Proceeding Paper

Uncertainty Quantification at the Microscale: A Data-Driven Multi-Scale Approach †

José Pablo Quesada-Molina [1,2,*] and Stefano Mariani [1]

[1] Department of Civil and Environmental Engineering, Politecnico di Milano, Piazza L. da Vinci 32, 20133 Milano, Italy; stefano.mariani@polimi.it

[2] Department of Mechanical Engineering, Universidad de Costa Rica, Rodrigo Facio Brenes Campus, Montes de Oca, 11501-2060 San José, Costa Rica; jose.quesadamolina@ucr.ac.cr

[*] Correspondence: josepablo.quesada@polimi.it; Tel.: +39-02-2399-6284

[†] Presented at the 9th International Electronic Conference on Sensors and Applications, Online, 1–15 November 2022; Available online: https://ecsa-9.sciforum.net/.

Abstract: Data-driven formulations are currently developed and can result extremely helpful to deal with the complexity of the multi-physics governing the response of micro-electro-mechanical systems (MEMS) to the external stimuli. Such devices are in fact characterized by a hierarchy of length- and time-scales, which are difficult to fully account for in a purely model-based approach. In this work, we specifically refer to a (single-axis) Lorentz force micro-magnetometer designed for navigation purposes. Due to an alternating current flowing in a slender mechanical part (beam) and featuring an ad-hoc set frequency, the micro-system is driven into resonance so that its sensitivity to the magnetic field gets improved. A reduced-order physical model was formerly developed for the aforementioned movable part of the device; this model was then used to feed and speed up a multi-physics and multi-objective topology optimization procedure, aiming to design a robust and performing magnetometer. The stochastic effects, which are responsible for the scattering in the experimental data at the microscale, were not accounted for in such a model-based approach. A recently proposed formulation is here discussed and further extended to allow for such stochastic effects. The proposed multi-scale deep learning approach features: at the material scale, a convolutional neural network adopted to learn the scattering in the mechanical properties of polysilicon, induced by its morphology; at the device scale, two feed-forward neural networks, one adopted to upscale the mechanical properties while the other to learn a microstructure-informed mapping between the geometric imperfections induced by the microfabrication process and the effective response of the movable part of the magnetometer. The data-driven models are linked through the physical model to provide a kind of hybrid solution to the problem. Results relevant to different neural network architectures are here discussed, along with a proposal to frame the approach as a multi-fidelity, uncertainty quantification procedure.

Keywords: data-driven model; multi-physics; micro-electro-mechanical systems (MEMS); Lorentz force micro-magnetometer; multi-scale; deep learning; neural network

1. Introduction

Most materials used in sophisticated technologies, such as in the case of MEMS, are characterized by a hierarchical internal structure with features at several length scales. These hierarchical structures completely determine their macroscale properties and performance characteristics. Consequently, important efforts in the field of materials informatics are put into developing novel data-driven approaches for mining high fidelity *process—structure—property* (PSP) linkages from large collections of experimental, modeling, and simulation datasets [8]. Advancements in machine learning and data science approaches, have played a key role in accelerating microstructure quantification and feature

extraction tasks, based on several methods such as *n*-point spatial correlations (*n*-point statistics), for the micro-structural quantification and principal component analysis (PCA) for feature recognition [9–11]. Recently, deep learning approaches have facilitated the extraction of higher-order feature information for the establishment of PSP links [12]. Moreover, spatial statistics can be combined with deep learning-based extracted features to build unified models, leading to improved accuracy in terms of properties prediction [13,14].

Building upon a formerly developed artificial neural network (NN)-based framework, see [4–7], in this work we propose an improved methodology in terms of composition of the datasets, optimization of the architectural hyperparameters and overall training strategy. This data-driven protocol allows to account for material related uncertainties governing the performance of polycrystalline structures at the microscale, while incorporating microfabrication-induced geometric uncertainties, to produce accurate structure-property mappings for a (single-axis) Lorentz force micro-magnetometer [2,3]. The response of the device is characterized in terms of the maximum oscillation amplitude of its polysilicon resonant structure.

The remainder of this work is organized as follows. Methodological details are discussed in Section 2. Results are reported and analyzed in Section 3. Finally, concluding remarks and insights for the future research work are collected in Section 4.

