International Journal of Solids and Structures 321 (2025) 113539

Contents lists available at ScienceDirect

International Journal of Solids and Structures

journal homepage: www.elsevier.com/locate/ijsolstr

Accounting for uncertainties in ML-based design of shape-morphing elements

Silvia Monchetti [a], Roberto Brighenti [a],*, Noy Cohen [b]

[a] Department of Civil and Environmental Engineering, University of Florence, Via di S. Marta 3, 50139 Firenze, Italy
[b] Department of Materials Science and Engineering, Technion - Israel Institute of Technology, Haifa 3200003, Israel

ARTICLE INFO

Keywords:
Uncertainties in material design
Bayesian approach
Artificial neural network
Inverse controlled shape change design
Confined swelling

ABSTRACT

Shape-morphing structures deform from one configuration to another in response to an external stimulus. In order to achieve a target shape, inverse design algorithms that enable one to compute the initial state of the system are required. Thanks to advances in 3D printing technologies, the realization of shape-morphing structures was demonstrated in a variety of recently published works. Commonly there are several sources of uncertainties that can influence the design. Examples include code inputs and outputs, model inadequacy, and the mechanical properties of 3D-printed materials. In this paper, we present an effective design of shape-morphing structures that accounts for these errors. We integrate a probabilistic approach to characterize model-form uncertainties in the inverse design of shape-morphing elements based on Machine Learning (ML) approach. The proposed approach relies on an Approximate Bayesian Computation (ABC) model where the parameter space is extended through the definition of the uncertainties involved in the process. To demonstrate the merit of this approach, we consider a system of a heterogeneous elastic tube embedding a gel core. The gel swells, and the swelling-induced forces lead to the deformation of the elastic tube, resulting in dilation and a change in the shape of the system. The proposed algorithm receives a target shape as input and determines the required spatial distribution of material properties in the heterogeneous ring. It is quantitatively shown how the system is sensitive to various sources of uncertainty: parameter uncertainty, model inadequacy, and observation errors. In particular, the effect of the parameter uncertainties has been investigated in terms of posterior distributions. In general, this work provides insight into the role of uncertainties in shape-controlled problems, and specifically, it allows for improving the reliability of the target shape inverse design in shape morphing elements.

1. Introduction

The inverse design problem of shape morphing systems is common in many fields, in particular when a prescribed target shape is provided and the properties required to achieve it are unknown. Shape change capabilities can be easily obtained from polymer-based materials. Nowadays, polymers are widely used in a variety of applications thanks to their low production cost, lightweight, broad range of physical-mechanical properties, controllable synthesis, etc (Caló and Khutoryanskiy, 2015; Revete et al., 2022; Chirani et al., 2015; Louf et al., 2021; Parodi et al., 2016; Hoffman, 2013). These constitute the ideal starting point for functional materials, in which the microstructure tends to host smart molecules or compounds with responsive characteristics (Brighenti and Cosma, 2020). However, other mechanisms, typically based on the deformation capabilities of their polymer network, can be exploited for such a purpose; relevant examples are represented by the nematic-isotropic phase transition in liquid crystal elastomers (LCE) (Brighenti et al., 2021), and by the controlled fluid uptake in polymers with a good affinity with specific fluids. Among the latter group, hydrogels are highly hydrophilic, *i.e.* they are capable of uptake a huge amount of fluid in their network without any chemical dissociation. The fluid absorbed in the polymer network is responsible for a large volume increase (swelling); the solid-fluid equilibrium condition is achieved once a chemo-mechanical balance is attained (Wu et al., 2004). Thanks to the large volume expansion taking place upon swelling, gels can be exploited to obtain shape change in gels-made elements (Li et al., 2018). Advanced applications exploit the specific responsiveness of some classes of gels to environmental stimuli (such as mechanical, chemo-mechanical, temperature, variations in pH, light, biological, etc.) (Brighenti and Cosma, 2020; Zhao et al., 2016; Herrmann et al., 2021; Brighenti and Cosma, 2022; Zu et al., 2022; Brighenti et al., 2023) that can be precisely spatiotemporally controlled. Further, coupling of gels with other elastic structures enables obtaining interacting forces that can be useful to induce controlled deformations (Cohen, 2019; Velders et al., 2017; Liu et al., 2021; Ravi

---

* Corresponding author.
E-mail address: roberto.brighenti@unifi.it (R. Brighenti).

https://doi.org/10.1016/j.ijsolstr.2025.113539
Received 4 February 2025; Received in revised form 19 June 2025; Accepted 21 June 2025

List of symbols

C	Cost function of the ANN
C_s, v_s	Solvent concentration, and its molar volume, respectively
$\bar{d}_{ANN}, \bar{d}_{mode}$	Mean distances quantifying the errors of the resulting shape with respect to the target one. The resulting shape is obtained by using the shear modulus provided by the ANN or the mode values with the ABC procedure, respectively
$\boldsymbol{F}, \boldsymbol{J}$	Deformation gradient tensor and its determinant, respectively
k_B, T	Boltzmann's constant and absolute temperature, respectively
$\boldsymbol{P}$	First Piola stress tensor
$\bar{\boldsymbol{r}}_{target}, \bar{\boldsymbol{r}}_{FEM}$	Dimensionless radius vector of the target shape and that provided by FEM analysis, respectively
$\boldsymbol{w}^{(i)}, \boldsymbol{b}^{(i)}$	Weight matrix and bias vector of the ANN, respectively
$\boldsymbol{x}, \boldsymbol{y}$	Input and output vector, respectively
Δ	Value of the prior distribution width
$\mu(\theta), \bar{\mu}(\theta)$	Shear modulus distribution of the elastic tube and its dimensionless counterpart, respectively
μ_c	Chemical potential
π	Hydrostatic pressure
θ	Angle identifying the position along the elastic tube perimeter
χ	Flory–Huggins interaction parameter
$\Psi, \Psi_{net}, \Psi_{mix}$	Total energy density, polymer network energy density, and mixing energy density, respectively
$\boldsymbol{\xi} = \mu_l$	Vector of the unknown quantities, represented by elastic tube shear modulus values at discrete angles, $(\theta = \theta_l, l = 1, \ldots, p)$

et al., 2002; Levin and Cohen, 2023; Plummer et al., 2024; Brighenti et al., 2018; Zhang et al., 2024b). It is worth noticing that coupling mechanisms can be exploited to produce energy out of dielectric elastomers (Hanuhov et al., 2024), by using piezoelectric polymers, or through the deformation of multilayer structures (Du et al., 2023).

The design process of swelling-induced shape changes in gels requires the development of theoretical or computational approaches capable of reproducing the involved physical phenomena. It is typically overcome by employing advanced finite-element routines or through analytical models. Recent works propose to employ ML-based methods to increase the efficiency of inverse design algorithms. The ML-based methods have gained large applications in the engineering field due to rapidly growing data, computing power, and the development of advanced algorithms (Dimiduk et al., 2018). An example is provided in Liu et al. (2021) where several engineering applications of ML approach based on the use of ANN models in composites constitutive modeling are reviewed, or in Argatov (2019) where the ANN is used as a modeling technique in tribology. Another valuable work (Xiao et al., 2023) provides a review of ANNs by focusing on forward and inverse problems and their applications in computational materials science.

The main computational cost associated with ML relies on the training of the algorithm, while afterward physically complex problems can be quickly solved (Bhuwal et al., 2023; Boley et al., 2019; Anitescu et al., 2023; Bai et al., 2023; Wang et al., 2025). From an applicative viewpoint, shape morphing systems can be 3D-printed. However, the properties of the available materials that can be printed typically depend and are limited depending on the 3D printer's capabilities and characteristics. In addition, variations between 3D-printed specimens can be present and, moreover, uncertainties in computer models need to be considered. In this context, quantifying the role of uncertainties in modeling gels-related parameters becomes crucial to provide reliable and accurate predictions. Even though the relevance of uncertainty quantification is widely recognized, its proper formalization is still an open issue. Some optimization methods have been developed by considering the presence of uncertainties; among them, RBDO (Reliability-Based Design Optimization) represents an interesting example (Hu et al., 2024; Dawei et al., 2021). However, these methods focuses on optimization design and do not allow for explicit modeling of epistemic and aleatory uncertainties. In this scenario, Bayesian inference is proposed as a credible solution for inverse problem by proving this efficiency for a wide range of applications (Madireddy et al., 2015; Wang et al., 2024; Pepi et al., 2020; Hamdia and Ghasemi, 2022), such as in additive manufacturing applications (Kim et al., 2022; Mahadevan et al., 2022), in the printed polymer process (Espino et al., 2023) and in the design of reinforced polymer composite (Thomas et al., 2022). This approach allows the updating of the prior knowledge of model parameters by considering the modeled sources of uncertainties, based on some observed or target data, through the proper definition of the likelihood function. The likelihood function represents the probability of the observed or target data vs the parameters of interest (the model's parameters). However, as underlined in Beaumont (2019), many of the employed models turn out to have intractable likelihood functions, and necessarily methods based on simulations, such as Approximate Bayesian Computation (ABC) (Rubin, 1984), must be employed. This likelihood-free approach requires only a *simulator*, *i.e.* a computer code (such as, in this case, a Finite Element Model) which takes in input the uncertain parameters and provides as outputs the simulated data. As a drawback, ABC is typically computationally intensive, and a surrogate model is usually needed (Monchetti et al., 2024c,b). The main advantage of performing the process of swelling-induced shape changes in gels in a Bayesian way is the availability of an accuracy assessment of the parameter estimates according to the modeled uncertainties. Moreover, we demonstrate that the prediction capabilities of ML-based inverse design tools can be improved through a Bayesian integrated approach. Indeed, the proposed framework allows for exploiting, on the basis of the defined uncertainties, a larger space of parameter variability. The posterior distribution represents the more reliable prediction of the random variable, and it includes the expert judgment, the defined uncertainties, and the observed or target data.

In this work, we derive an ML-Bayesian-based integrated approach to account for the uncertainties associated with the fabrication of systems and with the adopted modeling tools. As an example, we focus on the swelling-induced controlled shape change of confined gels. In particular, we consider the role of uncertainties of mechanical parameters on the elastic structure-swollen gels coupled response from a quantitative viewpoint and we extend the approach in Monchetti et al. (2024a) by integrating a probabilistic-based strategy in order to improve the predictive capability of the inverse design of shape-morphing elements. This topic is greatly relevant to the precise design of shape change in swelling-driven deformation, where a target shape is required to be achieved by the system despite the dispersion of the mechanical parameter values of the involved material.

The manuscript is organized as follows: Section 2 surveys the Machine Learning-based (ML) approaches in programming shape changes in deformable solids, and outlines the role of uncertainties of the involved parameters in this context. Section 3 illustrates the proposed ANN surrogate method for analyzing the shape change in swollen gels and the effect of mechanical parameters uncertainties on the obtained shape. Section 4 presents several numerical results related to swelling-driven shape change of elastic structures, where the Approximate Bayesian Computation (ABC) model is used to improve the reliability of the solution, and Section 5 concludes the paper.

2. Programming shape change in materials

Designing the shape change in materials can be obtained by properly tuning their physical properties. This inverse-design problem is common in many applications and engineering fields and it can be solved by employing ML, which requires a reasonable computational cost to achieve a solution after properly training the algorithm (Liu et al., 2024). It is pointed out that the training process is computationally expensive, and the number of training sets strongly affects the predictive capability of the algorithm. Moreover, we observe that the application of ML to solve inverse-design problems provides a deterministic solution. Therefore, the network architecture or a limited number of training samples can affect the outcome of the ML algorithm (Quek et al., 2025). In the classical approaches, no uncertainties associated with the design process are considered in the algorithm. Integrating the ML with the Bayesian inference algorithm increases the reliability of the final result, as it captures both the most probable solution and the associated uncertainties. This approach provides a twofold advantage: *(i)* it enables classifying the various sources of uncertainties and *(ii)* it allows defining the probability distributions of the model inputs and outputs to explore a wider parameter space and obtain a more reliable design tool.

