Reconstruction of a constitutive law for rubber from in silico experiments using Ogden's laws

Maya De Buhan, Antoine Gloria, Patrick Le Tallec, Marina Vidrascu

RECONSTRUCTION OF A CONSTITUTIVE LAW FOR RUBBER FROM IN SILICO EXPERIMENTS USING OGDEN'S LAWS

MAYA DE BUHAN, ANTOINE GLORIA, PATRICK LE TALLEC, AND MARINA VIDRASCU

Abstract. This article deals with the following data assimilation problem: construct an analytical approximation of a numerical constitutive law in three-dimensional nonlinear elasticity. More precisely we are concerned with a micro-macro model for rubber. Macroscopic quantities of interest such as the Piola-Kirchhoff stress tensor can be approximated for any value of the strain gradient by numerically solving a nonlinear PDE. This procedure is however computationally demanding. Hence, although conceptually satisfactory, this physically-based model is of no direct practical use. The aim of this article is to circumvent this difficulty by proposing a numerical strategy to reconstruct from in silico experiments an accurate analytical proxy for the micro-macro constitutive law.

MSC2010 subject classifications: 74B20, 74E35, 74Q05, 93A30, 93B30.
Key words and phrases: paramater estimation, multiscale modelling, rubber elasticity, polymer physics.

1. Introduction

In this article we address a problem which exhibits at the same time very standard and rather unusual features: data assimilation in nonlinear elasticity for a micro-macro constitutive law.

Let us start with the features which make our problem a standard one in mechanical engineering: we wish to reconstruct an analytical constitutive law from a set of experiments. In particular, we aim at reconstructing a function (the energy density) from a set of samples (the experiments). Such problems are quite standard in elasticity. We refer the reader to the review paper [5]. Most of the constitutive laws used in rubber elasticity (or more generally in computational mechanics) are phenomenological (see for instance [16, 3, 17, 7]): the law is supposed to have a specific analytical form characterized by some explicit parameters. For the constitutive law to be of any use, these parameters have to be fitted. This is where things get complicated. On the one hand, the more parameters the more accurate the phenomenological law. On the other hand, the more parameters the more difficult the data assimilation problem. There is a wide choice of measurements which can be used to estimate the parameters, see for instance [5] and the references therein. For nonlinear materials, few theoretical results are available, and parameter identification methods are often based on direct measurements of the stress associated with a homogeneous strain and give satisfactory results only for a very small number of parameters. It is for instance rather well-admitted that Ogden's laws have the potential to model rubber elasticity quite well [16, 3, 7]. Yet for reasonable sets of experimental data, there may be several possible sets of fitted parameters which give similar results on the set of data but which yield completely different behaviors in other regimes of interest, as shown in [18]. Hence, although methods have been developed to fit parameters in Ogden's laws [16, 20, 9, 18], Ogden's laws are not so used in practice. The associated inverse problem is indeed often ill-posed: the observations obtained by mechanical experiments are too partial to characterize the constitutive law. Although this inverse problem is standard, there is up to now no fully satisfactory way to solve it in practice.

Date: January 13, 2014.

Let us turn now to the unusual features of this problem. Unlike purely phenomenological constitutive laws, the model under investigation here is based on some physical grounds at the scale of the polymer-chain network [11]. In particular the micro-macro constitutive law is obtained by a rigorous thermodynamic limit starting from a physically-based (without phenomenological parameters) model [2]. As a by-product of the analysis, we learn that the associated micro-macro energy density satisfies some formula, which involves the solution of a nonlinear elasticity problem on a sequence of domains of increasing size (see Theorem 1 in Section 2). This is the so-called cell-problem in stochastic homogenization. Although this energy density is not analytical, it can be numerically approximated at any deformation gradient, as shown in [11]. Hence it seems there is no need for data assimilation. Things are unfortunately not that simple and the solution method used to approximate the energy density is computationally very demanding. One cannot afford to include it into a nonlinear elasticity software. This is where data assimilation comes into the picture again: given a set of data generated by solving numerically the cell-problem we wish to construct an analytical approximation of the micro-macro constitutive law.

In this article we shall show that the specific features of our problem make the "standard inverse problem" recalled above much nicer. Data assimilation in rubber elasticity may be an ill-posed problem because the sets of data which are available are often too partial (engineering stress for uniaxial and biaxial tractions for instance). In particular all the regimes cannot be tested by mechanical experiments. On the contrary, for the micro-macro model and the numerical approximation method of the associated energy density we are dealing with here, any strain gradient can be considered: we have at our disposal an arbitrary amount of data at arbitrary values of the strain gradient. This opens the door to the use of reliable and efficient data assimilation methods. In addition, the analysis of the model and of its thermodynamic limit performed in [2] is a very good guide to restrict the class of admissible energy densities in which to solve the inverse problem — which is the aim of our study.

The article is organized as follows. In Section 2 we briefly recall the micro-macro model for rubber, its structural properties and the link with analytical constitutive laws. In the following section, we describe the inverse problem to be solved and the numerical solution method used (an evolutionary algorithm). Section 4 is dedicated to the calibration and test of the method, for different analytical energy densities and with exact and noisy data. In the last section, we apply the method to the micro-macro model of interest. The numerical results are very good, both qualitatively and quantitatively. In particular they draw the link between a physical model based on "first principles" (the micro-macro model) and phenomenological constitutive laws for rubber.

2. Model and parametrization

2.1. **Homogenization of a discrete model for rubber.** In this subsection we recall the discrete stochastic homogenization results of [2] which have allowed us to study the thermodynamic limit of a discrete model for rubber in [11]. To this aim, we have to define the notion of stochastic lattice, and make precise the associated energy functional.

Definition 1. *We say that a stochastic point process $\mathcal{L}$ in $\mathbb{R}^3$ (that is a sequence of random points in $\mathbb{R}^3$) is admissible if:*

- *(regularity) There exist $\mathbf{r} \geq r > 0$ such that almost surely:*
  - *any two points of $\mathcal{L}$ cannot be closer than r (hard-core property),*
  - *any ball of radius $\mathbf{r}$ contains at least one point of $\mathcal{L}$ (non-empty space property);*
- *(stationarity) $\mathcal{L}$ and $x + \mathcal{L}$ have the same statistics for all $x \in \mathbb{R}^3$;*
- *(ergodicity) $\mathcal{L}$ is ergodic.*

We further assume that the Delaunay tessellation $\mathcal{T}$ of $\mathbb{R}^3$ into tetrahedra associated with $\mathcal{L}$ (that is, the vertices of $\mathcal{T}$ are given by $\mathcal{L}$) is almost surely unique (see [8]*).*

For rigorous definitions of admissible stochastic lattices and their Delaunay tessellations, we refer the reader to [1] and [12], and to the references therein. Let us also introduce a rescaling of $\mathcal{L}$ and $\mathcal{T}$. For all $\varepsilon > 0$, we set $\mathcal{L}_\varepsilon := \varepsilon\mathcal{L}$, which satisfies Definition 1 with εr and $\varepsilon\mathbf{r}$ in place of r and $\mathbf{r}$, and with $\mathcal{T}_\varepsilon := \varepsilon\mathcal{T}$ in place of $\mathcal{T}$.

Given a tessellation $\mathcal{T}_\varepsilon$ of $\mathbb{R}^3$, one may define the space $\mathcal{S}(\mathcal{T}_\varepsilon)$ of continuous and piecewise affine deformations u_ε on $\mathcal{T}_\varepsilon$. Such deformations u_ε are such that their gradients ∇u_ε are piecewise constant on $\mathcal{T}_\varepsilon$. In particular, for every element T_ε (tetraedron) of the tessellation $\mathcal{T}_\varepsilon$, $\det \nabla u_{\varepsilon|T_\varepsilon}$ measures the ratio of volume between $u_\varepsilon(T_\varepsilon)$ and T_ε.

We are now in position to associate an energy with any deformation field $u_\varepsilon \in \mathcal{S}(\mathcal{T}_\varepsilon)$, on an open bounded domain D of $\mathbb{R}^3$. We consider two contributions: an energy associated with the changes of length of the edges of the tessellation, and an energy associated with the changes of volume of the elements of the tessellation. More precisely, denoting by $\mathcal{E}_d$ the set of edges of $\mathcal{T}$, we define the energy of $u_\varepsilon \in \mathcal{S}(\mathcal{T}_\varepsilon)$ on D by

$$F_\varepsilon(u_\varepsilon, D) = \varepsilon^3 \sum_{e\in\mathcal{E}_d, e\subset D/\varepsilon} W_{nn}\Big(|e_1 - e_2|, \frac{|u_\varepsilon(\varepsilon e_1) - u_\varepsilon(\varepsilon e_2)|}{\varepsilon|e_1 - e_2|}\Big) + \sum_{T\in\mathcal{T}, T\subset D/\varepsilon} |\varepsilon T| W_{\rm vol}(\det \nabla u_{\varepsilon|T_\varepsilon}), \tag{1}$$

where $e = (e_1, e_2)$ (e_1 and e_2 are the two vertices of the edge e), $W_{nn} : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ is the energy of the deformed edges, and $W_{\rm vol} : \mathbb{R} \to \mathbb{R}^+$ is the volumetric energy. Denote by $\mathbb{M}^3(\mathbb{R})$ the set of 3×3-real matrices. We make the following assumptions on W_{nn} and $W_{\rm vol}$:

Hypothesis 1. *There exist $p > 1$ and a positive constants C such that for all $r \le \gamma \le \mathbf{r}$, $\lambda \ge 0$, and $\Lambda \in \mathbb{M}^3(\mathbb{R})$,*

$$\frac{1}{C}\lambda^p - C \le W_{nn}(\gamma, \lambda) \le C(\lambda^p + 1), \tag{2}$$

$$W_{\rm vol}(\det \Lambda) \le C(|\Lambda|^p + 1). \tag{3}$$

We then have the following convergence result (see [1, Theorem 5]).

Theorem 1. *For all $\varepsilon > 0$, let $\mathcal{L}_\varepsilon$ and $\mathcal{T}_\varepsilon$ be the rescaled stochastic point process and the associated Delaunay tessellation of Definition 1. For every open bounded subset D of $\mathbb{R}^3$ with a Lipschitz-continuous boundary, we consider the energy $F_\varepsilon(u_\varepsilon, D)$ defined by* (1) *for $u_\varepsilon \in \mathcal{S}(\mathcal{T}_\varepsilon)$, and extended by $+\infty$ on $W^{1,p}(D) \setminus \mathcal{S}(\mathcal{T}_\varepsilon)$, for $p > 1$, W_{nn}, and $W_{\rm vol}$ as in Hypothesis 1.*
Then the functional $F_\varepsilon(\cdot, D)$ $\Gamma(L^p(D))$-converges on $W^{1,p}(D)$ as $\varepsilon \to 0$ to the functional $F_V(\cdot, D)$ defined by

$$F_{\rm V}(u, D) = \int_D W_{\rm V}(\nabla u(x))\, dx, \tag{4}$$

where $W_V : \mathbb{M}^3(\mathbb{R}) \to \mathbb{R}^+$ is quasiconvex, satisfies a standard growth condition of order p, and is given by the asymptotic homogenization formula: for all $\Lambda \in \mathbb{M}^3(\mathbb{R})$,

$$W_{\rm V}(\Lambda) = \lim_{R\to\infty} \frac{1}{R^3} \inf\{F_1(u, Q_R), u \in \mathcal{S}(\mathcal{T}), u(x) = \Lambda \cdot x \text{ if } \operatorname{dist}(x, \partial Q_R) \le 2\mathbf{r}\}, \tag{5}$$

with $Q_R = (-R/2, R/2)^3$, almost surely.
In addition, if $\mathcal{L}$ is isotropic in the sense that for all rotations $\mathcal{R} \in SO(3)$, $\mathcal{L}$ and $\mathcal{R}(\mathcal{L})$ have the same statistics, then $W_{\rm V}$ is an isotropic energy density.

Let us comment on Theorem 1. When W_{nn} and $W_{\rm vol}$ are properly chosen (see [11]), the discrete model corresponds to a (simplified) thermodynamic description of polymer-chain networks. At the thermodynamic limit (that is when the typical size of the network goes to zero), the description of the material becomes continuous, and the obtained model is hyperelastic and characterized by the energy density W_V. Formula (5) is the starting point for numerical approximations of W_V (see Subsection 5.1). For the unfamiliar reader, from the above "Γ-convergence result" (and its generalization to the case when Dirichlet boundary conditions are considered, see [2]), we deduce that given a boundary value problem on D and a minimizer u_ε of the discrete energy at scale $\varepsilon > 0$, u_ε converges in $L^p(D)$ as $\varepsilon \to 0$ to a minimizer $u \in W^{1,p}(D)$ of the continuous energy with energy density $W_{\rm V}$, the same boundary conditions, and external loads. This fully justifies the passage to the limit as $\varepsilon \to 0$ in the static setting: not only the energy does converge, but also the minimizers. As shown in [11] the model obtained at the thermodynamic limit is in very good agreement with the classical mechanical experiments by Treloar.