2. Methodology

2.1. Data-Driven Multiscale Approach

Figure 1 provides a general description of the adopted model. Two approaches are proposed to address the upscaling of the mechanical property of interest, in this case the in-plane Young's modulus $\bar{E}$ of the polysilicon film. A homogenization procedure carries the information from the length scale of the squared $h \times h$ domains, representing stochastic volume elements (SVEs) in the Model 1, to the length scale of the resonant structure of the device, i.e., the scale of the polysilicon beam of size $L \times h$ in Model 2. Dataset 1 defines the set of values h while Dataset 2 defines a set of values $L \times h$ where $n = L/h$ must be defined a priori and defines the size of the input layer of Model 2. Consequently, the combined use of Model 1 and Model 2 establishes a data-driven homogenization framework for all the possible beam geometries $L \times h$ for which $L/h = n$. In this work, Dataset 1 is composed by elements described by the set $h = \{2 \ \mu m, 3 \ \mu m, 5 \ \mu m \text{ and } 10 \ \mu m\}$ and $n = 20$ has been defined as the dimensionality of the input vector to Model 2. Moreover, only $h = 2 \ \mu m$ has been considered for assembling the elements of Dataset 2, so that all the $L \times h$ structures correspond to samples featuring $40 \ \mu m \times 2 \ \mu m$.
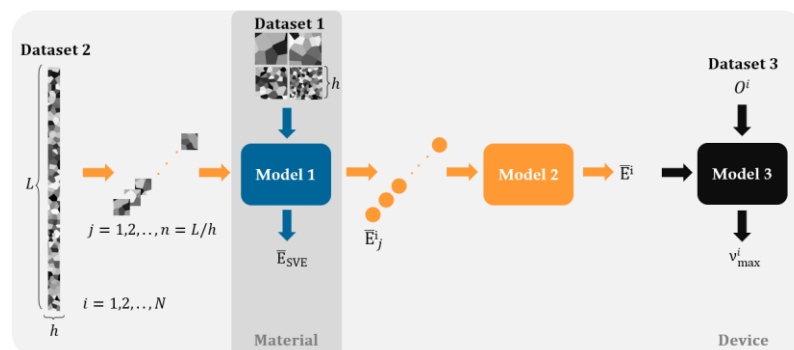
Figure 1. Overall model schematic, composed of three data-driven sub models to account for material- and geometry-related uncertainties at different (termed material and device) length scales.

The homogenization framework can therefore be exploited to produce microstructure-informed inference models for the polycrystalline structures, whose characteristic length scale is compatible with the scale defined by Model 2. This solution can be the case

for typical structures of MEMS devices. An example of it is represented by Model 3, which employs the output of Model 2, together with an additional variable (representing a microfabrication defect termed over-etch, O) as inputs to predict the maximum oscillation amplitude v_{max} of the resonant structure in a Lorentz force MEMS magnetometer, leveraging available analytical models in [2,3] to generate the ground truth data. O has been modeled according to the statistical distribution reported in [15,16].

2.2. Implementation Details

To find the optimal model architecture, a search space strategy is adopted and the KerasTuner hyperparameter optimization framework [17] is employed with Random-Search as the tuning algorithm. To explore the search space, a maximum number of 50 trials, with one execution per trial is established. Model 1 consists of a convolutional NN with the following arrangement of layers: 2D.Conv (filters = {32,64,96,128}, kernel_size = {3,5}) + Max.Pooling (pool_size = (2,2)) + 2D.Conv (filters = {32,64,96,128}, kernel_size = {3,5}) + Max.Pooling (pool_size = (2,2)) + Flatten followed by a single neuron output layer, Dense (units = 1). Adamax (learning rate = {0.01,0.001,0.0001}) has been selected as the optimizer. Here, values specified within curly brackets represent the elements of the associated search space.

Similarly, both Model 2 and Model 3 consist of feed forward NN which feature an arrangement of 3 hidden layers of the type Dense (units = {50, 100, 150}), followed by a single neuron output layer Dense (units = 1). The same configuration for Adamax was used again. Moreover, mean squared error (MSE) loss function, batch size equal to 10 and early stopping with a patience equal to 50 were used in all the models.

2.3. Input Data Generation

2.3.1. Dataset 1

Stochastic volume elements were digitally generated via the regularized Voronoi tessellation procedure described in [18] and formerly adopted in [4–7]. Associated target values of the homogenized Young's modulus $\bar{E}_{SVE}$ were obtained via standard FE simulations. Exploiting the anisotropic elasticity of silicon [19], gray level values between 0 and 1 encoded the in-plane lattice orientation of the crystalline domains, measured with respect to a global reference axis. These squared $h \times h$ domains were representative of the texture of polysilicon thin films at the microscale. For the current work, a finite set of characteristic values, namely h = {2 μm, 3 μm, 5 μm and 10 μm} was considered. Specifically, the same dataset reported in [7] for the training of the *material-level neural network* was adopted for the fitting and assessment of Model 1, only differing in the selection of a lower image resolution of 64 × 64 pixels, rather than the original 128 × 128 pixels. A comparison between these two resolutions is depicted in Figure 2.