In this section, we summarize the main features of the selected ML approach as well as the adopted uncertainty quantification process.

2.1. Machine learning-based design

In the following, we employ an ML approach based on the use of ANN models (Chakraverty and Jeswal, 2021; Xiao et al., 2023; Liu et al., 2021; Argatov, 2019). Thanks to its ability to learn complex patterns and relationships, ANN represents a powerful approach widely used for the solution of a variety of problems (Wright et al., 2022; Park et al., 2023; Pogorelko et al., 2024).

The typical structure of an ANN is made of interconnected nodes (artificial neurons) arranged in layers through which the information flows, starting from the input layer up to the final (output) one. Mathematically, the transformation of the information in the ith layer is written as

$$\begin{aligned} &\boldsymbol{y}^{(N)} = \sigma_N(\boldsymbol{z}^{(N)}) \\ &\boldsymbol{z}^{(i)} = \boldsymbol{w}^{(i)}\sigma_{i-1}(\boldsymbol{z}^{(i-1)}) + \boldsymbol{b}^{(i)}, (i = N-1, N-2, \dots, 2) \\ &\boldsymbol{z}^{(1)} = \boldsymbol{w}^{(1)}\boldsymbol{x} + \boldsymbol{b}^{(1)} \end{aligned} \tag{1}$$

where $\boldsymbol{w}^{(i)}$ and $\boldsymbol{b}^{(i)}$ indicate the weight matrix and the bias vector, respectively, required to transform the information from layer $i-1$ to layer i and σ is the so-called *activation function* accounting for the nonlinearities of the problem. The ANN algorithm requires an initial data set, i.e. a vector of values of the design parameters $\boldsymbol{x} = [x_1, x_2, \dots, x_m]^T$ (used to define the initial vector $\boldsymbol{z}^{(1)}$) and the corresponding target vector $\boldsymbol{y} = [y_1, y_2, \dots, y_p]^T$ to be compared through a cost function C with the output vector provided by the network, $\boldsymbol{y}^{(N)}$. Throughout the paper, we indicate the input vector as $\boldsymbol{x}$ and the final output vector as $\boldsymbol{y}^{(\boldsymbol{N})}$. By minimizing the cost function $C : R^n \rightarrow R$, quantifying the performance of the ANN

$$\boldsymbol{w}^{(i)}, \boldsymbol{b}^{(i)} = \underset{\boldsymbol{w},\boldsymbol{b}}{\operatorname{argmin}}\, C(y^N(\boldsymbol{x}, \boldsymbol{w}, \boldsymbol{b})). \tag{2}$$

Through a back-propagation algorithm based on the evaluation of the cost function gradient with respect to the weights $dC/dw^{(i)}_{jk}$ and the biases $dC/db^{(i)}_j$, the ANN determines the matrix of the weights $\boldsymbol{w}_i$ and the vector of the bias $\boldsymbol{b}_i$. This leads to an increasingly good approximation of the output, i.e. $y^{(N)}_l \approx y_l, l = 1, \dots, p$. Finally, the weights and the biases of the ANN are updated using the standard gradient descent algorithm

$$\begin{aligned} \Delta w^{(i)}_{kj} &= -\eta \frac{dC}{dw^{(i)}_{jk}} = -\eta \delta^{(i)}_j y^{(i-1)}_k, \\ \Delta b^{(i)}_j &= -\eta \frac{dC}{db^{(i)}_j} = -\eta \delta^i_j, \end{aligned} \tag{3}$$

where $\eta > 0$ is the learning rate.

The ANN is capable of dynamically adjusting the weights and biases through an iterative calculation involving: *(i)* a forward process aimed at propagating and transforming data along the layers; *(ii)* an error evaluation through the assessment of the cost function; *(iii)* a back-propagation for adjusting the weights and biases to reduce the overall error. It is worth mentioning that among the initial dataset of design vectors, a certain percentage (here 80% of the initial dataset of design vectors) is used for training the ANN, while the remaining elements of the initial design vectors dataset are used for the ANN validation (Anitescu et al., 2023).

2.2. Uncertainty quantification through the Bayesian-based approach

In common engineering design practice, model parameters are typically characterized by uncertainties (scarce or poor-quality experimental data, inaccuracy of measurements, difficulties in collecting data, etc.) which can affect the design reliability; thus uncertainty quantification must be accounted for to maximize the reliability and quality of designs. A well-known approach in the literature for handling uncertainties (Gray et al., 2022) is represented by the Reliability-Based Design Optimization approach (RBDO). This method is not deterministic and aims at providing the optimum design parameters setting by making use of extreme values and suitable safety factors for handling uncertainties. This offers a more realistic and robust approach (Hu et al., 2024; Dawei et al., 2021). It is worth recalling that the ABC approach exploits the statistical inference, rather than using an optimization algorithm as in RBDO, and is more flexible in handling any uncertainty involved in the problem, and allows for the adoption of any probability distribution. The programming of shape changes in structures can be considered an inferential problem in which we want to learn the features of the physical problem in order to ultimately achieve a target shape. The problem has a statistical nature when the relationship between the unknown quantities ξ and the observed or target data y_0 includes sources of uncertainty. To learn ξ starting from y_0, at least two different approaches exist: *(i)* the frequentist approach and *(ii)* the Bayesian approach. The frequentist approach neglects any uncertainties or randomness of the system, and ξ is considered as a vector of parameters, *i.e.* fixed but unknown quantities. The Bayesian approach, instead, aims at deriving a probability distribution for ξ in order to formalize and quantify the uncertainty. It follows that ξ is considered a random vector. This latter approach is herein adopted by using the formalization of the uncertainty quantification in model predictions proposed in Berger and Smith (2019), Kennedy and O'Hagan (2001).

Let us consider a mechanical model M representing a physical phenomenon. The real process is represented by the function $y^{(R)}(\xi)$, which maps the mechanical properties of interest (collected in the vector ξ) to some observable quantities, $y_0(\xi)$. In order to model and compute the response of the system, a simulator (such as a FE framework) is required. The simulator receives as input the vector of variables ξ, representing the object of the inference, and provides as outputs the prediction of the real process, $y^{(M)}(\xi)$. The discrepancy between the model output and the real process is termed *model bias*, $b(\xi)$, and describes the inadequacy of the model. Consequently, $b(\xi) = y^{(R)}(\xi) - y^{(M)}(\xi)$. Moreover, the observation, $y_0(\xi)$, contains an experimental measurement error e which must be considered. Accordingly, $y_0 = y^{(R)}(\xi) + e$. To sum up, we can write the following equation:

$$y_0 = y^{(R)}(\xi) + e = y^{(M)}(\xi) + b(\xi) + e. \tag{4}$$

Bayesian inference aims to define the posterior probability $p(\xi|y_0)$ representing the distribution of the unknown parameters (ξ) given the observed or target data (y_0), by using the well-known Bayes' theorem (Bayes, 1763).

$$p(\xi \mid y_0) = \frac{p(\xi)p(y_0 \mid \xi)}{\int_\xi p(\xi)p(y_0 \mid \xi)d\xi} \tag{5}$$

where $p(\xi)$ is the prior distribution, $p(y_0 \mid \xi)$ is the likelihood function, and the denominator is known as the marginal likelihood that acts as a normalizing constant. In the following, we point out a specific framework where the uncertainty quantification is integrated into the design of controlled shape changes in gels. We demonstrate that this can improve the predictive capability of the ANN algorithm proposed in Monchetti et al. (2024a). Thus, this paper represents a sequel to such research and shows how it can be easily extended to other engineering applications.

3. Swelling-driven shape changes in gels

Controlled shape change is a relevant example of smart responsiveness of solids and structures whose final shape needs to be precisely controlled to fulfill specific design requirements. Among several possibilities offered by materials (Oliver et al., 2016), shape change can be easily obtained by harnessing the large volume increase experienced by gels in the presence of a suitable solvent fluid, taking place in free or constrained conditions (Marcombe et al., 2010). Controlled swelling can be achieved by properly designing the gel's multiphase network microstructure (Zhang et al., 2024a), or, more easily, by exploiting a properly designed confining elastic structure interacting with the gel (Monchetti et al., 2024a). Despite the simplicity of the considered example, uncertainties can strongly affect the desired morphing design.

In this section, we consider the controlled swelling of gels under heterogeneous elastic constraints. A soft, heterogeneous elastic tube filled with gel is supposed to represent a specific constraint of the system. We show that the swelling-induced forces can deform the initially cylindrical tube into a non-circular one if its elastic modulus is non-uniform along its perimeter, Fig. 1. Note that due to the heterogeneity of the ring, it is difficult to obtain analytical solutions.

The physical problem of a swollen gel interacting with an elastic element requires fulfilling the kinematic constraint and the stress equilibrium at the interface, *i.e.* the normal displacement impenetrability condition and the stress continuity:

$$\begin{aligned} &\left[\boldsymbol{u}^{+}(\boldsymbol{X},t)-\boldsymbol{u}^{-}(\boldsymbol{X},t)\right]\cdot\boldsymbol{n}(\boldsymbol{X},t)\leq 0 \quad \text{for } \boldsymbol{X}\in\Gamma, t\in(0,t_f)\\ &\left[\boldsymbol{\sigma}^{+}(\boldsymbol{X},t)-\boldsymbol{\sigma}^{-}(\boldsymbol{X},t)\right]\cdot\boldsymbol{n}(\boldsymbol{X},t)=\boldsymbol{0} \quad \text{for } \boldsymbol{X}\in\Gamma, t\in(0,t_f) \end{aligned} \tag{6}$$

where the superscripts $'+'$ and $'-'$ indicate the quantities related to the gel and elastic layer interface Γ, respectively, $\boldsymbol{u}$ is the displacement vector, $\boldsymbol{\sigma}$ is the stress tensor, $(0,t_f)$ is the time interval where the swelling phenomenon occurs, and $\boldsymbol{n}$ is the unit outward normal to the interface. It must be recalled that if the gel is just in contact without sticking to the elastic structure, it can exert only a compressive normal stress $p(\theta) = (\boldsymbol{\sigma}\boldsymbol{n})\cdot\boldsymbol{n}$ on the elastic tube, i.e. no traction is allowed to occur at the interface, thus enabling the creation of convex or nearly convex morphed shapes only. The controlled swelling problem is firstly solved by using an ANN trained with FEM results and, secondly, by integrating the uncertainty quantification framework. The numerical results shown in the next section will demonstrate the benefits of this latter approach. The mathematical formulation of the mechanics of gel swelling is reported in [Appendix].

A qualitative demonstration of the shape control of hydrogels obtainable by harnessing an elastic constraint, is shown in Fig. 2; the surrounding initially circular elastic layer has been obtained through 3D printing of a thermoplastic polyurethane polymer (TPU) by using the fused deposition modeling (FDM) technology. The variation of the elastic bending stiffness of the elastic TPU tube has been obtained by properly varying the layer thickness along the perimeter of the cylindrical element. The elastic layer has been filled with an unswollen hydrogel element, which, upon swelling, interacts with the layer, leading to the final shapes shown in the figure.

3.1. ANN used in shape change inverse design or as a surrogate model in swelling of confined gels

The ANN developed for solving the problem under study is trained using the results provided by finite element simulations of a confined swelling gel embedded in an elastic layer with various stiffness distributions. Once properly trained, an ANN is suitable to be used as an inverse design tool or a surrogate model of the physical problem in turn, according to the way the ANN is trained (see Fig. 3).