As stated in Theorem 1, we do not only have an existence result for W_V, but also some qualitative properties of W_V: it is quasiconvex in the sense of Morrey (see for instance [7] for the application of quasiconvexity to nonlinear elasticity), and isotropic provided the stochastic network is isotropic. To complete this picture it is desirable to further characterize W_V, and ideally obtain an analytical formula (or at least proxy) which could be used in standard nonlinear elasticity softwares. As we shall see in the following subsection, one may derive additional properties for W_V, and restrict the general class of functions it belongs to.

2.2. **Structure properties of the homogenized energy density.** In [2], we have proved that provided the discrete model is frame-invariant and the stochastic network statistically isotropic, the homogenized model is hyperelastic, and the associated energy density W_V frame-invariant and isotropic. Energy densities satisfying these properties do have a specific structure. This is the famous Rivlin-Eriksen representation theorem (see for instance [7, Theorem 3.6-1]):

Theorem 2. *The energy density of a frame-invariant and isotropic hyperelastic material is a function of the principal invariants of the Cauchy-Green strain tensor. In particular, it is a symmetric function of the principal stretches of the strain tensor.*

As a consequence of Theorem 1, the energy density W_V is quasiconvex. Quasiconvexity is a property which ensures the lower-semicontinuity of integral functionals with respect to the weak topology of $W^{1,p}(D)$. Although it is a crucial property for abstract existence theories, quasiconvexity is a very difficult property to handle since it does not admit any local characterization (see [14]). We make a first simplification here by looking for approximations of W_V in a restricted class of quasiconvex functions: the class of polyconvex functions. Polyconvexity is a notion introduced by Ball in [3] which plays a fundamental role in the existence theory for nonlinear elasticity. It is an intermediate notion between quasiconvexity and convexity defined as follows.

Definition 2. *A function $W : \mathbb{M}^3(\mathbb{R}) \to \mathbb{R}^+$ is said to be polyconvex if $\Lambda \mapsto W(\Lambda)$ can be written as a convex function of Λ and of its minors.*

Recall that convex functions are polyconvex, and polyconvex functions are quasiconvex. Since a polyconvex function can be seen as a convex function in a larger space and since convexity is local (a twice-differentiable function is convex if and only if its Hessian is non-negative), polyconvexity is a local property, and is therefore easier to handle than quasiconvexity. As emphasized in [10] the specific form of the discrete model for rubber considered here (see Subsection 5.1) is closely related to polyconvexity. In particular if the affine assumption held, that is if the infimum in (5) was attained for $u(x) = \Lambda \cdot x$, W_V would be a polyconvex function. Yet, as illustrated in [11] the affine assumption does not hold.

Furthermore, unlike quasiconvexity, polyconvexity is not preserved by homogenization (see for instance [6, 4]). Hence it is not clear whether replacing quasiconvexity by polyconvexity is justified (see however the conjecture in [10]).

At this stage we would like to find an analytical approximation of W_V in the class of polyconvex, isotropic, and frame-invariant energy densities. To proceed we need a characterization of this manifold (that we shall denote by $\mathcal{P}$). Using the density of convex polynomials in the set of convex functions on bounded domains for the norm of the supremum, we directly deduce the density of polyconvex polynomials in the set of polyconvex functions. Yet it is not an easy task to characterize the set of polyconvex polynomials in terms of their coefficients, which is a huge handicap for the numerical practice of parameter identification. This is where Ogden's laws come into the picture.

2.3. **From homogenized energy densities to Ogden's constitutive laws.** As a first practical example of polyconvex functions, Ball considers in his seminal paper [3] the case of the Ogden laws introduced in [16] to model frame-invariant isotropic rubber materials. Relying on the Rivlin-Eriksen representation theorem, Ogden has proposed a restricted class of energy densities:

$$W_{\rm og}(\Lambda) = \sum_{i=1}^{k_1} a_i(\lambda_1^{\alpha_i}+\lambda_2^{\alpha_i}+\lambda_3^{\alpha_i})+\sum_{j=1}^{k_2} b_j\big((\lambda_1\lambda_2)^{\beta_j}+(\lambda_2\lambda_3)^{\beta_j}+(\lambda_3\lambda_1)^{\beta_j}\big)+W_3(\lambda_1\lambda_2\lambda_3), \tag{6}$$

where λ_1, λ_2 and λ_3 are the square-root of the eigenvalues of the Cauchy-Green strain tensor $\Lambda^T\Lambda$ (or singular values of Λ), $k_1, k_2 \in \mathbb{N}$, $a_i, b_i, \alpha_i, \beta_i \in \mathbb{R}$, and W_3 is a convex function. This class of constitutive laws is rather large, although its intersection with $\mathcal{P}$ is not dense in $\mathcal{P}$ for the topology of the local uniform convergence (see Section 4.2). The interest of Ogden's laws is the following: Ball has obtained in [3] (see also [7, Theorem 4.9-2]) a rather simple set of conditions which ensures the polyconvexity of Ogden's laws:

$$\forall\, i \in \{1,\dots,k_1\},\ a_i > 0, \alpha_i \geq 1, \quad \forall\, j \in \{1,\dots,k_2\},\ b_j > 0, \beta_j \geq 1. \tag{7}$$

This is of utmost interest to solve the identification problem with the constraint of polyconvexity (at least in this subclass of polyconvex Ogden's laws).

The second constraint requires that $W_{\rm og}$ be minimal at identity:

$$W_{\rm og}({\rm Id}) \;=\; \inf W_{\rm og}. \tag{8}$$

The class of functions in which we shall approximate $W_{\rm hom}$ is the following: Ogden's laws (6) satisfying the conditions (7) and (8), with $W_3 : (0, +\infty) \to \mathbb{R}$ given by

$$W_3(t) \;=\; K_1 t^2 - 2K_2 \log t \tag{9}$$

for some $K_1, K_2 \geq 0$, which is a variant of the Helmholtz energy density (replacing a single constant K by two constants K_1 and K_2 ensures that one can impose the identity to be a natural state of (6)). This manifold is rather complex. Let k_1 and k_2 be arbitrary yet fixed, and set $n = 2(k_1 + k_2 + 1)$. Given $p = (p_1, p_2) \in \mathbb{R}^n$ with $p_1 = (\alpha_1, \cdots, \alpha_{k_1}, \beta_1, \cdots, \beta_{k_2})$ and $p_2 = (K_1, K_2, a_1, \cdots, a_{k_1}, b_1, \cdots, b_{k_2})$, we denote by $W_{\rm og}^p$ the associated Ogden law.

3. Inverse problem and numerical solution method

In this section, we present a numerical method to solve the identification problem: we introduce a cost function and describe the evolutionary algorithm chosen to minimize this cost function.

3.1. **Available data, cost function and integration rule.** In order to find the Odgen constitutive law (6) which best approximates $W_{\rm V}$, we have to identify the corresponding set of parameters $p = (p_1, p_2) \in \mathbb{R}^n$. This is done by minimizing a well-chosen cost function $\mathcal{E}$ depending on the difference between quantities associated with the energy density $W_{\rm og}^p$ (seen as a function of p) and the corresponding quantities obtained by "in silico experiments" (namely numerical approximations of $W_{\rm V}$).

Since we are interested in boundary value problems, the important quantity is not the energy density $W_{\rm V}$ itself, but rather its derivatives. The numerical approximation method of the discrete model allows one to evaluate $W_{\rm V}(\Lambda)$ at any deformation gradient Λ. Since it relies on a Newton algorithm, this method provides as outputs approximations of:

$$\Pi_{\rm V}(\Lambda) = \frac{\partial W_{\rm V}}{\partial \Lambda}, \qquad \text{the first Piola-Kirchhoff stress tensor and}$$

$$H_{\rm V}(\Lambda) = \frac{\partial^2 W_{\rm V}}{\partial \Lambda^2}, \qquad \text{the elasticity tensor (Hessian).}$$

These approximations are the observations for the inverse problem. Since the material is isotropic and frame-invariant, $W_{\rm V}$ is characterized by its values on diagonal matrices $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)$ and we only consider such deformation gradients. In particular, the associated Piola-Kirchhoff stress tensor is diagonal and we denote by $\{\tilde{\Pi}_i\}_{1\le i\le 3}$ its diagonal entries. We also denote by $\{\tilde{H}_{ij}\}_{1\le i,j\le 3}$ the entries $\frac{\partial^2 W_{\rm V}}{\partial \Lambda_{ij}^2}$ of the Hessian (which is positive for strongly elliptic materials).

With this preliminary, we are in position to introduce the cost function we shall consider:

$$c(p;\lambda_1,\lambda_2,\lambda_3) := \frac{\sum_{i=1}^3 \left(\Pi_i^p - \tilde{\Pi}_i\right)^2}{\sum_{i=1}^3 \tilde{\Pi}_i^2} + \eta \sum_{i=1}^3 \sum_{j=1}^3 \left(\frac{H_{ij}^p - \tilde{H}_{ij}}{\tilde{H}_{ij}}\right)^2, \tag{10}$$

where $\eta \ge 0$ is a (small) regularization parameter, and Π^p and H^p stand for the first Piola-Kirchhoff stress tensor and Hessian associated with the Ogden law $W_{\rm og}^p$ of parameter p.

In the cost function, we have chosen to restrain the values of the admissible deformations $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)$ to entries typically in the intervals $(1/6, 6)$ (deformation up to 500%) and $(1/3, 3)$ (deformation up to 300%). In the rest of this paragraph we focus on the interval $(1/6, 6)$. In order to give a bigger weight to strain gradients in small deformations (λ_i close to 1), we introduce the following importance measure on $(1, 6)$:

$$\mu_\kappa(x) = K_\kappa (x-1)^\kappa, \tag{11}$$

where $\kappa > -1$ and K_κ is a normalization factor such that :

$$\int_1^6 \mu_\kappa(x)dx = 1, \quad \text{that is} \quad K_\kappa = \frac{1+\kappa}{5^{1+\kappa}}. \tag{12}$$

The parameter κ can be chosen such that, for some fixed $x^0 \in (1, 6)$, we have

$$\int_1^{x^0} \mu_\kappa(x)dx = \int_{x^0}^6 \mu_\kappa(x)dx, \quad \text{which implies} \quad \kappa = \frac{\log(2)}{\log(5) - \log(x^0 - 1)} - 1, \tag{13}$$

that is the weights given to the intervals $(1, x^0)$ and $(x^0, 6)$ are the same. This is a convenient way to give more importance to the small deformation regime. For the applications, we set $x_0 = 1.5$. On Figure 1, the weight function $(x-1)^\kappa$ is plotted for different values of κ. The formula (11) defines μ_κ on $(1, 6)$ and we extend this function to $(1/6, 1)$ by setting $\mu_\kappa(1/x) = \mu_\kappa(x)$ for all $x \in (1, 6)$. Thus, the measure μ_κ gives equal weight to compression and extension.