Figure 2. Pairs of images featuring original resolution (left) and reduced resolution (right), for all the considered values of *h*.

To quantify the variation of information associated with the adopted resolution, the Kullback-Leibler divergence was computed to measure the distance between the discrete probability distributions of the pixel values for the images at larger and lower resolutions, respectively represented via P and Q. These distributions were obtained by normalizing the histograms of the pixel values. Specifically, the reverse KL i.e., KL(Q‖P) was computed, with Q regarded as an approximation of P. Results showed that even for the samples displaying the largest KL(Q‖P) distance, an acceptable match of the modes of P was obtained by Q, thus validating the adopted image resolution adjustment.

2.3.2. Dataset 2

As introduced in Section 2.1, Model 2 was trained based on a dataset composed of N samples of size $L \times h$, where $N = 1050$, $L = 40$ μm and $h = 2$ μm, so that $n = 20$. At variance with former works [4–7], where each slender beam was considered as a concatenation of a single individual domain of size $h \times h$; in the current work a microstructure was digitally generated for the entire geometry of the beam following the same procedure specified in Section 2.3.1. Figure 3 illustrates one of the generated beam microstructures, displaying a resolution of 64 × 1280 pixels.
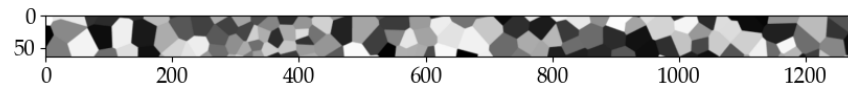
Figure 3. Exemplary 2 μm × 40 μm beam structure of Dataset 2.

Each beam indexed by i ($i = 1, 2, \ldots, N$) is partitioned into a set of j ($j = 1, 2, \ldots, n$) 64 × 64 pixels subdomains, that are fed to the already trained Model 1 to generate a vector of predictions that constitute the input of Model 2. Model 2 is trained to map the vector of predictions $\bar{E}^i_j$ from Model 1 into the homogenized Young's modulus $\bar{E}^i$ of the entire beam structure. Additionally, following a format given by *number of samples*, *mean* and *standard deviation* of the target values, Dataset 2 was composed by: *Training set = 800, 149.5 GPa, 1.3 GPa, Validation set = 150, 149.6 GPa, 1.3 GPa, Test set = 100, 149.6 GPa, 1.4 GPa*. An additional test set, namely, *Test set 2 = 91, 145.0 GPa, 26.3 GPa*, consisted of 91 additional data points simulating beams of a hypothetically homogeneous material (so that no real microstructures need to be generated), displaying values of Young's modulus in the range [100 GPa, 190 GPa]. This set was used to assess the generalization capabilities of Model 2 outside the range of target values characterizing the polysilicon beam microstructures, given by [144.9 GPa, 154.0 GPa].

2.3.3. Dataset 3

To generate the ground truth data, i.e., the maximum oscillation amplitude v_{max} of the resonant structure, this work relies on the analytical model of the single axis Lorentz force MEMS magnetometer derived in [3], accounting for the weakly coupled thermo-electro-magneto-mechanical multi-physics governing the vibrations of the slender clamped-clamped resonant beam. The specific geometric parameters i.e., the length, width, and thickness of the beams, were respectively given by 40 μm, 2 μm and 5 μm and were considered in the analytical model to assemble Dataset 3, featuring: *Training set = 800, 1.67 nm, 0.032 nm, Validation set = 100, 1.67 nm, 0.029 nm, Test set = 100, 1.67 nm, 0.032 nm.*

3. Results

3.1. Homogenization of the Stochastic Volume Elements (SVE)

Within the tested configurations from the search space, the best performing model was obtained for the combination of hyperparameters given by {128, 5, 128, 3, 0.0001}. With this configuration, Model 1 consisted of a total of 176,001 parameters. A very good performance was obtained for the homogenization at the SVE scale, improving the results reported in [7]. The obtained MSE values correspond to 0.2037 GPa², 0.0907 GPa², 0.0874 GPa², 1.4589 GPa², 0.1802 GPa² for the Training set ($h = 2, 10$ μm), Validation set ($h = 2, 10$ μm), Test set 1 ($h = 2, 10$ μm), Test set 2 ($h = 3$ μm) and Test set 3 ($h = 5$ μm) sets, respectively.

