilon}$$

Where λ are the external forces on the system nodes formulated as a Lagrange multiplier.

By separating the external from the internal nodes, the tangent stiffness matrix along with the internal forces can be written as:

$$K = \begin{bmatrix} K_{aa} & K_{ab} \\ K_{ba} & K_{bb} \end{bmatrix} \qquad f = \begin{bmatrix} f_a \\ f_b \end{bmatrix}$$

The linearized system equations then take the form:

$$\begin{bmatrix} K_{aa} & K_{ab} \\ K_{ba} & K_{bb} \end{bmatrix} \begin{Bmatrix} \Delta u_a \\ \Delta u_b \end{Bmatrix} + \begin{Bmatrix} f_a \\ f_b \end{Bmatrix} = \begin{Bmatrix} 0 \\ \Delta\lambda \end{Bmatrix}$$

The procedure to solve the nonlinear equation in linear increments is:

1) For the first iteration of the increment where the microscopic equilibrium is satisfied a displacement increment Δu_b is imposed as $\Delta u_b = D_n^T \Delta\bar{\varepsilon}$ resulting in a $\Delta u_a = K_{aa}^{-1} K_{ab} \Delta u_b$
2) For the following iterations where Δu_b =0 the internal deformations are updated as $\Delta u_a = -K_{aa}^{-1} f_a$ until convergence has been achieved in the sense that $|f_a| \simeq 0$
3) If convergence has been reached the macroscopic stress and the macroscopic tangent moduli are calculated:

   $\bar{\sigma} = \frac{1}{|V|} D\delta$, $\bar{\mathrm{C}} = \frac{1}{|V|} D\tilde{K}_{bb} D^T$, where $\tilde{K}_{bb} = K_{bb} - K_{ba} K_{aa}^{-1} K_{ab}$
4) Continue to the next increment

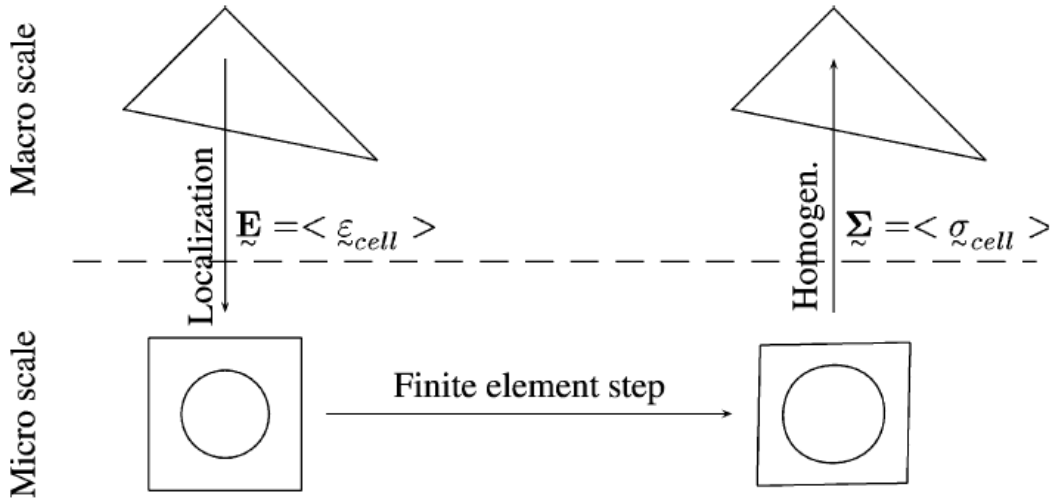
FE^2 solution scheme

FE2 is an integrated multiscale method proposed by [2] where simultaneous computations of the mechanical response at two different scales take place. The macroscopic scale represents the whole structure while the microscopic is the RVE at each macroscopic numerical integration point (i.e. Gauss-Legendre quadrature). With this method all the macroscopic phenomenological relations are not required even in the nonlinear case, since these relations are emerging fully

from the microscopic scale and affect directly the mechanical properties of the macroscopic system.

The three main steps that characterize the FE2 procedure are:

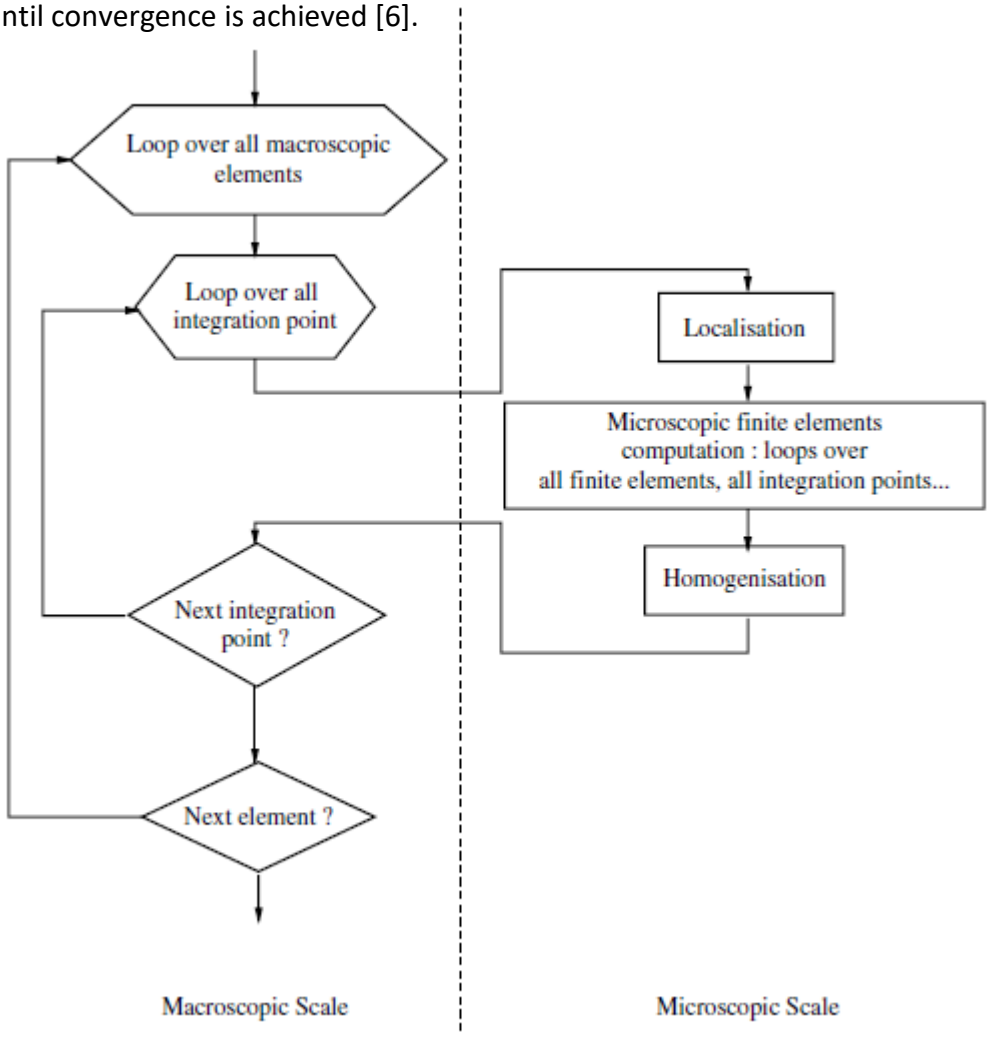
(1) a modeling of the mechanical behavior at the RVE and the discretization of it.
(2) a localization rule which determines the local solutions inside the RVE, for any given overall strain.
(3) a homogenization rule giving the macroscopic stress tensor, knowing the micromechanical stress state.



1.3 Multiscale steps with FE^2 procedure

In a more detailed way, the procedure followed here can be described as follows. The macroscopic model, which consists of a microscopically heterogeneous material is analyzed simultaneously on two separate levels. On the macroscopic level, the model is analyzed as if it consisted of a homogeneous material with nonlinear behavior, using a standard Newton–Raphson iterative procedure. The FE^2 solution scheme dictates that material properties at the macroscopic level are not calculated using a constitutive law, but by solving an appropriate boundary value problem at the microscopic level, using a homogenization technique. The steps of the method can be summarized as follows: The macroscopic structure is discretized and an appropriate RVE is chosen to represent the microstructure, which is subsequently discretized as well. For the first step of the Newton–Raphson algorithm, zero displacements are assigned to all

the nodes of the macrostructure. The macroscopic nodal displacements are used to calculate the macroscopic strain ε at every Gauss Point, using the macroscopic shape function derivatives B. Next, the macroscopic strain is used to apply appropriate boundary displacements to every RVE, according to the localization rule of the homogenization scheme. After the solution of the resulting boundary value problem, the macroscopic stress σ and the macroscopic tangent modulus C at every macroscopic Gauss Point are calculated according to the homogenization rule of the homogenization scheme. These are then used to calculate the internal nodal force vector Q and tangential stiffness matrix K of the macroscopic structure. If the internal nodal force vector is in equilibrium with the external one, the procedure is terminated. Else, the macroscopic nodal displacements are updated according to the Newton–Raphson algorithm and the procedure is repeated, until convergence is achieved [6].



1.4 Flow chart of FE^2 algorithm

Representation of CNTs

At the lowest scale the interatomic relations between carbon atoms are governed by a quadratic force field potential. With the use of the mMSM approach [21] the C-C covalent bonds are substituted by beam elements, making feasible the modeling of the atomic lattice of the carbon nanotube with a space frame structure at the nanoscale.

The C-C bond potential energy is calculated as a function of the bond stretching, bond in-plane bending, torsion and out of plane bending. Thus, the relations for the calculations of the beam element properties are derived:

$$k_r = \frac{EA}{L}, \quad k_\theta = \frac{E(I_y'' + I_z)}{L}, \quad k_\omega = \frac{3EI_y'}{L}, \quad k_\tau = \frac{GJ}{L}$$

$$I_y = I_y' + I_y'' = \left(\frac{k_\omega}{3k_\theta} + \tan\phi\right) I_z$$

Although an efficient modeling of the carbon atom interactions has been achieved with the mMSM approach the formed space-frame structure is still significantly complex and requires high computational effort. In order to simplify this structure without substantial loss of the mechanical properties the mMSM model is projected to an equivalent beam element (EBE) with linear behavior. This model transition can be achieved by subjecting the mMSM structure of length L_0 to an axial displacement u_x, a transverse displacement u_y and a torsional displacement ϕ respectively at the one end while the other end is fixed. By calculating the resulting forces in the fixed end for each loading case the mechanical properties of the EBE can be derived [4]:

1.5 FE mesh and boundary conditions for different types of loading

$$(EA)_{eq} = \frac{F_x L_0}{u_x}, \quad (EI)_{eq} = \frac{F_y}{3u_y} L_0^3, \quad (GI)_{eq} = \frac{T}{\phi} L_0$$