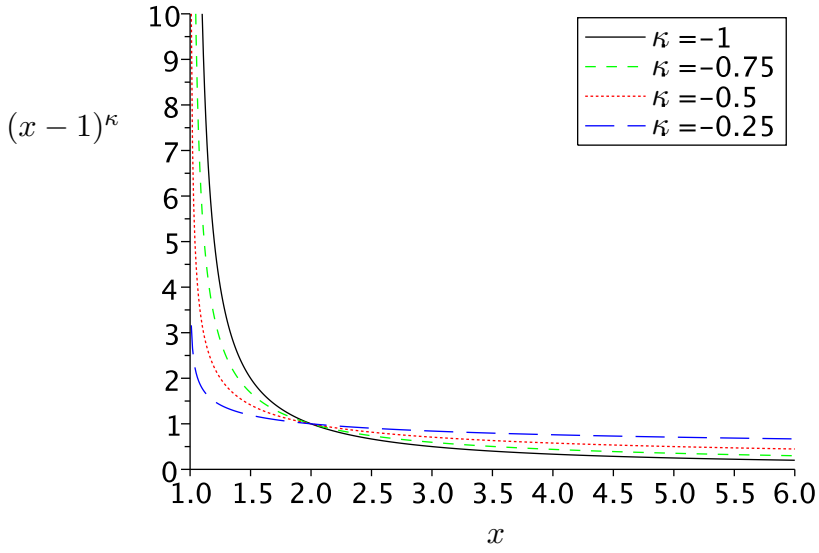


FIGURE 1. Jacobi weight function $x \mapsto (x-1)^\kappa$ (for $\kappa \in \{-1, -0.75, -0.5, -0.25\}$)

In the case of quasi-incompressible materials, the nonlinear constraint $\det \Lambda = \lambda_1\lambda_2\lambda_3 \simeq 1$ has to be taken into account. To this end, we consider the reduced principal strains

$$J = \lambda_1\lambda_2\lambda_3, \qquad \tilde{\nu}_1 = \frac{\lambda_1}{J^{1/3}}, \quad \tilde{\nu}_2 = \frac{\lambda_2}{J^{1/3}}, \quad \text{and} \quad \tilde{\nu}_3 = \frac{\lambda_3}{J^{1/3}}.$$

As primary variables we take J, $\nu_1 = \frac{1}{\tilde{\nu}_1}$ and $\nu_2 = \tilde{\nu}_2$, and restrict J to $(1-\delta, 1+\delta)$, for some small $\delta > 0$. Finally, we define the following global cost function:

$$\mathcal{F}(p) = \left(\frac{\displaystyle\int_1^6 \int_1^6 \int_{1-\delta}^{1+\delta} c\big(p; J^{1/3}\frac{1}{\nu_1}, J^{1/3}\nu_2, J^{1/3}\frac{\nu_1}{\nu_2}\big)\mu_\kappa(\nu_1)\mu_\kappa(\nu_2)dJd\nu_1d\nu_2}{\displaystyle\int_1^6 \int_1^6 \int_{1-\delta}^{1+\delta} \mu_\kappa(\nu_1)\mu_\kappa(\nu_2)dJd\nu_1d\nu_2} \right)^{1/2}. \tag{14}$$

Note that, by symmetry of c, we take into account all the possible deformation gradients $\Lambda = J^{1/3}\mathrm{diag}(\tilde{\nu}_1, \tilde{\nu}_2, \tilde{\nu}_3)$ with $\nu_1, \nu_2, \nu_3 \in (1/6, 6)$ such that $\tilde{\nu}_1\tilde{\nu}_2\tilde{\nu}_3 = 1$. The identification problem consists in finding $p \in \mathbb{R}^n$ that minimizes $\mathcal{F}$ under the constraints (7) and (8).

We then approximate the integral in J with a standard three points integration rule and the integrals in ν_1 and ν_2 by the Jacobi integration rule of order $m \in \mathbb{N}$, that is for all f continuous,

$$\int_1^6 f(x)(x-1)^\kappa dx \simeq \sum_{k=1}^m \omega_k f(x_k), \tag{15}$$

where x_k, for $k \in \{1, \dots, m\}$, are the roots of the $(0,\kappa)$-Jacobi polynomial of degree m (after mapping $(-1,1)$ to $(1,6)$) and ω_k are the corresponding weights.

The minimization problem we are considering is then as follows:

$$\inf \Big\{ \mathcal{F}(p) \,\Big|\, p = (p_1, p_2) \in \mathbb{R}^n, p_1 = (\alpha_1, \cdots, \alpha_{k_1}, \beta_1, \cdots, \beta_{k_2}), \\ p_2 = (K_1, K_2, a_1, \cdots, a_{k_1}, b_1, \cdots, b_{k_2}) : \\ \alpha_i, \beta_i \geq 1, a_i, b_i, K_1, K_2 \geq 0,\ W^p_{\mathrm{og}}(\mathrm{Id}) = \inf W^p_{\mathrm{og}} \Big\}.$$

3.2. **Numerical method.** The minimization problem considered here has the following three properties:

- the cost function $\mathcal{F}$ is not convex and defines a rather complex energy landscape with numerous local minima;
- the cost function $\mathcal{F}$ is twice-differentiable;
- the set of parameters is a nonlinear implicitly defined manifold.

Due to the complexity of the energy landscape, Newton-type algorithms get trapped into local infima. To illustrate this difficulty, we display in Table 1 the outputs of a Newton algorithm for the reconstruction of an Ogden law (without the nonlinear constraint that the function be minimal at identity), depending on the initial guess. As can be seen, if the initial guess is close to the solution, then the algorithm converges. If not, the algorithm does not always converge, and when it does, it is not necessarily to the right solution.

b_1^0	β_1^0	Newton algorithm	K_1	K_2	a_1	α_1	a_2	α_2	b_1	β_1
		exact coefficients	10.	10.	4.8	1.8	0.24	5.1	0.05	2.3
0.05	3.0	converges to the global minimum	10.	10.	4.8	1.8	0.24	5.1	0.05	2.3
0.05	5.3	converges to a local minimum	9.9	7.2	0.9	4.3	10^{-7}	4.3	10^{-7}	4.4
0.05	4.0	does not converge in the admissible set	11.4	6.7	10^{-7}	1.0	10^{-7}	1.0	10^{-7}	1.0
0.04	5.3	diverges	∞	∞	∞	∞	∞	∞	∞	∞

TABLE 1. Outputs of a Newton algorithm for the reconstruction of an Ogden law depending on the initial values b_1^0 and β_1^0

In order to circumvent (or at least reduce) this difficulty, we shall appeal to a stochastic optimization procedure — namely, an evolutionary algorithm.

The general procedure is as follows. Assume momentarily that we minimize $\mathcal{F}$ without the constraints (7) and (8), and that the minimizer $\bar{p}$ of $\mathcal{F}$ on $\mathbb{R}^n$ is unique. Instead of directly looking for $\bar{p}$ in $\mathbb{R}^n$, we look for a probability measure $\bar{\mu} \in \mathcal{M}(\mathbb{R}^n)$ which minimizes $\mu \mapsto \int_{\mathbb{R}^n} \mathcal{F}(p) d\mu(p)$. Of course, by uniqueness of the minimizer, $\bar{\mu} = \delta_{\bar{p}}$, the Dirac mass at $\bar{p}$. The strategy is now to approximate $\bar{\mu}$ by a sequence of Gaussian measures G_k, characterized by their means $m_k \in \mathbb{R}^n$, their covariance matrices $C_k \in SO_n(\mathbb{R})$, and their standard deviation $\sigma_k \in \mathbb{R}^+$. The sequence G_k is an approximation of $\bar{\mu}$ if $\lim_{k\to\infty} G_k = \bar{\mu}$ weakly in the sense of measures.

The evolutionary algorithm is characterized by its updating procedure, that is the construction of G_{k+1} knowing G_k. Let $S > s > 0$ be integers. Given a sampling $p_1^k, \dots, p_S^k$ of G_k, we select the s best search points p_i^k (that is those s points among the S search points which yield the s minimal values of $\mathcal{F}$). The mean m_{k+1} of G_{k+1} is then obtained by taking a (suitable) weighted average of the s best search points, the covariance matrix C_{k+1} is chosen so that the s search points are a suitable sampling of a Gaussian measure with this covariance matrix C_{k+1}. It remains to choose the standard deviation σ_{k+1}. The larger σ_{k+1}, the more regions of the energy landscape will be potentially visited. Yet, $G_k \to \bar{\mu}$ requires $\sigma_k \to 0$. The choice of σ_{k+1} is therefore crucial. Once G_{k+1} is defined, one generates randomly S samples p_i^{k+1}. We use the Covariance Matrix Adaptation Evolutionary Strategy (CMA-ES) algorithm. For the precise update of G_k, we refer the reader to the Appendix and to [13].

Let us now describe precisely how the CMA-ES algorithm is *used* in our context. We first neglect the constraint (8) that the Odgen energy density be minimal at identity. The cost function $\mathcal{F}$ has a very specific structure, and the parameters $p_1 = (\alpha_1, \cdots, \alpha_{k_1}, \beta_1, \cdots, \beta_{k_2})$ and $p_2 = (K_1, K_2, a_1, \cdots, a_{k_1}, b_1, \cdots, b_{k_2})$ do have different roles. In particular, since the dependence of $\mathcal{F}$ upon p_2 is quadratic (after taking $\mathcal{F}$ to the square), it makes sense to consider the *reduced* cost function

$$\mathcal{F}_r(p_1) := \inf_{p_2 \geq 0} \mathcal{F}(p_1, p_2).$$

The infimum can indeed be effectively computed by deterministic methods (recall that we have neglected the constraint (8) that $W^p_{\rm og}$ is minimal at identity). Since

$$\inf_{p\in\mathbb{R}^n \text{ s. t. } (7)} \mathcal{F}(p) \;=\; \inf_{p_1 \text{ s. t. } (7)} \mathcal{F}_r(p_1),$$

one may either apply the CMA-ES algorithm to $\mathcal{F}$ or $\mathcal{F}_r$. There are two main differences: the nonlinearity of the functionals ($\mathcal{F}_r$ is much more nonlinear than $\mathcal{F}$ since the minimizers p_2 are themselves nonlinear functions of p_1) and the dimension of the parameters (k_1+k_2 for $\mathcal{F}_r$, $2(k_1+k_2+1)$ for $\mathcal{F}$). In both cases we impose the constraint (7) on the parameters p_1 by penalization so that the search space remains the linear space $\mathbb{R}^{k_1+k_2}$, and not $\{p\in\mathbb{R}^{k_1+k_2} \text{ such that } (7)\}$.

This picture would be complete if we did not have to deal with the constraint (8) that $W^p_{\rm og}$ be minimal at identity. In order to take this constraint into account, we use a splitting method, and add a projection step to the algorithm: we proceed by prediction-correction to take into account the constraint (8). This amounts to minimizing a different functional $\overline{\mathcal{F}}_r$ defined as follows. Given p_1, we let $\tilde{p}_2$ be a minimizer of $\mathcal{F}(p_1,\cdot)$ on $[0,+\infty)^{k_1+k_2+2}$, and set $\tilde{p}=(p_1,\tilde{p}_2)\in\mathbb{R}^n$. If $W^{\tilde{p}}_{\rm og}$ satisfies (8), we set $\overline{\mathcal{F}}_r(p_1) := \mathcal{F}_r(p_1)$. Otherwise, we "project" $W^{\tilde{p}}_{\rm og}$ on the set of Ogden laws satisfying (8). To this aim, we let $\gamma>0$ be the unique minimizer of $t\mapsto W^{\tilde{p}}_{\rm og}(t{\rm Id})$ on $\mathbb{R}$, which we may compute by a Newton algorithm (the problem is strictly convex), and finally define

$$\overline{\mathcal{F}}_r(p_1) \;:=\; \mathcal{F}(p_1,p_2),$$

where $p_2 := (\gamma^6 K_1, K_2, \gamma^{\alpha_1}a_1, \cdots, \gamma^{\alpha_{k_1}}a_{k_1}, \gamma^{2\beta_1}b_1, \cdots, \gamma^{2\beta_{k_2}}b_{k_2})$. Setting $p=(p_1,p_2)$, this ensures that $W^p_{\rm og}$ satisfies the minimality condition at identity (8).

The splitting method amounts to minimizing the functional $p_1\mapsto\overline{\mathcal{F}}_r(p_1)$ on $[1,+\infty)^{k_1+k_2}$ by the evolutionary algorithm. For all $p_1\in\mathbb{R}^{k_1+k_2}$, $\overline{\mathcal{F}}_r(p_1)\geq\mathcal{F}_r(p_1)$ by definition. Hence it is not clear whether minimizing $\overline{\mathcal{F}}_r$ is equivalent to minimizing $\mathcal{F}_r$. This is the case if any minimizer p_1 of $\mathcal{F}_r$ satisfies the identity

$$\inf_{p_2 \text{ s.t. } (7)\&(8)} \mathcal{F}(p_1,p_2) \;=\; \inf_{p_2 \text{ s.t. } (7)} \mathcal{F}(p_1,p_2). \tag{16}$$

When this condition does not hold, the splitting procedure only gives an upper bound on the minimum.

Let us make a further comment: nothing prevents the algorithm to split a term $a_i(\lambda_1^{\alpha_i}+\lambda_2^{\alpha_i}+\lambda_3^{\alpha_i})$ into two terms $\tau a_i(\lambda_1^{\alpha_i}+\lambda_2^{\alpha_i}+\lambda_3^{\alpha_i})+(1-\tau)a_i(\lambda_1^{\alpha_i}+\lambda_2^{\alpha_i}+\lambda_3^{\alpha_i})$ for any $\tau\geq 0$. This issue is more pronounced for the minimization of the $\mathcal{F}$ than for the minimization of $\mathcal{F}_r$.

Since the algorithm is stochastic, one may perform several independent realizations. To enhance the precision, one may use the values of the parameters obtained at convergence of the evolutionary algorithm as initial guess for a Newton algorithm. The CMA-ES algorithm would then somehow only be used as a "black-box" to select the right region of the energy landscape to focus on — this coupling of the deterministic and evolutionary algorithms is minimal.