In particular, in inverse design, we train the ANN by providing as the input vector $x_h = \bar{r}_h$, where $\bar{r}_h$ represents the dimensionless values of the radius at $h = 1,\dots,m$ discrete points used to describe the shape of the external boundary of the elastic layer surrounding the gel. The output vector $y_l^{(N)} = \bar{\mu}_l^{(N)}$ denotes the dimensionless shear modulus at positions $l = 1,\dots,p$ along the perimeter of the elastic ring (Fig. 3a). In this case, the loss function providing the error of the ANN prediction is defined as

$$\ell_{inv} = \text{mean}(\boldsymbol{L}_{inv})\ , \text{ with } \boldsymbol{L}_{inv} = (\ell_{inv,1}, \ell_{inv,2}, \dots \ell_{inv,h}, \dots, \ell_{inv,N_T}) \\ \text{and } \ell_{inv,j} = \sum_{i=1}^{p}(\bar{\mu}_{ANN(i)} - \bar{\mu}_{FEM(i)})^2 \tag{7}$$

where N_T is the number of training cases used for the ANN, while p indicates the angular positions at which the difference $\bar{\mu}_{ANN} - \bar{\mu}_{FEM}$ is evaluated.

On the other hand, the ANN used as a surrogate model needs to be trained by providing as input vector $x_h = \bar{\mu}_h$, where $\bar{\mu}_h$ are the dimensionless values of the shear modulus at $h = 1,\dots,p$ discrete points along the external boundary of the elastic layer. The output vector $y_l^{(N)} = \bar{r}_l^{(N)}$ denotes the dimensionless radius at positions $l = 1,\dots,m$ along the perimeter of the elastic ring describing its final shape (Fig. 3b). In this case, the loss function of the ANN is defined as

$$\ell_{sur} = \text{mean}(\boldsymbol{L}_{sur})\ , \text{ with } \boldsymbol{L}_{sur} = (\ell_{sur,1}, \ell_{sur,2}, \dots \ell_{sur,j}, \dots, \ell_{sur,N_T}) \\ \text{and } \ell_{sur,j} = \sum_{i=1}^{m}(\bar{r}_{ANN(i)} - \bar{r}_{FEM(i)})^2 \tag{8}$$

where N_T is the number of training case, while m indicates the angular positions at which the difference $\bar{r}_{ANN} - \bar{r}_{FEM}$ is evaluated.

In the present study, we adopt an ANN characterized by $N = 6$ layers with $q = 72$ neurons each, while $m = q = p = 72$. As for the activation function σ, we assume the well-known $ReLU$ (rectified linear unit, $ReLU(z_i) = \max(0, z_i)$, (Chakraverty and Jeswal, 2021; Dubey et al., 2022)) for the layers $1, 2, \dots, N-1$, while the sigmoid function $(\sigma(z_i) = (1-e^{z_i})^{-1})$ is used for the last (Nth) layer of the ANN; due to its simplicity, efficiency, and speed of training, $ReLU$ is typically adopted for hidden layers, while the output layer's activation function depends on the prediction goal of the ANN, and – among others – the sigmoid one is often used (Dubey et al., 2022).

3.2. Effect of uncertainty in shape change of confined swelling

The uncertainty quantification in a swelling-induced controlled shape change of confined gels is herein formalized by exploiting the information-driven by the required target shape. The target shape is discretized through the vector $\boldsymbol{y_0} = \bar{\boldsymbol{r}}_{target} = \bar{r}_h$, $(h = 1,\dots,m)$, to infer the spatial stiffness distribution of the elastic ring which encapsulates the gel, $\boldsymbol{\xi} = \bar{\boldsymbol{\mu}} = \mu_l$, $(l = 1,\dots,p)$, *i.e.* the spatial distribution of the shear modulus. The input and the output vectors are specified along the perimeter of the elastic ring over the θ angle, $\theta \in [2\pi/q - 2\pi]$ where q represents the number of discretization steps along the perimeter. By following the formalization of Rubin (1984) in Bayesian inference, the known quantities can be treated as observed values of a random variable while the unknown quantities are unobserved random variables. In this perspective, the posterior distribution (Eq. (5)) is treated as the probability of $\boldsymbol{\xi}$ under the conditioning to a specific target $\boldsymbol{y_0}$ over the space of the plausible values of unknowns (represented by the prior

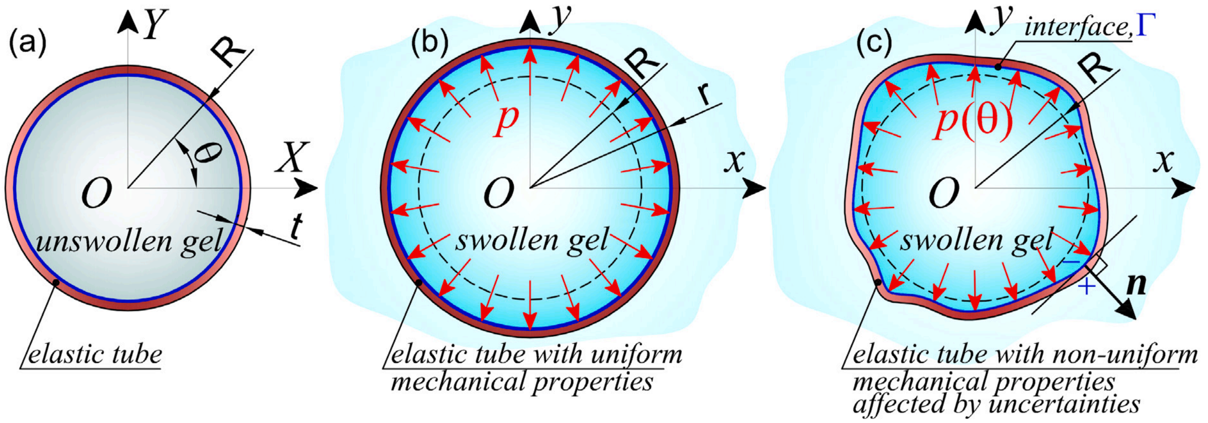

Fig. 1. Gel-filled elastic tube: unswollen state (a) and swollen states for a homogeneous elastic tube (b) and for an elastic tube with non-uniform mechanical properties (c). The deformation of the elastic tube is driven by the pressure $p(\theta)$ exerted by the filling swollen gel.

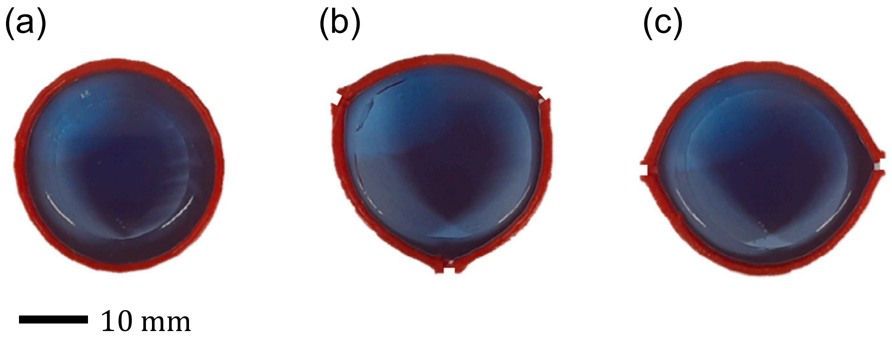

Fig. 2. Examples of experimentally obtained shapes of a gel-filled elastic tube (outer red layer) upon complete swelling of the inner gel (blue core region): uniform elastic layer providing a circular shape (a), a triangular-like shape (b), and an eye-like shape (c). The elastic layer has been obtained by 3D printing of a TPU thermoplastic polymer.

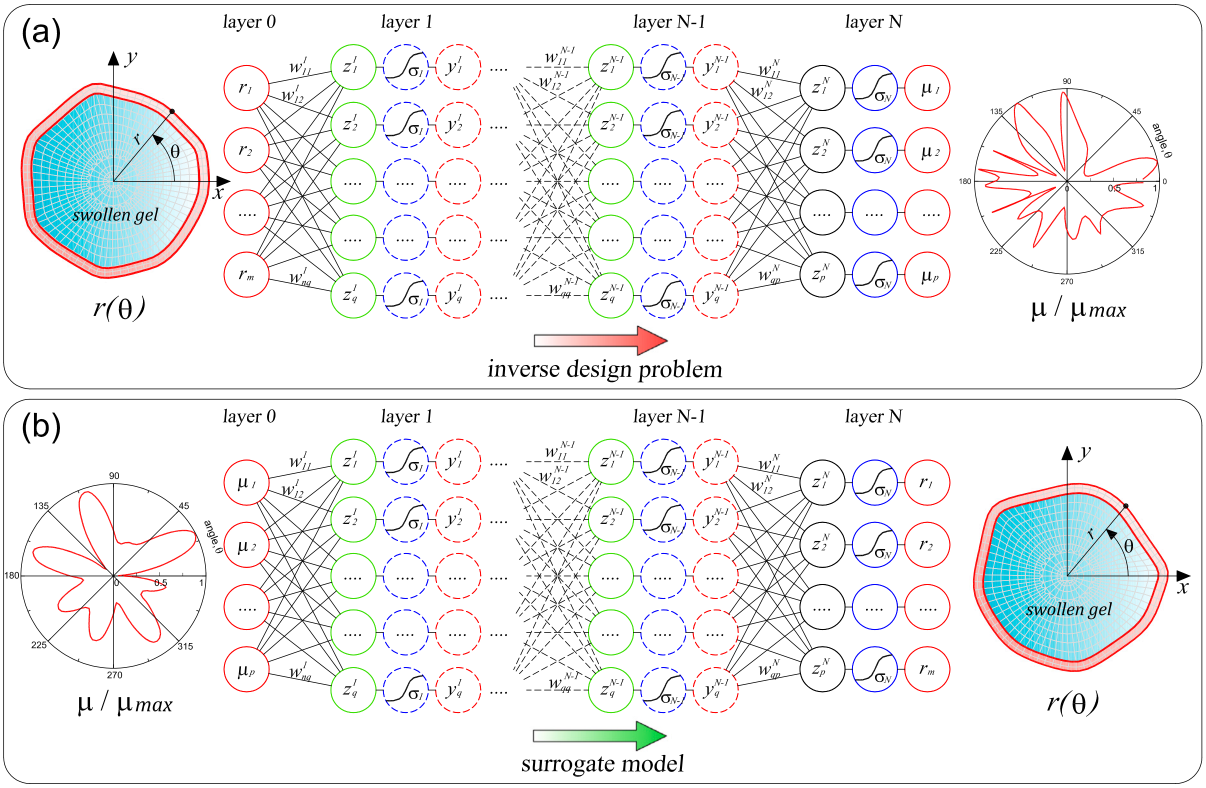

Fig. 3. Schematic of the neural network algorithm used for solving (a) the inverse design problem (given a target swollen shape, find the shear modulus distribution) or (b) the surrogate model (given the shear modulus distribution, find the corresponding swollen shape).

distribution of ξ). Note that, $\boldsymbol{y_0}$ represents the desired target shape, no observation data are required. The uncertainty quantification process can be summarized as follows. The selected target shape represents the input of the ANN used in the inverse design, $\bar{\boldsymbol{r}}_{target}$. The output of the inverse design tool is the shear modulus along the perimeter of the elastic ring, $\bar{\boldsymbol{\mu}}_{ANN}$. This result represents only an approximation of the effective, but unknown, spatial distribution of the shear modulus ($\bar{\boldsymbol{\mu}}_{target}$) along the perimeter of the elastic ring which exactly reproduces the selected target shape. Consequently,

$$\begin{aligned} &\boldsymbol{y}^{(R)}(\xi) = \bar{\boldsymbol{r}}_{target}(\bar{\boldsymbol{\mu}}_{target}) \text{ and} \\ &\boldsymbol{y}^{(M)}(\xi) = \bar{\boldsymbol{r}}_{FEM}(\bar{\boldsymbol{\mu}}_{ANN}). \end{aligned} \tag{9}$$

The previously introduced bias term represents the discrepancy between the target shape and the FE output (Fig. 4). As an example, in real cases, where the elastic ring is supposed to be obtained by AM procedure, $\bar{\boldsymbol{\mu}}_{ANN}$ represents the optimal spatial distribution of the shear modulus, which unavoidably will differ from the result of the AM process. This discrepancy can be modeled through the experimental measured error, e. Note that, this error will reflect the specific performance of a certain printer in reproducing the shear modulus distribution along the perimeter of the elastic ring. All these uncertainties are responsible for discrepancies with respect to the programmed target shape.