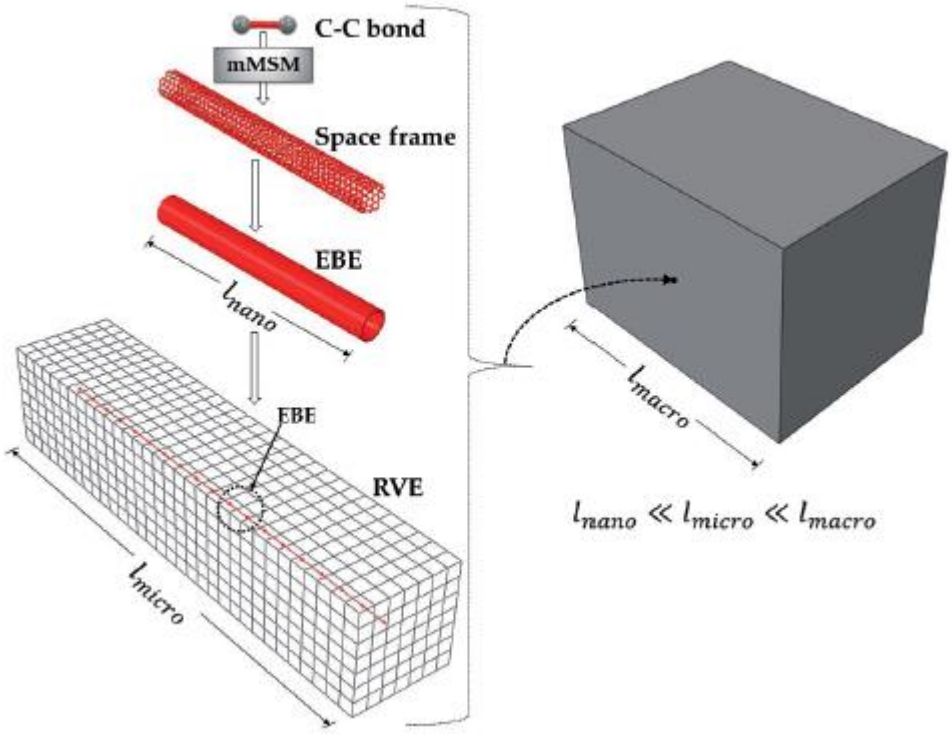
The profile section of the beam is considered to be a solid cylinder with diameter $D_{eq.}$

The section properties are given by:

$$A_{eq} = \frac{\pi}{4} D_{eq}^2 \qquad I_{eq} = \frac{\pi}{64} D_{eq}^4$$

Consequently, the EBE's diameter can be written in terms of these properties as:

$$\frac{(EI)_{eq}}{(EA)_{eq}} = \frac{1}{16} D_{eq}^2$$



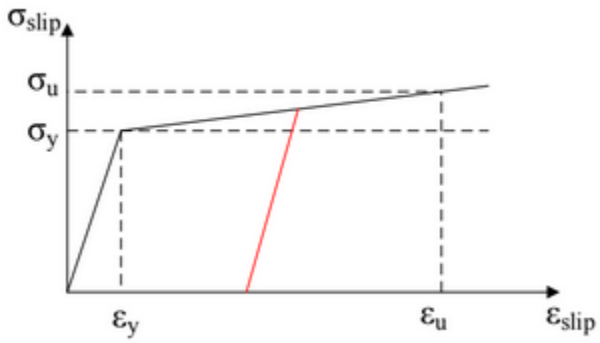
1.6 Hierarchical multiscale homogenization step

CNT-Polymer interaction

The CNT-Polymer interaction plays a major role in the overall response of the composite structure. There are several traction-separation approaches that can describe this interaction. A bilinear bond-slip with kinematic hardening approach is easy to be implemented without losing much accuracy compared to similar approaches such as an exponential.

The bilinear behavior of the slippage can be approximated according to the following relation:

$$D = \begin{cases} D_{slip} & , \epsilon_{slip} < \epsilon_y \\ (1/10)D_{slip} & , \epsilon_{slip} > \epsilon_y \end{cases}$$



1.7 Bond-Slip constitutive law

In order to simulate this interaction an analytical method that connects the CNT with the polymer surface must be considered. Such a method is the Cohesive Zone Model (CZM) [22] which is widely used in the fracture mechanics. According to CZM the domain Ω can be divided according to the existing discontinuity and create a new boundary surface Γ_{coh}. The boundary conditions over this surface can be denoted as:

$$\sigma_{ij}^{+} n_j^{+} = T_i^{+} = -T_i^{-} = -\sigma_{ij}^{-} n_j^{-} \quad \text{on} \quad \Gamma_{coh}$$

Where + is the positive (upper) part and – is the negative (lower) part of the discontinuity.

Writing the equilibrium equation in the weak form by applying the principle of virtual work the following is obtained:

$$\int_{\Omega+} b_i \delta u_i d\Omega + \int_{\partial\Omega+} \tilde{T}_i \delta u_i d\Gamma + \int_{\Gamma_c^+} T_i^+ \delta u_i^+ d\Gamma = \int_{\Omega+} \sigma_{ij} \delta E_{ij} d\Omega$$
$$\int_{\Omega-} b_i \delta u_i d\Omega + \int_{\partial\Omega-} \tilde{T}_i \delta u_i d\Gamma + \int_{\Gamma_c^-} T_i^- \delta u_i^- d\Gamma = \int_{\Omega-} \sigma_{ij} \delta E_{ij} d\Omega$$

These can be stated as one equation as:

$$\int_{\Omega} b_i \delta u_i d\Omega + \int_{\partial\Omega} \tilde{T}_i \delta u_i d\Gamma = \int_{\Omega} \sigma_{ij} \delta E_{ij} d\Omega - \int_{\Gamma_c^+} T_i^+ \left(\delta u_i^+ - \delta u_i^-\right) d\Gamma$$

By assuming a mid-surface between the two subdomain boundary surfaces, using the + surface as reference and transforming the coordinate system from local to global the equation is transformed to:

$$\int_{\Omega} b_i \delta u_i d\Omega + \int_{\partial\Omega} \tilde{T}_i \delta u_i d\Gamma = \int_{\Omega} \sigma_{ij} \delta E_{ij} d\Omega + \int_{\bar{\Gamma}_c} R_{ij} T_j \left(\delta u_i^+ - \delta u_i^-\right) d\Gamma$$

The only part of the equation that is going to be simulated with the CZM is:

$$\delta W^{int,ce_k} = \int_{ce_k} R_{ij} T_j \left(\delta u_i^+ - \delta u_i^-\right) d\Gamma$$

Implementing the standard FE procedure in order discretize the above equation to a finite number of dofs a matrix-vector notation is generated:

$$\left\{\mathbf{r}^{\mathbf{int,ce_k,top}}\right\} = \int_{-1}^{1} [\mathbf{N}]^T [\mathbf{R}] \{\mathbf{T}\} J d\xi$$
$$\left\{\mathbf{r}^{\mathbf{int,ce_k,bot}}\right\} = -\int_{-1}^{1} [\mathbf{N}]^T [\mathbf{R}] \{\mathbf{T}\} J d\xi$$

where $r^{\mathrm{int,ce_k,top}}$ and $r^{\mathrm{int,ce_k,bot}}$ are the internal force vectors of the top and bottom surfaces in the cohesive zone respectively.

Writing these equations in a first-order incremental scheme the [12x12] tangent stiffness takes the form:

$$Kcoh = \int_{-1}^{1} \begin{bmatrix} M & -M \\ -M & M \end{bmatrix} J d\xi \simeq \sum_{i=1}^{n} w_i \begin{bmatrix} M(\xi_i) & -M(\xi_i) \\ -M(\xi_i) & M(\xi_i) \end{bmatrix} J$$

where,

$$M = N_{beam}^{T} R_m D_{\tan} R_m^{T} N_{beam}$$

and

$D_{\tan}$ is the tangent elasticity matrix

while the [12x1] current internal force vector is:

$$f_{coh} = \begin{bmatrix} f'_{\text{int}} \\ f''_{\text{int}} \end{bmatrix}$$

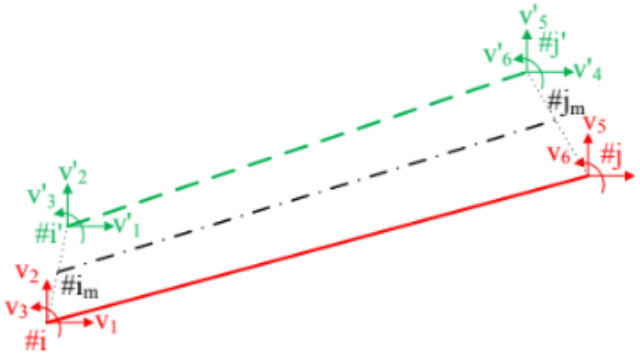
where,

$$f'_{\text{int}} = \int_{-1}^{+1} N_{beam}^{T}(\xi) R_m \sigma J d\xi \simeq \sum_{i=1}^{n} w_i N_{beam}^{T}(\xi_i) \mathrm{R}_m \sigma J$$

$$f''_{\text{int}} = -\int_{-1}^{+1} N_{beam}^{T}(\xi) R_m \sigma J d\xi \simeq -\sum_{i=1}^{n} w_i N_{beam}^{T}(\xi_i) \mathrm{R}_m \sigma J$$

, ξ_i being the integration points

$$\mathbf{N}_{beam} = \begin{bmatrix} N_1^b(\xi) & 0 & 0 & N_2^b(\xi) & 0 & 0 \\ 0 & N_1^b(\xi) & 0 & 0 & N_2^b(\xi) & 0 \end{bmatrix}$$



1.8 Mid-surface according to upper/lower surface of the cohesive zone

In order to correlate the dofs of the cohesive element with the dofs of the surrounding matrix the embedding technique is implemented. This yields the final form of the cohesive element tangent stiffness matrix:

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

, where T_{constr} is the embedding matrix and it has the form:

	u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8
v_1	$N_1(x_i)$	0	$N_2(x_i)$	0	$N_3(x_i)$	0	$N_4(x_i)$	0
v_2	0	$N_1(x_i)$	0	$N_2(x_i)$	0	$N_3(x_i)$	0	$N_4(x_i)$
v_3	$-N_{1,y}(x_i)$	$N_{1,x}(x_i)$	$-N_{2,y}(x_i)$	$N_{2,x}(x_i)$	$-N_{3,y}(x_i)$	$N_{3,x}(x_i)$	$-N_{4,y}(x_i)$	$N_{4,x}(x_i)$
v_4	$N_1(x_j)$	0	$N_2(x_j)$	0	$N_3(x_j)$	0	$N_4(x_j)$	0
v_5	0	$N_1(x_j)$	0	$N_2(x_j)$	0	$N_3(x_j)$	0	$N_4(x_j)$
v_6	$-N_{1,y}(x_j)$	$N_{1,x}(x_j)$	$-N_{2,y}(x_j)$	$N_{2,x}(x_j)$	$-N_{3,y}(x_j)$	$N_{3,x}(x_j)$	$-N_{4,y}(x_j)$	$N_{4,x}(x_j)$

with v_i the cohesive beam dofs, u_i the matrix dofs and N_i the FE shape functions

After having fully characterized the cohesive zone element the complete tangent stiffness matrix and internal forces of the system can be written:

$$\mathbf{K}_T = \begin{bmatrix} \mathbf{K}_{solid} + \mathbf{K}^{*11}_{coh} & \mathbf{K}^{*12}_{coh} \\ \mathbf{K}^{*21}_{coh} & \mathbf{K}^{*22}_{coh} + \mathbf{K}_{beam} \end{bmatrix}$$

and

$$\mathbf{f}_I = \mathbf{f}_{solid} + \mathbf{T}^{T}_{constr}\mathbf{f'}_{int}$$

$$\mathbf{f}_{II} = \mathbf{f}_{beam} + \mathbf{f''}_{int}$$

$$\mathbf{F} = [\mathbf{f}_I, \mathbf{f}_{II}]^T$$

Bayesian updating

With the rapid growth of information and data collection technologies a lot of knowledge can be gathered around engineering systems. The need to take into consideration this information and based on it modify the current uncertainty of the system parameters led to the development of mathematical and computational models that update the beliefs as new data emerge.