4. Calibration and first tests of the method

4.1. **Exact law.** In this section we test the data assimilation algorithms on the example of Ogden's laws (6). In particular, we consider an Odgen law, generate the associated data and try to recover the coefficients of this law by minimizing $\mathcal{F}$ and $\mathcal{F}_r$ using the evolutionary algorithm. The aim of this section is to illustrate four facts:

- The minimization of the reduced error functional $\mathcal{F}_r$ yields more precise results than the minimization of $\mathcal{F}$, and allows one to recover more coefficients than with $\mathcal{F}$ (recall that for the same law, $\mathcal{F}_r$ has less unknowns than $\mathcal{F}$).

- For general laws with less than 14 coefficients to be fitted, an integration rule with $m = 3$ is sufficient.
- The use of the Hessian significantly improves the reconstruction.
- The inverse problem is stable in the sense that it does not amplify errors (errors on the recovered coefficients are of the same order than errors on the sampling of the law).

In a first series of tests, we compare the minimization of $\mathcal{F}_r$ and $\mathcal{F}$, using $\eta = 0$ and $m = 3$. In Tables 2 and 3, the first line displays the coefficients of the exact law to recover. The row error gives the value of $\mathcal{F}$ (or $\mathcal{F}_r$) at convergence, in function of the number of coefficients k_1 and k_2 used for the reconstruction. The complete procedure is as follows: we iteratively increase k_1 and k_2 in order to reduce the error $\mathcal{F}$ (or $\mathcal{F}_r$), and stop when the error stops decreasing (and at comparable final error we choose the law with the less coefficients). As can be seen on Tables 2 and 3, the minimization of $\mathcal{F}_r$ indeed yields more precise coefficients and avoids the issue of splitting a single term of the law into two terms (as it is the case in the last line of Table 2 for α_2 and α_3).

error $\mathcal{F}$	k_1	k_2	K_1	K_2	a_1	α_1	a_2	α_2	a_3	α_3	b_1	β_1	b_2	β_2
exact	2	1	5.86	10.0	4.09	1.80	$1.52 \cdot 10^{-1}$	5.10			$3.32 \cdot 10^{-2}$	2.30		
$2 \cdot 10^{-1}$	1	0	6.12	8.21	1.00	3.92								
$1 \cdot 10^{-1}$	1	1	4.25	12.6	$3.53 \cdot 10^{-1}$	4.61					7.44	1.00		
$8 \cdot 10^{-3}$	2	0	5.90	10.0	4.14	1.80	$1.54 \cdot 10^{-1}$	5.10						
$7 \cdot 10^{-10}$	2	1	5.86	10.0	4.09	1.80	$1.52 \cdot 10^{-1}$	5.10			$3.32 \cdot 10^{-2}$	2.30		
$8 \cdot 10^{-10}$	2	2	5.86	10.0	4.09	1.80	$1.52 \cdot 10^{-1}$	5.10			$3.32 \cdot 10^{-2}$	2.30	$2.13 \cdot 10^{-7}$	1.20
$9 \cdot 10^{-10}$	3	1	5.86	10.0	4.09	1.80	$1.51 \cdot 10^{-1}$	5.10	$1.96 \cdot 10^{-3}$	5.16	$3.32 \cdot 10^{-2}$	2.30		

TABLE 2. Data assimilation on an Ogden law by minimizing $\mathcal{F}$ ($\eta = 0$ and $m = 3$).

error $\mathcal{F}_r$	k_1	k_2	K_1	K_2	a_1	α_1	a_2	α_2	a_3	α_3	b_1	β_1	b_2	β_2
exact	2	1	5.86	10.0	4.09	1.80	$1.52 \cdot 10^{-1}$	5.10			$3.32 \cdot 10^{-2}$	2.30		
$2 \cdot 10^{-1}$	1	0	6.12	8.21	1.00	3.92								
$1 \cdot 10^{-1}$	1	1	4.25	12.6	$3.53 \cdot 10^{-1}$	4.61					7.44	1.00		
$8 \cdot 10^{-3}$	2	0	5.90	10.0	4.14	1.80	$1.54 \cdot 10^{-1}$	5.10						
$6 \cdot 10^{-10}$	2	1	5.86	10.0	4.09	1.80	$1.52 \cdot 10^{-1}$	5.10			$3.32 \cdot 10^{-2}$	2.30		
$6 \cdot 10^{-10}$	2	2	5.86	10.0	4.09	1.80	$1.52 \cdot 10^{-1}$	5.10			$3.32 \cdot 10^{-2}$	2.30	$7.59 \cdot 10^{-9}$	4.51
$6 \cdot 10^{-10}$	3	1	5.86	10.0	4.09	1.80	$1.52 \cdot 10^{-1}$	5.10	$3.69 \cdot 10^{-6}$	4.64	$3.32 \cdot 10^{-2}$	2.30		

TABLE 3. Data assimilation on an Ogden law by minimizing $\mathcal{F}_r$ ($\eta = 0$ and $m = 3$).

On Tables 5 and 6 we report the errors $\mathcal{F}_r$ and $\mathcal{F}$ at convergence and the maximum and the mean of the relative errors on the coefficients of the best approximation, in function of the number of terms $2(1 + k_1 + k_2)$ of the exact law. Coefficients reported on Table 4 are chosen such that for all $1 \leq i \leq k_1$ and for all $1 \leq j \leq k_2$, the terms $a_i x_0^{\alpha_i}$ and $b_j x_0^{\alpha_j}$ are of the same order of magnitude. As expected, minimizing $\mathcal{F}_r$ instead of $\mathcal{F}$ allows one to recover correctly Ogden laws with larger k_1 and k_2, although we cannot go beyond $k_1 = k_2 = 4$ (that is, 18 coefficients) in practice.

K_1	a_1	a_2	a_3	a_4	a_5	a_6	b_1	b_2	b_3	b_4	b_5	b_6
2.46	5.60	3.63	2.37	1.646	0.956	0.632	2.04	1.30	0.557	0.209	0.0783	0.0548
K_2	α_1	α_2	α_3	α_4	α_5	α_6	β_1	β_2	β_3	β_4	β_5	β_6
9.22	1.43	2.5	3.55	4.46	5.79	6.81	1.96	2.52	3.56	4.77	5.98	6.42

TABLE 4. Coefficients used for validation tests.

k_1	k_2	error $\mathcal{F}$	error max	mean error	error $\mathcal{F}_r$	error max	mean error
1	1	$6 \cdot 10^{-10}$	0.0%	0.0%	0	0.0%	0.0%
2	2	$6 \cdot 10^{-10}$	0.0%	0.0%	0	0.0%	0.0%
3	3	$7 \cdot 10^{-7}$	31%	4.6%	$3 \cdot 10^{-11}$	0.3%	0.04%
4	4	$1 \cdot 10^{-6}$	65%	21%	$3 \cdot 10^{-7}$	92%	18%
5	5	$8 \cdot 10^{-7}$	88%	18%	$1 \cdot 10^{-7}$	100%	32%

TABLE 5. Data assimilation on an Ogden law by minimizing $\mathcal{F}$ and $\mathcal{F}_r$ (with $\eta = 0$ and $m = 3$).

k_1	k_2	error $\mathcal{F}$	error max	mean error	error $\mathcal{F}_r$	error max	mean error
1	1	$3 \cdot 10^{-10}$	0.0%	0.0%	$6 \cdot 10^{-8}$	0.0%	0.0%
2	2	$3 \cdot 10^{-10}$	0.0%	0.0%	$2 \cdot 10^{-8}$	0.0%	0.0%
3	3	$2 \cdot 10^{-10}$	0.0%	0.0%	$1 \cdot 10^{-8}$	0.0%	0.0%
4	4	$8 \cdot 10^{-7}$	99%	13%	$1 \cdot 10^{-8}$	0.0%	0.0%
5	5	$5 \cdot 10^{-7}$	98%	21%	$3 \cdot 10^{-7}$	75%	19%

TABLE 6. Data assimilation on an Ogden law by minimizing $\mathcal{F}$ and $\mathcal{F}_r$ (with $\eta = 1$ and $m = 5$).

In a second series of tests we have checked the influence of m and η on the recovery of the coefficients. To this aim we have chosen an Ogden law with $k_1 = 4$ and $k_2 = 4$ (that is with the 18 coefficients given in Table 4). We report on Table 7 the maximum of the relative errors on the coefficients obtained by minimizing $\mathcal{F}_r$ in function of m and η for $m \in \{3, 5\}$ and $\eta \in \{0, 10^{-2}, 10^{-1}, 1\}$. Both parameters have a significant influence on the precision of the reconstruction.

m \ η	0	10^{-2}	10^{-1}	1
3	92%	2.5%	0.4%	0.3%
5	32%	0.3%	0.2%	0.0%

TABLE 7. Relative error on the coefficients by minimizing $\mathcal{F}_r$ in function of η and m (for $k_1 = k_2 = 4$).

In order to check the robustness of the algorithm, we have performed the same tests with the addition of noise on the data (by independent and identically distributed random variables on each data, of the order of a few percents). This is an important property since the data we shall use to reconstruct the micro-macro law are noisy (in the sense that we have random realizations and numerical errors). The results for typical realizations are diplayed in Table 8. These tests show that the algorithm is rather robust, and that the problem is stable.

noise	error $\mathcal{F}_r$	K_1	K_2	a_1	α_1	a_2	α_2	b_1	β_1
	exact	5.86	10.0	4.09	1.80	$1.52 \cdot 10^{-1}$	5.10	$3.32 \cdot 10^{-2}$	2.30
0	$4 \cdot 10^{-10}$	0%	0%	0%	0%	0%	0%	0%	0%
2%	$9 \cdot 10^{-3}$	0.6%	0.8%	2.5%	1%	1.8%	0.2%	0.2%	1.9%
5%	$3 \cdot 10^{-2}$	1.6%	0.5%	1.7%	1.6%	5.1%	0.7%	35%	2.5%
10%	$5 \cdot 10^{-2}$	1.7%	1.2%	0.2%	2.8%	14.6%	1.7%	45%	17%

TABLE 8. Recovery process with noisy data (with $m = 3$, $\eta = 0$, $k_1 = 2$ and $k_2 = 1$).

4.2. **Arruda-Boyce model.** As a first example of an energy density which does not belong to the subspace $\mathcal{P}$ of Ogden laws, we consider the Arruda-Boyce model — which is known to be in rather good agreement with the Treloar data on natural rubber. This constitutive law is analytical and given by

$$\mathbb{M}^3(\mathbb{R}) \ni \Lambda \mapsto W_{\rm AB}(\Lambda) = \frac{n}{\beta}\left(\frac{1}{2}I_1 + \frac{1}{20N}I_1^2 + \frac{1}{1050N^2}I_1^3 + \frac{1}{7000N^3}I_1^4 + \frac{1}{673750N}I_1^5\right) + W_3(\det\Lambda), \tag{17}$$

where $I_1 = \operatorname{trace}\Lambda^T\Lambda$ and W_3 is as in (9) for some $n, \beta, N \in \mathbb{R}^+$, and $K_1, K_2 \geq 0$.

Typical coefficients are displayed on Table 9 (recall that the coefficients are such that Id is a natural state)

K_1	K_2	n/β	N
2.439	2.575	0.2652	26.98
2.442	2.575	0.2653	254.4

TABLE 9. Coefficients for the Arruda-Boyce model.

Before we turn to the numerical tests, let us mention that the first two terms of the Arruda-Boyce constitutive law (17) can be written in the form of an Ogden law since $I_1 = \lambda_1^2 + \lambda_2^2 + \lambda_3^2$ (where the λ_i are the eigenvalues of the matrix $\Lambda^T\Lambda$). The coefficients are given by

$$\alpha_1 = 2,\ a_1 = \frac{n}{2\beta},\ \alpha_2 = 4,\ a_2 = \frac{n}{\beta}\frac{1}{20N},\ \beta_1 = 2,\ b_1 = \frac{n}{\beta}\frac{1}{10N}.$$

In particular, the higher N is, the closer the Arruda-Boyce is to an Ogden law. We therefore expect the coefficients obtained by the reconstruction procedure to be close to these coefficients (at least for large N). Since we do not expect the function

$$(\lambda_1, \lambda_2, \lambda_3) \mapsto \lambda_1^2\lambda_2^4 + \lambda_1^2\lambda_3^4 + \lambda_2^2\lambda_1^4 + \lambda_2^2\lambda_3^4 + \lambda_3^2\lambda_1^4 + \lambda_3^2\lambda_2^4$$

to be well-approximated by positive linear combinations of terms of the type $\lambda_1^\alpha + \lambda_2^\alpha + \lambda_3^\alpha$ and $(\lambda_1\lambda_2)^\beta + (\lambda_2\lambda_3)^\beta + (\lambda_3\lambda_1)^\beta$, it is likely that the Arruda-Boyce model cannot be approximated *arbitrarily* close in $\mathcal{P}$ for small to moderate N — recall that

$$(\lambda_1, \lambda_2, \lambda_3) \mapsto I_1^3 = \lambda_1^6+\lambda_2^6+\lambda_3^6+3\lambda_1^2\lambda_2^4+3\lambda_1^2\lambda_3^4+3\lambda_2^2\lambda_1^4+3\lambda_2^2\lambda_3^4+3\lambda_3^2\lambda_1^4+3\lambda_3^2\lambda_2^4+3\lambda_1^2\lambda_2^2\lambda_3^2.$$

This could be a limitation of the choice of the Ogden laws.