The Bayesian inference improves on ANN-approach in two respects. First, the predictions can allow for any sources of uncertainty. Second, we can quantitatively assess the effects of uncertainties on the matching between the target and the actual obtained shape. However, a critical concern can be related to evaluating the posterior distribution (Eq. (5)), which often involves methods based on simulations. Indeed, it requires numerical approximations to overcome the resolution of integrals or summations over high dimensional spaces when the physical problem includes several uncertainties.

To overcome this issue, a likelihood-free method is used in this research (see, Fig. 5). In particular, among the approaches that allow the evaluation of the posterior distribution overcoming the exact likelihood computation, the ABC approach is integrated. It consists of replacing the evaluation of the likelihood function with a Monte Carlo (MC) estimate (Rubin, 1984). The basic ABC algorithm follows a rejection scheme made of the steps displayed in Fig. 6 and requires the availability of a numerical model (*i.e.*, ANN surrogate tool) that takes the parameters object of the inference as input ($\bar{\boldsymbol{\mu}}^{(n)}$), performs stochastic calculations and outputs the pseudo data ($\bar{\boldsymbol{r}}_n^*$). The algorithm is based on a positive tolerance threshold ε and a distance function $d(\bar{\boldsymbol{r}}_{\boldsymbol{target}}; \bar{\boldsymbol{r}}^*)$.

The input vector $\bar{\boldsymbol{\mu}}^{(n)}$ is sampled from the selected prior distribution. The emulator, *i.e.* the ANN tool of the surrogate model, computes the swollen shape for each $s \in 1, \ldots, S$ sampled vector of $\bar{\mu}$, where $S = 10^5$. The output of the emulator combined with the error function represents the *generative* model $\bar{\boldsymbol{r}}^* = \bar{\boldsymbol{r}}_{ANN}(\bar{\boldsymbol{\mu}}^{(n)}) + f(\bar{\boldsymbol{\mu}}^{(n)})$. The error function represents the measurement errors, e, and could be correlate with the difference between $\bar{\mu}_l$ and the nearby values $\bar{\mu}_{l-1}, \bar{\mu}_{l+1}$ to simulate the capability of the printer to reproduce specific peaks in the spatial distribution of the shear modulus of the 3D-printed materials. The ABC algorithm retains the $\bar{\boldsymbol{r}}^*$ vector such that $d(r_{target}; r^*) \le \varepsilon$. The distance d is the Euclidean norm evaluated as follows:

$$d(\boldsymbol{r}_{target}; \boldsymbol{r}^*) = \sqrt{(\boldsymbol{r}_{target,1} - \boldsymbol{r}_1^*)^2 + (\boldsymbol{r}_{target,2} - \boldsymbol{r}_2^*)^2 + \cdots + (\boldsymbol{r}_{target,m} - \boldsymbol{r}_m^*)^2} \tag{10}$$

The tolerance threshold ε is a positive value equal to the $\alpha - th$ quantile of the distance d; in the following, we adopt $\alpha = 0.001$.

Note that, the vector $\xi = \mu_l$ $(l = 1, \ldots, p)$ contains the shear modulus values at l discrete points located on the external elastic layer. For each value of the angular position θ, a prior distribution of the shear modulus values is assumed; it reflects the range of plausible values of ξ intended as constraints (Fig. 7). Non-informative and informative priors can be

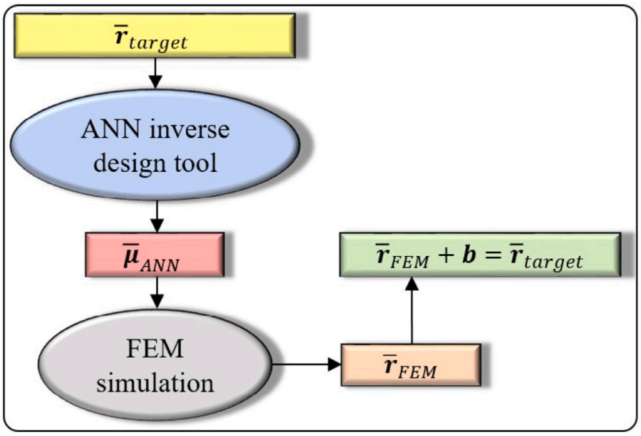

Fig. 4. Flow chart of the model inadequacy: the selected target shape ($\bar{r}_{target}$) is the input of the ANN inverse design tool, whose output is the shear modulus along the perimeter of the elastic ring ($\bar{\mu}_{ANN}$) to be used as input in FEM simulation to reproduce the desired shape. Some discrepancies, described by the bias term ($\boldsymbol{b}$) quantifying the inadequacy of the model, are unavoidably expected.

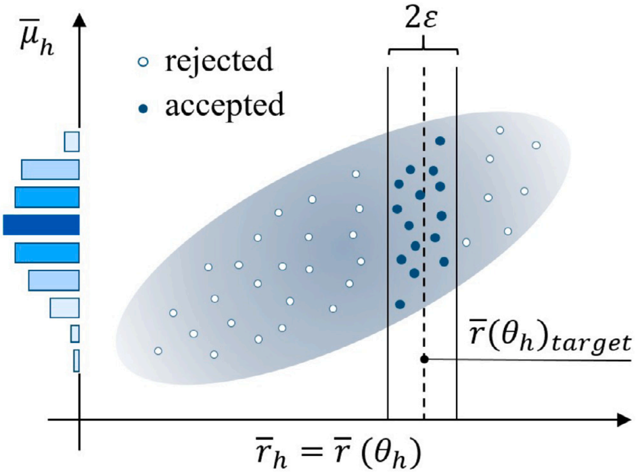

Fig. 5. Schematic of the ABC approach: a set of parameter points is sampled from the prior distribution ($\bar{\mu}_h$); given the sampled parameter points, a data set is simulated ($\bar{r}_h$). When the generated data are close to the target ones, the sampled parameters are accepted (when a given tolerance ε is satisfied), otherwise they are discarded (Beaumont, 2019).

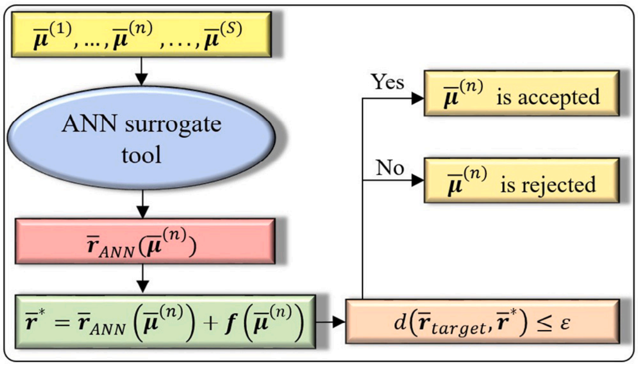

Fig. 6. Flow chart of the rejection strategy in the ABC approach: the shear modulus values ($\bar{\mu}^{(S)}$) are sampled from the prior distribution, and are used as the input of the ANN surrogate tool, which reproduces the swollen shape. The output of the ANN tool is combined with the error function to obtain the generative model, which computes the radius $\bar{r}^*$. The ABC algorithm accepts the shear modulus values which lead to a discrepancy – with respect to the desired target shape – lower than a positive tolerance threshold ε.

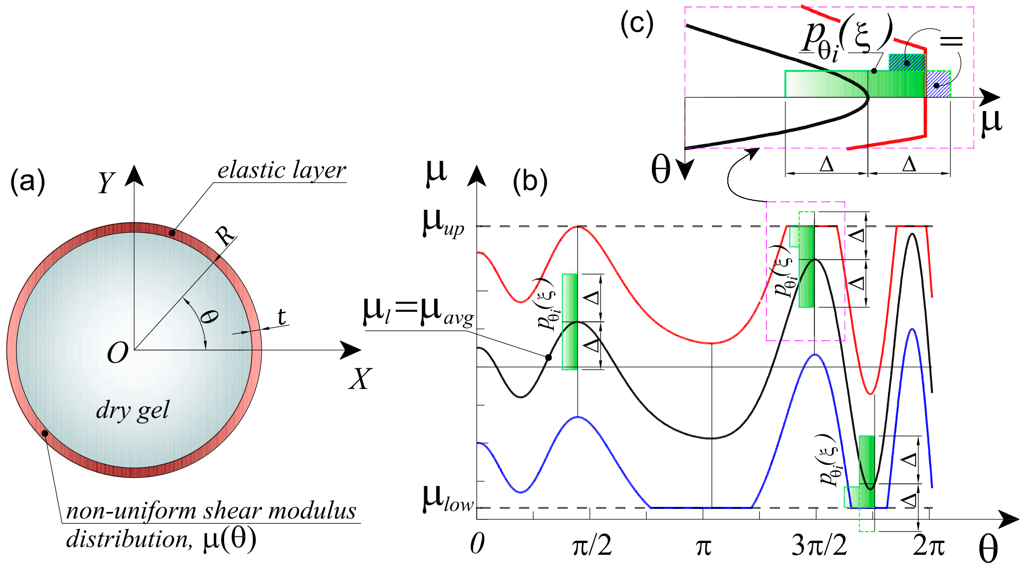

Fig. 7. Elastic tube filled with a dry gel (a). Schematic of the prior distributions of the shear modulus assumed at each angular position of the elastic ring with the reflecting barrier approach. The lower and upper limits allowed for the shear modulus values (b). The detail illustrates the meaning of the reflecting barrier concept (c).

selected to represent the initial level of knowledge about the values of the uncertain model parameters. When a non-informative prior is defined, its contribution is negligible with respect to the observed or target data. When a so-called informative prior is defined, information and subjective beliefs are included about the model parameters. This aspect, on the one hand, introduces a source of subjectivity in the process; nevertheless, it can improve the optimization process. As an example of informative prior, we have considered the output of the ANN used in inverse design given a specific target shape, $\bar{\boldsymbol{\mu}}_{ANN}$, as the mean vector values of the prior distribution of ξ. When a uniform distribution $\mathcal{U}_{[a_l,b_l]}$ is selected, considering $a_l = \mu_{ANN,l} - \Delta$, and $b_l = \mu_{ANN,l} + \Delta$ (Eq. (11)), a Multivariate Reflecting Barrier strategy (Hoff, 2009) can be used to avoid having sampled shear modulus values, μ^*, falling outside the admissible range defined by the upper (μ_{up}) and lower (μ_{low}) limit of such a quantity, see Fig. 7. Indeed, μ^* is the sampled shear modulus value from the uniform distribution $\mathcal{U}_{[a_l,b_l]}$ and it is reflected in the admissible range following Eq. (12). The application of the reflection method is useful when the random variable has bounded support such as in the present case where we consider, 33 kPa $\leq \mu_l(\theta) \leq$ 660 kPa, corresponding to Young's modulus values E_{up} = 2 MPa and E_{low} = 0.1 MPa. As detailed in Fig. 7c, the uniform distribution is modified to account for the upper values of the considered parameter.

$$\mu^* \mid \mu_{ANN} \sim \mathcal{U}_{[a_l,b_l]} \tag{11}$$

$$\mu = \begin{cases} \mu^* & \text{if } \mu^* \in [\mu_{low}, \mu_{up}] \\ 2\mu_{up} - \mu^* & \text{if } \mu^* > \mu_{up} \\ 2\mu_{low} - \mu^* & \text{if } \mu^* < \mu_{low} \end{cases} \tag{12}$$

This strategy represents only one possible probability distribution that we can integrate in order to consider the prior information about a certain model parameter. As a general comment, the uncertainty quantification strategy proposed herein allows for including any sources of uncertainties and relaxes some of the assumptions usually needed to implement the likelihood function. More complex probability distributions can be assumed to represent the uncertainties since the ABC approach only requires the ability to simulate them.