3.2. Homogenization of the Resonant Structure

The best performing model was obtained for the combination of hyperparameters given by {50, 50, 50, 0.001}, resulting in an architecture with a total of 6201 fitting parameters for Model 2. Figure 4 illustrates the good agreement between predicted and ground-truth data obtained by Model 2 on the test sets. The equation for the linear least squares (LLS) fitting of the predictions is reported in the parity plots, as well as the associated coefficient of determination R^2.
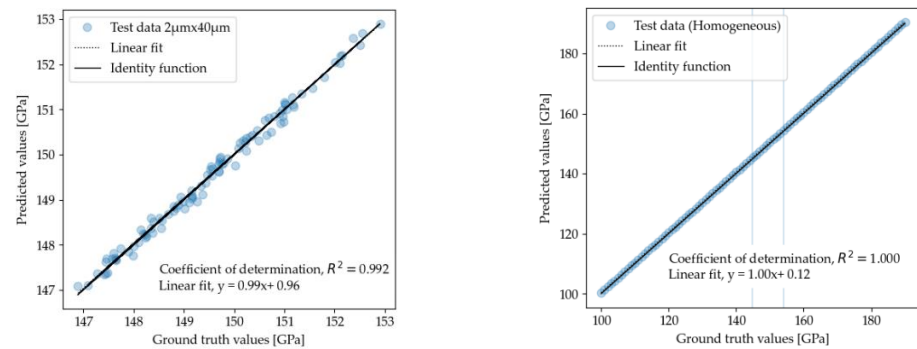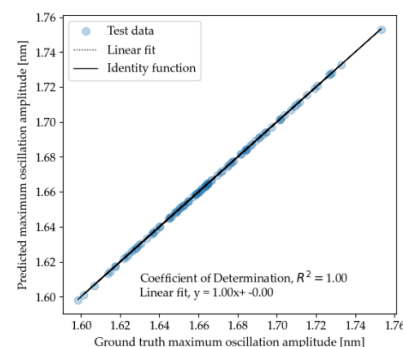

Figure 4. Performance of Model 2 on the test sets specified in Section 2.3.2.

The absolute percentage error associated to the mean and standard deviation computed from the set of predictions, corresponds to 0.0% and 7.7% for the Training set, 0.1% and 0.0% for the Validation set, 0.0% and 0.0% for the Test set and 0.28% and 0.0% for Test set 2. Moreover, the MSE values associated are 0.0222 GPa^2, 0.0149 GPa^2, 0.0163 GPa^2 and 0.1755 GPa^2, respectively. These results confirm that Model 2 is capable to generate predictions consistent with an effective homogenization scheme, reproducing the *trivial* expected output for the simulated beams featuring hypothetical homogeneous materials.

3.3. Maximum Oscillation Amplitude of the Lorentz Force MEMS Magnetometer

The best performing model was obtained for the combination given by {50, 150, 50, 0.01}. With these hyperparameters, Model 3 consists of a total of 15,401 fitting parameters. Figure 5 summarizes the performance of Model 3 on the test set, which can produce exact one-to-one predictions. In this case the MSE values associated to each of the sets correspond to 2.099×10^{-8} nm^2, 2.065×10^{-8} nm^2, 1.966×10^{-8} nm^2 for the Training, Validation, and Test sets, respectively.

Figure 5. Performance of Model 3 on the test set specified in Section 2.3.3.

4. Conclusions

A data-driven modelling scheme was implemented and allowed the effective upscaling of the properties across the identified hierarchy of length scales in MEMS devices, accounting for the microstructure of complex textured structural films (e.g., made of pol-

ycrystalline materials). The proposed protocol can be readily exploited to model the response of intricate systems, such as MEMS devices, by incorporating additional relevant variables as inputs of a *device-level model*, capable to learn complex structure-property mappings in a microstructure-informed data-driven manner.

Future research activities will focus on adapting the proposed homogenization framework to become more general and therefore less constrained by the specific device geometry.

