

Neural networks based surrogate modeling for efficient uncertainty quantification and calibration of MEMS accelerometers

Filippo Zacchei^a, Attilio Frangi^b, Andrea Manzoni^a

^a*Politecnico di Milano, MOX - Department of Mathematics, p.za Leonardo da Vinci, 32, Milano, 20133, Italy*

^b*Politecnico di Milano, Department of Civil and Environmental Engineering, p.za Leonardo da Vinci, 32, Milano, 20133, Italy*

Abstract

This paper addresses the computational challenges inherent in the stochastic characterization and uncertainty quantification of Micro-Electro-Mechanical Systems (MEMS) capacitive accelerometers. Traditional methods, such as Markov Chain Monte Carlo (MCMC) algorithms, are often constrained by the computational intensity required for high-fidelity (e.g., finite element) simulations. To overcome these limitations, we propose to use supervised learning-based surrogate models, specifically artificial neural networks, to effectively approximate the response of MEMS capacitive accelerometers. Our approach involves training the surrogate models with data derived from initial high-fidelity finite element analyses (FEA), providing rich datasets to be generated in an offline phase. The surrogate models replicate the FEA accuracy in predicting the behavior of the accelerometer under a wide range of fabrication parameters, thereby reducing the online computational cost without compromising accuracy. This enables extensive and efficient stochastic analyses of complex MEMS devices, offering a flexible framework for their characterization. A key application of our framework is demonstrated in estimating the sensitivity of an accelerometer, accounting for unknown mechanical offsets, over-etching, and thickness variations. We employ an MCMC approach to estimate the posterior distribution of the device's unknown fabrication parameters, informed by its response to transient voltage signals. The integration of surrogate models for mapping fabrication parameters to device responses, and subsequently to sensitivity measures, greatly enhances both backward and forward uncertainty quantification, yielding accurate results while significantly improving the efficiency and effectiveness of the characterization process. This process allows for the reconstruction of device sensitivity using only voltage signals, without the need for direct mechanical acceleration stimuli.

Keywords: deep learning, uncertainty quantification, Bayesian analysis, MEMS, computational mechanics

1. Introduction

Micro-Electro-Mechanical Systems (MEMS) [1, 2] have emerged as one of the most prevalent technologies in today's world, manifesting in various forms across numerous devices we use daily. These systems integrate micro-structural and electronic components to miniaturize a range of functions, such as accelerometers, gyroscopes, magnetometers, and pressure sensors, down to the micron scale. However, the small scale characteristic of MEMS leads to complex fabrication processes that are influenced by multiple sources of uncertainty. These uncertainties must be thoroughly accounted for and characterized to accurately calibrate each device and effectively harness its measurements [3–6], for instance leading to geometric discrepancies with respect to the nominal layout prescribed before fabrication. Fabrication variances can result in up to a 10% deviation from the intended design in standard MEMS [7, 8], significantly modifying their operational characteristics and thus requiring a comprehensive calibration process to ensure accurate measurements.

Over the past few decades, significant advancements in computational power and the development of algorithms have established numerical simulation as a cornerstone in engineering and applied sciences [9]. This is particularly relevant for solving complex differential problems, such as those governing the dynamics of MEMS [10]. As a result, it has become feasible to explore the impact of uncertainties on various design

parameters [11, 12]. However, the substantial computational resources required by high-fidelity models often limit their practical utility, especially for tasks like Uncertainty Quantification (UQ), which are inherently multi-query – that is, they require repeated queries to a numerical solver. Consequently, lower-fidelity surrogate models have been increasingly adopted to enhance computational efficiency. While less resource-intensive, these models still provide an acceptable level of accuracy and expedite repeated computations. Surrogate models can be categorized into various types [13]: data-fit models (e.g., Kriging models [7, 14–16]), hierarchical models (e.g., simplified physics models or coarse discretizations [17]), and reduced-order models [18–23]. In the domain of structural mechanics, the most widely used techniques currently rely on reduced-order models that involve the use of modal decompositions or reduced bases. Modes’ selection can be based on physical insights or system identification procedures [21], such as Proper Orthogonal Decomposition [19, 23–26]. Alternatively, approaches are formulated within the framework of Nonlinear Normal Modes [22, 27, 28], with several developments that have generalized the method to arbitrary order expansions and can be directly applied to finite element models, using the so-called direct parametrization of invariant manifolds (DPIM) [29, 30]. However, in recent years, a plethora of approaches for constructing data-driven surrogate models [31] have been proposed, often relying on supervised learning [14], in computational mechanics. Several methodologies have successfully integrated data-driven deep learning techniques for parameter estimation and uncertainty quantification across various domains, including cardiac mechanics [32], solid mechanics [33], structural health monitoring [34], and aeroelastic systems [35], among others. In the field of MEMS devices and uncertainty quantification, the core challenge lies in determining the relationship between selected device features and the system’s behavior. Recent results have shown the feasibility of the use of deep learning-based reduced order models [36, 37] in order to simulate the mechanical behavior of MEMS also involving nonlinearities [38], ultimately comparing the DPIM method with more recent model discovery strategies based on the sparse identification of nonlinear dynamics [39, 40].

In a production environment for MEMS, various tests are usually carried out to assess each device’s response and performance under external stimuli [41], also aiming at real-time auto calibration [4]. However, inferring device characteristics directly from collected data relying on model-based methodologies often yields ill-posed problems, whose solution is even more difficult due to the unavoidable measurements’ noise. Consequently, a probabilistic framework, enabled by Bayesian model updating, emerges as a natural solution to address these issues. Bayesian model updating, in particular, provides a set of general and robust strategies for parameter estimation when dealing with physics-based models. Nonetheless, the complexity of evaluating posterior probability density functions usually calls into play computationally demanding Markov Chain Monte Carlo (MCMC) methods [42–44]. These methods require simulating device responses across different parameter sets through several (thousands of) iterations for each new device. To mitigate the computational burden, the integration of data-driven surrogate models proves highly advantageous, substantially reducing the cost per iteration. Alternative strategies to lower the computational cost include decreasing the number of required iterations, as seen in transitional MCMC [45], accelerating the convergence of the chain, e.g., through Hamiltonian MCMC [46] or the No-U-Turn sampler [47], or employing multi-level simulation fidelities, like in the case of Multi-Level MCMC methods [48]. The usage of a probabilistic learning framework for quantifying model-form uncertainties in MEMS has been proposed in [49]. An approach based on Bayesian analysis for sensor calibration has been proposed in [50]. Finally, the combination of deep learning techniques with Bayesian model update has been greatly investigated in [51].

To focus on a real-life practical case in microsystems’ design, this work aims to characterize and identify a series of geometrical uncertainties inherent in MEMS axial capacitive accelerometers. The investigation begins with the design of a numerical test, which involves applying an active voltage to a portion of the accelerometer’s electrodes and determining the device’s response based on its characteristics. A comprehensive electro-mechanical model of the accelerometer is developed, and finite element analysis (FEA) is carried out to generate a robust dataset of simulations. These simulations serve as the basis for training a supervised learning surrogate model, exploiting artificial neural networks (ANNs) [52], to learn the map between input features and the device electro-mechanical response to the stimulus. Subsequently, a Bayesian procedure relying on MCMC methods is employed to infer the posterior distribution of the device’s input features from noisy data. A second configuration is then subsequently implemented to construct another

surrogate model, this time focusing on predicting the device’s sensitivity, here intended as the device electrical response to a 1g acceleration, based on its features. This approach enables uncertainty quantification of the device’s sensitivity, grounded in the posterior distribution of its features. The novelty of this work lies in the application of the proposed methodology, meticulously adapted to MEMS accelerometers, and the specifically designed numerical tests to assess device characteristics. Bayesian analysis for MEMS calibration is indeed still an area to be greatly investigated [5]. The approach outlined here paves the way for integrating data-driven deep learning models with Bayesian model updating in the physical characterization of devices. Employing a data-driven mapping from parameters to responses, developed within a supervised learning framework, facilitates a significant acceleration in the evolution of the MCMC algorithm. This enhancement is attributed to the immediate assessment of likelihood for each newly generated sample. In contrast, utilizing a Finite Element Analysis (FEA) simulation based on the Euler-Bernoulli approximation [53] would necessitate minutes per sample, while a comprehensive FEA simulation could extend this requirement to hours for each instance. In addition the accuracy characterizing the developed surrogate model, as assessed in the work, allows to recover the same performance of the high fidelity models. Eventually, the two stage procedure, based on the development two distinct surrogate models, allows to easily forward propagate uncertainties on additional quantities of interests based on the geometrical parameters of the device. The proposed methodology is broadly applicable across a wide range of MEMS devices. However, for the purpose of simplifying result analysis, our study is restricted to the simpler, yet of strong industrial relevance, of one-dimensional accelerometers.

The paper is organized as follows. Section 2 introduces the case of one dimensional capacitive accelerometer, highlighting their functioning and the role of fabrication inaccuracies. Section 3 details the methodology employed in this study, starting from the formulation of the model and elaborating on the development of the surrogate model and Bayesian model updating. Section 4 presents the results of our numerical experiments. After introducing a full-order model for the numerical simulation of a class of devices, we use it to assess the devices’ behavior under varying unknown features, and to construct a comprehensive numerical dataset. Then, we discuss the training of the surrogate models and their accuracy. The identification results, obtained from noisy data using Markov Chain Monte Carlo (MCMC) methods, and the forward reconstruction of the device’s sensitivity are finally illustrated. Section 5 provides a summary of the work, some concluding remarks, and further perspectives on possible future developments.

2. Capacitive Accelerometer: Problem Description

This section briefly outlines the operational principle of a MEMS capacitive accelerometer and examines the impact of geometric uncertainties on its performance. Capacitive sensing is the predominant mechanism in contemporary micro-accelerometers. The geometry of the x-axis MEMS accelerometer is presented in Figure 1. The device comprises a movable mass, shown in orange, which functions as a rotor and is anchored to the substrate by two supports, depicted in blue. The device’s body is connected to the anchors through springs composed of folded beams, allowing for motion and compliance. The mass is maintained at a ground voltage. Parallel to the plates of the movable mass are placed two sets of electrodes acting as a stator and fixed to the electrode layer of the MEMS. These are denoted as the Left (in red) and Right (in gray) electrode groups, with an arbitrary voltage V_l and V_r . The capacitances of these groups, denoted as C_l and C_r respectively, vary over time in response to the movement of the accelerometer’s body.

With the only purpose of elucidating the operational mechanism of this type of device, we introduce a simplified model of a 1D capacitive x-accelerometer, as depicted in Figure 2. This model features a proof mass. Attached to this proof mass is an electrode, positioned equidistantly, at a distance d , from two conductive plates. These plates, referred to as the left and right electrodes, have an identical area A and are maintained at potentials V_l and V_r , respectively. If the gap d is much smaller than the dimensions of the plates, their capacitance may be approximated with the analytical formula $C = \epsilon_0 \epsilon_r A / d$. Any displacement u of the proof mass along the gap direction induces differential changes in capacitance. If C_l and C_r denote

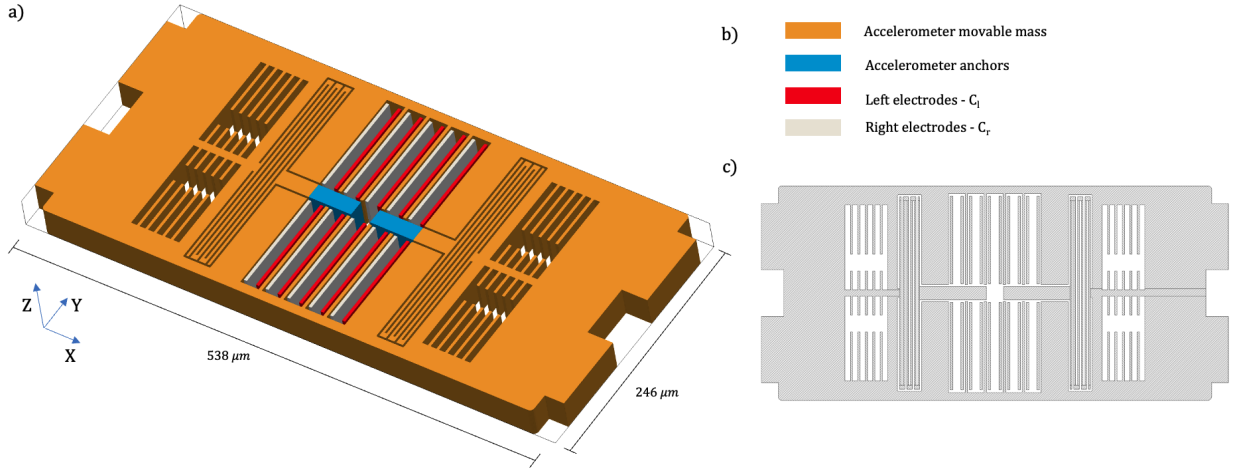


Figure 1: X-axis MEMS accelerometer: (a) Device geometry, (b) Color-coded representation of device regions, (c) Device top view.

the capacitance of the left and right capacitors, respectively, and u is assumed positive to the right, one has:

$$\Delta C_r = \epsilon_r \epsilon_0 A \left[\frac{1}{d-u} - \frac{1}{d} \right], \quad \Delta C_l = \epsilon_r \epsilon_0 A \left[\frac{1}{d+u} - \frac{1}{d} \right]. \quad (1)$$

It is worth stressing that these formulas are based on the assumption of infinitely large plates, an approximation that is generally justified when the plate gap is significantly smaller than their area. For more accurate analytical capacitance estimations, fringe field effects, as described by Palmer's formula [54], must be considered. In general, however, to calculate the capacitance variation for general devices and arbitrary configurations, numerical coupled electro-mechanical simulations are required, solving for both the displacement and electric field, and this approach will be adopted and discussed in Section 3.2. The hardest obstacles arising in the three dimensional case concern the coupled evaluation of the deformation of the geometry and the computation electric field, which is needed since the analytical approximation for the capacitance is not accurate enough in our case and Fringe fields play an important role.