4. Numerical results

The proposed framework, definitions, and notation in this paper are tested in a concrete application. In particular, we propose the numerical results of the swelling of gels under elastic constraints. As previously mentioned, any function can be selected to represent the uncertainties in the problem under study. We explore a possible scenario with the aim of demonstrating the effectiveness of the uncertainty quantification integration in the ANN inverse design in swelling of confined gels. The posterior marginal probability density functions of the k-dimensional random vector $\bar{\boldsymbol{\mu}}$ are thus estimated using the ABC method following the procedure discussed in the previous section. The updating framework aims at modifying the prior probability distributions of the random vector $\bar{\boldsymbol{\mu}}$, and the results are shown by considering different levels of uncertainties. In particular, different informative prior distributions are considered (Eqs. (11) and (12)) by defining two amplitudes of the uniform distribution, Δ_1 = 330 kPa and Δ_2 = 165 kPa. Any other value can be selected to simulate the prior information about the random variables. These values are arbitrarily chosen to represent two different levels of knowledge regarding the elastic modulus of the gel-filled tube. The larger value, Δ_1, corresponds to a lower reliability of the prior information. On the contrary, by assuming a smaller value, Δ_2, we are assuming that the information about the spatial distribution of elastic modulus is more reliable and accurate.

The numerical results are illustrated with reference at two target shapes, assumed to have the meaning of observed data. The selected target shapes, A and B, are shown in Fig. 8. In the same figure, we can observe a schematic representation of the measurement error, e. The measurement error is herein assumed as a zero-mean Gaussian distribution $\mathcal{N}(0,\sigma^2)$ with σ = 0.15 (in terms of dimensionless shape size).

With these sources of uncertainties, the posterior distributions of the dimensionless shear modulus are shown in Fig. 9 through point-wise credible intervals, *i.e.* the range of the possible shear modulus values. In particular, we propose the representation of the posterior distributions with the quantile-based intervals by considering as lower bound the 5-quantile, $\bar{\mu}_{q5,l}$, and as upper bound the 95-quantile, $\bar{\mu}_{q95,l}$. The output of the ANN inverse tool is represented with a red line, while the blue line connects the lower bound values $\bar{\boldsymbol{\mu}}_{q5}$, the green line connects the upper bounds $\bar{\boldsymbol{\mu}}_{q95}$, and the brown line connects the medians $\bar{\boldsymbol{\mu}}_{median}$. In terms of resulting shapes, Fig. 10 shows the comparison between the target shapes and the obtained ones when $\bar{\boldsymbol{\mu}}_{q95}$, $\bar{\boldsymbol{\mu}}_{median}$ and $\bar{\boldsymbol{\mu}}_{q5}$ are used as input in the FE model. By comparing these two figures (Fig. 9 and Fig. 10), we observe how the discrepancy between the target shapes and the obtained shapes is correlated to the capability of producing – in terms

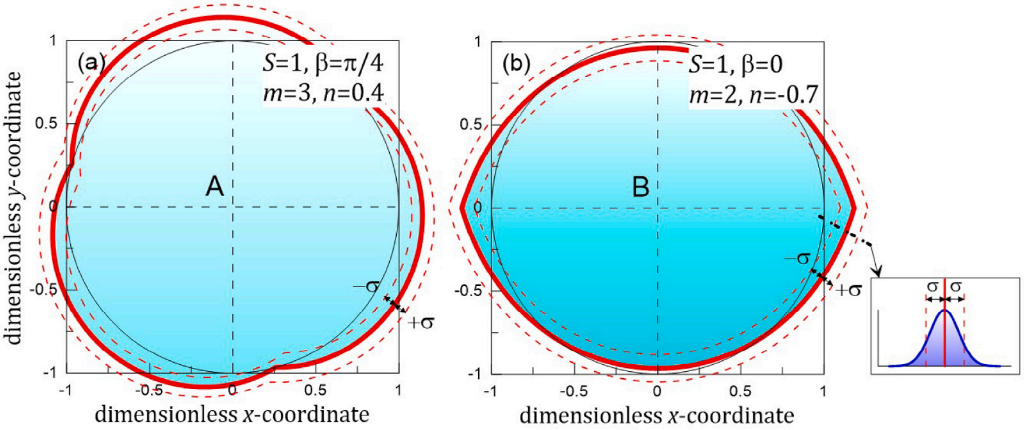

Fig. 8. Target shapes A (a) and B (b) adopted for the inverse design problem with uncertainties, see Monchetti et al. (2024a) for details about the mathematical description of these shapes. The shape discrepancy σ (positive and negative) is the allowed difference (dashed lines) – for the cases with shear modulus uncertainties – with respect to the desired shape. The insert represents the probability distribution of the measurement error e.

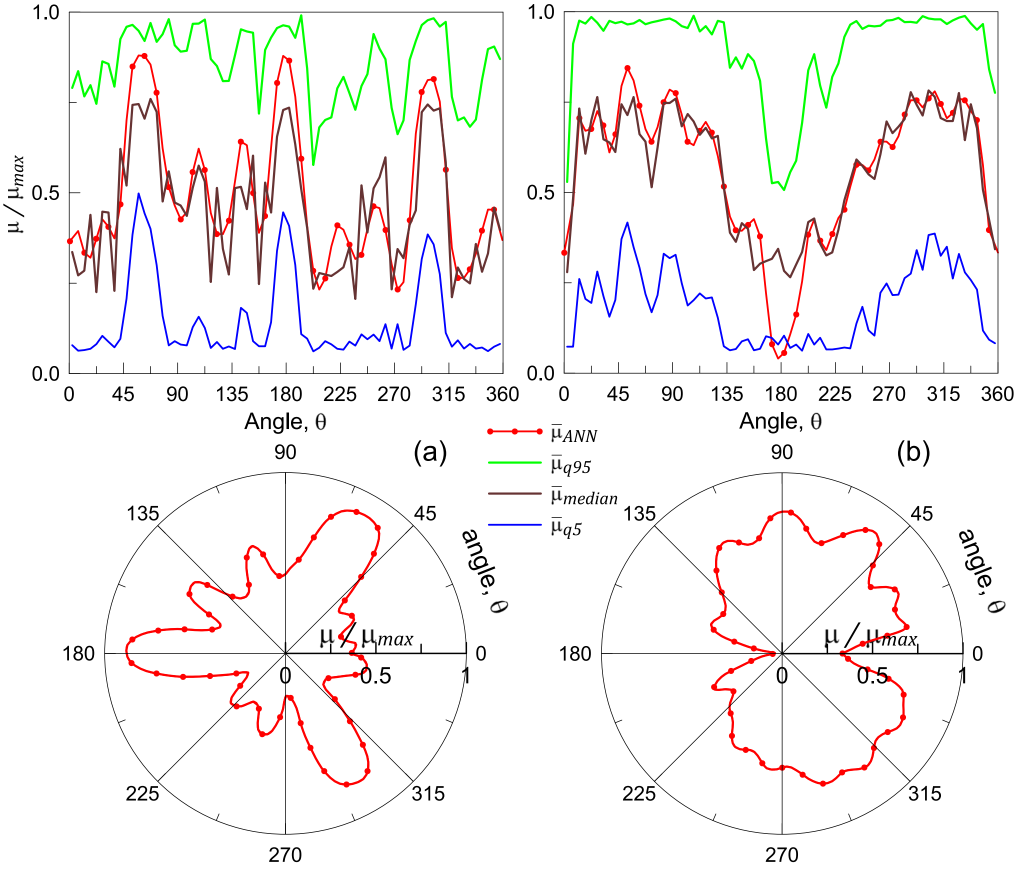

Fig. 9. Point-wise credible intervals of the dimensionless shear modulus ($\bar{\mu} = \mu/\mu_{max}$) distributions *vs* the angle θ provided by the ANN and posterior distributions provided by the ABC algorithm corresponding to the Δ_1-prior for the target shapes A (a) and B (b), see Fig. 8. The bottom figures illustrate the dimensionless distribution of the shear modulus $\bar{\mu}$ along the perimeter of the elastic layer.

of dimensionless shear modulus spatial distribution – a trend similar to that provided by the ANN. Indeed, as an example, observing the brown curve of the target shape B, despite the median of the posterior distribution of $\bar{\boldsymbol{\mu}}$ is close to the ANN output, the result in terms of the obtained shape differs significantly with respect to the ANN one. This is due to the lack of concentrated low values of the median of the posterior distribution of $\bar{\boldsymbol{\mu}}$ in correspondence to $\theta = 180°$ required to obtain the desired shape. Notably, the red line corresponds to the case when all the uncertainties are neglected; $\Delta = 0$ and $\sigma = 0$. This entails two strong assumptions: firstly, $\bar{\boldsymbol{\mu}}_{target}$ is supposed to be equal to $\bar{\boldsymbol{\mu}}_{ANN}$, and secondly the FE model is supposed to reproduce the result of the real process exactly. However as shown in Fig. 10, $\bar{\boldsymbol{r}}_{target} \neq \bar{\boldsymbol{r}}_{ANN}$. The result of the deterministic approach in terms of the discrepancy between $\bar{\boldsymbol{r}}_{target}$ and $\bar{\boldsymbol{r}}_{ANN}$, although being acceptable, is herein improved by integrating the ANN approach with the Bayesian inference.

The proposed procedure is consistent with the high predictive quality of the ANN developed in Monchetti et al. (2024a); indeed, the

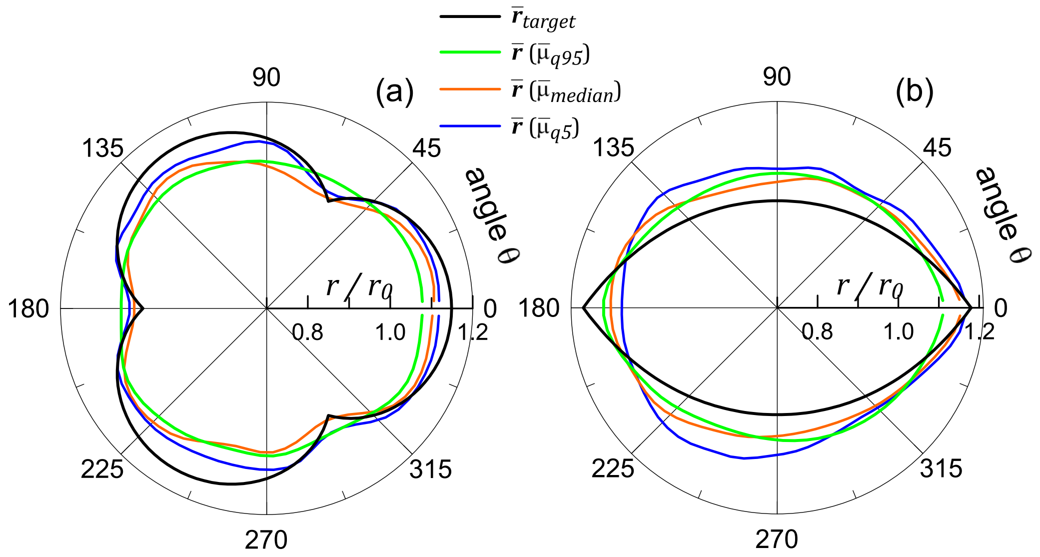

Fig. 10. Target shapes A (a) and B (b) vs shapes of the confined swelling problem obtained by using as input of the FE model $\bar{\boldsymbol{\mu}}_{q95}$, $\bar{\boldsymbol{\mu}}_{median}$ and $\bar{\boldsymbol{\mu}}_{q5}$, provided by the ABC algorithm, see Fig. 8.