The results of the data assimilation algorithm for the Arruda-Boyce constitutive laws of Table 9 are displayed on Tables 10 and 11. The tests confirm that the higher N, the better the reconstruction. We observe that the quality of the reconstruction saturates for $k_1 = 2$ and $k_2 = 1$ (and we do not display the results for higher k_1, k_2).

In order to illustrate the error on mechanically meaningful quantities, we have plotted the stress tensor of the Arruda-Boyce constitutive law and of the associated Ogden law (obtained by the reconstruction procedure) for two typical mechanical experiments: uniaxial deformation (see Figure 2) and pure shear (see Figure 3). The results are in excellent agreement and Ogden's laws capture very well the behavior of the Arruda-Boyce model, even for moderate N (whereas this is not guaranteed theoretically).

error $\mathcal{F}_r$	k_1	k_2	K_1	K_2	a_1	α_1	a_2	α_2	a_3	α_3	b_1	β_1
$2 \cdot 10^{-2}$	1	0	2.44	2.57	$1.25 \cdot 10^{-1}$	2.10						
$1 \cdot 10^{-2}$	1	1	2.44	2.58	$1.18 \cdot 10^{-1}$	2.12					$1.40 \cdot 10^{-2}$	1.00
$9 \cdot 10^{-5}$	2	1	2.44	2.57	$1.33 \cdot 10^{-1}$	2.00	$4.30 \cdot 10^{-4}$	4.08			$9.70 \cdot 10^{-4}$	2.02

TABLE 10. Approximation of the Arruda-Boyce model (with $n/\beta = 0.2652, N = 26.98, K_1 = 2.439, K_2 = 2.575$) by Ogden laws, using $m = 3$ and $\eta = 0$.

error $\mathcal{F}_r$	k_1	k_2	K_1	K_2	a_1	α_1	a_2	α_2	a_3	α_3	b_1	β_1
$2 \cdot 10^{-3}$	1	0	2.44	2.57	$1.32 \cdot 10^{-1}$	2.01						
$2 \cdot 10^{-3}$	1	1	2.44	2.58	$1.31 \cdot 10^{-1}$	2.01					$1.41 \cdot 10^{-3}$	1.00
$2 \cdot 10^{-9}$	2	1	2.44	2.57	$1.33 \cdot 10^{-1}$	2.00	$5.15 \cdot 10^{-5}$	4.01			$1.04 \cdot 10^{-4}$	2.00

TABLE 11. Approximation of the Arruda-Boyce model (with $n/\beta = 0.2653, N = 254.4, K_1 = 2.442, K_2 = 2.575$) by an Ogden laws, using $m = 3$ and $\eta = 0$.

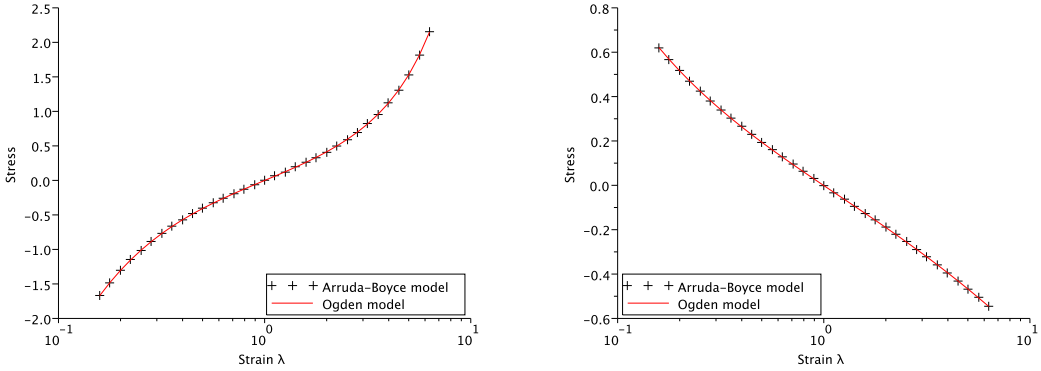


FIGURE 2. *Comparison of Arruda-Boyce and Ogden models (N=26.98) in uniaxial deformation* $\Lambda = diag(\frac{1}{\sqrt{\lambda}}, \frac{1}{\sqrt{\lambda}}, \lambda)$. *Piola Kirchhoff stress components* Π_3 *(on the left) and* $\Pi_1 = \Pi_2$ *(on the right) in function of* λ *in log-scale.*

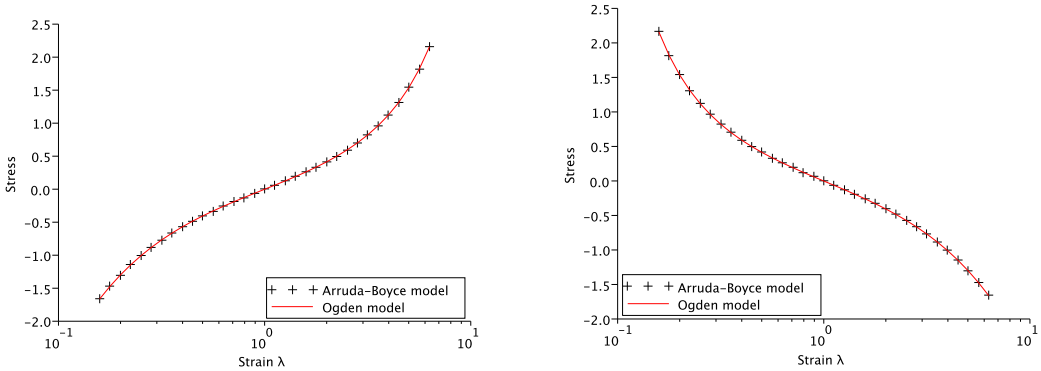


FIGURE 3. *Comparison of Arruda-Boyce and Ogden models (N=26.98) in pure shear* $\Lambda = diag(\lambda, \frac{1}{\lambda}, 1)$. *Piola Kirchhoff stress components* Π_1 *(on the left) and* Π_2 *(on the right) in function of* λ *in log-scale.*

5. Application to the micro-macro model for rubber

In this section we turn to the main application of the data assimilation procedure introduced in this article: the analytical reconstruction of the micro-macro constitutive law of [11].

5.1. **Generation of the data.** Our aim is to approximate $\Lambda \mapsto W_{\rm V}(\Lambda)$ with the help of (a variant of) formula (5):

$$W_{\rm V}(\xi) = \lim_{R\to\infty} \frac{1}{|B(R)|} \inf\{F_1(u,B_R), u \in \mathcal{S}(\mathcal{T}), u(x) = \xi\cdot x \text{ if } \mathrm{dist}(x,\partial B_R) \le 2\mathbf{r}\}, \tag{18}$$

with W_{nn} given by

$$W_{nn}(|e|,\lambda) = \frac{n}{\beta} N_e \left(\frac{\lambda}{\sqrt{N_e}} \theta\left(\frac{\lambda}{\sqrt{N_e}}\right) + \log \frac{\theta\left(\frac{\lambda}{\sqrt{N_e}}\right)}{\sinh \theta\left(\frac{\lambda}{\sqrt{N_e}}\right)} \right), \tag{19}$$

where θ is the inverse of the Langevin function, and N_e related to e through $N_e = |e|^2$, $W_{\rm vol}(\det\Lambda) = W_{\rm Helm}(\Lambda)$ given by

$$W_{\rm Helm}(\Lambda) = K(\det(\Lambda)^2 - 1 - 2\log(\det(\Lambda))), \tag{20}$$

and where $B(R)$ is the ball of radius $R>0$ centered at the origin. The physical parameters to be fixed are the typical length $\overline{e}$ of the edges of the network (that is, the average of the lengths of all the edges of the tessellation), the typical number of monomers $\overline{N}$ per chain (related to the characteristic length l of a monomer through the identity $\overline{N} = \left(\frac{\overline{e}}{l}\right)^2$) and $\frac{n}{\beta}$ in (19), and K in (20). We refer the reader to [11] for the physical and mechanical justifications of the model.

Typical values in the simulations are chosen in function of R so that the ratio of the contributions of the chains and of the volumetric effect is constant (this does not change the value of $W_{\rm V}$ at the limit). We denote by N_R and by $|B(R)|$ the number of edges and the volume of $B(R)$, respectively. We will take:

$$\overline{N} \in \{26.5, 250\}, \quad \frac{n}{\beta} = \frac{|B(R)|}{N_R} \times \frac{26.5}{\overline{N}} \times 0.27\,\mathrm{MPa}, \quad K = 2.575\,\mathrm{MPa}. \tag{21}$$

This choice of $\frac{n}{\beta}$ makes the contribution to the energy of the $N_R\overline{N}$ monomers be of order $|B_R|0.27\,\mathrm{MPa}$ (such a rescaling makes the number of monomers per unit volume to fixed in the simulations — which is needed to compare polymers with different $\overline{N}$ since the network is not at the packing limit, see [11, Remark 2]). These correspond to the typical values considered in the Arruda-Boyce model to fit Treloar's experiments.

For the numerical approximation procedure one has to pick a large — though finite — $R \gg 1$, and minimize $F_1(\cdot, B_R)$ on the set of continuous and piecewise constant functions on $\mathcal{T}\cap B_R$ whose values on the boundary coincide with the affine deformation $x \mapsto \Lambda\cdot x$. To this aim, one needs to know $\mathcal{T}\cap B_R$. Yet, $\mathcal{T}$ is a Delaunay tessellation associated with a point process on the whole space $\mathbb{R}^3$. We therefore need to approximate the point process itself on the domain B_R. The point process we consider on $\mathbb{R}^3$ is the thermodynamic limit of the random parking measure associated with unit balls (that we will refer to as the "parking lattice"), which is rigorously defined as the limit of point processes on finite domains in [19]. As shown in [12], the parking lattice is stationary, ergodic, isotropic, almost surely general (the associated Delaunay tessellation is unique), and satisfies the hard-core and non-empty space conditions required by Theorem 1. Hence, formula (18) makes sense. The approximation of the parking lattice on finite domains B_R is as follows. For all $R>0$, we make a uniform mesh of the sphere S_R of radius R (with triangles of side ~ 1) and consider a hard-core Poisson point process (with minimal distance 1) in B_{R-1} up to the packing limit (that is, until one cannot add any other point: due to the

hard-core constraint, there is an easy upper bound on the number of points which can be accepted). In particular, for such a point process, any two points are at least at distance 1, and any ball of radius 1 has at least one point. We denote by $\mathcal{T}_R$ a Delaunay tessellation associated with the mesh of S_R and the points in B_{R-1}. As proved in [12], although $\mathcal{T}_R$ does not coincide with $\mathcal{T} \cap B_R$, we still have

$$\lim_{R\to\infty} \frac{1}{|B(R)|} \inf\{F_1(u, B_R), u \in \mathcal{S}(\mathcal{T}_R), u(x) = \xi \cdot x \text{ if } \mathrm{dist}(x, \partial B_R) \le 2\mathbf{r}\} \;=\; W_{\mathrm{V}}(\xi). \tag{22}$$

This is the final variant of (5) we consider, and which has the advantage to be *practically computable*.

The numerical approximation of (22) is made in two steps:

- We first generate the deterministic set of points on ∂B_R and a realization of the stochastic set of points in B_{R-1}. The latter is generated iteratively. Points are randomly picked in B_{R-1}. The first point is accepted. When another point is picked, either it is at distance less than 1 from a point which has already been accepted and it is discarded, or it is at distance at least 1 from all the other points and it is accepted. The algorithm stops when B_{R-1} is packed, that is, when no additional point can be accepted. Given the deterministic set of points on ∂B_R and the realization of the random set of points in B_{R-1} we then construct an associated Delaunay tessellation of B_R. Note that this choice characterizes the relation between the number N_R of chains in the volume $B(R)$, and $N_R/|B(R)|$ tends to a deterministic limit as $R \to \infty$.
- In a second step, we solve the minimization problem associated with (22) for R finite and the Delaunay tessellation of B_R (well-defined as the minimization of a smooth coercive function on a finite-dimensional space) by a Newton algorithm. It is classical in nonlinear elasticity (see for instance [15], and [21]) provided the addition of the energy of the edges (which are "non-standard" one-dimensional elements). Continuation methods are also used to ensure the convergence of the Newton algorithm.