Once the capacitance variation is known, the sensitivity S of the accelerometer can be computed, defined as the ratio between the capacitance change ΔC and the external acceleration a_{0x} :

$$S := \frac{\Delta C}{a_{0x}} = \frac{\Delta C_r - \Delta C_l}{a_{0x}}. \quad (2)$$

The sensitivity of a capacitive accelerometer is significantly affected by the material, structural, and geometric features of the MEMS design. Variations introduced during the fabrication process can alter the structural configuration of the accelerometer. These variations can influence both the electrode capacitance and the device's elastic stiffness, thereby affecting the accelerometer's overall sensitivity.

The final objective of our research is the development of a cost-effective testing method to be employed during the production phase of MEMS accelerometers for sensitivity calibration. Typically, sensitivity is determined by conducting mechanical tests, where the devices are subjected to external accelerations and their displacements are recorded. However, these tests are labor-intensive and require to move accelerometers to specialized equipments. In contrast, tests based on electrical activations can be more conveniently carried out due to easier access to electrode potentials. While mechanical testing directly yields sensitivity values, electrical activation tests can be used to assess first fabrication uncertainties which, in a second stage, can then be used to deduce the accelerometer's sensitivity. The methodology specifically developed for addressing this issue is elaborated in the subsequent section.

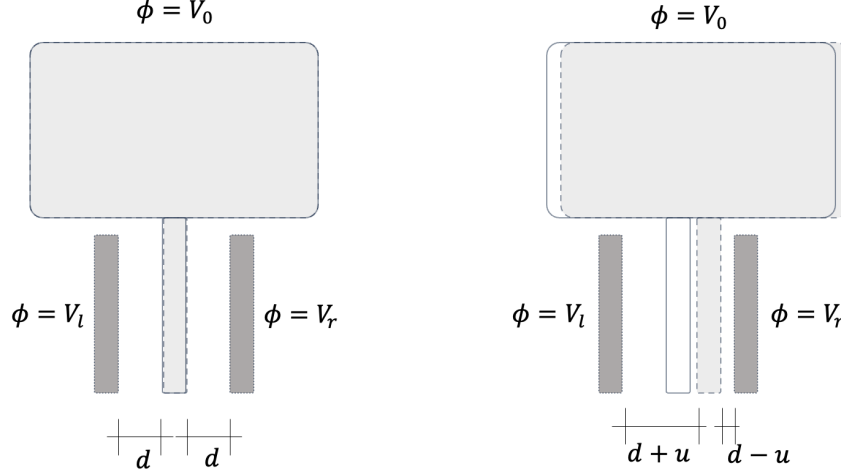


Figure 2: Parallel-plate sensing model

2.1. The Role of Fabrication Uncertainties

In our study’s current setup, we identify three potential sources of fabrication uncertainty in the accelerometer, as illustrated in Figure 3. The first aspect deals with the mechanical offset U of the device (see Fig. 3-a). This offset can arise from process tolerances, such as uneven gaps between the differential parallel plates or mechanical stresses induced by wafer bending, which may occur due to temperature gradients during processing or operational stresses from wafer bonding. Such factors can cause the accelerometer’s proof mass to deviate from its nominal centered position, moving it closer to one set of external electrodes. This displacement can lead to a non-zero output ($\Delta C \neq 0$) even in the absence of external accelerations. In our analysis, we disregard residual stresses within the accelerometer, concentrating solely on geometrical offsets.

As a second possible fabrication defect, we consider the over-etching of the device (see Fig. 3-b). This is defined as a deviation in the in-plane geometry of the movable structure, stemming from variations in temperature and etchant concentration, as outlined in [55]. Such fluctuations result in the device’s actual geometry differing from its intended design, typically yielding a final structure marginally smaller or larger than its nominal dimensions in all directions. For the sake of simplicity, we assume a uniform over-etch O on each side of the nominal geometry. This parameter not only affects the device’s capacitance by increasing the gap between electrodes, but also alters its mechanical characteristics. In particular, it reduces the width of the commonly incorporated folded beams, thereby adjusting the device’s stiffness and consequently changing its eigen-frequencies.

Lastly, we consider unknown the device’s thickness T , defined as the height of the mechanical layer of the MEMS device and we include it as a further uncertain parameter (Fig. 3-c).

3. Methodology

In this section, we detail the methodology adopted for characterizing a MEMS accelerometer from a mechanical point of view. Our primary objective is to investigate the impact of uncertain geometric parameters on the device’s performance relying on the solution of a parameter estimation problem, enhanced by the use of neural network (NN) based surrogate models.

To build a set of snapshots required to train our surrogate models, a series of finite element analysis (FEA) simulations has been performed for two distinct scenarios: in the former, we examine the electro-mechanical response of the device under varying electric voltage stimuli. In this case, the response coincides with the dynamic variation of capacitance between the left and right electrodes of the device during the external stimulus; in the latter, we evaluate the sensitivity of the device, defined as its response, i.e. the capacitance

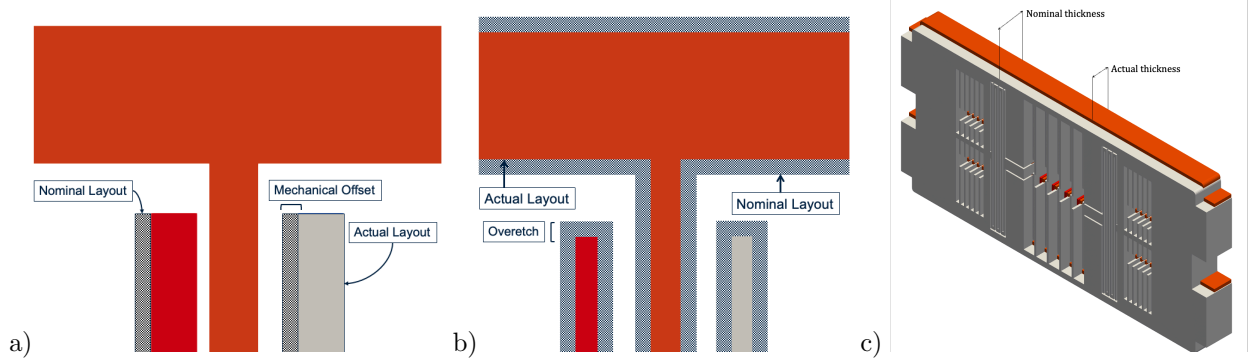


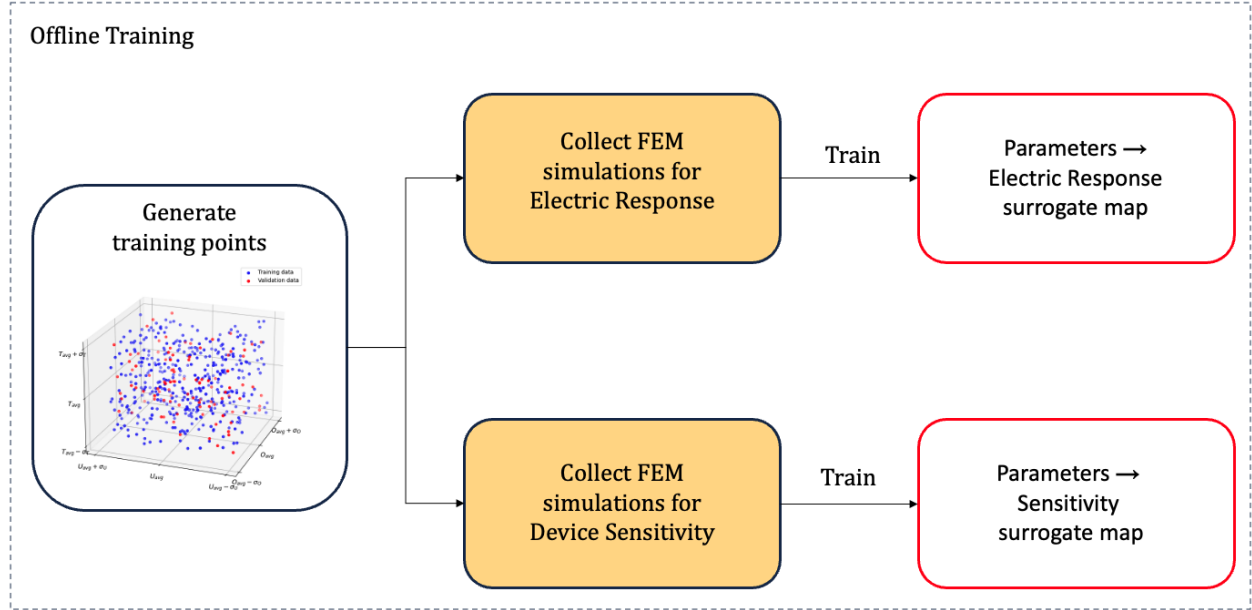
Figure 3: Fabrication uncertainties in the MEMS accelerometer: (a) Mechanical offset, (b) Over-etching, (c) Thickness variability.

variation among left and right electrodes, to a standard mechanical input of $1g$ inducing a movement of the mass. Based on these data, we develop two separate surrogate models to approximate the functional relationships between the device parameters and its electro-mechanical responses. Once trained, the first surrogate model has been employed within a Bayesian model updating procedure to infer the unknown input parameters from observed device responses in the case of active voltage stimulus. Compared to a least-square estimation, a Bayesian framework allows to treat input parameters as random variables themselves, therefore providing their posterior probability distributions informed by the data, and not only a point estimate. This is of key importance not only for the chance to derive any desired statistics from the posterior distribution (like, e.g., its mean, its covariance, the maximum a posteriori, to name a few examples) but also – even more importantly – for the possibility to sample from the posterior distribution and propagate the uncertainty from the (estimated) parameters to any output quantity of interest such as, e.g., the sensitivity of the device, in the spirit of forward uncertainty quantification. This latter task requires to (i) sample from the posterior distribution and (ii) exploit the second surrogate model – once trained – for the sake of output evaluation. Overall, the very rapid evaluation capability of both surrogate models significantly reduces the computational time required for each sample generation from the posterior distribution, and the subsequent forward UQ analysis. This efficiency is crucial, considering the potentially high number of iterations needed to accurately capture the posterior distribution’s characteristics, and the need to quantify uncertainty in (possibly, several) output quantities of interest. A schematic description of the methodology is shown in Fig. 4.