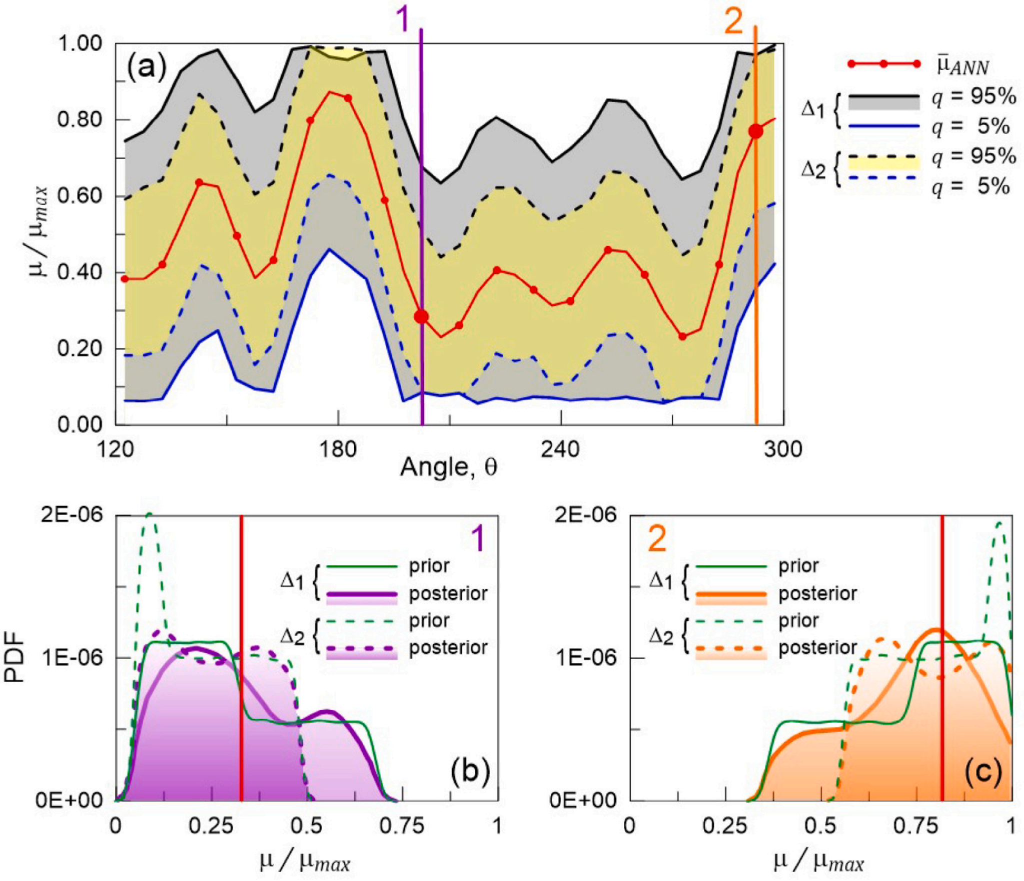

Fig. 11. Detail of the point-wise credible intervals of dimensionless shear distributions in the range $120° \leq \theta \leq 300°$ for the shape A (see Fig. 8). Comparison of the prior and posterior distributions of the dimensionless shear modulus at two different angles $\theta = 202.5°$ (b) and $\theta = 297.5°$ (c), for two different values of the prior distribution widths Δ_1, Δ_2. The vertical thick lines 1, 2 in (b,c) correspond to $\bar{\mu}_{ANN,l}$, with l corresponding to the above-cited angles, respectively.

posterior distributions have median values close to the output of the ANN, as can be appreciated from the comparison between the brown and the red lines in Fig. 9. This result confirms and strengthens the advantages obtainable by using Bayesian inference. Moreover, the proposed procedure is able to take into account an extended parameter space of $\bar{\mu}_l$, excluded by the deterministic procedure, which has a higher probability of reproducing the target shape. In general terms, the posterior distributions reflect the uncertainties introduced in the framework. The point-wise credible interval of the posterior distributions of $\bar{\mu}_l$ can be efficiently used by adopting, for each value of the angle θ, their mode value. Starting from the prior distributions of $\bar{\mu}_l$, the posterior distributions show the probability of obtaining $\bar{r}_{target,l}$

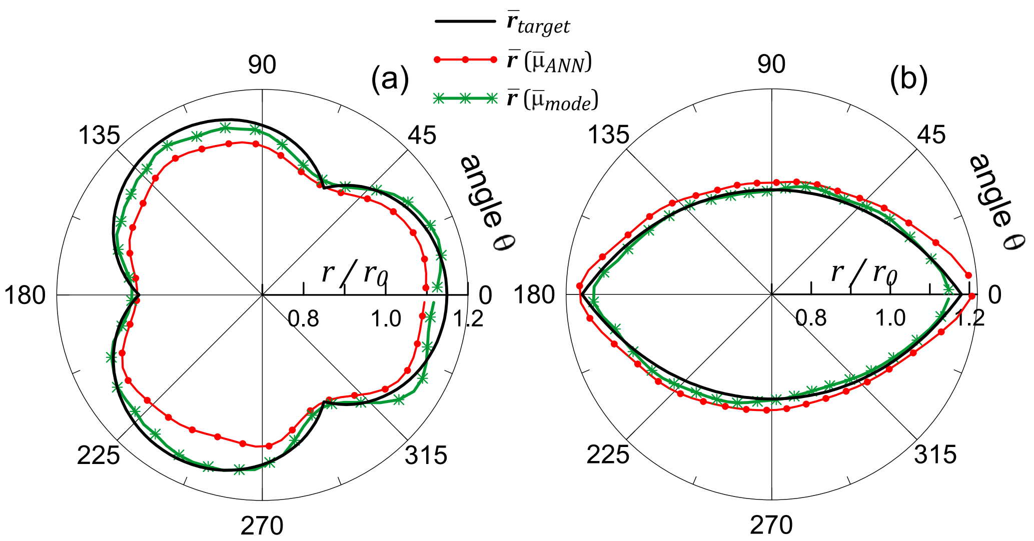

Fig. 12. Shapes of the confined swelling problem provided by FEM analyses whose mechanical property distributions are obtained by the ANN-ABC integrated approach (green line) and the ANN (red line) for the target shapes A (a) and B (b), see Fig. 8. In particular, the green line is obtained by using the mode of the posterior distributions of the ANN-ABC algorithm using the output of the ANN as the mean value of the prior distribution with amplitude value Δ_2. The errors are reported in Table 1.

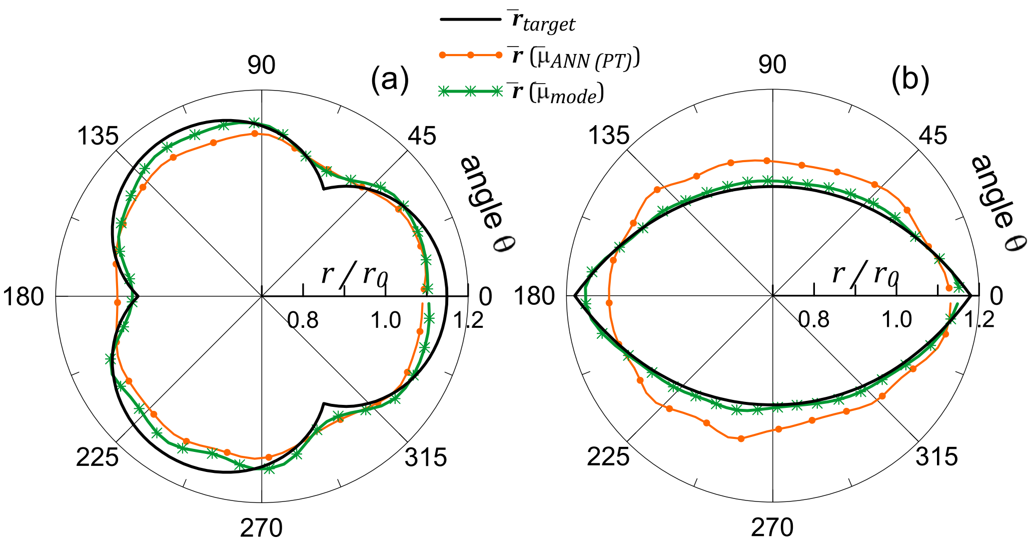

Fig. 13. Shapes of the confined swelling problem provided by FEM analyses whose mechanical property distributions are obtained by the ANN-ABC integrated approach (green line) and the poorly trained ANN (orange line) for the target shapes A (a) and B (b). In particular, the green line is obtained by using the mode of the posterior distributions of the ANN-ABC algorithm using the output of the poorly trained ANN as the mean value of the prior distribution with amplitude value Δ_1 = 330 kPa. The errors of the obtained shapes are reported in Table 1.

Table 1

Errors $\bar{d}$ (calculated according to Eq. (10)) quantifying the difference between the two considered target shapes (A and B) and the FEM results obtained by using the shear modulus distribution provided by: 1) the ANN inverse design tool ($\bar{d}_{ANN}$), the ANN-ABC integrated approach considering 2) Δ_1 and 3) Δ_2 amplitude values ($\bar{d}_{mode(\Delta 1)}$ and $\bar{d}_{mode(\Delta 2)}$, respectively), 4) the poorly trained ANN inverse design tool ($\bar{d}_{ANN(PT)}$), and the ANN(PT)-ABC integrated approach considering 5) Δ_1 and 6) Δ_2 amplitude values ($\bar{d}_{mode(PT,\Delta 1)}$ and $\bar{d}_{mode(PT,\Delta 2)}$, respectively).

target shape	1) $\bar{d}_{ANN}$	2) $\bar{d}_{mode(\Delta 1)}$	3) $\bar{d}_{mode(\Delta 2)}$	4) $\bar{d}_{ANN(PT)}$	5) $\bar{d}_{mode(PT,\Delta 1)}$	6) $\bar{d}_{mode(PT,\Delta 2)}$
A	2.1×10^{-3}	4.05×10^{-4}	2.4×10^{-4}	3.2×10^{-3}	5.2×10^{-4}	1.4×10^{-3}
B	6.5×10^{-4}	3.05×10^{-4}	1.3×10^{-4}	4.1×10^{-3}	1.7×10^{-4}	3.1×10^{-3}

associated with each value of $\bar{\mu}_l$, see Fig. 11. This aspect represents the main advantage of the proposed framework. In particular, Fig. 11 shows the point-wise credible intervals of $\bar{\mu}_l$ in the range $120° \le \theta \le 300°$ corresponding to two different values of Δ; in Fig. 11 (a, b) the comparison between prior and posterior distributions of $\bar{\mu}_l$ (at $l = [41, 60]$ corresponding to $\theta = [202.5°, 297.5°]$), respectively, is shown. Some interesting considerations can be traced: *(i)* the prior distribution significantly influences the posterior distribution; *(ii)* despite the prior distributions having equally probable intervals on the distribution's support, one or multiple modes characterize the posterior distributions. The strong influence of the prior distribution is an advantage; the prior distribution represents the initial belief on the uncertainty parameters of the physical system ($\bar{\mu}$) defined within specific bounds. Changing the interval length of the prior distribution entails modifying the probability within the distribution's support. Thus, we must expect that the posterior distributions show different trends and modes. Moreover, by selecting from the posterior distributions of $\bar{\mu}_l$ the mode for the various angles θ, the optimal spatial distribution of the shear modulus $\bar{\boldsymbol{\mu}}_{mode}$ is obtained. When $\bar{\boldsymbol{\mu}}_{mode}$ is used as input in the FE model, it provides the shape $\bar{\mathbf{r}}(\bar{\boldsymbol{\mu}}_{mode})$ represented in Fig. 12 (see the green curves). In Table 1, it can be appreciated that the error of $\bar{\mathbf{r}}(\bar{\boldsymbol{\mu}}_{mode})$ with respect to the target shape $\bar{\mathbf{r}}_{target}$, *i.e.* $\bar{d}_{mode}$, is lower compared to that corresponding to $\bar{\mathbf{r}}(\bar{\boldsymbol{\mu}}_{ANN})$, *i.e.* $\bar{d}_{ANN}$. Indeed, when the output of the ANN is used as input in the FE model, the obtained shape – represented by the vector $\bar{\mathbf{r}}(\bar{\boldsymbol{\mu}}_{ANN})$ – provides just an approximation of $\bar{\mathbf{r}}_{target}$. In particular, it represents the best possible approximation of $\bar{\mathbf{r}}_{target}$ when all the uncertainties are neglected. This result is herein improved by integrating the ANN tool with a Bayesian-based approach; the corresponding result represents the optimal spatial distribution of the shear modulus $\bar{\boldsymbol{\mu}}_{mode}$ for a specific target shape. Note that when a properly trained ANN is used to obtain the mean value of the prior distribution, we are using reliable information. Indeed, using the prior distribution with amplitude Δ_2 in the ABC procedure is more efficient than using an amplitude $\Delta_1 > \Delta_2$ (see in Table 1 the corresponding errors $\bar{d}_{mode(\Delta 1)}$ and $\bar{d}_{mode(\Delta 2)}$).