A robust and effective technique for combining new observations with existing models is based on the Bayes theorem and it is called Bayesian analysis. With this, prior probabilistic information about the uncertain parameters are updated according to the new available data based on real and in most cases also uncertain measurements of the mechanical system's response. The uncertainty of the measurements plays a major role in the resulted uncertainty of the updated probabilistic model. With the shift in the prior beliefs critical decisions can be made on the engineering systems concerning the reliability and the risk assessments.

Random variables

Random variable is a measurable function which assigns probabilities in a sample space Ω which is the set of all possible outcomes of an experiment. Random variables can take the form of discrete functions, where they can be described by a probability mass function which assigns probabilities for every discrete outcome or they can take the form of continuous functions named as probability density functions (PDF), where in this case the probability is assigned to continuous intervals.

A PDF can be fully characterized by its moments. A moment is a measure of a specific aspect of a function. By taking into consideration all the relevant moments the shape of a function can be described. The most popular moments are the mean value and the standard deviation.

$$\mu_X = E[X] = \int_{\Omega} x f_X(x) dx$$

$$\sigma_X = \sqrt{E[(X-\mu_X)^2]} = \sqrt{\int_{\Omega} (x-\mu_x)^2 f_X(x) dx}$$

The most common PDF which is widely used in many statistical applications is the Gaussian (normal) distribution. The two moments mentioned above are enough to fully describe this function of the form:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2} \right] \quad -\infty < x < +\infty$$

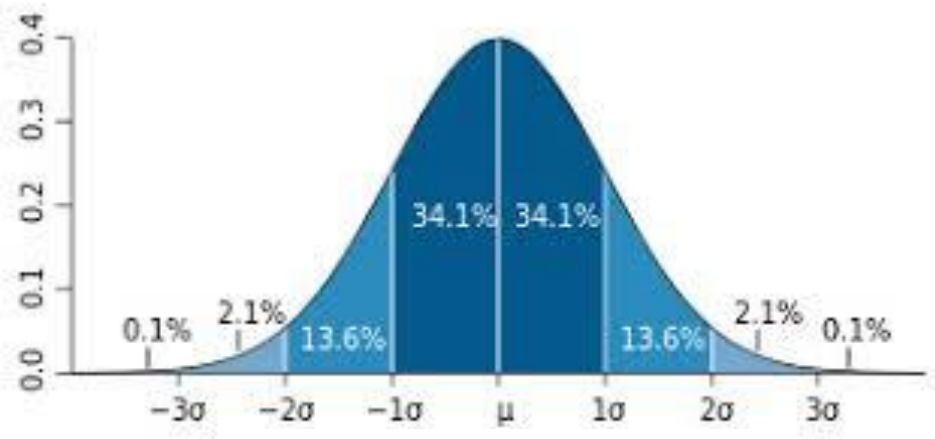
The main properties of a Gaussian distribution can be described as followed:
1. The linear functions of Gaussian random variables remain Gaussian distributed. If $X \sim N(\mu_X, \sigma^2_X)$ then $Y = aX + b \sim N(\mu_Y = a\mu_X + b, \sigma^2_Y = a^2\sigma^2_X)$

2. If X_1, X_2 are independent Gaussian random variables with mean values μ_1, μ_2 and standard deviations σ_1, σ_2, respectively, then the random variable $X = X_1 + X_2$ is also Gaussian with mean value $\mu = \mu_1 + \mu_2$ and standard deviation $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2}$

3. The nth moment of the standard Gaussian random variable Z can be computed from:

$$E[Z^n] = \int_{-\infty}^{+\infty} z^n f_Z(z) dz = \int_{-\infty}^{+\infty} z^n \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{z^2}{2} \right] dz$$

4. A Gaussian random variable X with mean μ_X and standard deviation σ_X can be obtained from the standard Gaussian random variable: $X = \sigma_X Z + \mu_X$. Then the central moments of X are given from: $E[(X - \mu_X)^n] = E[(\sigma_X Z)^n] = \sigma^n_X E[Z^n]$.



2.1 Probability intervals of normal distribution

When there is more than one random variable in a system, a PDF where also the correlation between those must be employed. This is the joint PDF and has the form:

$$F_{XY}(x,y) = \int_{-\infty}^{y}\int_{-\infty}^{x} f_{XY}(\xi_X,\xi_Y)\mathrm{d}\xi_X\mathrm{d}\xi_Y$$

The moment that describes the relation of the random variables in a joint PDF is the covariance and is written as:

$$\mu_{11} = E[(X-\mu_X)(Y-\mu_Y))] = \int_{-\infty}^{y}\int_{-\infty}^{x}(\xi_X-\mu_X)(\xi_Y-\mu_Y)\,f_{XY}(\xi_X,\xi_Y)d\xi_X d\xi_y =$$

$$= E[XY]-\mu_X\mu_Y$$

While the normalized covariance is:

$$\rho = \frac{\mu_{11}}{\sigma_X\sigma_Y}$$

Which is equal to zero in the case of uncorrelated random variables and unity in the case of them being fully correlated.

Bayes' Theorem

For two events A and B, according to the Bayes' Rule, the possibility of A happening while knowing that B is true is given by the following relation:

$$P(A|B) = \frac{P(B|A)\cdot P(A)}{P(B)}$$

In the context of random variables this relation can be expressed as:

$$f(\vec{X}|\,Data) = \frac{L(\vec{X}|\,Data)\cdot f(\vec{X})}{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\dots\int_{-\infty}^{+\infty} L(\vec{X}|\,Data)\cdot f(\vec{X})\,dX_1 dX_2\dots dX_n}$$

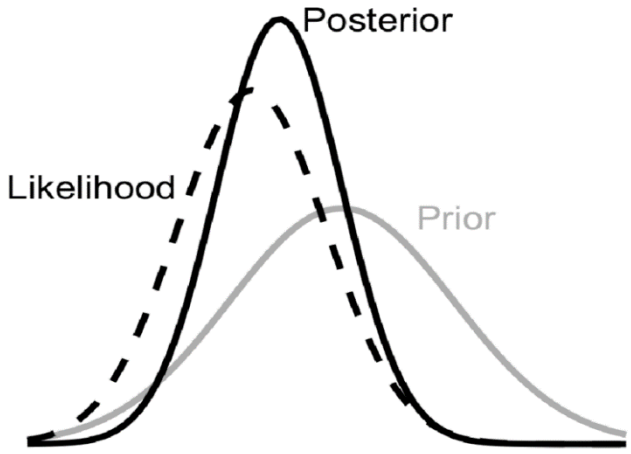
, where $\vec{X}$ are the system's random variables.

The joint PDF $f(\vec{X})$ is called prior distribution and represent the beliefs of the probabilistic model before new information are known. The likelihood function $L(\vec{X}\mid Data)$ represent the beliefs of the probabilistic model based solely on the emerged data. By multiplying these two functions one can acquire a proportional function to the posterior distribution $f(\vec{X}\mid Data)$ which is the transformed probabilistic model after taking into consideration the new data. In order to calculate the exact posterior distribution, one has to calculate the integral (normalization factor) which lies in the denominator of the relation. This is a very hard task to solve analytically as more random variables exist in the system and the need to resort to other approaches arise.

The likelihood can represent the deviation of the observations from the statistical model predictions in a form:

$$L_i(\mathbf{x}) = f_{\epsilon_i}[y_i - h_i(\mathbf{x})].$$

, where y_i are the observations, $h_i(x)$ are the model predictions and ε_i is the deviation($y_i - h_i(x) = \varepsilon_i$) which may be influenced from measurement or model errors.



2.2 Component distributions of a Bayesian model

Markov Chain Monte Carlo (MCMC)

The introduction of the MCMC methods were motivated by the impossible in most cases task of finding the exact posterior distribution. Instead they can work with a proportional function of the posterior: $f(\vec{X} \mid Data) \propto L(\vec{X} \mid Data)\, f(\vec{X})$

Stochastic processes with the Markovian property are those in which every future state can be predicted by the knowledge of the present state while it does not get affected by previous states and that is why it is also called memoryless process. A subcategory of these processes is the Markov Chain which is described in the space of discrete random variables $\{X_n : n \in N\}$ and discrete state space S:

$$P(X_{n+1} = j \mid X_0, X_1, \ldots, X_n) = P(X_{n+1} = j \mid X_n)$$

If the Markov Chain is also independent of the time that the transition n→n+1 took place then the stochastic process is also called homogeneous.

The basic properties of a Markov Chain are:

1)Reducibility: A state j is said to be communicate with another state i if there is a positive probability for the chain to reach the state j from the state I and vice versa. If this condition holds for all of the pairs of chains then the Markov chain is called irreducible.

2)Periodicity: A state i has a period d if the chain is returned to state i in multiples of d time steps and it is expressed as the greatest common divisor $d_i = \gcd\{n : P_{ii}^n > 0\}$. In the case where d=1 an irreducible Markov Chain is said to be aperiodic.

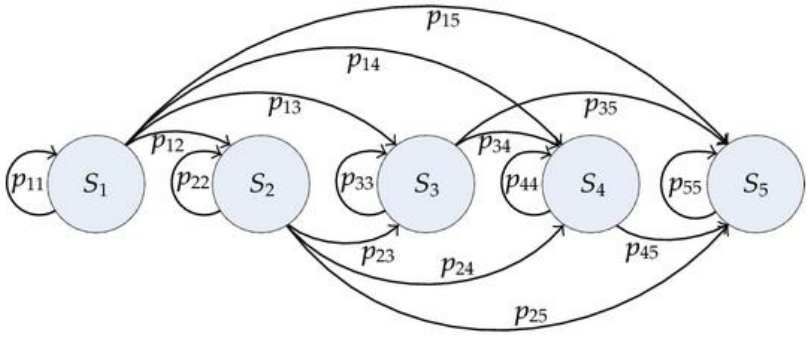
3)Recurrence: A state i is called recurrent if the probability of returning to i in finite or infinite chain steps is non-zero and it is expressed as:

$$\sum_{n=1}^{\infty} f_{ii}^{(n)} = \sum_{n=1}^{\infty} \Pr(T_i = n \mid X_0 = i) = 1$$

In the case that the above relation equals to less than unity the state is called transient. When the return happens in finite steps the state is called positive recurrent otherwise it is called null recurrent.