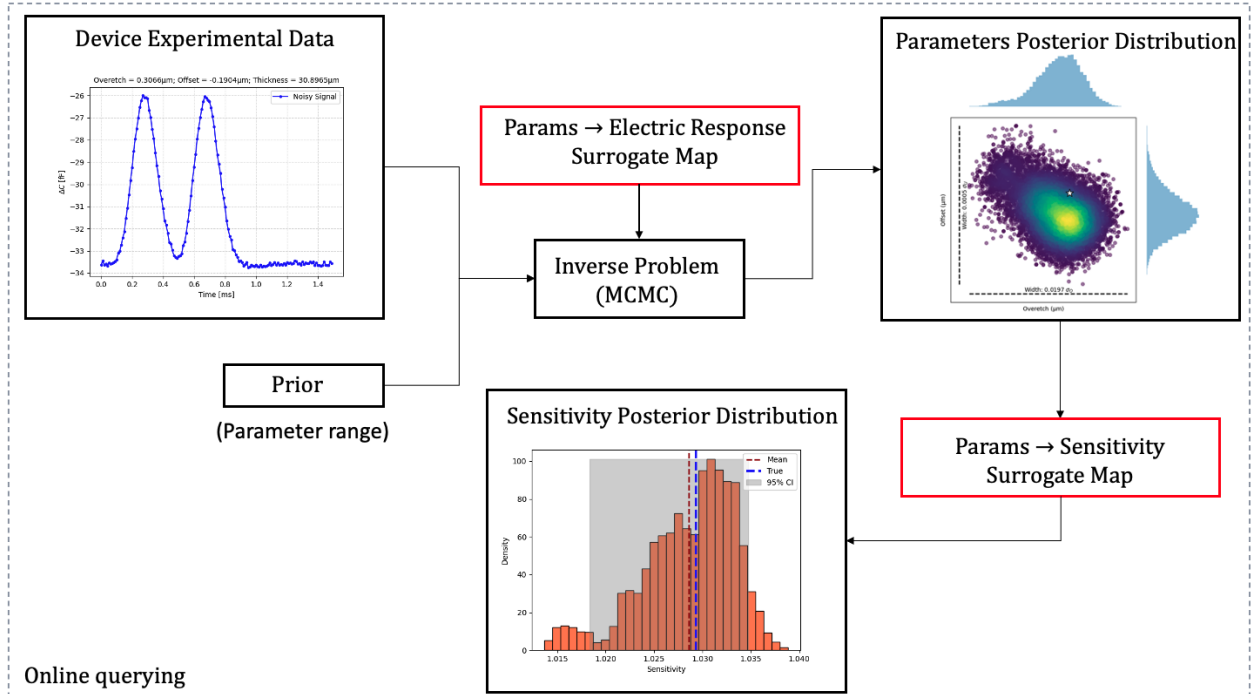
We observe that an alternative methodology might involve the direct formulation of a surrogate model connecting the device’s sensitivity to electro-mechanical response to the active voltage. However, due to the presence of multiple and distinct combinations of geometric parameters to yield identical sensitivities despite differing responses to voltage stimuli, a direct mapping from sensitivity to voltage response is impracticable. Consequently, the construction of a singular-stage surrogate model, along with its associated identification routine, is unviable. The two-stage approach we propose is therefore favoured not only for its operational flexibility but also for its congruence with the simulation workflow.

3.1. Problem formulation

We start by presenting an overview and establish the formal notation for the modeling principles and assumptions that constitute the high-fidelity model used throughout our work. Our focus is the electro-mechanical problem associated to a MEMS device under external stimuli. We start by considering a device occupying a spatial volume, denoted as $\Omega_0 \subset \mathbb{R}^3$, at a null electric potential, specifically $V_0 = 0$ [V]. The device’s motion is analyzed in a non-inertial reference frame attached to the MEMS system. The spatial domain in the reference configuration, represented by the material coordinates \mathbf{X} , undergoes deformation characterized by the displacement field $\mathbf{x} = \boldsymbol{\chi}(\mathbf{X}) = \mathbf{X} + \mathbf{u}(\mathbf{X}, t)$. This deformation results in an actual configuration denoted by Ω , see Figure 5. The boundary of the reference configuration, $\partial\Omega_0$, is subdivided into two components: the Dirichlet boundary $\partial\Omega_{0D}$ and the Neumann boundary $\partial\Omega_{0N}$. Dirichlet boundary



(a) Development and training of supervised learning surrogate models



(b) Bayesian model updating with experimental data

Figure 4: The overall methodology of probabilistic identification framework for device geometric features and sensitivity using surrogate two-stage surrogate modelling.

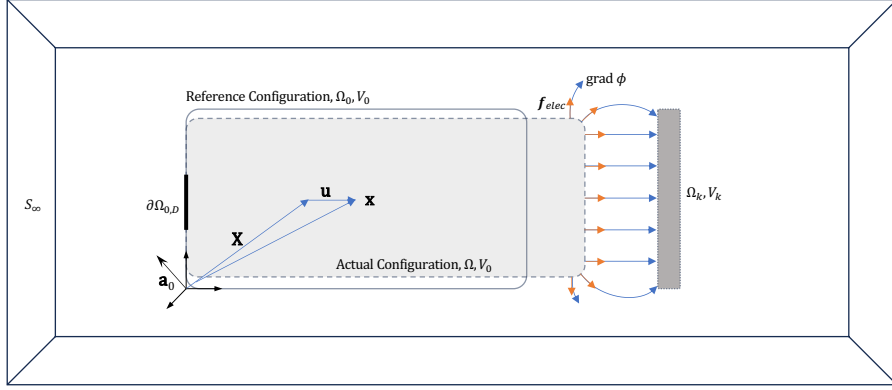


Figure 5: Electro-mechanical problem coordinate scheme and reference quantities describing the PDE system.

conditions are characterized by the imposition of a null displacement, while Neumann conditions are associated with the application of electric forces. Hereon, we denote by \mathbf{N} and \mathbf{n} the normal vector to the device surface in the material and the current configuration, respectively. Electrostatic potentials are applied to the external electrodes of the device, thus making it subject to an electric field. The device is enclosed within an unbounded volume Ω_∞ , where Gauss's law is prescribed, and a zero electric flux condition is assumed on its surface S_∞ . The governing equations of the system are expressed as follows:

$$\rho_0 \ddot{\mathbf{u}}(\mathbf{X}, t; \boldsymbol{\mu}) + \mathbf{C} \dot{\mathbf{u}}(\mathbf{X}, t; \boldsymbol{\mu}) - \nabla_{\mathbf{X}} \cdot \mathbf{P}(\mathbf{u}(\mathbf{X}, t; \boldsymbol{\mu}); \boldsymbol{\mu}) = -\rho_0 \mathbf{a}_0 \quad \text{in } \Omega_0 \times \mathcal{T}, \quad (3)$$

$$\mathbf{P}(\mathbf{u}(\mathbf{X}, t; \boldsymbol{\mu}); \boldsymbol{\mu}) \cdot \mathbf{N}(\mathbf{X}) = \mathbf{f}_{elec}(\mathbf{X}) \quad \text{on } \partial\Omega_{0N} \times \mathcal{T}, \quad (4)$$

$$\mathbf{u}(\mathbf{X}, t) = \mathbf{0} \quad \text{on } \partial\Omega_{0D} \times \mathcal{T}, \quad (5)$$

$$\mathbf{u}(\mathbf{X}, 0) = \mathbf{0} \quad \text{in } \Omega_0, \quad (6)$$

$$\dot{\mathbf{u}}(\mathbf{X}, 0) = \mathbf{0} \quad \text{in } \Omega_0, \quad (7)$$

$$\text{div}(\text{grad}\phi(\mathbf{x})) = 0 \quad \text{in } \Omega_\infty \setminus \Omega \times \mathcal{T}, \quad (8)$$

$$\phi(\mathbf{x}) = V_k(t) \quad \text{on } \partial\Omega_k, \quad (9)$$

$$\text{grad}\phi(\mathbf{x}) \cdot \mathbf{n} = 0 \quad \text{on } S_\infty. \quad (10)$$

Equation (3) represents the conservation of momentum, where ρ_0 represents the initial density in the reference configuration. The vectors $\mathbf{u}, \dot{\mathbf{u}}, \ddot{\mathbf{u}} \in \mathbb{R}^3$ denote displacement, velocity, and acceleration, respectively. The term \mathbf{a}_0 is the acceleration of the non-inertial frame of reference, as observed from an external inertial viewpoint. The terms \mathbf{C} and \mathbf{P} are the damping matrix (built according to the Rayleigh model) and the Piola-Kirchhoff stress tensor, respectively. The parameter vector $\boldsymbol{\mu} \in \mathcal{P} \subset \mathbb{R}^p$ encapsulates the set of input parameters of the problem, which include all the geometrical fabrication uncertainties we want to focus on. These parameters belong to a bounded set representing the parameter space, with p denoting the number of parameters. Equations (4) and (5) represent the Neumann and Dirichlet boundary conditions, respectively.

The electrostatic pressure on the conductor surface, \mathbf{f}_{elec} , is expressed in the material configuration, originating from the following spatial formulation:

$$\tilde{\mathbf{f}}_{elec}(\mathbf{x}) = \frac{1}{2\epsilon_0\epsilon_r} \left(\epsilon_0\epsilon_r \frac{\partial\phi(\mathbf{x})}{\partial\mathbf{n}} \right)^2 \mathbf{n}, \quad (11)$$

where ϵ_0 denotes the vacuum dielectric permittivity constant and ϵ_r is the relative permittivity coefficient of air, and ϕ the electrostatic potential scalar field. The initial conditions for displacement and velocities are encapsulated in Equations (6) and (7), assumed to be null, indicating the device's commencement from a

rest position. Equation (8) represents Gauss’s law for the external vacuum surrounding the device, whereas Equations (9) and (10) pertain to the active potentials applied to the fixed electrodes of the device and the condition of zero electric flux at a surface infinitely distant from the device, respectively.

In the context of our study, we adopt the hypothesis of small strains and displacements: hence displacements, along with the shape and volume variations of the solid body, are negligible compared to the body’s dimensions. This hypothesis is well described by the Saint Venant-Kirchoff constitutive model:

$$\mathbf{S}(\mathbf{X}, t) = \mathcal{A}(\mathbf{X}) : \mathbf{E}(\mathbf{X}, t) \quad \text{for } (\mathbf{X}, t) \text{ in } \Omega_0 \times \mathcal{T} \quad (12)$$

where \mathbf{S} is the second Piola-Kirchoff stress, \mathbf{E} is the Green-Lagrange strain tensor, which is expressed as $\mathbf{E} = 1/2 (\nabla_{\mathbf{X}} \mathbf{u}(\mathbf{X}, t) + \nabla_{\mathbf{X}}^T \mathbf{u}(\mathbf{X}, t) + \nabla_{\mathbf{X}}^T \mathbf{u}(\mathbf{X}, t) \cdot \nabla_{\mathbf{X}} \mathbf{u}(\mathbf{X}, t))$ and \mathcal{A} is the fourth-order elasticity tensor, which, in the context of isotropic materials, is determined by specifying the values for the Young’s modulus E and Poisson’s ratio ν . Consequently, the first Piola-Kirchoff stress tensor is $\mathbf{P} = (\mathbf{I} + \nabla_{\mathbf{X}} \mathbf{u}(\mathbf{X}, t)) \cdot \mathbf{S}(\mathbf{X}, t)$.

The problem described so far can be discretized (in both space and time) to obtain a high fidelity solver relying, e.g., on the finite element method. It is noteworthy that our formulation utilizes material coordinates to describe mechanical equilibrium and spatial coordinates for electrostatic equilibrium, thus adding a further layer of complexity to the development of intrusive reduced order models. Despite in principle possible, the use of projection-based ROMs exploiting, e.g., the proper orthogonal decomposition for the construction of a low-dimensional subspace and a (Petrov-)Galerkin projection for the generation of the reduced-order problem, becomes rather involved in the case of coupled problems, due to the intrusive character of the method, the need to account for every variable of the problem in the ROM, and the usually complex parameter dependencies [56]. Recent strategies relying on projection-based ROMs in computational mechanics can be found, e.g., in [19, 24–26, 57–59]. For the aforementioned reasons, and given the specific need of approximating only input-output relationship in the calibration context at hand, we rather opt for data-driven surrogate models to approximate the output quantities of interest in this study. Indeed, for the case at hand, the output quantities of interest depend quite smoothly on the input parameters, thus making the approximation of this map a rather simple task. Nevertheless, the availability of a set of snapshots from the high-fidelity solver, and the relatively fast training process of the surrogate models at hand, shall make the construction of further models to approximate different output quantities of interest still depending on the electro-mechanical behavior of the MEMS at hand quite straightforward.