The proposed ANN-ABC integrated approach is also useful when the ANN tool cannot be trained with a sufficiently large number of cases; in fact, a poorly trained (PT) ANN does not provide a reliable mean value of the quantity of interest (see Table 1, $\bar{d}_{ANN(PT)}$). Thus, the use of the ABC approach in these cases allows for overcoming this drawback if a sufficiently large amplitude of the prior distribution is adopted. In Fig. 13, the output of a poorly trained ANN (only 300 cases have been considered for the training) is compared with the target shapes. In this case, the mean value of the prior distribution is the output of the poorly trained ANN, $\bar{\mu}_{ANN(PT)}$. The above-mentioned mean value is less reliable due to the low number of training cases; consequently, the result of the ABC approach provides lower discrepancy with respect to the target shape when a wider amplitude of the prior distribution is considered. This is confirmed comparing the values of $\bar{d}_{mode(PT,\Delta 1)}$ and $\bar{d}_{mode(PT,\Delta 2)}$ shown in Table 1. It is worth noticing that the errors are lower when the ANN is properly trained, despite this corresponds to higher computational efforts. However, the proposed ANN-ABC integrated approach is capable of reducing the computational cost, as its high efficiency demonstrates even if the results of the poorly trained ANN are used.

5. Conclusions

In this paper, we have proposed a method to quantitatively evaluate the uncertainties in the inverse design of shape-morphing elements. In particular, an ANN-based method has been integrated with a Bayesian-based approach. We applied this combined framework to swelling-driven shape change in confined gels; interesting results, that are useful to guide the design of shape change elements accounting for the uncertainties of their synthesis and production process, have been obtained. Given a specific target shape, the present method can evaluate the probability of obtaining such a desired shape. Irrespective of the assumed prior distribution – which strongly affects the posterior one – the median value of the posterior distribution remains close to the output of the ANN prediction. In particular, for the considered swelling-driven shape change problem, the posterior distribution of the shear modulus at different angles θ entails obtaining the most likely value leading to the target's closest shape. The results reveal that the predictive capability of an ANN-based inverse design tool can be significantly improved by integrating it within a Bayesian-based framework. The proposed integrated approach is of particular interest when the material – whose morphing has to be obtained – is produced using AM technologies, which usually provide materials with a discrete variation of the mechanical parameters of the printed parts.

Future perspectives of the proposed integrated procedure should address the enhancement *(i)* of the informativeness prior distribution and *(ii)* of the distance function. In the present study, the sampled prior values of the shear modulus at different angles θ are assumed to be independent of each other. Improvements in the latter aspect can be obtained by introducing a dependence between the mechanical parameter value at a specific angle and the values at the neighboring locations, as suggested by the physical problem in turn. Moreover, the same dependence could also be introduced to improve the distance function by using statistical parameters reflecting the scatter between the mechanical parameter values. As a final consideration, an interesting sequel to this paper will be related to the integration of an experimental setup, which can strengthen the effectiveness of the proposed framework in many engineering application fields.

CRediT authorship contribution statement

Silvia Monchetti: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Conceptualization. **Roberto Brighenti:** Writing – review & editing, Writing – original draft, Methodology, Data curation, Conceptualization. **Noy Cohen:** Writing – review & editing, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no conflict of interest.

Acknowledgments

The authors thank Prof. Cecilia Viscardi for helpful discussions.

Appendix

A polymer possessing a sufficiently high affinity with a fluid is keen to uptake a significant amount of such a fluid within its network. Both the polymer network and the fluid are assumed to be individually incompressible. The polymer network occupies the region B_0 with boundary ∂B_0, while the material points in the dry polymer (reference configuration) are identified by the vector $\mathbf{X}$. In the present problem, the network is placed in a cylindrical ring and is in contact with an appropriate solvent; thanks to the diffusion of the solvent molecules into the network, the gel swells, showing a significant volume increase. In the deformed swollen configuration, the gel occupies the region B with boundary ∂B, while the material points in this configuration are denoted by $\mathbf{x}$. The deformation gradient quantifying the change of configuration between the reference to the deformed swollen one is given by $\mathbf{F} = \partial\mathbf{x}/\partial\mathbf{X}$, while the volumetric deformation is provided by $J = \det\mathbf{F} = 1 + C_s v_s$, where C_s is the uptake concentration and v_s is the molar volume of the solvent. The latter relationship indicates that the increase in volume of the gel-fluid system is proportional to the molecules that diffuse into the network; such a volume increase represents a volumetric constraint that must be fulfilled by the deformation gradient at any time during the swelling process. The equilibrium state of the gel corresponds to the stationarity of the following energy-density function (Hong et al., 2008; Chester and Anand, 2010; Cohen and McMeeking, 2019)

$$\Psi = (\Psi_{net} + \Psi_{mix}) + \pi(1 + C_s v_s - J) \tag{A.1}$$

where Ψ_{net} is the energy density related to the deformation of the network (here assumed to follow the Neo-Hookean hypothesis,

$$\Psi_{net} = \frac{\mu}{2}\left(I_1 - 3\right), \tag{A.2}$$

being $I_1 = \mathrm{tr}\left(\mathbf{F}\mathbf{F}^T\right)$ the first invariant and μ the shear modulus); finally, Ψ_{mix} is the solid-fluid mixing energy density (Huggins, 1942), (Flory, 1942)

$$\Psi_{mix} = \frac{k_B T}{v_s}\left[(J-1)\ln\frac{v_s C_s}{J} + v_s C_s \frac{\chi}{J}\right], \tag{A.3}$$

where k_B is the Boltzmann constant, T is the temperature, and χ is the network-fluid dimensionless interaction parameter quantifying the gel-solvent affinity. Finally, π is a pressure-like term, necessary to fulfill the volumetric constraints, which can be determined from the boundary conditions of the problem. The zero variation of Eq. (A.1) with respect to the involved field quantities $\boldsymbol{F}, C_s$ identifies the equilibrium condition of the network-fluid system:

$$\begin{aligned}\delta\Psi =& \frac{\partial\Psi_{net}}{\partial\boldsymbol{F}} : \delta\boldsymbol{F} - \pi\frac{\partial J}{\partial\boldsymbol{F}} : \delta\boldsymbol{F} + \pi\upsilon_s\delta C_s + \frac{\partial\Psi_{mix}}{\partial C_s}\delta C_s = \\ =& [\boldsymbol{P} - \pi J\boldsymbol{F}^{-T}] : \delta\boldsymbol{F} + (\mu_c + \pi\upsilon_s)\delta C_s = 0\end{aligned} \tag{A.4}$$

where $\mu_c = \partial\Psi_{mix}/\partial C_s$ is the so-called chemical potential, while the symbol $:$ indicates the double contraction operator. It is worth noticing that the hydrostatic pressure π can be seen as a further term affecting the chemical potential μ_c, so an amended pressure-dependent chemical potential μ_c' can be introduced:

$$\mu_c' = \mu_c + \pi\upsilon_s \tag{A.5}$$

It appears that π plays the role of hindering the solvent molecules from entering the polymer. In this regard, the presence of an elastic layer confining the gel limits the volumetric expansion and promotes the pressure to increase, further hindering the swelling process. This corresponds to an equilibrium state whose solvent concentration is lower than that achievable in the free swelling case. The above-stated variational problem can be solved numerically by adopting a discretization framework such as the finite element (FE) method. Problems involving the coupling of various physics, such as fluid diffusion and network mechanics, can be solved numerically in a staggered or monolithic way; the first approach is simpler and, if the time step adopted is sufficiently small, the solution is relatively accurate. It is based on the concept of freezing the fluid diffusion problem during a small time step while the mechanical problem is solved, and by assuming the previously converged mechanical configuration of the system to be fixed, while the diffusion problem is advanced further in the time domain. The staggered approach is adopted in the present computational framework. More details can be found in Monchetti et al. (2024a).

Data availability

Data will be made available on request.

References

Anitescu, C., İsmail Ateş, B., Rabczuk, T., 2023. Physics-informed neural networks: Theory and applications. In: Machine Learning in Modeling and Simulation: Methods and Applications. Springer, pp. 179–218.

Argatov, I., 2019. Artificial neural networks (ANNs) as a novel modeling technique in tribology. Front. Mech. Eng. 5, 30.

Bai, J., Rabczuk, T., Gupta, A., Alzubaidi, L., Gu, Y., 2023. A physics-informed neural network technique based on a modified loss function for computational 2D and 3D solid mechanics. Comput. Mech. 71 (3), 543–562.

Bayes, T., 1763. LII. An essay towards solving a problem in the doctrine of chances. By the late rev. Mr. Bayes, FRS communicated by Mr. Price, in a letter to John Canton, AMFR s. Philos. Trans. R. Soc. Lond. (53), 370–418.

Beaumont, M.A., 2019. Approximate bayesian computation. Annu. Rev. Stat. Appl. 6, 379–403.

Berger, J.O., Smith, L.A., 2019. On the statistical formalism of uncertainty quantification. Annu. Rev. Stat. Appl. 6, 433–460.

Bhuwal, A.S., Pang, Y., Ashcroft, I., Sun, W., Liu, T., 2023. Discovery of quasi-disordered truss metamaterials inspired by natural cellular materials. J. Mech. Phys. Solids 175, 105294.

Boley, J.W., Van Rees, W.M., Lissandrello, C., Horenstein, M.N., Truby, R.L., Kotikian, A., Lewis, J.A., Mahadevan, L., 2019. Shape-shifting structured lattices via multimaterial 4D printing. Proc. Natl. Acad. Sci. 116 (42), 20856–20862.

Brighenti, R., Artoni, F., Vernerey, F., Torelli, M., Pedrini, A., Domenichelli, I., Dalcanale, E., 2018. Mechanics of responsive polymers via conformationally switchable molecules. J. Mech. Phys. Solids 113, 65–81.

Brighenti, R., Cosma, M.P., 2020. Swelling mechanism in smart polymers responsive to mechano-chemical stimuli. J. Mech. Phys. Solids 143, 104011.

Brighenti, R., Cosma, M.P., 2022. Mechanics of multi-stimuli temperature-responsive hydrogels. J. Mech. Phys. Solids 169, 105045.

Brighenti, R., Cosma, M.P., Cohen, N., 2023. Mechanics and physics of the light-driven response of hydrogels. Mech. Res. Commun. 129, 104077.