Transition matrix $P(m) = p_{ij}^{(m)}$ is called the probability of moving from state i to state j in m steps. Stationary distribution π of a Markov Chain is a distribution, which independently of the n implementations of the transition matrix on it, remains the same:

$\pi = \pi \mathrm{P}^n$, where $\sum_i \pi_j = 1$



$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} & p_{15} \\ 0 & p_{22} & p_{23} & p_{24} & p_{25} \\ 0 & 0 & p_{33} & p_{34} & p_{35} \\ 0 & 0 & 0 & p_{44} & p_{45} \\ 0 & 0 & 0 & 0 & p_{55} \end{bmatrix} \text{ where } p_{ij} \geq 0 \text{ and } \sum_j p_{ij} = 1$$

2.3 A Marcov Chain with its transition matrix

There can be different chains having the same stationary distribution.

An irreducible chain has a stationary distribution only if it is positive recurrent.

The MCMC methods, by adopting these properties that a Marcov Chain has, can draw efficiently samples from the posterior distribution. They work by finding after some number of steps the transition that has as a stationary distribution the posterior. They perform better than simple Monte Carlo methods such as Acceptance Rejection Sampling (ARS) in most cases, especially in higher dimension spaces. The Markov Chain procedure though adds some drawbacks in the whole process. There is correlation between subsequent samples which needs to be taken into consideration in order to draw efficiently from the posterior. That can be achieved by lengthening the algorithm steps but another issue rises with this step shift, which is the chain trying to exit the highest probability region. Also, some samples must be discarded at the start of the procedure, the number depending on how close is the proposal transition to the true transition of the stationary-posterior. Consequently, the optimal adjustments must be made so to overcome these issues.

Metropolis-Hastings algorithm (MH)

The most popular algorithm for the implementation of a MCMC procedure is a random walk algorithm called Metropolis-Hastings which later has led to the creation of alternative or modified algorithms but with the same general concept.

Two conditions must be met in order for MH algorithm to work:

1. Existence of a stationary distribution $\pi(x)$, which requires that every transition x→ x' is reversible, meaning that the probability of being in state x and transitioning to state x' must me equal to the probability of being in state x' and transitioning to state x.

$$\pi(x)P(x'|x) = \pi(x')P(x|x')$$

2. Uniqueness of the stationary distribution $\pi(x)$, which required that every state x must be aperiodic and positive recurrent

The approach is so to decompose the transition probability $P(x'|x)$ in two sub-probabilities. The proposal distribution $g(x'|x)$ is the probability of moving from a state x to x' and the acceptance ratio $A(x',x)$ is the probability to accept the proposed state x'. Rewriting the above equation after this change yields:

$$\frac{A(x',x)}{A(x,x')} = \frac{P(x')}{P(x)} \frac{g(x|x')}{g(x'|x)}$$

, where

$$P(x'|x) = g(x'|x)A(x',x)$$

The next step is to choose an appropriate acceptance ratio. The Metropolis choice is either $A(x',x)$=1 or $A(x,x')$=1. This can be written as:

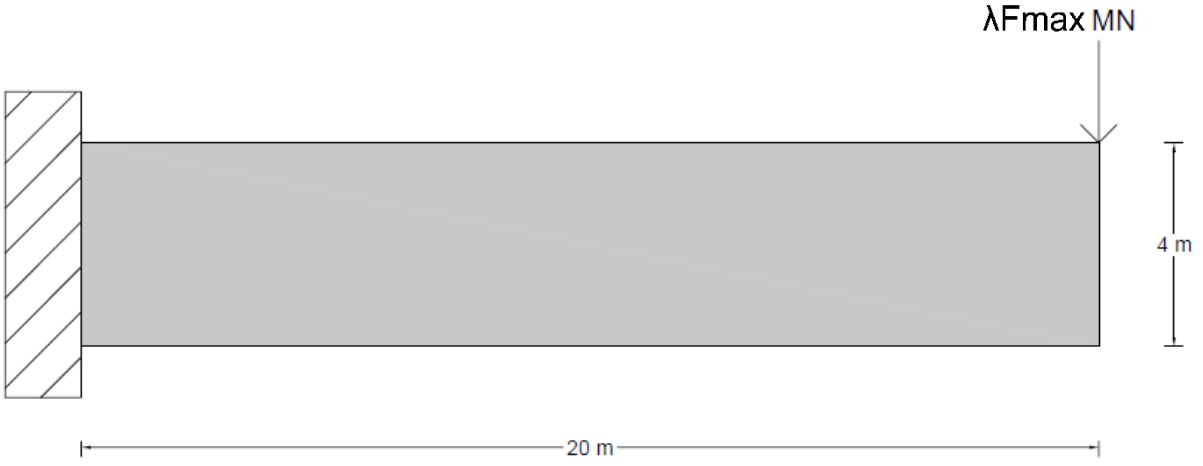
$$A(x',x) = \min\left(1, \frac{P(x')}{P(x)} \frac{g(x|x')}{g(x'|x)}\right)$$

The MH algorithm can be described as followed:

1. Choose an initial starting point x_0 and a proposal density $g(x'|x)$. A common proposal density is the normal distribution and the x_0 can be the mean value of the prior distribution.
2. Get a random proposed sample x' from $g(x'|x_t)$, where t is the number of the iteration.
3. Calculate the acceptance ration $A(x',x)$ from the relation above
4. Generate a random number u with uniform distribution in the interval [0,1]
5. Accept the proposed sample if $u \leq x'$ / Reject the proposed sample if $u > x'$
6. Go to step 2 for the next sample / End if the required samples have been gathered

Numerical Example

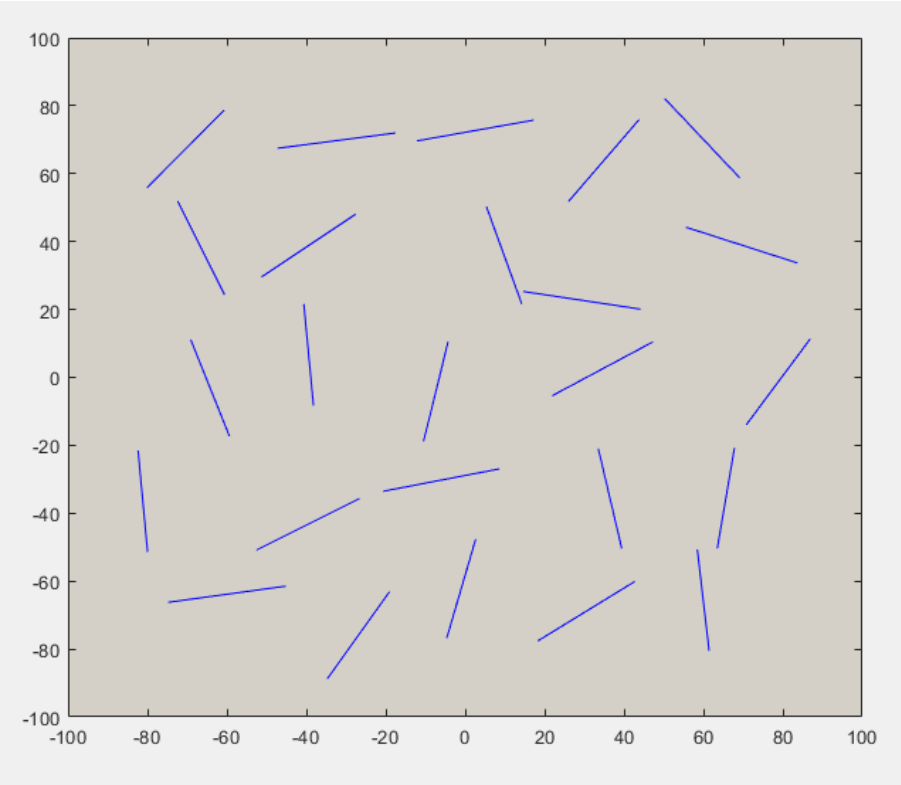
In this section a numerical example which demonstrates the procedure of Bayesian update with MCMC in a nonlinear multiscale model is implemented. The coding for this example was developed entirely on MATLAB. Specifically, a pushover analysis was carried out on a beam constructed with a composite material. The composite material can be described as a carbon nanotube-reinforced polymer matrix.



3.1 Beam element in the macroscopic scale

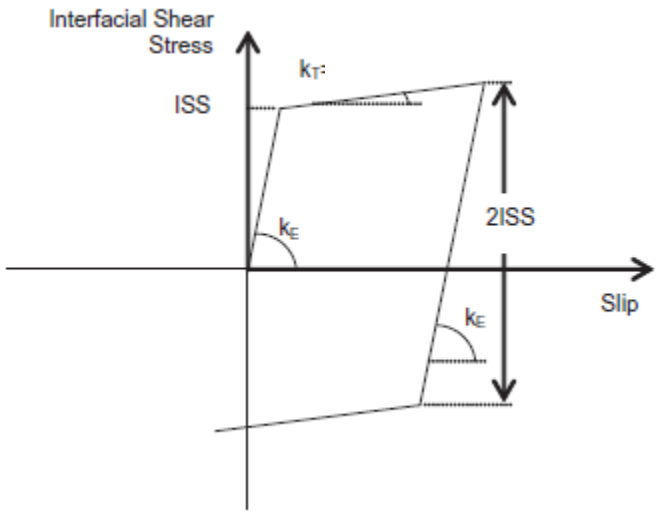
The interaction between the CNTs and the surrounding polymer matrix was examined in the microstructure. The modulus of elasticity of the polymer is $E_p = 2.79GPa$ and Poisson's ratio $\nu = 0.4$. Implementing the EBE procedure the CNTs were formed as beam elements with a solid circular profile section with mean equivalent diameter $D_{eq} = 4.57nm$ and modulus of elasticity $E_C = 1051GPa$. The dimensions of the RVE are 200x200 nm^2 in which 25 CNTs with length

$L = 30nm$ with random positioning and orientation were placed. The number and the length of the CNTs where chosen for a weight fraction of $wf = 1\%$.