In practice, we also consider several independent realizations of the stochastic set of points and make an empirical average of the approximations of W_{V} obtained. This enhances the convergence with respect to the randomness. In the tests reported here, we have taken 10 independent realizations of the random parking measure with 59,500 points.

For notational convenience, we define $\mathcal{W}_{nn} : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^+$ by $\mathcal{W}_{nn}(h, \zeta) := W_{nn}(h, |\zeta|)$. At the end of the minimization algorithm, the homogenized energy W_{V} is approximated by the spatial average on B_R of the energy density of the minimizer u_R which has been numerically obtained:

$$\begin{aligned}
W_{\mathrm{V}}(\Lambda) \;&\simeq\; \frac{1}{|B_R|} F_1(u_R, B_R) \\
&=\; \frac{1}{|B_R|}\Big(\sum_{e\in\mathcal{E}_d, e\subset B_R} W_{nn}\Big(|e_1 - e_2|, \frac{|u_R(e_1) - u_R(e_2)|}{|e_1 - e_2|}\Big) \\
&\qquad\qquad + \sum_{T\in\mathcal{T}, T\subset B_R} \int_T W_{\mathrm{vol}}(\det \nabla u_R)\Big) \\
&=\; \frac{1}{|B_R|}\Big(\sum_{e\in\mathcal{E}_d, e\subset B_R} \mathcal{W}_{nn}\Big(|e_1 - e_2|, \frac{|u_R(e_1) - u_R(e_2)|}{|e_1 - e_2|}\Big) \\
&\qquad\qquad + \sum_{T\in\mathcal{T}, T\subset B_R} \int_T W_{\mathrm{Helm}}(\nabla u_R)\Big).
\end{aligned}$$

The Piola-Kirchhoff stress tensor is given by the spatial average on B_R of the associated local Piola-Kirchhoff stress tensor (provided the minimizer is isolated, and the local Hessian strongly elliptic):

$$D_\Lambda W_{\rm V}(\Lambda) \simeq \frac{1}{|B_R|}\Big(\sum_{e\in\mathcal{E}_d, e\subset B_R} D_\zeta \mathcal{W}_{nn}\Big(|e_1-e_2|, \frac{|u_R(e_1)-u_R(e_2)|}{|e_1-e_2|}\Big)\otimes \frac{e_1-e_2}{|e_1-e_2|} + \sum_{T\in\mathcal{T}, T\subset B_R} \int_T D_\Lambda W_{\rm Helm}(\nabla u_R)\Big).$$

The approximation of the Hessian $D^2_{\Lambda^2} W_{\rm V}(\Lambda)$ is obtained from the implicit definition of $W_{\rm V}$ by solving in addition 9 linear systems (see [11]).

5.2. **Data assimilation in moderate deformation.** In this paragraph we test the data assimilation procedure for deformations up to 200%, that is, we take admissible deformations in the set of diagonal matrices with entries in $(1/3, 3)$ and consider $\overline{N} = 250$ in (21). We first observe that the quality of the reconstruction saturates for $k_1 = 2$ and $k_2 = 1$ (and we do not display the results for higher k_1, k_2). The results of the data assimilation procedure are displayed in Table 12 for $\eta = 0$ and $m = 3$, where Π is the relative error on the Piola stress tensor, and H is the relative error on the Hessian. As can be seen, the Hessian is better approximated than the Piola stress tensor, so that the regularization procedure with $\eta > 0$ does not improve the reconstruction. To illustrate these results on mechanical quantities of interest, we have proceeded as for the Arruda-Boyce model and plotted the stress tensor of the micro-macro constitutive law and of the associated Ogden law (obtained by the reconstruction procedure) for two typical mechanical experiments: uniaxial deformation on Figure 4 and pure shear on Figure 5. The reconstruction is in very good agreement in moderate deformation.

error Π	error H	k_1	k_2	K_1	K_2	a_1	α_1	a_2	α_2	b_1	β_1
$4.267\cdot 10^{-3}$	$1.903\cdot 10^{-4}$	1	0	2.563	2.577	$1.400\cdot 10^{-2}$	2.013				
$3.170\cdot 10^{-3}$	$1.385\cdot 10^{-4}$	1	1	2.562	2.577	$1.388\cdot 10^{-2}$	2.014			$1.181\cdot 10^{-4}$	1.846
$3.170\cdot 10^{-3}$	$1.386\cdot 10^{-4}$	2	1	2.562	2.577	$1.389\cdot 10^{-2}$	2.014	$5.748\cdot 10^{-8}$	2.450	$1.160\cdot 10^{-4}$	1.855

TABLE 12. Approximation of the micro-macro model (with $n/\beta = 0.27, \overline{N} = 250, K = 2.575$) by Ogden laws, using $m = 3$, $\eta = 0$ (reconstruction on (1/3,3)).

5.3. **Data assimilation in large deformation.** In this paragraph we consider deformations up to 500%, that is, we take admissible deformations in the set of diagonal matrices with entries in $(1/6, 6)$. We first observe that the quality of the reconstruction saturates for $k_1 = 2$ and $k_2 = 1$ (and we do not display the results for higher k_1, k_2). The results are qualitatively different for $\overline{N} = 26.5$ and $\overline{N} = 250$ (see the two sets of data in (21)). In particular, whereas the Hessian is better approximated than the stress tensor for $\overline{N} = 250$ and $\eta = 0$, it is not the case for $\overline{N} = 26.5$, see Tables 13 and 14. In the latter, a regularization using the Hessian with $\eta > 0$ allows one to improve the approximation of the Hessian, without compromising the approximation of the stress tensor (see Table 15).

The stress tensors of the micro-macro constitutive law and of the associated Ogden law (obtained by the reconstruction procedure) are displayed for uniaxial deformation on Figures 6 and 8 and for pure shear on Figures 7 and 9. Note that the reconstruction is better in large deformation for $\overline{N} = 250$ than for $\overline{N} = 26.5$. This is due to the finite extensibility in the model: For $\overline{N} = 26.5$ the energy of the polymer chains blows up faster than for $\overline{N} = 250$, which makes the range of approximation of the Ogden laws much smaller.

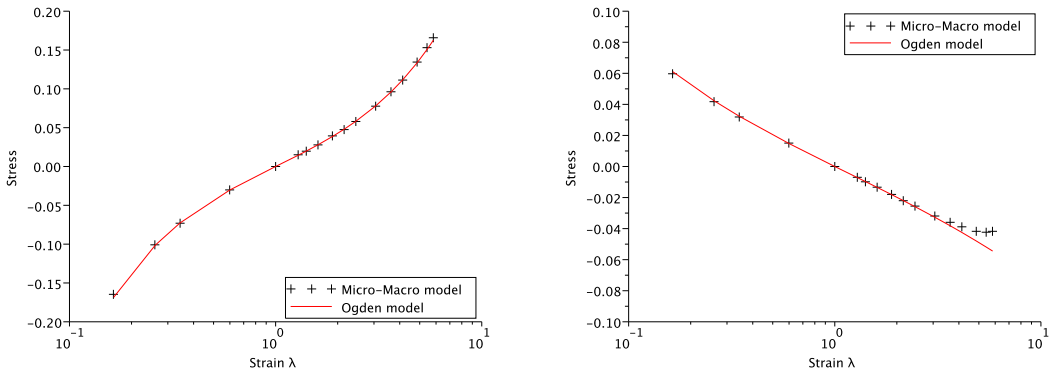


FIGURE 4. *Comparison between the micro-macro model ($\overline{N} = 250$) and the associated Ogden law (reconstruction on (1/3,3)) in uniaxial deformation $\Lambda = diag(\frac{1}{\sqrt{\lambda}}, \frac{1}{\sqrt{\lambda}}, \lambda)$. Piola Kirchhoff stress components Π_3 (on the left) and $\Pi_1 = \Pi_2$ (on the right) with respect to λ in log-scale.*

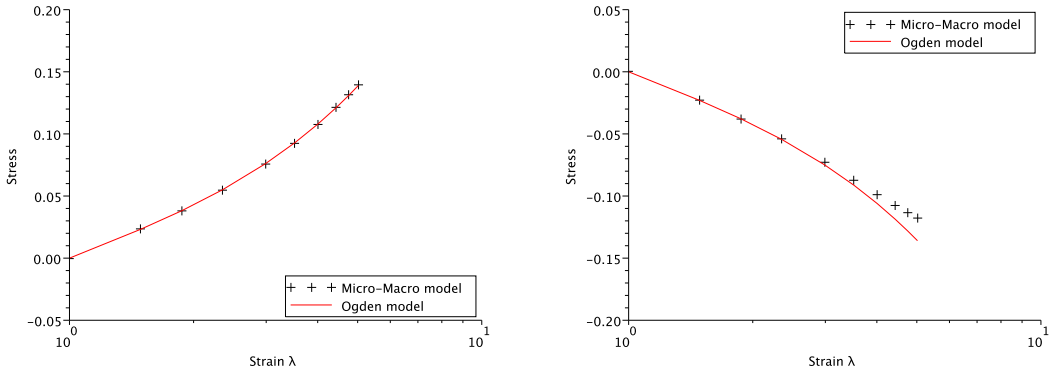


FIGURE 5. *Comparison between the micro-macro model ($\overline{N} = 250$) and the associated Ogden law (reconstruction on (1/3,3)) in pure shear $\Lambda = diag(\lambda, \frac{1}{\lambda}, 1)$. Piola Kirchhoff stress components Π_1 (on the left) and Π_2 (on the right) with respect to λ in log-scale.*

5.4. **An example of ill-posed problem for the reconstruction using Treloar's data.** In this last subsection we come back to the problem of estimating parameters of an Ogden law starting from standard mechanical measurements, such as the Treloar data (the engineering stresses in uniaxial compression, uniaxial traction, and pure shear). The parameters of the Arruda-Boyce ($n/\beta = 0.265, \overline{N} = 26.98, K_1 = 2.439, K_2 = 2.575$) and of the micro-macro ($n/\beta = 0.27, \overline{N} = 26.5, K = 2.575$) models considered here have been chosen to match Treloar's experiments (see [11]). On Figure 10, we have plotted the outputs of these experiments (engineering stress for three types of deformation) for the Ogden's law approximating the Arruda-Boyce model, and for the Ogden's law approximating the micro-macro model. As can be seen on these graphs, the two Odgen's laws yield very close results.

However, the engineering stress does not characterize the law. In particular two laws may at the same time be close in terms of Treloar's experiments and quite different on the full Piola-Kirchhoff stress tensor. We have plotted the stress tensor of the two Odgen's laws for two typical mechanical experiments: uniaxial deformation (see Figure 11) and pure

error Π	error H	k_1	k_2	K_1	K_2	a_1	α_1	a_2	α_2	b_1	β_1
$2.204 \cdot 10^{-2}$	$1.437 \cdot 10^{-3}$	1	0	2.563	2.577	$1.327 \cdot 10^{-2}$	2.046				
$1.993 \cdot 10^{-2}$	$1.285 \cdot 10^{-3}$	1	1	2.563	2.577	$1.305 \cdot 10^{-2}$	2.050			$2.511 \cdot 10^{-4}$	1.663
$1.993 \cdot 10^{-2}$	$1.285 \cdot 10^{-3}$	2	1	2.563	2.577	$1.305 \cdot 10^{-2}$	2.050	$2.283 \cdot 10^{-9}$	2.146	$2.512 \cdot 10^{-4}$	1.663

TABLE 13. Approximation of the micro-macro model (with $n/\beta = 0.27, \overline{N} = 250, K = 2.575$) by Ogden laws, using $m = 3$ and $\eta = 0$ (reconstruction on (1/6,6)).

error Π	error H	k_1	k_2	K_1	K_2	a_1	α_1	a_2	α_2	b_1	β_1
$4.049 \cdot 10^{-2}$	$8.587 \cdot 10^{-2}$	1	0	2.453	2.569	$1.033 \cdot 10^{-1}$	2.240				
$3.491 \cdot 10^{-2}$	$8.128 \cdot 10^{-2}$	1	1	2.448	2.574	$9.525 \cdot 10^{-2}$	2.275			$1.432 \cdot 10^{-2}$	1.248
$2.901 \cdot 10^{-2}$	$4.766 \cdot 10^{-2}$	2	1	2.450	2.572	$1.040 \cdot 10^{-1}$	2.198	$4.766 \cdot 10^{-8}$	8.991	$4.722 \cdot 10^{-3}$	1.716

TABLE 14. Approximation of the micro-macro model (with $n/\beta = 0.27, \overline{N} = 26.5, K = 2.575$) by Ogden laws, using $m = 3$ and $\eta = 0$ (reconstruction on (1/6,6)).