3.2. The development of surrogate models

We rely therefore on Finite Element Analysis (FEA) as high-fidelity numerical solver to generate snapshots data, forecasting the response of our device under various external stimuli. To this aim, we design a numerical experiment that involves dynamically varying the active potential, following a predetermined voltage profile on one set of the device’s electrodes. The resulting electrostatic forces prompt the accelerometer to move toward the activated electrodes, causing a time-varying capacitance change ΔC . This output response is influenced by the device’s uncertain parameters $\boldsymbol{\mu}$ (offset, overetch, thickness). We aim to learn a map $(\boldsymbol{\mu}, t) \mapsto \Delta C$, in order to assess this relationship efficiently, for several combinations of our input parameters. To this aim, we employ *artificial neural networks* (ANN) trained using the numerical data derived from our high-fidelity simulations. Once trained, this surrogate model will then be exploited inside a Bayesian sampler for a more rapid – yet accurate – evaluation of the input-output map for any new input parameters’ value.

Simultaneously, we conduct a numerical simulation to assess the accelerometer’s sensitivity, characterized by its unknown parameters. This is done under a constant external acceleration \mathbf{a}_0 and uniform voltage across all electrode sets. Through this approach, we construct a new input-output relationship $(\boldsymbol{\mu}) \mapsto S$, for which a second surrogate model is developed. The primary application of this second surrogate is to facilitate forward uncertainty quantification of the device’s sensitivity, utilizing the posterior distribution obtained from Bayesian sampling.

3.2.1. Artificial Neural Networks

In this study, we focus on a neural network (NN)[52] architecture tailored for scalar map regression. This architecture is composed of L hidden layers, each containing N_{n_l} neurons, where the variable l specifies the

layer number within the range $1 \leq l \leq L$. The network employs a nonlinear activation function ψ_l for each layer. The input vector $\mathbf{x}^{in} \in \mathbb{R}^{d_{in}}$ is processed through these layers to produce a scalar output $x^{out} \in \mathbb{R}$, prescribing $N_L = 1$. The inter-layer connectivity is defined by the weights W_{ij}^l , where j and i represent the neuron indices in the $(l-1)$ -th and l -th layers, respectively. Bias parameters are denoted as b_i^l for the i -th neuron in the l -th layer, with indices satisfying $1 \leq j \leq N_{l-1}$ and $1 \leq i \leq N_l$. The neural network's functional mapping from \mathbf{x}^{in} to x^{out} is established through iterative linear transformations and nonlinear activations across each layer. The output of each neuron in layer l is a combination of the weighted sum of outputs from the previous layer $(l-1)$ and the bias, subsequently passed through the nonlinear activation function ψ_l . Mathematically, this process is represented as follows:

$$x_i^0(\mathbf{x}^{in}) = x_i^{in} \quad 1 \leq i \leq d_{in} \quad (13)$$

$$x_i^l(\mathbf{x}^{in}) = \psi_l \left(b_i^l + \sum_{j=1}^{N_{l-1}} W_{ij}^l x_j^{l-1}(\mathbf{x}^{in}) \right) \quad 1 \leq i \leq N_{n_l}, 1 \leq l < L \quad (14)$$

$$x^{out} = b^L + \sum_{j=1}^{N_{L-1}} W_j^L x_j^{L-1}(\mathbf{x}^{in}). \quad (15)$$

For the problems at hand, we aim to approximate a multivariate scalar function, denoted by $x^{out} = f(\mathbf{x}_{in})$, using a neural network-based surrogate model, $\mathbf{f}^{NN}(\cdot; \mathbf{W}, \mathbf{b}) : \mathbb{R}^{d_{in}} \rightarrow \mathbb{R}^{d_{out}}$. This model is trained on a set of input-output pairs $\{(\mathbf{x}^{(k)}, y^{(k)})\}_{k=1}^{N_{data}}$, where \mathbf{W} and \mathbf{b} encapsulate all the network's weights and biases, respectively. The parameter N_{data} represents the total number of pairs, derived from Finite Element Analysis (FEA). In our case, this number coincides with the multiplication of the total simulations N_{sim} count by the number of time steps per simulation N_t for the parameters to electric response surrogate map, and just N_{sim} for the sensitivity surrogate map.

3.3. Parameter Identification and Uncertainty Quantification

We now introduce and formalize the application of a Bayesian model updating strategy, specifically tailored for estimating the sensitivity of MEMS accelerometers and providing uncertainty quantification. This strategy is grounded on device data acquired during experiments that monitor the electro-mechanical response of the device under electrical input stimuli. Central to our approach are the two surrogate models previously developed. Initially, we employ a Markov Chain Monte Carlo (MCMC) analysis to derive the posterior distribution of the device parameters. This analysis utilizes experimental data, acknowledging the inherent noise in the observed responses, and the (first) surrogate model. Subsequently, repeated queries to the (second) surrogate model are performed drawing samples from the posterior distribution of the parameters, in order to recover the predictive posterior distribution of the device's sensitivity, based on its electro-mechanical response. This two-stage approach – first focusing on parameter estimation, and then on uncertainty propagation – enables a comprehensive uncertainty quantification of the MEMS device at hand.

3.3.1. Bayesian Inference

In our study, we adopt a Bayesian inference approach to infer the posterior distribution of device parameters. This process relies on the choice of a prior distribution on the input parameters, a set of noisy, observational data (of the capacitance change) and a likelihood function that assimilates those data. The posterior distribution $\pi(\boldsymbol{\mu} \mid \boldsymbol{\Delta C})$ of the input parameters $\boldsymbol{\mu}$, given the data $\boldsymbol{\Delta C}$, is obtained through the Bayes theorem as

$$\pi(\boldsymbol{\mu} \mid \boldsymbol{\Delta C}) = \frac{\pi(\boldsymbol{\Delta C} \mid \boldsymbol{\mu}) \pi(\boldsymbol{\mu})}{\pi(\boldsymbol{\Delta C})}, \quad (16)$$

where $\boldsymbol{\Delta C}$ is the experimental data vector collecting the recorded measurements on the capacitance change for a tested device, that in our case it coincides with the vector $\boldsymbol{\Delta C}(\boldsymbol{\mu}) := \{\Delta C(t; \boldsymbol{\mu})\}_{t=t_1, \dots, t_{N_t}}$. The value $\pi(\boldsymbol{\mu})$ represents the prior distribution, $\pi(\boldsymbol{\Delta C} \mid \boldsymbol{\mu})$ the likelihood function, while $\pi(\boldsymbol{\Delta C})$ is the marginal distribution, acting as a normalization factor.

Regarding the prior, we assume a uniform distribution on the three parameters, within a suitable bound. Based on the assumptions that measurement errors are present, and that they are independent and identically normally distributed with zero mean and fixed variance σ^2 accounting for noise in measurement acquisitions, the likelihood function takes the following form [60]:

$$\pi(\Delta \mathbf{C} \mid \boldsymbol{\mu}) := \left(\frac{1}{2\pi\sigma^2} \right)^{\frac{n}{2}} \exp \left[-\frac{SS_q}{2\sigma^2} \right], \quad (17)$$

where

$$SS_q := \sum_{i=1}^{N_t} [\Delta C(t_i; \boldsymbol{\mu}) - f(\boldsymbol{\mu}, t_i)]^2 \quad (18)$$

is the sum of the squares error. Here $\Delta C(t_i; \boldsymbol{\mu})$ denotes the data at time step t_i and $f(\boldsymbol{\mu}, t_i)$ is the response given by the surrogate model.

To compute the posterior distribution $\pi(\boldsymbol{\mu} \mid \Delta \mathbf{C})$ without directly determining $\pi(\Delta \mathbf{C})$, we rely on a Monte Carlo Markov Chain method, specifically the Metropolis-Hastings algorithm. This algorithm sequentially generates a sample series to approximate the desired distribution, with each new sample based on the previous one. At each iteration i , the algorithm proposes a candidate $\boldsymbol{\mu}'$, based on the current sample $\boldsymbol{\mu}^i$. This candidate is either accepted—forming the basis for the next iteration—or rejected, in which case $\boldsymbol{\mu}^i$ is retained. The acceptance probability is determined by the ratio:

$$\alpha = \frac{\pi(\Delta \mathbf{C} \mid \boldsymbol{\mu}') \pi(\boldsymbol{\mu}')}{\pi(\Delta \mathbf{C} \mid \boldsymbol{\mu}^i) \pi(\boldsymbol{\mu}^i)}, \quad (19)$$

comparing the proposed and current samples in relation to the desired distribution.

4. Results

In this section, we detail the results and validation of our methodology on an x-axis capacitive accelerometer. The investigation is structured as follows: initially, we describe in detail the full order finite element model (FEM) simulations. Next, we detail the development and accuracy assessment of the surrogate models. Lastly, we present the findings from the parameter identification process and uncertainty quantification concerning the device's sensitivity.

4.1. Problem Description and FOM Results

Two distinct simulation configurations were employed to characterize the device's response to an active voltage signal and to determine its sensitivity. The first configuration represents an electro-mechanical test, which could be utilized at the end of the device fabrication process for product characterization. In this setup, an active voltage is applied to the right electrodes, denoted as C_r , while the left electrodes and the device body are maintained at ground potential. This active voltage induces an electrostatic force, causing the device to move towards the active electrodes and consequently altering the values of both C_l and C_r . The active voltage profile is designed as two complete sine wave periods, varying from 0 V to 1.8 V. The mathematical representation of this profile is as follows:

$$V_r = 0.9 \times (1 - \cos(2\pi ft)) \quad [\text{V}] \quad \text{for } 0 \leq t \leq 2T \quad [\text{s}], \quad (20)$$

$$V_r = 0 \quad [\text{V}] \quad \text{for } t > 2T \quad [\text{s}], \quad (21)$$

$$V_l = 0 \quad [\text{V}] \quad \text{for } t \geq 0 \quad [\text{s}], \quad (22)$$

where f is the impulse frequency set to 2500 Hz (close to resonant frequency) and $T = 1/f$. This specific profile was chosen as an optimal balance between simplicity, as each period of the sine wave necessitates simulation time, and generated information. The choice of a two-periodic input is instrumental in enhancing the identifiability of the device's frequency response, which is highly affected by the device overetch. A single

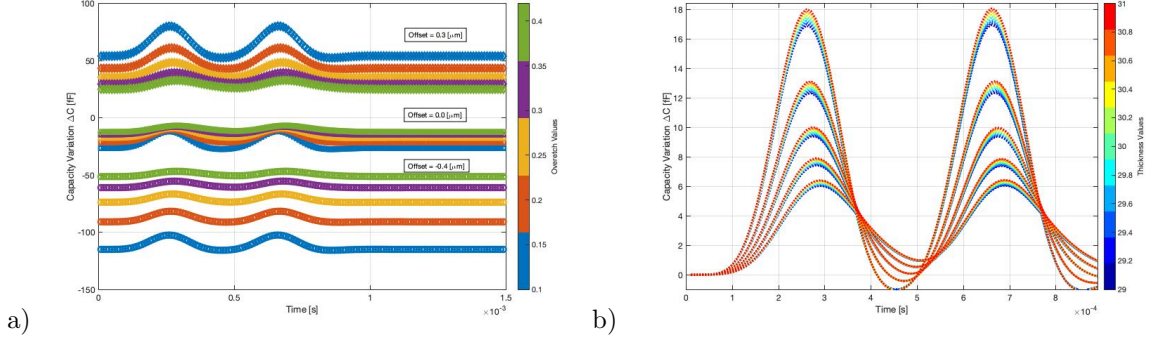


Figure 6: Numerical results obtained using Coventor MEMS+ for the first simulation configuration: (a) fixed thickness $T = 30.0\mu\text{m}$ (b) fixed offset $U = 0.0\mu\text{m}$

period often proves insufficient for the system to reach a steady operational regime, hence the necessity of spanning two complete periods. Additional periods result in additional information but augmenting the cost of each simulation and the complexity of a surrogate model. As a final output of the simulation, we retain the capacitance difference in time between right and left electrodes, i.e. $\Delta C(t) := \Delta C_r(t) - \Delta C_l(t)$, for t ranging from 0 to 1.5s , therefore leaving time to the device to come back to rest position after the periodic impulse. A second configuration is designed to mimic operating conditions, where both electrode groups are set at 0.69 V and an external axial acceleration of $a_0 = 1g$ is applied to the MEMS accelerometer. This setup is aimed at determining the device's sensitivity, which is defined as $S := (\Delta C_{final} - \Delta C_0)/1g$ $[F/N]$, where ΔC_0 stands for the capacitance difference before the acceleration is applied and ΔC_{final} at the equilibrium position when $1g$ acceleration is present.