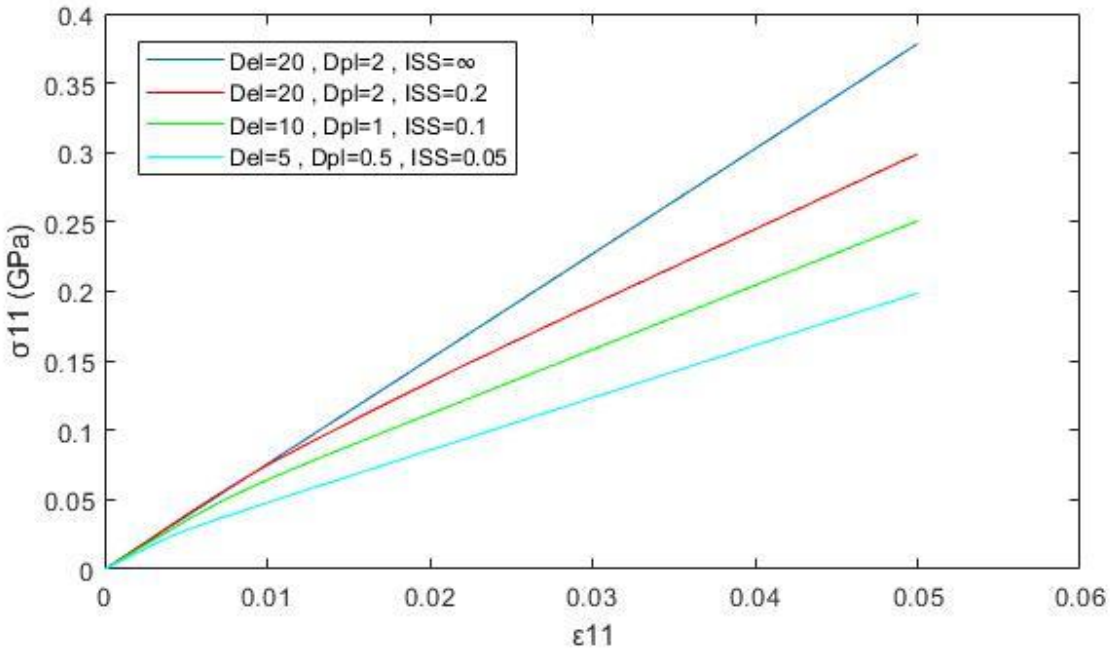
3.2 Representation of the RVE

The interface between the polymer and the CNTs was simulated with a bond-slip kinematic hardening and the connection between the dofs was done with a cohesive zone model and the embedding technique. The parameters than define the interface are the Interfacial Shear Stress (ISS) which is the stress at which the slip happens, the relation of traction-separation in the full bonded case (D_{el}) and the same relation after the maximum separation before the slip has been reached (D_{pl}).

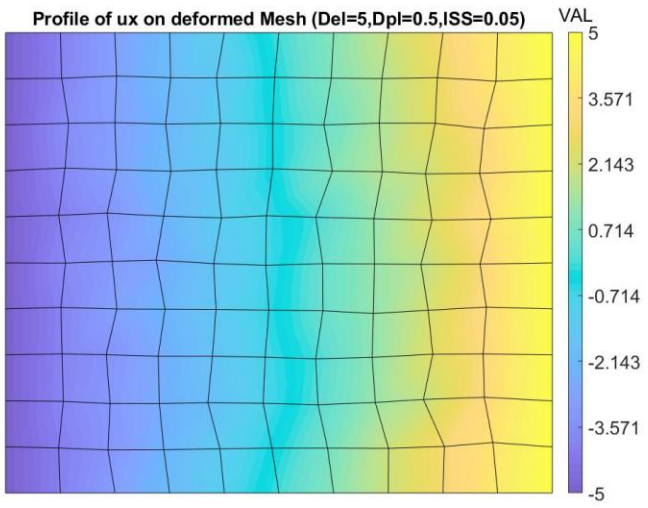


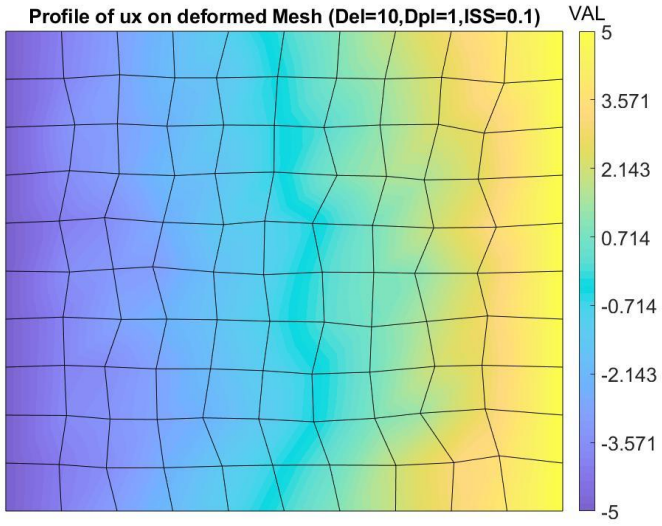
3.3 Bilinear slippage constitutive law

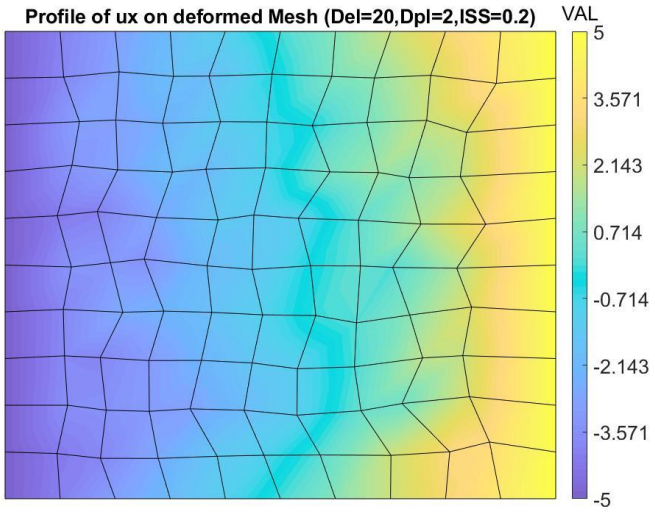
Firstly, a microscopic testing was carried out in order to observe how a single RVE is responding to different values of the interface parameters. A homogenization scheme with the linear boundary displacement method was applied to the discretized RVE and was solved with a FE model. The RVE was subjected to increasing values of macroscopic strains of the form $\bar{\varepsilon} = [\varepsilon_{11} \quad \varepsilon_{22} \quad 2\varepsilon_{12}]^T = [\varepsilon \quad -0.4\varepsilon \quad 0]^T$ until a maximum strain of $\varepsilon_{\max} = 5\%$. Below shows the relation $\sigma_{11} - \varepsilon_{11}$ for different values of interface parameters (D_{el} and D_{pl} in GPa/nm, ISS in GPa).



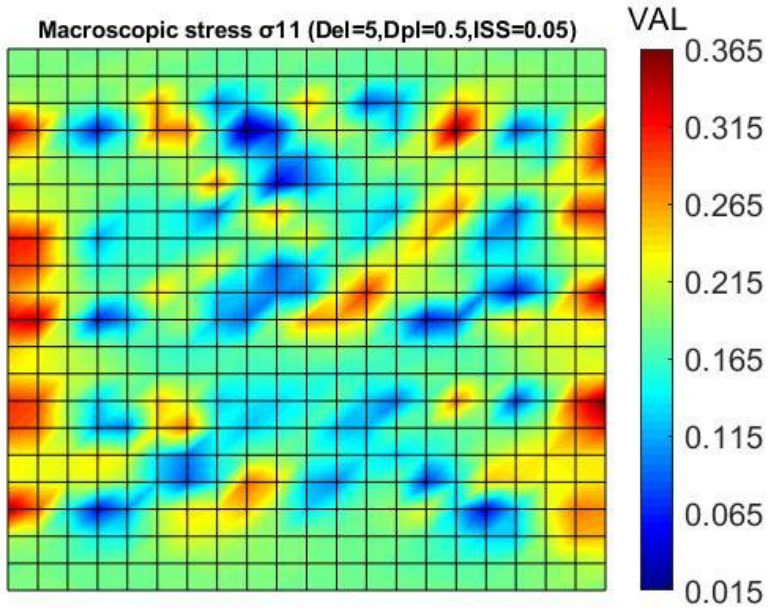
3.4 Macroscopic strain – Homogenized stress relation