η	error Π	error H	k_1	k_2	K_1	K_2	a_1	α_1	a_2	α_2	b_1	β_1
0	$2.901 \cdot 10^{-2}$	$4.766 \cdot 10^{-2}$	2	1	2.450	2.572	$1.040 \cdot 10^{-1}$	2.198	$4.766 \cdot 10^{-8}$	8.991	$4.722 \cdot 10^{-3}$	1.716
0.01	$2.908 \cdot 10^{-2}$	$2.558 \cdot 10^{-2}$	2	1	2.449	2.572	$1.041 \cdot 10^{-1}$	2.196	$3.689 \cdot 10^{-7}$	7.838	$4.809 \cdot 10^{-3}$	1.706
0.1	$2.926 \cdot 10^{-2}$	$1.525 \cdot 10^{-2}$	2	1	2.448	2.571	$1.042 \cdot 10^{-1}$	2.193	$1.903 \cdot 10^{-6}$	6.937	$4.941 \cdot 10^{-3}$	1.693
1	$2.959 \cdot 10^{-2}$	$1.304 \cdot 10^{-2}$	2	1	2.447	2.570	$1.049 \cdot 10^{-1}$	2.182	$9.777 \cdot 10^{-6}$	6.122	$5.275 \cdot 10^{-3}$	1.663

TABLE 15. Influence of η on the approximation of the micro-macro model (with $n/\beta = 0.27, \overline{N} = 26.5, K = 2.575$) by Ogden laws, using $m = 3$ (reconstruction on (1/6,6)).

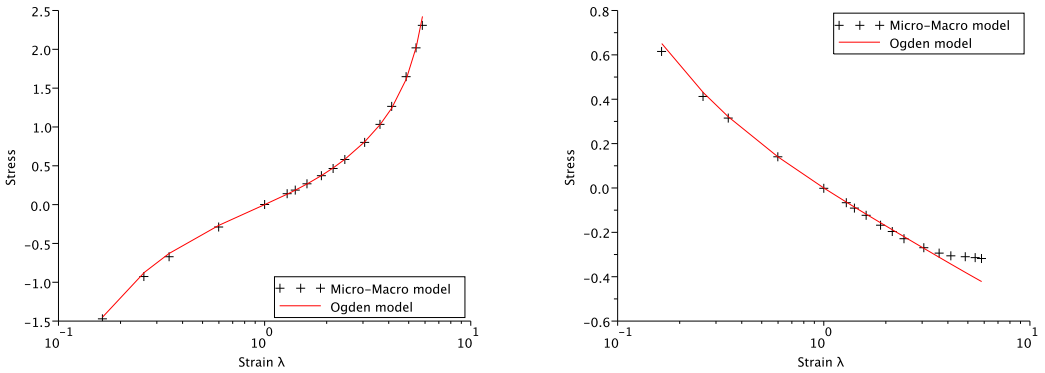


FIGURE 6. *Comparison between the micro-macro model ($\overline{N} = 26.5$) and the associated Ogden law (reconstruction on (1/6,6)) in uniaxial deformation $\Lambda = diag(\frac{1}{\sqrt{\lambda}}, \frac{1}{\sqrt{\lambda}}, \lambda)$. Piola Kirchhoff stress components Π_3 (on the left) and $\Pi_1 = \Pi_2$ (on the right) with respect to λ in log-scale.*

shear (see Figure 12). As can be seen, they do not match as well as for the engineering stress.

This illustrates that the engineering stress usually measured in mechanical experiments is not sufficient to identify a single empirical Ogden's law, and completes the analysis of [18]. The micro-macro model considered in this article does not suffer from this since we have access to the full Piola stress tensor, and the reconstruction procedure is efficient.

To conclude, we have presented a method to reconstruct an analytical constitutive law starting from the numerical approximation of an implicit constitutive law obtained by

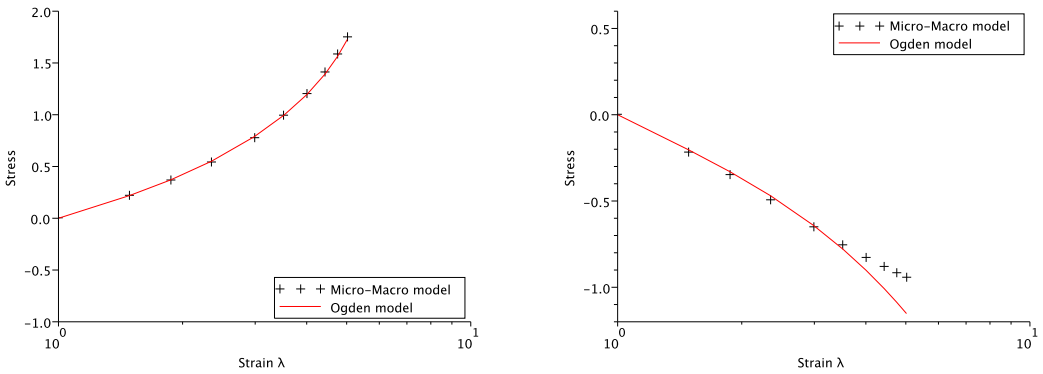


FIGURE 7. *Comparison between the micro-macro model ($\overline{N} = 26.5$) and the associated Ogden law (reconstruction on (1/6,6)) in pure shear $\Lambda = diag(\lambda, \frac{1}{\lambda}, 1)$. Piola Kirchhoff stress components Π_1 (on the left) and Π_2 (on the right) with respect to λ in log-scale.*

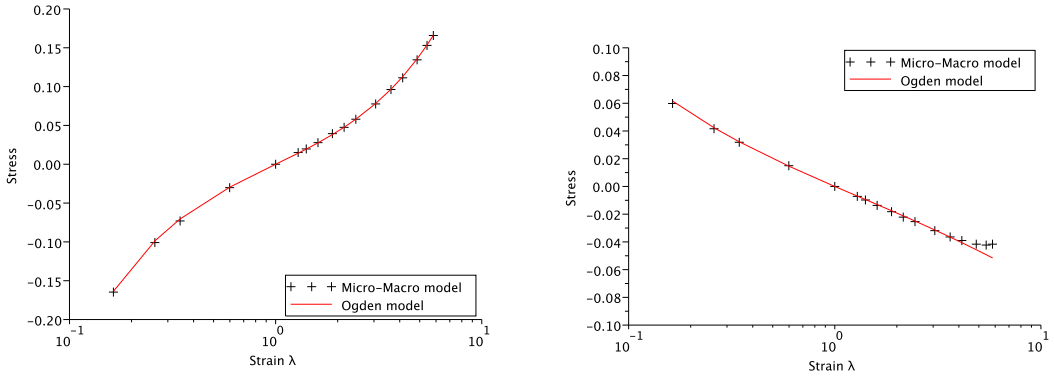


FIGURE 8. *Comparison between the micro-macro model ($\overline{N} = 250$) and the associated Ogden law (reconstruction on (1/6,6)) in uniaxial deformation $\Lambda = diag(\frac{1}{\sqrt{\lambda}}, \frac{1}{\sqrt{\lambda}}, \lambda)$. Piola Kirchhoff stress components Π_3 (on the left) and $\Pi_1 = \Pi_2$ (on the right) with respect to λ in log-scale.*

a rigorous upscaling of a polymer physics model. The choice of the parametrized space of constitutive laws for the reconstruction procedure is crucial. In this article we have considered the explicit subspace of *polyconvex* Ogden laws constructed by Ball [3]. The results are qualitatively and quantitatively convincing. They could however be improved in two respects. First, it is already clear from the approximation of the Arruda-Boyce model that the Ogden laws do not reproduce equally well the behavior of elastomers depending on the physical parameter N (the typical number of monomers per polymer-chain). Indeed, for small to moderate N, the energy of polymer-chains grows faster than any polynomial close to the finite extensibility limit, which reduces the range of validity of Ogden's laws. A second aspect concerns more subtle properties such as the Rivlin effect, which is a nonlinear effect at small deformation. Whereas this effect is not captured by the Arruda-Boyce model, it is captured to some extent by the micro-macro model (see [11]). The reconstructed Ogden law for the micro-macro model does not reproduce this effect. In order to further improve the reconstruction procedure, the parametrized space of Ogden laws needs to be enriched. The difficulty of this promising program lies in the

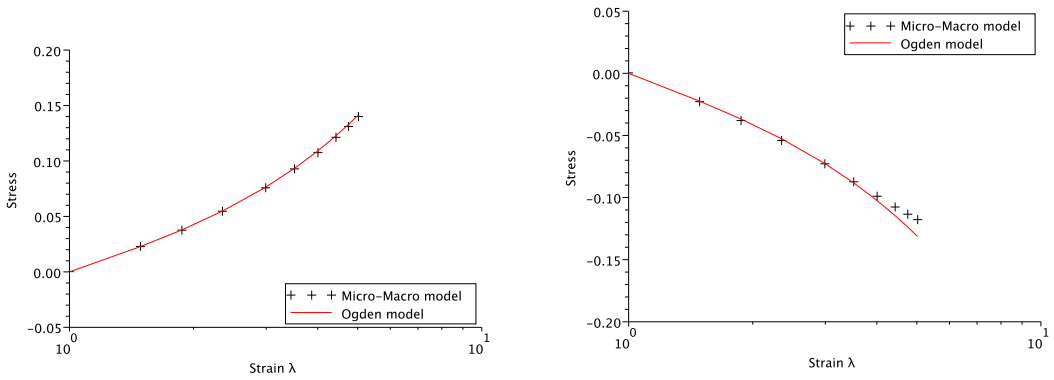
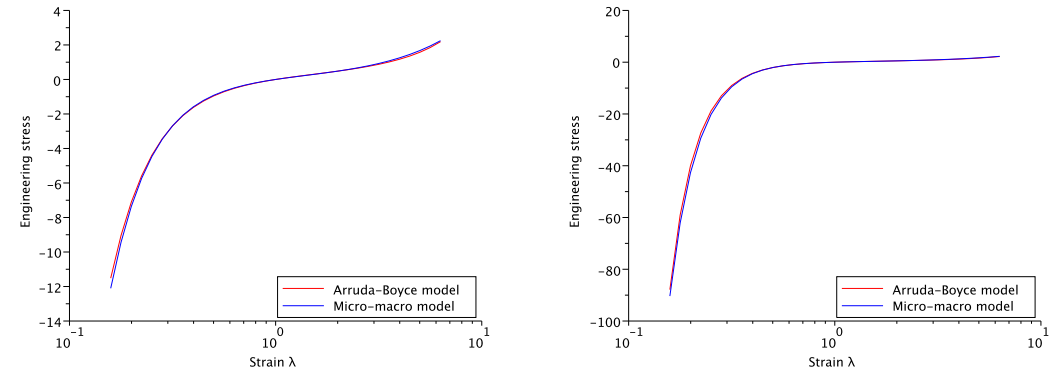


FIGURE 9. *Comparison between the micro-macro model ($\overline{N} = 250$) and the associated Ogden law (reconstruction on (1/6,6)) in pure shear $\Lambda = diag(\lambda, \frac{1}{\lambda}, 1)$. Piola Kirchhoff stress components Π_1 (on the left) and Π_2 (on the right) with respect to λ in log-scale.*



(a) (b)

FIGURE 10. *Comparison between the Ogden approximations of the micro-macro and Arruda-Boyce models ($\overline{N} = 26.5$) on the Treloar experiments. (a) uniaxial compression/traction $\Lambda = diag(\lambda, \frac{1}{\sqrt{\lambda}}, \frac{1}{\sqrt{\lambda}})$ and (b) pure shear $\Lambda = diag(\lambda, \frac{1}{\lambda}, 1)$. Engineering stress with respect to the strain λ.*

qualitative properties we required on the laws (and in particular quasiconvexity in the form of polyconvexity).