References

1. Mariani, S.; Ghisi, A.; Corigliano, A.; Martini, R.; Simoni, B. Two-scale simulation of drop-induced failure of polysilicon MEMS sensors. *Sensors* **2011**, *11*, 4972–4989. https://doi.org/10.3390/s110504972.
2. Bagherinia, M.; Bruggi, M.; Corigliano, A.; Mariani, S.; Lasalandra, E. Geometry Optimization of a Lorentz Force, Resonating MEMS Magnetometer. *Microelectron. Reliab.* **2014**, *54*, 1192–1199. https://doi.org/10.1016/j.microrel.2014.02.020.
3. Bagherinia, S.M. Mariani. Stochastic effects on the dynamics of the resonant structure of a Lorentz force MEMS magnetometer. *Actuators* **2019**, *8*, 36. https://doi.org/10.3390/act8020036.
4. Quesada-Molina, J.P.; Mariani, S. A two-scale multi-physics deep learning model for smart MEMS sensors. *J. Mater. Sci. Chem. Eng.* **2021**, *9*, 41–52. https://doi.org/10.4236/msce.2021.98004.
5. Quesada-Molina, J.P.; Mariani, S. Two-Scale Deep Learning Model for Polysilicon MEMS Sensors. *Comput. Sci. Math. Forum* **2022**, *2*, 12. https://doi.org/10.3390/IOCA2021-10888.
6. Quesada-Molina, J.P.; Mariani, S. Hybrid Model-Based and Data-Driven Solution for Uncertainty Quantification at the Microscale. *Micro Nanosyst.* **2022**, *14*, 281–286. https://doi.org/10.2174/1876402914666220328123601.
7. Quesada-Molina, J.P.; Mariani, S. Deep Learning-based Multiscale Modelling of Polysilicon MEMS. In Proceedings of 2022 23rd International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems (EuroSimE), St. Julian, Malta, 24–27 April 2022. https://doi.org/10.1109/EuroSimE54907.2022.9758899.
8. Kalidindi, S.R. Materials, Data, and Informatics. In *Hierarchical Materials Informatics*; Elsevier: Amsterdam, The Netherlands, 2015; pp. 1–32. https://doi.org/10.1016/b978-0-12-410394-8.00001-1.
9. Kalidindi, S.R.; Niezgoda, S.R.; Salem, A.A. Microstructure informatics using higher-order statistics and efficient data-mining protocols. *JOM* **2011**, *63*, 34–41. https://doi.org/10.1007/s11837-011-0057-7.
10. Gupta, A.; Cecen, A.; Goyal, S.; Singh, A.K.; Kalidindi, S.R. Structure-property linkages for non-metallic inclusions/steel composite system using a data science approach. *Acta Mater* **2015**, *91*, 239–254. https://doi.org/10.1016/j.actamat.2015.02.045.
11. Çeçen, A.; Fast, T.; Kumbur, E.C.; Kalidindi, S.R. A data-driven approach to establishing microstructure-property relationships in porous transport layers of polymer electrolyte fuel cells. *J. Power Sources* **2014**, *245*, 144–153. https://doi.org/10.1016/j.jpowsour.2013.06.100.
12. Bock, F.E.; Aydin, R.C.; Cyron, C.J.; Huber, N.; Kalidindi, S.R.; Klusemann, B. A Review of the Application of Machine Learning and Data Mining Approaches in Continuum Materials Mechanics. *Front. Mater.* **2019**, *6*, 110. https://doi.org/10.3389/fmats.2019.0011.
13. Cecen, A.; Dai, H.; Yabansu, Y.C.; Klidindi, S.R.; Song, L. Material structure-property linkages using three-dimensional convolutional neural networks. *Acta Mater* **2018**, *146*, 76–84. https://doi.org/10.1016/j.actamat.2017.11.053.
14. Ganapathi, A.S. Prediction of two-phase composite microstructure properties through deep learning of reduced dimensional structure-response data. *Comp. Part B Eng.* **2021**, *225*, 109–282. https://doi.org/10.1016/j.compositesb.2021.109282.
15. Mirzazadeh, R.; Mariani, S. Uncertainty Quantification of Microstructure-Governed Properties of Polysilicon MEMS. *Micromachines* **2017**, *8*, 248. https://doi.org/10.3390/mi8080248.
16. Mirzazadeh, R.; Eftekhar Azam, S.; Mariani, S. Mechanical Characterization of Polysilicon MEMS: A Hybrid TMCMC/POD-Kriging Approach. *Sensors* **2018**, *18*, 1243. https://doi.org/10.3390/s18041243.
17. O'Malley, T.; Bursztein, E.; Long, J.; Chollet, F.; Jin, H.; Invernizzi, L. Keras Tuner. 2019. Available online: https://github.com/keras-team/keras-tuner (accessed on).

18. Mariani, S.; Martini, R.; Ghisi, A.; Corigliano, A.; Beghi, M. Overall elastic properties of polysilicon films: A statistical investigation of the effects of polycrystal morphology. *Int. J. Mult. Comp. Eng.* **2011**, *9*, 327–346. https://doi.org/10.1615/IntJMultCompEng.v9.i3.50.
19. Hopcroft, M.A.; Nix, W.D.; Kenny, T.W. What is the Young's Modulus of Silicon? *J. Micro. Syst*. **2010**, *19*, 229–238. https://doi.org/10.1109/JMEMS.2009.2039697.