Brighenti, R., McMahan, C.G., Cosma, M.P., Kotikian, A., Lewis, J.A., Daraio, C., 2021. A micromechanical-based model of stimulus responsive liquid crystal elastomers. Int. J. Solids Struct. 219, 92–105.

Caló, E., Khutoryanskiy, V.V., 2015. Biomedical applications of hydrogels: A review of patents and commercial products. Eur. Polym. J. 65, 252–267.

Chakraverty, S., Jeswal, S.K., 2021. Applied Artificial Neural Network Methods for Engineers and Scientists: Solving Algebraic Equations. World Scientific.

Chester, S.A., Anand, L., 2010. A coupled theory of fluid permeation and large deformations for elastomeric materials. J. Mech. Phys. Solids 58 (11), 1879–1906.

Chirani, N., Yahia, L., Gritsch, L., Motta, F.L., Chirani, S., Farè, S., et al., 2015. History and applications of hydrogels. J. Biomed. Sci. 4 (02), 1–23.

Cohen, N., 2019. Programming the equilibrium swelling response of heterogeneous polymeric gels. Int. J. Solids Struct. 178, 81–90.

Cohen, N., McMeeking, R.M., 2019. On the swelling induced microstructural evolution of polymer networks in gels. J. Mech. Phys. Solids 125, 666–680.

Dawei, Z., Jinyu, Z., Chunqiu, L., Zhiling, W., 2021. A short review of reliability-based design optimization. In: IOP Conference Series: Materials Science and Engineering. vol. 1043, (3), IOP Publishing, 032041.

Dimiduk, D.M., Holm, E.A., Niezgoda, S.R., 2018. Perspectives on the impact of machine learning, deep learning, and artificial intelligence on materials, processes, and structures engineering. Integr. Mater. Manuf. Innov. 7, 157–172.

Du, P., Wang, F.-F., Wang, J., 2023. On a simplified multi-layered plate model of growth: Asymptotic analyses and numerical implementation. Thin-Walled Struct. 191, 111100.

Dubey, S.R., Singh, S.K., Chaudhuri, B.B., 2022. Activation functions in deep learning: A comprehensive survey and benchmark. Neurocomputing 503, 92–108.

Espino, M.T., Tuazon, B.J., Espera, Jr., A.H., Nocheseda, C.J.C., Manalang, R.S., Dizon, J.R.C., Advincula, R.C., 2023. Statistical methods for design and testing of 3D-printed polymers. MRS Commun. 13 (2), 193–211.

Flory, P.J., 1942. Thermodynamics of high polymer solutions. J. Chem. Phys. 10 (1), 51–61.

Gray, A., Wimbush, A., de Angelis, M., Hristov, P.O., Calleja, D., Miralles-Dolz, E., Rocchetta, R., 2022. From inference to design: A comprehensive framework for uncertainty quantification in engineering with limited information. Mech. Syst. Signal Process. 165, 108210.

Hamdia, K.M., Ghasemi, H., 2022. Quantifying the uncertainties in modeling soft composites via a multiscale approach. Int. J. Solids Struct. 256, 111959.

Hanuhov, T., Brighenti, R., Cohen, N., 2024. Energy harvesting with dielectric elastomer tubes: active and (responsive materials-based) passive approaches. Smart Mater. Struct. 33 (5), 055004.

Herrmann, A., Haag, R., Schedler, U., 2021. Hydrogels and their role in biosensing applications. Adv. Heal. Mater. 10 (11), 2100062.

Hoff, P.D., 2009. A First Course in Bayesian Statistical Methods. vol. 580, Springer.

Hoffman, A.S., 2013. Stimuli-responsive polymers: Biomedical applications and challenges for clinical translation. Adv. Drug Deliv. Rev. 65 (1), 10–16.

Hong, W., Zhao, X., Zhou, J., Suo, Z., 2008. A theory of coupled diffusion and large deformation in polymeric gels. J. Mech. Phys. Solids 56 (5), 1779–1793.

Hu, W., Cheng, S., Yan, J., Cheng, J., Peng, X., Cho, H., Lee, I., 2024. Reliability-based design optimization: a state-of-the-art review of its methodologies, applications, and challenges. Struct. Multidiscip. Optim. 67 (9), 168.

Huggins, M.L., 1942. Thermodynamic properties of solutions of long-chain compounds. Ann. New York Acad. Sci. 43 (1), 1–32.

Kennedy, M.C., O'Hagan, A., 2001. Bayesian calibration of computer models. J. R. Stat. Soc. Ser. B Stat. Methodol. 63 (3), 425–464.

Kim, J.Y., Garcia, D., Zhu, Y., Higdon, D.M., Hang, Z.Y., 2022. A Bayesian learning framework for fast prediction and uncertainty quantification of additively manufactured multi-material components. J. Mater. Process. Technol. 303, 117528.

Levin, M., Cohen, N., 2023. Swelling under constraints: Exploiting 3D-printing to optimize the performance of gel-based devices. Adv. Mater. Technol. 8 (14), 2202136.

Li, J., Li, X., Xu, G., Zheng, Z., Deng, J., Ding, X., 2018. A self-deformable gel system with asymmetric shape change based on a gradient structure. Chem. Commun. 54 (82), 11594–11597.

Liu, Y., He, H., Cao, Y., Liang, Y., Huang, J., 2024. Inverse design of TPMS piezoelectric metamaterial based on deep learning. Mech. Mater. 198, 105109.

Liu, X., Tian, S., Tao, F., Yu, W., 2021. A review of artificial neural networks in the constitutive modeling of composite materials. Compos. Part B: Eng. 224, 109152.

Louf, J.F., Lu, N.B., OConnell, M.G., Cho, H.J., Datta, S.S., 2021. Under pressure hydrogel swelling in a granular medium. Sci. Adv. 7 (7), eabd2711.

Madireddy, S., Sista, B., Vemaganti, K., 2015. A Bayesian approach to selecting hyperelastic constitutive models of soft tissue. Comput. Methods Appl. Mech. Engrg. 291, 102–122.

Mahadevan, S., Nath, P., Hu, Z., 2022. Uncertainty quantification for additive manufacturing process improvement: Recent advances. ASCE- ASME J. Risk Uncertain. Eng. Syst. Part B: Mech. Eng. 8 (1), 010801.

Marcombe, R., Cai, S., Hong, W., Zhao, X., Lapusta, Y., Suo, Z., 2010. A theory of constrained swelling of a pH-sensitive hydrogel. Soft Matter 6 (4), 784–793.

Monchetti, S., Brighenti, R., Hanuhov, T., Cohen, N., 2024a. Controlled swelling-induced shape change of soft gel filled structures. Thin-Walled Struct. 204, 112280.

Monchetti, S., Pepi, C., Viscardi, C., Gioffrè, M., 2024b. Approximate Bayesian computation for structural identification of ancient tie-rods using noisy modal data. Probabilistic Eng. Mech. 77, 103674.

Monchetti, S., Viscardi, C., Betti, M., Clementi, F., 2024c. Comparison between Bayesian updating and approximate Bayesian computation for model identification of masonry towers through dynamic data. Bull. Earthq. Eng. 22 (7), 3491–3509.

Oliver, K., Seddon, A., Trask, R.S., 2016. Morphing in nature and beyond: a review of natural and synthetic shape-changing materials and mechanisms. J. Mater. Sci. 51, 10663–10689.

Park, S., Marimuthu, K.P., Han, G., Lee, H., 2023. Deep learning based nanoindentation method for evaluating mechanical properties of polymers. Int. J. Mech. Sci. 246, 108162.

Parodi, A., Khaled, S.M., Yazdi, I.K., Evangelopoulos, M., Furman, N.E.T., Wang, X., Urzi, F., Hmaidan, S., Hartman, K.A., Tasciotti, E., 2016. Smart hydrogels. In: Bhushan, B. (Ed.), Encyclopedia of Nanotechnology. Springer Netherlands, Dordrecht, pp. 3735–3747.

Pepi, C., Gioffre, M., Grigoriu, M., 2020. Bayesian inference for parameters estimation using experimental data. Probabilistic Eng. Mech. 60, 103025.

Plummer, A., Adkins, C., Louf, J.-F., Košmrlj, A., Datta, S.S., 2024. Obstructed swelling and fracture of hydrogels. Soft Matter 20 (7), 1425–1437.

Pogorelko, V., Mayer, A., Fomin, E., Fedorov, E., 2024. Examination of machine learning method for identification of material model parameters. Int. J. Mech. Sci. 265, 108912.

Quek, A., Ouyang, N., Lin, H.-M., Delaire, O., Guilleminot, J., 2025. Enhancing robustness in machine-learning-accelerated molecular dynamics: A multi-model nonparametric probabilistic approach. Mech. Mater. 105237.

Ravi, N., Mitra, A., Hamilton, P., Horkay, F., 2002. Characterization of the network properties of poly (ethylene glycol)–acrylate hydrogels prepared by variations in the ethanol–water solvent composition during crosslinking copolymerization. J. Polym. Sci. Part B: Polym. Phys. 40 (23), 2677–2684.

Revete, A., Aparicio, A., Cisterna, B.A., Revete, J., Luis, L., Ibarra, E., Segura González, E.A., Molino, J., Reginensi, D., 2022. Advancements in the use of hydrogels for regenerative medicine: properties and biomedical applications. Int. J. Biomater. 2022 (1), 3606765.

Rubin, D.B., 1984. Special, invited paper. Ann. of" Stat. 12 (4), 1151–1172.

Thomas, A.J., Barocio, E., Bilionis, I., Pipes, R.B., 2022. Bayesian inference of fiber orientation and polymer properties in short fiber-reinforced polymer composites. Compos. Sci. Technol. (228), 109630.

Velders, A.H., Dijksman, J.A., Saggiomo, V., 2017. Hydrogel actuators as responsive instruments for cheap open technology (HARICOT). Appl. Mater. Today 9, 271–275.

Wang, Q., Li, C., Zhang, S., Zhou, C., Zhou, Y., 2025. Physics-informed extreme learning machine framework for solving linear elasticity mechanics problems. Int. J. Solids Struct. 309, 113157.

Wang, Y., Zong, Y., McCreight, J.L., Hughes, J.D., Tartakovsky, A.M., 2024. Bayesian reduced-order deep learning surrogate model for dynamic systems described by partial differential equations. Comput. Methods Appl. Mech. Engrg. 429, 117147.

Wright, L.G., Onodera, T., Stein, M.M., Wang, T., Schachter, D.T., Hu, Z., McMahon, P.L., 2022. Deep physical neural networks trained with backpropagation. Nature 601 (7894), 549–555.

Wu, S., Li, H., Chen, J.P., Lam, K., 2004. Modeling investigation of hydrogel volume transition. Macromol. Theory Simulations 13 (1), 13–29.

Xiao, S., Li, J., Bordas, S.P.A., Kim, T.Y., 2023. Artificial neural networks and their applications in computational materials science: A review and a case study. Adv. Appl. Mech. 1–33.

Zhang, L., Zhao, T., Liu, M., 2024a. Bioinspired multiphase gels using spatial confinement strategy. Accounts Mater. Res. 5 (1), 48–63.

Zhang, K., Zhou, Y., Zhang, J., Liu, Q., Hanenberg, C., Mourran, A., Wang, X., Gao, X., Cao, Y., Herrmann, A., et al., 2024b. Shape morphing of hydrogels by harnessing enzyme enabled mechanoresponse. Nat. Commun. 15 (1), 249.

Zhao, Q., Yang, X., Ma, C., Chen, D., Bai, H., Li, T., Yang, W., Xie, T., 2016. A bioinspired reversible snapping hydrogel assembly. Mater. Horizons 3 (5), 422–428.

Zu, S., Wang, Z., Zhang, S., Guo, Y., Chen, C., Zhang, Q., Liu, T., Liu, Q., Zhang, Z., 2022. A bioinspired 4D printed hydrogel capsule for smart controlled drug release. Mater. Today Chem. 24, 100789.