APPENDIX A. CMA-ES

The Covariance Matrix Adaptation Evolution Strategy, developed by Hansen [13], is a stochastic method for real-parameter (continuous domain) optimization. It is an attractive option for non-linear optimization, if classical search method, *i.e.* quasi-Newton methods (BFGS) and/or conjugate gradient methods, fails due to a non-convex search landscape. It is also particularly well-suited to non-differentiable problems or problems for which $\frac{d\mathcal{F}}{dp}$ is not easily available. Yet, as many optimization methods, it can yield a local optimum, and not a global one. We quickly describe the algorithm :

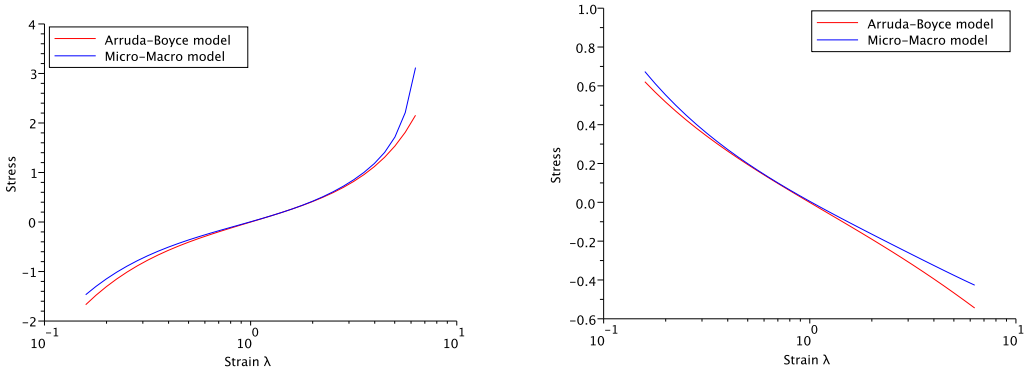


FIGURE 11. *Comparison between the Ogden approximations of the micro-macro and Arruda-Boyce models ($\overline{N} = 26.5$) in uniaxial deformation $\Lambda = diag(\frac{1}{\sqrt{\lambda}}, \frac{1}{\sqrt{\lambda}}, \lambda)$. Piola Kirchhoff stress components Π_3 (on the left) and $\Pi_1 = \Pi_2$ (on the right) with respect to λ in log-scale.*

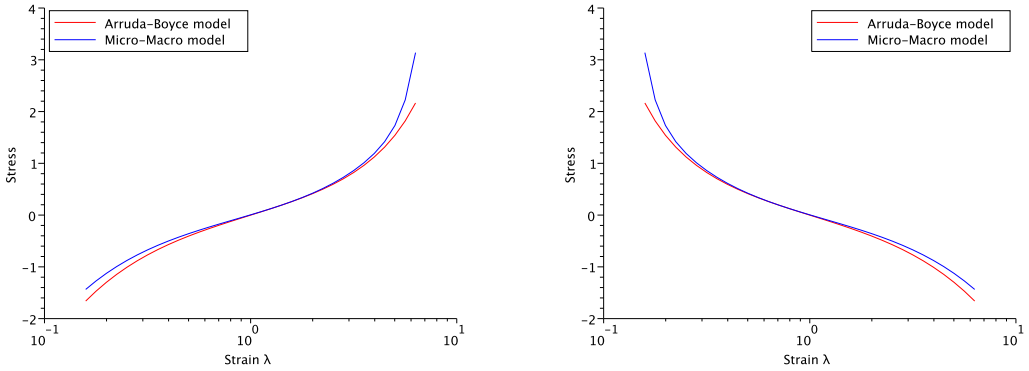


FIGURE 12. *Comparison between the Ogden approximations of the micro-macro and Arruda-Boyce models ($\overline{N} = 26.5$) in pure shear $\Lambda = diag(\lambda, \frac{1}{\lambda}, 1)$. Piola Kirchhoff stress components Π_1 (on the left) and Π_2 (on the right) with respect to λ in log-scale.*

- Generation
  The population of individuals (search points p) is generated by sampling a multivariate normal distribution as follows :
  $$p_k^{(g)} \sim G^{(g)} = m^{(g)} + \sigma^{(g)} \mathcal{N}(0, C^{(g)}), \quad k = 1, .., S, \quad g \geq 0,$$
  where S is the population size (small population size usually lead to faster convergence, large population sizes help to avoid local optima), $p_k^{(g)} \in \mathbb{R}^n$ is the k-th individual from generation g, $m^{(g)} \in \mathbb{R}^n$ is the mean value of the search distribution at generation g, $\sigma^{(g)} \in \mathbb{R}_+$ is the standard deviation of the search distribution at generation g and $\mathcal{N}(0, C^{(g)})$ is a multivariate normal distribution with zero mean and covariance matrix $C^{(g)}$.
- Selection
  We select $s \leq S$ individuals out of the population $p_1^{(g)}, ..., p_S^{(g)}$. Let us denote by $p_{i:S}^{(g)}$ the i-th best individual (with respect to the cost function), that is :
  $$\mathcal{F}(p_{1:S}^{(g)}) \leq \cdots \leq \mathcal{F}(p_{s:S}^{(g)}).$$

Now, we are going to see how to calculate $m^{(g+1)}, \sigma^{(g+1)}$ and $C^{(g+1)}$ for the next generation $g+1$ from these s selected parents.

- Updating the mean value
The new mean $m^{(g+1)}$ of the search distribution is defined as a weighted average of the s best individuals :

$$m^{(g+1)} = \sum_{i=1}^{s} w_i p_{i:S}^{(g)},$$

where $w_i \in \mathbb{R}_+$ are positive weight coefficients satisfying $w_1 \geq \cdots \geq w_s > 0$ and $\sum_{i=1}^{s} w_i = 1$. We also define the measure :

$$\mu = \left(\sum_{i=1}^{s} w_i^2 \right)^{-1}.$$

- Updating the covariance matrix
First case (rank-s-update): the population contains enough information to reliably estimate a covariance matrix from only one generation, that is $S \geq 4n$. In this case, we define :

$$C_s^{(g+1)} = \sum_{i=1}^{s} w_i \left(\frac{p_{i:S}^{(g)} - m^{(g)}}{\sigma^{(g)}} \right) \left(\frac{p_{i:S}^{(g)} - m^{(g)}}{\sigma^{(g)}} \right)^T.$$

The matrix $C_s^{(g+1)}$ is an estimator for the distribution of selected individuals, *i.e.* sampling from $C_s^{(g+1)}$ tends to reproduce selected individuals, giving a justification for what a better covariance matrix means.
Second case (cumulation with evolution path): information from previous generations must be used. We introduce the evolution path at generation g:

$$e_c^{(g+1)} = (1 - c_c) e_c^{(g)} + \sqrt{c_c(2 - c_c)\mu} \frac{m^{(g+1)} - m^{(g)}}{\sigma^{(g)}},$$

with $e_c^{(0)} = 0$ and $c_c \leq 1$. Here, $\sqrt{c_c(2 - c_c)\mu}$ is a normalization factor. Then, an estimator for the distribution is given by:

$$C_1^{(g+1)} = \left(e_c^{(g+1)} \right) \left(e_c^{(g+1)} \right)^T.$$

Final choice: combining rank-s-update and cumulation. Finally, we update the covariance matrix as follows:

$$C^{(g+1)} = (1 - c_\mu - c_1) C^{(g)} + c_s C_s^{(g+1)} + c_1 C_1^{(g+1)},$$

where $c_1 \leq 1$ and $c_s \leq 1 - c_1$ are learning rates. The choice of c_s is crucial. Small values lead to slow learning, too large values lead to a failure, because the covariance matrix degenerates.

- Updating the standard deviation
We introduce the conjugate evolution path :

$$e_\sigma^{(g+1)} = (1 - c_\sigma) e_\sigma^{(g)} + \sqrt{c_\sigma(2 - c_\sigma)\mu} (C^{(g)})^{-1/2} \frac{m^{(g+1)} - m^{(g)}}{\sigma^{(g)}},$$

where $c_\sigma < 1$ and $\sqrt{c_\sigma(2 - c_\sigma)\mu}$ is a normalization factor. To decide whether the standard deviation is long or short, we compare the length of the path $\|e_\sigma^{(g+1)}\|$ with its expected length under random selection $E\|\mathcal{N}(0, I)\|$. Thus, the new standard deviation is calculated as follows:

$$\sigma^{(g+1)} = \sigma^{(g)} \exp\left(\frac{c_\sigma}{d_\sigma} \left(\frac{\|e_\sigma^{(g+1)}\|}{E\|\mathcal{N}(0, I)\|} - 1 \right) \right),$$

where d_σ is a damping parameter.

References

[1] R. Alicandro, M. Cicalese, and A. Gloria. Convergence analysis of the Böl-Reese discrete model for rubber. In *Proceedings of the 11th International Symposium on Continuum Models and Discrete Systems*, 2008.

[2] R. Alicandro, M. Cicalese, and A. Gloria. Integral representation results for energies defined on stochastic lattices and application to nonlinear elasticity. *Arch. Ration. Mech. Anal.*, 200(3):881–943, 2011.

[3] J.M. Ball. Convexity conditions and existence theorems in nonlinear elasticity. *Arch. Rat. Mech. Anal.*, 63:337–403, 1977.

[4] M. Barchiesi. Loss of polyconvexity by homogenization: a new example. *Calc. Var. Partial Differential Equations*, 30(2):215–230, 2007.

[5] M. Bonnet and A. Constantinescu. Inverse problems in elasticity. *Inverse Problems*, 21(2):R1–R50, 2005.

[6] A. Braides. Loss of polyconvexity by homogenization. *Arch. Rat. Mech. Anal.*, 127:183–190, 1994.

[7] P.G. Ciarlet. *Mathematical elasticity. Volume I: three-dimensional elasticity*, volume 20 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 1988.

[8] B. N. Delone, N. P. Dolbilin, M. I. Štogrin, and R. V. Galiulin. A local test for the regularity of a system of points. *Dokl. Akad. Nauk SSSR*, 227(1):19–21, 1976.

[9] A. S. Gendy and A. F. Saleeb. Nonlinear material parameter estimation for characterizing hyperelastic large strain models. *Comput. Mech.*, 25:66–77, 2000.

[10] A. Gloria. Strong ellipticity of nonlinear elastic materials and homogenization of periodic and stochastic discrete systems. In preparation.

[11] A. Gloria, P. Le Tallec, and M. Vidrascu. Foundation, analysis, and numerical investigation of a variational network-based model for rubber. *Continuum Mech. Thermodyn.*, 2013.

[12] A. Gloria and M.D. Penrose. Random parking, Euclidean functionals, and rubber elasticity. *Comm. Math. Physics*, 321(1):1–31, 2013.

[13] N. Hansen. The CMA evolution strategy: a comparing review. In J.A. Lozano, P. Larranaga, I. Inza, and E. Bengoetxea, editors, *Towards a new evolutionary computation. Advances on estimation of distribution algorithms*, pages 75–102. Springer, 2006.

[14] J. Kristensen. On the non-locality of quasiconvexity. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 16(1):1–13, 1999.

[15] P. Le Tallec. Numerical methods for nonlinear three-dimensional elasticity. In *Handbook of numerical analysis, Vol. III*, pages 465–622. North-Holland, 1994.

[16] R. W. Ogden. Large deformation isotropic elasticity: on the correlation of theory and experiment for incompressible rubberlike solids. *Proc. R. Soc. Lond. A*, 326:565–584, 1972.

[17] R. W. Ogden. Recent advances in the phenomenological theory of rubber elasticity. *Rubber Chem. Technol.*, 59:361–383, 1986.

[18] R. W. Ogden, G. Saccomandi, and I. Sgura. Fitting hyperelastic models to experimental data. *Comp. Mech.*, 34:484–502, 2004.

[19] M.D. Penrose. Random parking, sequential adsorption, and the jamming limit. *Commun. Math. Phys.*, 218:153–176, 2001.

[20] E. H. Twizell and R. W. Ogden. Non-linear optimization of the material constants in ogden's stress-deformation relation for incompressible isotropic elastic materials. *J. Aust. Math. Soc.*, 24:424–434, 1983.

[21] M. Vidrascu. Solution of non-linear elasticity problems using the *continu* software. *Inria Report Research (http://www.inria.fr/rrrt/rr-4128.html)*, 2001.

(Maya de Buhan) CNRS, UMR 8145, MAP5, Université Paris Descartes, Sorbonne Paris Cité, France

E-mail address: maya.de-buhan@parisdescartes.fr

(Antoine Gloria) Département de mathématique, Université Libre de Bruxelles, Belgium, and MEPHYSTO team, INRIA Lille - Nord Europe, Villeneuve d'Ascq, France

E-mail address: agloria@ulb.ac.be

(Patrick Le Tallec) LMS, École polytechnique, Palaiseau, France

E-mail address: patrick.letallec@polytechnique.edu

(Marina Vidrascu) Project-team REO, Inria Paris-Rocquencourt, Le Chesnay, France

E-mail address: marina.vidrascu@inria.fr