Both simulations were conducted using Coventor MEMS+TM [61]. This software employs Euler-Bernoulli elements to model the beams and conformal mappings for describing the electrostatic forces, utilizing eight integration points per electrode. For the first configuration we use forward Euler scheme damping the solution every 0.01ms , for a total of $N_t := 150$ time steps. It is important to note that the predicted displacements in these simulations are minimal relative to the size of the elements. This allows us to neglect geometric nonlinearities, as the nonlinear effects in this scenario are predominantly induced by electromechanical interactions. The use of classical beam theory in Coventor MEMS+ accelerates the simulations by up to 100 times compared to a full finite element simulation. This acceleration is particularly advantageous in scenarios like ours, where the device's behavior needs to be tested across several geometric parameters to generate an extensive training dataset. However, the simulation times are still significant, limiting the feasibility of using this software for repeated iterations, such as in the MCMC scenario, highlighting the necessity of employing a surrogate model.

4.1.1. Coventor Results

The results depicting the time-dependent variation of the capacitance (ΔC) for a range of thickness (T), offset (U), and overetch (O) values are illustrated in Figure 6. The analysis clearly demonstrates the significant influence of the offset parameter on the resting value of ΔC , as well as the pronounced impact of overetch on the capacitance response amplitude to the impulse. Conversely, the thickness parameter exhibits a relatively minor effect on the output within the investigated range. This may lead to challenges in its identification, particularly in scenarios with high noise levels.

Snapshots of the solution for the displacement and voltage field for the second configuration of the simulation at equilibrium, considering a device with $0.5\mu\text{m}$ overetch, $-0.1\mu\text{m}$ offset and $30\mu\text{m}$ thickness are plotted in Figure 7.

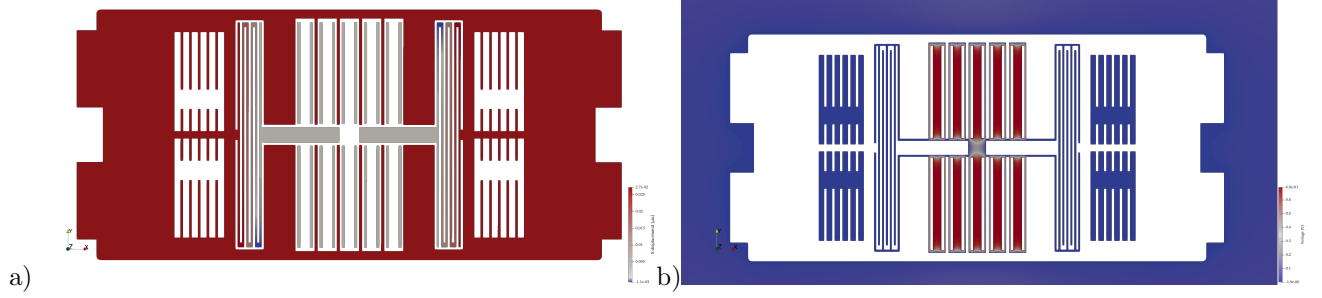


Figure 7: Snapshots of the numerical solution obtained using Coventor MEMS+ for the second simulation configuration: (a) displacement (b) electric field.

4.2. Surrogate model results

In this section, we present the numerical results obtained from the surrogate models applied to two simulation configurations. For these simulations, fully connected artificial neural networks (FCNN) are employed as the surrogate model. The training dataset comprises 800 unique combinations of three parameters: mechanical offset U , unilateral over-etch O , and thickness T . The training parameter space is defined as follows: for U , it is centered around $U_{\text{avg}} = 0.0 \mu\text{m}$ with a range of $[U_{\text{avg}} - \sigma_U, U_{\text{avg}} + \sigma_U]$; for O , it is centered around $O_{\text{avg}} = 0.0 \mu\text{m}$ with a range of $[O_{\text{avg}} - \sigma_O, O_{\text{avg}} + \sigma_O]$; and for T , it is centered around $T_{\text{avg}} = 30.0 \mu\text{m}$ with a range of $[T_{\text{avg}} - \sigma_T, T_{\text{avg}} + \sigma_T]$. The beam width that constitutes the springs, adjusted for the mean overetch value, measures $2.1 \mu\text{m}$. Sensitivity metrics are normalized against a baseline sensitivity, which corresponds to a device characterized by mean values of overetch, offset, and thickness. This normalization is represented by the ratio S , defined as the sensitivity of the device relative to the baseline sensitivity. Sampling points within these ranges are generated using the Latin hypercube sampling technique. Due to confidentiality agreements with the company providing the accelerometer layout, specific boundary values for these parameters are not disclosed. Neural networks were trained using the Adam optimization algorithm, and a total number of epochs equal to 100000 which was tuned by hand. The dataset was subdivided in a 80-20 training-validation split to facilitate hyperparameter tuning within the training-validation framework. Batch training is employed using 40 mini-batches. The architectures of the neural networks and the L_2 regularization coefficient λ were calibrated using grid search optimization aimed at minimizing the validation loss, considering both the mean and variance. Parameters for training were selected through fine hand tuning, as they demonstrated a secondary influence on the overall model accuracy.

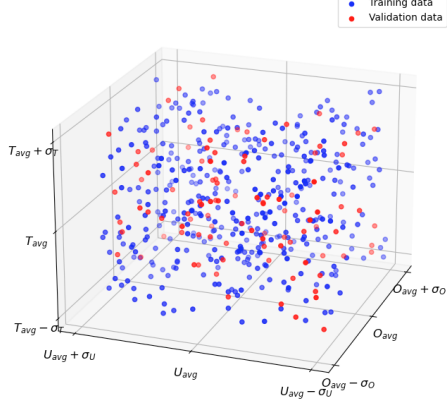
4.2.1. Configuration I: Electro-mechanical Response to an Active Voltage Signal

In this configuration, the training dataset comprised $N_{\text{sim}} = 800$ simulation samples, each generated using Coventor MEMS+. The dataset is visualized in Figure 8a. Each sample captures the time-dependent capacitance variation ΔC , considering time as an additional input parameter for the neural network. Consequently, the network functions as the mapping:

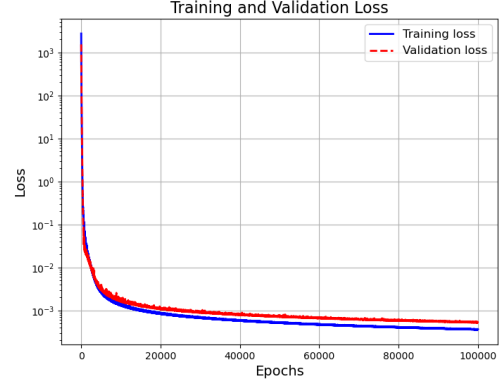
$$f_{NN}^{(I)} : (u_i, \mu_i, h_i; t^k) \mapsto \Delta C_i(t^k) \quad \text{for } i = 1, \dots, N_{\text{sim}}; k = 1, \dots, N_t.$$

The training errors for both the training and validation datasets are depicted in Figure 8b, highlighting the absence of overfitting. Predictions from the surrogate model for two sets of input parameters outside the training/validation dataset are illustrated in Figure 9, demonstrating a close match with the testing data. Using an additional linearly spaced grid of samples for testing, we obtain a Pearson correlation coefficient of 99.92%. A high number of epochs and accurate fine tuning is needed in particular when very low noise signals will be the starting point of the parameter identification routine, i.e. the noise is the same order of the error of the surrogate model.

The neural network's performance was compared against two different supervised learning strategies: Gaussian Process Regression (GPR) and Long Short-Term Memory networks (LSTM). The latter retains the



(a) Simulation samples used during the training process.



(b) Loss function evolution during the training process considering the training and the validation dataset.

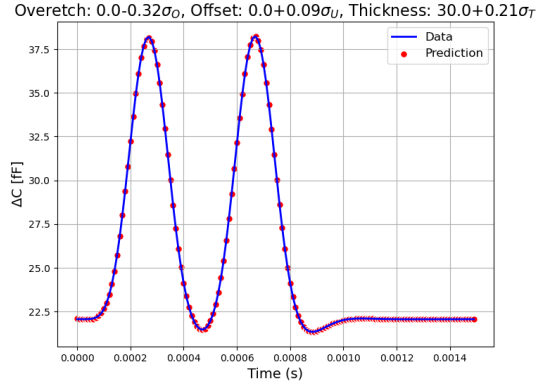
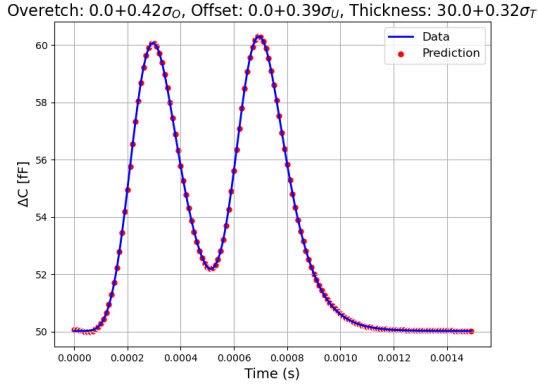


Figure 9: Comparison between testing data and surrogate model predictions for two input parameter combination unseen during training process.

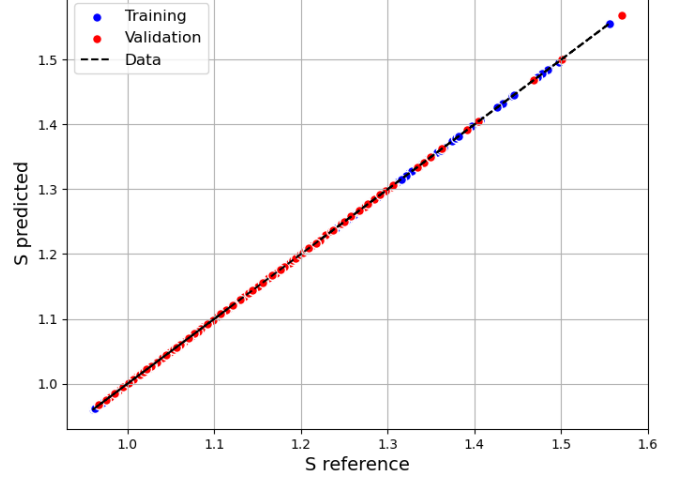
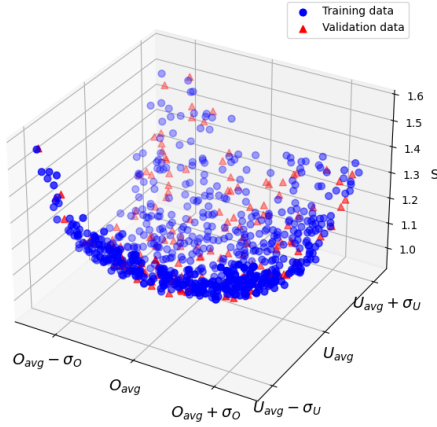
same network structure and generally shows more resistance to overfitting, albeit at a higher computational cost due to the need for solving a linear system at each evaluation. On the other hand, LSTM leverages time-series data during predictions, implemented by integrating an LSTM cell layer within the network and inputting the complete sequence of sample time values. Performance comparisons among the three models revealed minimal differences, thereby favoring the use of a Fully Connected Neural Network (FCNN) for its faster evaluation time relative to both GPR and LSTM.

This implementation utilized the TensorFlow Keras library for deep learning [62]. Detailed information regarding the final structure is documented in Table 1.

4.2.2. Configuration II: Electro-mechanical response for an external 1g acceleration

In this configuration, we train a surrogate model using the same samples as in the previous setup, but with the objective of inferring the sensitivity of the device. It is observed that the mapping from parameters to sensitivity is not bijective, unlike the case in the parameter-response map to active electric voltage. This remark justifies our adoption of a two-stage approach for uncertainty quantification. Accordingly, we design a second surrogate model to represent the mapping:

$$f_{NN}^{(II)} : (u_i, \mu_i, h_i) \mapsto S_i \quad \text{for } i = 1, \dots, N_{\text{sim}}.$$



(a) Scatter plot of Sensitivity vs overetch and offset. (b) Comparison of the predictions of the model and the data.

The training dataset is split in the same 80-20 ratio as before. Post-training predictions, shown in Figure 10b, indicate an almost perfect alignment between the surrogate model's predictions and the actual data. The architecture of this neural network is detailed in Table 2.