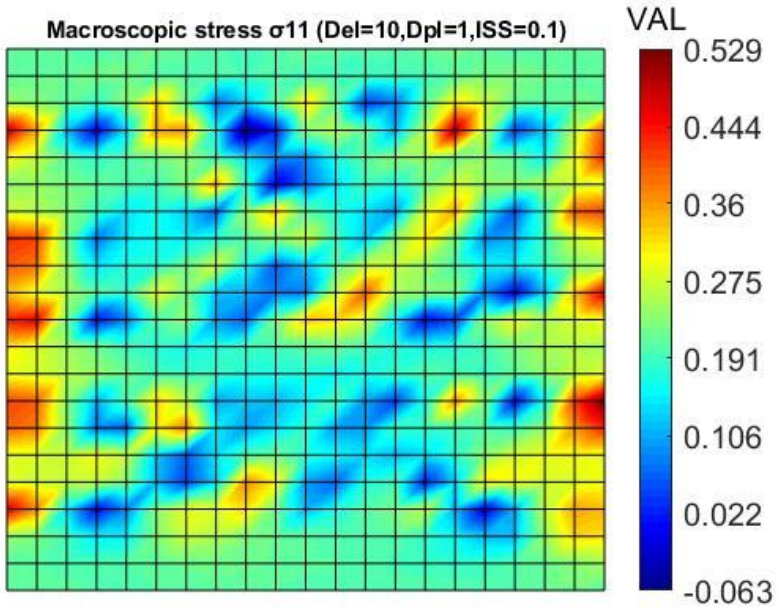


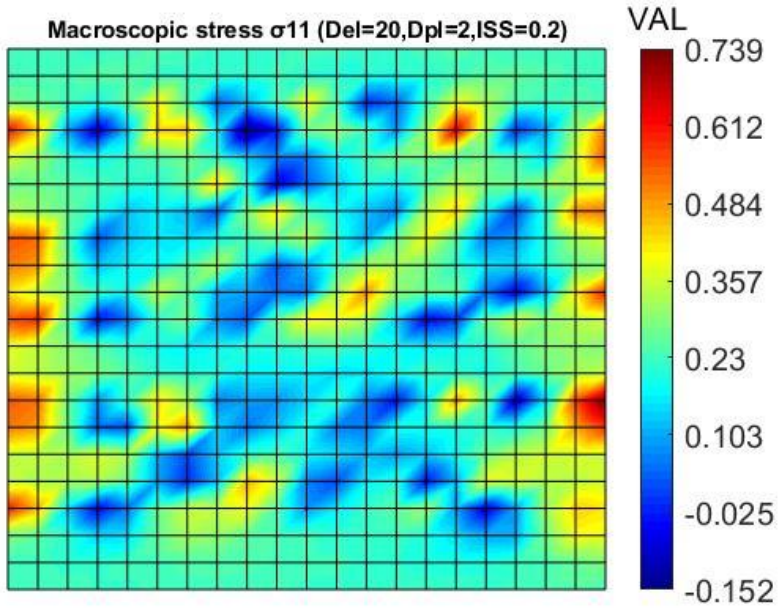




3.5 Deformation ux on RVE from linear displacement boundaries

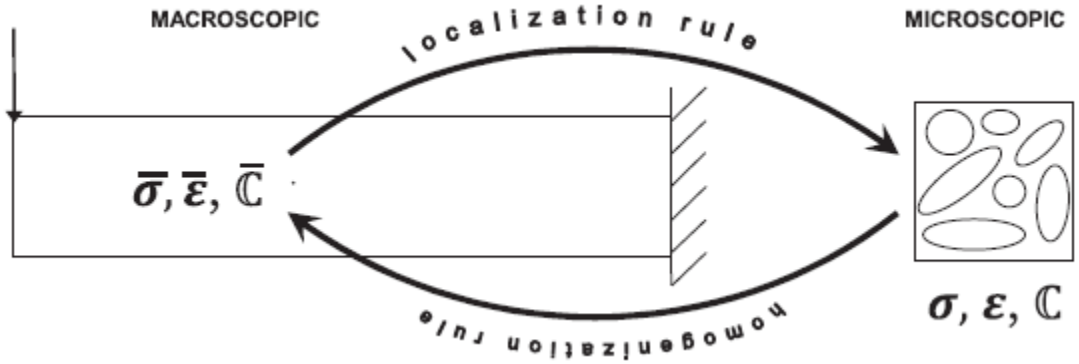






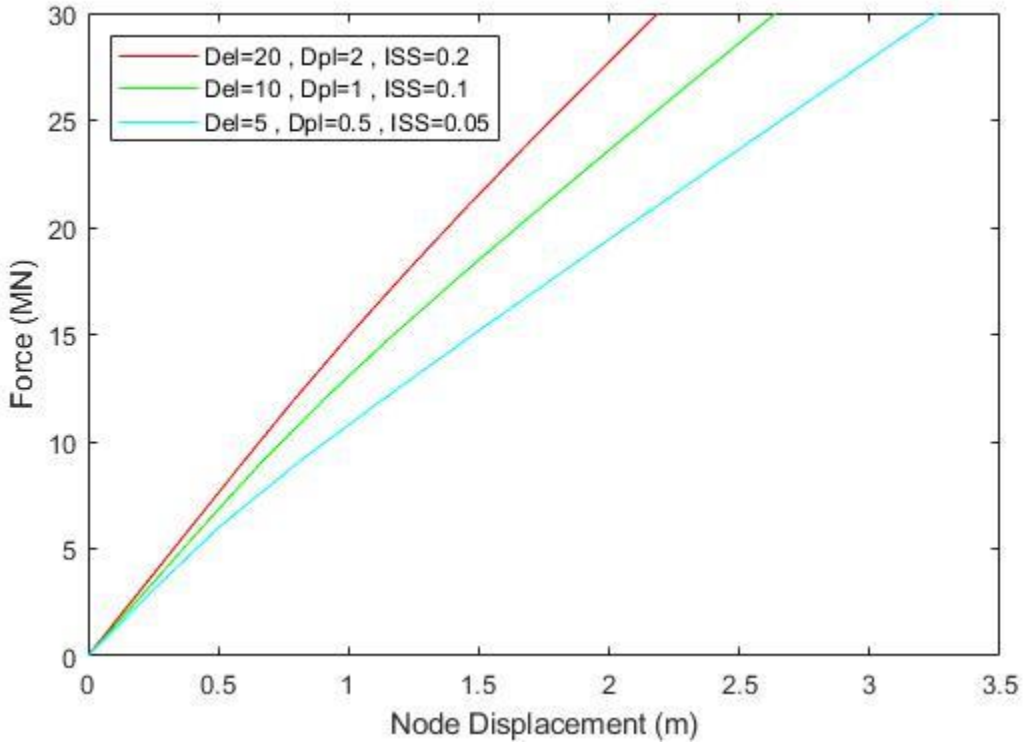
3.6 Stress σ11 on RVE from linear displacement boundaries

A macroscopic testing was performed with the FE2 procedure. The beam with 20m length and 4m height was fixed in the left end while in the right end was subjected to a vertical load F_{incr} of increasing magnitude from the unloaded case until it reaches F_{max}=30MN which was achieved with the introduction of a scale factor λ ($F_{incr}=\lambda F_{max}$). The macrostructure was discretized and then was connected with the microstructure with the localization and homogenization rule. The detailed algorithm procedure was described in paragraph 2.2

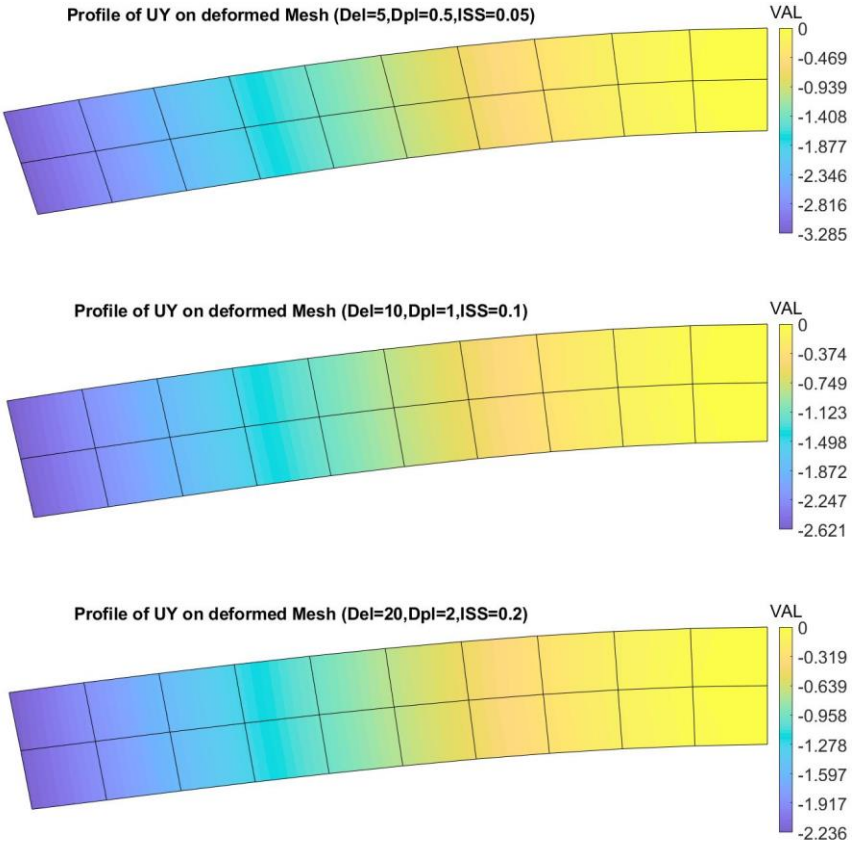


3.7 Semi-concurrent multiscale approach

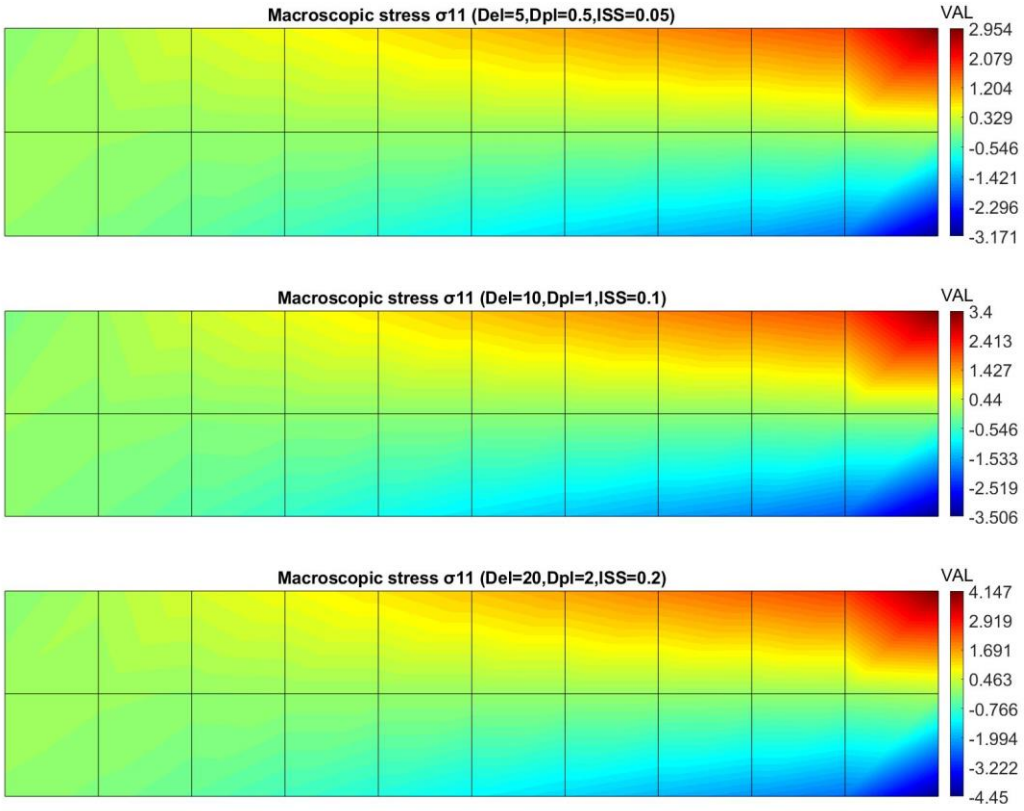
The response of the macrostructure was tested for the same values of the CNT-polymer interaction parameters with those in the microscopic testing. In the pushover analysis the macroscopic displacement of the node A in the bottom right corner (maximum displacement point) was observed and it is shown below.



3.8 Macroscopic displacement uy on node A



3.9 Macroscopic displacement uy on beam



3.10 Macroscopic stress σ11 on beam

Having observed the responses of the micro and the macro structure for several deterministic values of the interface parameters a probabilistic model is implemented. In this model the three parameters are considered to be random variables.

It is assumed that a real experiment has been implemented on the composite beam by imposing the same loading condition as described before. The measurement that has been observed for F=30MN is u_A=3.1m, where A is the node defined before. The prior distributions P_{pr} are considered to be normal for the three parameters. The mean value and standard deviation for each parameter is presented below:

$$D_{el} \rightarrow N \sim (10, 2)$$
$$D_{pl} \rightarrow N \sim (1, 0.2)$$
$$ISS \rightarrow N \sim (0.1, 0.02)$$

The likelihood function has the form:

$$L(\vec{X}) = f_{\varepsilon}[u_A - u(\vec{X})]$$

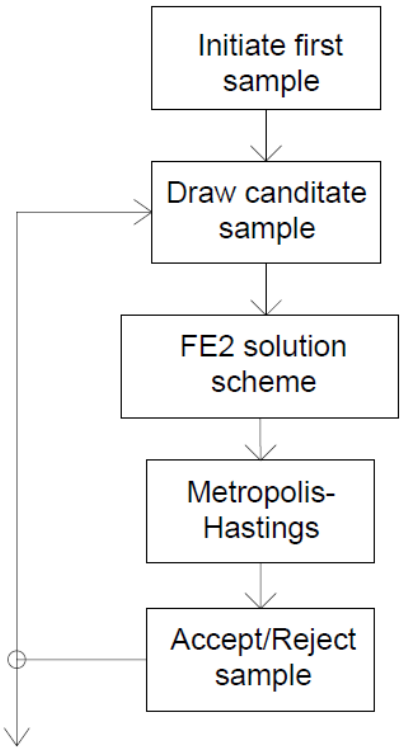
Where $\vec{X}$ is the vector of the random variables and $u(\vec{X})$ is the displacement which resulted from the multiscale model solution, while ε is the uncertainty of the measurement.

The uncertainty of this measurement is expressed with a standard deviation $\sigma_m = 0.06$, so the likelihood function is assumed to be a normal variable defined as:

$$u_A \to N \sim (3.1, 0.06)$$

The MCMC procedure is followed in order to sample from the posterior distribution:

It has to be pointed out that in order to accept a candidate sample from the posterior the whole FE^2 algorithm must be solved in order to find the displacement at A, making the application a heavy computational task. The Metropolis Hastings algorithm was used as the MCMC method. The process can be shown schematically with the flow chart below.