4.3. Bayesian analysis

In the Bayesian inference framework applied to the first configuration, we commence with experimental data $\Delta \mathbf{C}_{1, \dots, N_{\text{obs}}}^{\text{exp}}$, where $N_{\text{obs}} = N_t$, representing a time series of the device's response to an electrical signal. The surrogate model $f_{NN}^{(I)}$ is employed to approximate the posterior probability density function $\pi(\boldsymbol{\mu} \mid \Delta \mathbf{C}_{1, \dots, N_{\text{obs}}}^{\text{exp}})$. For experimental observations, we use testing data samples generated via Coventor MEMS+ that were not included in the training process. These data samples are numerically corrupted by the addition of Gaussian white noise, uncorrelated with time, characterized by zero mean and a standard deviation defined as $\sigma := n_l S \sqrt{B_W}$. Here, n_l denotes the noise level (in $N/\sqrt{\text{Hz}}$), B_W the total bandwidth of the device (in Hz), and S its sensitivity (in F/N). Conservatively, we assume B_W to be 200 Hz and S to be the reference one, in accordance with the device features.

Although an informative prior can offer valuable insights into the device parameters and statistical tools can be employed to design it, a uniform prior is assumed for all the device parameters $\boldsymbol{\mu}$ to prevent potential biases in the results. For monitoring the convergence of the Markov Chain, the Gelman and Rubin's

Layer	Nr. of neurons	Activation Function
1	4	<i>tanh</i>
2	64	<i>tanh</i>
3	64	<i>tanh</i>
4	64	<i>tanh</i>
5	64	<i>tanh</i>
6	64	<i>tanh</i>
7	64	<i>tanh</i>
8	64	<i>tanh</i>
9	1	<i>linear</i>

Table 1: Features of dense layers in the FCNN $f_{NN}^{(I)}$. The total number of parameters is 20800, kernel regularization coefficient is $\lambda = 1e - 12$.

Layer	Nr. of neurons	Activation Function
1	3	<i>tanh</i>
2	32	<i>tanh</i>
3	32	<i>tanh</i>
4	32	<i>tanh</i>
5	1	<i>linear</i>

Table 2: Features of dense layers in the FCNN $f_{NN}^{(II)}$. The total number of parameters is 2176, kernel regularization coefficient is $\lambda = 1e - 9$.

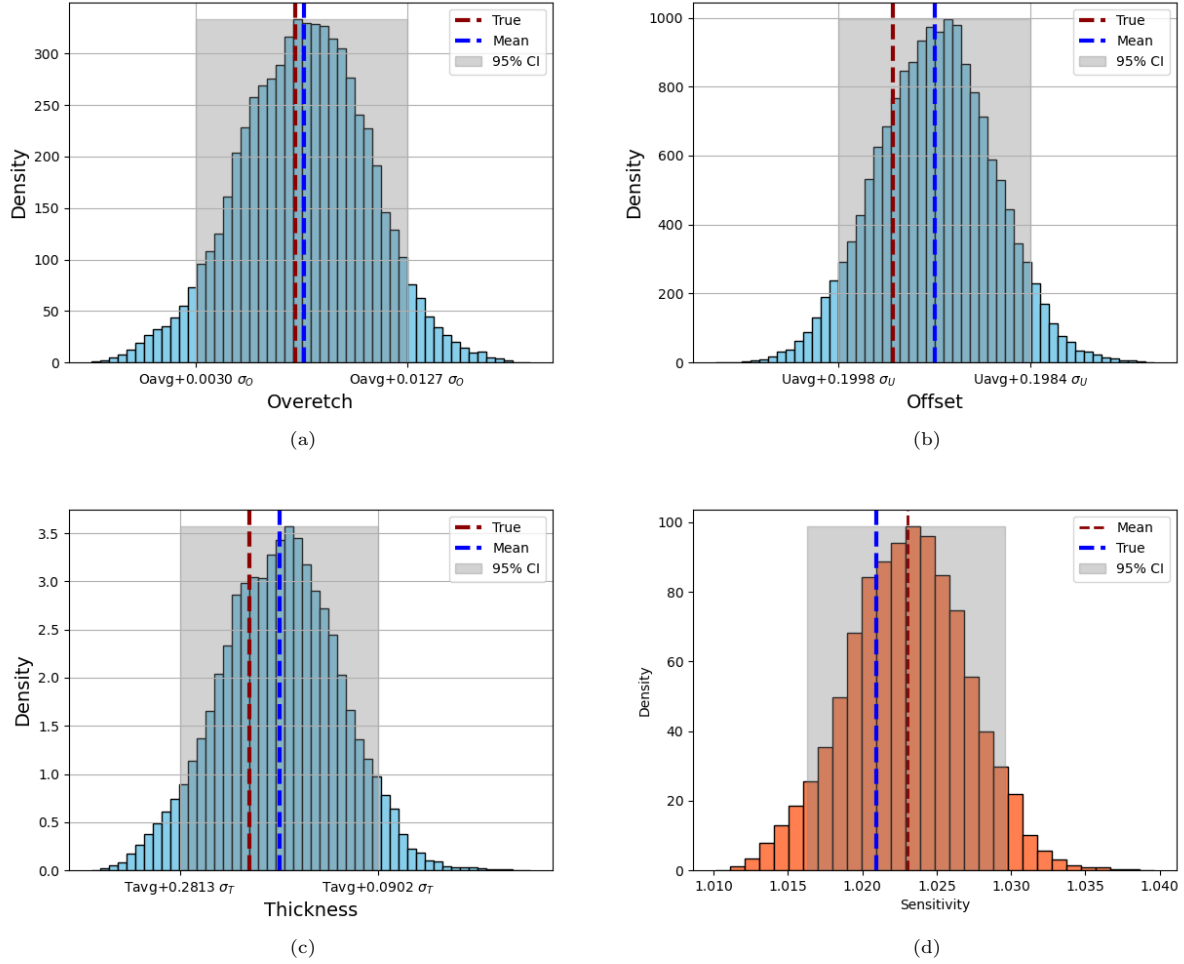


Figure 11: (a)-(c) Histogram representing densities of the posterior distribution for each input parameter; (d) Histogram of the density of the posterior distribution of sensitivity of the device.

convergence diagnostic \hat{R} [63] is employed. This statistic is computed from n independent Markov Chains initialized at different starting points μ^0 . The \hat{R} metric assesses the convergence of a multivariate chain by comparing the ratio of the between-chain variance estimate of samples to the average within-chain variance. Convergence is considered to be achieved when \hat{R} falls below a predetermined threshold, typically set at 1.1.

Once the MCMC samples are obtained, we leverage the second surrogate model $f_{NN}^{(II)}$ to infer the distribution of the sensitivity, by evaluating the model for each sample of the chain.

4.3.1. Results

The Bayesian procedure adopted in this study is implemented utilizing CUQIpy [64], a comprehensive and open-source Python package designed specifically for computational UQ in inverse problems.

In our Markov Chain Monte Carlo (MCMC) simulations, each chain commences with $N_{\text{obs}} = N_t$ observations of the time-dependent capacitance difference ΔC , associated with a specific set of input parameters $\mu = [o, u, h]$. To ensure comprehensive exploration of the parameter space, we initiate five independent random chains from distinct starting points. The total number of samples in each chain is adjusted to achieve a Gelman-Rubin statistic \hat{R} below 1.1, resulting in a collective sample count of approximately 500,000. To mitigate the effects of initial transients, the first 20% of the samples from each chain are discarded as part of

the burn-in process. Furthermore, to reduce the auto-correlation inherent in the chains, we thin the samples by retaining only every fifth sample post burn-in. The proposal distribution within the MCMC algorithm is a Gaussian random walk, characterized by a covariance matrix with diagonal elements. These elements are set to one hundredth (and squared) of the parameter range width, ensuring a finely tuned exploration of the parameter space. Additionally, the width of the random walk is dynamically adjusted throughout the simulation to maintain an acceptance rate between the optimal thresholds of 0.2 and 0.4. The results, derived under a noise factor of $n_f = 1000\mu\text{g}/\sqrt{\text{Hz}}$, are depicted in Figure 11. A notable observation from the top figure is the remarkably narrow confidence intervals for the overetch and offset parameters under this noise factor, indicating precise parameter estimation within the specified noise constraints. Looking at Figure ?? we see that the reconstruction error for the sensitivity is very low, and we provide again a narrow range for the confidence interval confirming the accuracy of the proposed method. The obtained results have been investigated for ten different devices.

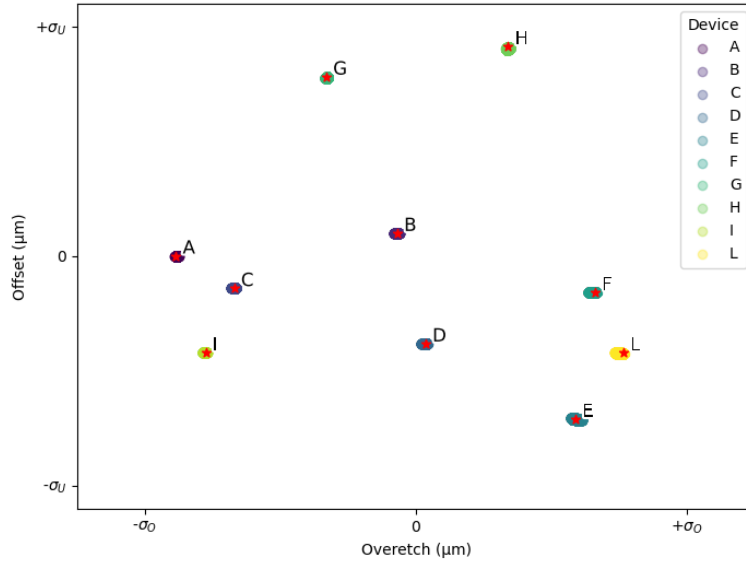


Figure 12: Final posterior distributions of the MCMC samples in the O-U domain, as obtained for ten of the tested devices.

5. Conclusion

This study provided a comprehensive exploration of the computational characterization and uncertainty quantification of Micro-Electro-Mechanical Systems (MEMS) capacitive accelerometers. By integrating supervised learning-based surrogate models, namely artificial neural networks, with traditional Markov Chain Monte Carlo (MCMC) methods, we established an effective characterization routine. This approach facilitated the reconstruction of device sensitivity directly from the time-dependent capacitance variance signal in response to a prescribed electric voltage.

- Utilizing high-fidelity finite element analysis (FEA) as training data, our surrogate models accurately predicted MEMS capacitive accelerometer behavior under a diverse array of fabrication parameters.
- Significant Reduction in Computational Costs: Our method achieved a remarkable decrease in computational demands without sacrificing result accuracy, enabling comprehensive and efficient stochastic analyses.

<i>Parameter</i>	$n_f [\mu g/\sqrt{Hz}]$	<i>True Value</i>	<i>Mean Value</i>	<i>95% C.I.</i>
Over-etch	100	0.0306	0.0306	[0.0304, 0.0308]
	1000	0.0306	0.0300	[0.0277, 0.0317]
	10000	0.0306	0.0245	[0.0131, 0.0363]
	100000	0.0306	0.0734	[-0.0285, 0.1757]
Offset	100	0.00608	0.00608	[0.00607, 0.00609]
	1000	0.00608	0.00602	[0.00593, 0.00612]
	10000	0.00608	0.00547	[0.00476, 0.00618]
	100000	0.00608	0.00856	[0.00152, 0.01630]
Thickness	100	30.515	30.513	[30.492, 30.533]
	1000	30.515	30.500	[30.218, 30.671]
	10000	30.515	30.341	[29.706, 30.961]
	100000	30.515	30.136	[29.165, 30.921]
S	100	1.0298	1.0279	[1.0287, 1.0302]
	1000	1.0298	1.0285	[1.0182, 1.0350]
	10000	1.0298	1.0204	[0.9978, 1.0450]
	100000	1.0298	1.0453	[0.9837, 1.1516]

Table 3: Posterior distribution predictions results with respect to different noise levels for device feature parameters and sensitivity.

- Precise Sensitivity Estimation: The models adeptly estimated accelerometer sensitivity, accounting for variables such as unknown mechanical offsets, over-etching, and thickness variations.

Future research directions could include broadening the spectrum of studied parameters, delving into various MEMS device types, and incorporating more sophisticated machine learning techniques for enhanced model robustness. Promising developments might involve enriching the surrogate model with more intricate multi-physics elements and augmenting the MCMC setup by employing multi-fidelity levels for sample generation.