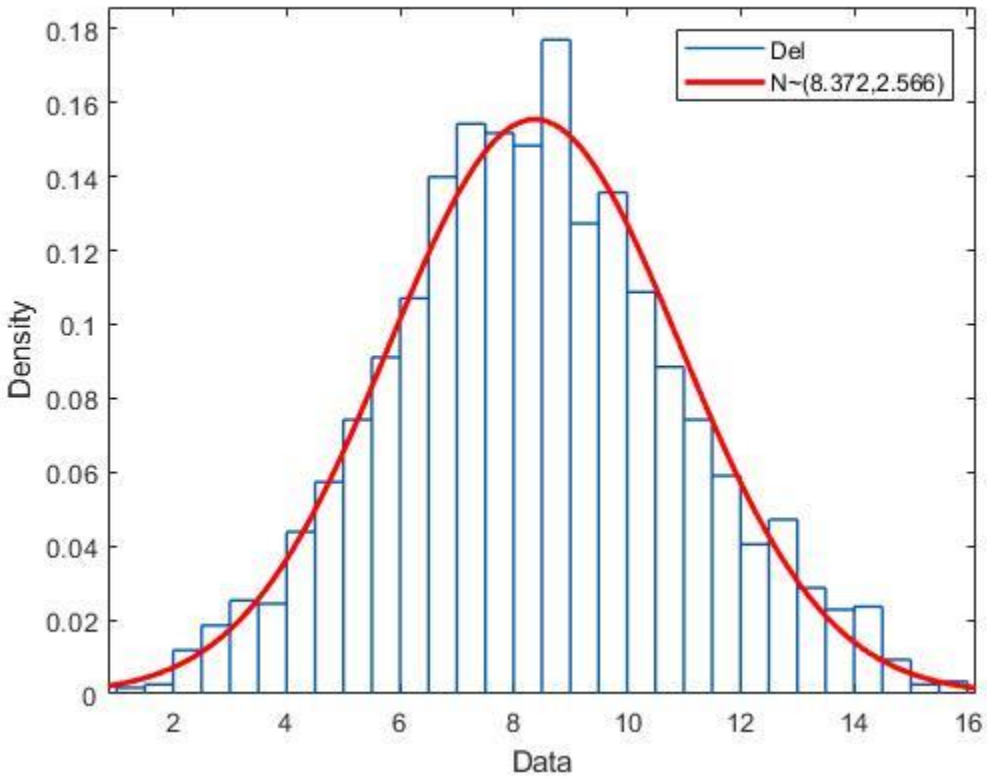
3.11 Flow chart of complete procedure

For the MH algorithm only a burn-in period of 20 samples was chosen in order the Markov Chain to converge to the target distribution. Additionally, none of the accepted samples is thrown away for the sake of correlation issues. That is because from a Bayesian update perspective the problem is not considered complex and the target distribution is expected to have similar shape and values as the prior.

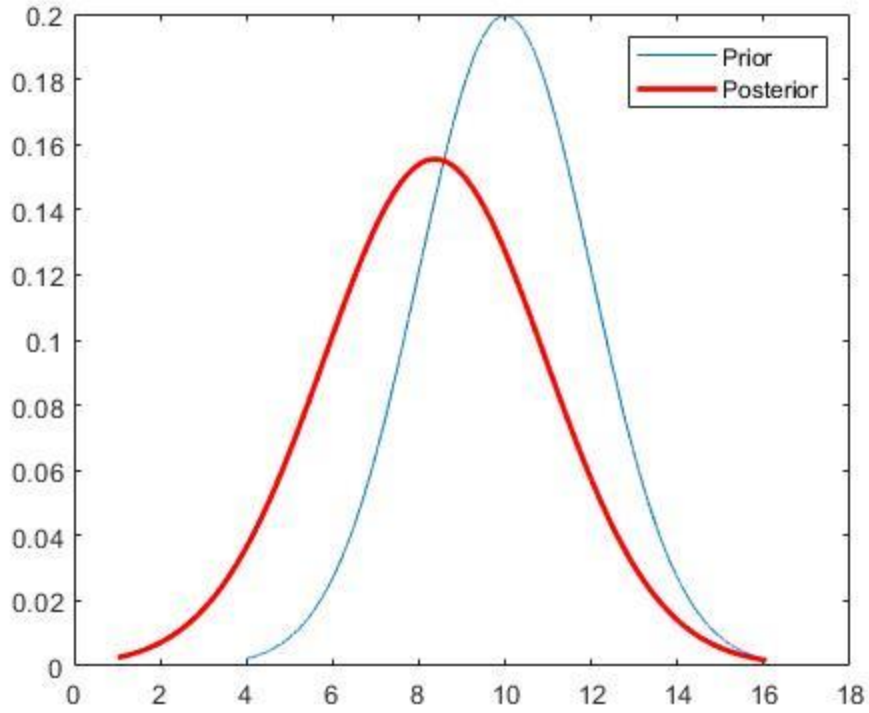
After the whole process has finished and the necessary samples have been gathered it is possible to create the histogram of the three random variables and fit an appropriate distribution to them which will be the posterior/target distribution.

If we consider the target distribution to be normal the transformed statistical parameters after the Bayesian update for each random variable are:

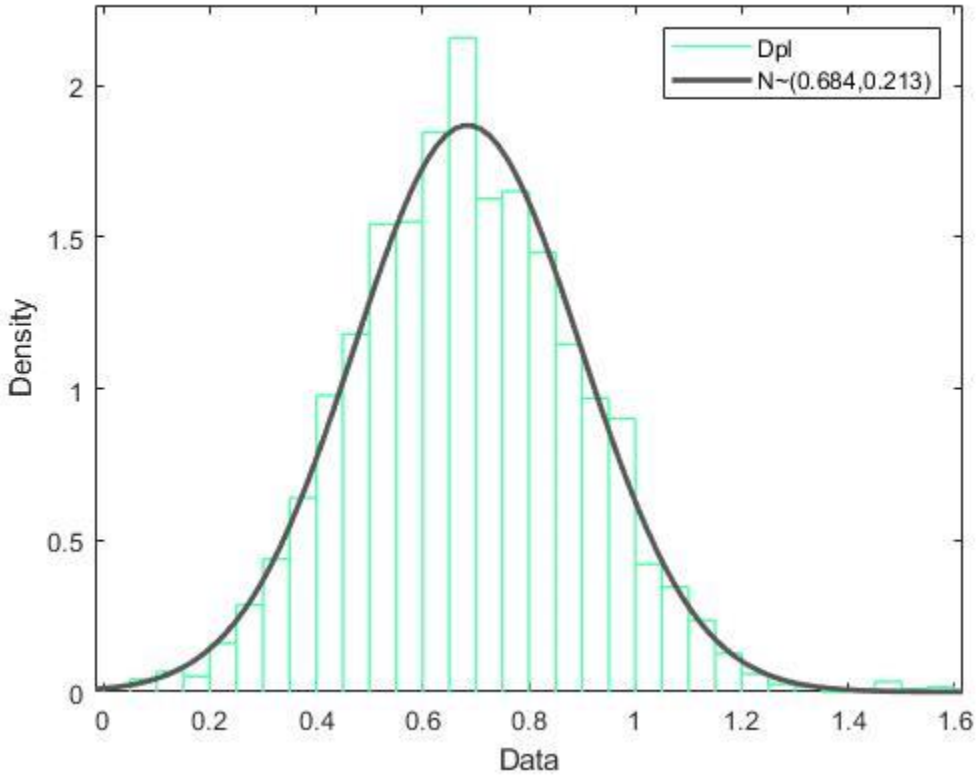
$$D_{el} \rightarrow N \sim (8.372, 2.566)$$
$$D_{pl} \rightarrow N \sim (0.684, 0.213)$$
$$ISS \rightarrow N \sim (0.745, 0.023)$$



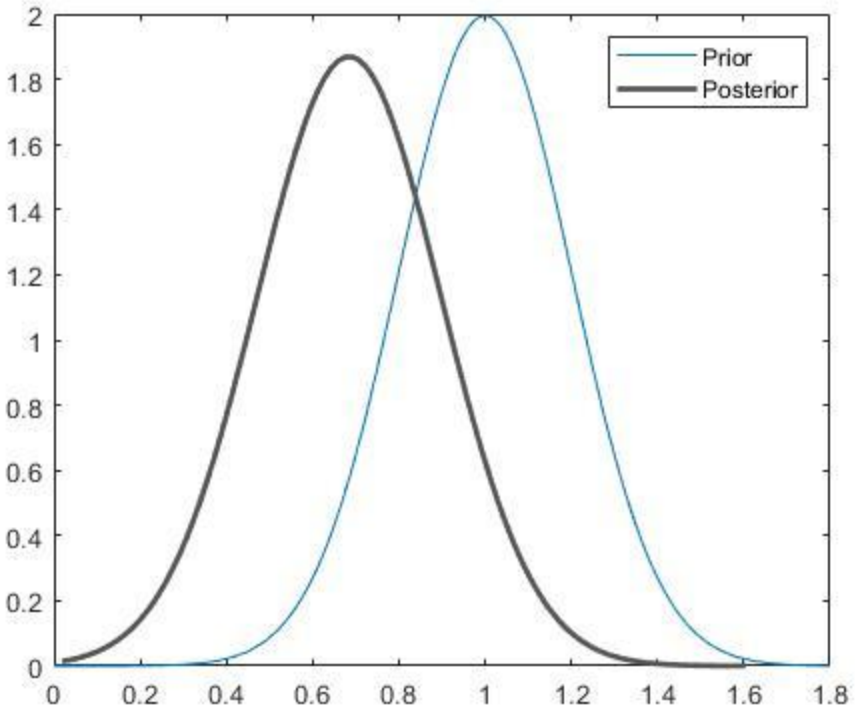
3.12 Posterior Histogram for Del



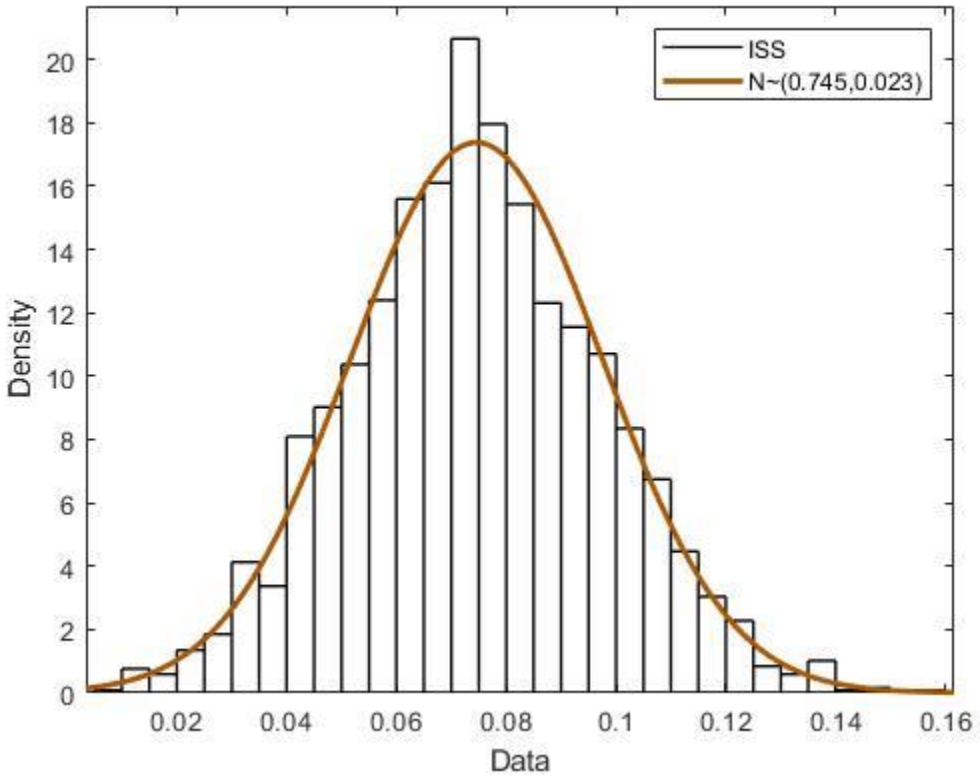
3.13 Comparison Prior-Posterior distribution for Del