References

- [1] A. Corigliano, R. Ardito, C. Comi, A. Frangi, A. Ghisi, and S. Mariani, *Mechanics of microsystems*. John Wiley & Sons, 2018.
- [2] J. W. Judy, “Microelectromechanical systems (MEMS): Fabrication, design and applications,” *Smart Materials and Structures*, vol. 10, no. 6, p. 1115, Nov. 2001, ISSN: 0964-1726. DOI: 10.1088/0964-1726/10/6/301. (visited on 12/11/2023).
- [3] G. A. Aydemir and A. Saranlı, “Characterization and calibration of MEMS inertial sensors for state and parameter estimation applications,” *Measurement*, vol. 45, no. 5, pp. 1210–1225, Jun. 2012, ISSN: 0263-2241. DOI: 10.1016/j.measurement.2012.01.015. (visited on 12/11/2023).
- [4] M. Glueck, D. Oshinubi, P. Schopp, and Y. Manoli, “Real-Time Autocalibration of MEMS Accelerometers,” *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 1, pp. 96–105, Jan. 2014, ISSN: 1557-9662. DOI: 10.1109/TIM.2013.2275240. (visited on 12/11/2023).
- [5] O. Dürr, P.-Y. Fan, and Z.-X. Yin, “Bayesian calibration of mems accelerometers,” *IEEE Sensors Journal*, 2023.
- [6] A. Harindranath and M. Arora, “A systematic review of user-conducted calibration methods for mems-based imus,” *Measurement*, p. 114001, 2023.
- [7] R. Mirzazadeh, S. Eftekhari Azam, and S. Mariani, “Mechanical Characterization of Polysilicon MEMS: A Hybrid TMCMC/POD-Kriging Approach,” *Sensors*, vol. 18, no. 4, p. 1243, Apr. 2018, ISSN: 1424-8220. DOI: 10.3390/s18041243. (visited on 11/14/2023).

- [8] Y. S. Hong, J. H. Lee, and S. H. Kim, “A laterally driven symmetric micro-resonator for gyroscopic applications,” *Journal of Micromechanics and Microengineering*, vol. 10, no. 3, pp. 452–458, Sep. 2000, ISSN: 0960-1317, 1361-6439. DOI: 10.1088/0960-1317/10/3/322. (visited on 02/12/2024).
- [9] A. Frangi, C. Cercignani, S. Mukherjee, and N. Aluru, *Advances in Multiphysics Simulation and Experimental Testing of Mems*. Imperial College Press, 2008. DOI: 10.1142/p522. eprint: <https://www.worldscientific.com/doi/pdf/10.1142/p522>. [Online]. Available: <https://www.worldscientific.com/doi/abs/10.1142/p522>.
- [10] A. Corigliano, R. Ardito, C. Comi, A. Ghisi, and S. Mariani, *Mechanics of Microsystems*. Feb. 2018, ISBN: 9781119053835. DOI: 10.1002/9781119053828.
- [11] N. Agarwal and N. R. Aluru, “A data-driven stochastic collocation approach for uncertainty quantification in MEMS,” *International Journal for Numerical Methods in Engineering*, vol. 83, no. 5, pp. 575–597, 2010, ISSN: 1097-0207. DOI: 10.1002/nme.2844. (visited on 12/11/2023).
- [12] J. P. Quesada-Molina and S. Mariani, “Uncertainty quantification at the microscale: A data-driven multi-scale approach,” *Engineering Proceedings*, vol. 27, no. 1, p. 38, 2022.
- [13] D. Allaire and K. Willcox, “Surrogate Modeling for Uncertainty Assessment with Application to Aviation Environmental System Models,” *Karen Willcox*, Aug. 2010, ISSN: 1533-385X. (visited on 12/11/2023).
- [14] M. Guo and J. S. Hesthaven, “Reduced order modeling for nonlinear structural analysis using Gaussian process regression,” *Computer Methods in Applied Mechanics and Engineering*, vol. 341, pp. 807–826, Nov. 2018, ISSN: 0045-7825. DOI: 10.1016/j.cma.2018.07.017. (visited on 11/14/2023).
- [15] L. Cicci, S. Fresca, M. Guo, A. Manzoni, and P. Zunino, “Uncertainty quantification for nonlinear solid mechanics using reduced order models with Gaussian process regression,” *Computers & Mathematics with Applications*, vol. 149, pp. 1–23, Nov. 2023, ISSN: 0898-1221. DOI: 10.1016/j.camwa.2023.08.016. (visited on 10/24/2023).
- [16] R. B. Gramacy, *Surrogates: Gaussian process modeling, design, and optimization for the applied sciences*. CRC press, 2020.
- [17] B. Ganapathysubramanian and N. Zabaras, “Modeling diffusion in random heterogeneous media: Data-driven models, stochastic collocation and the variational multiscale method,” *Journal of Computational Physics*, vol. 226, no. 1, pp. 326–353, Sep. 2007, ISSN: 0021-9991. DOI: 10.1016/j.jcp.2007.04.009. (visited on 12/11/2023).
- [18] P. Astrid, S. Weiland, K. Willcox, and T. Backx, “Missing Point Estimation in Models Described by Proper Orthogonal Decomposition,” *IEEE Transactions on Automatic Control*, vol. 53, no. 10, pp. 2237–2251, Nov. 2008, ISSN: 1558-2523. DOI: 10.1109/TAC.2008.2006102. (visited on 12/11/2023).
- [19] G. Kerschen, J. Golinval, A. Vakakis, and L. Bergman, “The method of proper orthogonal decomposition for dynamical characterization and order reduction of mechanical systems: An overview,” *Nonlinear Dynamics*, vol. 41, no. 1, pp. 147–169, 2005.
- [20] A. Quarteroni, A. Manzoni, and F. Negri, *Reduced Basis Methods for Partial Differential Equations: An Introduction* (UNITEXT). Springer International Publishing, 2015, ISBN: 978-3-319-15431-2.
- [21] A. Frangi and G. Gobat, “Reduced order modelling of the non-linear stiffness in MEMS resonators,” *International Journal of Non-Linear Mechanics*, vol. 116, pp. 211–218, Nov. 2019, ISSN: 0020-7462. DOI: 10.1016/j.ijnonlinmec.2019.07.002. (visited on 12/11/2023).
- [22] C. Touzé, “Normal form theory and nonlinear normal modes: Theoretical settings and applications,” in *Modal analysis of nonlinear mechanical systems*, Springer, 2014, pp. 75–160.
- [23] A. Corigliano, M. Dossi, and S. Mariani, “Domain decomposition and model order reduction methods applied to the simulation of multi-physics problems in MEMS,” *Computers & Structures*, vol. 122, pp. 113–127, 2013, ISSN: 0045-7949. DOI: 10.1016/j.compstruc.2012.12.012.

- [24] A. Radermacher and S. Reese, “POD-based model reduction with empirical interpolation applied to nonlinear elasticity,” *Int. J. Numer. Meth. Engrg.*, vol. 107, no. 6, pp. 477–495, 2016.
- [25] D. Bonomi, A. Manzoni, and A. Quarteroni, “A matrix deim technique for model reduction of nonlinear parametrized problems in cardiac mechanics,” *Comput. Methods Appl. Mech. Engrg.*, vol. 324, pp. 300–326, 2017.
- [26] G. Gobat, A. Opreni, S. Fresca, A. Manzoni, and A. Frangi, “Reduced order modeling of nonlinear microstructures through Proper Orthogonal Decomposition,” *Mechanical Systems and Signal Processing*, vol. 171, p. 108 864, May 2022, ISSN: 0888-3270. DOI: 10.1016/j.ymssp.2022.108864. (visited on 02/13/2024).
- [27] G. Kerschen, M. Peeters, J. Golinval, and A. Vakakis, “Nonlinear normal modes, Part I: A useful framework for the structural dynamicist,” *Mechanical Systems and Signal Processing*, vol. 23, no. 1, pp. 170–194, 2009, ISSN: 0888-3270. DOI: 10.1016/j.ymssp.2008.04.002.
- [28] L. Renson, G. Kerschen, and B. Cochelin, “Numerical computation of nonlinear normal modes in mechanical engineering,” *Journal of Sound and Vibration*, vol. 364, pp. 177–206, 2016, ISSN: 0022-460X. DOI: 10.1016/j.jsv.2015.09.033.
- [29] X. Cabré, E. Fontich, and R. de la Llave, “The parameterization method for invariant manifolds III: Overview and applications,” *Journal of Differential Equations*, vol. 218, no. 2, pp. 444–515, Nov. 2005, ISSN: 0022-0396. DOI: 10.1016/j.jde.2004.12.003. (visited on 02/09/2024).
- [30] A. Frangi, A. Colombo, A. Vizzaccaro, and C. Touzé, *Reduced order modelling of fully coupled electro-mechanical systems through invariant manifolds with applications to microstructures*, Dec. 2023. DOI: 10.48550/arXiv.2312.14803. arXiv: 2312.14803 [cs, math]. (visited on 02/09/2024).
- [31] S. Fresca and A. Manzoni, “POD-DL-ROM: Enhancing deep learning-based reduced order models for nonlinear parametrized PDEs by proper orthogonal decomposition,” *Computer Methods in Applied Mechanics and Engineering*, vol. 388, p. 114 181, Jan. 2022, ISSN: 0045-7825. DOI: 10.1016/j.cma.2021.114181. (visited on 12/11/2023).
- [32] L. Cai, L. Ren, Y. Wang, W. Xie, G. Zhu, and H. Gao, “Surrogate models based on machine learning methods for parameter estimation of left ventricular myocardium,” *Royal Society Open Science*, vol. 8, no. 1, p. 201 121, Jan. 2021. DOI: 10.1098/rsos.201121. (visited on 12/11/2023).
- [33] E. Haghighat, M. Raissi, A. Moure, H. Gomez, and R. Juanes, “A physics-informed deep learning framework for inversion and surrogate modeling in solid mechanics,” *Computer Methods in Applied Mechanics and Engineering*, vol. 379, p. 113 741, Jun. 2021, ISSN: 0045-7825. DOI: 10.1016/j.cma.2021.113741. (visited on 12/11/2023).
- [34] M. Torzoni, A. Manzoni, and S. Mariani, “A multi-fidelity surrogate model for structural health monitoring exploiting model order reduction and artificial neural networks,” *Mechanical Systems and Signal Processing*, vol. 197, p. 110 376, Aug. 2023, ISSN: 0888-3270. DOI: 10.1016/j.ymssp.2023.110376. (visited on 11/14/2023).
- [35] M. McGurk, A. Lye, L. Renson, and J. Yuan, *Data-Driven Bayesian Inference for Stochastic Model Identification of Nonlinear Aeroelastic Systems*. Oct. 2023. DOI: 10.21203/rs.3.rs-3408685/v1.
- [36] S. Fresca, G. Gobat, P. Fedeli, A. Frangi, and A. Manzoni, “Deep learning-based reduced order models for the real-time simulation of the nonlinear dynamics of microstructures,” *International Journal for Numerical Methods in Engineering*, vol. 123, no. 20, pp. 4749–4777, 2022.
- [37] G. Gobat, A. Baronchelli, S. Fresca, and A. Frangi, “Modelling the Periodic Response of Micro-Electromechanical Systems through Deep Learning-Based Approaches,” *Actuators*, vol. 12, no. 7, p. 278, Jul. 2023, ISSN: 2076-0825. DOI: 10.3390/act12070278. (visited on 11/21/2023).
- [38] G. Gobat, S. Fresca, A. Manzoni, and A. Frangi, “Reduced Order Modeling of Nonlinear Vibrating Multiphysics Microstructures with Deep Learning-Based Approaches,” *Sensors*, vol. 23, no. 6, p. 3001, Jan. 2023, ISSN: 1424-8220. DOI: 10.3390/s23063001. (visited on 02/12/2024).