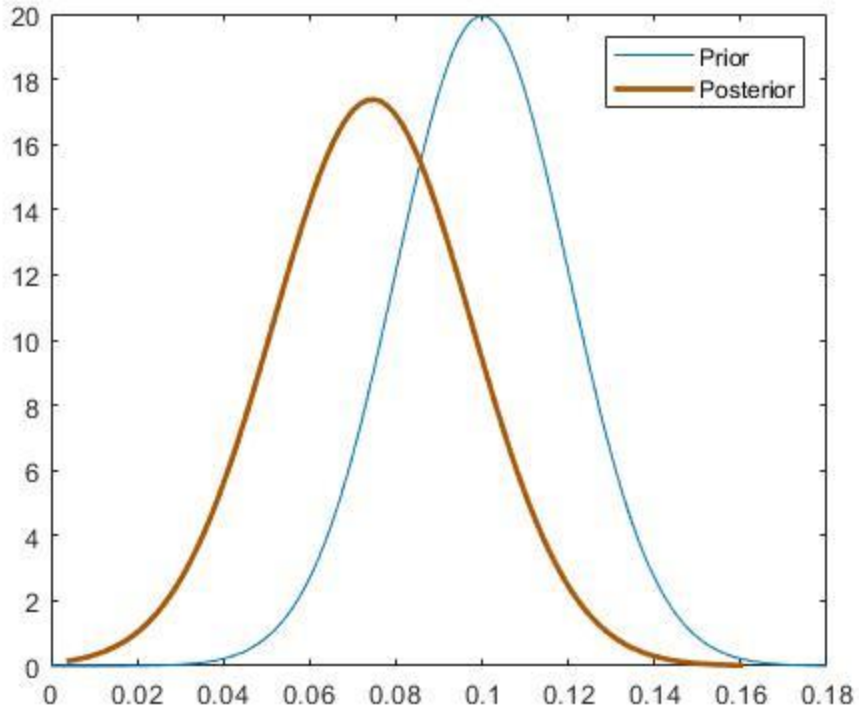
3.14 Posterior Histogram for Dpl



3.15 Comparison Prior-Posterior distribution for Dpl

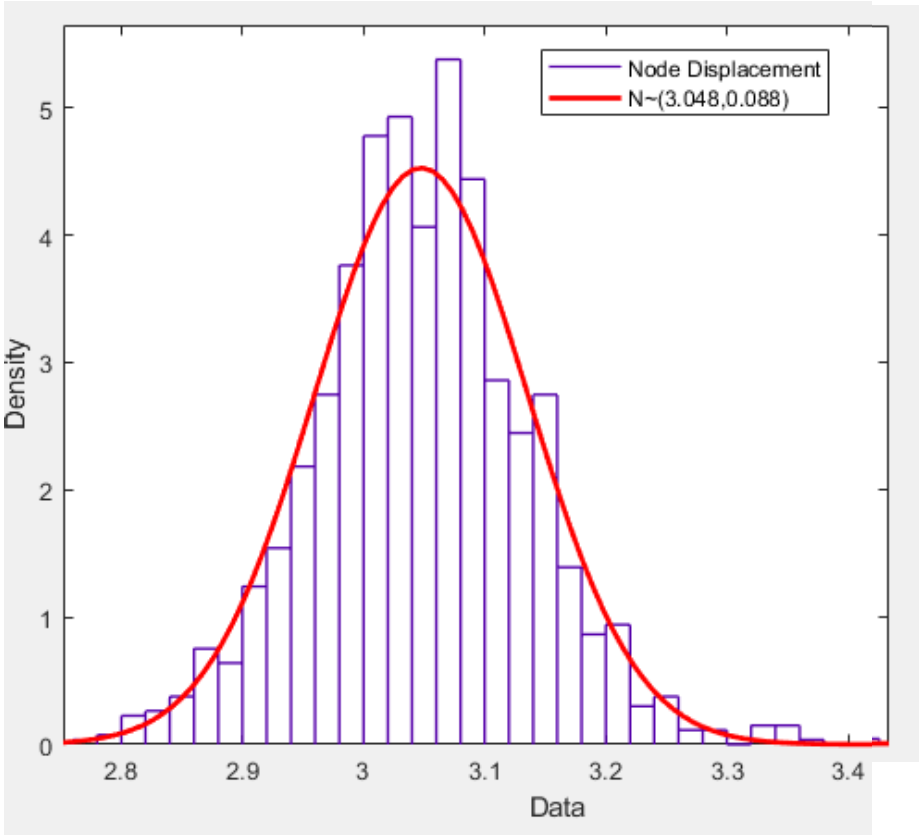


3.16 Posterior Histogram for ISS



3.17 Comparison Prior-Posterior distribution for ISS

The distribution of the node displacement can also be derived.



3.18 Posterior Histogram for u_A

It can be seen that all of the distributions have lower mean values, while the standard deviations are approximately the same. This is to be expected considering the results are fully compatible with the deterministic solutions of the multiscale model conducted before.

Conclusions

A numerical example has been developed where the importance of Bayesian updating in a complex engineering system such as the response of a highly heterogeneous material on a specific loading condition has been highlighted. By taking into consideration new emerged data on the macroscopic structure such as a deformation observation, the update of the probabilistic structure of the parameters that lie in the microscopic structure has been achieved. This is of high importance since these parameters are hard and most of the times costly to be examined directly.

Specifically, the MCMC procedure with the Metropolis Hastings algorithm has been used on a semi-concurrent multiscale method (FE^2) in order to draw samples from the target distribution of the microscopic parameters that describe the CNT-polymer interaction, after a real measurement of the macroscopic deformation on the edge of a composite beam is known. Comparing these updated samples with deterministic solutions of the model for fixed microscopic parameters reveals the efficient implementation of the Metropolis Hastings algorithm making it a very promising methodology for the investigation of updated microscopic uncertainties in a multiscale model.

References

[1] Miehe, C., & Koch, A. (2002). Computational micro-to-macro transitions of discretized microstructures undergoing small strains. Archive of Applied Mechanics 72 (2002) 300 – 317 Springer-Verlag 2002

[2] Feyel F. (2003). A multilevel finite element method (FE2) to describe the response of highly non-linear structures using generalized continua. Comput. Methods Appl. Mech. Engrg. 192 (2003) 3233–3244

[3] Kouznetsova, V. (2002). Computational homogenization for the multi-scale analysis of multi-phase materials

[4] D.N. Savvas, V. Papadopoulos, M. Papadrakakis (2012). The effect of interfacial shear strength on damping behavior of carbon nanotube reinforced composites. International Journal of Solids and Structures 49 (2012) 3823–3837

[5] D.N. Savvas, V. Papadopoulos (2014). Nonlinear multiscale homogenization of carbon nanotube reinforced composites with interfacial slippage

[6] Tavlaki M., Papadopoulos V. (2016). The impact of interfacial properties on the macroscopic performance of carbon nanotube composites. A FE2-based multiscale study

[7] Papadrakakis, M. (2001). Αναλυση Φορέων με τη Μέθοδο των Πεπερασμένων Στοιχείων. Αθήνα: Παπασωτηρίου.

[8] Hill, R. (1972). On constitutive macro–variables for heterogeneous solids at finite strain. Proc R Soc London A 326 131–147

[9] Straub D., Papaioannou I. (2014). Bayesian Updating with Structural Reliability Methods. Journal of Engineering Mechanics 141(3):04014134

[10] Beck, J. L. and S. K. Au (2002). Bayesian updating of structural models and reliability using Markov chain Monte Carlo simulation. Journal of Engineering Mechanics-Asce 128(4): 380-391.

[11] Beck, J. and L. Katafygiotis (1998). Updating Models and Their Uncertainties. I: Bayesian Statistical Framework. Journal of Engineering Mechanics 124(4): 455-461. Framework. Journal of Engineering Mechanics 124(4): 455-461. Carlo simulation. Journal of Engineering Mechanics-Asce 128(4): 380-391.

[12] Bolstad W. M. (2010). Understanding Computational Bayesian Statistics. Wiley

[13] Fisher, R. A. (1922). On the Mathematical Foundations of Theoretical Statistics. Philosophical Transactions of the Royal Society of London. Series A 222(594-604): 309-368.

[14] Papadopoulos V., Giovanis D.G. (2018). Stochastic Finite Element Methods. Springer

[15] Gilks, W. R., S. Richardson and D. J. Spiegelhalter (1998). Markov chain Monte Carlo in practice. Boca Raton, Fla., Chapman & Hall.

[16] Papadopoulos V., Seventekidis P., Sotiropoulos G. (2017). Stochastic multiscale modeling of graphene reinforced composites. Engineering Structures, Volume 145, 15 August 2017, Pages 176-189

[17] Biswal, B.; Øren, P.-E.; Held, R. J.; Bakke, S.; Hilfer, R. (2007). Stochastic multiscale model for carbonate rocks, Physical Review E, volume 75, issue 6

[18] Achim A., Bezerianos A., (2001). Novel Bayesian multiscale method for speckle removal in medical ultrasound images. IEEE Transactions on Medical Imaging, Volume 20, Issue 8

[19] Bouman, C.A.; Shapiro, M. (1994). A multiscale random field model for Bayesian image segmentation. IEEE Transactions on Image Processing, Volume 3, Issue 2

[20] Choi, H.; Baraniuk, R.G. (2001). Multiscale image segmentation using wavelet-domain hidden Markov models. IEEE Transactions on Image Processing, Volume 10, Issue 9

[21] Cheng, H.-C., Liu, Y.-L., Hsu, Y.-C., Chen, W.-H., 2009. Atomistic-continuum modeling for mechanical properties of single-walled carbon nanotubes. Int. J. Solids Struct. 46 (7–8), 1695–1704

[22] Barenblatt G.I. (1962). The mathematical theory of equilibrium cracks in brittle fracture. Advances in applied mathematics, volume 7