- [39] S. L. Brunton, J. L. Proctor, and J. N. Kutz, “Discovering governing equations from data by sparse identification of nonlinear dynamical systems,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 113, no. 15, pp. 3932–3937, Apr. 2016, ISSN: 0027-8424. DOI: 10.1073/pnas.1517384113. (visited on 02/12/2024).
- [40] P. Conti, G. Gobat, S. Fresca, A. Manzoni, and A. Frangi, “Reduced order modeling of parametrized systems through autoencoders and SINDy approach: Continuation of periodic solutions,” *Computer Methods in Applied Mechanics and Engineering*, vol. 411, p. 116072, Jun. 2023, ISSN: 0045-7825. DOI: 10.1016/j.cma.2023.116072. (visited on 02/12/2024).
- [41] Y. Wang, S. Adams, J. Thorp, N. MacDonald, P. Hartwell, and F. Bertsch, “Chaos in MEMS, parameter estimation and its potential application,” *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 45, no. 10, pp. 1013–1020, Oct. 1998, ISSN: 1558-1268. DOI: 10.1109/81.728856. (visited on 12/11/2023).
- [42] C. J. Geyer, “Introduction to Markov Chain Monte Carlo,” in *Handbook of Markov Chain Monte Carlo*, Chapman and Hall/CRC, 2011, ISBN: 978-0-429-13850-8.
- [43] J. Kaipio and E. Somersalo, *Statistical and computational inverse problems* (Applied Mathematical Sciences). Springer Science+Business Media, Inc., 2005, vol. 160.
- [44] A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation*. Philadelphia, PA: SIAM, 2004.
- [45] Ching Jianye and Chen Yi-Chu, “Transitional Markov Chain Monte Carlo Method for Bayesian Model Updating, Model Class Selection, and Model Averaging,” *Journal of Engineering Mechanics*, vol. 133, no. 7, pp. 816–832, Jul. 2007. DOI: 10.1061/(ASCE)0733-9399(2007)133:7(816). (visited on 12/11/2023).
- [46] R. Neal, “MCMC Using Hamiltonian Dynamics,” in *Handbook of Markov Chain Monte Carlo*, S. Brooks, A. Gelman, G. Jones, and X.-L. Meng, Eds., vol. 20116022, Chapman and Hall/CRC, May 2011, ISBN: 978-1-4200-7941-8 978-1-4200-7942-5. DOI: 10.1201/b10905-6. (visited on 12/11/2023).
- [47] M. D. Hoffman and A. Gelman, “The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo,” *Journal of Machine Learning Research*, vol. 15, no. 47, pp. 1593–1623, 2014, ISSN: 1533-7928. (visited on 12/11/2023).
- [48] M. B. Giles, “Multilevel Monte Carlo methods,” *Acta Numerica*, vol. 24, pp. 259–328, May 2015, ISSN: 0962-4929, 1474-0508. DOI: 10.1017/S096249291500001X. (visited on 12/11/2023).
- [49] C. Soize and C. Farhat, “Probabilistic learning for modeling and quantifying model-form uncertainties in nonlinear computational mechanics,” *International Journal for Numerical methods in engineering*, vol. 117, no. 7, pp. 819–843, 2019.
- [50] M. Berger, C. Schott, and O. Paul, “Bayesian sensor calibration,” *IEEE Sensors Journal*, vol. 22, no. 20, pp. 19384–19399, 2022.
- [51] S. Lan, S. Li, and B. Shahbaba, “Scaling up bayesian uncertainty quantification for inverse problems using deep neural networks,” *SIAM/ASA Journal on Uncertainty Quantification*, vol. 10, no. 4, pp. 1684–1713, 2022. DOI: 10.1137/21M1439456. eprint: <https://doi.org/10.1137/21M1439456>. [Online]. Available: <https://doi.org/10.1137/21M1439456>.
- [52] J. J. Hopfield, “Neural networks and physical systems with emergent collective computational abilities,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 79, no. 8, pp. 2554–2558, Apr. 1982, ISSN: 0027-8424. (visited on 11/27/2023).
- [53] S. Timoshenko, *History of strength of materials: with a brief account of the history of theory of elasticity and theory of structures*. Courier Corporation, 1983.
- [54] H. B. Palmer, “The capacitance of a parallel-plate capacitor by the schwartz-christoffel transformation,” *Electrical Engineering*, vol. 56, pp. 363–368, 1937. [Online]. Available: <https://api.semanticscholar.org/CorpusID:51633055>.

- [55] X. Xiong, Y.-L. Wu, and W.-B. Jone, “A dual-mode built-in self-test technique for capacitive MEMS devices,” *IEEE Transactions on Instrumentation and Measurement*, vol. 54, no. 5, pp. 1739–1750, Oct. 2005, ISSN: 1557-9662. DOI: 10.1109/TIM.2005.855094. (visited on 11/23/2023).
- [56] A. Corigliano, M. Dossi, and S. Mariani, “Domain decomposition and model order reduction methods applied to the simulation of multi-physics problems in mems,” *Computers & Structures*, vol. 122, pp. 113–127, 2013.
- [57] J. S. Hesthaven and S. Ubbiali, “Non-intrusive reduced order modeling of nonlinear problems using neural networks,” *Journal of Computational Physics*, vol. 363, pp. 55–78, Jun. 2018, ISSN: 0021-9991. DOI: 10.1016/j.jcp.2018.02.037. (visited on 02/13/2024).
- [58] K. Lu *et al.*, “Review for order reduction based on proper orthogonal decomposition and outlooks of applications in mechanical systems,” *Mechanical Systems and Signal Processing*, vol. 123, pp. 264–297, 2019.
- [59] G. KERSCHEN and J. GOLINVAL, “Physical interpretation of the proper orthogonal modes using the singular value decomposition,” *Journal of Sound and Vibration*, vol. 249, no. 5, pp. 849–865, 2002, ISSN: 0022-460X. DOI: <https://doi.org/10.1006/jsvi.2001.3930>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0022460X01939306>.
- [60] R. C. Smith, *Uncertainty Quantification: Theory, Implementation, and Applications*. SIAM, 2013, vol. 12.
- [61] *CoventorMP*, <https://www.lamresearch.com/coventormp/>. (visited on 12/05/2023).
- [62] F. Chollet, *Keras: Deep Learning for humans*, <https://keras.io/>, 2015.
- [63] A. Vehtari, A. Gelman, D. Simpson, B. Carpenter, and P.-C. Bürkner, “Rank-normalization, folding, and localization: An improved \widehat{R} for assessing convergence of MCMC,” *Bayesian Analysis*, vol. 16, no. 2, Jun. 2021, ISSN: 1936-0975. DOI: 10.1214/20-BA1221. arXiv: 1903.08008 [stat]. (visited on 12/11/2023).
- [64] N. A. B. Riis *et al.*, *CUQIpy – Part I: Computational uncertainty quantification for inverse problems in Python*, May 2023. DOI: 10.48550/arXiv.2305.16949. arXiv: 2305.16949 [cs, math]. (visited on 12/11/2023).

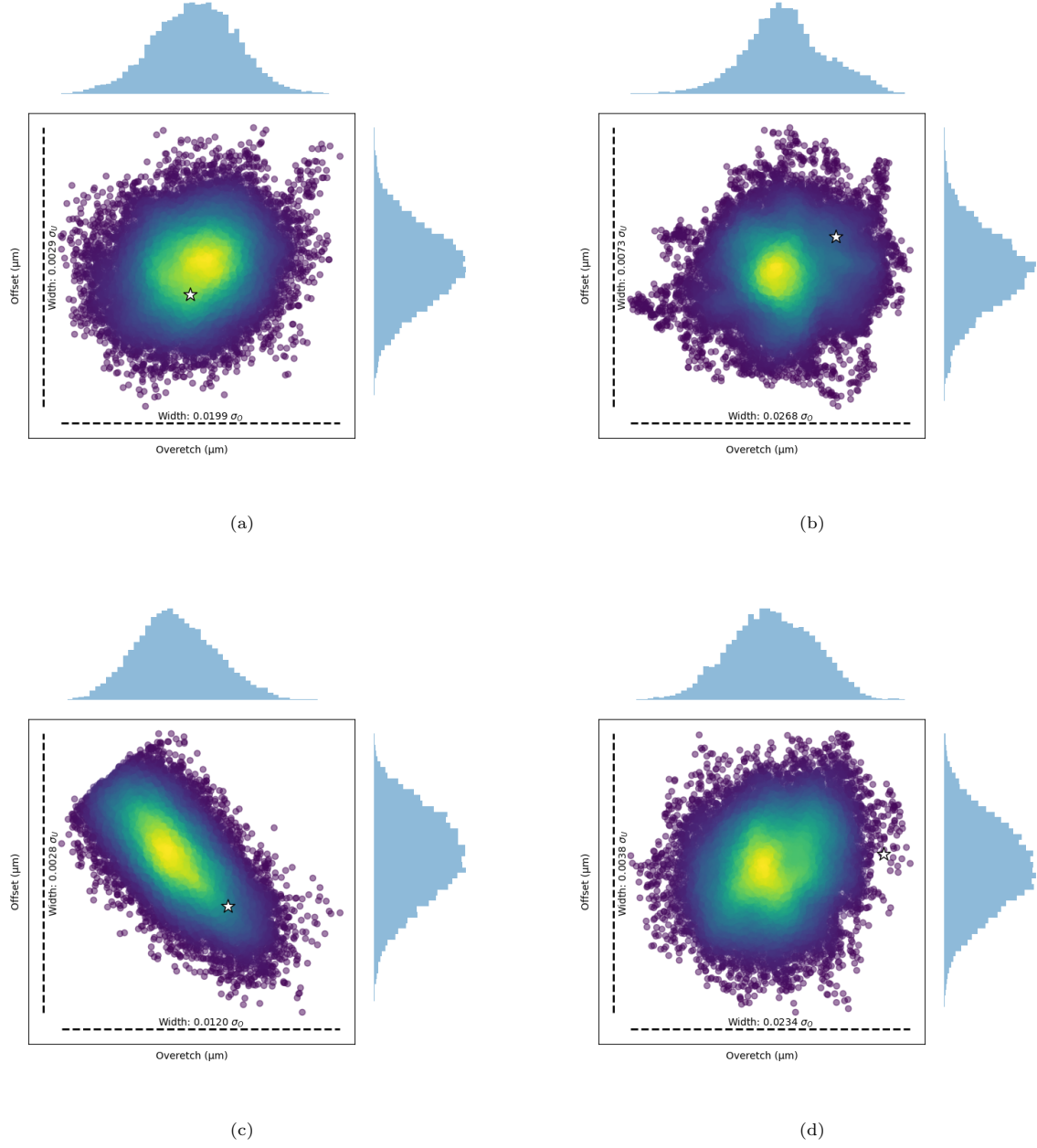


Figure 13: Scattering of the MCMC samples, and corresponding histograms (reported on the right and top sides of each plot) of the model parameter distributions for devices (a) A to (i) D.

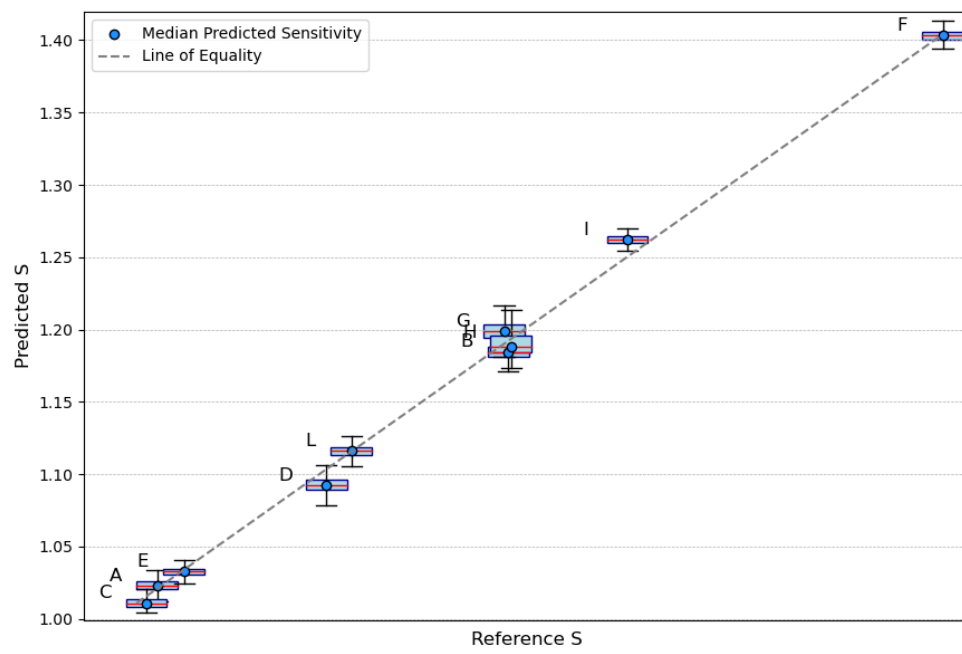


Figure 14: Posterior distribution of the sensitivities for ten of the tested devices.